Table of Contents

Abstracts

Abstracts

Accelerated Design and Deployment of Metal Alloy Surfaces for Chemoresponsive Liquid Crystals

Lead Investigator: Nicholas L. Abbott (nla34@cornell.edu)
Participating Institutions: Cornell University, University of Wisconsin-Madison, Kent State University
Website: http://ls-staging.doit.wisc.edu/liquid-crystals/
Keywords: Liquid crystals; chemical-responsive; surfaces; catalysis; soft matter

Project Scope. Our project explores liquid crystals (LCs) supported on metal surfaces as a novel materials platform for amplification of atomic-scale events involving elementary surface reactions and catalytic reactions *in situ* into macroscopic optical signals. The project involves a convergence of soft matter science and atom-scale understanding of chemical reaction mechanisms at metal alloy surfaces from the surface science/catalysis community. It also aims to move chemoresponsive LCs along the materials development continuum by focusing on industrially important analytes (e.g., Cl_2, H_2, NO_2 and CO) with sufficiently weak binding energies that they cannot be detected by previous designs of chemoresponsive LCs.

Relevance to MGI. The project is establishing an iterative theoretical and experimental methodology for the rapid design of chemically responsive LCs. *Generations* of theory, synthesis, and experimental evaluation drive the accelerated deployment of new designs of LCs that are based on metal and metal alloy-supported thin films. Specifically, electronic structure calculations are guiding designs of new mesogens with tailored interactions with metal surfaces, which are then synthesized and experimentally validated. In some cases, the experiments reveal additional considerations that need to be incorporated into the computational models, leading to model refinements and improved predictive ability. As detailed below, this iterative process (Fig. 1) has led to two *Generations* of designs based on atomic displacement events and catalytic reactions on surfaces.

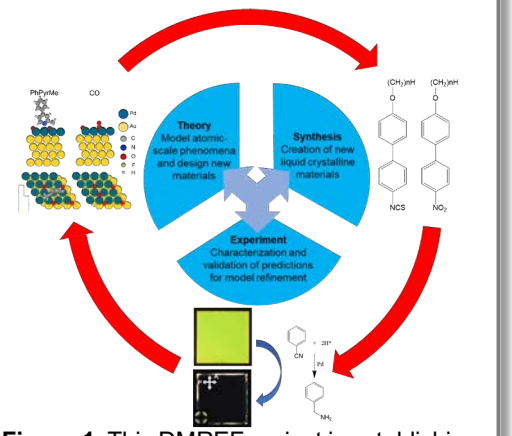


Figure 1. This DMREF project is establishing an iterative theoretical and experimental methodology for the accelerated design of chemically responsive LCs. Generations of theory, synthesis, and experimental evaluation drive the rapid deployment of new designs of LCs supported on metal and metal alloy thin films.

Technical Progress. Technical progress is organized into two *Generations* of designs, both informed by fundamental knowledge related to surface science and catalysis. The first *Generation* of designs focuses on chemoresponsive LCs that employ elementary surface reactions involving *displacement events* at the atomic scale. Electronic structure calculations of displacement events guided design and synthesis of new classes of mesogens with tailored interactions with metal surfaces. Feedback from experiments informed refined computational models, ultimately yielding chemoresponsive LCs that respond to CO, NH_3, NO_2 based on metal alloy (Pd on Au) and metal oxide (AgO) surfaces. Additionally, knowledge emerging from these first generation designs were fed forward into a second *Generation* of designs of chemoresponsive LCs in which the thermochemical pathways of complex *catalytic reactions* were computationally predicted as the basis of a chemical response in a LC system. For example, through a combination of experiment and computation, we designed a LC system driven by sequential reactions on Pd surfaces, the first reaction being dissociation of H_2, and the second being a hydrogenation reaction that involves the mesogen. Additionally, we benchmarked kinetic information emerging from the dynamic LC response against conventional catalysis experiments and microkinetic models, revealing that the dynamic responses of LCs on metal/alloy surfaces can yield fundamental information of relevance to heterogeneous catalysis. More broadly, this accomplishment constitutes the first design of a chemoresponsive LC based on a catalytic reaction. It also emphasizes the large upside potential that our DMREF team seeks to realize going forward by exploring the convergence of catalysis and soft materials design via iterative cycles of computation, synthesis and characterization (MGI approach).

Future Plans. The future goals of our team include exploration (via iterations of computational prediction, synthesis and experiment) of (i) an expanded range of chemical functional groups tailored for interaction with metal alloy surfaces (NO_2, NCS, fluorinated mesogens), and (ii) studies aimed at understanding the role of reorganization of the metal substrate in the ordering response of LCs to surface catalytic reactions (e.g., Pd on Au), (iii) the use of metal nanoparticle-decorated surfaces (e.g., Au, Pd or Pt on TiO_2 substrates) as catalytic surfaces for design of chemoresponsive LCs, (iv) additional metal and metal alloy surfaces containing Pt, Ag, Ru or Cu. Additionally, future electronic structure calculations will focus on elucidating the role of adsorbed mesogens on modulating the catalytic reactions catalyzed by the metal/alloy surfaces.

Broader Impacts and Workforce Development. This DMREF project provides a rich environment for training of the next generation of scientists and engineers in accelerated materials design. The entire DMREF team meets via Zoom every two weeks, and additionally, frequent emails and phone calls between team members lead to a constant exchange of ideas between computation, synthesis and characterization efforts. The Kent group is remarkably productive in synthesizing new mesogens, and many of those mesogens have been evaluated in experiments by the Cornell group. Two Mavrikakis group members and one Abbott group member who worked on this project have joined the academic ranks. Although COVID-19 made in-person participation in public outreach events challenging, our team has made strides towards its vision of chemoresponsive LCs.

Data Management and Open Access. We have continued to expand the capabilities of our website and database (http://ls-staging.doit.wisc.edu/liquid-crystals/). Progress includes an expanded range of surfaces and a LC synthesis section. Because the website contains data not yet patented, parts of it are password protected.

Advancing Along the Materials Development Continuum and Partnerships to Translation. Our DMREF team has validated computationally-guided designs of chemoresponsive LCs based on metal alloy surfaces that provide access to functional properties (range of chemical targets, sensitivity, and selectivity) that have not been possible with prior designs of chemoresponsive LCs. The efforts have been guided by exchanges with *Platypus Technologies LLC* and several research groups developing robotic systems. Our DMREF team filed two patent applications in 2021 based on the outcomes of this DMREF project (US Patent 10,928,306 ad US Patent App. 17/353,271).

Publications and References

1) H. Yu, J. I. Gold, T. J. Trent, N. Bao, R. J. Twieg, M. Mavrikakis, N. L. Abbott., Synthetic liquid crystalline systems that are rapidly and reversibly actuated by chemical catalysis. *Submitted*.
2) T. Szilvási, H. Yu, J. I. Gold, N. Bao, T. J. Wolter, R. J. Twieg, N. L. Abbott, M. Mavrikakis, Coupling the chemical reactivity of bimetallic surfaces to the orientations of liquid crystals. *Mater. Horizons.* 8, 2050–2056 (2021).
3) N. Bao, J. I. Gold, T. Szilvási, H. Yu, R. J. Twieg, M. Mavrikakis, N. L. Abbott, Designing chemically selective liquid crystalline materials that respond to oxidizing gases. *J. Mater. Chem. C.* 9, 6507–6517 (2021).
4) K. Wang, M. S. Rahman, T. Szilvási, J. I. Gold, N. Bao, H. Yu, N. L. Abbott, M. Mavrikakis, R. J. Twieg, Influence of multifluorophenyloxy terminus on the mesomorphism of the alkoxy and alkyl cyanobiphenyl compounds in search of new ambient nematic liquid crystals and mixtures. *Liq. Cryst*. 48, 672–688 (2021).
5) H. Yu, K. Wang, T. Szilvási, K. Nayani, N. Bao, R. J. Twieg, M. Mavrikakis, N. L. Abbott, Design of chemoresponsive soft matter using hydrogen-bonded liquid crystals. *Materials* 14, 1–17 (2021).
6) J. Gold, T. Szilvási, N. L. Abbott, M. Mavrikakis, Binding of organophosphorus nerve agents and their simulants to metal salts. *ACS Appl. Mater. Interfaces.* 12, 30941–30953 (2020).
7) K. Nayani, Y. Yang, H. Yu, P. Jani, M. Mavrikakis, N. Abbott, Areas of opportunity related to design of chemical and biological sensors based on liquid crystals. *Liq. Cryst. Today*. 29, 24–35 (2020).
8) K. Wang, T. Szilvási, J. Gold, H. Yu, N. Bao, P. Rai, M. Mavrikakis, N. L. Abbott, R. J. Twieg, New room temperature nematogens by cyano tail termination of alkoxy and alkylcyanobiphenyls and their anchoring behavior on metal salt-decorated surface. *Liq. Cryst.* 47, 540–556 (2020).
9) K. Wang, M. Jirka, P. Rai, R. J. Twieg, T. Szilvási, H. Yu, N. L. Abbott, M. Mavrikakis, Synthesis and properties of hydroxy tail-terminated cyanobiphenyl liquid crystals. *Liq. Cryst.* 46, 397–407 (2019).
10) H. Yu, T. Szilvási, K. Wang, J. I. Gold, N. Bao, R. J. Twieg, M. Mavrikakis, N. L. Abbott, Amplification of elementary surface reaction steps on transition metal surfaces using liquid crystals: dissociative adsorption and dehydrogenation. *J. Am. Chem. Soc.* 141, 16003–16013 (2019).

Low Cost, High Strength and Ductile Mg Alloys

Lead Investigator: Sean R. Agnew, agnew@virginia.edu
Participating Institutions: University of Virginia, Cornell University
Website: none.
Keywords: Lightweight, Alloy design, Precipitation, GP zones, Crystal plasticity.

Project Scope

As the lightest structural metals, Mg alloys provide unique opportunities to enhance energy efficiency in the transportation sector. Nevertheless, wrought Mg alloys remain underutilized, due to high processing costs and inadequate mechanical properties. This project aims to simultaneously improve strength and ductility, while being cognizant of cost. The new materials design strategy utilizes GP-zones composed of elements widely available in the earth's crust (e.g., Mg, Al, Ca). It is hypothesized that the thermally activated dislocation-GP-zone interactions increase strain rate sensitivity, leading to improved ductility, in addition to the obvious strengthening effect.

Relevance to MGI

This project combines computational and experimental techniques from atomic length scales to the macroscale (Figure 1). Given the macroscopic stress-strain response and the microstructure, the single crystal strength and strain hardening parameters can be inferred from crystal plasticity simulations. This "top-down" approach relates the macroscale mechanical properties to dislocation-precipitate interactions within individual crystals. The "bottom up" approach involves prediction of precipitate structures and fundamental physical parameters such as coherency strain due to lattice mismatch, stacking fault energies, and elastic moduli using density functional theory (DFT) and statistical mechanics-based mixed-space cluster expansion (CE) models. The dislocation-precipitate interaction mechanisms and the temperature and strain rate-dependence of these processes can be predicted using molecular dynamics (MD) and reaction rate theory. Together, the single crystal strength can be obtained from first principles and compared with the ones obtained from crystal plasticity simulations. Once validated, this comprehensive solution to the forward modeling problem provides the basis for the inverse problem represented by goal-driven materials design.

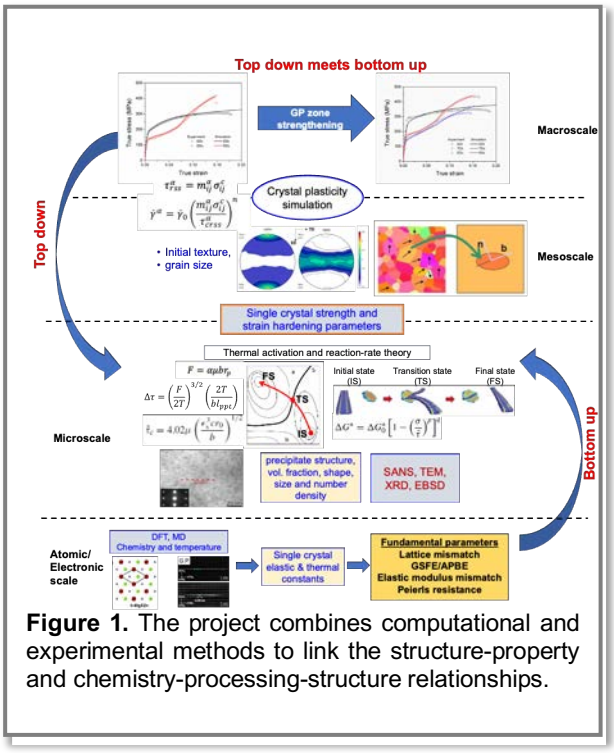


Figure 1. The project combines computational and experimental methods to link the structure-property and chemistry-processing-structure relationships.

Technical Progress

Atomic structures of ordered monolayer Guinier-Preston (GP) zones in Mg-Zn-X (X = Ca, Nd) systems have been predicted based on first-principles calculations combined with the statistical-mechanical approach of cluster expansion (CE). Simulated TEM images and diffraction patterns of the predicted structures agree well with prior experimental observations. Notably, the determined structure is distinct from the one which was proposed and remained uncontested in the literature for nearly two decades.

Comprehensive first-principles calculations based on DFT, CE and Monte Carlo (MC) simulations are being performed to study the crystal structures and stabilities of phases in Mg-Zn system, including GP zones, β_1', β_2' and stable β phase. Mixed-space cluster expansion (MSCE) is generalized to systems with arbitrary crystal symmetry and multiple sublattices by formulating compatible reciprocal space interactions and developing a crystal-symmetry-agnostic algorithm for the calculation of constituent strain energy. A neural network-based approach to describe the interatomic potential of Mg-Zn-Ca is near completion, and the process has already enabled predictions of GP zone stiffness and generalized stacking fault and anti-phase boundary energies. Furthermore, both experimental and simulation-based results have shown that the evolution from GP zone to metastable precipitate to

equilibrium phase involves enrichment in the more slowly diffusing species, Zn.

The age-hardening response of 4 candidate alloys has been characterized, including one in the Mg-Al-Ca-Mn and 3 in the Mg-Zn-Ca-Zr quaternary systems. Overall metallographic investigation of the 4 alloys in various thermomechanical processing conditions has been performed. Advanced TEM/STEM imaging techniques have provided atomic resolution imaging with atomic number contrast. The images have been analyzed using a Python package which uses affine transforms and GPU acceleration to rapidly correct linear drift. Small angle neutron scattering (SANS) investigation has been performed using the GP-SANS instrument at the High-Flux Isotope Reactor (HFIR) at Oak Ridge National Laboratory. A major result of the TEM and SANS investigations so far is the observation of a lower number density of GP zones/precipitates than was previously reported despite similar hardness. This suggests that the precipitates impart a higher resistance on the motion of dislocations.

Continuum mechanical models of the effects of coherency, modulus mismatch, and order strengthening are being developed. Preliminary calculations suggest that coherency strains impact the motion of basal dislocations most strongly, while order strengthening places an additional barrier on the motion of prismatic dislocations. These strength differences unfortunately contribute to the intrinsic plastic anisotropy of Mg alloys.

Future Plans

The atomic structures and thermodynamic stabilities of precipitates in Mg-Al-Ca alloys will be thoroughly investigated with the MSCE and MC, parameterized by DFT. The development of neural network-based interatomic potential will enter the training-testing-revising cycle and the parameter will be finalized. The analysis of the SANS data will be completed to provide information about the GP zone microstructure and precipitation kinetics. The interaction between the dislocations and the GP zones, accounting for the coherency strain fields and mismatch between the elastic modulus will be performed using a FFT based discrete dislocation dynamics code. Low energy (80 kV) atomic resolution imaging and spectroscopy will be used to determine the atomic structure of the GP zones. Electron energy-loss near-edge and extended fine-structure analysis will be performed at different stages of GP zone formation to understand the contribution of alloying elements during early-stage clustering.

Broader Impacts and Workforce Development

The broader impacts of the project are development of computational tools for the prediction of multicomponent alloy structure and performance and a new paradigm for GP zone-strengthened alloy development. The project is also educating a diverse group of graduate students and postdoctoral fellows including members of underrepresented groups: PhD students, Du Cheng and Bassel Khoury; PhD student/postdoc, Eric Hoglund, MEng student, Oluwaniyi Ajiteru (now gainfully employed by Micron Technologies, Manassas, VA); research scientist, Jishnu Bhattacharyya, and postdoc, Kang Wang. Monthly meetings provide the students and postdocs with deep exposure to the collaborative nature of advanced materials development research.

Data Management and Open Access

The developed computational package implementing the MSCE and MC will be made open-source and uploaded to the online platform GitHub. The predicted and experimentally verified atomic structures of GP zones and metastable precipitates are uploaded as supplementary materials of publications. The DFT-generated training set for neural network-based atomic potential will be uploaded to online repositories, such as Materials Cloud.

Advancing Along the Materials Development Continuum and Partnerships to Translation

Based on the predicted atomic structures of GP zones and their thermodynamic stabilities, the diffusivities of solutes, anti-phase boundaries energies, estimated strengthening effect, new alloy compositions with earth-abundant elements and processing parameters have been proposed and will be tested in the coming months.

Publications and References

[1] D. Cheng, E. R. Hoglund, K. Wang, J. M. Howe, S. R. Agnew, B. C. Zhou, *Atomic structures of ordered monolayer GP zones in Mg-Zn-X (X= Ca, Nd) systems*, Scripta Materialia **216**, 114744 (2022).

[2] D. Cheng, K. Wang, B. C. Zhou, *A Comprehensive First-Principles Investigation of the Phase Structures and Stabilities in Mg-Zn Alloys*, submitted to Acta Materialia.

[3] K. Wang, D. Cheng, B. C. Zhou, *Generalization of the mixed space cluster expansion method for arbitrary lattices*, submitted to npj Computational Materials.

DMR2118678: Designer 3D Mesoscale Materials Synthesized in the Self-Assembly Foundry

Lead Investigator: Alfredo Alexander-Katz, aalexand@mit.edu. Caroline A Ross, caross@mit.edu. Vincenzo Vitelli, vitelli@uchicago.edu.
Participating Institutions: MIT, U.Chicago
Website: None yet.
Keywords: dual self-assembly line, multi-layer interconnection, reversible metamaterials, metal infiltration, topological self-assembly

Project Scope

This project (started 9/1/21) is to develop paths to synthesize 3D designer mesoscale materials by self-assembly. It will be accomplished via a new concept: the "dual self-assembly line" that combines artificial intelligence and dissipative particle dynamic (DPD) simulations with physical approaches to making functional mesoscale materials from programmable block copolymer assemblies. The following experimental 3D design will be validated based on "topological self-assembly" and demonstrated by polystyrene (PS)-*block*-poly(2-vinylpyridine) (P2VP) systems in combination with reversible multilayer metallization.

Relevance to MGI

This project will demonstrate the dual assembly line based on a limited number of physical process that may be repeated to build up a 3D nano/mesostructure. At each stage of the line, there is feedback between metrology and modeling. The physical fabrication process is followed from beginning to end, and macroscopic and microscopic information is obtained as input to the computational line as well as to the shared artificial intelligence (AI) engine. By running both physical and computational processes simultaneously, we will train the virtual line to closely reproduce the experimental conditions, enabling optimization and discovery to be performed, and accelerating the process by massively parallelizing the search for a particular target structure. In the past, our team have not only reparametrized the DPD force field to reproduce all the different phases observed in diblock copolymer thin films [1], but we have also demonstrated its use in experimental quantification of processing conditions.[2]

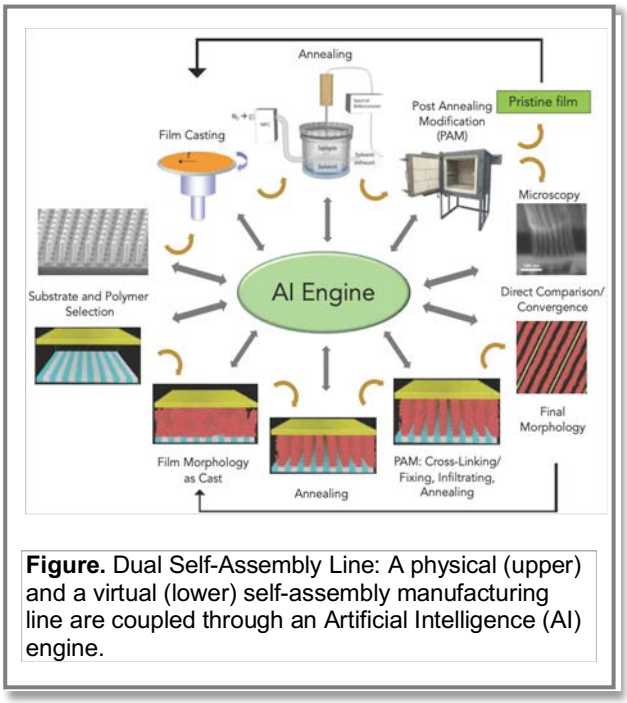


Figure. Dual Self-Assembly Line: A physical (upper) and a virtual (lower) self-assembly manufacturing line are coupled through an Artificial Intelligence (AI) engine.

Technical Progress

Progress in the first year took advantage of a multimechanism directed self-assembly (MMDSA) method we developed [2] to generate complex 3D nanostructures from self-assembly and metallization of a block copolymer (BCP), including well-aligned conductive nanomeshes and multicomponent 2D/3D hybrid metallic nanopatterns. The MMDSA process enables the fabrication of various geometries with a range of compositions via a simple self-assembly process by taking advantage of three different mechanisms, trench wall guidance, edge nucleation and underlayer guidance. Structure-information-propagation relationships in layer-by-layer self-assembly was then revealed when the period of the consecutive two layers is incommensurate and the topographic modulation of the underlayer is large enough, allowing control over the angle between the patterns in each layer.

From the findings of [2] we then developed a reversible metallization method in which metal species (e.g., Pt) infiltrate the BCP thin film to constrain its swelling behavior and subsequent morphological evolution under solvent vapor annealing. This locking effect could be reduced by unloading with appropriate HCl aqueous solution to reversibly remove most of the infiltrated metal species. This facilitates development of programmable metamaterial systems.

We extended layer-by-layer structure propagation to other BCPs. We demonstrated the synthesis and self-

assembly of a high interaction parameter silicon-containing liquid crystalline BCP, where the liquid crystal block provides additional degrees of freedom [3]. By tuning the side chains, a thermally reversible hexagonally packed cylindrical—double gyroid—orthorhombic *Fddd* phase transition along with a liquid crystalline phase transition could be easily induced by a simple temperature change.

Beyond the above-mentioned successive layer-by-layer stacking methods, we discovered a simple intrinsic nanoconfinement strategy, driven by macromolecular design, to generate unprecedented 3D mesh morphologies. [4] We designed triblock bottlebrush polymers with two Janus domains: one perpendicular and one parallel to the polymer backbone. The former enforces a lamellar superstructure that intrinsically confines intra-layer self-assembly of the latter, giving rise to a series of substructures with excellent long-range order.

Besides those experimental achievements, we successfully established a reparametrized DPD simulation model [1] which is powerful in tracking morphology evolution, predicting stable structures and taking out metastable states. It provides a path towards the dual physical and virtual self-assembly line capable of automated decision-making for discovery and optimization of mesoscale designer 3D materials.

Future Plans

In the next year, the group will continue to develop the experimental unit processes needed for the self-assembly line, while refining the DPD modeling and setting up an AI engine to run the virtual self-assembly line. Data from the experimental line (e.g. morphologies as a function of processing conditions; 3D stacking results from different BCPs; reversible metallization processes) will provide the input to the AI engine. Correlations between the virtual and experimental assembly lines will be established.

Broader Impacts and Workforce Development

The proposed research will form the basis for training a new generation of materials scientists where new concepts in assembly and function will be brought to the forefront. The convergence of academia and industry community colleges, computation and experiment, different mentoring perspectives, and the high-level view of the overall self-assembly manufacturing process provides a rich environment for the participants to develop new knowledge, skills and abilities. We have established and 8-10 week in-person internship in summer 2022 for an undergraduate student from Bunker Hill Community College (BHCC). The student will be part of the team and will interact with the PIs, graduate students and industrial collaborators.

Data Management and Open Access

A public website will be established in order to share the experimental data such as TEM images and videos, as well as simulation codes. Apart from this, a platform in the cloud for data storage and sharing will also be developed where we will be able to seamlessly interchange data across our groups, and scientists and engineers at large.

Advancing Along the Materials Development Continuum and Partnerships to Translation

By training our proposed dual self-assembly line, the optimization given by the variational autoencoder could greatly accelerate the process for finding a particular target structure much faster than conventional experiments. This proposal will develop a new, potentially transformative, approach to manufacturing capabilities through coupled computational and experimental research. Our PIs will also engage with an industrial partner, Tokyo Electron Ltd. (TEL), a manufacturer of processing equipment for the semiconductor industry. The principles underlying 3D nanomanufacturing established by this work will thus impact a range of industries including microelectronic device fabrication, 3D printing, NEMS, membrane filtration, photovoltaics and others.

Publications and References

1. H Huang, and A Alexander-Katz, *Dissipative particle dynamics for directed self-assembly of block copolymers*, Journal of Chemical Physics **151 (15)**, 154905 (2019).
2. R Liu, H Huang, Z Sun and A Alexander-Katz, C Ross. *Metallic Nanomeshes Fabricated by Multimechanism Directed Self-Assembly*. ACS nano **15(10)**, 16266-16276(2021).
3. L Weng, M Ma, C Yin, Z Fei, and C Ross, K Yang, and L Shi. *Synthesis and Self-assembly of Silicon-Containing Azobenzene Liquid Crystalline Block Copolymers.* (submitted, 2022)
4. Z Sun, R Liu, T Su, H Huang, K Kawamoto, R Liang, B Liu, M Zhong, and A Alexander, C Ross, J Johnson. *Emergence of mesh-like network morphologies through intrinsic molecular confinement self-assembly.* (submitted, 2022)

Center for PRedictive Integrated Structural Materials Science (PRISMS Center)

Lead Investigator: John Allison (Director), johnea@umich.edu
Participating Institutions: University of Michigan and University of California, Santa Barbara
Website: http://prisms-center.org
Keywords: Integrated Multi-Scale Modeling, Magnesium, Microstructural Evolution, Mechanical Behavior

Project Scope

The primary objective of the PRISMS Center is to develop and deploy a unique scientific platform for accelerating predictive materials science. The platform is comprised of advanced open-source software, an information repository and scientific use cases. The suite of four major open-source multi-scale computational tools is designed to predict microstructural evolution and mechanical behavior of metals. The Materials Commons is a sophisticated, user-friendly information repository and collaboration space. The PRISMS Center platform is demonstrated in scientific use cases which accelerate the quantitative and predictive understanding of magnesium alloys.

Relevance to MGI

The PRISMS Center platform of open-source software, Materials Commons and scientific use cases has been designed to form an important and extensible component of the MGI Materials Innovation Infrastructure. Our scientific use cases demonstrate the value of the integration of simulation, experiment, and theory for accelerating predictive understanding important phenomena of metallic materials.

Technical Progress

Over 4700 unique clones have been made of PRISMS Center software (see Fig. 1) with over 1300 downloads over the past year. The codes are updated with new features annually and represent state-of-art computational efficiency and capability. The four major codes are DFT-FE (real space density functional theory), CASM (Cluster Approach to Statistical Mechanics), PRISMS-Plasticity (crystal plasticity finite element) and PRISMS-PF (phase field). Communities of open-source developers are forming around each code.

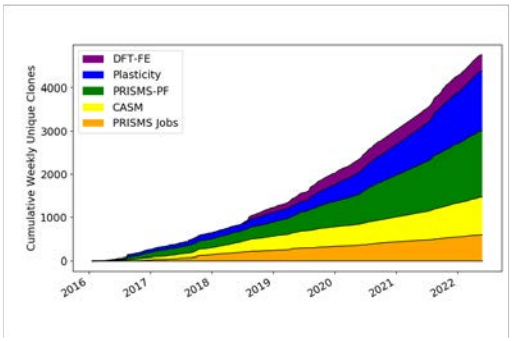


Figure 1. Growing number of cumulative downloads (unique clones) of PRISMS Center open-source software since 2016.

Materials Commons 2.0 debuted in 2021 and represents a complete rework of this information repository which is continually updated with new features. There are over 630 registered users of Materials Commons and it stores over 3 million data files (20TB). Materials Commons use doubled over the past year. Published datasets have been viewed over 14,500 times and downloaded more than 7000 times.

Our fundamental science is conducted in "scientific use cases" which include grain boundary strengthening, precipitate evolution, monotonic and cyclic deformation, recrystallization, corrosion and fatigue behavior. Significant progress has been made in all areas and are documented in over 100 archival papers. In the past year, 19 papers have been published and an additional 11 submitted. Scientific highlights over the past year include:

Grain boundary (GB) strengthening: Used DFT-FE to identify a novel mechanism for GB nucleation of <c+a> dislocations. Used high-resolution EBSD to quantify the local field ahead of a slip band blocked by GB and micro-Hall Petch coefficients. Incorporated parameters into PRISMS-Plasticity simulations and used to study texture.

Cyclic deformation and fatigue: Used PRISMS-Plasticity to model the influence of alloying and texture on cyclic twinning and detwinning. Coupling this with new in-situ SEM (2D) and HEDM (3D) experimental and analytic techniques enabled quantification of alloying effect on CRSS for twinning, detwinning and hcp slip modes. Established twin simulation framework integrating CASM, PRISMS-Plasticity and PRISMS-PF. An advanced fatigue simulation tool, PRISMS-Fatigue, was developed and disseminated in collaboration with Georgia Tech.

Texture evolution: Used a new PRISMS-Plasticity feature and experimental results to predict the influence of Ca,

Zn and Ca/Zn on deformation texture. Determined that alloying with >1wt% Zn+Ca hardens basal slip and decreases basal texture. HEDM experiments of recrystallization phenomena showed that early in recrystallization, grains with a moderately weak basal texture were observed to have a significant and persistent nucleation advantage.

Future Plans: Detailed future plans have been developed for all PRISMS Software, Materials Commons and the scientific use case.

Broader Impacts and Workforce Development

PRISMS Center holds an annual workshop to disseminate our scientific results, establish a collaboration network and train users. Over 320 researchers have attended this workshop and been trained in use of one or more PRISMS Tools. In addition, a PRISMS Center You-Tube channel has been established with training for PRISMS software and have been viewed nearly 6000 times. Training workshops have been conducted at major international conferences most recently at the 2022 TMS ICME World Congress. PRISMS software has also been integrated into graduate courses at UM and Georgia Institute of Technology, as well as short courses at UM and Texas A&M.

Data Management and Open Access

PRISMS Center has four open-source major codes and ancillary codes for use and co-development by the materials community. The Materials Commons information repository functions as a digital laboratory notebook, collaboration platform and public data repository. All data and simulation results from PRISMS Center research are made available to the public via the Materials Commons. Codes and data can be accessed at the following sites:

- DFT-FE: https://github.com/dftfeDevelopers/dftfe
- CASM: https://github.com/prisms-center/CASMcode
- PRISMS-PF: https://github.com/prisms-center/phaseField
- PRISMS-Plasticity: https://github.com/prisms-center/plasticity
- PRISMS-Fatigue: https://github.com/prisms-center/Fatigue
- PRISMS Toolbox: https://github.com/prisms-center/prisms-toolbox
- Materials Commons: https://materialscommons.org

Advancing Along the Materials Development Continuum and Partnerships to Translation

Industrial and national lab researchers have been trained on PRISMS software tools and are using one or more of these codes. The scientific outcomes of our fundamental research projects provide understanding and data (via Materials Commons) that can lead to/used in applied research and development projects as opportunities arise.

Publications and References

1. M Andani, A Lakshmanan, V Sundararaghavan, J Allison, A Misra, "Estimation of micro-Hall-Petch coefficients for prismatic slip in Mg-4Al as a function of grain boundary parameters", Acta Materialia, 226 (2022) 117613.
2. R Roumine, S Lee, T Berman, K Shanks, J Allison, A Bucsek, "Dynamics of recrystallized grains during static recrystallization in Mg-Zn-Ca alloy using in-situ far field HEDM" Acta Materialia, 234 (2022) 118039
3. K Stopka, M Yaghoobi, J Allison and D McDowell, "Simulated effects of sample size and grain neighborhood on the modeling of extreme value fatigue response," Acta Materialia, 224 (2022) 117524.
4. J Yoo, M Adani, A Misra, "Correlating dislocation structures to basal and prismatic slip bands near grain boundaries in tensile-tested Mg–4Al using a multiscale electron microscopy approach", MSEA, 840 (2022) 142993.
5. M Yaghoobi, Z Chen, A Murphy-Leonard, V Sundararaghavan, S Daly, J Allison, "Deformation twinning and detwinning in extruded Mg-4Al: in-situ experiment and crystal plasticity simulation" I J Plasticity, 155 (2022) 103345.
6. J Thomas, A Natarajan, A Van der Ven, "Comparing crystal structures with symmetry and geometry," NPJ Computational Materials, (2021) 7:164.
7. M Higgins, J Kang, G Huang, D Montiel, N Lu, H Liu, Y Shen, P Staublin, J Park, J Almer, P Kenesei, P Sanders, R Suter, K Thornton, and A Shahani, "Anomalous strain-energy-driven macroscale translation of grains during non-isothermal annealing". Physical Review Materials, 5 (2021) L070401.
8. Y Hu, V Menon, L Qi, "Formation of I1 Stacking Fault by Deformation Defect Evolution from Grain Boundaries in Mg", Journal of Magnesium and Alloys, in review, 2022
9. K. Bhattacharya, V. Gavini, M. Ortiz, M. Ponga, P. Suryanaryana, Accurate approximations of density functional theory for large systems with applications to defects in crystalline solids, in Density Functional Theory, Editors: E. Cancés and G. Friesecke, in press, 2022.
10. B Puchala, J Thomas, A Natarajan, J Goiri and AVan der Ven, "CASM – a software package for first-principles based study of multi-component crystalline solids," submitted to Computational Materials Science, 2022.
11. S Das, P Motamarri, V Subramanian, D Rogers, V Gavini, "DFT-FE 1.0: A massively parallel hybrid CPU-GPU density functional theory code using finite-element discretization", in review, 2022.

DMREF: AI-Guided Accelerated Discovery of Multi-Principal Element Multi-Functional Alloys

Lead Investigator: Raymundo Arróyave, rarroyave@tamu.edu
Participating Institutions: Texas A&M University
Website: none
Keywords: Bayesian Optimization, Material Discovery, High throughput, Low Throughput, Shape Memory Alloy
Co-PIs: Ibrahim Karaman, Xiaoning Qian

Project Scope

In this project we aim to discover Multi-Principal Element Multi-Functional Alloys (MPEMFA) with extreme property combinations, such as ultra-high temperature Martensitic Transformations (MTs) with low hysteresis, and stable reversible shape change under stress. Our framework addresses the shortcomings of conventional and existing ML/AI-driven alloy development frameworks as it: (i) employs novel physics-informed ML and data-driven search algorithms to efficiently identify the feasible regions amenable to optimization; (ii) exploits integrative modeling to fuse simulations and experiments to better train ML models for efficient discovery; (iii) develops new Bayesian Optimization (BO) with the consideration of multiple objectives and constraints.

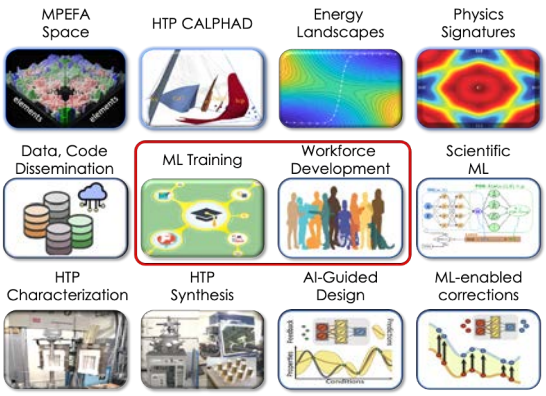


Figure 1. Essential elements of the proposed project aiming to explore, exploit, and understand the Multi-Principal Element Multi-Functional Alloys (MPEMFA) space.

Relevance to MGI

Significant aspects of the proposed project are depicted in **Fig. 1**. They include a combination of computational tools (high-throughput, HTP, determination of phase stability, energy landscapes, and the identification of physics-based signatures for MT), ML/AI methods (including frontier research in Automated Feature Engineering, Bayesian Active Learning for Phase Diagram Identification, Multi-Objective BO, Physics-informed ML), as well as simulation and experimental platforms (including highly integrated HTP synthesis/characterization workflows). These tools are integrated into an AI-driven closed loop. Container technology will be leveraged to disseminate data and ML workflows to ensure Repeatability and Reproducibility, and will deposit datasets in Jarvis and contribute data to the database on Shape Memory Materials Database recently developed by NASA. Research with training/workforce development are performed through the D^3EM program at Texas A&M.

Technical Progress

Materials Discovery: Using a preliminary version of the SMA database that is being developed in this work, a new AI-enabled materials discovery framework has been developed. The information from the database is featurized not only in terms of properties mapped to the composition of the alloy but also incorporating information about thermo-mechanical processing protocols used. The database incorporates information from the literature as well as in-house developed databases. A sophisticated neural network was trained against the data and was used to rapidly explore a small sector of the SMA design space. Using this ML model**, a new alloy with the narrowest hysteresis under loading to date was predicted and confirmed** via synthesis and characterization experiments [**1**].

Framework Development: In the context of accelerated materials discovery, most approaches developed to date rely on Bayesian Optimization (BO). BO provides with an algorithmic approach to solve expensive black box optimization problems. The main ingredients for any BO approach include the type of ML models used to predict the outcome of observation yet to be made, as well as a utility function that evaluates the value of any potential point in the design space relative to the objective at hand. In most BO implementations, the ML models used are based on Gaussian Processes (GPs). In work supported by this award, we have demonstrated that other types of ML models can be more effective than GPs, particularly when the design space is complex and multi-dimensional [**2**].

Future Plans

To navigate the vast chemical space associated to MPEMFA a new framework will be developed that (i) employs novel physics-informed ML to efficiently identify the feasible regions amenable to optimization; (ii) fuses simulations and experiments to obtain efficient ML models; (iii) develops new Batch (parallel) Bayesian Optimization (BO) strategies to make globally optimal iterative experimental design; and (iv) is capable of simultaneously considering multiple objectives and constraints. The aim is to go beyond accelerated discovery, seeking to address questions about the underlying factors responsible for the multi-functional behavior in MPEMFAs. To efficiently identify optimal MPEMFAs, we will use a High Throughput (HTP) synthesis and characterization approach. This approach will yield 1) novel, optimized MPEMFAs through physics-informed ML approaches and 2) the most extensive, highest quality SMA/MPEMFA database to date. This workflow will use synthesis and characterization methods that correlate best with scalable, bulk materials to avoid size effects on MTs.

Broader Impacts and Workforce Development

The PI established the Data-Enabled Discovery and Development of Energy Materials (D^3EM) program, supported by the NSF NRT program, DGE-1545403. The D^3EM program, already institutionalized as an official Interdisciplinary Graduate Certificate at Texas A&M, is one of the very few interdisciplinary programs in the US that directly address the workforce development needs to be brought about by the MGI. To date, 50 Ph.D. and 2 MS students from seven academic departments have been enrolled in the program. Students participating in this project will be involved in the D^3EM program. Moreover, data generated from this project will be used in courses associated with the D^3EM curriculum, including a Materials Informatics course. Another strategy for integrating the proposed research and our workforce development efforts is through the Materials Design Studio (MDS) course created and taught by the PI. MDS is a unique project-based course modeled after Engineering UG capstone design courses in which students solve a real-life materials discovery or design problem by leveraging ML/AI-based frameworks. Throughout the project, the PI and Co-PIs will formulate challenging issues derived from the work carried out in this project for students to take on as final projects in MDS.

Data Management and Open Access

We are proposing several mechanisms to disseminate the data and code generated as part of this work, as detailed in the Data Management Plan. Our plans include, for example, to contribute to the DFT and experimental data repositories in the new Joint Automated Repository for Various Integrated Simulations (JARVIS) for data-driven materials design. While sharing data and meta-data in the spirit of the FAIR principles is a minimum requirement of the DMREF program, we recognize that data alone is not sufficient to create vibrant MGI communities, and code and analysis workflows must be part of the dissemination ecosystem. We will leverage container-based platforms to disseminate code as well as representative data. Specifically, we will partner with Code Ocean, a cloud-based research collaboration platform that enables researchers to create end-to-end workflows geared towards reproducibility Our discovery framework and expertise, once fully developed, can be utilized in other materials systems. Thus, the impact of the proposed research will be widespread and go beyond the proposed MPEMFAs, including other MPE alloys.

Advancing Along the Materials Development Continuum and Partnerships to Translation

The PIs have experience in commercializing their designed alloys through start-up companies, and thus, they are familiar with engineering and translational challenges. In the last 15 years, they have been collaborating with industry, and government labs to bring their discoveries to deployment. Notably, in the past decade, two startup companies have been created by Co-PI Karaman's former students based on discoveries made in the Shape Memory Alloy space. The PI (Arróyave) and Co-PI (Karaman) have joint IP on the control of microstructure in SMAs to achieve Invar effect. In this DMREF project the impact of composition complexity on Invar and Elinvar effects will be one of the specific research thrusts to be investigated.

Publications supported by this work

1. Trehern, W., et al. "Data-driven shape memory alloy discovery using Artificial Intelligence Materials Selection (AIMS) framework." Acta Materialia 228 (2022): 117751.
2. Lei, Bowen, et al. "Bayesian optimization with adaptive surrogate models for automated experimental design." npj Computational Materials 7.1 (2021): 1-12.

Molecular Automated Discovery of Novelty Enabled by Synthetic Systems (MADNESS)

Lead Investigator: Alán Aspuru-Guzik, University of Toronto
Co-Lead Investigators:
Martin Burke, University of Illinois, Urbana-Champaign
Leroy Cronin, University of Glasgow
Bartosz Grzybowski, Allchemy Inc.
Jason Hein, University of British Columbia
Participating Institutions: University of Toronto; University of Illinois, Urbana-Champaign; University of British Columbia; University of Glasgow; Allchemy Inc.
Website: None
Keywords: Self-driving lab; Organic laser materials; Artificial intelligence; Automated experiments; Closed-loop discovery

Project Scope

MADNESS aims to accelerate the discovery of molecular materials through the development of self-driving labs: automated experimentation guided by machine learning (ML). To achieve this goal, we are developing AI algorithms for molecular design and experiment planning, software for robotics control and data management, as well as (iterative) coupling schemes for efficient molecular synthesis. Using this technology, we aim to discover novel high-performing emitter materials for organic solid-state laser devices (OSLD) and identify the underlying design principles along the way – demonstrating that our SDL platforms are able to identify dozens of new high-performing materials by synthesizing and characterizing them at higher throughput and lower cost.

Relevance to MGI

The MADNESS OSLD materials discovery workflow integrates AI recommendations for synthesizing candidate laser molecules with automated synthesis, analysis and purification, and characterization experiments. The data obtained from these experiments is automatically uploaded to a central database accessible for the AI and all the groups within our MADNESS team. Consequently, we have been able to set up a closed-loop, asynchronous and delocalized molecular discovery campaign that is distributed across multiple countries, time zones and continents. Additionally, by having replicates of some of our synthesis machines at different sites, we are able to validate that synthesis conditions and outcomes are translatable between labs and different automated synthesis platforms. The software and hardware technologies that we are developing are open source with the goal of being reusable by other groups and researchers. Achieving success in discovering state-of-the-art organic laser molecules will demonstrate the potential of these spatially and temporally delocalized laboratories for accelerating materials discovery. We are working toward the rapid discovery of organic laser molecules with state-of-the-art optoelectronic properties such as optical cross-section gain. By analyzing the outcomes obtained by our self-driving labs, we are also able to learn important fundamental design principles for e.g. organic laser molecules, as described in our recent preprint[1].

Technical Progress

Our self-driving labs are up and running, and we anticipate further publications in the near future. To date, we have published a number of papers and preprints related to the MADNESS project. Below are summaries of each.

- Materials acceleration platform for organic laser discovery[1]: We recently published a preprint describing the automated synthesis and characterization capabilities that we have developed for our self-driving lab. We used these platforms to identify two novel state-of-the-art organic laser molecules.
- RouteScore[2]: We developed a framework for evaluating the cost of combined manual and automated synthetic routes. This framework was used in a multi-objective optimization of potential laser molecules to

show that it can find molecules that are easier to synthesize without sacrificing important optoelectronic properties.

- Computer algorithm for discovering iterative reaction sequences[3]: We designed a computer algorithm that identifies reactions and functional groups which enable iterative reaction sequences – similar to those used by nature to synthesize complex molecules such as proteins, sugars and DNA/RNA. We showed that a small set of simple reactions can be used to synthesize highly complex molecules.
- Automated synthesis with real-time feedback[4]: We developed an automated synthesis platform that performs actions based feedback from real-time reaction monitoring. This provides automated synthetic chemistry with the opportunity to move away from the traditional "dump and stir" approach to more targeted and efficient synthetic methods based on real-time monitoring of the reaction.
- Platform for discovery of phosphorous catalyst ligands (Kraken)[5]: We embedded data-driven workflows in a library of potential catalysts to build predictive models for catalyst performance and design. Using these workflows, we were able to predict the properties of over 300,000 new catalyst ligands, which could be readily employed in our iterative cross-coupling sequences.
- Digitizing chemistry through custom reactionware[6]: We combined process chemistry principles with custom 3D-printed reactionware to demonstrate a new way of digitizing chemistry that provides an unambiguous link between experiments and the code used to generate experimental conditions.
- Limitations of ML for reactivity prediction[7]: We demonstrated that the abundance of data within the chemical literature may be insufficient to build accurate and predictive ML models of chemical reactivity. Using the example of Suzuki-Miyaura coupling reactions, we showed that a variety of ML methods fail to perform significantly better than simply relying on the frequency of conditions within the literature.
- Autonomous reactivity optimization[8]: We integrate our Bayesian optimization algorithms with a Chemspeed automated synthesis platform to identify ligands that simultaneously improve the yield and stereoselectivity of a challenging Suzuki-Miyaura coupling reaction.
- Automated solubility screening using computer vision[9]: We demonstrated a robotic platform for closed-loop solubility screening, with automated solvent titration informed by monitoring solution turbidity with computer vision.

Future Plans

Our objectives for the future are to demonstrate that self-driving labs can accelerate the development of a new class of materials, in this case for organic semiconductor laser devices. We are partnering with collaborators outside this grant to test our most promising materials in solid-state devices. To achieve this aim, we are developing frameworks for asynchronous, cloud-based, delocalized synthesis and characterization guided by ML algorithms. Additionally, many enabling technologies must be developed to realize our self-driving labs. Some of the requisite technologies include: robust coupling conditions for iterative Suzuki-Miyaura cross-coupling reactions, efficient automated experimentation workflows and protocols, online monitoring and feedback of reaction progress, and a high-level ontology that can be used to replicate syntheses between different laboratories and automated synthesis platforms.

Broader Impacts and Workforce Development

This project provides training to the next generation of scientists from a wide variety of backgrounds. Our students and postdoctoral fellows come from many different fields including both experimental and computational chemistry, electrical engineering, materials science, and computer science. Through this collaboration, they are learning how to work at the interface of disparate fields and effectively communicate with colleagues who do not share their technical background. This is a critical skill for communicating science to the public at large, as well as within the workforce with colleagues who do no share their area of expertise. Most importantly, our student and postdoctoral fellows are learning how to design, build and execute self-driving labs, which we regard as the technology of the future for research and development. Furthermore, the results of our research have been presented at a number of conferences, as well as the Kavli Lecture of the 2022 Spring Meeting of the ACS.

Data Management and Open Access

Free access to code and data is crucial to enhancing the repeatability and reproducibility of science. As we publish papers and release software, we are making the data and code freely accessible. Free access to code through common community tools such as GitHub and GitLab, and access to data through the same repositories, or through well-known data repositories such as Zenodo, makes it easy for other members of our community to re-use our code and data, and verify or build upon our results.

Code and data repositories:

- Database for cloud-based, asynchronous, and delocalized discovery of laser molecules, built on top of molar: https://github.com/aspuru-guzik-group/molar
- RouteScore:
    - Code: https://github.com/aspuru-guzik-group/routescore
    - Data: https://zenodo.org/record/5106659
- Web application for iterative reaction sequences: https://iterator.allchemy.net/
- Code for adaptive auto-synthesizer and automated solubility screening are available at https://gitlab.com/heingroup
- Kraken:
    - Web application: https://kraken.cs.toronto.edu/
    - Kraken code: https://github.com/aspuru-guzik-group/kraken
- ChemSCAD: https://pypi.org/project/chemscad/

Advancing Along the Materials Development Continuum and Partnerships to Translation

Lasers are used across a wide range of industries including manufacturing, medicine, communication, scientific research, and a variety of consumer technologies. OSLDs also have potential in display technologies, where their narrower emission spectra will enable better display colors. However, little is known about the molecular design principles of these materials. We are in the process of demonstrating that our self-driving labs can accelerate the development of this class of materials by substantially reducing the cost and time required to synthesize and characterize new molecules. Additionally, we can more efficiently optimize molecular properties using automated experiments guided by ML algorithms. Finally, we can use the workflows we develop as a blueprint to accelerate the discovery of materials for technologies beyond OSLDs, such as organic redox flow batteries.

In the process of building our self-driving labs, we collaborate with a number of automation manufacturers including Chemspeed Technologies, North Robotics, Mettler Toledo, and others. We are also working to commercialize self-driving lab technologies in the form of start-ups out of our groups such as DeepMatter, Kebotix and Telescope Innovations.

Publications and References

1. T. C. Wu, A. A. Granda, K. Hotta, S. A. Yazdani, R. Pollice, J. Vestfrid, H. Hao, C. Lavigne, M. Seifrid, N. Angello, F. Bencheikh, J. E. Hein, M. Burke, C. Adachi, A. Aspuru-Guzik, A Materials Acceleration Platform for Organic Laser Discovery (2022), doi:10.26434/chemrxiv-2022-9zm65.
2. M. Seifrid, R. J. Hickman, A. Aguilar-Granda, C. Lavigne, J. Vestfrid, T. C. Wu, T. Gaudin, E. J. Hopkins, A. Aspuru-Guzik, Routescore: Punching the Ticket to More Efficient Materials Development. *ACS Cent. Sci.* **8**, 122–131 (2022).
3. K. Molga, S. Szymkuć, P. Gołębiowska, O. Popik, P. Dittwald, M. Moskal, R. Roszak, J. Mlynarski, B. A. Grzybowski, A computer algorithm to discover iterative sequences of organic reactions. *Nat Synth*. **1**, 49–58 (2022).
4. J. Liu, Y. Sato, F. Yang, A. J. Kukor, J. E. Hein, An Adaptive Auto-Synthesizer using Online PAT Feedback to Flexibly Perform a Multistep Reaction. *Chemistry–Methods*, e202200009 (2022).
5. T. Gensch, G. dos Passos Gomes, P. Friederich, E. Peters, T. Gaudin, R. Pollice, K. Jorner, A. Nigam, M. Lindner-D'Addario, M. S. Sigman, A. Aspuru-Guzik, A Comprehensive Discovery Platform for Organophosphorus Ligands for Catalysis. *J. Am. Chem. Soc.* **144**, 1205–1217 (2022).
6. A. Bubliauskas, D. J. Blair, H. Powell-Davies, P. J. Kitson, M. D. Burke, L. Cronin, Digitizing Chemical Synthesis in 3D Printed Reactionware. *Angewandte Chemie International Edition*. **61**, e202116108 (2022).
7. W. Beker, R. Roszak, A. Wołos, N. H. Angello, V. Rathore, M. D. Burke, B. A. Grzybowski, Machine Learning May Sometimes Simply Capture Literature Popularity Trends: A Case Study of Heterocyclic Suzuki–Miyaura Coupling. *J. Am. Chem. Soc.* **144**, 4819–4827 (2022).

8. M. Christensen, L. P. E. Yunker, F. Adedeji, F. Häse, L. M. Roch, T. Gensch, G. dos Passos Gomes, T. Zepel, M. S. Sigman, A. Aspuru-Guzik, J. E. Hein, Data-science driven autonomous process optimization. *Commun Chem*. **4**, 1–12 (2021).
9. P. Shiri, V. Lai, T. Zepel, D. Griffin, J. Reifman, S. Clark, S. Grunert, L. P. E. Yunker, S. Steiner, H. Situ, F. Yang, P. L. Prieto, J. E. Hein, Automated solubility screening platform using computer vision. *iScience*. **24**, 102176 (2021).
10. L. Wilbraham, S. H. M. Mehr, L. Cronin, Digitizing Chemistry Using the Chemical Processing Unit: From Synthesis to Discovery. *Acc. Chem. Res.* **54**, 253–262 (2021).
11. Y. Shi, P. L. Prieto, T. Zepel, S. Grunert, J. E. Hein, Automated Experimentation Powers Data Science in Chemistry. *Acc. Chem. Res.* **54**, 546–555 (2021).
12. R. Pollice, G. dos Passos Gomes, M. Aldeghi, R. J. Hickman, M. Krenn, C. Lavigne, M. Lindner-D'Addario, A. Nigam, C. T. Ser, Z. Yao, A. Aspuru-Guzik, Data-Driven Strategies for Accelerated Materials Design. *Acc. Chem. Res.* **54**, 849–860 (2021).

Theory Aware Machine Learning for Polymers

Lead Investigator: Debra J. Audus, debra.audus@nist.gov
Participating Institutions: National Institute of Standards and Technology
Website: https://github.com/usnistgov/TaML
Keywords: polymers, machine learning, transfer learning, theory

Project Scope

Machine learning as applied to polymer physics has recently shown immense progress---mostly in areas where there are existing large datasets or where datasets can be generated quickly. However, there are numerous interesting problems where the dataset sizes are too small, improved extrapolation is desired or the need to understand the physics behind the machine learning prediction is essential. Here, we tackle these three challenges by incorporating domain knowledge, in the form of polymer theories, into machine learning models. Success is measured via demonstrations of improved interpolation and extrapolation for small dataset sizes when compared to machine learning models without domain knowledge.

Relevance to MGI

The goal of the MGI is to bring new materials to market faster and at a fraction of the cost. Machine learning relying computational data, experimental data or both has shown much promise towards this goal for data rich problems. However, there are many problems of interest where such data is not available or improved extrapolation is essential. In these cases, the integration of theory into the machine learning models can improve interpolation (allowing for more accurate models), extrapolation (so new materials can be identified faster), and explainability (providing concrete learning lessons to further accelerate the process). Our initial efforts utilize computational data, but future work will rely on experimental data and then both experimental and simulation data. Ultimately, we are interested in simultaneously leveraging experimental data, simulation data and theory, with the theory directly embedded in the machine learning.

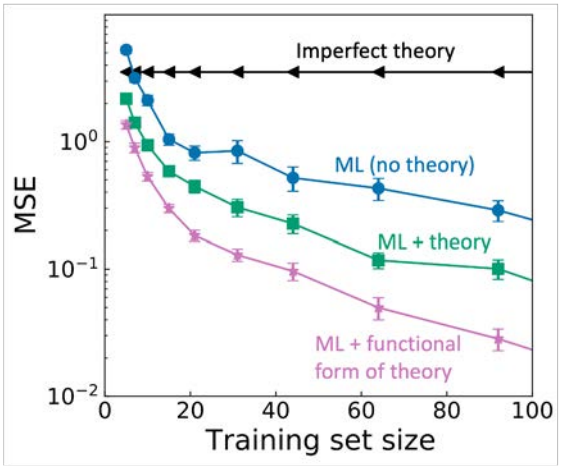


Figure 1. Performance as measured by mean squared error (MSE) of two methods for incorporating theory into machine learning along with benchmarks of no theory and no machine learning. Test dataset is composed entirely of outlier data.

Technical Progress

We explored a variety of different methods for incorporating theory into machine learning in conjunction with three different machine learning models. We started with the canonical problem of predicting the size of a single polymer chain in solution as a function of the chain length and solvent quality. Specifically, we made use of simulation to generate a training dataset, as well as a testing dataset composed entirely of outliers. Using the mean squared error as a function of data set size as a metric for performance, we demonstrated that for a given machine learning model, the incorporation of an imperfect theory improved performance under both interpolation and extrapolation (see Fig. 1). However, the best method is dependent on the machine learning model. As expected, we also find that some machine learning models perform better than others. We highlight the use of Gaussian Process Regression (GPR) with heteroscedastic noise as it incorporates known uncertainties in the training data and then correctly propagates them to yield good predictions with more accurate uncertainty estimates. For GPR with heteroscedastic noise, we also explore incorporating not only numeric values of theory, but also the functional form explicitly. This resulted in further improved performance---in some cases, requiring half as much data to achieve the same uncertainty compared to a machine learning model without any knowledge of theory (see Fig. 1). We also explored explainability by either using the

theory as a benchmark or understanding how physical parameters change as a function of input.

Future Plans

We plan to extend our work to experimental data where additional complicating factors such as larger uncertainty are present, as well as more complicated theories including more parameters. In the process, we will also test other methods for including theory that are leverage the structure of these more complicated theories.

Broader Impacts and Workforce Development

A key output of our work is data and code that are publicly available. More specifically, we focused on creating user-friendly, stand-alone Jupyter notebooks that illustrate the key concepts such that others can build upon our work. The PI also has a history of co-organizing machine learning workshops (2020 American Physical Socitey Division of Polymer Physics Short Course and the 2020 Tutorial Workshop: Machine Learning in Materials Science hosted by the Data Science Institute, Columbia University).

Data Management and Open Access

A software repository (https://github.com/usnistgov/TaML) that provides Jupyter notebooks illustrating key concepts and the source code to test improved interpolation and extrapolation is publicly available. All the testing data, training data, and machine learning performance data will be available on NIST's Science Data Portal (https://data.nist.gov/sdp/#/) soon.

Advancing Along the Materials Development Continuum and Partnerships to Translation

The outcomes of this project focus on methods for incorporating theory into machine learning that others can easily adopt to achieve MGI goals. Thus, we have focused on making those methods easily accessible to others especially in the form of Jupyter notebooks that can trivially be run using Google Colab. Interested parties, including those from industry, can prototype to see if the methods we developed may be a good fit for their use-case.

Physics-informed meta-learning for design of complex materials

Lead Investigator: Stephen Baek, baek@virginia.edu
Participating Institutions: University of Virginia, University of Iowa
Website: None
Keywords: Physics-informed machine learning, meta-learning, multi-scale, AI, materials by design.

Project Scope

This multidisciplinary project will harness recent innovations in artificial intelligence (AI) to establish a novel design and discovery cycle for complex materials that will dramatically accelerate material innovations. A wide class of complex materials, including solid-solid composites, porous solids, foams, biological materials, etc., have complex microstructures, which play a dominant role in determining their properties. We will create new methodologies through which human scientists and AI will collaborate to discover optimal microstructural designs of such materials for targeted properties and performance. A particular focus will be on energetic materials (EM), which are the archetype of materials with strong microstructure influence.

Relevance to MGI

Our long-term vision is to achieve disruptive innovation for material discovery, through convergence of computational design and AI. The scientific premise is that microstructure and geometry at meso-scale—as much as chemistry at a molecular/atomic level—play a critical role in the properties and performance of a material. Hence, advanced machine cognition can effectively discern the interplay of complex microstructural features to assimilate structure-property-performance linkages, and design novel microstructures with engineered characteristics. Our team has demonstrated that AI can assimilate surrogate models for multiscale simulations,[1] generate synthetic microstructures with controlled morphologies,[2] learn linkages between geometry and physics,[3] and design material microstructures for targeted properties.[4] As a next step in this thread of work, this project promises rich scientific advances in data driven design of next-generation EMs. Albeit the focus on EMs, this project will deliver techniques and tools generally applicable to a broad range of complex materials.

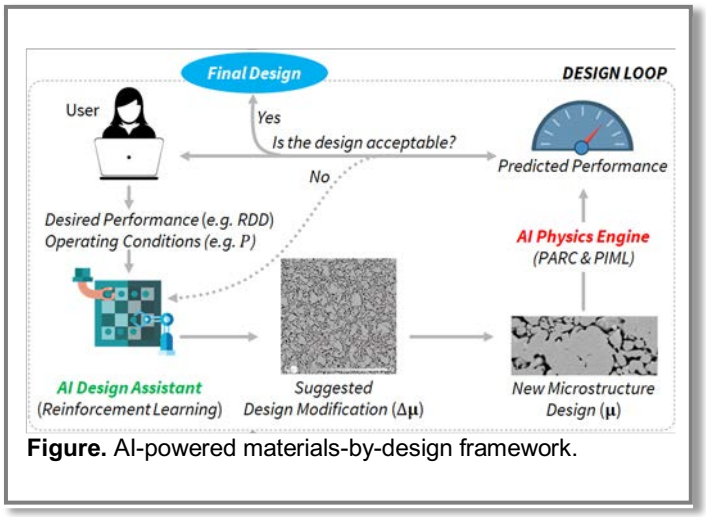


Figure. AI-powered materials-by-design framework.

Technical Progress

This is a new project aiming to build the AI-driven design and discovery framework illustrated in the above Figure. We will develop new techniques to extend our previous work on one particular EM species, HMX, to all known species of CHNO based EMs, encompassing materials critical to DoD and DoE, such as PETN, TNT, RDX, TATB, and CL-20. However, such a leap from a single material (HMX) to the comprehensive space of CHNO EM species presents foundational challenges. For example, the physics-aware recurrent convolution (PARC) model for HMX[1] from our previous work has demanded years of computational time to generate a large ensemble of simulation data. Repeating this process for all other species of CHNO explosives is not a feasible enterprise (problem of expensive data). Instead, we will find a route to use the HMX-trained model as a springboard to accelerate development of burn models for new species. Moreover, we will make significant advances in the deep learning model itself, as the current approach requires toggling of parameters in the model architecture, making them less reliable (problem of robustness) and difficult to interpret (problem of interpretability).

Future Plans

Solving these problems will require (1) a transformative shift from the current status of applying "off-the-shelf" AI models by making fundamental breakthroughs driven by materials research use cases, and (2) close interactions between EM modeling expert (Udaykumar), AI/ML expert (Baek), and manufacturing expert (Song). In this project, we will bring to bear extensive previous work and complementary expertise to develop, test, and validate a new framework called physics-informed meta-learning (PIML) for the design of EMs. The project is organized into three core scientific research thrusts: 'Thrust 1: Establish the CHNO energetic material space and training/testing datasets,' 'Thrust 2: Physics-informed meta-learning from small, sparse, and multi-fidelity datasets,' 'Thrust 3: Validation and loop closure.'

Broader Impacts and Workforce Development

We will engage with the AFOSR and AFRL to initiate a student exchange program between the university research groups and one or more of their research units. Regular meetings will be organized with the researchers at AFOSR and AFRL to form a critical iterative feedback loop. These meetings will also facilitate translation of the proposed findings of this fundamental research projects closer toward real-world applications critical to DoD missions. We will also develop tutorial programs to serve students and researchers in academia, national laboratories, Air Force and DoD research labs, and broader materials research communities to help them acquiring new AI-driven research methodologies and computational techniques. We will focus on delivering the core technical concepts in AI and scientific ML by using our open source software library 'MaterIQ' which will be developed from this project. We will build upon actual use case examples from our research to develop PowerPoint slides, handouts, and exercises, alongside hands-on coding examples. These tutorial programs will be presented in conferences (e.g. the shock conference, ASME meetings, JANNAF meetings), as well as self-hosted workshops.

Data Management and Open Access

In this DMREF project, a great emphasis will be put on creating a shared data platform among the energetic materials (EM) research community. Our goal in terms of data management is to create a data repository where the design, manufacturing, and experimental data generated from the proposed research activities are stored and shared among EM researchers. Due to the sensitive nature of EM data and their criticality for the national security, we will closely collaborate with the AFOSR/AFRL researchers to implement a safe and reliable EM database.

Advancing Along the Materials Development Continuum and Partnerships to Translation

The new AI-powered materials-by-design loop will speed up material innovation by multiple orders of magnitude. Advanced machine cognition can effectively discern the interplay of complex microstructural features and design novel microstructures with engineered characteristics. The development of an open source software library 'MaterIQ' will accelerate the AI & data integration in the materials research community. It will benefit the materials community by supplying advanced computational frameworks to discover, optimize, and design material systems multiple times faster and at a fraction of the cost.

Publications and References

1. Phong C.H. Nguyen, Joseph B. Choi, Yen-Thi Nguyen, Pradeep K. Seshadri, H.S. Udaykumar, and Stephen Baek, *Physics-Aware Recurrent Convolutional (PARC) Neural Networks to Assimilate Meso-scale Reactive Mechanics of Energetic Materials*, ArXiv Preprint, arXiv:2204.07234 (2022).
2. Sehyun Chun, Sidhartha Roy, Yen Thi Nguyen, Joseph B. Choi, H. S. Udaykumar, and Stephen S. Baek, *Deep learning for synthetic microstructure generation in a materials-by-design framework for heterogeneous energetic materials*, Scientific Reports **10**, 13307 (2020).
3. Zhiyu Sun, Ethan Rooke, Jerome Charton, Yusen He, Jia Lu, and Stephen Baek, *ZerNet: Convolutional Neural Networks on Arbitrary Surfaces Via Zernike Local Tangent Space Estimation*, Computer Graphics Forum **39**, 204-216 (2020).
4. Phong C. H. Nguyen, Nikolaos N. Vlassis, Bahador Bahmani, WaiChing Sun, H. S. Udaykumar & Stephen S. Baek, *Synthesizing controlled microstructures of porous media using generative adversarial networks and reinforcement learning*, Scientific Reports **12**, 9034 (2022).

CSSAS: The Center for the Science of Synthesis Across Scales EFRC

Lead Investigator: François Baneyx, baneyx@uw.edu
Participating Institutions: University of Washington, Pacific Northwest National Laboratory, University of Chicago, University of California, San Diego, Oak Ridge National Laboratory
Website: https://www.cssas-efrc.com/
Keywords: order, hierarchy, hybrid materials, proteins, peptoids

Project Scope

Hierarchical materials exploit the coupling between heterogeneities distributed across length scales to harness emergent properties that are unavailable in homogeneous materials. By making it possible to control the transport of charge, mass and spin, the localization of sequential chemical reactions, and dissipative responses to external stimuli, hierarchical materials will enable dramatic advances in the complex functions required for energy production, storage, and use. Inspired by the complex functionality of biomolecular materials, the Center for the Science of Synthesis Across Scales (CSSAS) seeks to achieve rational design of hierarchical materials for complex energy functions based on a predictive understanding of materials synthesis across scales. CSSAS's three interconnected, hypothesis-driven research thrusts are underpinned by a clear set of scientific goals and uniquely enabled by the EFRC approach. *Thrust* 1 aims at predicting how the chemical and sequence information of defined building blocks translates into the emergence of order and outcomes of assembly. *Thrust 2* seeks to master the free energy landscapes of disparate building blocks in complex environments and at surfaces to control assembly into hierarchical units, and inorganic morphogenesis and organization. *Thrust 3* aims at integrating data science, *in situ* characterization, and simulations to adaptively control synthetic outcomes and access metastable states of matter.

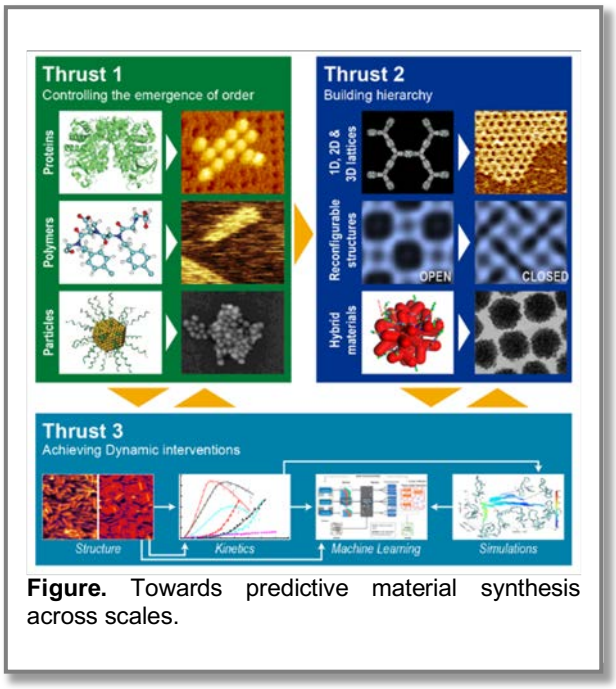


Figure. Towards predictive material synthesis across scales.

Relevance to MGI

CSSAS was established in 2018 under the auspices of the DOE-BES EFRC program. It is housed at the University of Washington and operates in close integration with Pacific Northwest National Laboratory, Oak Ridge National Laboratory, the University of California San Diego, and the University of Chicago.

Technical Progress

Building on an existing array of collaborations and forging new synergies between key personnel, postdoctoral research associates and graduate students, the CSSAS team is developing the ability to use the details of molecular structure to predict how order emerges in macromolecular systems, how hierarchy develops from complex and disparate building blocks, and how to intervene dynamically to capture non-equilibrium states and redirect macroscopic outcomes of synthesis. Beyond shape complementarity and packing considerations, CSSAS seeks to understand, codify, and generalize how ionic, solvent and surface effects, along with designed interactions and conformational flexibility, drive emergent phenomena in the self- and co-assembly of high-information content building blocks. Major breakthroughs achieved over the past four years include: (1)

groundbreaking progress in understanding how sidechain chemistry, solution conditions, and interfacial properties influence the self-assembly of proteins and peptoids; (2) a growing understanding of the molecular interactions that underpin the behavior of *de novo*-designed, site-modified, and solid-binding proteins at interfaces with profound implications for self-assembly, biomineralization, and structural, opto- electronic, catalysis and charge and momentum transfer applications; (3) the synthesis of next-generation precision materials from high information content building blocks; and (4) major advances in theory and machine learning (ML) that will enable dynamical reconfigurations of materials and access to non- equilibrium or metastable states of matter.

Future Plans

Exploiting the EFRC mechanism at its fullest, we will merge theory, computation, experiments, ML, and characterization to advance a unifying mission of harnessing the complex functionality of hierarchical materials by mastering the design of high information content macromolecular building blocks that predictively self-assemble into responsive, reconfigurable materials, and/or direct the formation and organization of inorganic components. Major research goals include. (**RG1**) Determine the molecular-scale distribution and response of sidechains, solvent, and ions in the interfacial region of approaching building blocks, and atomic and nanoscale inorganic components targeted for assembly or directed nucleation. Understand how these distributions and responses give rise to the resulting interaction potentials that orchestrate materials formation at different scales. (**RG2**) Realize 2D and 3D hierarchical and hybrid materials by understanding how the interplay of interactions between disparate blocks, surfaces, solvent, and electrolytes determined in RG1 defines the energy landscapes across which hierarchy develops and inorganic nucleation proceeds. (**RG3**) Achieve adaptive control of synthetic outcomes and access non-equilibrium and metastable states of matter by integrating the tools of data science with *in situ* characterization and simulations, and by using external fields and localized changes in solution chemistry.

Broader Impacts and Workforce Development

CSSAS is a vibrant research community in which every individual is valued. Graduate students and postdocs (collectively referred to as trainees) are routinely co-advised, both formally and informally, by two or more key personnel. Trainees are encouraged to connect with their peers within and across sites, and to develop scientific, leadership and communication skills through a variety of programs and activities. These include professional development resources and opportunities available at each institution (e.g., the UW Office of Postdoctoral Affairs), a student-led *Journal Club*, and presentations and Q&A sessions at CSSAS *Biomineralization & Hybrid Materials*, *Assembly & Dynamic Interventions*, *All-Hands*, and *Annual Retreat* meetings (multiple trainees often work together to craft compelling team science talks). Poster and oral presentations at professional societies meetings (MRS, ACS, AIChE, Gordon Research Conferences…) and at biannual EFCR PI meetings foster the practice of soft skills, while short talks and brainstorming sessions during focused *Project meetings* hone scientific acumen, intellectual curiosity and an appreciation for interdisciplinarity. Team-building activities such as hikes and social outings cement relationships and instill a sense of belonging and greater purpose. Beyond core scientific endeavors, trainees are encouraged to practice cultural competence harnessing the offerings of all institutions involved, and to produce content (e.g., video, podcasts, art) to educate the public, contribute to the nationwide *EFRC Community Website,* and participate in network-level activities (e.g., service on the editorial board and producing content for the *Frontiers in Energy Research Newsletter*).

Data Management and Open Access

CSSAS generates a variety of research data that can be stored digitally and accessed electronically, such as charts, figures, images, and the underlying data used to generate them: codes, software, and algorithms used to generate or analyze research data; text; numeric information; images; video; audio; and associated metadata. Data is preserved and shared to enable validation through reproduction of calculated results using publicly available software, reproduction of experimental results for standard samples, re- analysis of data to verify findings, and comparison with new, and complementary experimental or calculated results to expand understanding.

Advancing Along the Materials Development Continuum and Partnerships to Translation

N.A.

Publications and References

CCSAS publications can be found at: https://www.cssas-efrc.com/publications

Microstructure by Design: Integrating Grain Growth Experiments, Data Analytics, Simulation and Theory

Lead Investigator: Katayun Barmak, kb2612@columbia.edu.
Participating Institutions: Columbia Univ., Illinois Inst. Technol., Lehigh Univ., Univ. of Utah
Website: None.
Keywords: Grain Growth, Theory, Experiments, Numerical Simulation, Modeling and Analysis, Data Analytics.

Project Scope

The principal objective of this project is to quantify statistical measures of grain growth obtained from experimental data garnered from metallic films and to employ these measures to inform mesoscale models of grain growth. This link forged between experiment and theory via analysis and simulations will result in improved models of grain coarsening that properly reflect the underlying physics. Data analytics will be employed to study the evolution of experimental and computational microstructures, regarded here as a collection of interacting grain triple junctions, to validate and refine the models and guide future experiments.

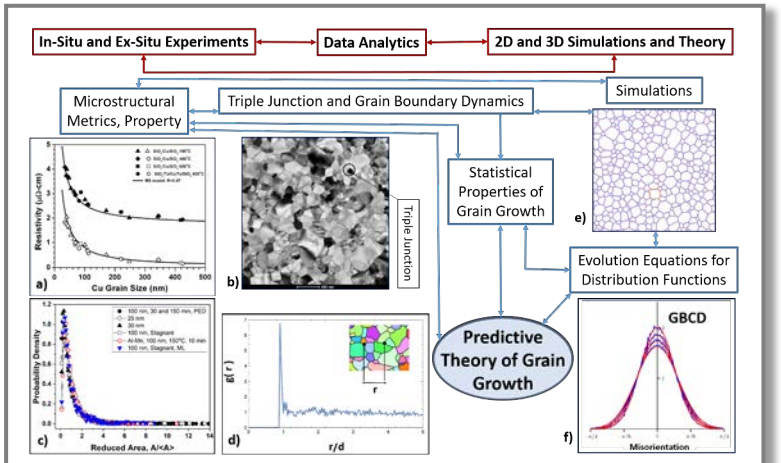


Fig. 1. (a) Electrical resistivity versus grain size in Cu films; (b) Bright-field transmission electron micrograph of a polycrystalline Pt film from an instance of an in-situ experiment (scale bar 100 nm); (c) Grain area distribution for Al films from ex-situ experiments; (d) The triple-junction pair correlation function, g(r), vs. triple junction separation, r, for a 40-nm thick W film microstructure. (Inset - A schematic of separated triple junctions); (e) Example of time instance of microstructure from the 2D grain growth simulation and (f) evolution of grain boundary character distribution (GBCD) (blue) from 2D simulations versus solution to Fokker-Planck equation (red).

Relevance to MGI

The MGI goal of rapid discovery and deployment of advanced materials requires, in the case of polycrystalline materials, an enhanced understanding of microstructural development to guide subsequent materials design. Thus, this project integrates experiment, theory and simulation to obtain a better, quantitative understanding of grain growth mechanisms, using thin metallic films as the experimental testbed. In particular, the data extracted from time sequences of experimental micrographs, via machine learning techniques, are used to inform new models of grain growth that describe a microstructure as an interacting assemblage of triple junctions. Figure 1 illustrates the four interrelated components of this project, namely: experiments, data analytics, numerical simulations and theory.

Technical Progress

Experiments: Bright-field TEM images, such as Fig. 1b, display very complex contrast, and prior efforts at image analysis based on more conventional filters have had limited success. However, in the past year, we have shown the successful use of a modified version of U-net convolutional neural network to automate grain boundary detection and to demonstrate good agreement with prior results for grain size distribution (Fig. 1c). In addition, experimentally obtained correlations between grain boundary character distribution (GBCD) and grain boundary energy distribution (GBED) in nanocrystalline tungsten and aluminum films imply that the force balance at the triple junction (Herring condition) does not fully specify the triple junction geometry.

Modeling, Simulation and Theory: In [1,2,4], we developed new models for the evolution of the 2D grain-boundary network with finite mobility of the triple junctions and with dynamic lattice misorientations. We established well-posedness results, as well as large time asymptotic behavior for the models. In particular, our results included obtaining explicit energy decay rate for the system in terms of mobility of the triple junction and the misorientation parameter. In [3], we also presented and discussed relevant experimental results of grain growth in thin films. Moreover, the developed grain growth models and theory are consistent with very recent thin films experimental findings that the triple junctions in thin films do not always obey the Herring condition. Our work

also contributes to the development of structure-preserving algorithms for grain growth simulation in 2D and 3D.

Data Analysis: We have developed statistical tools, including pair correlation, quaternion pair correlation and triple-junction distribution functions, to characterize the spatial distribution of triple junctions in experimental micrographs of both tungsten and aluminum thin films (see Fig. 1d). Our aim is to determine the effective interaction between spatially separated triple junctions and to monitor the evolution of developing film texture which will provide key inputs to our models.

Future Plans

Experiments: Ex-situ experiments will be conducted at a series of temperatures, with lower temperatures to allow for more time steps prior to stagnation, and higher temperatures for long periods to determine whether there is indeed stagnation of grain growth. The latter experiments will also allow us to determine whether the triple junctions reach a state wherein the Herring condition of force balance is satisfied. In-situ experiments in the TEM will be used to obtain dynamic data on the motion of boundaries and triple junctions, as well as to examine the evolution of grain disorientations. The dynamic data are of particular importance to theory development.

Modeling, Analysis and Numerical Simulations: We will continue our work on deriving thermodynamically consistent models for deterministic evolution, as well as stochastic dynamics of grain boundaries. We will continue to study systematically the mathematical structures of the derived systems, as well as we will design corresponding high-resolution numerical algorithms to validate the analytical results and to compare with experiments.

Data Analysis: With the availability of in-situ experimental data, we will quantify triple junction motion using a density-density correlation function. This approach is motivated by a similar treatment of the dynamics of interacting fluid particles that is employed in atomic-level computer simulations. The correlation function embodies the dynamical properties of the ensemble of triple junctions comprising the microstructure.

Broader Impacts and Workforce Development

Undergraduate students, J. Lopez, S. Toderas and J. Eckstein, worked under the guidance of Barmak and a collaborator, S. Levine (Duquesne Univ.), employed a modified convolutional neural network machine learning approach to automate detection of grain boundaries in bright-field transmission electron microscopy (TEM) images.

The postdoc Chang (Kamala) Liu (Utah) has worked under the guidance of Y. Epshteyn and C. Liu on the analysis and validations of the nonlinear Fokker-Planck models with energy-dissipation structures that appear as macroscopic models for grain boundary dynamics in polycrystalline materials [5].

Data Management and Open Access

Codes developed in this program will be made publicly available. For example, the modified code for automated grain boundary detection based on the U-Net CNN can be found at: https://github.com/xkstein/Grain-U-Net. Experimental and synthetic microstructures will also be made publicly available via curated repositories.

Advancing Along the Materials Development Continuum and Partnerships to Translation

It is unlikely that a non-integrated approach would be successful in this context as a predictive theory of grain growth requires considerable experimental input, which in turn necessitates statistical analyses of experimental data. Looking forward, successful industrial implementation of the knowledge base acquired here will depend critically on the acquisition of additional time-dependent grain growth data needed for model refinement.

Publications and References

1. Y. Epshteyn, C. Liu, M. Mizuno, *Large Time Asymptotic Behavior of Grain Boundaries Motion with Dynamic Lattice Misorientations and with Triple Junctions Drag*, Communications Math. Sci. **19**,1403 (2021).
2. Y. Epshteyn, C. Liu, M. Mizuno, *Motion of Grain Boundaries with Dynamic Lattice Misorientations and with Triple Junctions Drag*, SIAM J. Math. Anal. **53**, 3072 (2021).
3. K. Barmak, A. Dunca, Y. Epshteyn, C. Liu, M. Mizuno, *Grain Growth and the Effect of Different Time Scale*s, Springer AWM series volume "Research in the Mathematics of Materials Science", (2022), in press.
4. Y. Epshteyn, C. Liu, M. Mizuno, *A Stochastic Model of Grain Boundary Dynamics: a Fokker-Planck Perspective* (2021), submitted for publication, arXiv:2106.14249.
5. Y. Epshteyn, C. Liu, C. Liu, M. Mizuno, *Local Well-Posedness of a Nonlinear Fokker-Planck Model*, (2022), submitted for publication, arXiv:2201.09117.

Data science driven quantum chemistry for reactive chemistry controlled by stimuli

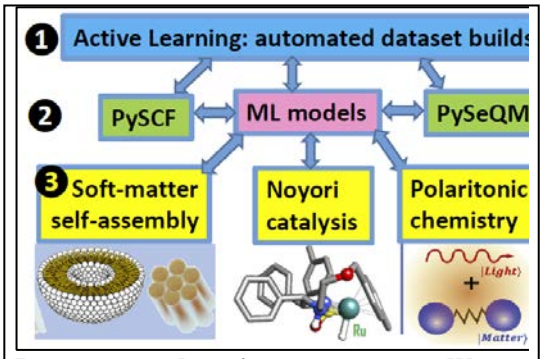
Lead Investigator: Kipton Barros, Theoretical Division, Los Alamos National Lab, kbarros@lanl.gov.
Sergei Tretiak, Theoretical Division, Los Alamos National Lab, serg@lanl.gov.
Garnet K. Chan, Division of Chemistry and Chemical Engineering, California Institute of Technology, gkc1000@gmail.com.

Keywords: Computational quantum chemistry, interatomic potentials, polaritonics, ionic liquids

Project scope: The overarching goal of this project is to leverage Artificial Intelligence (AI) and Machine Learning (ML) to develop advanced simulation methodologies that will enable the study of a variety of materials and chemical processes of technological relevance. In particle, we aim to develop atomistic simulation techniques that are applicable to highly non-equilibrium processes such as chemical reactions, catalytic or light-induced processes. We employ an active learning framework to iteratively collect simulation data, which is used to train and refine physics-based ML models of interatomic potentials.

Relevance to MGI: Computational atomistic modeling is a crucial element in our discovery and understanding of natural and man-made materials. Large-scale molecular dynamics (MD) simulations are a core tool in theoretical efforts. A fundamental limitation of classical MD simulations is in its description of atomic interactions in the interatomic potentials. The advancement of ML-potentials unlocks the possibility of large-scale simulations with unprecedented accuracy. This project aims to significantly extend the domain of applicability of such potentials by incorporating more physical knowledge into their design. Our work includes several thrusts:

- *An active learning procedure to enable the automated collection of training data through careful design of new quantum simulations*. High-quality training data is essential to the success of any ML model. We employ active learning to iteratively collect a training dataset by applying the existing ML models to perform MD simulation of highly non-equilibrium processes, and looking for uncertainty in the ML model predictions.
- *Physics inspired design of next-generation ML models and semi-empirical quantum mechanics.* We formulate extensions of ML models to predict, e.g., excited state properties and highly non-equilibrium processes involving charge dynamics and bond breaking.
- *Applications to challenging chemical processes, including soft matter self-assembly, homogeneous catalysis, and polaritonic chemistry.*



Interconnected project components. We are developing computational tools to understand and control chemical reactions under varying thermodynamic conditions and external stimuli.

Technical Progress: Since the inception of our project in 2021, we have made progress in a number of key areas. Some highlights of our progress are listed below.

- *Advances in sampling methodologies for active learning*. The key to developing accurate and general ML models is the collection of high-fidelity training datasets that span the extremely large space of chemical configurations. It is hard to predict, *a priori*, what configurations will be most relevant. We have been using our ML-potentials in conjunction with sampling methods such as transition path sampling to probe reaction barriers and surfaces; the accurate modeling of these regions of configuration space is essential to obtaining correct predictions of, e.g., reaction rates. In some cases, the possibly relevant reaction paths are not known. Here, we have been developing a new approach to sampling that biases the dynamical trajectories towards regions of chemical configuration space for which the ML model reports larger uncertainty. This procedure accelerates the collection of training data that is most likely to yield significant improvements to the ML model.
- *Methodologies for training to diverse datasets that are multi-fidelity and multi-modal in character*. Traditional ML potentials are trained to homogeneous datasets including a consistent set of observables, such as energies and forces. It is of great interest to incorporate inhomogeneous data, e.g., quantum simulations at varying levels of theory (DFT, MP2, CCSD, …) as well as experimental measurements, which has an entirely different

character, and comes with different forms of uncertainty. We have been exploring a *meta-learning* approach to train models that are highly adaptable, and capable of incorporate these disparate sources of information.

- *Development of dynamically responsive chemical Hamiltonians with semi-empirical quantum mechanics*. This approach goes beyond traditional modeling of interatomic potentials by using ML to predict parameters for a semi-empirical quantum Hamiltonian that is solved self-consistently. The subtle aspect of the work-flow is the procedure to iteratively train the ML model parameters. A standard gradient descent procedure can be used to train the model parameters. This approach demonstrates the many virtues of incorporating physics-based structure into ML models: The resulting models yield highly accurate predictions that are both transferable and interpretable.

Future Plans: We will continue with the various theoretical and methodological efforts, in particular: (1) Improvements to the sampling techniques used in active learning to develop training datasets that yield more robust and transferable ML potentials, (2) Methods for incorporating more information from inhomogeneous datasets, e.g., quantum simulations calculated at multiple levels of theory, as well as experimental data of a different modality, (3) Further development of physics-based ML model architectures, which includes dynamical parameterization of semi-empirical quantum Hamiltonians, as well as new ways to incorporate long-range physical effects such as self-consistent charges within more traditional ML potentials, (4) The development of new empirical quantum chemistry codes that are highly adaptable for integration with ML-based parameterization. These advances in methodology will unlock new applications that were previously inaccessible, including reactive chemistry, activated processes such as polaritonic chemistry, and soft-matter self-assembly driven by long-range electrostatic interactions.

Data Management and Open Access: A large part of our project is the application of active learning to develop datasets of high-fidelity simulations that will be released to the ML community. Our ML codes are being developed openly on Github, such as the HippyNN project (https://github.com/lanl/hippynn). Finally, we will be providing open-access versions of our published manuscripts.

Broader Impacts and Workforce Development: Training the next-generation workforce is a key imperative in the present effort. In order to foster a new generation of researchers, we hired three postdoctoral fellows at LANL and engaged several students at Caltech. Our goal is for the postdocs and student to work at the intersection of the different aspects of the project and acquire marketable skills in data science, machine learning, electronic structure theory, dynamics, and high performance computing.

Advancing Along the Materials Development Continuum: Our project aims to build simulation tools at the atomic scale providing combination of accuracy and efficiency. These tools will enhance the fidelity of traditional MD simulations, as well as enable new types of simulations, such as nonequilibrium polaritonic processes, where matter-light interactions can provide a mechanism for controlling chemical processes, including energy harvesting, transport, photochemistry, and quantum information processing. We are also developing ML models that accurately capture reactive pathways; this will enable, e.g., the study of catalytic process that are key to unlocking more efficient production of fuels and chemicals.

Publications:

1. Guoqing Zhou, Nicholas Lubbers, Kipton Barros, Sergei Tretiak, Ben Nebgen, *"Deep Learning of Dynamically Responsive Chemical Hamiltonians with Semi-Empirical Quantum Mechanics,"* ***PNAS***, (2022, to appear).
2. Nikita Fedik, Roman Zubatyuk, Maksim Kulichenko, Nicholas Lubbers, Justin S. Smith, Benjamin Nebgen, Richard Messerly, Ying Wai Li, Alexander I. Boldyrev, Kipton Barros, Olexandr Isayev and Sergei Tretiak, *"Extending machine learning beyond interatomic potentials for predicting molecular properties,"* ***Nature Reviews Chemistry***, (2022 to appear).
3. Maksim Kulichenko, Nicholas Lubbers, Justin S. Smith, Benjamin Nebgen, Ying Wai Li, Kipton Barros, and Sergei Tretiak, *"Accelerating active learning by biasing towards regions of model uncertainty,"* In preparation.
4. Michael Chigaev, Justin S. Smith, Kipton Barros, and Nicholas Lubbers, *"Lightweight and Effective Tensor Sensitivity for Atomistic Neural Networks,"* In preparation.

Collaborative Research: DMREF: Switchable Underwater Adhesion through Dynamic Chemistry and Geometry

Lead Investigator: Michael Bartlett (mbartlett@vt.edu), Bruce Lee (bplee@mtu.edu), Rong Long (Rong.Long@colorado.edu), Grace Gu (ggu@berkeley.edu)
Participating Institutions: Virginia Tech, Michigan Tech, University of Colorado Boulder, University of California Berkeley
Website: https://dmref.org/projects/1577
Keywords: Adhesion, Soft materials, Mussel adhesive chemistry, Machine learning, Contact mechanics

Project Scope

We will combine dynamic underwater adhesive chemistry with active adhesive geometry to determine how adhesion can be switched in dry and wet conditions (Figure 1). This will be accomplished by incorporating mussel-inspired adhesive chemistry on an octopus-inspired adhesive structure to rapidly switch adhesion, which will allow us to simultaneously leverage the strengths of both organisms to develop new design principles to transform state-of-the-art adhesives. We will establish for the first time a switchable adhesive framework for cooperative design that amplifies the benefits of chemistry and active geometry for rapidly switchable yet strong underwater adhesives.

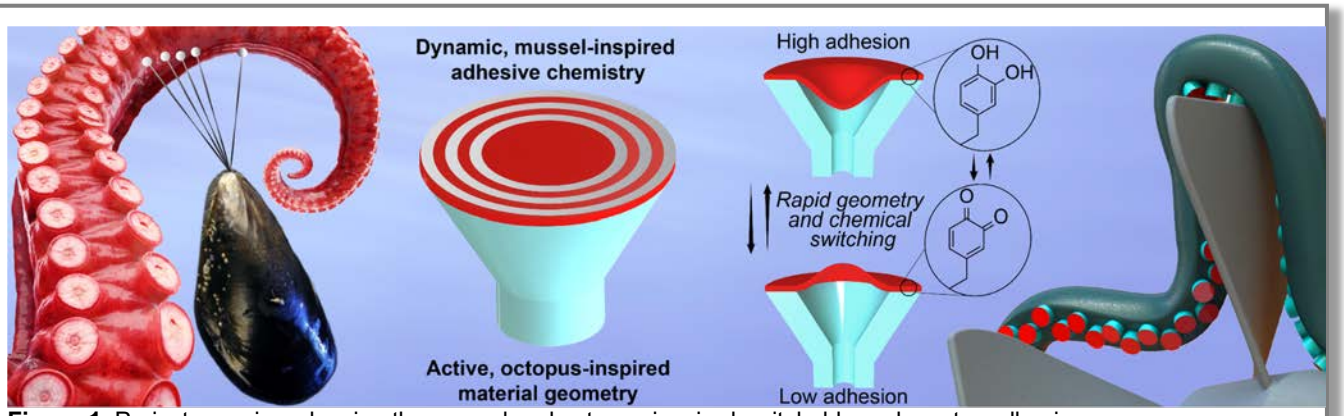


Figure 1. Project overview showing the mussel and octopus inspired switchable underwater adhesive

Relevance to MGI

Strong adherence to underwater or wet surfaces is a significant challenge, especially when rapid attachment with high capacity and easy release is required. To overcome this, we will systematically determine the geometry of the adhesive structure through feedback between mechanistic modeling and machine learning (ML) algorithms to create a virtual prototyping environment. This will feed into digital fabrication techniques to create materials and active membrane geometries enabling addressable and rapidly programmable adhesion strength. We will first determine the mechanisms behind octopus-inspired active geometry and mussel-inspired chemistry and then how they synergize to control adhesion. This provides unique opportunities to speed up the switching of adhesive chemistry by releasing the interface with active materials and enhancing the adhesion strength with engineered adhesive chemistry. Taken together, our approach will aim to simultaneously increase the switching ratio and strength of adhesion, and to reduce the switching time. By uniquely combining experiments, simulations, and ML, we will accelerate and build fundamental knowledge of how chemical, geometric, and material properties influence switchable adhesion to engineer rapidly switchable adhesives for wet and dry environments.

Technical Progress

Thus far in the project the team has established fabrication procedures for controlling adhesive geometry and chemical patterns. Characterization tools are also being established to evaluate adhesive characteristics rapidly and systematically in dry and wet environments. Finite element (FE) models have been established and systematic investigation of geometry has begun. Machine learning has created a fully connected neural net trained on simulated data points with random adhesive geometry. Apart from a few outliers, initial results describe model behavior well.

Future Plans

We will combine wet adhesive chemistry with active geometry to create a combined mussel-inspired and octopus-inspired switchable underwater adhesive. By creating a framework to bring together experiments, modeling, and ML based design, we will establish the fundamental mechanisms behind underwater switchable adhesives. We will first determine how octopus-inspired and mussel-inspired adhesive mechanisms function (Objective 1 and 2) and then how they couple together to control adhesion (Objective 3). In contrast to prior work which studied how chemistry or geometry independently influences adhesion, we will uniquely determine how dynamic chemistry and active materials combine to control adhesion, providing new paradigms in adhesive design.

Objective 1 will determine how materials and geometry of an underwater adhesive element affect its adhesion performance against objects with different stiffness and curvature. We will establish fabrication techniques and characterization approaches, create mechanistic models, and use ML design over large geometric and material space. In Objective 2 we will elucidate the effect of chemistry patterning on the active membrane on the performance of underwater adhesion. Chemistry patterning will offer a new dimension of control over the adhesion strength and switching ratio by guiding the membrane detachment process against highly curved or soft substrate. This will create a library of catechol-based adhesives and predictive modeling tools for interface design that will leverage ML inverse design to accelerate pattern exploration. In Objective 3, we will combine dynamic adhesive chemistry with active membrane geometry for rapid and effective switchable adhesion. This includes electrochemical reactions for adhesion control, using ML to determine coupling relationships between chemical patterns and geometry, and experimental implemented to determine the combined effect of active chemistry and geometry on adhesion switching ratio, switching time, and adhesion strength.

Broader Impacts and Workforce Development

In addition to training strong Ph.D. candidates, we will develop K-12 outreach activities with the Girls in Engineering program to develop bioinspired adhesive robotic gripping kits for hands-on activities. The Co-PI's will also each give online guest lectures in courses taught by the other members of the team. Adhesive design that combines strategies utilized by the octopus and mussel will be of great interests to a diverse group of students. We will develop a workshop at the Adhesion Society Annual Meeting to disseminate the concepts of ML and underwater adhesion. We will further host a career development panel and invite Adhesion Society members from academia, industry, national labs, and start-up companies as panelists to inform and inspire future workforce leaders in adhesion science and engineering.

Data Management and Open Access

Code for computational modeling and ML design and procedures for chemical and experimental operations will be released for public access as the project develops.

Advancing Along the Materials Development Continuum and Partnerships to Translation

As the project develops partnerships for translation will be explored to deploy the rapidly switchable underwater adhesives.

Publications and References

None to report yet.

Systematic Discovery of Materials Platforms for Spin-Light Quantum Interfaces

Lead Investigator: Lee C. Bassett, lbassett@seas.upenn.edu; Michael Flatte, michael-flatte@uiowa.edu; Rashid Zia, Rashid_Zia@brown.edu
Participating Institutions: University of Pennsylvania; University of Iowa; Brown University
Website: none
Keywords: quantum information science; solid-state defects; spin qubits; light-matter interface

Project Scope

This project aims to accelerate the discovery of defects, dopants, and host materials optimized as physical platforms for quantum information science. The project combines theoretical and experimental approaches for:

- Prediction of new quantum defects
- Systematic discovery and efficient characterization
- Refinement, generalizations, and in-depth study

The project's research goals are further coupled with a mission to educate students and the public about quantum information science, and to generate useful open-source resources for the larger scientific community.

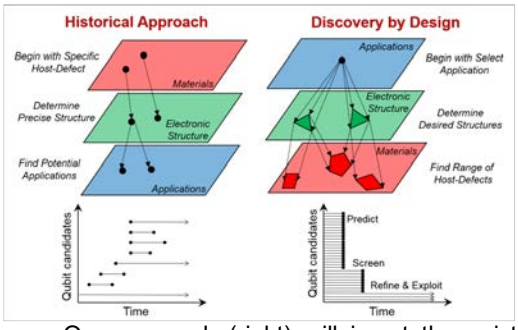


Figure. Our approach (right) will invert the existing paradigm (left) to identify defects optimized for specific applications in quantum technology.

Relevance to MGI

The number of potential quantum defect systems is vast. This project posits that improved computational and experimental tools can accelerate the discovery of quantum defects that are optimized for particular applications.[1] To achieve this, we are pursuing a new approach to efficiently predict defect properties based on a combination of analytical group theory and *ab initio* calculations. The goal is to create tools that allow for rapid screening of host-material/defect combinations with sufficient quantitative accuracy to enable prediction of relevant properties for quantum applications. In parallel, we are developing new measurement and data-analysis tools that facilitate rapid sample screening[2,4] and efficient characterization of quantum-optical properties.[5] These efforts are made possible through a close collaboration between team members with complementary expertise in each of these domains.

Technical Progress

In this project to date, we have focused on a few model systems to develop and refine our methodology. These include rare-earth (RE) and transition-metal (TM) impurities in three-dimensional crystals, and visible quantum defects in two-dimensional hexagonal boron nitride (h-BN). The RE and TM systems are naturally suited to group-theory analysis, since their electronic levels can be qualitatively predicted using a symmetrized crystal field associated with a particular host. However, *ab initio* calculations have historically been challenging for the *f*-level electron systems intrinsic to RE ions. To validate and assess the machine learning approaches for lanthanides we have examined charge transfer for La in M-type hexaferrites[6] and we have established calculations of the crystal field splittings for lanthanides in GaN and Y_2O_3. Once a validated calculation has been completed for Er in Y_2O_3, we will extend these simulations to many other lanthanides to provide a large, validated data set for machine learning identification of preferred defect-host combinations for spin-photon interfaces.

We have also performed extensive theoretical and experimental investigations of copper impurities in zinc sulfide crystals. Zinc sulfide's wide bandgap, low spin-orbit coupling, and dilute nuclear spin background make it a promising host material for quantum defects, but it is barely explored for this purpose. Copper-doped ZnS is well known in the literature as a phosphor system with visible fluorescence and characteristic electron-spin resonance signatures, however the origin of this fluorescence remains under debate. Through extensive experiments and computational modeling, we have established the red fluorescence of Cu:ZnS as resulting from a set of copper-vacancy complexes. We have also begun experiments in collaboration with Sandia National Laboratory to create these complexes using focused ion beam implantation in arrays that allow for rapid characterization of their creation efficiency. These experiments will be facilitated by new techniques we have developed to rapidly detect, count, and classify quantum emitters in large-area photoluminescence maps.[2,4]

In parallel, we are studying visible quantum emitters in h-BN, which are promising for applications in quantum photonics and quantum sensing. Two-dimensional materials can be readily interfaced with photonic and electronic devices, and h-BN's quantum emitters show unique properties as bright single-photon emitters and optically addressable spins. Despite extensive study, however, the chemical and electronic structure of these emitters remain unknown. Using photon emission correlation spectroscopy,[5] we have uncovered important features of the emitters' optical dynamics, including the existence of transient excited states and multiple metastable states associated with distinct spin and charge manifolds in the excitation and emission cycle.[3]

Future Plans

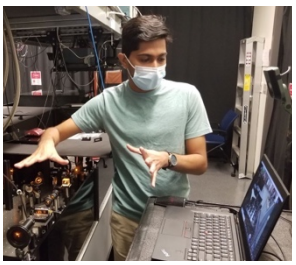
We intend to build on our current research efforts in technical Thrusts I-III, continue developing outreach and education resources, and engage with the scientific community. Our plans particularly include:

- Applying our understanding of Cu:ZnS to make theoretical predictions for other TM impurities in ZnS.
- Exploring TM and RE-based defect centers in SiC and MgO as potential quantum defects.
- Closing the gap between analytical group theory treatments of RE impurities.
- Establishing quantum control of single spin defects in hBN and/or ZnS.

Broader Impacts and Workforce Development

In support of the broad MGI goals to train a new generation of researchers, we have:

- Designed and implemented new courses in quantum technology and quantum engineering specifically targeted to materials scientists and engineers.
- Conducted outreach programs aimed to inspire secondary-school and undergraduate students to pursue studies in quantum engineering.
- Organized community symposia and workshops around the topic of materials for quantum information science.



DMREF-supported PhD Student Raj Patel giving a virtual lab demonstration in 2021.

Data Management and Open Access

A long-term goal of this project is to establish an open-source codebase and searchable database for the calculation of defect properties based on analytical group theory and *ab initio* calculations. We are also developing new methods for the efficient analysis of experimental data to aid the identification and characterization of quantum defects. These methods are described in three papers,[2,4,5] and the associated codes are in public repositories.[1]

Advancing Along the Materials Development Continuum and Partnerships to Translation

Although quantum technology remains nascent, industry is playing an increasingly important role in developing and refining commercial applications. Our team maintains connections with research teams in industry (Northrop Grumman, IBM) and at national laboratories (Sandia, Ames, NRL), who will be able to capitalize on discoveries to realize new generations of quantum sensors and quantum communication links.

Publications and References

1. L. C. Bassett, A. Alkauskas, A. L. Exarhos, and K.-M. Fu, *Quantum defects by design.* Nanophotonics **8**, 1867 (2019).
2. S. A. Breitweiser, A. L. Exarhos, R. N. Patel, J. Saouaf, B. Porat, D. A. Hopper, and L. C. Bassett, *Efficient optical quantification of heterogeneous emitter ensembles* ACS Photonics **7**, 288 (2020).
3. R. N. Patel, D. A. Hopper, J. A. Gusdorff, M. E. Turiansky, T-Y. Huang, R. E. K. Fishman, B. Porat, C. G. Van de Walle and L. C. Bassett, *Probing the Optical Dynamics of Quantum Emitters in Hexagonal Boron Nitride*, (Manuscript under review), Preprint available at arXiv:2201.08881 [cond-mat.mes-hall].
4. L. R. Narun, R. E. K. Fishman, H. J. Shulevitz, R. N. Patel, and L. C. Bassett, *Efficient Analysis of Photoluminescence Images for the Classification of Single-Photon Emitters.* (Manuscript under review), Preprint available at arXiv:2112.05654 [cond-mat.mes-hall].
5. R. E. K. Fishman, R. N. Patel, D. A. Hopper, T.-Y. Huang, L. C. Bassett, *Photon emission correlation spectroscopy as an analytical tool for quantum defects.* (Manuscript in revision), Preprint at arXiv:2111.01252 [quant-ph].
6. C. Bhandari, M. E. Flatté, D. Paudyal, *Enhanced magnetic anisotropy in lanthanum M-type hexaferrites by quantum-confined charge transfer*, Physical Review Materials **5**, 094415 (2021).

[1] https://github.com/penn-qel/emitter-quant
https://github.com/penn-qel/PL-classify
https://github.com/penn-qel/photon-emission-correlation-spectroscopy

Turning Carbon Dioxide into 3D-Printed Concrete via Integrated Machine Learning, Simulations, and Experiments

Lead Investigator: Mathieu Bauchy, bauchy@ucla.edu
Participating Institutions: University of California, Los Angeles (UCLA)
Website: none
Keywords: Carbon utilization, waste-to-resource approach, concrete, 3D-printing

Project Scope

This project aims to decipher the fundamental science required to accelerate the design of a 3D-printable alternative concrete allowing CO_2 uptake. First, we will understand, control, and optimize the rheology of concentrated portlandite suspensions to enable printability. Second, we will refine portlandite carbonation routes at ambient temperature and elevated pressures to maximize CO_2 uptake. Third, we will benefit from 3D-printing to optimize the strength-to-weight ratio of the printed structures.

Relevance to MGI

By systematically integrating experiments and simulation/machine learning, this project goes significantly beyond traditional paradigms (wherein simulations are solely used to understand experiments and/or experiments are conducted to simply validate simulations). Rather, each of the activities conducted in this project is based on a closed-loop iterative process wherein computational and experimental activities mutually inform and advance each other. Specifically, experiments, multi-scale simulations, machine learning, and optimization methods are integrated to (i) optimize the rheology of portlandite suspensions to enable printability, (ii) identify optimal conditions for portlandite carbonation, and (iii) guide the design of optimal multi-material 3D-printed structures.

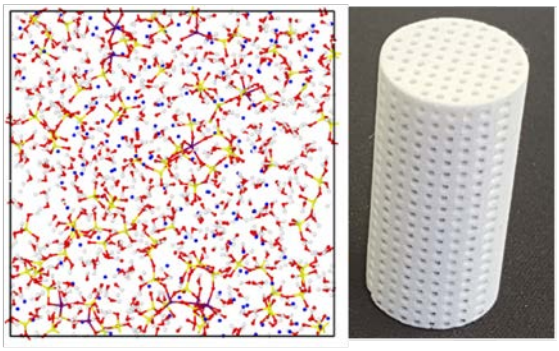


Figure. (left) Reactive molecular dynamics simulation of portlandite carbonation. (right) Picture of a 3D-printed 4-cm-long portlandite cylinder.

Technical Progress

1) We explored the relationship between packing density, solution properties, temperature, propensity for particulate aggregation, and rheology in aqueous suspensions of portlandite. We found that, in contrast to the viscosity of the suspending medium, the viscosity of portlandite suspensions increases upon increasing temperature—which arises from the temperature-induced aggregation of polydisperse portlandite particulates because of the diminution of electrostatic repulsive forces upon increasing temperature. This understanding is needed to control the rheological properties of colloidal portlandite suspensions that feature strong charge screening behavior and specifically how to account for temperature changes that affect the engineering-scale processability of these suspensions. This is key to enable their 3D-printing.

2) We explored how the temperature, relative humidity (RH), and CO_2 concentration influence mineralization reactions at the particulate and monolith scales at atmospheric pressure. We found that, in general, increasing RH significantly impacts the rate and thermodynamics of carbonation reactions, leading to high(er) conversion regardless of prior exposure history. This mitigated the effects (if any) of allegedly denser, less porous carbonate surface layers formed at lower RH. This study improved our understanding of the influence of gas processing conditions on CO_2 mineralization reactions—for both particulate and monolithic portlandite geometries—to promote rapid CO_2 mineralization within cementing components.

3) We conducted some reactive molecular dynamics simulations to investigate the mechanism of the nucleation of amorphous calcium carbonate following the dissolution of portlandite precursors. We showed that the gelation

reaction is driven by the existence of some competing local molecular stresses within the Ca and C precursors, which progressively get released upon gelation. Such stress acts as an elastic energy penalty that thermodynamically promotes the polymerization of the calcium carbonate gel.

4) We developed a machine learning algorithm aiming to predict the compressive strength of concrete materials. Our approach offers a state-of-the-art accuracy and is able to predict the 28-day compressive strength of concrete materials with an accuracy of ±4.4 MPa based on the sole knowledge of their mixture proportions.

Future Plans

We plan to (i) use machine learning to uncover the effect of the binder's composition and aggregates' physical characteristics on the rheology and strength of concretes, (ii) develop an "inverse design" generative approach combining high-throughput continuum simulations and deep learning to discover optimal structures maximizing strength while minimizing material usage, and (iii) refining the 3D-printing process of the present binder.

Broader Impacts and Workforce Development

We developed a short course on "Machine Learning for Materials," which was offered online in collaboration with the American Ceramic Society and the International Materials Research Congress. Going forward, the course is expected to be augmented and hosted on nanoHUB.org. The short course is also available on the Youtube channel of the PARISlab laboratory (https://www.youtube.com/channel/UCT6I2JS2k5tMYV9zg1Yq98A).

Data Management and Open Access

All simulation and machine learning codes have been made available on the shared GitHub depository of the PARISlab laboratory (https://github.com/parislab).

Advancing Along the Materials Development Continuum and Partnerships to Translation

As a key achievement, our team won the $7.5M NRG COSIA Carbon XPRIZE competition, which was initiated to develop breakthrough technologies to convert CO_2 emissions into usable products. The jury noted that our solution is a first-of-its-kind, eco-friendly approach for taking carbon dioxide emissions directly from power plants and other industrial facilities—emissions that would otherwise go into the atmosphere.

Publications and References

1. B. Ouyang, Y. Song, Y. Li, G. Sant, M. Bauchy, *EBOD: An ensemble-based outlier detection algorithm for noisy datasets*, Knowledge-Based Systems **231**, 107400 (2021).
2. S. Xiao, H. Liu, E. Bao, E. Li, C. Yang, Y. Tang, J. Zhou, M. Bauchy, *Finding defects in disorder: Strain-dependent structural fingerprint of plasticity in granular materials*, Applied Physics Letters **119**, 241904 (2021).
3. B. Ouyang, Y. Song, Y. Li, F. Wu, H. Yu, Y. Wang, Z. Yin, X. Luo, G. Sant, M. Bauchy, *Using machine learning to predict concrete's strength: Learning from small datasets*, Engineering Research Express **3**, 015022 (2021).
4. T. Du, H. Liu, L. Tang, S. S. Sørensen, M. Bauchy, M. M. Smedskjaer, *Predicting fracture propensity in amorphous alumina from its static structure using machine learning*, ACS nano **15**, 17705–17716 (2021).
5. L. Tang, H. Liu, G. Ma, T. Du, N. Mousseau, W. Zhou, M. Bauchy, *The energy landscape governs ductility in disordered materials*, Materials Horizons **8**, 1242–1252 (2021).
6. E. C. La Plante, I. Mehdipour, I. Shortt, K. Yang, D. Simonetti, M. Bauchy, G. N. Sant, *Controls on CO_2 Mineralization under Ambient Conditions*, ACS Sustainable Chemistry & Engineering **9**, 10727–10739 (2021).
7. G. Falzone, I. Mehdipour, N. Neithalath, M. Bauchy, D. Simonetti, G. Sant, *New insights into the mechanisms of carbon dioxide mineralization by portlandite*, AIChE Journal **67**, e17160 (2021).
8. H. Liu, S. Xiao, L. Tang, E. Bao, E. Li, C. Yang, Z. Zhao, G. Sant, M. M. Smedskjaer, L. Guo, M. Bauchy, *Predicting the early-stage creep dynamics of gels from their static structure by machine learning*, Acta Materialia **210**, 116817 (2021).
9. Y. Song, K. Yang, J. Chen, K. Wang, G. Sant, M. Bauchy, *Machine learning enables rapid screening of reactive fly ashes based on their network topology*, ACS Sustainable Chemistry & Engineering **9**, 2639–2650 (2021).
10. R. A. Arnold, S. Dong, L. Tang, D. Prentice, M. Collin, J. C. Vega-Vila, A. Hernandez, E. C. La Plante, K. Ellison, A. Kumar, *Dissolution Amplification by Cavitational Stimulation*, The Journal of Physical Chemistry C **126**, 3432–3442 (2022).

Ada: A Self-Driving Laboratory for Clean Energy Materials Research

Lead Investigator: Curtis Berlinguette, cberling@chem.ubc.ca
Participating Institutions:
1. The University of British Columbia, Departments of Chemistry and Chemical & Biological Engineering
2. Stewart Blusson Quantum Matter Institute
3. Canadian Institute for Advanced Research (CIFAR)
Website: berlinguettegroup.com; projectada.ca
Keywords: flexible automation, Bayesian optimization, electrochemical reactors, thin films, clean energy

Project Scope

The discovery of new high-performance, low-cost materials is integral to clean energy technology development. However, such discoveries can take decades using traditional methods. Self-driving laboratories present an opportunity to explore vast materials spaces faster than ever before. Our self-driving laboratory, "Ada", accelerates materials discovery by leveraging flexible automation and machine learning to autonomously discover and optimize new materials. Ada automatically makes and characterizes thin films and coatings relevant to advanced solar cells and CO_2 electrolyzers. Ada uses a Bayesian optimizer to iteratively design experiments that maximize the information gain per sample.

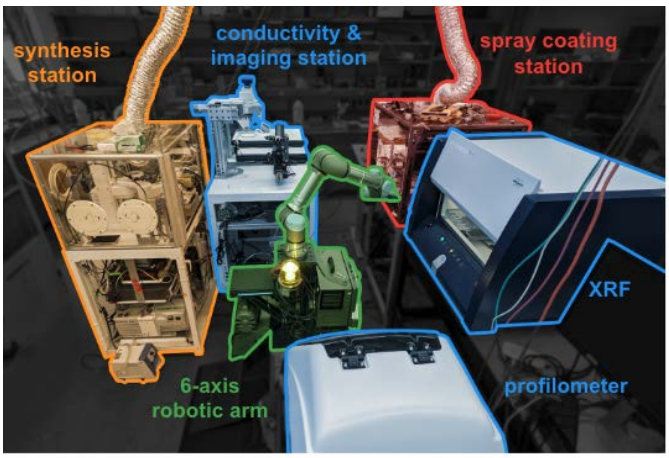


Figure 1: Ada is a self-driving laboratory that makes and tests thin-film materials. It consists of a robotic arm surrounded by various material synthesis, processing, and characterization stations.

Relevance to MGI

Ada enables researchers to address questions that would be difficult or impossible to answer through conventional experimentation. Our interdisciplinary team pairs experts in chemistry and materials science with experts in automation and machine learning. We collect large amounts of experimental data that we use to build machine learning models for predicting material properties. Our algorithms then design experiments to study commercially relevant materials that must satisfy complex and sometimes conflicting objectives. The development of both new and improved autonomous research tools and methodologies enable the materials science community to build better self-driving laboratories that have direct relevance to MGI.

Technical Progress

Project Ada, which launched in 2018, has made significant contributions to the materials community[1–7]. Our first publication describing the Ada self-driving laboratory showed how a small laboratory robot could perform a complex workflow involving multiple materials synthesis and characterization tasks[1]. We coupled this robot with a Bayesian optimization algorithm to enable Ada to autonomously optimize the hole mobility of spiro-OMeTAD, an organic hole-transport material common to perovskite solar cells. This paper has been cited nearly 200 times and has influenced the development of many subsequent self-driving laboratories and research initiatives.

Project Ada has also highlighted how flexible automation is enabling widespread automation of materials science experiments[4]. We have leveraged flexible automation to upgrade Ada with additional capabilities for inorganic film synthesis and characterization. A common challenge in the solution synthesis of inorganic films is the tradeoff between film properties and processing temperature. Many film properties are maximized by using high processing temperatures, however higher temperatures are generally undesirable because they may damage other layers in a device. To address this challenge, we developed a multi-objective optimization methodology based on a hypervolume improvement algorithm. We demonstrated this methodology and the expanded Ada platform by performing a multi-objective optimization study on combustion-synthesized palladium films relevant to hydrogen

purification, sensing, and catalysis. Ada discovered new synthesis conditions that yielded conductive films at low processing temperatures through multi-objective optimization, thereby maximizing conductivity while minimizing processing temperature of palladium films formed through combustion synthesis[5].

Future Plans

CO_2 electrolyzers can enable the conversion of waste CO_2 into valuable fuels and chemicals. The commercialization of CO_2 electrolyzers requires the development of durable membrane electrode assemblies (MEAs) that yield high product purity, energy efficiency, and carbon efficiency. These assemblies consist of gas diffusion electrodes (GDEs) with thin catalyst layers, and polymeric membranes. To obtain meaningful performance data, these materials must be tested simultaneously within the complex electrochemical environment of an operating electrolyzer. This requirement makes the development of high-throughput experiments challenging. We have addressed this challenge in two ways. First, we equipped Ada with the capacity to deposit catalyst layers using an industrially-relevant spray coating method. Second, we built a semi-automated CO_2 electrolyzer test station for rapidly characterizing MEA performance. These tools have enabled Ada to robotically fabricate and characterize GDEs, interface them with membranes to form MEAs, and rapidly test them under conditions representative of an operating electrolyzer. Future work will involve using a combination of automation and artificial intelligence to discover and test new industrially-relevant catalyst materials for CO_2 conversion.

Broader Impacts and Workforce Development

To date more than 90 people have worked on Project Ada, including students, postdoctoral fellows, and staff engineers and scientists. We are continuing to grow and pursue ambitious projects with industrial and government partners that will accelerate the development and deployment of clean energy technologies.

Data Management and Open Access

Project Ada is committed to data transparency and open communication. Datasets from published papers are provided in the SI for each paper (see references section). We make code available through a public repository (https://gitlab.com/ada-chem).

Advancing Along the Materials Development Continuum and Partnerships to Translation

Project Ada is constantly evolving and adapting to pursue new research directions. We pursued materials related to advanced solar cells and CO_2 electrolyzers, as well as collaborated with 3M to discover new adhesive formulations[7]. Flexible automation allows us to readily develop new tools and techniques for our robotic platform when new experiments are proposed. The discovery of new materials is essential for accelerating commercialization, but discovery is only part of the story. To commercialize new materials, lab techniques used during the research phase must be scaled up. We are working to overcome this hurdle by incorporating scalable materials processing methods and representative testing methods early in the research process. This approach will further accelerate the translation of research results from lab to market.

Publications and References

1. MacLeod, B. P. *et al.* Self-driving laboratory for accelerated discovery of thin-film materials. *Sci Adv* **6**, eaaz8867 (2020) DOI: 10.1126/sciadv.aaz8867.
2. Taherimakhsousi, N. *et al.* Quantifying defects in thin films using machine vision. *npj Computational Materials* **6**, 1–6 (2020) DOI: 10.1038/s41524-020-00380-w.
3. Taherimakhsousi, N. *et al.* A machine vision tool for facilitating the optimization of large-area perovskite photovoltaics. *npj Computational Materials* **7**, 1–10 (2021) DOI: 10.1038/s41524-021-00657-8.
4. MacLeod, B. P., Parlane, F. G. L., Brown, A. K., Hein, J. E. & Berlinguette, C. P. Flexible automation accelerates materials discovery. *Nat. Mater.* (2021) DOI:10.1038/s41563-021-01156-3.
5. MacLeod, B. P. *et al.* A self-driving laboratory advances the Pareto front for material properties. *Nat. Commun.* **13**, 1–10 (2022) DOI: 10.1038/s41467-022-28580-6.
6. MacLeod, B. P., Parlane, F. G. L., Brown, A. K., Hein, J. E. & Berlinguette, C. P. 4 Flexible automation for self-driving laboratories. in *Accelerated Materials Discovery* 105–122 (De Gruyter, 2022).
7. Rooney, M. B. *et al.* A self-driving laboratory designed to accelerate the discovery of adhesive materials. *Digital Discovery* (2022) DOI:10.1039/D2DD00029F.

DMREF: Hydrogel-actuated cellular soft robotic materials with programmable mechanical properties

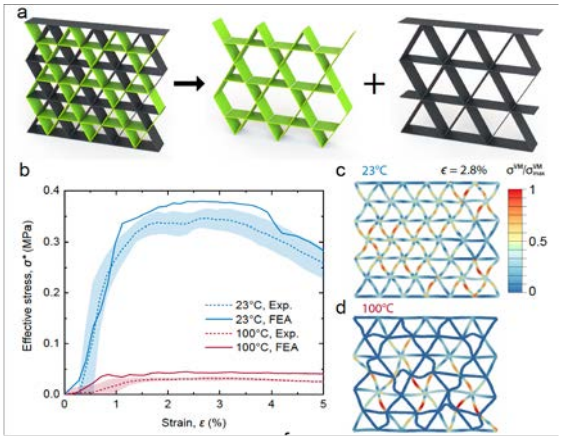
Lead Investigator: Katia Bertoldi
Participating Institutions: Harvard University and MIT
Website: none
Keywords: programmable materials, soft robotic materials, mechanical properties

Project Scope

This project introduces and investigates a new class of environmentally responsive soft robotic materials with programmed mechanical properties based on elastomeric cellular structures. The objective is to establish robust and computationally efficient methods to capture their highly non-linear response, to synthesize soft materials that generate large deformations in response to external stimuli, to identify cellular architectures that amplify and direct the response of the active materials, to develop 3D printing strategies that enables fabrication of computationally identied material designs, and to solve the inverse problem of identifying realizable layouts that form soft robotic matter with the desired behavior.

Relevance to MGI

The interaction between theory, simulations and experiments is at the core of our project, since developing and understanding new materials with programmable properties requires a balance of all three components. Theoretical and numerical analyses allow deeply understanding of the experimental results; on the other hand, experiments stimulate the development of analysis able to capture the behavior observed during the tests. In our workflow analytical algorithms, computational methods are routinely used to characterize the phase space of possible material solutions. Our design rules are then scrutinized by fabrication of material prototypes that are thoroughly tested experimentally to evaluate their performance. This allows us, through iterative feedback between theory, computation and experiment, to refine the material design rules and computational environments.



Active multimaterial lattices. (a) Multimaterial triangular lattice composed of an active material (green), which substantially weakens upon heating, and a passive material (black), whose mechanical response is temperature independent over the experimental conditions of interest. (b) Stress-strain response as measured in experiments (dashed line) and predicted by FE simulations (solid line) at T=23°C (blue lines) and T=100°C (red lines). (c,d) Numerically predicted deformation at e =2:8% and T=23°C (c) and T=100°C (d).

Technical Progress

To achieve the ambitious goal of realizing environmentally responsive soft robotic materials with programmed mechanical responses, our team has made progress in several directions. Origami Robotic structures: Origami-inspired metamaterials are realized by folding initially at sheets along predefined crease patterns andhave been used to design a wide range of innovative devices. This year *Bertoldi* has realized inflatable and bistable origami structures comprising rigid triangular facets connected by stretchable hinges at the meter-scale. Further, *Mahadevan* has shown that a marching algorithm can be applied to develop a quadrilateral origami crease pattern that has a rigid embedding at a single target shape.
LCE robotic structures: *Aizenberg* and *Bertoldi* have recently fabricated cellular microstructures out of liquid crystalline elastomers (LCEs) with an arbitrary, user-defined liquid crystal (LC) mesogen orientation encrypted by a weak magnetic field. They then demonstrated that a palette of symmetry breakings can be realized in substrate-attached LCE cellular microstructures by independently programming the anisotropy at the molecular and structural

scales. Active lattices: *Bertoldi* and *Lewis* have created a versatile computational and experimental framework to realize thermally programmable lattice architectures (see Figure). These lattices are composed of two polymeric materials with disparate glass transition temperatures, with distributions defined computationally and arranged deterministically via 3D printing. By tailoring the local composition and structure, we have produced architected lattices with tunable stiffness, Poisson's ratio, and deformation modes that are controlled through changes in the thermal environment. Color-changing materials: *Kolle* is working on an approach for producing large-area, structurally colored sheets with a rich and easily controlled design space of color patterns, spectral properties, angular scattering characteristics, and responses to mechanical stimuli.

Future Plans

Building on our recent results, Mahadevan will generalize the marching algorithm to design the entire span of rigid to floppy structures, with a spectrum of multiscale multistable structures in between. Bertoldi and Mahadevan will then apply such generalize algorithm to identify deployable structures able to switch between targeted stable states. Further, Bertoldi and Lewis will focus on integrating mechanosenory elements within 3D polymeric lattices to locally measure strain and approximate their shape change. Finally, Bertoldi and Mahadevan will work to establish a computational framework able to capture and predict the optical properties of the manufactured color-changing materials manufactured by Kolle.

Broader Impacts and Workforce Development

An important outcome of this proposal is the promotion of interdisciplinary research and teaching and an increase in the interactions between materials scientists, chemists, mathematicians, physicists, and engineers. Students of all levels (high-school, undergraduate, graduate and post-doctoral) are involved in the research process and co-advised by a combination of PIs. Our educational efforts teach students how to make broad connections embracing concepts in a multidisciplinary framework.

Data Management and Open Access

All analyzed experimental data (such as digital photographs and video) collected as part of this project have been made available in peer-reviewed journal publications and disseminated at topical conferences. The corresponding raw data, which is not necessarily published, as well as the numerical codes developed as part of this project are made available in public repositories.

Advancing Along the Materials Development Continuum and Partnerships to Translation

Our team is in close contact with Harvard Office of Technology Development to identify opportunities for commercialization and it is currently filling a few patents.

Publications and References

1. D. Melancon, B. Gorissen, C.J. García-Mora, C. Hoberman, and K. Bertoldi. *Multistable inflatable origami structures at the metre scale*. Nature **592**, 545 (2021).

2. S. Li, G. Librandi, Y. Yao, A.J. Richard, A. Schneider-Yamamura, J. Aizenberg and K. Bertoldi. *Controlling Liquid Crystal Orientations for Programmable Anisotropic Transformations in Cellular Microstructures*. Advanced Materials **33**, 2105024 (2021).

3. M.C. Fernandes, J. Aizenberg, J.C. Weaver and Bertoldi, Katia. *Mechanically robust lattices inspired by deep-sea glass sponges*. Nature Materials **20**, 237 (2020)

4. S. Li, B. Deng, A. Grinthal, A. Schneider-Yamamura, J. Kang, R.S. Martens, C.T. Zhang, J. Li, S. Yu, K. Bertoldi and J. Aizenberg. *Liquid-induced topological transformations of cellular microstructures*. Nature **592**, 386 (2021)

6. Mueller J, Lewis JA, Bertoldi K. *Architected Multimaterial Lattices with Thermally Programmable Mechanical Response*. Advanced Functional Materials **32**, 2105128 (2021)

Intrinsic local symmetry breaking in functional materials

Lead Investigator: Simon Billinge, sb2896@columbia.edu.
Participating Institutions: Columbia University, University of Colorado
Website: https://pdfitc.org
Keywords: Local structure, symmetry breaking, perovskites

Project Scope

This project applies a combined theoretical and experimental approach to discovering and understanding materials that show spontaneous *local* symmetry breaking that leads to significant modifications of their properties. We refer to these materials as polymorphous materials. With a combined novel polymorphous-DFT, synthesis, total scattering and diffraction approach we aim to discover new polymorphous materials (including discovering unknown polymorphous behavior in known materials) and to understand the origin of the polymorphous behavior. Our initial focus has been on perovskites as a technologically important class of materials. We are developing cloud-based tools (pdfitc.org) for analyzing polymorphous data and storing poly-DFT results.

Relevance to MGI

In order to realize the dream of data-driven materials discovery it is important to have the data in FAIR (findable, accessible, interoperable, reusable) repositories, and to have software services that can interact with those data. This has been an important focus of the DMREF project with the development of a database backed pdfitc.org website that hosts software applications that provide capabilities for people studying local structure and polymorphous materials. Users upload measured pair distributions (PDFs, which encode local structure), and the apps provide useful information such as candidate structures and space-group symmetries, whilst saving the data into a FAIR database backend. We are now working to integrate the results of the DFT calculations in this workflow. We are exploring the storage of the results of Poly-DFT calculations in the MPcontribs section of the Materials Project (MP), with an integration into MP, as well as storing the poly-DFT energy minimized structures themselves at pdfitc.org. Candidate polymorphous materials are fed to the synthesis and characterization groups in the project to obtain samples for local structural PDF measurements that are carried out at x-ray synchrotrons and neutrons sources to validate the theoretical predictions.

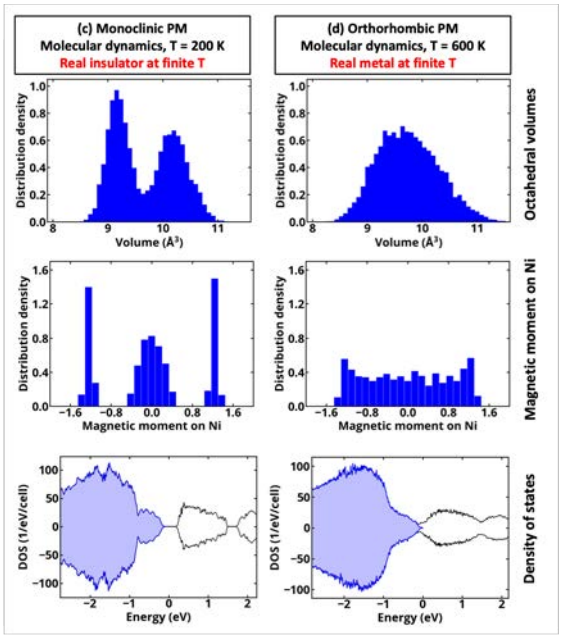


Figure. Polymorphous behavior discovered in $YMnO_3$ using poly-DFT and validated using neutron PDF measurements.

Technical Progress

The technology for carrying out polymorphous DFT has been refined and improved. We developed for the first time an approach for modeling collective rigid distortions in a material from PDF data. These distortions are often the basis of the local symmetry breaking. The pdfitc website was launched and updated with new apps for using applied math and machine learning methods for comparing large numbers of datasets. We extended the range of materials that we have applied these methods to cubic nickelates, niobates and titanates.

Future Plans

Moving forward we will pivot to our first MGI activity of searching for previously unknown polymorphous materials, not limited to the cubic phase of ABO_3. We will complete the work of building databases of the DFT

results from earlier in the project. On the experimental side the DFT predictions in the cubic nickelate phases are being investigated using total scattering and PDF methods.

Broader Impacts and Workforce Development

The project has supported 10 students and post-docs either full or part time. Regular project-wide zoom calls allow and late career participants to interact and learn from each other across the disciplines. A student led project to build a platform for providing DFT services in sub-Saharan Africa obtained a $60,000 Google Innovation Prize.

Data Management and Open Access

Software developed in the project is made available as web-based services at pdfitc.org and also released on GitHub. For example, the code that allows collective rotational distortions to be refined in the perovskites is at https://github.com/sandraskj/glazer_fitting. PDF data from our experiments are uploaded to pdfitc.org and stored in the database there. In upcoming work we plan to allow users of that website to publish their data at the click of a button which will greatly facilitate the sharing of data and their associated metadata. However, these developments are beyond the scope of the current project. As discussed above, we are working to make the results of DFT calculations available using the MPcontribs infrastructure.

Advancing Along the Materials Development Continuum and Partnerships to Translation

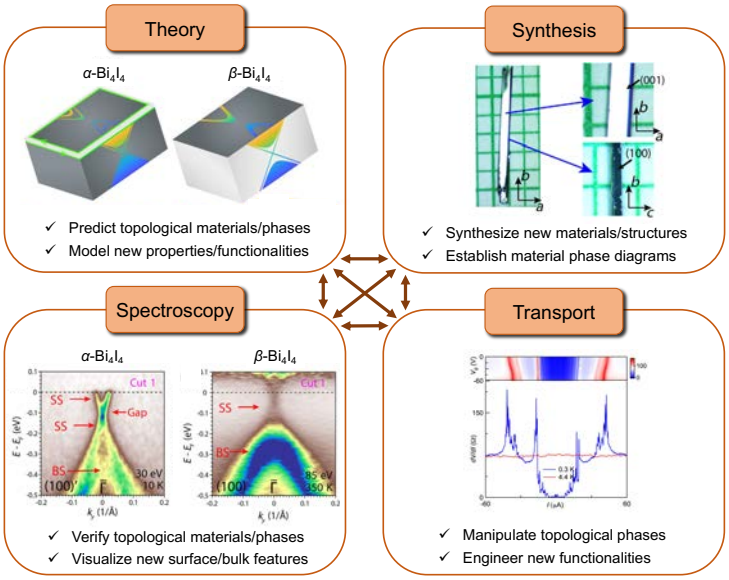
The ultimate practical application of this work would be to speed up the discovery of polymorphous materials that may have improved properties.

Publications and References

1. Xin-Gang Zhao,Oleksandr I. Malyi,Simon J.L. Billinge,Alex Zunger, *Intrinsic local symmetry breaking in the nominally cubic paraelectric phase of BaTiO3,* Phys. Rev. B, **105**, pp. 224108 (2022) 10.1103/PhysRevB.105.224108
2. Lan, L., Liu, C.-H., Du, Q., Billinge, S.J.L., 2022. *Robustness test of the spacegroupMining model for determining space groups from atomic pair distribution function data*. J Appl Cryst **55**, 626–630 (2022) https://doi.org/10.1107/S1600576722002990
3. Sandra H. Skjaervoe,Martin A. Karlsen,Riccardo Comin,Simon J. L. Billinge *Refining perovskite structures to pair distribution function data using collective Glazer modes as a basis*, arXiv:2203.00127 (2022).
4. Choudhary, K., DeCost, B., Chen, C., Jain, A., Tavazza, F., Cohn, R., Park, C.W., Choudhary, A., Agrawal, A., Billinge, S.J.L., Holm, E., Ong, S.P., Wolverton, C.. *Recent advances and applications of deep learning methods in materials science*. npj Comput Mater **8**, 1–26. https://doi.org/10.1038/s41524-022-00734-6 (2022)
5. Stach, E., DeCost, B., Kusne, A.G., Hattrick-Simpers, J., Brown, K.A., Reyes, K.G., Schrier, J., Billinge, S., Buonassisi, T., Foster, I., Gomes, C.P., Gregoire, J.M., Mehta, A., Montoya, J., Olivetti, E., Park, C., Rotenberg, E., Saikin, S.K., Smullin, S., Stanev, V., Maruyama, B.. Autonomous experimentation systems for materials development: A community perspective. Matter. https://doi.org/10.1016/j.matt.2021.06.036 (2021)
6. Billinge, S.J.L., Skjaervoe, S.H., Terban, M.W., Tao, S., Yang, L., Rakita, Y., Frandsen, B.A.. *Local structure determination using total scattering data*, in: Reference Module in Chemistry, Molecular Sciences and Chemical Engineering. Elsevier. https://doi.org/10.1016/B978-0-12-823144-9.00040-6 (2021)
7. Aktas, O., Kangama, M., Linyu, G., Catalan, G., Ding, X., Zunger, A., Salje, E.K.H.. Piezoelectricity in nominally centrosymmetric phases. Phys. Rev. Research **3**, 043221. https://doi.org/10.1103/PhysRevResearch.3.043221 (2021).
8. Zhao, X. G., Wang, Z., Malyi, O. I., & Zunger, A *Effect of static local distortions vs. dynamic motions on the stability and band gaps of cubic oxide and halide perovskites*. Materials Today, 49, 107-122 . (2021).
9. Zhi Wang,Xin-Gang Zhao,Robert Koch,Simon J. L. Billinge,Alex Zunger,, *Understanding electronic peculiarities in tetragonal FeSe as local structural symmetry breaking*, Phys. Rev. B, **102**, pp. 235121, 10.1103/PhysRevB.102.235121 (2020).

DMREF Collaborative Research: Establishing the Platform of Quasi-one-dimensional Topological Insulators with Emergent Functionalities

Lead Investigator: R. J. Birgeneau[1], C. N. Lau[2], Bing Lv[3], Ming Yi[4], Fan Zhang[3]
Participating Institutions: 1. UC Berkeley, 2. Ohio State Univ., 3. UT Dallas, 4. Rice Univ.
Website: https://dmref-q1dtm.wixsite.com/home/team **Twitter:** q1dtm
Keywords: topological insulator, quasi-one-dimensional material, phase transition



Project Scope

Our research scope includes discovering new topological materials, phases, and phenomena, controlling phases transitions, enabling superior functionalities, and fostering new technologies. Quasi-1D materials form an important family of the landscape, and this project is advancing the understanding and discovering of topological materials and phases, building on our team's expertise (theory, synthesis, fabrication, transport, spectroscopy). Establishing the materials platform and database is preparing future scientists to implement discovery, invention, and innovation. Our educational activities integrated into the research include increasing the participation of underrepresented groups, performing public outreach, mentoring undergraduate students, and comprehensive training of students and postdocs.

Relevance to MGI

To achieve the goal of efficient discovery and optimization of novel topological materials, our program operates on an iterative closed loop from theoretical prediction of topological phases to crystal synthesis to electronic structure characterization to transport realization via device fabrication back to inputs for better theoretical modeling. A successful example is our identification of the monolayer of Bi_4I_4 as a 2D Z_2 topological insulator (TI), serving as a fundamental building block for constructing exotic topological phases realized via distinct stacking patterns of such layers in real materials. From this design principle, we have predicted, realized, and confirmed weak TI and higher-order TI, which are new and rare phases of matter. The success of the feedback loop is especially apparent in our recent realization of an ideal weak TI where the complete set of topological surface state of a pair of Dirac cones and van Hove singularities are optimized to be exposed in a large global bulk gap at all temperatures. It is impossible to efficiently discover or optimize the topological materials without the iterative feedback loop.

Technical Progress

For our focus, Bi_4X_4 (X=I,Br), we predicted the β phase to be a prototypical weak TI – the only bulk-insulating one to date, and that the two α phases are intrinsic and extrinsic higher-order TIs – the only bulk/surface-insulating ones to date. We successfully synthesized Bi_4I_4, unambiguously demonstrated the weak TI for the first time using angle-resolved photoemission spectroscopy (ARPES), and found key signatures consistent with a higher-order TI. We further discovered a room-temperature topological transition between the two phases in electric transport, heat capacity, X-ray diffraction, and ARPES measurements. These results established a novel TI paradigm.

We nanofabricated ultrathin Bi_4I_4 devices and observed gate-tunable longitudinal/Hall resistances. We further achieved a Bi_4I_4 Josephson junction and observed a gate-tunable supercurrent. We predicted that Bi_4I_4/Bi_4Br_4 offer a unique platform toward achieving room-temperature quantum spin Hall effect and topological many-body physics. Our collaborators provided compelling evidence for the former using scanning tunneling microscope (STM).

We developed new synthetic routes and achieved high-quality large-size single crystals of $Bi_4I_{4-x}Br_x$ ($0 \leq x \leq 4$;

16 different values) customized for the needs of transport, ARPES, and STM characterizations. We obtained the temperature dependent structural phase diagram, with an optimized ideal weak TI and many new phases.

We also synthesized many other quasi-1D materials that potentially have topological properties based on our preliminary analyses. These include $TaTe_4$, $NbTe_4$, $(TaSe_4)_2I$, $NbTe_4I$, $TaTe_4I$, $K_2Mo_3As_3$, $Rb_2Mo_3As_3$, $Cs_2Mo_3As_3$, Nb_4FeTe_4, $BaFe_2Se_4$, Ba_2FeSe_3, $Ba_9Fe_4Se_{16}$, $Ba_5Fe_9S_{18}$, and $K_3Fe_2Se_4$. We carried out explorative syntheses to look for new quasi-1D materials/structures potentially with topological properties and found $BaPt_4Se_6$, $Zr_{6.5}Pt_6Se_{19}$, β-$BaCu_2As_2$, $BaCu_6Sn_2As_{4-x}$, $BaCu_6Sn_2P_{4-x}$, and Verbeekite-$PdSe_{1.2}Te_{0.8}$. Some of them were characterized.

Future Plans

(i) We will predict, demonstrate, and characterize the electronic structures and (new) topological properties of our synthesized 16 samples of $Bi_4I_{4-x}Br_x$ ($0 \le x \le 4$). (ii) We will nanofabricate multi-terminal Bi_4Br_4 devices to demonstrate the predicted room-temperature helical edge/hinge states in non-local quantum transport experiments. (iii) We will synthesize thicker $Bi_4I_{4-x}Br_x$ crystals with larger side surfaces to study the unprecedented side surface state properties and strain-induced topological phase transitions. (iv) We will continue to identify and characterize new quasi-1D topological/correlated materials (such as $TaTe_4$, $NbTe_4$, $BaFe_2Se_4$, $K_3Fe_2Se_4$, $(TaTe_4)_4I_2(TaI_6)$, and $(TaTe_4)_6I_4(TaI_6)$). (v) We will continue to be active through our established DMREF twitter account to bring the excitement of MGI, the progress of our research, and other significant scientific news to the public and the next generation of scientists. (vi) We will host a workshop on quasi-1D topological materials in late 2022. All the talks and brainstorming discussions will be open to the public worldwide and posted in our YouTube channel. (vii) We will document all our discovered/identified quasi-1D topological/correlated materials data to our DMREF website.

Broader Impacts and Workforce Development

Our results demonstrate $Bi_4I_{4-x}Br_x$ ($0 \le x \le 4$) as a unique tunable material platform for exploring the rich interplay of geometry, symmetry, topology, and interaction – a fundamental theme of physics. Significantly, they not only establish a new topological insulator paradigm that unifies the first and second order topological insulators but also imply the possible existence of many topological materials beyond the scope of our current knowledge, thereby accelerating the discovery of a new generation of functional materials in the topological age. An integral effort of our program is the education and training of graduate and undergraduate students as well as career preparation for postdocs. For example, in 2021 three of our students obtained their physics Ph.D., including one female, one LGBT, and one who won the Steven Weinberg Research Award from American Physical Society. Our regular interactions between the groups via team meetings provide junior participants a broad training covering materials chemistry and synthesis, materials physics and characterizations, and materials modeling and predictions.

Data Management and Open Access

Our research results have been disseminated to the research communities through journal publications, Ph.D. dissertations, and conference presentations. In addition to free versions online, all these are available from us upon request. In the following year, we will document all our quasi-1D topological/correlated materials data from both experiment and computation to our DMREF website and make them searchable and free.

Advancing Along the Materials Development Continuum and Partnerships to Translation

Operating on an iterative closed loop from theoretical prediction to crystal synthesis to material characterization to device functionality back to inputs for better theoretical modeling, we have efficiently discovered and optimized new topological materials in the $Bi_4I_{4-x}Br_x$ ($0 \le x \le 4$) family and other quasi-1D families. Our identified topological phase transition in Bi_4I_4 can be leveraged for a conceptually new, room-temperature, topological switch between robust surface and hinge conduction channels, provided that this device can pass the durability test. The Bi_4Br_4 monolayer hosts a large band gap (~200 meV) and a helical edge state, and potentially this is an ideal platform for exploring low-dissipation electronics and Majorana-based quantum information.

Our Publications (*10 Most Closely Related and Other Significant*)

1. Y. Liu, R. Chen, Z. Zhang, M. Bockrath, C. N. Lau, Y.-F. Zhou, C. Yoon, S. Li, X. Liu, N. Dhale, B. Lv, F. Zhang, K. Watanabe, T. Taniguchi, J. Huang, M. Yi, J. S. Oh, and R. J. Birgeneau, *Gate-tunable transport in quasi-one-dimensional α-Bi_4I_4 field effect transistors*, Nano Lett. **22**, 1151 (2022).
2. H. Wu, A. M. Hallas, X. Cai, J. Huang, J. S. Oh, V. Loganathan, A. Weiland, G. T. McCandless, J. Y. Chan, S.-K. Mo, D.

Lu, M. Hashimoto, J. D. Denlinger, R. J. Birgeneau, A. H. Nevidomskyy, G. Li, E. Morosan, and M. Yi, *Nonsymmorphic symmetry-protected band crossings in a square-net metal $PtPb_4$*, npj Quantum Mater. **7**, 31 (2022).
3. J. Huang, S. Li, C. Yoon, J. S. Oh, H. Wu, X. Liu, N. Dhale, Y.-F. Zhou, Y. Guo, Y. Zhang, M. Hashimoto, D. Lu, J. Denlinger, X. Wang, C. N. Lau, R. J. Birgeneau, F. Zhang, B. Lv, and M. Yi, *Room-temperature topological phase transition in quasi-one-dimensional material Bi_4I_4*, Phys. Rev. X, **11**, 031042 (2021).
4. S. Li, Y. Zhang, H. Wu, H. Zhai, W. Liu, D. P. Petit, J. S. Oh, J. D. Denlinger, G. T. McCandless, J. Y. Chan, R. J. Birgeneau, G. Li, M. Yi, and B. Lv, *Transport anomalies in the layered compound $BaPt_4Se_6$*, npj Quantum Mater. **6**, 80 (2021).
5. W. Liu, M. R. Osanloo, X. Wang, S. Li, N. Dhale, H. Wu, M. L. Van de Put, S. Tiwari, W. G. Vandenberghe, and B. Lv, *New Verbeekite-type polymorphic phase and rich phase diagram in the $PdSe_{2-x}Te_x$ system*, Phys. Rev. B **104**, 024507 (2021).
6. X. Liu, K. M. Taddei, S. Li, W. Liu, N. Dhale, R. Kadado, D. Berman, C. Dela Cruz, and B. Lv, *Canted antiferromagnetism in the quasi-one-dimensional iron chalcogenide $BaFe_2Se_4$*, Phys. Rev. B **102**, 180403(R) (2020).
7. H. Wu, H. Zhai, S. Li, M. Sorolla, G. T. McCandless, D. P. Petit, J. Y. Chan, and B. Lv, *Crystal structure and electronic properties of new compound $Zr_{6.5}Pt_6Se_{19}$*, Inorg. Chem. **59**, 8196 (2020).
8. C. Ma, S. Yuan, P. Cheung, K. Watanabe, T. Taniguchi, F. Zhang, and F. Xia, *Intelligent infrared sensing enabled by tunable moiré quantum geometry*, Nature, **604**, 266 (2022).
9. C. N. Lau, M. W. Bockrath, K. F. Mak, and F. Zhang, *Reproducibility in the fabrication and physics of moiré materials*, Nature **602**, 41 (2022).
10. F. R. Geisenhof, F. Winterer, A. M. Seiler, J. Lenz, T. Xu, F. Zhang, and R. T. Weitz, *Quantum anomalous Hall octet driven by orbital magnetism in bilayer graphene*. Nature, **598**, 53 (2021).

DMREF: Computational Chemistry to Accelerate Development of Long Wave Infrared Polymers

Lead Investigator: Jean-Luc Bredas, University of Arizona (all Co-PIs at UArizona)
Participating Institutions: none
Website: none (in progress, July 2022 launch date)
Keywords: infrared optics, infrared spectroscopy, optical polymers, machine learning

Project Scope

We propose to develop ***a new class of high refractive index, optical polymers*** that will enable the first use of ***low cost, plastic optics for long-wave infrared (LWIR) thermal imaging***. The key to accelerating materials design and optical-element fabrication will be the development of high-throughput computational methods combined with machine learning and the application of an integrated feedback loop with synthesis, processing, and optical characterization. The UArizona team will develop novel polymers derived from elemental sulfur that will possess enhanced LWIR transparency, along with robust thermomechanical properties to enable fabrication of viable IR plastic lenses for thermal imaging.

Relevance to MGI

The key to this advance will be the use of computational chemistry to rapidly simulate the infrared vibrational characteristics of candidate organic molecules, which when copolymerized with elemental sulfur, will enable creation of the desired IR optical polymers. Density functional theory (DFT), molecular mechanics (MM) and molecular dynamics (MD) will be utilized to create spectral databases from the "bottom up" with candidate molecules of interest that will possess limited vibrational absorption in the LWIR spectrum from 7-14 µm. Acceleration of candidate molecule discovery will be achieved using machine learning by accessing databases that contain millions of molecules. To understand the solubility parameters of molten liquid sulfur, computational studies on the chemical nature of molten elemental sulfur using a variety of methods, including Quantum Cluster Equilibrium (QCE) calculations are in progress. Promising predicted organic molecules of interest will then be synthesized in gram quantities to enable copolymerization with elemental sulfur to prepare new IR optical polymers. New high sulfur content polymers possessing the desired optical and thermomechanical properties will then be fabricated into optical elements, such as, lenses or windows and integrated into LWIR imaging systems for evaluation.

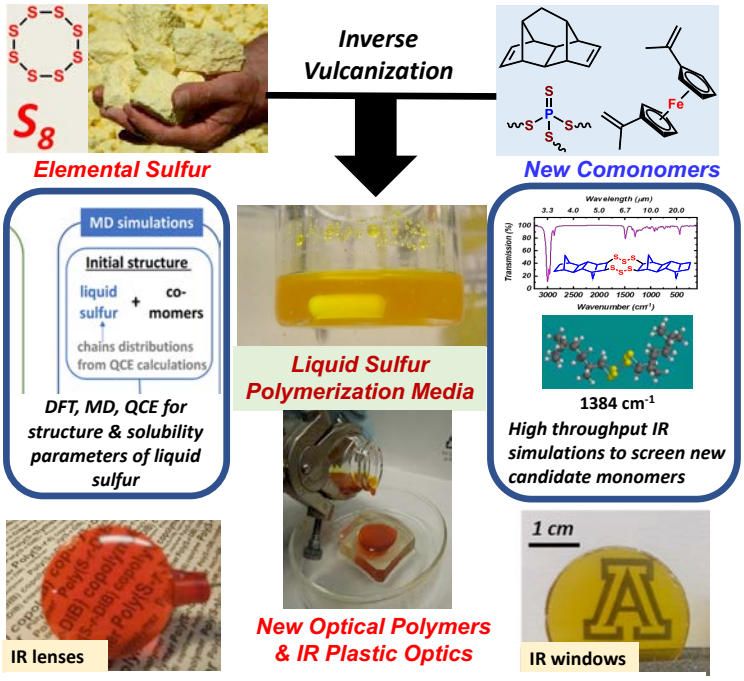


Fig. 1: Proposed research concept using computational chemistry & machine learning to accelerate the materials discovery of the first examples of LWIR polymers as optical components for LWIR thermal imaging. Computational efforts dedicated to understanding: (left) the chemical nature and solubility characteristics of liquid sulfur to expand scope of organic comonomers (right) high throughput IR simulations to predict and screen new candidate organic comonomers with improved IR transparency and solubility in sulfur for application for inverse vulcanization to prepare new IR optical polymers and fabricate IR plastic optical components such as lenses and windows.

Technical Progress

In Year-1 of the project, the UArizona team has made steady progress in the all of the proposed area of computations, organic synthesis, polymer synthesis and plastic optic fabrication with these new sulfur based materials. Below are summarized the key advances in achieved by appropriate technical grouping: ***Bredas &***

Licthenberger (Computations), ***Njardarson*** (organic synthesis), ***Pyun*** (polymer synthesis, processing), ***Norwood*** (optical element fabrication and optical characterization)

1) ***The Bredas group*** has developed a DFT model to describe for the first time a quantitative chemical description of photo-excited/relaxed speciation for liquid sulfur to describe the chemical origins of red coloration in polymeric sulfur and sulfur derived polymers from inverse vulcanization.
2) ***The Lichtenberger and Bredas groups*** have developed a new machine learning tool (Tucson Molecules to Materials Machine Learning, "T3ML", patent pending) for prediction of the LWIR transparency of organic molecules. Currently over 35,000 molecules have been computationally evaluated by T3ML for LWIR transparency, and over 130 1D and 2D constitutional, geometrical, auto-correlation, and intuitive descriptors have been integrated into this new ML model, with 3D descriptors in progress.
3) ***The Njardarson group*** has synthesized over 10 new organic compounds in the past year that include olefinic ferrocenes, organophosphorous compounds and cyclic olefinic monomers. Of particular interest has been the development of a new family of norbornene monomers derived from norbornadiene which are currently being evaluated with the Pyun group for copolymerization with elemental sulfur.
4) ***The Pyun group*** has explored a rigorous study on the chemical structure of sulfur derived polymers using NMR spectroscopic techniques in collaboration with ENI s.P.A as a means to interrogate the mechanistic aspects of the inverse vulcanization process. Collaborative efforts with the Bredas group has initiated computational calculations on possible intermediates and mechanistic pathways for the inverse vulcanization process
5) ***The Norwood group*** has developed polymer processing and nanofabrication methods for the first time with sulfur derived polymers to prepare integrated photonic devices and free standing optics. These methods are needed in anticipation of the availability of new DMREF optical polymers for the fabrication of LWIR optical elements. Processing methods to prepare free standing Fresnel lenses and microlens arrays have been demonstrated.

Future Plans

The immediate technical targets for Yr-2 of the project are the following:

1) ***Computational studies on liquid sulfur:*** the Bredas group will explore methods to characterize the chemical speciation of molten sulfur at temperatures required for inverse vulcanization to quantitate the solubility parameter for liquid sulfur.
2) ***Enhancement of the T3ML tool for IR screening:*** the Lichtenberger and Bredas groups will elaborate upon the existing ML model with more complex 3D descriptors and cross-reference predictions with known molecules to enable early deployment for new candidate monomer prediction
3) ***New LWIR polymers from inverse vulcanization:*** the Njardarson groups will continue rapid synthesis of new candidate molecules, an effort that has already demonstrated a five-fold increase in new candidate monomers that have been screened for inverse vulcanization than previously explored. The deployment of a new family of norbornadiene based sulfur polymers is anticipated for Yr-2.
4) ***On the mechanism of inverse vulcanization:*** the Pyun and Bredas groups will provide the first computational and experimental report on the mechanistic aspects of inverse vulcanization.
5) ***Prototype LWIR imaging systems and LWIR imaging standards:*** the Norwood group will create a new model LWIR imaging prototype system from new sulfur IR plastic lenses and commercially available LWIR microbolometer detectors to perform *in operando* characterization of new optical elements. There remains an urgent need to create the first LWIR imaging standard with set specifications for new prototype IR plastic optics.

Broader Impacts and Workforce Development

The UArizona DMREF team is strongly committed to the professional training of researchers on the project, the promotion of women/underrepresented minorities into STEM fields and wider communication of optical sciences to the community. From a training standpoint, the project is led by two women research scientists (Dr. Eunkyung Cho, Dr. Maliheh Tameh) and currently supports two women graduate students from the Pyun (Lindsey Holmen) and Norwood (Kate Newcomer) groups. In Year 1 of the project, the Lichtenberger group has focused on mentoring UArizona URM undergraduates (Taiya-Alvarez Williams, Victoria Munoz, Dorien Haney). In Summer 2022, two new UArizona undergraduates will be training on IR computational simulations (Satya Dulam, Sam Durfee) and one new high school student (Sophie Li) from the UArizona KEYS intern program. The impact of this collective

training will be significant as this group of students on the project will be presenting these results in both national and regional ACS meetings in 2023, and in the Southern Arizona Research, Science and Engineering Foundation (SARSEF) annual competitions for AZ high school students. The UArizona DMREF team website is under construction and anticipated to launch in July 2022, which will include tutorials on basic photonics concepts along with tutorials on IR thermal imaging.

Data Management and Open Access

New digital data and software for the T3ML tool is still being developed, so broader dissemination of this digital output is pending for 2023. We anticipate initially filing new IP on this software, followed by initial disclosure of results and data through publication in peer-reviewed scientific journals. The senior investigators retain the right to use the data before opening them up to wider use, with appropriate compliance with copyrights of published materials. For data generated covered under patents and other intellectual property, good faith negotiations with the senior investigators and the University of Arizona will be conducted to enable access to the data through licensing or other mechanisms. *It is the aim of the project to create new IR databases for use by the general public, which is planned to be posted on the UArizona DMREF site and linked to the NIST IR databases per future coordination.*

Advancing Along the Materials Development Continuum and Partnerships to Translation

The UArizona DMREF team is well-positioned for industrial engagement and translation of the work from this project. UArizona has a strong IP position on new optical materials through support of new IP by the State of Arizona through the UArizona technology transfer office, TechLaunch Arizona (TLA). Furthermore, the UArizona DMREF team has existing projects, or collaborations with multi-national companies from the optical and defense industrial sectors (PPG, LG CHEM, MOBASE, Raytheon, Hoya). Norwood and Pyun have also made new contacts with leading IR camera OEM's through NSF I-Corps to send samples of new prototype IR lenses and windows (II-VI, FLIR). Future industrial sponsors are also being pursued (Pyun will visit LG Chem, MOBASE, CLEX in Summer 2022). Finally, close partnerships with researchers at AFRL (Dr. Nicholas Godman) are in place to develop new IR plastic optics leveraging both AFRL facilities and access the wider AFRL network for potential partners for translation.

Publications and References

1. A. Nishant, K. Kim, S. Showghi, R. Himmelhuber, T.S. Kleine, T. Lee, J. Pyun, R.A. Norwood, *High Refractive Index Chalcogenide Hybrid Inorganic/Organic Polymers for Integrated Photonics*, Adv. Opt. Mater. *in press*, DOI:10.1002/adom.202200176 (2022)
2. C.F. Carrozza, S. Silvano, L. Boggioni, S. Losio, M. Guidolin, M. Tomaselli, M. Checchin, C. Olikagu, A. de Angelis, J. Pyun, W.O. Parker, Jr. *Structural Characterization of Sulfur Derived Copolymers from Inverse Vulcanization: Part 1: 1,3-diisopropenylbenzene*, Polym. Chem., submitted (2022)
3. J. Pyun, C.F. Carrozza, S. Silvano, L. Boggioni, S. Losio, M. Checchin, C. Olikagu, A. de Angelis, W.O. Parker, Jr., *Structural Characterization of Sulfur Derived Copolymers from Inverse Vulcanization: Part 2: Styrene*, J. Polym. Sci., submitted (2022)

DMREF: GOALI: Tetrahedral Ferroelectrics

Lead Investigator: Geoff Brennecka, geoff.brennecka@mines.edu
Participating Institutions: Colorado School of Mines, Carnegie Mellon University, The Pennsylvania State University, Broadcom **Website:** none yet
Keywords: Wurtzite, Tetrahedral, Ferroelectrics, Semiconductors, Polarization

Project Scope

Building off the recent demonstration of ferroelectric wurtzites, this project aims to design, develop, and deploy (with our GOALI partner) a new family of ferroelectric materials with tetrahedrally bonded crystal structures as candidates for direct integration into active semiconductor devices. The project tackles fundamental questions about the nature of polarization reorientation in tetrahedra (Fig. 1) through computation and thin film experimentation. A better understanding of the switching mechanism(s) will enable us to decouple lattice stiffness, coercive field, breakdown strength, and point/line defects to engineer ferroelectric function in tetrahedral materials.

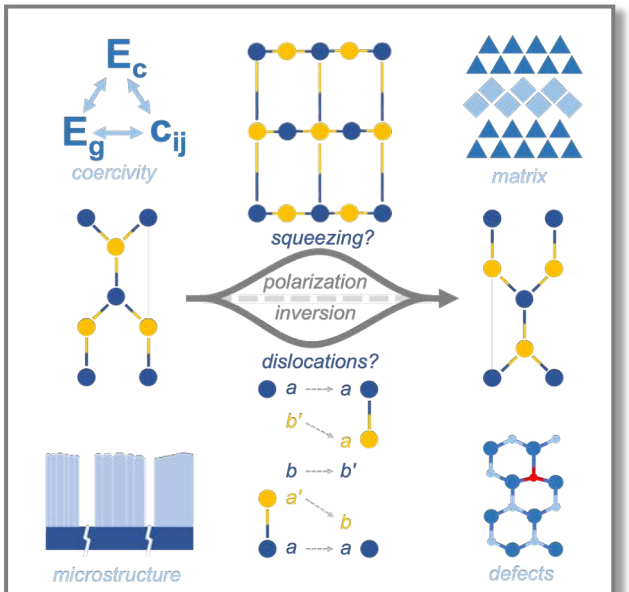


Figure 1. Tetrahedral polarization inversion is proposed to occur through either (top, center) a layered hexagonal prototype or (bottom, center) zinc blende ↔ wurtzite dislocation motion. Connecting switching mechanism(s) with (top corners) intrinsic and (bottom corners) extrinsic characteristics is key to engineering tetragonal ferroelectric semiconductors.

Relevance to MGI

This project operates in two modes that both integrate computation and experiment. Mode 1 targets knowledge generation with a focus on uncovering fundamental relationships. This mode hypothesizes new materials and/or structures via DFT-based calculations and chemical insight, then fabricates and characterizes them as sputtered thin films. The goal of Mode 1 is not to make the best material for any particular application but to understand specific relationships in polar tetrahedra to craft basic descriptors that will accelerate targeted development with close industry engagement in Mode 2. Trends in property descriptors are established using standard data science techniques (e.g., regression, unsupervised classifiers, and dimension reduction) as well as physical insights. These descriptors accelerate high-throughput predictions by turning large computational datasets into actionable guidance that feeds both combinatorial and targeted synthesis and testing, which in turn validate and/or refine the descriptors. Such descriptors are crucial for incorporating defect and alloy calculations that are essential to understanding measured properties of real materials. Promising candidate materials are further down-selected based on the expertise of our GOALI partner, and the pathway to industrial pilot testing is massively simplified and accelerated by close collaboration throughout the process.

Technical Progress

This project launched in June, 2022, but the team is already working closely together and extending preliminary results that launched the proposal. We are investigating bond ionicity as a key enabler of polarization inversion in tetrahedral ferroelectrics. Several prior reports have noted the relationship between Sc composition (x in $Al_{1-x}Sc_xN$) and the tetrahedral *c/a* ratio—or, equivalently, the *u* parameter denoting the cation position within the nitrogen tetrahedron—and it is frequently claimed that this *u* parameter is the primary descriptor of interest for reducing coercive field, E_c (and enabling polarization inversion when E_c falls below the breakdown field, E_b). Both Sc content and mechanical strain affect both *u* and E_c, but to date these have remained convoluted. However, calculations combined with growth and measurement of combinatorial thin films of $Al_{1-x}Sc_xN$ that exhibit similar *c/a* ratios for all compositions (Fig. 2) have enabled our team to decouple these factors and show that it is actually the local Sc-N bond ionicity that drives both the tetrahedral distortion and reduced coercive field rather than the Sc additions distorting the lattice and this distortion subsequently enabling ferroelectricity. This finding will be key to developing

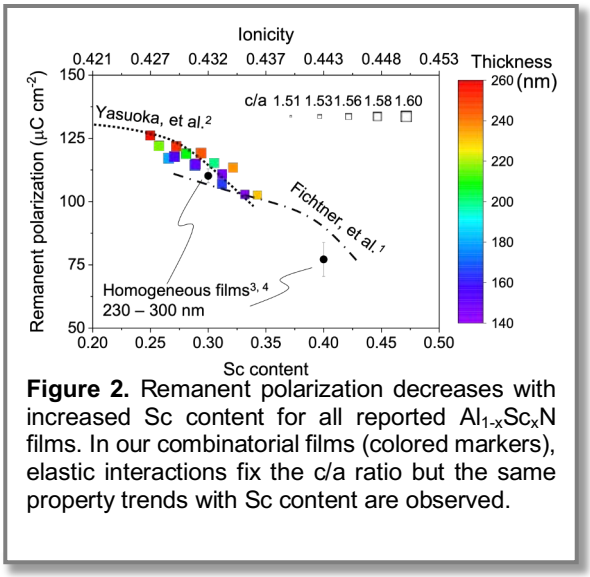


Figure 2. Remanent polarization decreases with increased Sc content for all reported $Al_{1-x}Sc_xN$ films. In our combinatorial films (colored markers), elastic interactions fix the c/a ratio but the same property trends with Sc content are observed.

new compositions of tetrahedral ferroelectrics and is a perfect example of the identification of a non-intuitive descriptor that can accelerate subsequent computation and discovery efforts.

Future Plans

We have identified three sets of fundamental scientific questions: (1) How is switchable polarization related to coercive field, a paraelectric transition, and/or stiffness? How is coercive field related to crystalline structure? (2) What is (are) the switching mechanism(s) in tetrahedral ferroelectrics? Do domain walls exist within single grains of tetrahedral ferroelectrics? (3) How do local structure, individual and clustered point defects, and synthesis affect switching and breakdown? Addressing these questions will expand our understanding of the fundamentals of ferroelectricity and of polar semiconductors alike.

One example of the kinds of fundamental relationships that we are investigating to enable practical deployment of tetrahedral ferroelectrics is the set of relationships among spontaneous polarization (P_s), lattice stiffness (c_{ijkl}), coercive field (E_c), and breakdown field (E_b). Reliable ferroelectric switching requires sufficient margin between E_c and E_b to allow operation without failure. Conventional wisdom is that stiff materials with high bond strengths exhibit both high E_b and E_c, and that structural defects reduce E_b but may increase or decrease E_c depending upon the dominant mechanisms at work. Showing a fundamental connection (or lack thereof) among such parameters will guide future searches for ferroelectric semiconductors and other such investigations.

Broader Impacts and Workforce Development

This project supports graduate student and/or post-doc training across three different institutions and will expose these students and the PIs in the full spectrum of the MGI philosophy from initial design and discovery through development to deployment via collaboration with our GOALI partner. This collaboration will also expose the students and post-docs to industrial R&D, preparing them for more informed employment decisions. Outcomes of the work will contribute to a greater understanding of the physics of ferroelectrics and polar semiconductors.

Data Management and Open Access

Materials data resulting from this work will be made available through the open MaterialsMine database (materialsmine.mines.edu) and the MDF. Datasets will also be indexed in the NIST Materials Data Repository. APIs, pre-trained ML models and associated training sets will be available on GitHub. Experimental data will be available through the NREL High Throughput Experimental Materials database.

Advancing Along the Materials Development Continuum and Partnerships to Translation

This project engages an industry leader from the start, helping to guide the work and to enable rapid transition of promising new materials. The involvement of an strong materials-enabled device manufacturer in the very early stages of materials discovery shows the power of the MGI approach; commercially deploying a new material by the end of this project would be a showcase win for MGI and DMREF.

Publications and References – no publications yet, as the project is less than a month old.

1. S. Fichtner, N. Wolff, L. Lofink, L. Kienle, and B. Wagner, *AlScN: A III-V semiconductor based ferroelectric*, J. Appl. Phys. **125**, 114103 (2019).

2. S. Yasuoka, T. Shimizu, A. Tateyama, M. Uehara, H. Yamada, M. Akiyama, Y. Hiranaga, Y. Cho, and H. Funakubo, *Effects of deposition conditions on the ferroelectric properties of* $(Al_{1-x}Sc_x)N$ *thin films*, J. Appl. Phys. **128**, 114103 (2020).

3. K. Yazawa, D. Drury, A. Zakutayev, and G L. Brennecka, *Reduced coercive field in epitaxial thin film of ferroelectric wurtzite* $Al_{0.7}Sc_{0.3}N$, Appl. Phys. Lett. **118**, 162903 (2021).

4. D. Drury, K. Yazawa, A. Mis, K. Talley, A. Zakutayev, and G. L. Brennecka, *Understanding Reproducibility of Sputter-Deposited Metastable Ferroelectric Wurtzite* $Al_{0.6}Sc_{0.4}N$ *Films Using In Situ Optical Emission Spectrometry*, PSS-RRL **15**, 2100043 (2021).

DMREF: A Data-Centric Approach for Accelerating the Design of Future Nanostructured Polymers and Composites Systems

Lead Investigator: L.C. Brinson
Participating Institutions: Duke University, University of Vermont, Northwestern University, Rensselaer Polytechnic Institute
Website: none
Keywords: Nanodielectric, MaterialsMine, Bayesian Optimization, Interface, Gaussian Process Modeling

Project Scope

The objective of this interdisciplinary DMREF is to use our unique data-driven design approach to create nanodielectrics able to support future technology in: e.g. flexible electronics, and efficient electrical generation and transmission. We have developed a multiscale modeling strategy that bridges length scales by using machine learning approaches, physics-based models, and includes a robust interphase model built using curated and custom generated data, as well as materials design methods and tools validated through test cases.

Relevance to MGI

The ability to design nanodielectrics would have far ranging impact on applications ranging from flexible electronics to electrical transmission. The number of design variables is large and there is a scarcity of data at some length scales. We have developed a design-based methodology using a data driven design approach grounded in physics-based models and experimental calibration. Specifically, we combine finite element modeling of dielectric constant and loss functions, and Monte Carlo and first principles simulation of carrier hopping for break down strength predictions which take the filler dispersion and interface properties explicitly into to account to compute objective functions for ideal nanodielectric insulators. Using Gaussian process meta-modeling, Bayesian optimization (BO), and multi-objective optimization on these computational predictions we can identify, for example, the Pareto frontier with respect to loading and dispersion of nanofillers for maximizing breakdown strength and minimizing the dielectric constant and loss tangents. The results of this design approach have provided insights into how to optimize properties and instigated new work with new filler geometries and filler modification with ligands that trap carriers.

Technical Progress

Our biggest accomplishments have been the integration of design, modeling, experiment, and interfacial control in an iterative knowledge discovery loop which accelerates materials discovery. The iterations can be seen in our publications that grew in sophistication in both modeling and design. For example, in our initial case study[1] for nanodielectric capacitors we demonstrated the integration of the processing-structure and structure-property relationships to design an optimal filler morphology for achieving multiple desired properties. (Figure 1) An inverse mapping back to optimal processing conditions was also demonstrated. The design scope was then extended to a more recent study[2] to concurrently identify optimal composition and microstructure morphology of insulating nanocomposite materials using a novel mixed-variable Latent Variable Gaussian Process (LVGP) approach combined with BO. We have used first-principles calculations of ensembles of amorphous polymer-nanofiller interfaces generated using molecular

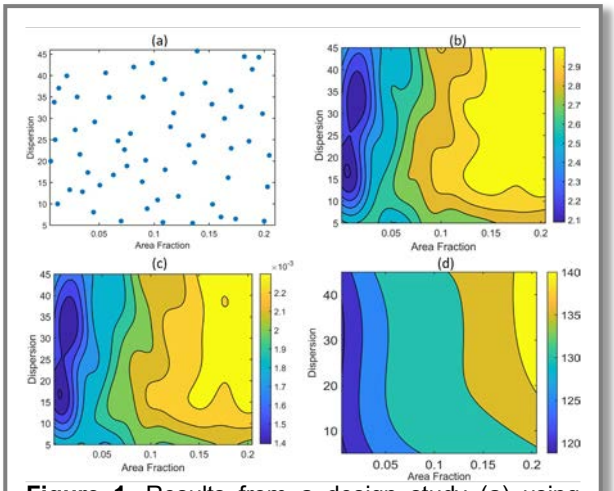


Figure 1. Results from a design study (a) using spherical nanofillers with several surface chemistries loadings and dispersion states showing the impact of dispersion on permittivity (b), loss (c), and breakdown strength (d). These results were used to create a pareto frontier and drive the next set of experiments.

dynamics to provide critical information regarding trap distributions[3]. We then perform Monte Carlo simulations of carriers hopping in an energy landscape determined by this trap distribution and the overall electrostatic potential in a given microstructure is computed using a finite-element method, as detailed in[4]. Importantly, this self-consistent electric field calculation accounts for the impact of filler dispersion and localized fields due to differences in matrix/filler permittivity. We have developed a novel experimental and machine learning approach to using electrostatic force microscopy (EFM)[5] to measuring local interfacial permittivity, and new machine learning approaches incorporating both experimental data and finite element simulations to reveal new relationships between microstructure, interfacial property gradients and macroscale mechanical and dielectric properties[6].

Future Plans

This project is nearing its conclusion. A critical challenge to be addressed is improved curation methods for incorporating existing data into MaterialsMine to increase the robustness and breadth of the design approach. New approaches that use artificial intelligence are needed in several junctures of the workflow.

Broader Impacts and Workforce Development

This project, which is a collaboration across four institutions and five research groups, has resulted in 14 Ph.D. students, 1 Post Doc, and 6 undergraduates learning, not only the importance and approach to MGI, but the skills for working in a truly interdisciplinary space. In addition, the results of this work have been presented in a range of venues and published in a broad set of journals. This has included presentations that include industry members and meetings with national laboratory scientists.

Data Management and Open Access

Our data management plan has resulted in adding significant data sets to MaterialsMine, and the storing of our codes either in Github or in a user-friendly version in MaterialsMine.

Advancing Along the Materials Development Continuum and Partnerships to Translation

The new design approach and critical new fundamental methods have significantly advanced the frontier for design of these complicated nanodielectrics. The integration with MaterialsMine ensures the data and approaches will be broadly available to the public. The work has been presented in DoD and industry venues. The methodology developed can be broadly used in nanocomposites.

Publications and References (out of 17 publications for this project)

1. W. Chen, L.S. Schadler, L.C. Brinson, Y. Wang, Y. Zhang, A. Prasad, X. Lin, A. Iyer, *Materials Informatics and Data System for Polymer Nanocomposites Analysis and Design*, chapter for Big, Deep, and Smart Data in the Physical Sciences, 2018.
2. A. Iyer, Y. Zhang, A. Prasad, P. Gupta, W. Tao, Y. Wang, L.C. Brinson, W. Chen, *Data centric nanocomposites design via mixed-variable Bayesian optimization*, Molecular Systems Design & Engineering ***5*(8),** 1376-1390 (2020).
3. A. Shandilya, L.S. Schadler, R. Sundararaman, *First-principles identification of localized trap states in polymer nanocomposite interfaces*, Journal of Materials Research ***35*(8)**, 931-939 (2020).
4. L.S. Schadler, W. Chen, L.C. Brinson, W. Sundararaman, P. Gupta, P. Prabhune, A. Shandilya, *A perspective on the data-driven design of polymer nanodielectrics*. Journal of Physics D: Applied Physics **53(33),** 333001 (2020).
5. P. Gupta, L.S. Schadler, R. Sundararaman, *Dielectric properties of polymer nanocomposite interphases from electrostatic force microscopy using machine learning*, Materials Characterization **173**, 110909 (2021).
6. Y Wang, M. Zhang, A. Lin, A. Iyer, A. Prasad, X. Li, Y. Zhang, L.S. Schadler, W. Chen, L.C. Brinson, *Mining structure–property relationships in polymer nanocomposites using data driven finite element analysis and multi-task convolutional neural networks*, Molecular Systems Design & Engineering **5**, 962-975 (2020).

Grain Interface Functional Design to Create Damage Resistance in Polycrystalline Metallic Materials

Lead Investigator: Curt A. Bronkhorst, cbronkhorst@wisc.edu.
Participating Institutions: Univ. of Wisconsin – Madison, Iowa State Univ., Univ. of New Hampshire.
Website: Under construction.
Keywords: Grain boundary, ductile damage mitigation, polycrystalline metal, interface, internal stress.

Project Scope

In high-purity tantalum, pores form at grain boundaries which are hypothesized to have a critical combination of high tensile stress and low tensile strength depending upon grain boundary type. Advanced single crystal theory will be developed to compute polycrystal internal stress state by coupling to an advanced macroscale ductile damage model. Grain boundary strength will be quantified via nanomechanical experiments and classical MD bicrystal calculations. Machine learning and uncertainty quantification tools will be developed to integrate experimental and computed data to prescribe probable metal forming pathways to reduce porosity by 30% in a high triaxiality sample.

Relevance to MGI

The workflow for our *accelerated material design cycle to develop advanced damage resistant material* is shown in Figure 1. *Task 1* includes material processing and sample production for grain interface functional design encompassing advanced multi-level experiment with a focus on single/polycrystal characterization and grain boundary/interface interrogation. Task 1 will receive processing recipes from Task 3 and numerical simulations from Task 2. *Task 2* includes multi-level physical theory and code development along with numerical simulations. Task 2 will receive physical data from Task 1 and uncertainty quantification from Task 3. *Task 3* provides the statistical and uncertainty framework to enable rapid integration of experimental and computational data sets and statistical model development. Task 3 will receive computational results from Task 2 and experimental results from Task 1. Throughout this iterative process, improvements to all aspects of each task will be pursued. Our *measure of success* is to design and produce grain interface functionally designed polycrystalline Ta demonstrating *30% reduction in porosity* at equal deformation for the proposed multi-level, high-triaxiality sample in comparison to this projects reference rod and plate materials. Information integration of both advanced experimental and advanced physics-based computational origin will be critical to accelerated design.

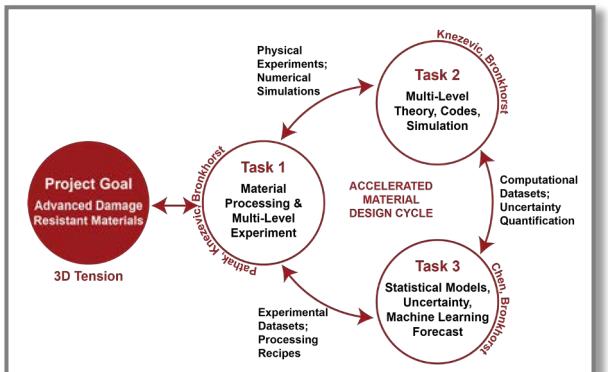


Figure 1. Flow of information between the four tasks to cyclically refine the design of damage resistant material. The proposed work combines both discovery and material design elements to achieve the project goal.

Technical Progress

All technical elements of the project are in place and progressing. Nanoindentation experiments have been conducted on single crystal tantalum to develop the proper surface preparation process and produce data to compare against simulations of indentation using our local and strain-gradient single crystal models for Ta. A strain-gradient single crystal model was implemented into ABAQUS to enable more accurate representation of stress states at grain boundaries for polycrystal simulations and also the length-scale effect in micro-pillar compression experiments. A bi-crystal MD data set was assembled from the literature assessing grain boundary strength under tensile states of stress. LAAMPS MD models for bi-crystal grain boundary strength simulations were acquired and further developed to enable the simulation of grain boundaries. The simulations will be compared with micro-tension experiments on grain boundaries for physical trend validation. A first series of rolling experiments on Ta were conducted at Ames National Laboratory to gain practical experience with the Materials Preparation Center and produce processed material samples for metallographic examination and compare against polycrystal simulations

of rolling which are underway. We have also established successful representation of polycrystal internal stress distribution evolution during complex loading history by the maximum entropy principal using the leading four moments of mean, variance, skewness, and kurtosis.

Future Plans

Characterization of rolling experiment samples is underway – four rolling passes with sample material collected after each pass with the final pass achieving -1.57 strain. Temperature-time recovery experiments will be conducted on rolling samples to assess recovery/recrystallization limits as a function of rolling intensity. Large polycrystal plane-strain compression simulations will be conducted of each rolling pass using both local and strain-gradient single crystal models for validation and quantify evolution of microstructural and internal stress distribution with deformation. A more advanced and physically improved local single crystal model for Ta is under development and will be engaged for this project. The strain-gradient single crystal model will continue development and will focus initially upon nanoindentation simulations in parallel with single crystal Ta experiments where strain-gradient influence is known to occur. Micro-tension experiments will also begin to probe the tensile strength of grain boundaries and validate physical trends observed via bicrystal MD simulations. MD simulations will continue for the life of the project to add computed results to this grain boundary tensile strength dataset. Statistical model development will continue with emphasis to efficiently represent the probability density function evolutions of internal stress state and grain boundary strength with deformation. This will continue by introducing grain boundary misorientation into the computed and experimental datasets as a potentially important variable for microstructure design. A macroscale porosity-based ductile damage model will be developed for low deformation rate loading and will be used to simulate macroscale experiments; including the targeted high-triaxiality tension sample. This model is being developed with a new statistically based pore nucleation model based upon work of this project. This model will also be used to provide evolving stress conditions for micromechanical study at locations in the samples where damage probability is high.

Broader Impacts and Workforce Development

The educational impact of this work will be integrated across the three associated institutions by organizing targeted student exchanges among the PIs. The PI and three co-PIs will sponsor guest lecturing rotations within the courses they teach, as a cross-disciplinary effort, the shared and complementary expertise will be brought to the classroom. Graduate students from each research group will spend time between the three different institutions as part of each of the Ph.D. programs as well as possible internships at AFRL. Additionally, the UG capstone senior design courses at each department would particularly benefit from this arrangement, where the students will be mentored by experts from modeling, mathematics and experiments. The PIs are eager to build a very strong long-term collaborative relationship with AFRL and have existing collaborations with multiple U.S. National Laboratories to provide the students in the team with direct access to state-of-the-art user facilities, including their world class training on safety and research, and world-class staff. The PIs plan to provide a multitude of research opportunities *for upper level undergraduate students* (~8 students/year) via direct funding.

Data Management and Open Access

Once available, all results will be archived to a project identified repository on the UW Libraries MINDS@UW system which provides long-term preservation and open access to digital materials. This project will follow the best practice protocols established for the MINDS@UW system which adhere to FAIR data management and stewardship principles. The project website will direct the community to locations of all data and code.

Advancing Along the Materials Development Continuum and Partnerships to Translation

The material production processes targeted in this project are generally available commercially and therefore the emphasis will be upon building the material design framework and methodology to enable more rapid manufacturing process design. The tools developed in this project are envisioned to be adapted in time to any material system or manufacturing process where both computational and experimental datasets are together used to determine high probability of success manufacturing procedures.

Publications and References

This project began Jan. 1, 2022 and publications have not yet ready for submission.

DMREF: Collaborative Research: DNA-based sensing, communicating, and phase-separating materials

Lead Investigator: Carlos Castro, castro.39@osu.edu
Participating Institutions: The Ohio State University, Duke University
Website: https://dmref.org/projects/1509
Keywords: DNA nanotechnology, DNA origami, single molecule dynamics, biomolecular materials.

Project Scope

This work aims to develop robotic DNA materials with dynamic sensing, communicating, and phase separating functions. The PIs have established approaches to actuate complex motion, tune conformational dynamics, and model the mechanical and dynamic properties of DNA origami (DO) structures. Building on this foundation, and guided by computational models and single-molecule characterization tools, we are developing robotic devices that leverage well-defined dynamic functions to measure forces with femtoNewton sensitivity, propagate mechanical signals over micrometer distances, and self-organize over assemblies of ~10 or more devices.

Relevance to MGI

This project integrates molecular modeling, DNA-based design, and single molecule methods to develop nanorobotic materials based on assemblies of dynamic DNA devices, where dynamic properties and interactions of devices drive collective behaviors of the assembly. The design of individual DNA nanodevices is guided by coarse-grained molecular dynamics simulations used to predict structure shape and properties such as motion or stiffness. These properties are characterized by transmission electron microscopy and single molecule methods like single molecule Förster Resonance Energy Transfer (smFRET), and experimental results are used to refine designs and modeling approaches. At the assembly scale, we closely couple experimental fabrication and characterization with a combination of coarse-grained molecular dynamics simulations, Brownian dynamics simulations, and theoretical modeling to achieve desired collective material behaviors (i.e. signal propagation, self-organization). We have developed opensource software tools, which advance the state-of-the-art in DO design. Using these design approaches, we are establishing the ability to propagate a signal via multiple reconfigurable components and assemblies that exhibit phase-transition behaviors to drive self-organization of many units.

Technical Progress

This work focuses on: 1) simulation guided DO design; 2) dynamic interactions and signal transmission in DNA devices; 3) coordinated behavior of many devices within assemblies to achieve long-range communication or domain formation; and 4) coarse-grained models of DO to access longer time scales of dynamic behavior and assembly. We rely on integration of design, fabrication, experiments, and modeling. We developed new design approaches embedded in software tools [1] to facilitate simple and rapid iterative design with feedback from coarse grained molecular dynamics simulations [2]. We recently advanced these approaches to enable design of structures with "freeform" shape (i.e. curvature in 3D) as illustrated for the trefoil knot (Fig 1A).

Using these design tools, we constructed DO devices with controllable mechanical properties to test mechanisms for dynamic functions, including two mechanisms for communication. The first relies on structural rearrangements of individual devices allowing triggering of a strand displacement reactions on a neighboring device. We demonstrated effective communication over neighboring devices and are optimizing communication over long-range assemblies up to ~10 μm long. A combination of Brownian dynamics and theoretical modeling [3] suggest these assemblies could speed up communication several-fold over prior systems that leverage immobilized DNA strands. We also developed signal transmission approach that exploits thermal fluctuations where a mobile component can influence the conformation

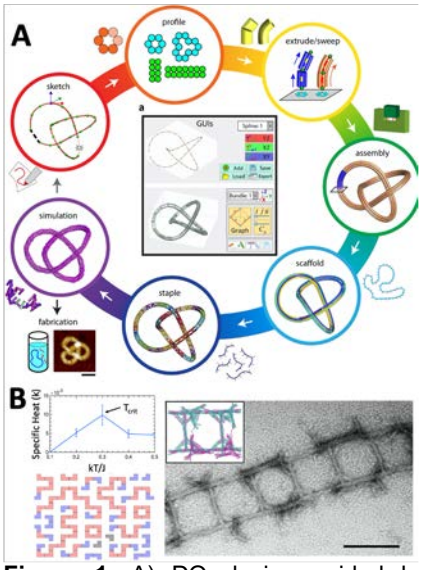


Figure 1. A) DO design guided by simulation. B) DO assemblies that exhibit cooperative behaviors such as domain formation can lead to material behaviors like phase transitions as predicted by simulations (left).

of another nearby mobile component. We demonstrated this steric communication in a single device consisting of a base platform with two "paddles," showing that the conformation of one paddle regulates the conformation of the neighboring paddle. Next steps will expand into arrays. Finally, we are developing DNA assemblies that exhibit cooperative behaviors. Specifically, we developed "spin" nanodevices that have a rotor component anchored to a base-platform. The base-platform allows for assembly in 1D (i.e. filaments) or 2D (i.e. sheets). The rotor can rotate 360 degrees and can interact with neighboring units in particular orientations. These local interactions drive collective material behaviors like phase transitions between ordered and disordered states (Figure 1B).

Currently, we have leveraged existing simulation tools to guide design of individual devices. However, these tools only capture millisecond time-scale behaviors at most, which is insufficient to describe dynamic interactions of multiple devices or study processes like assembly. Hence, our team is developing a new coarser model of DO structures that should allow modeling of processes that occur at timescales of up to ~0.1-10 seconds, which will allow direct comparison to experimental measurements, such as smFRET (time resolution of ~0.1 seconds).

Future Plans

In addition to publishing key studies, we will continue development of dynamic arrays that transmit signals via cascades of strand displacement reactions. Initially, we are using DNA strands as triggers for the signal transmission. We also plan to explore other triggers such as force or light. We will also translate the steric communication to arrays with a focus on demonstrating reversible signal transmission. For both systems we will quantify length and timing of communication. We also plan to build junctions and logic units to build towards more advanced structural computing systems. We also plan to finish the development of the phase changing assemblies and coarse-grained model of DO devices and assemblies.

Broader Impacts and Workforce Development

We developed DO design tools and held workshops to support use with researchers at OSU, Arizona State University, St. John's University, Georgia Tech, and Yale. We manage a Slack channel where users can get support on software (available on github) use. To support training of younger students, we ran a two-week Summer workshop where ~30 first year students at OSU learned to design DO and run simulations using online tools. Based on our work, we published a perspective article on dynamic DNA nanotechnology that won the "*Nanoscale Horizons* Outstanding Review 2020" [4]. Finally, we developed multiple modules for folding and analyzing DO structures that can be carried out in elementary, high school, or undergraduate classrooms [5].

Data Management and Open Access

All software is available on github (https://github.com/cmhuang2011/MagicDNA). Use has been supported with workshops and tutorials on youtube (https://www.youtube.com/channel/UCpI3shjsdy89Xg2iVt-ZYCw). We also manage communication among our team and with users through Slack. Several of our published device designs have also been uploaded to the DNA nanostructure design repository (e.g., https://nanobase.org/structure/70).

Advancing Along the Materials Development Continuum and Partnerships to Translation

We are developing materials well-suited for applications in biosensing, nanocomputing, nanomanufacturing, and nanorobotics. The PI is working towards translation of a biosensor device that is based on earlier NSF-supported developments. The work in progress through this grant can provide a number of key advancements in the areas of multiplexing, intelligent detection (i.e. signal processing), and signal amplification. To realize these advantages a critical future step will be integration of the DNA assemblies with other micro-fabricated or material systems.

Publications and References

1. C.M. Huang, A. Kucinic, J.A. Johnson, H.J. Su, and C.E. Castro, *Integrated computer-aided engineering and design for DNA assemblies*, Nature Materials, **20**: 1264-1271 (2021)
2. Z. Shi, and G. Arya. *Free energy landscape of salt-actuated reconfigurable DNA nanodevices*, Nucleic acids research **48.2**: 548-560 (2020).
3. S. Sensale, P. Sharma, and G. Arya. *Binding kinetics of harmonically confined random walkers,* Physical Review E **105**: 044136 (2022).
4. M. DeLuca, Z. Shi, C.E. Castro, and G. Arya, *Dynamic DNA nanotechnology: toward functional nanoscale devices*. Nanoscale Horizons, **5(2)**, 182-201 (2020)
5. P. E. Beshay, A. Kucinic, N. Wile, P.D. Halley, L. Des Rosiers, A. Chowdhury, J.L. Hall, C.E. Castro, M.W. Hudoba, *Translating DNA origami Nanotechnology to Middle School, High School, and Undergraduate Laboratories*, (submitted)

QMC-HAMM: High accuracy multiscale models from quantum Monte Carlo

Lead Investigator: David Ceperley (ceperley@illinois.edu)
Participating Institutions: University of Illinois at Urbana-Champaign, Rice University
Website: qmc-hamm.github.io
Keywords: quantum Monte Carlo, bilayer graphene, electronic structure, multiscale models

Project Scope

The QMC-HAMM project is focused on connecting highly accurate many-body quantum Monte Carlo (QMC) methods to larger length scales. We are focused on two main target problems: twisted bilayer graphene and hydrogen at high pressure. These problems share important aspects in that there is a subtle connection between atomic structure and electronic behavior, and that electron correlations are critical to being able to predict the behavior of the materials. Our objective is to provide high quality and vetted data and coarse-grained models that can be used to understand and predict the properties of these materials.

Relevance to MGI

This project is primarily computational in nature, and, in line with the Computational Materials Sciences award, is focused on developing tools and data to support other MGI projects.

Technical Progress

We have focused on two major types of advances: new techniques to link atomistic scale simulations to mesoscale models and publication of data that enables the above connection. In the figure we present a recent application of these ideas,[3] in which machine learning methods were used to derive tight-binding models for arbitrary geometric configurations of bilayer graphene. As shown in the figure, these models reproduce almost exactly the density functional theory band structure for twisted bilayer graphene, a configuration on which the model was not trained. The model, called local environment tight binding, is a many-body tight-binding model. It can be used in Python by simply typing "pip install bilayer_letb." We further used this model to assess the errors in approximate tight-binding models to analyze magic-angle twisted bilayer graphene, and found that conclusions change when using a more accurate model.

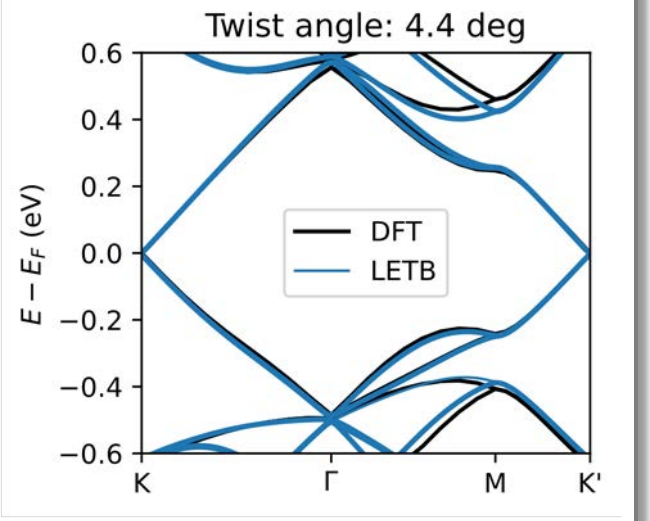


Figure. By basing larger length scale models on high accuracy first principles data, we can improve the relevance of these models to experiment. This is a tight-binding model derived using machine learning methods that obtains density functional theory accuracy for twisted bilayer graphene with

To enable high accuracy data collection, we have improved the quantum Monte Carlo suite of techniques, in particular optimization of complex wave functions for both ground[1] and excited states[2]. Excited states have long been an Achilles' heel of Monte Carlo technique, and as a result of our collaboration, we now have efficient methods to optimize compact and accurate wave functions, even for materials in which the electrons are strongly correlated.

We have also developed several models at the mesoscale for twisted bilayer graphene, which can be connected to

atomistic calculations, such as an efficient method for estimating the density of graphene[5] and an elastic model that explains the relative sizes of AA and AB sections in twisted graphene[4].

Future Plans

We are finalizing a holistic model for twisted bilayer graphene that builds on the LETB model above. This model includes data on the interlayer interaction derived from quantum Monte Carlo, which explicitly treats the complex van der Waals interaction between the layers, and changes the relative energy cost of different stackings compared to standard models.

For the hydrogen part of the project, we are preparing to publish a large database of QMC forces and energies, along with ML-based models of this data that will enable inexpensive and highly accurate simulations of hydrogen at high pressure.

Broader Impacts and Workforce Development

This project is training a cohort of graduate students and postdocs in the application of data to scientific problems. The graduate students are also being trained not just in the technical aspects of their respective disciplines, but also to communicate across disciplines, since our collaboration spans from physics to mechanical engineering. We believe that this type of training, enabled by this project, prepares the graduate students for future employment. We have sent graduate students from our collaboration to national labs, the Flatiron institute, and other postdoctoral positions, and they are carrying the principles of how to publish open data and code with them.

Data Management and Open Access

Our website is qmc-hamm.github.io, and our github repositories are located at github.com/qmc-hamm. As data and code are made available, they are published as repositories in our github organization. In many cases, we create a Python package that is easily installable, and provides a convenient interface to the data and models. For example, the LETB model generates a PythTB model given any bilayer structure.

Advancing Along the Materials Development Continuum and Partnerships to Translation

This project is quite fundamental and thus is several levels removed from commercialization. However, our data is of high interest for the purposes of materials design. We are paving the way to higher accuracy models that can then be used to design new materials, such as functional semiconductors, correlated electron materials, and potentially applications such as drug design.

Publications and References

1. S. Pathak, L.K. Wagner. *A light weight regularization for wave function parameter gradients in quantum Monte Carlo*. AIP Advances **10**, 085213 (2020).
2. S. Pathak, B. Busemeyer, J. Rodrigues, L.K. Wagner. *Excited states in variational Monte Carlo using a penalty method*. J. Chem. Phys. **154**, 034101 (2021)
3. S. Pathak, T. Rakib, R. Hou, A. Nevidomskyy, E. Ertekin, H. Johnson, L.K. Wagner. *Accurate tight-binding model for twisted bilayer graphene describes topological flat bands without geometric relaxation*. Phys. Rev. B **105** 115141 (2022)
4. E. Annevelink, H.T. Johnson, E. Ertekin. *Topologically derived dislocation theory for twist and stretch moire' superlattices in bilayer graphene*. Phys. Rev. B **102**, 184107 (2020)
5. T. Rakib, E. Ertekin, P. Pochet, H.T. Johnson. *An efficient computational framework for charge density estimation in twisted bilayer graphene*. Computational Materials Science **199**, 110746 (2021)
6. C.F.B. Lo, H.C. Po, A.H. Nevidomskyy. *Inherited topological superconductivity in two-dimensional Dirac semimetals*. Phys. Rev. B **105**, 104501 (2022)

Denoising and feature extraction in photoemission spectra with variational auto-encoder neural networks

Lead Investigator: Utpal Chatterjee
Participating Institutions: University of Virginia
Website: None
Keywords: Photoemission spectroscopy, denoising, feature extraction, variational auto-encoder, strongly correlated materials, Green's function

Project Scope

This particular work stems from our collaborative DMREF project "DMREF: Collaborative Research: Accelerated discovery of chalcogenides for enhanced functionality in magnetotransport, multiorbital superconductivity, and topological applications". The main objective of our DMREF project is to establish new paradigms for multi-band systems with moderate correlations and spin-orbit coupling. Some of the expected outcomes of this proposal are: (1) Discovery and optimization of novel electronic phases with unusual magneto-transport properties; (2) Discovery of topologically protected surface states in the background of textured magnetic phases; and (3) Emergence of unconventional superconductivity in insulators or multi-band low density transition metal dichalcogenides (TMDs).

Relevance to MGI

The spectral function $A(\vec{k},\omega)$, which is a function of momentum ($\vec{k}$) and energy (ω) provides direct access to the microscopic nature of the interactions in a material. This can be obtained from ARPES intensity. The main issue towards extraction of $A(\vec{k},\omega)$ from ARPES intensity is overcoming the challenge to deconvolve $A(\vec{k},\omega)$ out of resolution function $R(\vec{k}-\vec{k'},\omega-\omega')$. So far, various methods, such as resolution-broadened fermi function division, symmetrization and maximum entropy-based Lucy–Richardson method have been used to accomplish this. Current work using machine-learning based approach is an important step towards the requisite deconvolution using a procedure that is very different and hopefully more rigorous than the above-described methods, in most of which the effect of the convolution operation is entirely encapsulated by replacing actual temperature by an effective temperature. We anticipate that the stability of the current analysis can be enhanced by adding physical constraints through sum rules, which has not been tried so far for the analysis of ARPES data. If this succeeds, it will be very helpful towards quantitative comparisons between Quantum Monte Carlo based calculations and experimental data, an important step towards accomplishing close looped and iterative integration of experimental data, theory, and computational works in complex materials.

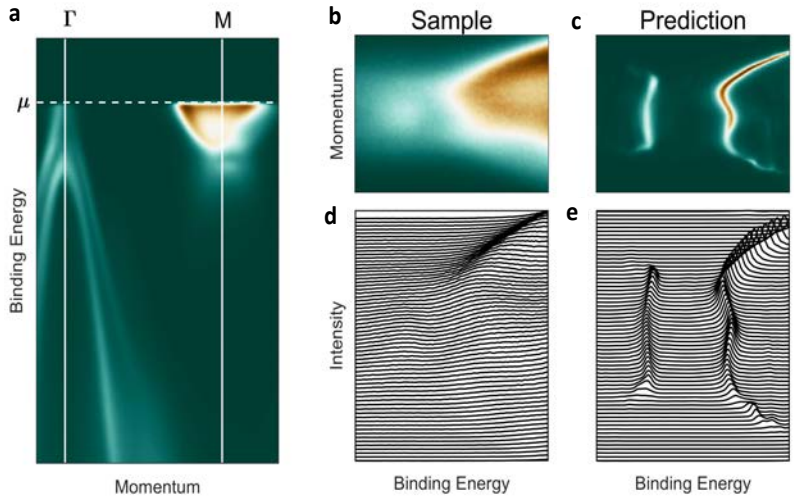


Fig. 1:Denoising and feature extraction from $1TiSe_2$ ARPES data: The proposed VAE produces sharper peaks, thus highlighting the dispersion of the upper branch and the backbending feature of the lower branch The VAE also reduces noise present in the raw data. (a)Wide energy-momentum dispersion in the $\Gamma-\bar{M}$ direction. The magenta square indicates the patch selected as input to the network. The original and predicted images are shown in panels (b) and (c), respectively (after a clockwise, 90° rotation). (d) – (e) Energy distribution curves corresponding to panels (b) and (c), showing the charge density wave gap and the dispersion more clearly, especially in panel (e). Figure is adopted from: *Review of Scientific Instruments* **93**, 065106 (2022).

Technical Progress

As a part of our multidisciplinary study of strongly correlated matter, we have developed analytical tools for data analysis based on neural networks. In particular, we proposed a shallow variational auto-encoder network to tackle noise and resolution-broadening issues in Angle Resolved Photoemission Spectroscopy (ARPES) data. While distinct neural networks have been designed in the past to address these two issues separately, the shallow network we propose can perform both tasks if trained separately for each. As to the best of knowledge, this work is the first one, where a single neural network was trained to

perform multiple tasks.
Our results indicate that denoising by itself is helpful in revealing subtle aspects of the electronic dispersion, like renormalizations of spectral line shape due to electronic correlations. We also note that denoising can be helpful for processing data which were collected with low statistics, due to either low photon flux or short acquisition times. Moreover, the reliability of data analysis procedures like Fourier transform, autocorrelation, and Kramers-Kronig transformations, which are routinely conducted on data from ARPES and other spectroscopic probes, depends critically on the noise level of the data. We also explored the possibility of simultaneously denoising and sharpening spectral peaks in the data, using the same neural network. Using ARPES data from the transition metal dichalcogenide 1T-$TiSe_2$ as a case study, we demonstrated that this is indeed possible. Our proposed VAE could sharpen the two Bogoliubov-like branches at the $\bar{M}$ point of the Brillouin Zone, which emerge because of the particle-hole mixing in the charge-density wave phase (see Fig. 1). We expect that the enhancement of such spectral features will be instrumental in elucidating the effects of electron interactions in these and other exotic materials. We also believe that the present work is particularly pertinent to ARPES data collected using soft and hard x-ray, where a lack of high-resolution data stymies reliable extraction of band parameters, making it possible to harness the potential of bulk probe ARPES to its fullest extent. This approach can be extended to the analysis of other spectroscopic data.

Future Plans
(i) It is naturally desirable to extend this work beyond peak sharpening to an exact energy deconvolution of the data. We believe this is possible by using real data during training instead of synthetic data, which we aim to accomplish in our future efforts. (ii) We plan to incorporate restricted Boltzmann machines with the goal to extract functional form of the self-energy from ARPES data. Such an approach, if successful, will be instrumental towards obtaining self-energy from the experimental data in a model-independent way. (iii) With the help of our theory collaboration, we also aim to extract real frequency dynamics from quantum Monte Carlo data on the single particle Green's function and the spin-spin correlation function in imaginary time using a machine learning artificial neural network that can be trained. Once this is done, we hope to compare them with available ARPES experimental data.

Broader Impacts and Workforce Development
In accordance with the NSF objectives, the overarching goal of our DMREF project has been to contribute in (a) the advancement of discovery and understanding while promoting teaching, training and learning, (b) the development of world class science, technology, engineering and mathematics (STEM) work force in the country, (c) the broadening of the participation of underrepresented groups. In the large and heavily researched field of metal-insulator transitions (MIT), our DMREF project introduced a novel mechanism, where MIT is driven in Pt doped 1T-$TiSe_2$ quasi-two-dimensional CDW materials by domain wall formation and by the creation of a network of strongly interacting Luttinger channels. As to the workforce development activities, we have provided research opportunities and educational trainings in the crosscutting areas of materials science and condensed matter physics to the participants in the project. Previously, Dr. Li, the postdoc in our group at UVA, developed a home-made Molecular Beam Epitaxy (MBE) instrument. Experience gained from this project was beneficial to him for finding a job in industry. Dr. Zhao, worked as a temporary researcher in this project, took a permanent position to work as a lab specialist at UVA physics instructional lab. For a short time, Dr. Francisco Restrepo worked on the machine learning-based data analysis project. Dr. Restrepo lead the machine-learning based ARPES data analysis project and developed necessary codes for the data analysis. Additionally, several undergraduate students were also actively involved in this DMREF project.

Data Management and Open Access
The neural network we have developed for denoising and peak sharpening was trained with synthetic data; that is, data numerically generated from suitable models, including resolution broadening and measurement noise. The instructions for generating such data sets are detailed in our publication referenced below [1]. The MATLAB codes for generating the training data and the PYTHON codes for defining and training the network can be found in the repository https://github.com/Francisco-UIC/Machine-learning-in-ARPES.

Advancing Along the Materials Development Continuum and Partnerships to Translation
Quantum Monte Carlo (QMC) has proved to be a crucial theoretical tool that can be used to obtain single particle spectral function and the spin-spin correlation function but only in imaginary time. How to obtain these dynamical

functions in real frequency that can be connected to experimental data remains an outstanding issue in the field of condensed matter physics and materials science. This particular work together with the planned activities for the future is expected to accelerate the development of appropriate machine learning-based algorithms to address this challenge.

Publications:

[1] F . Restrepo, J. Zhao, U. Chatterjee, "Denoising and feature extraction in ARPES spectra with variational auto-encoder neural networks", *Review of Scientific Instruments* **93**, 065106 (2022).
[2] K. Lee, J. Choe, D. Iaia, J. Li, J. Zhao, M. Shi, J. Ma, M. Yao, Z. Wang, C.-L. Huang, M. Ochi, R. Arita, U. Chatterjee, E. Morosan, V. Madhavan, and N. Trivedi, *Metal-to-Insulator Transition in Pt-Doped TiSe2 Driven by Emergent Network of Narrow Transport Channels*, Npj Quantum Ma- ter. **6**, 8 (2021).
[3] A. Wegner, J. Zhao, J. Li, J. Yang, A. A. Anikin, G. Karapetrov, K. Esfarjani, D. Louca, and U. Chatterjee, *Evidence for Pseudo–Jahn-Teller Distortions in the Charge Density Wave Phase of 1T−TiSe2*, Phys. Rev. B **101**, 195145 (2020).
[4] J. Zhao, K. Lee, J. Li, D. B. Lioi, G. Karapetrov, N. Trivedi, and U. Chatterjee, *Spectroscopic Fingerprints of Many-Body Renormalization in 1 T−TiSe2*, Phys. Rev. B **100**, 045106 (2019).

References specific to this work:

[1] H. Peng, X. Gao, Y. He, Y. Li, Y. Ji, C. Liu, S.A. Ekahana, D. Pei, Z. Liu, Z. Shen, and Y. Chen, Review of Scientific Instruments **91**, 033905 (2020).
[2] C. Dong, C.C. Loy, K. He, and X. Tang, IEEE Transactions on Pattern Analysis and Machine Intelligence **38**, 295 (2016).
[3] Y. Kim, D. Oh, S. Huh, D. Song, S. Jeong, J. Kwon, M. Kim, D. Kim, H. Ryu, J. Jung, W. Kyung, B. Sohn, S. Lee, J. Hyun, Y. Lee, Y. Kim, and C. Kim, Review of Scientific Instruments **92**, 073901 (2021).

Design and Synthesis of Novel Magnetic Materials

Lead Investigator: James R. Chelikowsky, jrc@utexas.edu
Participating Institutions: University of Texas, Iowa State University, and University of Nebraska
Website: https://novomag.physics.iastate.edu/.
Keywords: Permanent magnets, rare earths, adaptive genetic algorithms, magnetic materials databases

Project Scope
We implement transformative strategies for the design of novel magnetic materials with an emphasis on non-critical elements. Our work couples a strong experimental effort in synthesis and characterization with theoretical advances in quantum-modeling algorithms and software, machine learning (ML) and materials informatics, and high-performance hardware to accomplish our objective. In particular, we target new rare-earth-free magnetic compounds with high magnetization, high Curie temperatures, high spin polarization and high magnetic anisotropy. Materials with these properties will have significant applications in ultra-small spintronics devices, new high-density data-storage schemes and high-energy-product permanent-magnets.

Relevance to MGI
The discovery of new magnetic materials is a formidable problem, especially given the number of possible combinations of composition and structure. Our work represents a different paradigm to accelerate the discovery process. We use computationally driven phase-space explorations for composition-structure-property relationships coupled with experiment to identify candidates that have desirable properties. A ML-assisted adaptive genetic algorithm coupled to first-principles codes specifically designed for magnetic properties is used for property predictions. Our algorithm provides speed and efficacy for fast high-throughput explorations, while maintaining the accuracy of quantum-based calculations. We developed a comprehensive magnetic materials database to facilitate the materials informatics and ML guided materials design and discovery. Experimental research focuses on novel synthetic techniques and a comprehensive set of characterization methods. Nanoscale clusters and particles are produced by nonequilibrium inert gas-condensation techniques as well as sputtering and pulsed-laser-deposition methods. Ultra-fast quenching from the melt also produce bulk materials and new structures. These techniques are particularly important for fabrication of metastable phases. Structural characterization are performed with x-ray and neutron diffraction, and high-resolution electron microscopy; magnetic and electronic structure studies are pursued with magnetization, x-ray magnetic circular and dichroism.

Technical Progress
Finding appropriate compositions of ternary and quaternary alloys that can exhibit desired magnetic properties is difficult as the number of possible combinations of compositions and structures is enormous. Our research focuses on the discovery of new rare-earth free magnetic compounds with high magnetocrystalline anisotropy K_1, large high saturation magnetic polarization J_s, and high Curie temperature T_c. Recently researchers at Iowa and Texas have developed a machine learning (ML)-guided crystal graph convolutional neural network (CGCNN) to quickly examine the energetic stability of about 370,000 ternary compounds. This process has identified 4,342 promising structures (less than 1.2% of the total ML-screened structures). These structures were used for further investigation using density-functional-theory (DFT) calculations. The promising compositions suggested by ML and DFT calculations are used as input for an adaptive genetic algorithm (AGA) search, which uncovers more new low-energy structures.

The combined ML- and AGA-based search, as discussed above, has found several Fe-Co-B compounds with J_s larger than 10 kG (1 Tesla) and formation energy within 0.1 eV above convex hull (which can be easily synthesized using non-equilibrium fabrication methods). Magnetic anisotropy calculations were performed for those compounds and predicted about six compounds with K_1 larger than 10 Mergs/cm^3(1.0 MJ/m^3), which are suitable for permanent-magnet applications. Fe_3CoB_2 is one among the six compounds, which forms tetragonal crystal structure and exhibit K_1 = 12.1 Mergs/cm^3 and J_s = 14.0 kG. The formation energy of the Fe_3CoB_2 compound above the convex hull is only about 23 meV. Experimental synthesis of the predicted magnetic Fe_3CoB_2 compounds was completed using non-equilibrium cluster-deposition and melt-spinning methods at the University of Nebraska.

Future Plans

We will fabricate several Fe-Co-B compounds predicted by machine-learning and adaptive genetic algorithm searches. We also continue to work both theoretically and experimentally on Fe-Co-Si and Fe-Co-C ternary compounds. Our mining strategy will be based on: (a) screening based on available physical properties from databases and (b) refined screening by first-principles calculations. Database screening will narrow down the number of candidates to a computationally manageable set, which we can treat by more accurate first-principles calculations and adaptive genetic algorithm searches.

Broader Impacts and Workforce Development

The gender imbalance in STEM disciplines is a pressing concern. The Alice in Wonderland program attempts to address this issue. A disproportionately low number of women hold senior faculty level positions. The importance of successful women scientists as role models for graduate students has been recognized for some time. The Alice in Wonderland program targets high school girls, who participate in research in physics and related areas over the summer before they make decisions about colleges. Participants work in real research labs in the Departments of Physics or Chemical Engineering. In addition, a short course is offered at the start of the program in June. This informal course (no credit), is given by the University of Texas graduate students, covers subjects from computer modeling and quantum mechanics to scanning tunneling microscopy and thin film growth.

Data Management and Open Access

Databases created in this project will be placed in the public domain on a website dedicated to this project and as such will be available for other materials discovery efforts aimed at optimizing material properties. The database is constructed and maintained by the Iowa State group. https://www.novomag.physics.iastate.edu

Advancing Along the Materials Development Continuum and Partnerships to Translation

We are currently exploring a patent for materials related to our recent work on Fe-Co-B compounds, which were predicted by machine learning techniques and synthesized by experiment. These rare-earth free compounds have magnetic properties comparable to current rare earth magnets.

Publications and References

1. Y. Jin, R. Pathak, S.Valloppilly, P. Kharel, A. Kashyap, R. Skomski, and D. J. Sellmyer, "Unusual Perpendicular Anisotropy in the Heusler Compound Co_2TiSi," J. Phys. D: Appl. Phys. **52**, 035001 (2019).
2. R. Pahari, B. Balasubramanian, R. Pathak, M.C. Nguyen, S.R. Valloppilly, R. Skomski, A. Kashyap, C.-Z. Wang, K.-M. Ho, G. C. Hadjipanayis, and D. J. Sellmyer, "Quantum-Phase Transition and Ferromagnetism in Co_{1+x} Sn," Phys. Rev. B **99**, 184438 (2019).
3. M. Sakurai and J. R. Chelikowsky, "Metastable B-doped FeNi compounds for rare-earth-free permanent magnets," Physical Review Materials **04**, 084402 (2020).
4. H. Wang, B. Balasubramanian, R. Pahari, R. Skomski, Y. Liu, A. Haq, D.J. Sellmyer, and X. Xu, "Noncollinear Spin Structure in Fe_{3+x} $Co_{3-x}Ti_2$ (x= 0,2,3) from Neutron Diffraction,"' Phys. Rev. Materials **3**, 064403 (2019).
5. B. Balasubramanian, P. Manchanda, R. Pahari, Z. Chen, Wenyong Zhang, S. R. Valloppilly, X. Li, A.Sarella, L. Yue, A. Ullah, P. Dev, D. A. Muller, R. Skomski, G. C. Hadjipanayis, and D. Sellmyer, "Chiral Magnetism and High-Temperature Skyrmions in B20-Ordered Co-Si,"' Phys. Rev. Lett. **124**, 057201 (2020).
6. B. Balasubramanian, M. Sakurai, C. Z. Wang, X. Xu, and K.-M. Ho, J. R. Chelikowsky, and D. Sellmyer, "Synergistic Computational and Experimental Discovery of Novel Magnetic Materials," Mol. Syst. Des. Eng. **5**, 1098 (2020).
7. M. Sakurai, R. Wang, T. Liao, C. Zhang, H. Sun, Y. Sun, H. Wang, X. Zhao, S. Wang, B. Balasubramanian, X. Xu, D. J. Sellmyer, V. Antropov, J. Zhang, C. Z. Wang, K. M. Ho and J. R Chelikowsky, "Discovering Rare-Earth-Free Magnetic Materials through the Development of a Database," Phys. Rev. Materials **4**, 114408 (2020), Editors' Suggestion.
8. B. Balasubramanian, T. A. George, P. Manchanda, R. Pahari, A. Ullah, R. Skomski, and D. J. Sellmyer, ``Magnetic and Electron-Transport Properties of Co_2Si Nanomagnets," Phys. Rev. Materials **5**, 024402 (2020).
9. Y. Sun, Y. X. Yao, M. C. Nguyen, C. Z. Wang, K. M. Ho and V. P. Antropov, ``Spatial decomposition of magnetic anisotropy in magnets: application for doped $Fe_{16}N_2$," Phys. Rev. B **102**, 134429 (2020).
10. T. Liao, W. Y. Xia, M. Sakurai, R. H. Wang, C. Zhang, H. J. Sun, K. M. Ho, C. Z. Wang, and J. R. Chelikowsky, "Predicting magnetic anisotropy energies using site-specific spin-orbit coupling energies and machine learning: Application to iron-cobalt nitrides", Phys. Rev. Materials **6**, 024402 (2022), Editors' Suggestion.

Phase-field Models of Coupled Electronic and Structural Processes

Lead Investigator: Long-Qing Chen, lqc3@psu.edu.
Participating Institutions: The Pennsylvania State University
Website: https://www.mri.psu.edu/computational-mesoscale-materials-science
Keywords: Phase-field, mesoscale, electronic phase transitions, pattern formation, software

Project Scope

The goal of the project is to advance the mesoscale science of quantum and functional materials. Its main objectives are to (1) develop phase-field models for predicting the mesoscale pattern formation during coupled electronic and structural processes and implement them into an open-source software, (2) understand and predict the stability and response of mesoscale structures to external stimuli, and (3) experimentally validate and refine the computational models using atom-resolution synthesis together with cutting-edge characterization methods. Our ability to predict and control mesoscale pattern formation during simultaneous structural and electronic processes is critical to realizing their electronic device applications.

Relevance to MGI

This project is filling the critical gap of computational models and software at mesoscale for quantum and functional materials between the quantum mechanical electronic structure calculations and electronic device modeling. The project represents a transformational advance in the phase-field method for modeling electronic phase transitions and will lead to a world-unique software package for quantum and functional materials involving both electronic and structural processes including metal-insulator transitions and superconducting transitions. It involves a close integration of phase-field model development (Chen), experimental growth (Engel-Herbert) and characterization (Gopalan), DFT calculations of intrinsic electronic properties (Dabo), and numerical algorithm development (Xu). The close-loop feedbacks between computation and experiment have led to the discovery of an intrinsic non-capacitive insulator-metal phase self-oscillation in VO_2 and helped understand the extrinsic capacitive phase self-oscillations in VO_2 actuated by direct voltages (see Figure 1 for illustration with experimental data from Ramanathan at Purdue and Duwel and Callahan at Draper Lab). Our phase-field simulations predicted the formation of a transient metallic monoclinic phase during the voltage-driven insulator-metal transition in polycrystalline VO_2, validated by Lindenberg's group at SLAC. In collaboration with Zhu's group at BNL, our phase-field simulations helped understand the inherent stochasticity of the insulator-metal transition in VO_2 films. The nucleation temperatures at various Ga^+ irradiation doses obtained from our phase-field simulations are being validated by the experimental measurements by Wu's group at UC Berkeley.

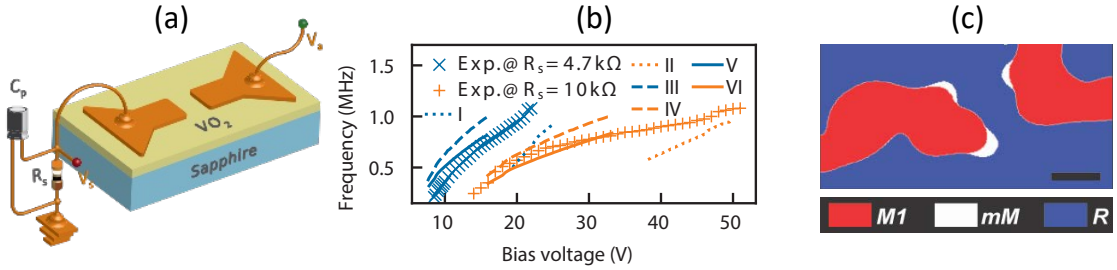


Figure 1 (a) A schematic of a VO_2 device; (b) Phase-field predicted (lines) and experimentally measured (markers) frequency scaling with the direct bias voltage of the voltage oscillation in VO_2; (c) A snapshot of the simulated phase distribution during the voltage-driven insulator-metal transition in VO_2, showing the metallic monoclinic (mM) phase stabilized between the equilibrium insulating monoclinic (M1) and metallic rutile (R) phases on a microsecond time scale. The scale bar is 20 nm.

Technical Progress

We made major progress in advancing phase-field models of coupled lattice and electronic processes and applying them to understand the mesoscale structure evolution in functional and quantum materials. For example, we developed a phase-field framework for modeling the simultaneous evolution dynamics of electronic carriers and

ferroelectric domains as well as an analytical model for the equations of motion for polar objects like vortices as a companion tool for analyzing phase-field simulation results. We established a phase-field model for the coupled thermal evolution of the electron and lattice subsystems of a crystal accompanied by carrier evolution and structural and electronic phase transitions at picosecond-to-nanosecond timescales. We constructed the software framework for the phase-field package Q-POP and created a set of multigrid-based finite-element solvers, based on which we are developing individual modules for different classes of materials. We are building *common utilities including a standard I/O interface,* a high-throughput first-principles-based software to provide thermodynamic and kinetic input data for phase-field simulations, and post-processing topological characterization routines of mesoscale structures in both real and Fourier spaces.

Future Plans

We will extend our current phase-field model of metal-insulator transitions by incorporating other possible electronic mechanisms and further develop the phase-field model of thermally coupled electron and lattice systems to incorporate both spin ordering and structural domain formation for predicting the coupled spatiotemporal evolution of electron, spin, and lattice. We will initiate the development of a phase-field model of superconducting transitions in the presence of strain in thin films. We will continue to implement scalable multigrid finite-element solvers for phase-field equations of metal-insulator transitions, electron carrier dynamics, polarization dynamics, spin dynamics, and implement them into the software package Q-POP.

Data Management and Open Access

The goal of the project is to develop phase-field models and software modules as well as tools for automating thermodynamic and kinetics inputs to the phase-field models and for characterizing the mesoscale topologies from phase-field simulations in both real and reciprocal spaces as well as material properties. All these will be made available to the General Public through a GitLab repository in the future (https://gitlab.com/doe-quantum-psu).

Advancing Along the Materials Development Continuum and Partnerships to Translation

The voltage oscillations during metal-insulator transitions have potential applications in neuromorphic computing, and our phase-field model can potentially be used to train surrogate models for device design. Although the Q-POP package will be open to the General Public, we may develop a commercial version of it in the future.

Publications and References

1. T. N. Yang and L.Q. Chen, Phase-field model of coupled carrier and structural domain dynamics. npj Computational Materials, 130 (2022)
2. Zorn, J.A., B. Wang, and L.Q. Chen, Q-POP-Thermo: A general-purpose thermodynamics solver for ferroelectric materials. Computer Physics Communications 275: 108302 (2022).
3. Y. Shi and L.Q. Chen, Intrinsic Insulator-Metal Phase Oscillations. Physical Review Applied 17, 014042 (2022)
4. T. N. Yang, C. Dai, V.A. Stoica, F. Xue, H. Wang, Y.Z. Ji, V. Gopalan, and L.Q. Chen, Computing diffraction patterns of microstructures from phase-field simulations. Acta Materialia, Under review (2022)
5. X. F. Xu, L. Zhang, Y. Shi, L.Q. Chen, and J.C. Xu, Integral Boundary Conditions in Phase-Field Models. Computers and Mathematics with applications, Under review (2022).
6. Y. Shi, A.E. Duwel, D.M. Callahan, Y.F. Sun, F.A. Hong, H. Padmanabhan, V. Gopalan, R. Engel-Herbert, S. Ramanathan, and L.Q. Chen, Dynamics of voltage-driven oscillating insulator-metal transitions. Physical Review B, 104, 064308, (2021).
7. Cheng, S.B., M.H. Lee, R. Tran, Y. Shi, X. Li, H. Navarro, C. Adda, Q.P. Meng, L.Q. Chen, R.C. Dynes, S.P. Ong, I.K. Schuller, and Y.M. Zhu, Inherent stochasticity during insulator-metal transition in VO_2. PNAS, **118**, e2105895118 (2021)
8. A. Sood, X.Z. Shen, Y. Shi, S. Kumar, S.J. Park, M. Zajac, Y.F. Sun, L.Q. Chen, S. Ramanathan, X.J. Wang, W.C. Chueh, and A.M. Lindenberg, Universal phase dynamics in VO_2 switches revealed by ultrafast operando diffraction. Science 373, 352 (2021)
9. Q. Li, V.A. Stoica, M. Pasciak, Y. Zhu, Y.K. Yuan, T.N. Yang, M.R. McCarter, S. Das, A.K. Yadav, S. Park, C. Dai, H.J. Lee, Y. Ahn, S.D. Marks, S.K. Yu, C. Kadlec, T. Sato, M.C. Hoffmann, M. Chollet, M.E. Kozina, S. Nelson, D.L. Zhu, D.A. Walko, A.M. Lindenberg, P.G. Evans, L.Q. Chen, R. Ramesh, L.W. Martin, V. Gopalan, J.W. Freeland, J. Hlinka, and H.D. Wen, Subterahertz collective dynamics of polar vortices. Nature 592, 376 (2021).
10. T. N. Yang, C. Dai, Q. Li, H.D. Wen, and L.Q. Chen, Condensation of collective polar vortex modes. Physical Review B 103, L220303 (2021)

Machine Learning Force Field Models for Dynamical Simulations of Functional Electronic Materials

Lead Investigator: Gia-Wei Chern, gchern@virginia.edu.
Participating Institutions: University of Virginia
Website: none.
Keywords: correlated electron systems, itinerant magnets, functional materials, machine learning, neural network.

Project Scope

The objective of this project is to develop a machine learning (ML) framework that enables multi-scale dynamical modeling of complex electronic systems, including strongly correlated electron materials and itinerant magnets. The central idea of this approach is to utilize the universal expressive capability of deep-learning neural networks to build efficient force-field models with the accuracy of state-of-the-art electronic structure methods and many-body techniques. Large-scale dynamical simulations enabled by the ML force-field models will shed light on the spatial-temporal dynamics of complex mesoscopic electronic textures, which in turn will lead to valuable insight on the discovery and design of novel functional electronic materials.

Relevance to MGI

MGI aims to utilize computational approaches to reduce the time and cost of developing new materials. Fundamental to this strategy is the need for efficient and accurate simulations of material propertiesNovel functionalities of complex electron materials, such as high-temperature superconductivity, colossal magnetoresistance, and resistive-switching, often originate from complex electronic textures that emerge at the mesoscale. A comprehensive dynamical modeling of these mesoscopic electronic structures is a challenging multi-scale task. On the one hand, large-scale dynamical simulations are necessary to properly describe the mesoscopic pattern formation process. On the other hand, sophisticated electronic structure methods and many-body techniques have to be employed to compute the electronic contribution to the driving forces of the mesoscale textures. In this project, a numerical framework based on modern machine learning methods is developed to overcome the difficulty of multi-scale dynamical modeling. Importantly, by comparing experiments with results from accurate large-scale dynamical simulations that are enabled by the ML force field models, one can gain crucial information on the mechanisms of novel material functionalities and how they can be controlled by external drives.

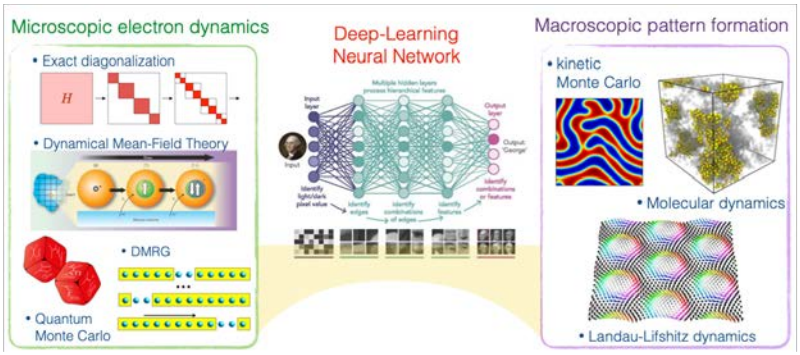


Figure. The machine-learning approach offers a bridge connecting microscopic electron dynamics and macroscopic pattern formation process, both of which are required for a comprehensive multi-scale dynamical modeling of functional electron materials.

Technical Progress

The ML force-field approach has been demonstrated on selected representative correlated electron models to enable large-scale dynamical simulations. In one case study, a multi-layer fully connected neural network is built to accurately predict the exchange forces in the well-studied double-exchange model [1], which plays a crucial role in the physics of colossal magnetoresistance observed in several manganites. By incorporating the ML model into

Landau-Lifshitz-Gilbert dynamics method, our large-scale simulations have uncovered an intriguing correlation-induced freezing behavior as doped holes are segregated from half-filled insulating background during equilibration. In another pilot study, a neural-network model is developed to enable large-scale kinetic Monte Carlo simulations of the Falicov-Kimball model [2], which is another canonical correlated electron system that exhibits complex phase-separation phenomena. We discover an unusual phase-ordering scenario where domain coarsening occurs simultaneously at two different scales: the growth of checkerboard clusters at smaller length scales and the expansion of super-clusters, which are aggregates of the checkerboard patterns of the same sign, at a larger scale. Glassy behaviors similar to those reported in these two works could be generic for other correlated electron systems.

Future Plans

The DE and FK models in our pilot studies discussed above are relatively simple in the sense that the electron subsystem can be solved by exact diagonalization, and it is relatively easier to obtain large number of training datasets with good accuracy. Armed with the tools and experiences obtained from these pilot projects, our goal in this proposal is to develop ML force-field models for electronic systems with nontrivial electron-electron interactions, such as the Hubbard repulsion. For such systems, the ML model is to be trained by datasets obtained using more sophisticated many- body techniques such as the Gutzwiller, DMFT methods or QMC simulations. Specifically, we aim to apply the ML framework to enable large-scale dynamical simulations of complex mixed-phase states in Hubbard or t-J type models. Another goal of the project in the next stage is to generalize the ML force field model for non-conservative forces that originate from highly nonequilibrium electrons [3]. In particular, such ML framework will enable efficient and accurate modeling of the spin transfer torques, which are a crucial component of the emerging field of spintronics.

Broader Impacts and Workforce Development

Our project has provided multi-disciplinary training to graduate and undergraduate students that address the modern challenges of a career in science. Four graduate students and two undergraduates were fully or partially supported by this project. Experiences gained from working on ML force field models have proven valuable in students' pursuit of career both in academia or industry. Results from this project will be of great technological implications for the development of next-generation electronic and multi-functional devices.

Data Management and Open Access

The C/Python codes, neural-network models, and sample training datasets developed in this project are publicly available at the GitHub repository: https://github.com/cherngroupUVA. An open-source package of general-purpose descriptor for condensed-matter lattice models is currently being developed.

Advancing Along the Materials Development Continuum and Partnerships to Translation

This project aims at the development and design stages of the materials development continuum. Without the aid of ML methods, comprehensive modeling of functional electronic materials is constrained by small-scale studies and empirical models. The ML force-field framework developed in this project will make possible computational software for large-scale device simulations based on novel functional materials. In particular, the ML-modeling of spin transfer torques has great potential to be a commercially relevant technique in the emerging field of spintronics.

Publications and References

1. Puhan Zhang and Gia-Wei Chern, Arrested Phase Separation in Double-Exchange Models: Large-Scale Simulation Enabled by Machine Learning, Phys. Rev. Lett. **127**, 146401 (2021).
2. Sheng Zhang, Puhan Zhang, and Gia-Wei Chern, Anomalous phase separation in a correlated electron system: Machine-learning-enabled large-scale kinetic Monte Carlo simulations, Proc. Natl. Acad. Sci. U.S.A. **119,** e2119957119 (2022).
3. Puhan Zhang and Gia-Wei Chern, Machine learning nonequilibrium electron forces for adiabatic spin dynamics, arXiv:2112.12124 (2021).

4. Yi-Hsuan Liu, Sheng Zhang, Puhan Zhang, Ting-Kuo Lee, and Gia-Wei Chern, Machine learning predictions for local electronic properties of disordered correlated electron systems, arXiv:2204.05967 (2022).
5. Sheng Zhang and Gia-Wei Chern, Nonequilibrium dynamics of gating-induced resistance transition in charge density wave insulators, arXiv:2201.02194 (2022).
6. Puhan Zhang, Sheng Zhang, and Gia-Wei Chern, Descriptors for machine learning model of generalized force field in condensed matter systems, arXiv:2201.00798 (2022).
7. Gia-Wei Chern, Spatio-temporal dynamics of voltage-induced resistance transition in the double-exchange model, arXiv:2105.11076 (2021).
8. Puhan Zhang, Preetha Saha, and Gia-Wei Chern, Machine learning dynamics of phase separation in correlated electron magnets, arXiv:2006.04205 (2020).
9. Lingnan Shen and Gia-Wei Chern, Cell dynamics simulations of coupled charge and magnetic phase transformation in correlated oxides, Phys. Rev. E **103**, 032134 (2021).

Understanding correlated quantum materials: the insight from *ab initio* dynamical mean field theory

Lead Investigator: Sangkook Choi[1], Patrick Semon[1], Andrey Kutepov[1], Yongxin Yao[3], Corey Melnick[1], Byungkyun Kang[4], Cedomir Petrovic[1], Christopher Homes[1], Gabriel Kotliar[1,2]
[1]Brookhaven National Lab, [2]Rutgers University, [3]Ames Lab, [4]University of Nevada Las Vegas

Quantum information science is a surging frontier of physical science. By creating quantum states and utilizing them as quantum bits, it promises vastly improved performance in computing, sensing, communication, and cryptography.

Quantum Materials are a class of materials that have the potential to be the "semiconductors" of quantum information science. One of the biggest challenges in this field is to understand and predict quantum materials properties, especially when the material is correlated and its quantum nature is many-body. Correlated quantum materials preclude simple explanations and computationally simple methods based on Landau's Fermi liquid theory, such as density functional theory and Hartree-Fock.

To this end, Comscope has been focused on developing and validating methodologies (theory, algorithm, and codes) based on *first principles* approaches and dynamical mean field theory (ab initio DMFT). In this presentation, we will introduce two *ab initio* DMFT flavors in Comsuite, an *ab initio* package targeting correlated quantum materials. We will also introduce our recent full GW+EDMFT implementation. Lastly, we will show our recent validation and prediction results on various classes of correlated quantum materials.

Machine-Assisted Quantum Magnetism

Lead Investigator: Sugata Chowdhury, sugata.chowdhury@howard.edu
Participating Institutions: Howard University, Northeastern University, SLAC, Stanford University
Website: https://physics.howard.edu/faculty-websites/sugata-chowdhury/
Keywords: Machine learning, quantum magnetism, quantum spin liquid, skyrmions, X-ray scattering.

Project Scope: This project combines the scientific domains of quantum magnetism, high-performance computing, and multimodal experimental spectroscopies to achieve real-time, machine-assisted control and analysis of x-ray scattering studies at LCLS. The goal is to significantly reduce the time to discover exotic magnetic quantum states in materials. Specific areas of interest include: (1) skyrmions in Fe/Gd multilayers, CrI_3 and other materials; (2) the pyroxene family of 1D Mott insulators and 1D-cuprate $CuGeO_3$; and (3) quantum spin liquid states in $RuCl_3$/graphene heterostructures. The project will give insight into areas of spontaneous fluctuations, quantum criticality, and related issues.

Relevance to MGI: Since materials with complex magnetic states are the core driver for further theoretical and experimental work in this project, a vigorous materials discovery effort to identify materials with interesting magnetic states is being pursued. Such materials can also be used to develop and test viable machine-learning models and the accuracy of first-principles schemes in modeling complex magnetic states for coupling with LCLS experiments. Our recent results in this direction include: (1) In a theory-experiment study, we showed that the metallic Fe_3Ga_4 displays a complex magnetic phase diagram that supports an ordered intermediate antiferromagnetic helical spin state at room temperature, which lies between two ferromagnetic phases.[1] (2) Using first-principles calculations, we showed that the ground state of SmB_6 involves many competing magnetic phases that lie very close in energy (Fig. 1). The topological Kondo state in SmB_6 is found to be robust regardless of its magnetic configuration.[2] (3) First-principles computations of $MnBi_2Se_4$ and related materials were shown to reveal that topological and electronic properties of this family depend on the direction of the magnetic spins and the chemistry of the materials.[3]

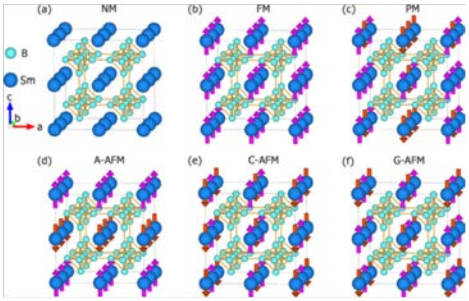


Fig 1. Rich landscape of magnetic configurations of SmB_6. (a) Nonmagnetic (NM), (b) ferromagnetic (FM), (c) special quasi-random paramagnetic structure (SQS-PM), (d) *A*-type antiferromagnetic (*A*-AFM), (e) *C*-type antiferromagnetic (*C*-AFM), and (f) *G*-type antiferromagnetic (*G*-AFM) configurations. [2]

Technical Progress: Our team has made substantial progress in advancing the five key pillars on which the project stands: (1) First-principles modeling to identify suitable candidate materials with interesting magnetic ground states and develop associated effective model Hamiltonians; (2) Higher-level modeling for accurate treatment of electron correlation effects, including variations in the strength of correlations and other model parameters; (3) Machine-learning modeling to identify the correct model Hamiltonian to capture magnetism; (4) Experimental X-ray scattering work on the most promising candidate materials; (5) Develop software tools to build a platform for real-time control of experiments guided by theory. Specific progress made is as follows: (1) A number of candidate materials that could support skyrmions, such as $NiPS_3$, EuAsAg, $EuAl_4$, and Gd_2PSi_3 have been identified, and some parallel experimental work has been carried out. In EuAgAs, magnetic susceptibility shows the signature of in-plane antiferromagnetism with possible magnetic canting below 12K and exhibits anomalous Hall resistivity below the magnetic transition. The Dzyaloshinskii-Moriya (DM) interaction in EuAgAs is found to be substantial, suggesting that the material could support skyrmions. We are also looking at MnNiGa, a sister compound in the symmetry group of EuAgAs, which shows a bi-skyrmion spin structure. (2) A machine-learning (ML) algorithm has been developed to recognize S(q,w) spectra and predict Heisenberg model Hamiltonian parameters. The algorithm correctly predicts phase transitions in the honeycomb lattice. A new experimental method was developed at the LCLS, using two pulses of equal energy, co-linear at ~100fs separation. Diagnostics incorporated a new real-time ML program that analyzed the data during the experiment.[4] (4) A new pipeline was developed that can take

experimental images as input and provide photon maps, which is the information needed to measure fluctuations with our experimental XFEL methodology. This new tool provides feedback 100x faster than the more traditional droplet-type algorithms, among other desirable features.[5]

Future Plans: Planned activities include: (1) Experimental work at the LCLS on the most promising candidate materials identified through our theory and modeling efforts, including $NiPS_3$, Gd_2PSi_3, and EuAgAs as potential candidate materials. New machine-learning algorithms and software tools will be developed. (2) First-principles efforts will continue to identify and unfold the electronic and magnetic structures of various systems of interest. We will construct effective (magnetic) Hamiltonians informed by exact diagonalization (ED) and density-matrix renormalization group (DMRG) studies. Atomistic Spin Dynamics simulations of skyrmions will also be pursued. The possibility of quantum spin liquid states in monolayer $ASiX_3$ (A = Cr, V, Co; X = Se, Te) will be explored. ED/DMRG datasets for various values of the on-site Coulomb interaction strength (U) and other parameters to refine ML models and combine them with the LCLD data. (3) Automated workflows with LCLS data systems to extract features from images and pass data to the Bayesian Optimizer (BO) will be implemented. We will incorporate the results of simulation/theory into the workflow and enable the passing of data to the BO.

Broader Impacts and Workforce Development: The project trains undergraduate, graduate, and postdoctoral students to pursue theoretical and experimental research in quantum materials, quantum information sciences, and data-driven analysis. Due to its location at Howard University, a leading HBCU, it directly impacts the training of underserved communities for careers in the quantum workforce for the industries of tomorrow. The project is also advancing the quantum infrastructure at Howard, which will impact the ability of the university to offer quantum training to students beyond the life of the project.

Data Management and Open Access: All collected data will be published in the open literature, including supplementary documentation giving details of the relevant parameters used and other archived materials. A project webpage will be maintained throughout the life of the project.

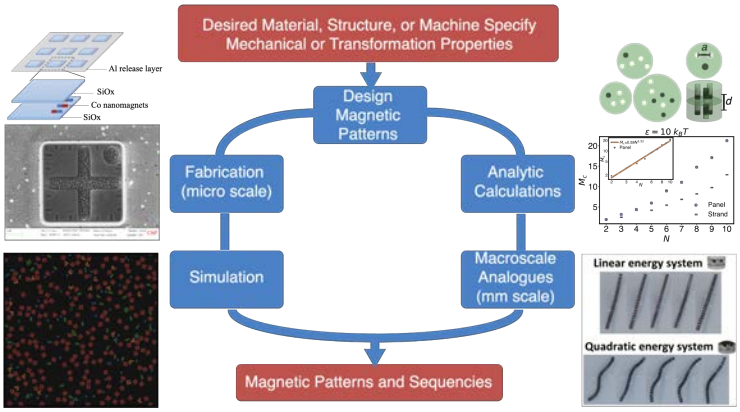
Advancing Along the Materials Development Continuum and Partnerships to Translation: The success of this research proposal will not only significantly accelerate the material discovery and exotic quantum states but also deepen our understanding of the interplay of nontrivial topology magnetism and electron correlations. Our team, with experts from material prediction and synthesis to measurements, is contributing to the subject of quantum magnetism and integrating it with educational opportunities for the students at various levels. All graduate and postdoctoral students in the project take part in bi-weekly discussions and planning meetings. The lead PI (Chowdhury) has been actively involved in designing new courses on quantum materials at Howard University in collaboration with IBM. As the COVID-19 restrictions are easing, the PI and Co-PIs plan to organize Summer Workshops at Howard University for middle and high school students to raise awareness of quantum and related technologies.

Publications and References:

1. Brandon Wilfong, Adrian Fedorko, Danil R. Baigutlin, Olga N. Miroshkina, Xiuquan Zhou, Gregory M. Stephen, Adam L. Friedman, Vaibhav Sharma, Omar Bishop, Radhika Baruah, Steven P. Bennett, Duck Young Chung, Mercouri G. Kanatzidis, Vasiliy D. Buchelnikov, Vladimir V. Sokolovskiy, Bernardo Barbiellini Arun Bansil, Don Heiman, Michelle E. Jamer, "Helical spin ordering in room-temperature metallic antiferromagnet Fe_3Ga_4," Journal of Alloys and Compounds 917, 165532 (2022).
2. Ruiqi Zhang, Bahadur Singh, Christopher Lane, Jamin Kidd, Yubo Zhang, Bernardo Barbiellini, Robert S. Markiewicz, Arun Bansil, and Jianwei Sun, "Critical role of magnetic moments in heavy-fermion materials: Revisiting SmB6," Physical Review B. 105, 195134 (2022).
3. Sugata Chowdhury, Kevin F. Garrity, and Francesca Tavazza, "Topological Phases of MnA_2X_4 (A=Bi, Sb; X = Se, Te) under Magnetic Field," arxiv 2202.07061, (2022).
4. T. A. Assefa, M. H. Seaberg, A. H. Reid, L. Shen, V. Esposito, G. L. Dakovski, W. Schlotter, B. Holladay, R. Streubel, S. A. Montoya, P. Hart, K. Nakahara, S. Moeller, S. D. Kevan, P. Fischer, E. E. Fullerton, W. Colocho, A. Lutman, F.-J. Decker, S. K. Sinha, S. Roy, E. Blackburn, J. J. Turner, "The Fluctuation-Dissipation Measurement Instrument at the Linac Coherent Light Source," Review of Scientific Instruments (2022, in press)
5. https://github.com/src47/CNN_XPFS.

Digital Handshake Materials, Structures, and Machines

Lead Investigator: Itai Cohen
Participating Institutions: Cornell University, Harvard University
Website: None
Keywords: Magnetic handshake materials, magnetic lithography, end to end optimization,
Project Scope



Our goal is to create a new class of smart materials and assemblies by combining modern magnetic information storage with ultrathin flexible materials. We encode panels with 'magnetic handshakes', microscopic patterns of magnetic dipoles binding with specific interactions. Using an integrated system of design, macroscale models, advanced simulations, and experiment, we master the programmed self-assembly of these panels. Ultimately, after microfabrication, the magnetic colloids, strands, or nets will release from the substrate into solution, and bend, move, and assemble according to their designed interactions to form new materials, structures and machines.

Relevance to MGI Our vision requires an integrated cycle between macroscopic experiment, microscopic experiment, theory and simulation, using a virtuous cycle that integrates design, analytic calculations, testing in macro-scale analogues, and simulation to enable nanoscale fabrication and self-assembly. Our simulation methodology uses recent advances in machine learning enabling end-end optimization of entire molecular dynamics simulation, allowing optimization over geometries, interactions and external fields of magnetic handshake materials for different materials properties and responses. Results iterate with experimental designs. Our cycle starts with the task of defining a material, structure, or machine with specified mechanical and/or transformation properties. The assembly is broken down into elementary constituents and magnetic handshakes are assigned to each element based on analytic calculations specifying the necessary strength and specificity of the interactions given the desired assembly path. Millimeter sized macroscale analogue assemblies are quickly fabricated to conduct proof of principle devices. Simulations are used to determine the assembly pathways. Finally, nanofabrication and microscopic characterization are used to implement and evaluate the design.

Technical Progress Our technical progress has significantly advanced magnetic handshake materials, through synergistic experiments, computation and theory. Advances include:

Design of low cross talk panels and experimental implementation (paper under review) Our theoretical framework uses an information theory driven approach to design magnetic patterns with high information capacity. This led to a simple design rule for generating magnetic patterns and designed patterns to self-assemble materials structures, which we validated with molecular dynamics simulations and macro-scale experiments. The scale invariance of the magnetic handshake materials platform means that these designs translate well towards the micro-scale.

Fabrication of microscale components, with implementation of low crosstalk designs We have nearly finished an implementation of these designs on the microscopic scale (manuscript in preparation). To test yield, we are developing an acoustic agitation approach to drive the magnetic particles exploring their assembling configuration space. The paper also includes a novel calculational approach to predict assembly yield at finite concentration. For this, we adopt a newly developed statistical mechanics framework to compute assembled-structure yield, and carry out molecular dynamics simulations to validate the theory. This theory is then used to identify optimal dipole strength for controlled assembly, where at high temperature, only finite squares form, and at lower temperatures, we achieve bulk growth.

Design of magnetic polymers We are in the final stages of completing a manuscript on this work. In this paper, we developed a novel experimental methodology for controlling the curvature of magnetic building blocks, by

controlling the shape, magnetic interactions and size of the monomers. We then create a theoretical model for computing the bending energy of magnetic monomers with different magnetic patterns and geometry, and develop first principle derivations to predict the persistence length based on bending energies. We carry out experiments at the macro-scale and then demonstrate that the theoretical and computational results well describe the experiments. The results provide insights on how to design polymer properties using magnetic handshake materials, allowing monomer design to program different natural curvature, persistence length, and responses from external stimuli.

Novel computational methods Our methodology employs the technology of machine learning (ML) to develop a way of doing end to end optimization of our component design. We did this by implementing a rigid body functionality in a newly developed molecular dynamics engine named JAX-MD, written in a state of the art ML framework, that allows users to take gradients throughout a whole molecular dynamics simulation. The development of rigid body functionality opens up new vistas for optimizing our components in conjunction with experiments, including optimizing over geometries, interactions and external fields of magnetic handshake materials for different materials properties and responses. The framework works for both equilibrium and nonequilibrium processes. We hope that it will give us the needed tools for designing nonequilibrium engines.

Future Plans Having completed the basic tools for designing magnetic handshake materials both experimentally and computationally, we now will use these tools to develop nonequilibrium engines, including with dynamic magnetic fields. Our recent progress on developing the microscopic platform means we will also be able to implement our designs microscopically using magnetic lithography. The tight coupling between theory, computation and experiment will let us design novel material constructs.We hope to accelerate the design and fabrication process of these microscopic self-assembly systems and build more delicate and complex microscopic structures from the bottom-up.

Broader Impacts and Workforce Development Due to Covid 19 our efforts at outreach have been somewhat hampered. Digital magnetics will be a powerful platform for exploring the science of nanoscale design, with unique features not found in DNA/RNA. It can also be fully integrated with other lithographic elements (electronics, optics, etc.) and will have broad applications in sensing, actuation, and microrobotics at the cellular scale. Our goal is to develop a broadly usable platform, which can be realized using available fab technology, for exploring the emerging scientific discipline of programmable self-assembly will open the door to developing these technologies. The macroscale analogues are being adopted into *lending kits* that will be used to explain the basic principles behind base pairing in DNA and its assembly into DNA origami structures. The kits are currently being developed and will be propagated through a library run by the Cornell MRSEC and the Ithaca Physics Bus - a roving exhibition hall that serves impoverished communities in upper Appalachia. The PIs are deeply involved in outreach based on their science. Cohen has written a textbook based on his course "Finding your voice". Brenner published his Science and Cooking textbook**,** which focuses on interplay between experiment and theory in science, and his HarvardX class has been taken by over 500k people, ramping significantly during Covid.

Data Management and Open Access The research data is being captured as text files from laboratory notes and ASCII files from instruments. Plots are being generated using MATLAB, Excel etc. The experimental data is being stored on the computers associated with this project as well as back-up servers and backup drives. Code modifications to the HOOMD Blue simulations and JAX-MD (current branch) are being shared with the community via GitHub.

Advancing Along the Materials Development Continuum and Partnerships to Translation The magnetic handshake platform has potential applications in fields ranging from adaptive metasurface optics to microrobotics for miniaturized medical instrumentation. Furthermore, the platform is general purpose. The design space is enormous, much too large to efficiently explore without iterated theory/experiment/simulation. We plan to patent and commercialize key aspects of the technology as they are developed.

Publications and References

1. CX Du, HA Zhang, T Pearson, J Ng, P McEuen, I Cohen, MP Brenner, *under review*, arXiv:2112.02614
2. Niu, R., Du, C. X., Esposito, E., Ng, J., Brenner, M. P., McEuen, P. L., & Cohen, I. (2019). Magnetic handshake materials as a scale-invariant platform for programmed self-assembly. *Proceedings of the National Academy of Sciences*, *116*(49), 24402-24407.

DMREF/Collaborative Research/GOALI/: High-Affinity Supramolecular Peptide Materials for Selective Capture and Recovery of Proteins

Lead Investigator: Honggang Cui; Email: hcui6@jhu.edu
Participating Institutions: The Johns Hopkins University, University of Chicago, Northwestern University, Bristol-Myers Squibb
Website: none
Other PIs: Matthew Tirrell, University of Chicago; Monica Olvera de la Cruz, Northwestern University; Xuankuo Xu, Bristol Meyers Squibb
Keywords: supramolecular materials; high-affinity materials; protein purification; phase separation, protein binding.

The separation and purification of therapeutic proteins from their biological resources pose a great limitation for industrial manufacturing of biologics in an efficient and cost-effective manner. Despite the high media cost and limited loading capacity, affinity chromatography remains the most widely used capture method for large-scale industrial protein purification. The rapid growth of upstream titers, due to advancements in mammalian cell culture and continuous process development, has further challenged the efficiency of downstream manufacturing. This DMREF/GOALI proposal aims to develop a peptide-based platform technology for non-chromatographic protein purification, with the ultimate goal of addressing the inherent production limit of affinity chromatography.

Project Scope

Our long-term goal is to develop a peptide-based platform technology for non-chromatographic purification of proteins and other biologics. The objective of this research is to lay out the foundation for this platform technology through the development of high-affinity supramolecular immunofibers (IF) for selective capture and recovery of monoclonal antibodies (mAbs). Three specific aims have been set, each covering a key step toward the successful development of a non-chromatographic protein purification technology: (1) selective capture, (2) phase separation, and (3) protein recovery (**Fig. 1**). We envision that upon further computational and experimental optimization of our molecular, supramolecular and materials design, the proposed supramolecular peptide materials can indeed serve as an efficient and economical alternative to traditional chromatography methods for purification mAbs and other proteins.

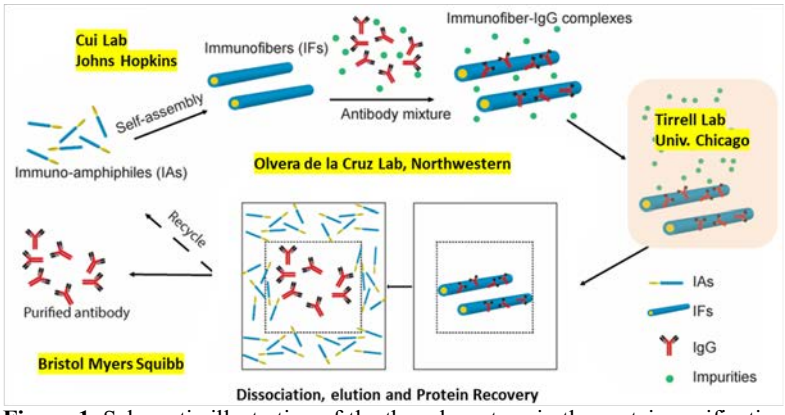


Figure 1. Schematic illustration of the three key steps in the protein purification process using self-assembling Immunofibers (IFs) as reversible affinity precipitants. Key step 1 (selective capture of proteins on IF surface): Rationally designed Peptide Amphiphile self-assemble into filamentous nanostructures with ligands present on their surface for high-affinity binding with target proteins such as IgG. Key step 2 (macroscopic phase separation): Following selective capture, binding-triggered phase separation is expected to occur, leading to coacervation/precipitation of the IFs bound with IgG from the impurities-containing supernatant. Key step 3 (protein recovery): Immunofiber-IgG complexes are then redissolved in elution buffer at lower pH, resulting in the dissociation of both IgG from the IFs and the IFs themselves. After dissociation, proteins are separated from the IAs and recovered in their purified state.

Relevance to MGI

The project aims to design, synthesize, and develop a class of self-assembling peptide materials that can specifically bind, selectively capture, and effectively separate proteins from their bio resources. To address the key fundamental challenges, an interdisciplinary research team is assembled, including computational, theoretical, and experimental expertise. The Olvera de la Cruz Lab serves as the core of the proposed project that guides our experimental studies

in the supramolecular design of peptide immunofibers for use in protein capture, separation, and recovery. Multiscale models are being developed to elucidate the thermodynamic assembly and kinetic pathways of immunofibers and provide guiding principles to the Cui lab on the amphiphile design and co-assembly strategies to optimize the ligand presentation for maximal protein capture. Theories and models are also being developed to understand the protein-binding triggered phase separation phenomena of immunofibers in solution and predict the influence of important experimental parameters validated by the Tirrell Lab on the kinetic and thermodynamic aspects of the phase separation process. Our industrial partner provides expertise in peptide ligand selection and performs experiments to assess the product quality attributes of our purified proteins.

Technical Progress

This is a new project, and the team is working to get things started. Over the past 8 months, we have made important progress in the design and synthesis of self-assembling peptide amphiphiles for high-affinity binding and separation of proteins of interests, as well as in the development of computation models and tools to simulate the binding events and the phase separation process.

Future Plans

Theoretical and multiscale models are being developed by the Olvera de la Cruz Lab, to provide important insight into the self-assembly of peptide amphiphiles into high-affinity immunofibers and the network structure of immunofibers in solution, and their impact on the protein binding and precipitation processes. The predictions from these molecular simulations and theoretical models will be used for new amphiphile chemistries and design rules for the Cui Lab, and to guide the experimental work conducted by the Tirrell Lab (**Figure 2**).

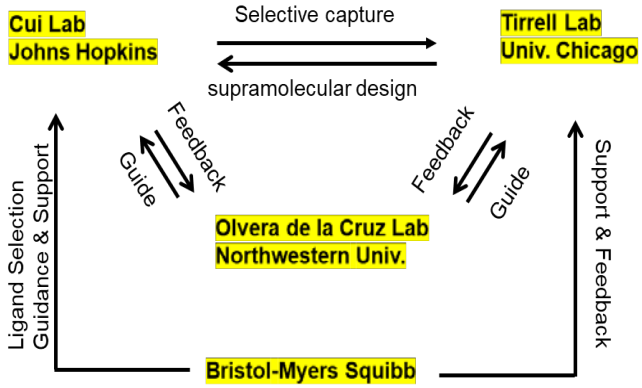


Figure 2. Overview of our research strategies to address the challenges in the development of supramolecular immunofibers for protein capture, separation, and recovery. The feedback loops are noted in the scheme to highlight the integration of experimental research with simulation and theoretical results. BMS plays an important role in helping our research team identify important peptide ligands (potentially markable) and contribute to assess the product quality and impurify levels of the purified proteins using the immunofibers to be developed under this support.

Broader Impacts and Workforce Development

Through our unique tetrahedral collaboration between JHU, UChicago, Northwestern and BMS, we plan to create a diverse, rotational undergraduate exchange program that allows students to experience highly different laboratory settings ranging from academia to industry in nature. we will also leverage our research collaborations to foster new educational opportunities for both Baltimore and Chicago Inner City students. A collaboration of this magnitude has the potential to make a remarkable impact on Inner City Students of the two major cities in the country by broadening the pipeline of underrepresented students into STEM professions. Content generated as part of this project will be adapted for use in educational outreach programs that the three PIs actively participate in.

Data Management and Open Access

We have created a OneDrive folder for data sharing among the four research labs involved in the project. Once we have more results collected, published, ready for public access, we will create an open access folder for the public.

Advancing Along the Materials Development Continuum and Partnerships to Translation

Through collaboration among the three research groups and an industrial partner, we can combine our respective expertise to understand the capture mechanisms and phase separation behavior between proteins and engineered filamentous nanomaterials from both a computational and experimental perspective. Such a collaboration would foster a flourishing learning environment that can streamline our research efforts directly into an industrial setting and have a profound real-world impact

Publications and References

Not available at this point of time.

Data-driven discovery of synthesis pathways and distinguishing electronic phenomena of 1D van der Waals bonded solids

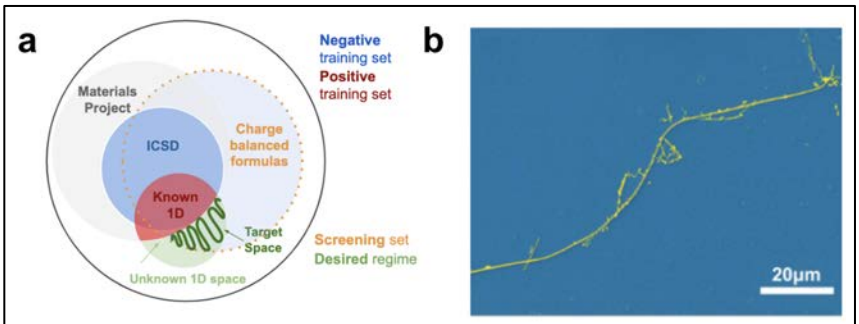
Lead Investigator: Felipe H. da Jornada, jornada@stanford.edu
Participating Institutions: Stanford University, UC Riverside
Website: none
Keywords: van der Waals, one-dimensional, machine learning, data-screening, CVT and CVD synthesis

Project Scope

One-dimensional van der Waals (vdW) bonded materials exhibit a range of scientifically interesting and technologically relevant properties involving their electrical transport, mechanical response, and stability. Our team at Stanford and UC Riverside employs a combination of materials growth, characterization, data mining, machine learning, electronic structure calculations, and database curation to identify, grow and characterize unique and useful 1D materials. We developed machine-learning methods to identify 1D materials from databases and ultimately expand available datasets to predict novel 1D materials that have not as-of-yet been characterized, synthesized, or even predicted. Some of these predictions have recently been confirmed by our team.

Relevance to MGI

The project employs a closed-loop collaboration between theory and computation (Stanford), together with synthesis and characterization (UC Riverside), to speed up the discovery of novel 1D vdW materials. Our unique combination of data mining and first-principles calculations has allowed us to predict a range of 1D materials not previously present in any database nor synthesized. We focused on electrical and vibrational properties, phase transformations of the native bulk phases, and emergent properties of the single wire phases. We demonstrated a closed-loop approach in the candidate material prediction MoI_3, which has been verified to indeed have a 1D form and potentially exhibit anomalous magnetic properties. We expect that our closed-loop collaboration will not only produce several novel quasi-1D materials but also serve as a blueprint for future materials discovery initiatives. In particular, our work has assembled valuable know-how on which computationally-accessible properties are relevant for identifying novel materials that are stable and can be experimentally synthesized and isolated.



(a) Machine learning search space: we start with the Materials Project database and seek 1D materials in the larger space of randomly generated compositions. (b) Experimental SEM image for the predicted 1D composition MoI_3.

Technical Progress

We predicted low-dimensional materials that have yet to be synthesized, with emphasis on materials that have the potential to exhibit multiple phases or charge density waves. Previous work has revealed that 1D vdW materials can exhibit experimentally feasible exfoliation energies, similar to their 2D counterparts. As the current number of such known compounds is only ~700, we seek to expand the range of such materials using machine learning methods. Using random forest and support vector machine (SVM) models, we predict 1D vdW compounds not yet synthesized in the space of binary and ternary compounds that are charge-balanced, since the majority of existing 1D vdW materials fall in this category. We train on classes of material that correspond to common CVD and CVT precursors – pnictogen, halogen, and chalcogen-containing compounds – and compositions with a transition metal. We found that a random forest model can yield precision up to 40 times better than a random-choice baseline, and a support vector machine model yields improvements up to 8 times better than random with a significant amount of overlap between the predicted positive, unsynthesized 1D wires. Various ML models trained on different subsets of the data recover similar unsynthesized 1D materials, suggesting the robustness of the model. A particular composition, MoI_3, which we predict to be 1D, is indeed confirmed to have wire-like subcomponents, and ongoing work is occurring with the UC Riverside experimental groups on this material. Our explicit density functional theory (DFT) and phonon calculations have corroborated experimental Raman and XRD measurements.

We performed additional calculations on quasi-1D and 1D materials studied in the Balandin and Bartels group, including members of the transition metal trichalcogenide family, and investigated the potential exfoliation of Weyl

semimetal Ta_2Se_8I. Balandin fabricated Ta_2Se_8I test structures for transport measurements, investigating current fluctuations and discovering a strong enhancement of the low-frequency noise at the phase transition point. Raman spectroscopy measurements coupled with the computational study of magnon and phonon dispersion in MoI_3 allowed us to shed light on the nature of this 1D material, and its phase transitions supported near room temperature.

Future Plans

We anticipate continued collaborations among the Stanford and UC Riverside groups for experimental synthesis and characterization of 1D vdW materials, aligning the in-lab measurements with data-driven predictions and electronic structure calculations. Experimental Raman spectroscopy data will be used for the validation of theoretical and computational models for phonon and magnon dispersion in 1D materials. Electrical measurement data, *e.g.* resistivity dependence on temperature, obtained at UC Riverside will be used for comparison with electronic band-structure calculations at Stanford, and for the identification of promising materials that preserve metallic conduction while scaled down to bundles of few atomic chains[1-3]. We will expand experimentally and computationally our understanding of how perturbation of 1D materials by a variety of methods (intercalation, substitutional doping, deviation from stoichiometry) affects the materials' properties and can ultimately be employed as a facile and predictive manner of their optimization. To this end, we started to collaborate with several experimental groups and companies beyond this project, offering us access to a wider set of materials and within one material, to a wider set of perturbations.

Broader Impacts and Workforce Development

Within each research group, we have mentored 1 high school and 2 undergraduate students under the framework of data analysis for materials discovery. This project saw several graduates taking their knowledge to the industry. Dr. Thomas Empante, who worked on metallic 2D nanowire growth in this project, moved to Applied Materials. Dr. Michelle Wurch accepted a position with Newport Optics. Dr. Kortney Almeida joined Applied Materials.

In the spring of 2021, a new class was offered at Stanford, "*Machine Learning for Materials Science and Chemical Engineering*," MSE 166, as a result of this project. We also offered a seminar in January on how Bayesian techniques can assess the likely validity of the information that comes through news outlets and social media. The previous lead PI Reed was a guest interviewee on the "*Future of Everything*" podcast hosted by Russ Altman – distributed by Sirius XM Satellite radio and Youtube (https://youtu.be/xWAP__HDllw). In Winter 2022, Bartels offered the graduate class "*Current Issues in Semiconductor Processing*" featuring 1D metallization schemes.

Data Management and Open Access

ML-predicted 1D compositions have been disseminated to experimental collaborators and will become available upon the submission of papers. Previous screening of 1D vdW materials has been made available on the groups' website and screening algorithms incorporated into the Pymatgen Python library for close integration with the existing Materials Project database.

Advancing Along the Materials Development Continuum and Partnerships to Translation

Our project has not only identified candidate 1D wires, but also learned practical issues to better translate relatively inexpensive computational searches into a selection of good candidate materials for experimental searches. For instance, we identified ways for our random forest and SVM ML approaches to yield a diverse set of candidates without being overly fitted or dependent on sparse DFT data. Our insights into how to estimate exfoliation energy and structural stability will also likely be helpful for future materials discovery initiatives. Finally, our groups are actively collaborating with industry partners within the Semiconductor Research Corporation.

Publications and References

1. F. Kargar, A. Krayev, M. Wurch, Y. Ghafouri, T. Debnath, D. Wickramaratne, T. T. Salguero, R. K. Lake, L. Bartels, and A. A. Balandin, "*Metallic vs. semiconducting properties of quasione-dimensional tantalum selenide van der Waals nanoribbons*", Nanoscale, 14, 6133-6143 (2022).
2. Z. Barani, F. Kargar, Y. Ghafouri, S. Ghosh, K. Godziszewski, S. Baraghani, Y. Yashchyshyn, G. Cywiński, S. Rumyantsev, T. T. Salguero, and A. A. Balandin, "*Electrically insulating flexible films with quasi-1D van der Waals fillers as efficient electromagnetic shields in the GHz and sub-THz frequency bands*", Advanced Materials, 33, 2007286, (2021).
3. A. A. Balandin, F. Kargar, T. T. Salguero, and R. Lake, "*One-dimensional van der Waals quantum materials"*, Mater. Today, 55, 74-91 (2022).
4. R. Freitas and E. J. Reed, "*Uncovering the effects of interface-induced ordering of liquid on crystal growth using machine learning*", Nat. Comm. 11, 3260 (2020).
5. H. W. Chung, R. Freitas, G. Cheon, and E. J. Reed, "*Data-centric framework for crystal structure identification in atomistic simulations using machine learning*", Phys. Rev. Mater. 6, 043801 (2022).

Artificial Intelligence Detection of Broken Symmetries at Phase Transitions

Lead Investigator: Adrian Del Maestro, Adrian.DelMaestro@utk.edu
Participating Institutions: University of Tennessee, Knoxville
Website: https://github.com/dagrawa2/ssb_detection_ising
Keywords: machine learning, symmetries, phase transitions, autoencoder

Project Scope

We introduce the group-equivariant autoencoder, a deep neural network method that locates phase boundaries in the Ising universality class by determining which symmetries of the Hamiltonian have spontaneously broken at different temperatures. This procedure produces a dramatic reduction in the number of free parameters such that the autoencoder size is independent of the system size. By examining the group representation by which the learned order parameter transforms, we can extract information about the associated spontaneous symmetry breaking. The code was released as open source and is generalizable to learning phase diagrams in a host of materials.

Relevance to MGI

With the growth of complexity of even minimal models describing materials with designer properties (e.g. magnetic, electronic, topological) it is essential to devise new machine learning tools that integrate theory, modelling, and simulation. The group equivariant autoencoder incorporates all high energy microscopic model symmetries directly into its machine learning architecture. It hybridizes a group theoretic mathematical analysis to harness the power of large-scale simulation data and is systematically validated by a comparison with existing state-of-the-art methods of learning phase transitions in two canonical models of magnetic materials. The main advantages of our approach include reduced model size, faster training, and the ability to learn and generalize from limited data sets using a symmetry and physics informed methodology. When generalized to a broader class of systems, it has the potential to accelerate the design and understanding of materials by allowing for the rapid investigation and prediction of phase diagrams that are essential to harnessing material functionalities.

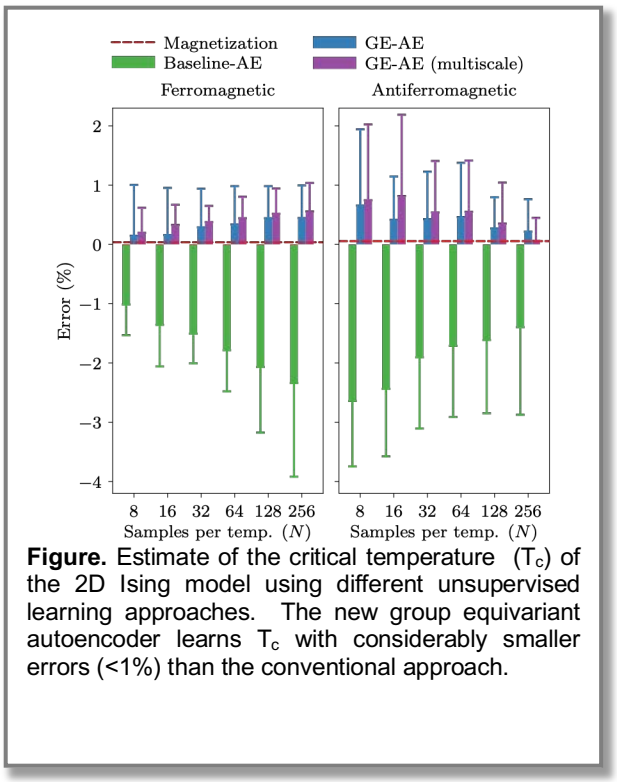


Figure. Estimate of the critical temperature (T_c) of the 2D Ising model using different unsupervised learning approaches. The new group equivariant autoencoder learns T_c with considerably smaller errors (<1%) than the conventional approach.

Technical Progress

The main contribution of this project is a new deep neural network method for identifying phase transitions, the group-equivariant autoencoder (GE-autoencoder). In contrast to previous unsupervised learning methods for phase transitions, the GE-autoencoder is designed to identify which symmetries of a given system are broken at each point of its phase diagram. The GE-autoencoder not only locates phase transitions, but gives insight into the underlying mechanism via the associated spontaneous symmetry breaking. Our method only assumes that we have knowledge of the symmetry group of the system Hamiltonian and that we have selected a latent dimensionality for the GE-autoencoder. We demonstrated its efficacy for the 2D classical ferromagnetic and antiferromagnetic Ising models,

finding that the GE-autoencoder (1) accurately determines which symmetries are broken at each temperature, and (2) estimates the critical temperature with greater accuracy and time-efficiency than the symmetry-agnostic autoencoders currently employed in unsupervised learning studies of phase transitions. Most importantly, the GE-autoencoder can be employed in a "multiscale" configuration which is independent of the system size and thus can be utilized to make predictions directly in the thermodynamic limit applicable to real systems. Having access to details on broken symmetries provides information on how to couple to the order parameter via a conjugate field – a requirement for probing associated phase transitions in experiments. Finally, the GE-autoencoder demonstrates robustness, both in terms of its ability to accurately learn magnetization without overfitting, and its greater sensitivity to the presence of an external symmetry-breaking magnetic field.

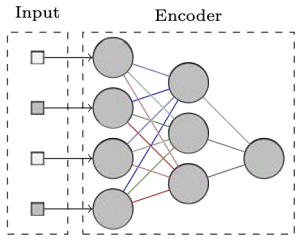


Figure. The group equivariant encoder is constrained to be invariant to the never-broken symmetries of the high-energy microscopic model describing the material system. Network edges of the same color and shade are constrained to have the same weight.

Future Plans

Recent progress on transfer learning from small to large lattice sizes via block decimation inspired by the renormalization group indicates that the scale-independent nature of the GE-autoencoder makes it well suited for this task. We intend to further improve our training regimen to reduce the need for computationally intensive large lattice sizes, providing a speedup over traditional Monte Carlo methods. The enhanced robustness of the GE-autoencoder also hints that it may operate favorably in learning phase diagrams in quantum systems with a sign problem, allowing for the exploration of more extreme regimes of temperature and doping in technologically important systems such as the single band Hubbard model. Knowledge of the spontaneously broken symmetry associated to a phase transition has value beyond phase detection and identifying relevant symmetries could help elucidate the mechanism driving subtle phase transitions, offering a means to control them in practice to realize desired material functionalities.

Broader Impacts and Workforce Development

The students and post-docs involved in this project and its future extensions work in a multidisciplinary research environment, receiving hands-on training in theoretical condensed matter physics and its interface with artificial intelligence, machine learning, data science, and high performance computing. We are also actively engaging with users at the national faculties to showcase the project's technology and train users in their usage.

Data Management and Open Access

All codes developed by this project are publicly available with complete documentation. All data sets and training details are included on public repositories (e.g. https://github.com/dagrawa2/ssb_detection_ising) ensuring reproducibility, rapid follow-ups, and broad dissemination to our communities of practice.

Advancing Along the Materials Development Continuum and Partnerships to Translation

The ability to detect and build a detailed understanding of phase transitions via unsupervised learning is crucially important to the goal of accelerating materials discovery and development from the growing body of experimental and simulation data produced by MGI. Techniques based on incorporating physical knowledge to reduce model complexity could be directly incorporated into commercial artificial intelligence approaches.

Publications and References

1. D. Agrawal, A. Del Maestro, S. Johnston, and J. Ostrowski, *A group-equivariant autoencoder for identifying spontaneously broken symmetries in the Ising model*, arXiv: 2202.06319 (under review at Phys. Rev. E), (2022).

Design of Superionic Conductors by Tuning Lattice Dynamics

Lead Investigator: Olivier Delaire (Duke University), olivier.delaire@duke.edu
Co-Principal Investigators: Veronica Augustyn (North Carolina State University) vaugust@ncsu.edu; Boris Kozinsky (Harvard) bkoz@seas.harvard.edu; Alexandra Zevalkink (Michigan State University) alexzev@msu.edu
Website: none.
Keywords: ionic diffusion, lattice dynamics, thermodynamics, disorder, machine-learning.

Project Scope

Superionic conductors (SICs) exhibit complex atomic dynamics, bridging extended time and length scales [1-4]. Their thermodynamic and transport properties remain poorly understood, as a result of insufficient characterization of their atomic dynamics and because of the challenge to computationally access all relevant time/length scales. This project aims to achieve atomistic understanding of the interaction between mobile ion and framework dynamics, effects of entropy and frustration, as well as spatial and temporal correlations, to develop a predictive description for superionic behavior and accelerate the discovery of new SICs. Design hypotheses and model superionic compounds are shown in Figure 1.

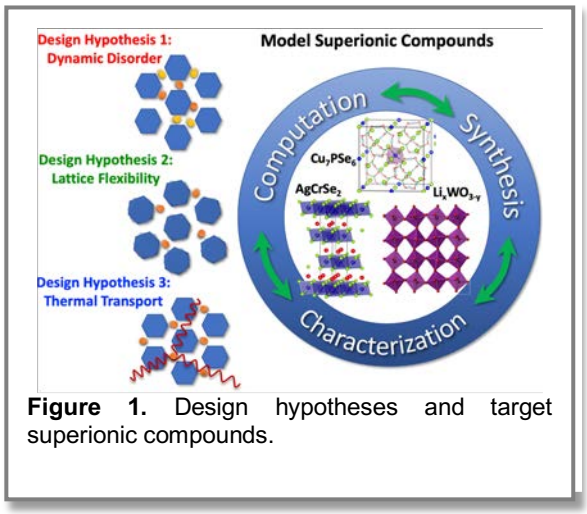


Figure 1. Design hypotheses and target superionic compounds.

Relevance to MGI

The proposed project, in the spirit of the Materials Genome Initiative, will establish a community data infrastructure and open-source tool ecosystem for automated dynamics simulations and analysis. The primary goal of this project is the development of an integrated computational and experimental approach capable of predicting diffusive and thermal properties of SICs, and to ultimately discover new materials. Our interdisciplinary team explores three Design Hypotheses summarized in Figure 1, featuring iterative synergy between simulations, experiments, and identification of key physical phenomena. Many parameters have been correlated with superionicity: unit cell volume, size of the diffusion bottleneck, anion chemistry, topology of the anion sublattice, as well as microstructure. Yet, previously proposed descriptors mostly ignore the collective sublattice dynamics and flexibility of the framework, despite emerging evidence for their importance. The details of this connection remain largely unknown because of the scarcity of atomically resolved measurements to benchmark microscopic theories of superionicity, and because of computational challenges. The project addresses this gap through a comprehensive approach integrating novel computational infrastructure with synthesis and characterization at state-of-the-art facilities, which, combined, allow us to capture all the relevant atomic-scale processes underlying thermodynamic and transport phenomena. This will reveal and chart a path toward controlling the interplay of lattice flexibility and ionic correlations in fast ion diffusion. The new understanding will be encapsulated in chemical and structural rules and descriptors, applicable to the data-driven understanding and design of new superionic conductors.

Technical Progress

To investigate the Design Hypotheses in Fig. 1, we use a multi-pronged approach fusing together state-of-the-art theoretical and experimental tools to build the fundamental knowledge base currently missing in SICs research. Our early work focused on argyrodite compounds, which allow for partial substitution of mobile cations (or vacancies) and in which the anionic framework can be tuned (chalcogen or halogen substitution) to control polarizability and stiffness. These highly tunable compounds are promising for both solid-state electrolytes and thermoelectric applications. The simulation tools developed in the project, including machine-learning augmented simulations, are being integrated into the analysis of neutron/x-ray spectroscopy data to isolate the relevant correlation functions and provide new insights into superionic conduction mechanisms. Neutron and x-ray scattering data have so far been performed on Li argyrodites (Li_6PS_5Cl) to benchmark predictions of atomic structure, dynamics and migration pathway. Early results suggest the importance of low-frequency lattice dynamics

of the host crystal. Synthesis and characterization of Li_xWO_{3-y} is ongoing.

Future Plans

Atomistic modeling will be used to probe ion and heat transport, free-energy landscapes, as well as lattice phonons and stability. Our theoretical modeling leverages state-of-the-art computer simulations of atomic dynamics with ab-initio and machine-learning accelerated MD simulations. Ab-initio and ML-accelerated dynamics simulations will be directly compared with measurements. Computed free energy landscapes will inform models of transport mechanisms, including ion-hopping pathways, and activation energies. ML-accelerated MD simulations will allow us to compute macroscopic diffusivities, connecting to experimental diffusion measurements, and to elucidate the mechanisms underlying the high conductivities of argyrodites and other SICs, highlighting effects of lattice phonons and ion-ion correlations. Our future plans include investigations of halogen substitutions in Li-argyrodites, as well as Ag-based argyrodites. In addition, we plan to investigate frustration and cooperative diffusion processes in 2D layered $MCrX_2$ compounds (M = Li, Na, Cu, Ag and X = S, Se). Their relatively simple structure and half-filled honeycomb mobile lattice make them ideal to study dynamic ionic correlations and superionic transition temperatures. They are of strong interest as both solid-state electrolytes and thermoelectrics. Simulations and characterization of Li_xWO_{3-y} compounds are also planned.

Broader Impacts and Workforce Development

The scientific outcomes of this project will be the basis for short online educational courses, helping to engage next generations of researchers in this multidisciplinary topic. The online lectures will be complemented with hands-on simulation exercises. To capitalize on the team's diverse expertise, we will hold annual 2-day hybrid workshops (in-person and virtual) organized by the PIs and their students. The workshops will be designed to introduce non-specialists to the expertise of each group. Co-PI Augustyn is the founder and faculty advisor of the SciBridge project at NC State, which develop experiment kits on renewable energy for partner universities in east Africa to further global engineering education. Co-PI Zevalkink is involved in K-12 outreach in mid-Michigan, encouraging K-12 girls to choose STEM by working with the Women in Engineering program at MSU. She has co-developed hands-on activities involving identification of single crystals from nature and the lab, deployed annually in engineering camps to reach ~200 students/year.

Data Management and Open Access

To automate calculations, collect and organize databases of results, we will use AiiDA, a workflow automation and provenance-centered data management platform designed by Co-PI Kozinsky. We will create tools for developing accelerated ML models, performing dynamics simulation and data analysis, emphasizing the reproducibility of the computed data as well as learned models and descriptors. Specifically, we will integrate the AIMD reference calculations, ML model training protocols and ML-MD trajectory generation and data analysis into automatable and reusable workflows in AiiDA and release these tools in open-source repositories. Further, we plan to develop a public online "inter-loan library" for physical samples resulting from NSF-funded research, allowing more ready collaboration between research groups.

Advancing Along the Materials Development Continuum and Partnerships to Translation

As part of this project, we will integrate ionic transport codes into the AiiDA platform, now used by tens of research groups worldwide to automate workflows and manage computed data. Co-PI Kozinsky is a part-time principal scientist at BOSCH, coordinating the transfer of computational methods to industrial research efforts Plans are in place for BOSCH Research to adopt and apply computational methods that will be developed in this project.

References

1. J.B. Boyce, B.A. Huberman, *Superionic Conductors - Transitions, Structures, Dynamics*, Phys. Rep.-Rev. Sec. Phys. Lett., **51,** 189-265 (1979).
2. M.B. Salamon, *Physics of Superionic Conductors*, Springer-Verlag, Berlin; New York (1979).
3. N. Kamaya, K. Homma, Y. Yamakawa, M. Hirayama, R. Kanno, M. Yonemura, T. Kamiyama, Y. Kato, S. Hama, K. Kawamoto, A. Mitsui, *A lithium superionic conductor*, Nat. Mater., **10**, 682-686 (2011).
4. J.L. Niedziela, D. Bansal, A.F. May, J.X. Ding, T. Lanigan-Atkins, G. Ehlers, D.L. Abernathy, A. Said, O. Delaire, *Selective breakdown of phonon quasiparticles across superionic transition in CuCrSe2*, Nat. Phys., **15**, 73 (2019)

DMREF: Collaborative Research: Computationally Driven Design of Tissue-Inspired Multifunctional Materials

Lead Investigator: Berkin Dortdivanlioglu, University of Texas at Austin, berkin@utexas.edu, Andy Sarles, University of Tennessee, Knoxville, ssarles@utk.edu, Robert Hickey, Pennsylvania State University, rjh64@psu.edu
Co-PI: Manish Kumar, University of Texas at Austin, manish.kumar@utexas.edu
Website: none.
Keywords: artificial tissues, biomimetic membranes, block copolymers, interface mechanics, 3D printing

Project Scope

Our objectives involve the computation-guided design, synthesis, characterization, and assembly of tissue-like materials made of cell-like compartmentalization. In an integrated computational and experimental framework, we focus on the integration across a multitude of length scales of self-assembling block copolymer microgels (BCP), biomimetic membranes (BMs), and stimuli-responsive biomolecules as bio-inspired materials. Physics-based simulation tools based on multi-field continuum models enriched with interface mechanics will provide insights into the compartment and membrane design at all scales and help build a rapid surrogate machine learning (ML) model for optimal microstructure exploration. We hypothesize that by coupling independent compartments through mechanical (i.e., BCP cytoskeletons) and transport (i.e., selective BMs) mechanisms, hierarchical tissues that match the performance of biological tissues can be created.

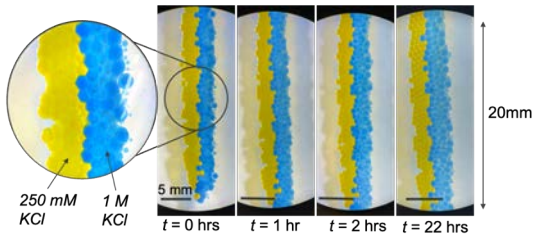


Figure. A printed tissue assembly. The yellow side contains volumes with 250mM KCl, whereas the blue droplets contain 1M KCl. The KCl concentration gradient results in water transport, with the goal of folding actuation response due to the differential shrinking (yellow) and swelling (blue) of the droplets.

Relevance to MGI

The project creates a synergy of expertise in computational modeling, polymer synthesis and physics, biomembranes, protein channels, and 3D printing. This synergy is designed to provide a quantitative knowledge base on cell-like compartmentalization, controlling transport via selective BMs, modulating stimuli-responsiveness, and exhibiting tunable mechanical properties with precise temporal and spatial resolution. To overcome scale-up challenges, we devised an iterative feedback loop to link simulation inputs and outputs to experimental design and characterization for validation of our high-fidelity models at various length scales. Trained on high-fidelity computations and experimental data, surrogate machine-learning models will explore rapid predictive designs that inform G-code instructions for 3D tissue-like printing and assembly. Given the vast design space across scales, the knowledge base created through our computationally driven experimental program is critical to elucidating key structural, mechanical, and transport-driven parameters needed to make rational choices to engineer materials that mimic structure-property relations of functional living tissues.

Technical Progress

Within the first year of the project, we have focused on investigating polymeric and hybrid membranes, developing protocols for printing droplet-based arrays, measuring membrane permeability, and establishing high- and low-fidelity simulation tools for bilayer mechanics and collective folding behavior of droplet networks. In particular, to enable membrane-divided tissues built from amphiphilic BCPs, we studied membrane formation between aqueous compartments in oil with various low-molecular weight diblocks, including poly(butadiene)-poly(ethylene oxide) (PBPEO, 1 kDa). We were able to form stable, homogeneously mixed hybrid membranes made with PBPEO (0-20 mol%) and DPHPC lipids. We measured lateral tensions and adhesive energy in hexadecane, leading to insights for improving the insertion of membrane active peptides. We developed electrophysiological setup for bilayer characterization and used it to demonstrate the large channel membrane proteins can be incorporated into droplet bilayers. For printing droplet-based arrays, we developed protocols, including custom G-codes, for using a commercial bioprinter (BioX from CellInk) to automatically dispense microdroplets (50-250 μm in diameter) into oil to build droplet tissues (**see Figure**). Folding behavior was likely inhibited by the shape of our reservoir used to contain the assembly. A new substrate has been designed to avoid

this issue. To enable quantitative predictions of droplet volume change due to osmosis, we developed a two-droplet experiment to characterize passive transport which informed our droplet network simulations. To guide printing of droplet assemblies, our droplet network simulator accounts for bilayer adhesion and folding through mass transport. To capture the mechanics of lipid membranes encasing the compartments, we developed a three-dimensional surface-enriched finite element model of curvature-dependent interface energetics at large deformations.

Future Plans

PIs will continue to improve and stabilize the 3D printing capabilities of compartments embedding BCP microgels, BCP-coated droplets, addressing the current limitations in self-actuation, bilayer formation, and embedding transmembrane proteins. We will study the effect of oil length on the self-assembly and adhesion between BCP-coated droplets. Using the two-droplet water transport approach, we will quantify the enhancement in membrane water permeability due to the insertion of pore-forming species, including water-selective aquaporin or gramicidin channels with a known MW cutoff for passive transport. The in-house capabilities for characterizing transport on protein-loaded bilayers will allow the design of optimal transport pathways on 3D networks**.** The transport properties will inform our computational models, accounting for deformations along with species diffusion. A milestone is to expedite water transport compared to lipid or BCP-only membranes to enable rapid actuation. We plan to print tissue constructs with self-actuating properties, guided by our droplet-network simulator. We will attempt to directly print self-assembly hydrogels made from diblock copolymers to further generalize our framework. Our objective for the future also involves micromechanical characterization of compartments and modeling and validating suitable membrane-solid models. Single compartments will be synthesized and tested under mechanical loading. For calibration, we will use a least-square approach in which the simulations will mimic the loading and interface failure tests and iteratively identify the effective unknown parameters. The extrapolative ability of our models will be tested by simulating the held-out dataset from experiments.

Broader Impacts and Workforce Development

The project has cross-disciplinary exchanges between materials science, biophysics, engineering, and computational sciences to design tissue-inspired multifunctional materials. This project is serving to train 4 PhDs and 3 undergraduate students from different disciplines at the 3 institutions. In addition to our regular online team meetings, several team members traveled to other campuses to exchange hands-on knowledge on 3D printing and material synthesis. A simulation tool developed to capture collective droplet behavior is made publicly available. We are developing short courses on the modeling, design, characterization, and testing of tissue-like materials at the ASME Smart Materials Adaptive Structures, and Intelligent Systems Conference in the upcoming year.

Data Management and Open Access

The PIs are working on post-processing current experimental data for disseminating it publicly. The software for predicting the folding behavior of droplet networks is freely available on https://github.com/cmsmlab/natedrop including a documentation and example scripts.

Advancing Along the Materials Development Continuum and Partnerships to Translation

We have successfully modified and continue to improve a commercial 3D bioprinter for dispensing customized cellular compartments. Our predictive computational tools are expected to provide G-codes including the optimal directions (where to travel and how to dispense) for this printer. A real-time feedback loop based on our surrogate ML models is promising to assess and accelerate the development of modular, tissue-like constructs. We envision to create G-code libraries for candidate designs. This framework based on a commercial printer offers a modular, generalizable materials solution, whereas manual or custom printers would otherwise limit the broader accessibility.

Publications and References

1. S. Koner, T. Tawfik, F. Mashali, K. B. Kennison, W.T. McClintic, F. A. Heberle, Y-M. Tu, M. Kumar, and S. A. Sarles, *Homogeneous hybrid droplet interface bilayers assembled from binary mixtures of DPhPC phospholipids and PB-b-PEO diblock copolymers*, BBA Biomembranes, accepted for publication (2022).
2. C. Lang, E. C. Lloyd, K. E. Matuszewski, Y. Xu, V. Ganesan, R. Huang, M. Kumar, and R. J. Hickey, *Nanostructured block copolymer muscles*, https://doi.org/10.1038/s41565-022-01133-0 (2022).
3. B. Dortdivanlioglu and A. Javili, *Plateau Rayleigh instability of soft elastic solids. Effect of compressibility on pre and post bifurcation behavior*, Extreme Mechanics Letters **55**, 101797 (2022).
4. A. Rastogi and B. Dortdivanlioglu, *Isogeometric analysis of energetic curvature-resisting material surfaces*, Computer Methods in Applied Mechanics and Engineering, submitted (2022).
5. C. Lang, M. Kumar, and R. J. Hickey, *Current status and future directions of self-assembled block copolymer membranes for molecular separations*, Soft Matter **17**, 10405-10415 (2021).

DMREF: Collaborative Research: Design and Synthesis of Novel Materials for Spin Caloritronic Devices

Lead Investigator: Gregory A. Fiete, g.fiete@northeastern.edu
Participating Institutions: Johns Hopkins University, University of Texas at Austin
Website: https://fiete.sites.northeastern.edu/spincaloritronics/
Keywords: Antiferromagnetism, canted spin structure, spin Seebeck effect, spin caloritronic devices

Project Scope

This project includes three ingredients: (i) theory-aided material design, (ii) synthesis and characterization, and (iii) device fabrication and measurement. We addressed the question of whether antiferromagnetic insulators with a canted spin structure can generate a spin current and exhibit new spin caloritronic effects (the ferrimagnet $Y_3Fe_5O_{12}$ has collinear spins). We studied a family of perovskites $RFeO_3$ (R=rare earth). Our important finding includes the first observations of the spin swapping effect and the vector spin Seebeck effect, related to the unique canted spin structure in perovskite $RFeO_3$. We have also developed a website which provides useful information for the spintronic community.

Relevance to MGI

Previously, spin caloritronics has been limited to ferromagnetic materials with collinear arrangement of moments. This DMREF project has greatly expanded spin caloritronics to the new realm of non-collinear canted antiferromagnets. The success of this DMREF team stems from the close interaction of its three essential components, crystal growth and thin film deposition, device fabrication and measurements theory and modelling. The $RFeO_3$ crystals were grown at UT-Austin and sent to Johns Hopkins University for fabrication of devices and measurements. These unusual results were analyzed and discussed regular zoom meetings during COVID with additional theoretical and model calculations at Northeastern University leading finally to the important new physics of spin swapping and the vector spin Seebeck effect. The theoretical components provide a clear guide as to which material properties are central to our experimental observations. Our project's success was thus the results of close interaction and integration of materials, devices, and theories, which propelled us to explore new frontiers in spin caloritronics, far beyond collinear ferromagnets. Having already made simple devices, we have demonstrated a proof concept which can now be readily transitioned out of the university to applications in technologies, in alignment with the goals of the MGI.

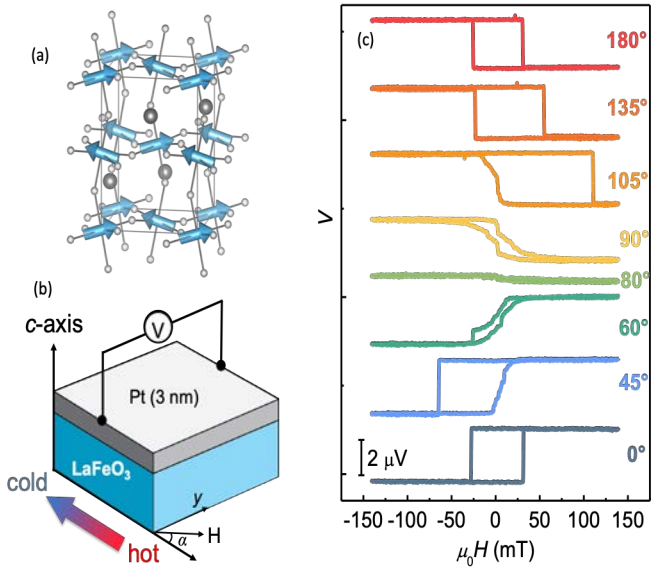


Figure 1. (a) The model of perovskite structure with canted spins (arrows); (b). the schematic of spin caloritronic device. (c) The magnetic field angle dependence of thermovoltage from the device in (b).

Technical Progress

Our high-quality perovskite $RFeO_3$ crystals were grown by the floating method. The magnetization measurements with a vector vibrating sample magnetometer indicate that $RFeO_3$ has a canted antiferromagnetic spin structure [Fig. 1(a)] with a small net magnetization along the *c*-axis and along the *a*-axis. The spin caloritronic

devices were fabricated by depositing a thin layer of heavy metals, either Pt or W, on the polished surface of the oriented crystals. The heavy metal layer is used to convert the spin current generated in $RFeO_3$ with a temperature gradient into an electric current and measured as an electrical voltage through the inverse spin Hall effect. The schematic of the device and the transvers configuration of the measurement is shown in Fig.1(b). A magneto-thermovoltage has been observed for both cases where the thermal gradient was applied vertically and in-plane with different crystal orientations. The magnetic field dependence of the thermovoltage in the transverse configuration is shown in Fig.1(c). In addition to the usual longitudinal efect, we observe the transverse spin Seebeck effect for the first time in an electrically insulating antiferromagnet. Moreover, our work reports the first experimental observation of the spin swapping effect, predicted by Lifshits and Dyakonov in 2009.

Future Plans

We have made full investigations on two members of the perovskite $RFeO_3$ family, $LaFeO_3$ and $LuFeO_3$, and already revealed two new phenomena in spin caloritronics: (1) spin swapping and (2) the vector spin Seebeck effect. In order to explore the full extent of the spin caloritronic phenomena, we need to map out the entire family of $RFeO_3$ which has 17 members. In particular, the effect of the exchange coupling between the rare-earth moment and iron moment on the spin Seebeck remains unknown. Since the scope of the proposed work exceeds the capacity of the three groups working in the "prior" mode, we plan to use the machine learning for device optimization without working on every member in the family. Moreover, since the thin films are necessary for application, we plan to grow thin films by the pulse laser deposition or the radio frequency sputtering technique.

Broader Impacts and Workforce Development

Three graduate students and two postdocs have been supported through this project. All participate in regular Zoom meetings and interact with one another across the theory, synthesis, and device fabrication components. As a result, they see the full scope of the research and gain experience in working on a research team. All PIs have public outreach components for their work, and there is a DMREF website (link above) that highlights key accomplishments.

Data Management and Open Access

All data from the project will be made available to the public for verification and formatted in a way usefule for machine learning. All experimental and computational data will be formatted in a standard form in the field and become searchable. Any interested party may directly contact any PI of the project.

Advancing Along the Materials Development Continuum and Partnerships to Translation

As we describe above, having already made simple devices, we have demonstrated a proof concept for spin caloritronic devices which can now be readily transitioned out of the university to applications in technologies, in alignment with the goals of the MGI. The fully integrated theory, synthesis, and device fabrication/characterization components were instrumental in our breakthrough work given in Ref.[1].

Publications and References

[1] Weiwei Lin, Jiaming He, Bowen Ma, Matthew Matzelle, Jinsong Xu, John Freeland, Yongseong Choi, Daniel Haskel, Bernardo Barbiellini, Arun Bansil, Gregory A. Fiete, Jianshi Zhou, and C. L. Chien, *Evidence for spin swapping in an antiferromagnet,* Nature Physics (2022) https://www.nature.com/articles/s41567-022-01608-w

[2] Jinsong Xu, Jiaming He, Jianshi Zhou, Danru Qu, Ssu-Yen Huang, and C.L. Chien, *Observation of Vector Spin Seebeck Effect in a Non-Collinear Antiferromagnet,* Phys. Rev. Lett, (under review).

DMREF: Designing Optical Materials with Small-Molecule Ionic Isolation Lattices (SMILES)

Lead Investigator: Amar Flood, aflood@indiana.edu
Participating Institutions: Indiana University
Website: none.
Keywords: crystal engineering, fluorescence, photon upconversion, quantum chemistry, cyberinfrastructure

Project Scope

We will undertake a multi-disciplinary project to create advanced materials with desired optical properties by computational design. This work solves a long-standing problem allowing the performance of fluorescent dyes to be translated to solid materials. The molecular materials are called small-molecule ionic isolation lattices (SMILES) and are easily made by mixing cationic dyes with structure-making anionic receptors, cyanostars. Our goals are to codify, test and apply the rules governing their self-assembly into functional materials. New cyberinfrastructure composed of a database of experimental and computed data will be generated and mined to create advanced optical materials with desired functions for future applications.

Relevance to MGI

We have developed a workflow that uses data mining and quantum chemistry computations to streamline identification of materials to synthesize and test the optical properties experimentally. We plan to use experimental validation of predictions as closed-loop feedback to improve data mining algorithms to accelerate materials creation. This is an interactive design-centric approach combining experiment, theory and computations with databases and data mining to create a central backbone to the materials design process.

Our material targets (blue emission, circularly polarized luminescence, photon upconversion) define specific and desired functions. To create them, we will use central ideas in the SMILES discovery to explore the >100,000 known dyes synthesized over the last 150 years for their use in materials. We will create a database with 1000s of dyes using natural language processing (NLP), and extend it to ~10^{4-5} using new quantum chemistry methods. We will learn the full set of rules governing the creation of SMILES materials and broaden the base of dyes and receptors that can be combined to make new materials. We expect to create new materials that far exceed the numbers of materials that can be imagined by domain expertise alone.

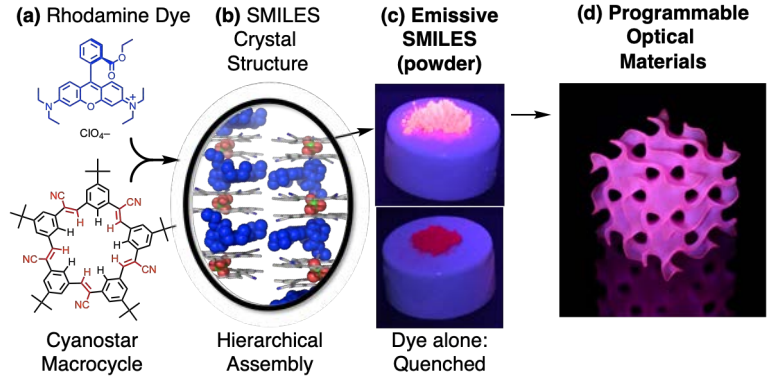


Figure. The basis for this DMREF project is the discovery of the brightest fluorescent materials by (a) the simple mixing of molecular dyes with cyanostar macrocycles. (b) The dyes are isolated allowing them to (c, top) shine brightly over (c, bottom) the dim versions seen when making a material using the dye alone. (d) This discovery creates the first seamless pipeline from synthetic dyes to programmable optical materials showcased here with the fluorescence from a 3D printed gyroid cube.

Technical Progress

The program has been running for 2.5 months since April 1, 2022. We have established regular meetings to facilitate the teaming between the theory, data and experimental teams. We have deployed web services for database generation with prototypes of web scraping tools and customization of natural language processing (NLP) software.

Future Plans

Our multidisciplinary project will establish the first seamless pipeline from fluorescent dyes to optical materials. We will develop a shared data- and theory-driven workflow to create materials (1) with multi-color fluorescence, (2) that combine circularly polarized luminescence with high brightness, and (3) solid-state upconversion with

superior efficiencies. The foundation of this work is the discovery of the brightest fluorescent materials ever reported. They are called small-molecule ionic isolation lattices (SMILES) and are formed by the charge-by-charge packing of cationic dyes alternating with the anionic receptor (cyanostar) complexes. This discovery opens up the first seamless pipeline of dyes from solutions to materials. The purpose of this project is to fully understand, explore and unleash the power of these brightest fluorescent materials.

In Aim 1, we will establish the cyberinfrastructure to predict optical materials by codifying the design rules governing the behavior of SMILES. Efforts include data extraction, enrichment, and extension, leading to a database with ~10^{4-5} dyes and ~10 complexes, and the creation of data mining algorithms to screen the ~10^{5-6} combinations for discovery and development of SMILES materials with specific properties. Once made, they will be tested, and the results fed back to iteratively improve algorithms for 2x speed up in materials creation. Aim 2 seeks to expand the universe of SMILES-compliant dyes. Dyes come in a diversity of shapes, charges and redox properties, and this effort will establish the rules for their use in SMILES. Fundamental knowledge on the structure-function relationships will be fed back to optimize mining algorithms. Aim 3 seeks to create advanced optical materials by applying the materials creation workflow to program blue emission to fill out the optical spectrum, discover circularly polarized luminescence using superior dyes, and undertake a high-risk/reward project to discover a solid-state platform for photon upconversion.

Broader Impacts and Workforce Development

MGI principles will be showcased by promoting workforce and data training, learning and team science in a multi-disciplinary and international collaboration, with outreach to K-12 using science museum exhibits and future career materials, in an undergrad lab on designing fluorescent materials using our workflow, encouraging participation of minority groups, providing open source and databases, and generating community buy-in by planned activities. Google summer of code framework is used to mentor students in cyberinfrastructure creation for generalized data models and user interfaces for services in data management. Success will be reflected in number of materials created, joint papers, survey feedback, students trained, usage statistics, and SMILES users.

Data Management and Open Access

Our digital data will accessed by a public SMILES Gateway with authenticated access for the team to share resources, experimental and computational databases and mining software, and an open database of SMILES materials to make data FAIR. We will exchange data with public databases and extract it from the literature. Our cyberinfrastructure components will be distributed through Github, SciGap or Apache Airavata for extensions to other science gateways to help the materials creation community with tools specific to optical materials. To promote community use, we will deploy gateway services, publish technical reports, YouTube videos on gateway use, and invite collaborations. Education materials and documentation will be published. We will quantify usage statistics, seek to improve them each year, and expect to see a multiplier effect as the SMILES community grows.

Advancing Along the Materials Development Continuum and Partnerships to Translation

We will advance technology transfer, SMILES deployment and workforce training by collaborating with start-up Halophore, Inc (https://www.halophore.com/). The CEO, and lead author on SMILES, Dr. Benson, will meet with the whole team bimonthly to provide feedback on the new technical capabilities emerging from the DMREF activities that have commercial value for medical lasers being pursued in a funded NIH SBIR-I, fluorescence calibration standards to be pursued in an NSF Phase II SBIR, and any new commercial opportunities. For example, obtaining specific emission wavelengths is often of value to meet the requirements of industry partners. Thus, with the creation of a workflow for programming the properties of SMILES-based optical materials in this DMREF, our connection to Halophore, Inc, extends that development pipeline from laboratory to market deployment.

Publications and References

1. J. Chen, S. G. Stenspil, L. Kacenauskaite, N. Lenngren, S. Kaziannis, M. Kloz, A. H. Flood and B. W. Laursen, Quantitative Energy Transfer and Antenna Effect in Organic Nanoparticles Based on Small-Molecule Ionic Isolation Lattices, *ACS Appl. Nano. Mat.* (2022) under revision.

Key paper on which the SMILES materials are based:

2. C. R. Benson, L. Kacenauskaite, K. L. VanDenburgh, W. Zhao, B. Qiao, T. Sadhukhan, M. Pink, J. Chen, S. Borgi, C. Chen, B. J. Davis, Y. C. Simon, K. Raghavachari, B. W. Laursen, and A. H. Flood, Plug-and-play Optical Materials from Fluorescent Dyes and Macrocycles, *Chem* **6**, 1978-1997 (2020)

Midwest Integrated Center for Computational Materials

Lead Investigator: Giulia Galli, gagalli@uchicago.edu
Participating Institutions: Argonne National Laboratory, University of Chicago, University of Notre Dame, University of California Davis
Website: http://miccom-center.org
Keywords: software development, electronic structure, quantum, molecular dynamics, perturbation theory

Project Scope

The Midwest Integrated Center for Computational Materials (MICCoM) develops and disseminates interoperable computational tools – open-source software, data, simulation templates, and validation procedures - that enable simulations and predictions of properties of functional materials for energy conversion and of solid-state materials for quantum information science. The distinctive features of the center are: (i) development of interoperable codes for simulation of materials at multiple length and time scales, (ii) focus on heterogeneous materials, inclusive of defects, interfaces, and building blocks, (iii) focus on spectroscopic and coherence properties.

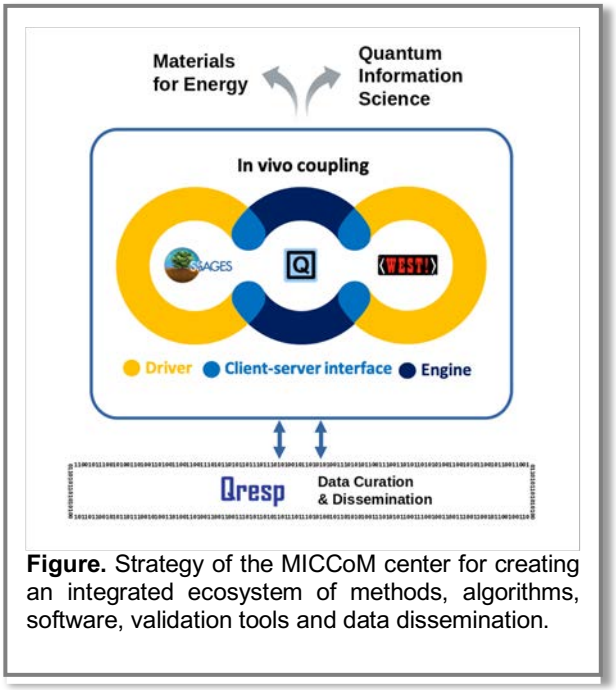


Figure. Strategy of the MICCoM center for creating an integrated ecosystem of methods, algorithms, software, validation tools and data dissemination.

Relevance to MGI

Computational materials science is an integral part of the innovative process leading to materials discovery and design. First-principles codes have a broad potential and they may be used to produce the necessary data to predict properties of vast classes of materials. In this context, MICCoM develops and disseminates interoperable and validated computational tools that enable a growing user community to simulate and predict properties of functional materials for energy conversion and of solid-state materials for quantum information sciences.

Technical Progress

The center develops interoperable codes for the simulations of materials at the atomistic, quantum mechanical level: a first principles molecular dynamics (FPMD) code, Qbox (http://qboxcode.org/), coupled with advanced sampling software, SSAGES (https://ssagesproject.github.io/), and with software that simulates spectroscopic properties, WEST (http://west-code.org/). The center has recently released the PyCCE code (https://pycce.readthedocs.io) for simulating the dynamics of spin qubits interacting with a spin bath using the cluster-correlation expansion method. The center has developed and implemented a new quantum embedding method, QDET, to compute strongly correlated states in periodic systems. QDET was used to simulate spin qubits in diamond and silicon carbide, with proven ability to carry out parts of the calculations on quantum computers. The center has pushed forward strategies to implement machine learning in electronic structure and enhanced sampling codes. The center has also developed several techniques to address the effect of the quantum vibronic coupling on the electronic properties of molecules, solids, and molecular crystals. A python package that simulates photoluminescence spectra was developed and validated against experiments. The center has carried out simulations of the mechanisms of spin defects' formation and stability using advanced sampling techniques coupled to first principles molecular dynamics. Finally, the center has carried out the porting of the codes to NVIDIA GPUs. The WEST code was demonstrated to scale up to the entire Summit supercomputer at OLCF.

Future Plans

The center will continue to innovate the field by developing new electronic structure and sampling methods, in particular to address the simulation of excited states and the stability of materials for energy and quantum information science. A defining strategy for the center is to push forward the coupling of codes so as to develop flexible and extensible complex workflows where codes cooperate, possibly operating also on different architectures, including both classical (CPUs and GPUs) and quantum computers (QPUs). The center seeks opportunities to integrate machine learning protocols in the codes, while defining protocols to rigorously verify and validate them. The center is addressing the challenge of porting the codes to exascale computers.

Broader Impacts and Workforce Development

The software activity of the center encompasses three interrelated areas summarized: (i) development of methods and algorithms for materials simulations and their proof-of-principle implementation by students and postdocs; (ii) optimization and delivery to the community in a user-friendly format of the codes developed in (i) by researchers; and (iii) research and development of scalable algorithms and code design strategies for porting to pre-exascale and exascale architectures with active engagement of DOE leadership class facilities and hardware vendors. The center is targeting the development of a workforce with distinctive skills in software development and capable of running exascale calculations by sponsoring training events (e.g., schools, courses, and hackathons).

Data Management and Open Access

A key guiding principle for materials research is to make data findable, accessible, interoperable, and reusable. We have developed the Qresp software (http://qresp.org) to facilitate the dissemination of data on a paper-by-paper basis and in a distributed manner. Qresp represents a technical infrastructure solution that facilitates the organization of research data for reuse and repurpose, as mandated by federally funded research projects.

Advancing Along the Materials Development Continuum and Partnerships to Translation

The center has developed strategies to engage hardware vendors (e.g., NVIDIA, Intel) with software developers. The computational infrastructure developed within the center is deployed in several other programs, including several EFRCs (CHOISE, NPQC, QMEENC, AMEWS) and the Q-Next national quantum research center, where tight connections with industry are in place.

Publications and References

1. W.-Z. V. Yu, and M. Govoni, *GPU acceleration of large-scale full-frequency GW calculations*, J. Chem. Theory Comput., under review (2022). arXiv:2293.05623.
2. N. Sheng, C. Vorwerk, M. Govoni, and G. Galli, *Green's function formulation of quantum defect embedding theory,* J. Chem. Theory Comput. **18**, 3512 (2022).
3. E. Lee, A. Yu, J. de Pablo, and G. Galli, *Stability and molecular pathways to the formation of spin defects in silicon carbide,* Nat. Commun. **12**, 6325 (2021).
4. M. Onizhuk, and G. Galli, *PyCCE: a python package for cluster correlation expansion simulations of spin qubit dynamics,* Adv. Theory Simul. **4**, 2100254 (2021).
5. Y. Jin, M. Govoni, G. Wolfowicz, S. Sullivan, J. Heremans, D. Awschalom, and G. Galli, *Photoluminescence spectra of point defects in semiconductors: validation of first principles calculations,* Phys. Rev. Materials **5**, 084603 (2021).
6. A. Kundu, M. Govoni, H. Yang, M. Ceriotti, F. Gygi, and G. Galli, *Quantum vibronic effects on the electronic properties of solid and molecular carbon,* Phys. Rev. Materials **5**, L070801 (2021).
7. S. Dong, M. Govoni, and G. Galli, *Machine learning dielectric screening for the simulation of excited state properties of molecules and materials,* Chem. Sci. **12**, 4970 (2021).
8. M. Govoni, J. Whitmer, J. de Pablo, F. Gygi, and G. Galli, *Code interoperability extends the scope of quantum simulations,* npj Comput. Materials **7**, 32 (2021).
9. E. Lee, T. Ludwig, B. Yu, A. Singh, F. Gygi, J. Nørskov, J. de Pablo, *Neural network sampling of the free energy landscape for nitrogen dissociation on ruthenium,* J. Phys. Chem. Lett. **12**, 2954 (2021).
10. K. Harmon, K. Letchworth-Weaver, A. Gaiduk, F. Giberti, M. Chan, F. Gygi, P. Fenter, and G. Galli, *Validating first-*

principles molecular dynamics calculations of oxide/water interfaces with x-ray reflectivity data, Phys. Rev. Materials **4**, 113805 (2020).

Fast and Accurate Prediction of Material Properties with Three-Body Tight-Binding Model for the Periodic Table

Lead Investigator: Kevin F. Garrity, Materials Measurement Lab, NIST, kevin.garrity@nist.go; Kamal Choudhary, Materials Measurement Lab, NIST, kamal.choudhary@nist.gov
Website: https://pages.nist.gov/ThreeBodyTB.jl/.
Keywords: Density functional theory, tight-binding, band structure, high-throughput

Project Scope

Parameterized tight-binding models fit to first principles calculations provide a minimal quantum mechanical model for materials properties, but existing models are of uncertain accuracy and are mostly hand-fit to a few materials. Here, we introduce a methodology including explicit two-body and three-body contributions and fit the model every pair of atoms in the periodic table (main group and transition metals), increasing the utility and accuracy of the method.

Relevance to MGI

First principles density functional theory (DFT) calculations are used throughout the MGI as a practical tool to understand materials properties, but the calculations remain computationally expensive and difficult to apply either to large systems or for systematic searches. Pure machine learning approaches can fill in some gaps, but due to the lack of an underlying physical model, they are limited when extrapolating to new phenomena or when large datasets are unavailable. In this work, we develop a physics-based tight-binding approach that incorporates basic quantum mechanics and electrostatics, which is combined with a large amount of data from high-throughput DFT calculations (>800,000 calculations) to fit a minimal model to describe chemical bonding. These tight-binding calculations for total energies, forces, etc., as well as electronic properties like the band structure, run hundreds of times faster than the underlying DFT calculations. This speed enables their application to many materials discovery and materials design applications that are otherwise difficult to attack. As a first example, we have showed that the models can provide elemental surface and vacancy energies with high accuracy, despite being fit to only small systems.

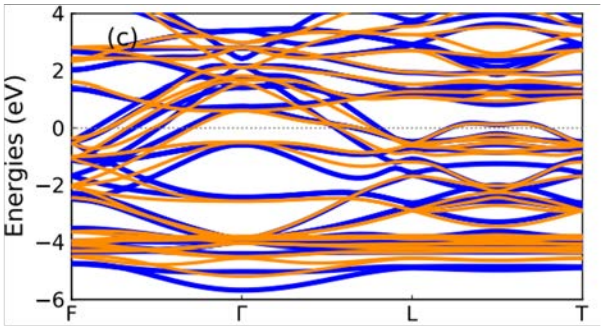


Figure. Example *out-of-sample* prediction of the band structure of Au_2Bi_8. Blue lines: DFT, Orange lines: tight-binding model (this work).

Technical Progress

Building on previous work in the field (see for example Ref. 1), we have developed our tight-binding methodology. We have demonstrated the increase in accuracy that comes from the explicit inclusion of three-body contributions to the tight-binding model, in addition to the typical two-body interactions. We have fit our model in a high-throughput manner to DFT calculations across the periodic table. We employ an iterative process where we 1) fit the model to a dataset, 2) use the model to search for new low energy structures, and then 3) include those structures into our new fitting set, repeating as necessary. This ensures that the model is fit to relevant structures for every possible chemistry. Our resulting model can predict total energies of bulk structures out-of-sample with an accuracy of 50 meV/atom, lattice constants with less than 1% error, as well as the band structure, forces, density of states, *etc.*

Future Plans

We are currently expanding our model to including magnetism, and future plans include adding fitting data from ternary systems, instead of just elemental and binary systems.

Data Management and Open Access

The code for the project is available at https://github.com/usnistgov/ThreeBodyTB.jl , written in the modern Julia programming language. It is also available directly from the Julia package manager. We also have a python interface at https://github.com/usnistgov/tb3py . The underlying DFT fitting data is available as part of the Jarvis project at NIST, at https://jarvis.nist.gov/jarvisqetb/ .

Advancing Along the Materials Development Continuum

This project enables fast and accurate quantum mechanics-based calculations of materials properties, which can be used for a variety of MGI-related projects. Simple calculations can be performed by a non-expert in seconds. Initial applications have looked at automated calculations of surface energies and defects, which are difficult to treat with current methodologies.

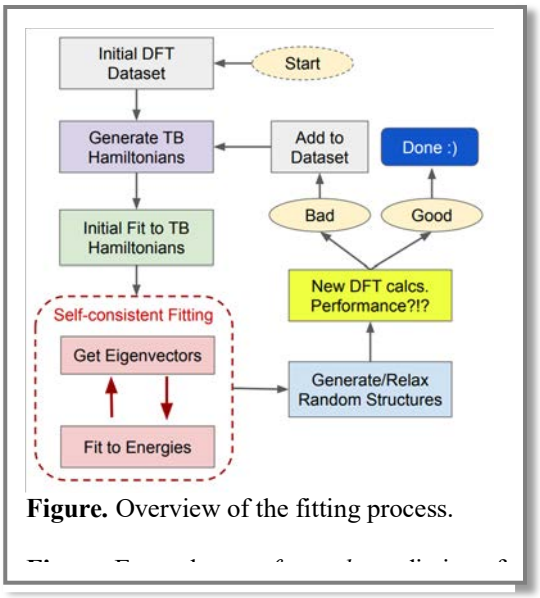


Figure. Overview of the fitting process.

References

1. P. Koskinen, V. Mäkinen, "Density-functional tight-binding for beginners", *Computational Materials Science*, **47,** 237 (2009)

Publications

1. K.F. Garrity, K. Choudhary, "Fast and Accurate Prediction of Material Properties with Three-Body Tight-Binding Model for the Periodic Table", arXiv: 2112.11585 (2021)

DMREF: Design of surface functionality through surface composition and structure

Lead Investigator: Andrew J. Gellman, gellman@cmu.edu
Participating Institutions: Carnegie Mellon University, Columbia University
Website: none
Keywords: alloys, surface chemistry, catalysis, segregation, structure sensitivity

Project Scope

The project objective is to accelerate the development of complex functional surfaces such as those of alloy catalysts. We are accelerating computational prediction of surface properties while concurrently developing high throughput measurements for prediction verification. These methods are spanning parameter spaces such as alloy composition and surface structure that are intractable to traditional methods of study.

Relevance to MGI

We focus on two parameter spaces: alloy composition, $A_xB_yC_{1-x-y}$, and surface structure, M(*hkl*). Both influence surface chemistry relevant to catalysis. Both are described by continuous variables; challenging to both experiment and computation. We address both challenges. For example, comprehensive measurement of surface segregation spanning composition space in $Cu_xPd_yAu_{1-x-y}$ and $Cu_xAg_yAu_{1-x-y}$ has been completed concurrently with computational simulations of segregation. Ultimately, this cooperation will lead to accurate computational predictions of surface segregation in reactive environments where experiments have no reach. Similarly, recent developments in strategies for relaxation of molecular structures on surfaces have accelerated the ability to study surfaces with complex structure. This will aid efforts to understand a dataset measuring the kinetics of O_2 adsorption on all Cu(*hkl*) surfaces lying within 12° of the Cu(111) plane.

Technical Progress

Surface Segregation. High throughput methods have been used to both measure and simulate surface segregation in $Cu_xPd_yAu_{1-x-y}$ and on $Cu_xAg_yAu_{1-x-y}$ alloys across their entire composition spaces.[1,2] These are far more compressive than any prior study. Initial work (Fig. 1A) revealed a discrepancy in the order of Pd and Cu segregation along the CuPd binary. Measurement shows preferential Cu segregation while the simulation shows otherwise. This has been traced to the modeling relying on simulation of an FCC(111) surface while the alloy films are polycrystalline. The predicted order of segregation of Pd versus Cu switches when the surface orientation is FCC(110).

Structure sensitivity. Using a curved Cu(*hkl*) single crystal with the (111) orientation at the center (Fig. 1B), we have mapped the kinetics of O_2 adsorption across surface orientation space with unparalleled fidelity and resolution. These reveal clearly that the rate of dissociative adsorption of O_2 is lowest on the Cu(111) surface. It increases monotonically as the step density increases towards the edges of the curved surface, however, it is also clear that (100) steps are much more active than (110) steps. We are

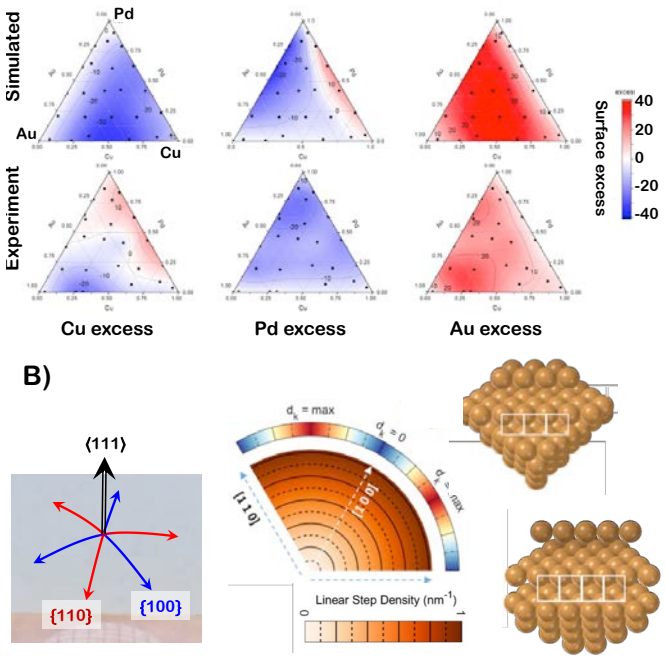


Figure 1. A) Segregation in $Cu_xPd_yAu_{1-x-y}$ measured and computed across ternary composition space. B) A surface structure spread Cu single crystal. The step densities and types vary across the curve surface.

now starting to explore the origins of these structure sensitive differences in reactivity using simulations. These simulations on complex high Miller index surfaces will benefit from some of the methods that we have developed for acceleration of geometry optimization.[3]

Degree of rate control. We have demonstrated the use of automated differentiation as an effective method for accurate calculation of derivatives such as those used in the definition of the degree of rate control.[4] This has been used to analyze a 17-step mechanism for propylene epoxidation, one of the catalytic processes being probed experimentally.[5,6] Our analysis has elucidated the three steps with dominant degrees of rate control. These three steps are the primary dictators of the overall reaction rate. The value of this approach lies in pointing modeling efforts towards those steps in a process whose acceleration will yield the greatest impact on overall reaction rate.

Future Plans

Our plans for the future include the application of accelerated simulation methods to the surface structure sensitivity of O_2 adsorption on Cu(*hkl*) surfaces. In addition, we will generate measurements of catalytic surface reaction kinetics spanning alloy composition space for reactions such as propylene and propyne hydrogenation.

Broader Impacts and Workforce Development

To date the program has graduated 4 students: Dr. Zhitao Guo (PhD, CMU), Dr. Yilin Yang (PhD, CMU), Mr. Matt Adams (MS, CMU) and Mr. Marcus Yu (BS, CMU joining Columbia ChE PhD program). Four more PhD students are currently being trained via this project. The subprojects studying 'alloy segregation' and developing the use of automated differentiation for determination of 'degrees of rate control' have been disseminated to the general public via YouTube videos (segregation - https://youtu.be/lbNANFj928k, degree of rate control - https://youtu.be/2PIEJp8CGNM). Three students have been allotted instrument time at the Center for Functional Nanomaterials at Brookhaven National Lab.

Data Management and Open Access

Data and software are being made openly available at: https://github.com/ulissigroup, https://ulissigroup.cheme.cmu.edu/software-data/, https://github.com/yilinyang1/NN-ensemble-relaxer

Advancing Along the Materials Development Continuum and Partnerships to Translation

Impact on materials R&D will arise from transfer and application of the methods (both experimental and computational) developed in the course of this project. The work so far has led to collaborations with researchers at Brookhaven NL and elsewhere.

Publications and References

1. C. Yin, Z. Guo, A.J. Gellman, "*Surface Segregation Across Ternary Alloy Composition Space: $Cu_xAu_yPd_{1-x-y}$*" Journal of Physical Chemistry – C **124**, 10605 – 10614 (2020) (10.1021/acs.jpcc.0c02058?ref=pdf)
2. Yilin Yang, Zhitao Guo, Andrew Gellman and John Kitchin, *Simulating segregation in a ternary Cu-Pd-Au alloy with density functional theory, machine learning and Monte Carlo simulations*, J. Phys. Chem. C, (2022). https://pubs.acs.org/doi/abs/10.1021/acs.jpcc.1c09647
3. Yang, Y., Omar A. Jiménez-Negrón, and Kitchin, J. R., *Machine-learning accelerated geometry optimization in molecular simulation*, Journal of Chemical Physics, **154**(23), 234704 (2021). http://dx.doi.org/10.1063/5.0049665
4. Yilin Yang, Siddarth K. Achar, John R. Kitchin, *Evaluation of the Degree of Rate Control via Automatic Differentiation*, AIChE Journal, (2022). https://doi.org/10.22541/au.162855056.60483260/v1.
5. W.N. Porter, Z. Lin and J.G. Chen, "*Experimental and theoretical studies of reaction pathways of direct propylene epoxidation on model catalyst surfaces*", Surface Science Reports **76** (2021) 100524
6. W.N. Porter, Z. Lin and J.G. Chen, "*Elucidating interactions of the epoxide ring on Pt(111) by comparing reaction pathways of propylene oxide and 1-epoxy-3-butene*". Journal of Vacuum Science and Technology A (Special issue honoring Pat Thiel) **39** (2021) 063214.
7. Yilin Yang, Mingjie Liu, John Kitchin, *Neural Network Embeddings based Similarity Search Method for Catalyst Systems*, submitted to Digital Discovery, 2022.
8. N. Shukla, M. Yu, A. Pradhan, Y. Han, A.J. Gellman, *Chirality Retention in Aqueous Propylene Oxide Hydration: Chirality of the Transition State* Israel Journal of Chemistry **61**, (2021), 743–749, (10.1002/ijch.202100098)

Unraveling the Photophysical Properties and Charge Transport Mechanisms in Hybrid Organic-Inorganic Nanomaterials

Lead Investigator: Raja Ghosh and Francesco Paesani
Participating Institutions: University of California San Diego
Website: none
Keywords: Photophysics, Charge Transport, Coherence, Metal Organic Frameworks, Perovskites

Project Scope

A fundamental understanding of the electronic excited state processes that govern the photophysical phenomenon and charge transport mechanisms of multifunctional materials is crucial for new scientific breakthroughs and future technological innovations. To achieve this goal, our project aims to develop a new theoretical framework based on the Holstein Hamiltonian[1-3] that will not only provide unambiguous interpretations of optical and transport signatures in hybrid organic-inorganic nanomaterials (HOINs) such as metal-organic frameworks and perovskites, but also unravel the underlying physical mechanisms driving energy and charge transport in HOINs. We envision that the synergistic integration of electronic structure methodologies and "first principles'' approaches based on the Holstein model will allow us to identify and establish new, transformative design strategies that will guide the synthesis and characterization of next generation HOINs optimized for a broad range of optoelectronic, spintronic, and photonic applications.

Relevance to MGI

The close agreement between experiment and theory is the primary focus of our work: An accurate theoretical model that quantitatively reproduces the lower energy mid infrared (< 1000 cm^{-1}) polaronic signatures and the UV-vis excitonic signatures for a wide range of organic and hybrid materials will result in atomistic-level characterization of structure-property relationships that will not only provide critical insights into the properties of known HOINs but also pave the way for the rational design of new hybrid materials suitable for various technological applications. Elucidating structure-phoptophysics-property relationships across a broad range of functional materials is non-trivial and requires our fundamental understanding of the intricate interplay among excitons (electron-hole pair), polarons (charges), bipolarons, phonons (vibrations), inter-layer stacking interactions, and different forms of structural and conformational defects. In parallel with electronic structure modeling and data-driven science that are routinely pursued; an accurate, computationally cheap, and physically sound theoretical model that consistently makes quantitative connections with conceptually complicated experimental observations in chemical physics and materials chemistry (Fig. 1) is equally important to further accelerate materials discovery.

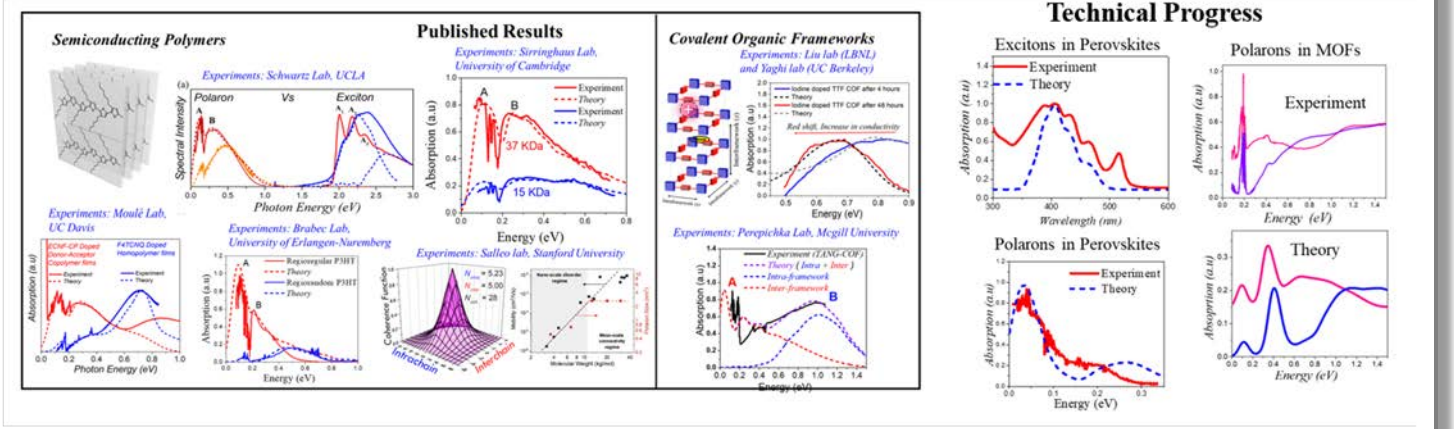


Figure1. Overview of the performance of the Holstein Model using the Multiparticle Approach.

Technical Progress

The newly proposed theoretical model, the "Frenkel-Wannier-Holstein" model is currently under development with promising results so far. The close agreement with experiments and theory as shown in Fig. 1 probes the local order of hybrid materials such as MOFs and perovskites with exciton and polaron delocalization.

Future Plans

1. Comprehensive investigation of the mid-infrared optical properties in redox-active MOFs and perovskites to establish direct correlations between polaronic spectral signatures and polaron coherence lengths.
2. Continued development of a new theoretical model, the "Frenkel-Wannier-Holstein" model, that elucidates structure-photophysics-property relationships in HOINs based on the interactions among photons, Frenkel excitons, charge-transfer excitons, Wannier-Mott excitons, and phonon modes.
3. Investigation of the excitonic spectral signatures to establish direct correlations between UV-vis spectral signatures, exciton delocalization, and energy transport in MOFs and perovskites.
4. Computational engineering of new organic cations for improved out-of-plane conductivity in perovskites.

Broader Impacts and Workforce Development

A broader audience will also be reached through periodic research reports on the group website and social media account as well as on the websites of the Physical Sciences Division and San Diego Supercomputer Center at UC San Diego, and on EurekAlert!, where our research has already been highlighted in the past. The simulation methodologies developed as part of the proposed research activities will be made available to the community through our group GitHub page. The development and application of our theoretical/computational methodology will involve the training and education of one graduate student and one postdoctoral fellow as well as several undergraduate students, who will acquire a solid foundation in theoretical physical chemistry and materials science, with particular focus on electronic structure and quantum dynamics, and high-performance computing. The interdisciplinary nature of the proposed project will provide a unique opportunity for all the students to establish bridges and inter-connections between quantum mechanics, statistical mechanics, solid state physics, and energy applications. Our outreach component also includes several initiatives for promoting the interest in STEM disciplines of high school and college students from underrepresented and underprivileged groups. In particular, every summer, the PI will be directly involved in the STARS (Summer Training Academy for Research Success) and ENLACE programs that offer high school, community college, undergraduate, and masters students the opportunity to be involved in an 8-week research period with UC San Diego faculty.

Data Management and Open Access

Data requests from interested researchers will be fulfilled upon request. Results obtained from the analyses of the data will be presented at conferences and published in peer-reviewed journals. Specific computational details necessary to reproduce the results of our publications, along with all the data of both figures and tables included in the publications, will be made available on our group's repository on GitHub: https://github.com/paesanilab/Data_Repository. Besides allowing any interested researchers to use and build upon our results, making the actual row data available, along with the input and output files associated with the corresponding calculations, will allow any interested researcher to reproduce and verify our findings.There are no restrictions on privacy or confidentiality associated with this project, and the data generated do not present a security concern.

Publications and References

1. R. Ghosh and F. Paesani, *Unraveling the effect of defects, domain size, and chemical doping on photophysics and charge transport in covalent organic frameworks*, Chem. Sci., **12,** 8373–8384
2. R. Ghosh and F. Paesani, *Topology Mediated Enhanced Polaron Coherence in Covalent Organic Frameworks*, *J. Phys. Chem. Lett.*, **12,** 9442–9448
3. R. Ghosh and F.C. Spano, *Excitons and Polarons in Organic Materials*, *Acc. Chem. Res.*, **53,** 2201–2211

Toward exascale computing of electron-phonon couplings for finite-temperature materials design

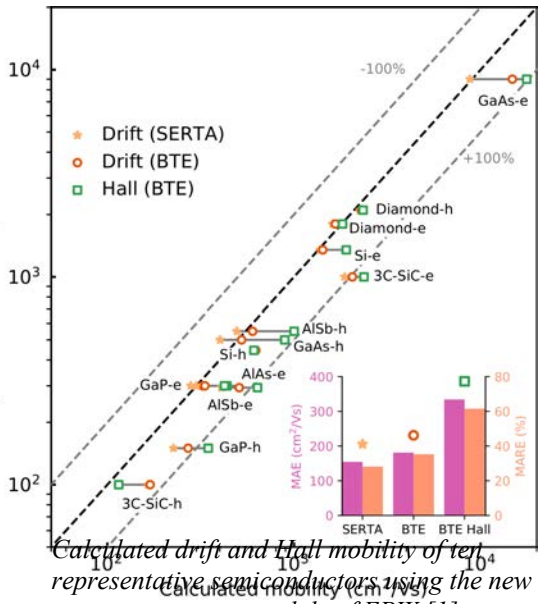
Lead Investigator: Feliciano Giustino, fgiustino@oden.utexas.edu
Participating Institutions: The University of Texas at Austin (Feliciano Giustino), University of Michigan at Ann Arbor (Emmanouil Kioupakis)
Website: https://epw-code.org
Keywords: Electron-phonon interactions, phonon-assisted quantum processes, exascale computing, open-source software development

Project Scope

The overarching goal of this project [DOE/BES CMS Award DE-SC0020129, Program Manager: Matthias Graf] is to enable *ab initio* calculations of the electronic, optical, and transport properties of solids including the effects of temperature and zero-point quantum fluctuations of the lattice. In density-functional theory and many-body perturbation theory approaxches, electrons are adiabatically decoupled from the atomic motion, and the atomic nuclei are described as classical particles. As a result, standard calculations do not take into account the quantum nature of lattice vibrations and the resulting phonon-mediated quantum processes. This project aims to overcome such limitations by systematically accounting for electron-phonon interactions in *ab initio* calculations. To achieve this goal we are advancing the theoretical foundations of electron-phonon physics, and we are developing novel *ab initio* computational methods and software for predicting related materials properties, in particular our flagship open-source software code EPW (https://epw-code.org).

Relevance to MGI

This project belongs to DOE's Computational Materials Sciences (CMS) program which started in 2015 as part of the MGI. The overarching goal of the CMS program is develop validated community codes and databases to underpin the predictive design of functional materials. Our team contributes to the DOE CMS and the MGI along two complementary directions: 1) To expand the predictive capabilities of the EPW software, for example by adding new modules to perform predictive *ab initio* calculations of optical spectra at finite temperature, as well as temperature-dependent carrier mobilities, conductivity, and resistivity in metals and semiconductors. 2) To refactor EPW in preparation of the exascale transition. These efforts are complemented by a strong engagement with the user community, via extensive documentation, online tutorials, YouTube lectures, a lively user forum, and the organization of multiple summer schools both virtual and in person. These advances in methodology, software, and training will accelerate the development of new materials for electronics, optoelectronics, energy harvesting, and energy-efficient lighting.



Calculated drift and Hall mobility of ten representative semiconductors using the new magnetotransport module of EPW [1].

Technical Progress

Among key scientific advances, we extended the implementation of the *ab initio* Boltzmann transport equation (aiBTE) to describe the effect of combined electric and magnetic fields for Hall mobility measurements. In Hall measurements the mobility is measured in the presence of a driving field in the direction of the current, while a magnetic field perpendicular to the current flow is simultaneously applied. We can now compute Hall effect mobilities, which can directly be compared with the majority of transport measurements. This implementation has been validated on several standard semiconductors, see figure above [1]. Complementary work includes the application of aiBTE to understand the mobility limits of materials for power electronics, such as cubic BN and diamond. This study identified and explained the fundamental

limitations to the performance of cBN and diamond in power-electronic devices, and made the cover of Applied Physics Letters [2]. On the front of software development, we refactored the EPW code to harness the capabilities of current and future supercomputing architectures. For example, we re-engineered the data structure in order to increase data locality, redesigned the code to implement prefetch-aware cache management, and introduced parallel I/O via HDF5. These changes led to a 10x speedup in specific modules of EPW, such as the module for calculating superconducting properties from first principles. To test these improvements we performed benchmarks on the entire Frontera supercomputer at the Texas Advanced Computing Center (448,448 Intel Cascade Lake cores) and we successfully executed EPW on up to 7,840 nodes, corresponding to 439,040 compute cores. For the largest test we achieved a speedup of 86%. This test constitutes the first demonstration of the execution of the code at 40 petaFLOPs, and indicates that the EPW code has reached the pre-exascale regime.

Future Plans

In the next reporting period we will prioritize the following tasks: (1) To release a major revision of the EPW code (v6) including the new multi-level MPI/OpenMP parallelization across all modules (transport, superconductivity, optics), simplified user interface, minimal I/O and re-designed data structure. (2) To publish the technical manuscript describing the new EPW v6, including contributions from all members of the EPW collaborations. (3) To finalize and publish ongoing calculations of polarons on several materials classes. (5) To finalize ongoing calculations of phonon-assisted optical absorption and Auger recombination. (6) To finalize calculations of ionized- impurity scattering in carrier mobilities and release the corresponding module in EPW.

Broader Impacts and Workforce Development

During June 2021 we organized a virtual summer school on electron-phonon phyiscs and the EPW code. This school consisted of 5 days of morning lectures and afternoon hands-on tutorials (06/16 to 06/18/2021), and hosted 177 participants from across the globe. All materials from this school have been made available through the EPW website, https://docs.epw-code.org/doc/School2021.html. During June 2022 we organized an in-person summer school on electron-phonon interactions from first principles (https://epw2022.oden.utexas.edu). The school lasted 8 days, from Sunday 13 to Sunday 19 June, and consisted of morning lectures, afternoon hands-on tutorials, and a weekend hackathon. We welcomed 93 participants to UT Austin (75 attendees and 18 instructors), and we had 155 participants registered for the Zoom streaming. All lectures, tutorials, videos have been made available through the school website (see schedule page: https://epw2022.oden.utexas.edu/74-schedule). The school covered codes such as Quantum ESPRESSO, wannier90, EPW, and BerkeleyGW.

Data Management and Open Access

All software developments of this project are made fully available to the community via two channels: 1) periodic releases of new software versions. We currently have two annual releases, in the Spring and in the Fall. The latest version EPW 5.5 was released on May 31 2022 (https://docs.epw-code.org/doc/Releases.html). Anyone can freely download and use the code without any registration or email requests. Users are encouraged to interact with developers via the user forum (https://forum.epw-code.org) which currently counts 3000+ posts and 650+ members. 2) The gitlab repository of the official version of the code is accessible to all withou limitation: https://gitlab.com/QEF/q-e/-/tree/develop/EPW. This version is updated in real time and can be used by more advanced users who are interested in the code structure. The code is released under GNU GPL license.

Advancing Along the Materials Development Continuum and Partnerships to Translation

This project enables advanced calculations of temperature-dependent materials properties. We expect that the new capabilties developed in this project will facilitate new partnerships at higher TRL levels, and we are currently in the process of exploring these opportunities.

Publications and References

[Including only two publications due to lack of space. Total project publications to date: 24]
1. S. Ponce', F. Macheda, E. R. Margine, N. Marzari, N. Bonini, F. Giustino, "First-principles predictions of Hall and drift mobilities in semiconductors ", Phys. Rev. Res., **3**, 043022 (2021).
2. N. Sanders, E. Kiupkakis, "Phonon- and defect-limited electron and hole mobility of diamond and cubic boron nitride: A critical comparison ", Appl. Phys. Lett., **119**, 062101 (2021).

Tuning liquid crystallinity in conjugated polymers to simultaneously enhance charge transport and control mechanical properties

Lead Investigator: Enrique Gomez, edg12@psu.edu
Participating Institutions: The Pennsylvania State University
Website: https://sites.google.com/site/gomezgroupatpsu/home/tg-of-polymers
Keywords: Organic electronics, nematic, rheology, phase behavior, entanglement

Project Scope

This program aims to develop tools capable of accelerating the prediction of mechanical and conductive properties of conjugated polymers. The combination of simulations and experiments will be crucial to generate accurate coarse-grained simulations capable of predicting liquid crystallinity by combining molecular dynamics simulations with self-consistent field theory calculations. This enables the systematic computational exploration of backbone and side chain architectures that are validated with selected synthesized model materials. Simulations and experiment are also crucial to incorporate nematic order in a unified theory of polymer entanglements, and thereby provide a tool capable of predicting rheological properties from the chemical structure.

Relevance to MGI

The proposed program will leverage progress so far to advance along the materials development continuum. The validation of computational tools that can predict mechanical and electrical properties of liquid crystalline polymers will move the field beyond discovery and development to property optimization and systems design and integration. Indeed, developing coarse-grained models and tight-binding descriptions will accelerate predictions of structure and macroscopic properties, thereby opening the door for computationally-guided property optimization and integration into devices.

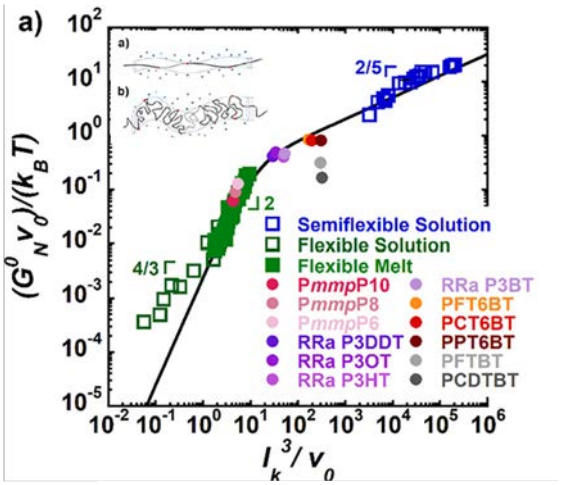


Figure. Conjugated polymer melts (circles) follow Everaers' scaling predictions with dimensionless plateau modulus G_N^0 vs dimensionless Kuhn monomer volume v_0. Flexible melt data (solid green squares), semiflexible solution data (open green squares), and semiflexible solution data were obtained from the literature (see Fenton et al.). The solid black line is a proposed crossover equation derived from our work. The more flexible P*mmp*P polymers fit best with the flexible melt scaling argument $(l_k^3/v_0)^2$ while P3AT, PFT6BT, and PCT6BT and PPT6BT fit best with the semiflexible scaling argument of $(l_k^3/v_0)^{2/5}$. The PFTBT and PCDTBT polymers lie well below the prediction; this is hypothesized to be due to lingering nematic domains slightly above their T_{NI}. The tube model for a semiflexible and flexible polymer can be seen in insets (a) and (b), respectively.

Technical Progress

We have made progress on developing predictive models of the properties of conjugated polymers. For example, our work builds on the relationship between Kuhn length l_k, Kuhn monomer volume v_0, and plateau modulus G_N^0, initially proposed by Graessley and Edwards for flexible polymers, and extended by Everaers. This correlation has a large gap in experimental data between the flexible and stiff regimes, and this gap prevents the prediction of mechanical properties from the chain structure for any polymer in this region. Given the chain architecture, including a semiflexible backbone and side chains, conjugated polymers are an ideal class of material to study this crossover region. Using small angle neutron scattering, oscillatory shear rheology, and the freely rotating chain model, we have shown that 12 polymers with aromatic backbones populate a large part of this gap. We also have shown that a few of these polymers exhibit nematic ordering, which lowers G_N^0. When fully isotropic, these polymers follow a relationship between l_k, v_0, and G_N^0, with a simple crossover proposed in terms of the

number of Kuhn segments in an entanglement strand N_e. In conjunction with linear viscoelastic descriptions (such as BoB), our work highlighted in the Figure can allow the generation of rheological spectra that would provide insight to other mechanical aspects of any polymer, such as relaxation and terminal behavior. For melts of strictly linear chains, G_N^0, M_e, and τ_e are the only material parameters needed once molecular weight distribution is measured. Ultimately, this work helps enable the prediction of mechanical properties of isotropic conjugated polymers to support the design of stretchable and biocompatible electronics.

Future Plans

Future work will test the hypotheses that side chain structure can affect liquid crystallinity and that backbone nematic order reduces entanglements. We will also use web-based seminar courses and remote mentoring techniques to expose commonwealth campus students to research, and continue our Polymer Materials Design Scholars Program (PolyMDSP) to involve students in remote computational work during the year, in addition to summer research at UP.

Broader Impacts and Workforce Development

Through this project, we have trained 6 graduate students and 5 undergraduates (see participants). Students have been trained in diverse areas. All graduate students have seen an overview and examples of data from rheology, MD simulations, SCFT calculations, dynamic/static light scattering, X-ray scattering, polymer synthesis, and electrical characterization. Most students specialize in at least two of these areas in their own work.

The integration of computational and experimental activities continues to provide a unique opportunity for student training. All-hands meetings (students and PIs) take place once a month, where one long presentation is followed by two brief presentations, with an emphasis on synthesis, characterization and theory/simulation (long versus short presentations rotate).

Data Management and Open Access

We have established a database that shows the Tgs and densities of conjugated polymers, available from the project website listed above. It provides much needed information to aid in the design and simulation of this class of materials.

Advancing Along the Materials Development Continuum and Partnerships to Translation

Our efforts develop computational tools to accelerate the design process, and thereby maximize the yield of materials with acceptable properties. The key to our approach is to leverage our ability to predict fundamental polymer properties, such as the persistence length, nematic-to-isotropic transition temperature and entanglement molecular weight in the isotropic phase, to develop computationally-efficient tools to predict phase diagrams through coarse-grained models, entanglements in liquid crystalline phases, and band structure through tight-binding descriptions.

Publications and References

Xie, Renxuan, Albree R. Weisen, Youngmin Lee, Melissa A. Aplan, Abigail M. Fenton, Ashley E. Masucci, Fabian Kempe et al. "Glass transition temperature from the chemical structure of conjugated polymers." *Nature communications* 11, no. 1 (2020): 1-8.

Fair, Ryan A., Renxuan Xie, Youngmin Lee, Ralph H. Colby, and Enrique D. Gomez. "Molecular Weight Characterization of Conjugated Polymers Through Gel Permeation Chromatography and Static Light Scattering." *ACS Applied Polymer Materials* 3, no. 9 (2021): 4572-4578.

Fenton, Abigail M., Renxuan Xie, Melissa P. Aplan, Youngmin Lee, Michael G. Gill, Ryan Fair, Fabian Kempe et al. "Predicting the plateau modulus from molecular parameters of conjugated polymers." *ACS Central Science* 8, no. 2 (2022): 268-274.

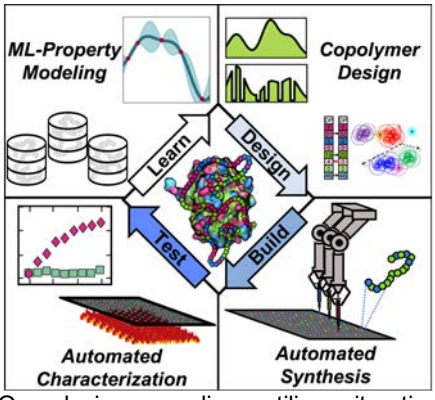
Zokaei, Sepideh, Donghyun Kim, Emmy Järsvall, Abigail M. Fenton, Albree R. Weisen, Sandra Hultmark, Phong H. Nguyen et al. "Tuning of the elastic modulus of a soft polythiophene through molecular doping." *Materials Horizons* 9, no. 1 (2022): 433-443.

Design of Stabilized Protein-Polymer Hybrids by Combinatorial Experimentation, Molecular Modeling, and Machine Learning

Lead Investigator: Adam Gormley, adam.gormley@rutgers.edu; Michael Webb, mawebb@princeton.edu
Participating Institutions: Rutgers University, Princeton University, Air Force Research Laboratory
Website: https://dmref.org/projects/1571
Keywords: enzyme, machine learning, single-enzyme nanoparticle, robotics, polymer-protein complex

Project Scope

Proteins are widely employed for diverse medicinal, industrial, and commercial applications. To increase durability in harsh environments, polymers are used as stabilizing components, but most only provide modest protection. We combine robotics and machine learning (ML) to discover new protein-polymer hybrids (PPHs). The hypothesis underlying this strategy is that copolymer chemistries can be tailored towards specific enzymes, enabling function beyond native conditions, and ML enables efficient, robust identification of relevant polymers within a combinatorial search space. Our long-term vision is to not only find polymers that effectively stabilize any target enzyme but deploy these PPHs as functional, assembled materials.



Our design paradigm utilizes iterative learn-design-build-test cycles to direct the design of stable protein-polymer hybrids.

Relevance to MGI

We utilize an iterative, closed-loop discovery process wherein ML models are trained to predict prospective experiments from existing data (learn), optimization proffers new PPHs (design), suggested PPHs are experimentally realized using robotics (build), and the resultant PPHs are characterized in high-throughput manner using automated assays (test). This data-driven discovery paradigm relies on seamless integration of experimental data into computational workflows that direct future experiments. Identifying high-performing systems provides new opportunities to understand their function with more extensive experimental characterization and molecular modeling. These detailed studies can then help refine our ML-guided design process for future discovery.

Technical Progress

We have established baseline protocols and workflows that provide proof-of-principle that our closed-loop discovery process will yield tailored PPHs. Given experimental bandwidth for polymer synthesis in 96-well plates, we have tested four distinct enzymes: chondroitinase ABC (ChABC),[1] horseradish peroxidase (HRP),[2] glucose oxidase (GOx),[2] and lipase (Lip).[2] In all cases, our design goal was maximizing retained enzyme activity (REA) following thermal stressing; for reference, proteins without any polymers exhibit REA = 0% post thermal stress and REA < 100% indicates fractional activity compared to native enzyme function. Ultimately, our data-driven discovery process yielded tailored copolymers that led to significant REA for each enzyme. The performance of these designed PPHs was particularly notable compared to more traditional brute-force, systematic chemical screens. The top-performing PPHs exhibited remarkable REA values of 141% (ChABC), 93.1% (HRP), 67.4% (GOx), and 107.9% (Lip). These numbers highlight the intriguing potential to enhance activity over native function but also raise questions about prospective limits on performance for any given protein. Cross-evaluation studies confirmed that designed copolymers uniquely stabilized the target enzyme[2]—proving our central hypothesis.

Initially designed thermostabilized PPHs have also accelerated applications and advanced scientific knowledge. Our top designed PPH for ChABC, which has potential therapeutic application for neural regeneration, was further evaluated for its long-term stability in artificial cerebrospinal fluid. This PPH exhibited measurable activity past one week time whereas ChABC in sugar solutions lost all activity in one day.[1] Meanwhile, biophysical characterization of our top-performing PPH for HRP revealed copolymer-assisted refolding of HRP as the likely mechanism of stabilization,[2] contrasting with the conventional idea of stabilization by surface-specific interactions.

In other work, we have devised and benchmarked various polymer featurization strategies that are appropriate for describing stochastic and/or sequence-defined copolymers;[3] the former is the primary design platform for this DMREF project. A key result is that composition-based ML models can perform better than sequence-based ML models, particularly if those representations are augmented with global polymer descriptors. This justifies the featurization employed in our proof-of-principle studies and provides a pathway forward for future enhancement.

Future Plans

In the next funding period, we will investigate sensitivity of designed PPHs, create assays and ML models for predicting stability against other stressors, and pursue the design of highly stable dehalogenases. Molecular modeling will be used to further interrogate the mechanism of stabilization in PPHs and provide feedback to ML models. Protein engineering will be used to promote enzyme activity and facilitate ML model generalizability; our goal in this direction will be to construct a ML model that uses both polymer and protein features as inputs. We will also benchmark alternative data selection and featurization strategies with implications on overall design efficiency.

Broader Impacts and Workforce Development

The project has supported in full or part the training of eight graduate students or postdoctoral researchers who regularly communicate and provide team updates. These trainees are firmly ingrained in all strategy meetings and are thus aware of the overarching philosophy and goals of the DMREF project. Rutgers is hosting multiple REU students in the upcoming summer who will be involved with the project, and Princeton is hosting two high school students with projects related to polymer machine learning, which is an integral part of our design. PI Webb also recently instructed a three-day workshop related to machine learning at the ACS middle Atlantic regional meeting (https://webbtheosim.github.io/ml-workshop-acsmarm2022/) with educational modules on Gaussian Process Regression and active learning, which are significantly leveraged in our materials design paradigm.

Data Management and Open Access

The DMREF team is fully committed to making key research outputs publicly accessible and useful to the community. Therefore, datasets generated from supported research have been published to DataSpace with registered DOIs for long-term preservation and access (10.34770/h938-nn26, 10.34770/chzn-mj42). The data is also hosted with accompanying code demonstrating the training and implementation of machine learning models on GitHub (github.com/webbtheosim/PPH_public, github.com/webbtheosim/featurization). These datasets will be further disseminated with Jupyter notebooks for publishing and further searchability on Materials Data Facility.

Advancing Along the Materials Development Continuum and Partnerships to Translation

Proof-of-principle studies demonstrated rapid and robust identification of efficacious PPHs by evaluating less than 0.1% of our current design space, which should dramatically accelerate development of future PPH-based materials. Early success demonstrated for PPHs of ChABC during *in vitro* studies led to Rutgers and Princeton pursuing a patent application on such formulations; PIs Gormley and Webb have applied for follow-on funding for further investigation. Gormley also founded Plexmer, Inc., which aims to develop polymer excipients for protein products. A major challenge will be connecting lab stability of PPHs to less controlled environmental conditions.

Publications and References

1. S. Kosuri, C. H. Borca, H. Mugnier, M. J. Tamasi, R. A. Patel, I. Perez, Z. Finkel, R. Schloss, L. Cai, M. L. Yarmush, M. A. Webb*, and A. J. Gormley*, *Machine-assisted Discovery of Chondroitinase ABC Complexes Towards Sustained Neural Regeneration*, *Advanced Healthcare Materials* **11,** 2102101 ***(2022)***.
2. M. J. Tamasi, R. A. Patel, C. H. Borca, S. Kosuri, H. Mugnier, R. Upadhya, N. S. Murthy, M. A. Webb*, and A. J. Gormley*, *Machine Learning on a Robotic Platform for the Design of Polymer-Protein Hybrids,* Advanced Materials, 2201809 *(2022)*
3. R. A. Patel, C. H. Borca, M. A. Webb*, *Featurization strategies for polymer sequence or composition design by machine learning,* Molecular Systems Design & Engineering **7**, 661-676 *(2022)*

Achieving Multicomponent Active Materials through Synergistic Combinatorial, Informatics-enabled Materials Discovery

Lead Investigator: Martha Grover, martha.grover@chbe.gatech.edu
Participating Institutions: Georgia Institute of Technology, Lehigh University, University of Kentucky
Website: none
Keywords: Organic Semiconductors, Conjugated Polymers, Transistors, Classification, High-Throughput

Project Scope

The objective of this project is to accelerate the optimization and insertion of multicomponent polymer blends into large-area flexible organic electronics. A tight feedback loop between simulations, experimental databases (including literature), and high-throughput gradient film experiments is accomplished with an informatics workflow, to traverse the composition and processing parameter space of these active materials. Electronic and mechanical properties are optimized while maintaining easy of processability and long-term stability.

Relevance to MGI

A foremost challenge in applying widespread, data-driven approaches to accelerate development of polymer semiconductors is the availability of experimental data that can yield the requisite knowledge necessary to inform robust performance and formulation precision at the manufacturing scale. Laboratory data is a necessary component in providing a closed loop of simulation, experimentation, and sequential design. Thus far, this project has successfully demonstrated the use of custom "small data analytics" techniques to identify solution concentration as a key performance indicator for high performance OFET devices fabricated from the p-channel donor acceptor polymer DPP-DTT [1]. To build on these findings, we are integrating simulations and high-throughput experimentation as techniques to uncover how this knowledge translates across the materials space (i.e., other organic semiconductors) and the process space (e.g., molecular weights, deposition process parameters, etc.). Additionally, we have provided the community with published viewpoints that draw attention to the need to interrogate the relevant organic semiconductor process-structure features in both solution and in thin film to provide for robust and shareable data, which is crucial in realizing the objectives of the Materials Genome Initiative [2].

Technical Progress

A dataset was initially constructed containing processing information and mobility data of OFETs from published literature for transistors fabricated using DPP-DTT [1]. A customized classification algorithm was applied to the dataset to identify the important processing parameters, to gain new insights and guide future experiments. The algorithm identified polymer solution concentration as a key factor in device performance and recommended a reduced design region (RDR) for this variable that would lead to improved hole mobility. Experiments performed to confirm the results from the classification analysis revealed a strong influence of solution concentration on the polymer chain excitonic interactions and electronic performance. As recommended by the classification algorithm, OFETs fabricated at the critical polymer chain overlap concentration (C*) resulted in improved hole mobility (Figure 1).

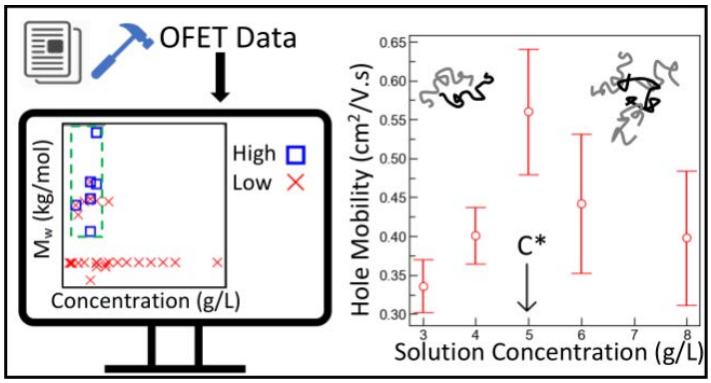


Figure 1. The figure represents the workflow adapted where a customized classification algorithm was applied to the DPP-DTT OFET dataset curated from literature. The algorithm identified a RDR (green box) for concentration and Mw that would lead to improved hole mobility. The blue squares represent devices with mobility > 1 cm2/V.s, while red crosses represent devices with mobility <= 1 cm2/V.s. Devices fabricated at this recommended design range showed improved hole mobility.

In tandem with experimental spectroscopic studies, density functional theory (DFT) and time-dependent DFT (TDDFT) calculations were performed on two families of DPP-DTT oligomers: (1) a series of single oligomers and (2) a series of bimolecular complexes of stacked trimers. These calculations reveal that charge delocalization between the two trimers widens the low-energy absorbance window due to an additional electronic

excitation. Natural transition orbital analysis was performed on select excited states, revealing considerable delocalization of the electron and hole and a small charge-transfer character to the excitation. Similar analyses were carried out for oxidized species, with low-energy absorption shifting into the near-IR region.

Our project team has also completed an initial demonstration of a custom-built gradient thin film deposition set-up for conjugated polymers [3]. The tool development involved (i) provisioning and validating an in-line herringbone microfluidic mixer, the first to our knowledge fabricated from steel and capable of operating at T ~ 130 °C to support solubilization, (ii) demonstrating the composition gradients in solution state and (iii) feeding gradient solutions into a distributor head/coating blade and demonstrating composition gradients in finished films. Our gradient films of poly(3-hexylthiophene) (P3HT) and polystyrene (PS) display morphological behavior that can be tied back to expected behavior of uniform composition films, which demonstrates the viability of our sample preparation technique to apply to other conjugated polymers of interest, including DPP-DTT.

Future Plans

We will study the influence of solution concentration on alternative classes of D-A conjugated polymers (e.g. N2200), as well as P3HT, to evaluate whether the critical chain overlap is a key parameter that universally defines polymer semiconductor performance. Additionally, the Mw and solvent will be varied to determine if the optima shifts with changes in these two variables. *Risko* will use atomistic MD simulations to determine the effects of chain molecular weight, solution concentration, and solvent choice on chain aggregation. These variables will be explored through a family of DPP-based donor–acceptor copolymers in chlorobenzene and chloroform. *Risko* has developed a procedure for preparing condensed-phase all-atom simulation systems with total system molecular weights in excess of 4500 kDa. This procedure has been applied to polymer families based on P3HT, DPP-4T, and benzodithiophene (BDT), each with experimentally relevant chain lengths (≈40 nm). By probing solution concentrations in line with those studied by *Reichmanis* and *Grover* and studying the solvated dynamics of individual long (>100 kDa) polymer chains, we seek to uncover the morphological basis for increased FET hole mobility at intermediate concentrations. Our custom gradient deposition tool generate libraries of P3HT and DPP-DTT films blended with PS that systematically explore the influence of (i) blend composition, (ii) solution pretreatment to induce aggregation, (iii) coating solution temperature, (iv) coating substrate temperature, (v) coating gap and rate on the polymer semiconductor performance and thin film microstructure (*Meredith*). In parallel, development of a high-throughput impedance screening tool will be used to measure hole transit times, which can be correlated to mobility, in a rapid, point-to-point fashion on libraries.

Broader Impacts and Workforce Development

Two graduate students supported by the project joined Georgia Tech's Graduate Certificate in Data Science for the Chemical Industries program, engaging in group projects with working professionals and facilitating their internships in materials informatics during the past year (Liu at BASF, Venkatesh at Dow). Undergraduate researchers also engaged in research on the team.

Data Management and Open Access

The DPP-DTT dataset along the Python code for the classification algorithm used to identify important processing variables for further investigation are available for download at:
https://github.com/rvenkatesh97/Reduced_Design_Region_Classification_DPPDTT_OFETs

Advancing Along the Materials Development Continuum and Partnerships to Translation

Georgia Tech partnered with Rieke Metals to help them understand issues of sensitivity and robustness in their P3HT products. Rieke Metals provided Georgia Tech with materials for performance characterization in transistors.

Publications and References

1. R. Venkatesh, Y. Zheng, C. Viersen, A. Liu, C. Silva, M. Grover, and E. Reichmanis, *ACS Materials Letters*, **3**(9), 1321–1327 (2021), DOI: 10.1021/acsmaterialslett.1c00320
2. C. P. Callaway, A. L. Liu, R. Venkatesh, Y. Zheng, M. Lee, J. C. Meredith, M. Grover, C. Risko, E. Reichmanis. *ACS Applied Materials & Interfaces*, **14**(3), 3613–3620(2022), DOI: 10.1021/acsami.1c20994
3. A. Liu, E. Dogan-Guner, M. McBride, R. Venkatesh, M. Gonzalez, E. Reichmanis, M. Grover, J. C. Meredith. Composition Gradient High-Throughput Polymer Libraries Enabled by Passive Mixing and Elevated Temperature Operability. *Chemistry of Materials* (2022), *submitted*.

Interatomic Potentials Repository and Tools

Lead Investigator: Lucas Hale, lucas.hale@nist.gov
Participating Institutions: NIST
Website: https://www.ctcms.nist.gov/potentials/
Keywords: Atomistic simulations, interatomic potentials, crystal defects.

Project Scope

The NIST Interatomic Potentials Repository aims to provide and develop novel computational infrastructure in support of the design, execution, and analysis of classical atomistic calculations for investigations of crystalline and crystal defect properties. The infrastructure consists of a database of interatomic potentials, open-source tools for designing, executing, and analyzing atomistic calculations, and reproducible property calculation methods that can be performed in high throughput. These resources make it easier for users to select an existing interatomic potential for their interests and perform standard property evaluations. Success is measured by the number of tool users, hosted potentials, and implemented calculation methods.

Relevance to MGI

Atomistic calculations using classical interatomic potentials provide a means of exploring how atomic-level structures influence larger scale dynamic properties and processes. The nature of classical interatomic potential models is such that predictions across different potentials of the same element are often substantially different. This makes the ability to evaluate and compare how different potentials perform for a set of standardized property tests important for users to determine which, if any, interatomic potentials will provide realistic predictions of their interests. Alternatively, considering each potential as a virtual element allows for the high throughput calculations to provide a substantial number of data points to better evaluate theorical models of structure-property relationships. The standard reproducible calculation methods also reduce the time and effort for developing and evaluating new interatomic potentials and emphasize the properties that are of interest to the community.

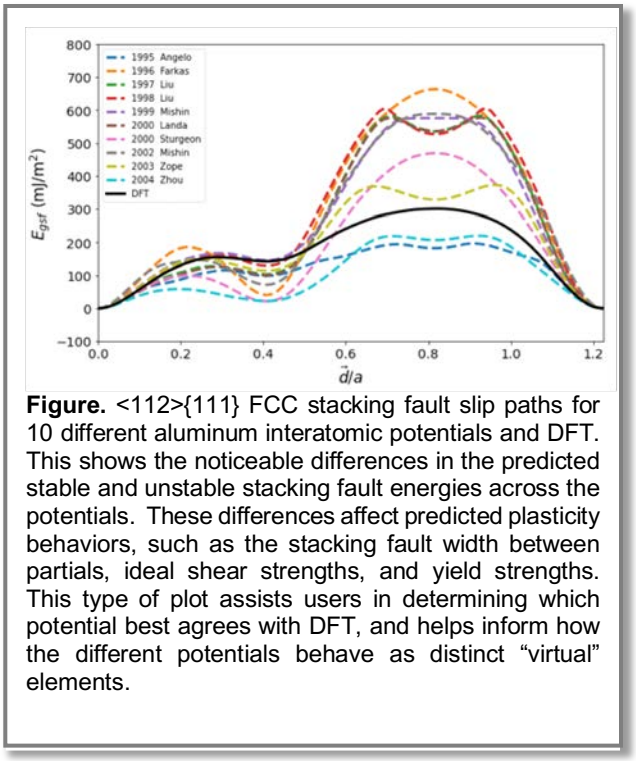


Figure. <112>{111} FCC stacking fault slip paths for 10 different aluminum interatomic potentials and DFT. This shows the noticeable differences in the predicted stable and unstable stacking fault energies across the potentials. These differences affect predicted plasticity behaviors, such as the stacking fault width between partials, ideal shear strengths, and yield strengths. This type of plot assists users in determining which potential best agrees with DFT, and helps inform how the different potentials behave as distinct "virtual" elements.

Technical Progress

Recent progress for the project has focused on improving the accessibility of the hosted data and tools. A major component of this was the development of a user-friendly API for accessing the data stored in the Repository. This API is available as the potentials Python package and makes it easy for users to search, explore, and download interatomic potentials from the repository. It also includes tools supporting the development of new interatomic potentials and allows for similar treatment of both native LAMMPS potentials and models from the OpenKIM project.

The underlying tools used by the potentials package for interacting with the repository and other databases was separated out to form the "yabadaba" Python package. The yabadaba package provides a mid-level abstraction layer allowing for entries to be accessed from multiple types of databases using a set of common interaction methods. Support is also provided for defining Record objects that interpret database entries of specific schemas into user-friendly data and Python objects. The result is that yabadaba makes it easy for generators of database content to create their own user-friendly Python API packages.

Work was also done in making the collection of reproducible calculations more accessible through the iprPy Python package. Users can easily install iprPy using pip or conda-forge, and then access the calculation methods either from a Python environment or solely through text files and command line arguments. This makes it easy for novice users to set up and perform any of the included calculations, and more advanced users to easily integrate any of the calculations into external workflows.

Future Plans

Short-term plans are focused on adding and performing more property calculations, notably dynamic crystal properties and dislocation core structures. These calculations are more computationally involved than the previous calculations but are important for characterizing how the interatomic potentials behave during complex simulations. The dynamic calculations reveal the stability and nature of different crystal phases as a function of temperature. The dislocation core structures have a direct influence on the plasticity mechanisms that the potential will predict.

Long-term plans are for increased connectivity with other projects, including better support for external atomistic tools and the integration of the developed calculation methods into alternative frameworks. This will make it easier for collaborators to develop new calculation methods and will help elevate the developed methods to be considered common standard tests.

Broader Impacts and Workforce Development

The repository and associated tools are all designed to make it easier for novice and expert users to perform and analyze atomistic calculations. Additionally, the potentials and yabadaba packages provide an example of how user-friendly APIs can be developed centered around different databases in which the end users can easily access the data without needing to know the database infrastructure specifics.

Data Management and Open Access

All data and the source code for the supporting Python packages are available from open-source locations
- Interatomic Potentials Repository: https://www.ctcms.nist.gov/potentials
- CDCS database with all Repository data: https://www.potentials.nist.gov
- cdcs Python API package: https://github.com/usnistgov/pycdcs
- yabadaba Python package: https://github.com/usnistgov/yabadaba
- potentials Python package: https://github.com/usnistgov/potentials
- atomman Python package: https://github.com/usnistgov/atomman
- iprPy Python package: https://github.com/usnistgov/iprPy

The Python packages are all also available through pip and conda-forge.

Advancing Along the Materials Development Continuum and Partnerships to Translation

There is an ongoing collaboration with the MolSSI SEAMM project https://molssi-seamm.github.io, which aims to create a graphical user interface to setting up atomistic calculation workflows for industrial use. The intent is to provide the tools and resources so that experts can easily perform these calculations alongside experiments. Workflow managers such as SEAMM rely on interfacing with modular components that evaluate a property of interest. The reproducible calculation methods collected in iprPy can be easily incorporated into SEAMM and other frameworks, thereby making the complex calculation methods accessible to non-experts.

Publications and References

1. C.A. Becker, F. Tavazza, Z.T. Trautt, and R.A. Buarque de Macedoc, *Considerations for choosing and using force fields and interatomic potentials in materials science and engineering*, Current Opinion in Solid State and Materials Science **17**, 277-283 (2013).
2. L.M. Hale, Z.T. Trautt, and C.A. Becker, *Evaluating variability with atomistic simulations: the effect of potential and calculation methodology on the modeling of lattice and elastic constants*, Modelling and Simulation in Materials Science and Engineering **26**, 055003 (2018).

A Computationally-Driven Predictive Framework For Stabilizing Viral Therapies

Lead Investigator: Caryn L. Heldt, heldt@mtu.edu (in person) and Sapna Sarupria, sarupria@umn.edu (virtual
Participating Institutions: Michigan Technological University, University of Minnesota, University of Massachusetts Amhurst
Website: none
Keywords: vaccine, excipients, glycol amino acids, active learning

Project Scope

Viral vaccines need a cold chain, which requires the vaccine to be cold throughout its distribution cycle. Excipients are often added to the formulation to keep the vaccine stable, and it has been hypothesized that many of these excipients stabilize viral vaccines by structuring water around of the virus. However, work in this area has been empirical, making it difficult to elucidate the subtle ways in which molecular structure affects water structure.

We hypothesize that experiments, modeling, and machine learning will identify molecular features/motifs that impart stability and use this framework to discover excipient mixtures for vaccine formulations.

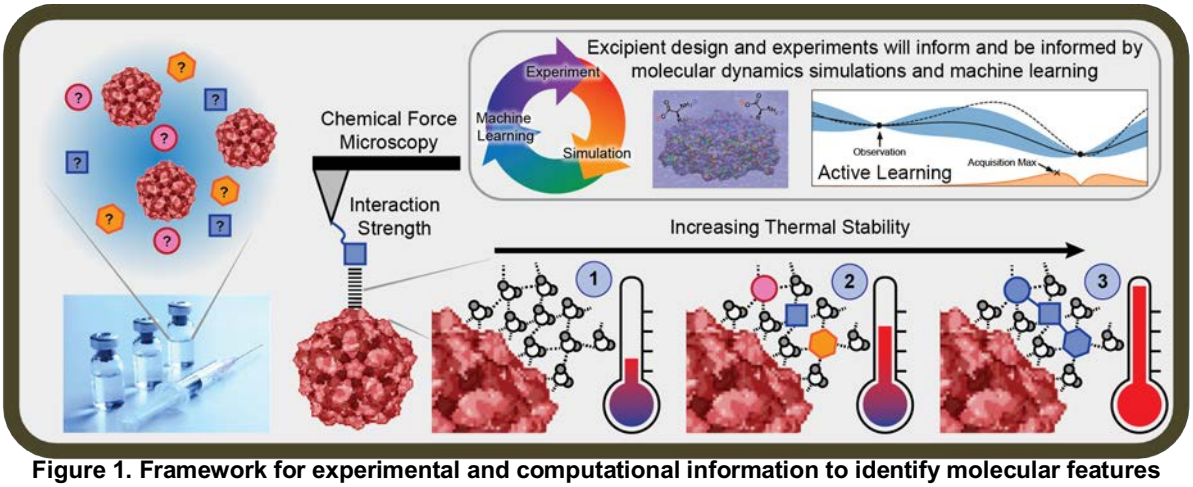


Figure 1. Framework for experimental and computational information to identify molecular features that impart stability.

Relevance to MGI

Understanding the subtleties of how the molecular structure of excipients affects water structure, and thus modulates the stability of biomolecules such as proteins and viruses remains a grand challenge that has implications for both the biophysical and biochemical communities. We aim to develop a framework, using both experimental and computational information, to identify molecular features imparting stability and discover complex excipient mixtures for increased vaccine thermal stabilization.

We will first study individual excipients experimentally and determine their ability to stabilize viral capsids against temperature denaturation. We will use common experimental techniques, like FTIR, DSC, and ITC. We will also use novel techniques, like chemical force microscopy (CFM) which is an atomic force microscopy (AFM) based method to measure how different chemistries and chemicals interact with a surface. The date will be explored with MD simulations to explore capsid-excipient interactions and then a ML algorithm will be tested to explore novel excipients. This will lead to complex designs of glycol amino acids and other excipients that will be synthesized and tested.

Technical Progress

We have begun to look at the interaction of long-chain linear alkane polymer (both coarse-grained [CG] and all-atom) to study the changes in the key features such as polymer folding and water structure affected by the presence of excipients. Currently for the CG model, we are estimating parameters which can clearly define both folded and unfolded states. In addition, we are also studying the excipient behavior in water to capture the individual contribution of the excipient to the water structure. We are now studying arginine as an excipient, as the team has previous data on the interaction of arginine with virus, but the overall goals are same irrespective of the choice of excipient.

Future Plans

Aim 1: We will identify molecular features/motifs present in excipients that impart stability to virus formulations. We will study experimentally and computationally how excipients interact with water and how they interact with virus. This is an important basis to understand the importance of water and virus structure on stability. Then we will proceed to Aim 2: Harness machine learning to understand and predict the performance of excipient oligomers and mixtures. The iterative process of experiments and molecular dynamics and machine learning will provide an unprecedented understanding of excipients that will lead to faster development of formulations and formulations that could help eliminate the cold chain.

Broader Impacts and Workforce Development

Community outreach will focus on using 3D printed viruses to demonstrate different aspects of viruses and vaccines. The first project will connect kinetics and thermodynamics to capsid assembly. Each PI will use local high school outreach programs as a platform to describe the importance of chemical engineering and chemistry to the stability of viral vaccines. The viruses have been printed and one demonstration has been completed. Future demonstrations are planned for each PI.

Data Management and Open Access

As we proceed, data and code will be opening available on our DMREF website.

Advancing Along the Materials Development Continuum and Partnerships to Translation

The goal of this project is to provide a computational framework to increase the discovery of excipients for viral vaccines. This framework will be available for industrial partners to use to help narrow their experimental space and increase the stability and time to market in the area of formulation. As we get closer to developing the framework, we will reach out to industrial partners to explore their interest in the tool.

Publications and References

None

Machine Learning Algorithm Prediction and Synthesis of Next Generation Superhard Functional Materials

Lead Investigator: Russell J. Hemley, rhemley@uic.edu; Sara Kadkhodaei, sarakad@uic.edu; Ravhi Kumar, ravhi@uic.edu; Michael Trenary, mtrenary@uic.edu; Eva Zurek, ezurek@buffalo.edu; Gary Cook, gary.cook.9@us.af.mil
Participating Institutions: University of Illinois Chicago, University at Buffalo, AFRL/RYDH
Website: https://dmref.org/projects/1580
Keywords: High pressure, diamond-anvil cells, surface science, machine learning, density-functional theory

Project Scope

The goal of this project is to combine newly developed crystal structure prediction techniques with experiments towards the discovery and synthesis of novel superhard materials, including those with additional functionality. The phases are predicted using an evolutionary algorithm that employs machine learning (ML) models to estimate a structure's hardness, and energies/enthalpies calculated with density functional theory (DFT). DFT calculations are employed to compute spectroscopic signatures, and properties of the predicted phases, and ways in which they may be synthesized. Advanced ML techniques guide the experiments by targeting precursor materials and synthesis pathways (Figure 1). Experiments involve high-pressure and surface science methods. Students are being trained in an interdisciplinary collaborative team of theoreticians and experimentalists with expertise in chemistry, physics, materials science and engineering.

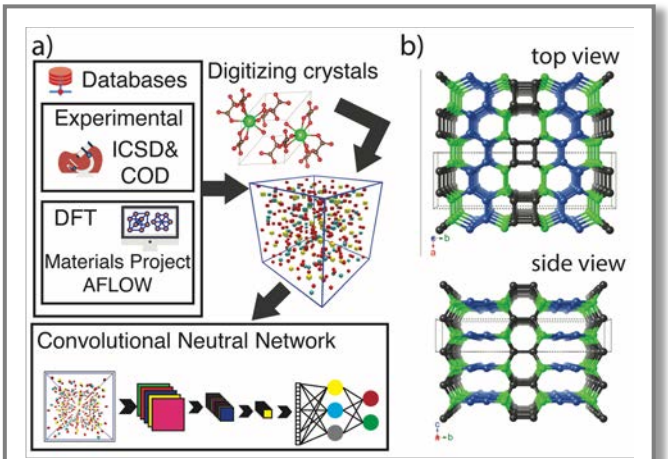


Figure 1. a) An overview of the ML model for high-throughput synthesizability prediction. b) The structure of superhard, superconducting 'DAC-carbon' predicted theoretically as an synthesis target.[1]

Relevance to MGI

Consistent with the MGI strategic plan, the project integrates experiment, computation, and theory in an iterative feedback loop among project tasks. The project is accelerating materials discovery and development of next generation functional superhard materials, as demonstrated by results from the past year, specifically in carbon- and boron-containing materials, including collaboration with other DMREF and MIP teams.

Technical Progress

The evolutionary algorithm XtalOpt developed in the Zurek group together with an ML model was used to predict novel, multifunctional carbon polymorphs. One such phase has mixed sp^2/sp^3 carbons and a Vickers hardness of 48 GPa and acts as a microscopic diamond anvil cell (DAC) by constraining the distance between sp^2 carbons in parallel cis-polyacetylene chains.[1] This phase is metallic with an estimated superconducting critical temperature similar to that of twisted bilayer graphene (1.6 K). DAC-carbon could be synthesized by on-surface polymerization of graphene nanoribbons, followed by pressurization of the resulting 2D sheets. Kadkhodaei's group developed an ML model for synthesizability prediction across different crystal structures and chemistries[2] using a novel representation of crystals by 3D voxel images. Visual features embedded in these images are learned by convolutional neural networks to predict the synthesizability of never-seen-before crystalline materials.

Experiments included a collaboration with the University of Florida (UF) DMREF team on high-pressure electrical resistivity and x-ray diffraction of WB_2 to 190 GPa, leading to the discovery superconductivity above 50 GPa with T_c reaching 17 K at 90 GPa.[3] High-pressure experiments provided new constraints on the P-V-T equations of state and possible phase transition in B_4C predicted by Zurek's group. X-ray Raman measurements were also

conducted to identify bonding changes in B_4C, BAs, and other boron-containing materials under pressure. Trenary and his students are exploring the surface chemistry associated with the growth of metal diboride thin films by chemical vapor deposition (CVD) with exceptionally high melting points, high hardness, and metallic electrical conductivity. The group synthesized $Zr(BH_4)_4$ which was used to create a hard ZrB_2 film on a Pd(111) surface.

Future Plans

The Zurek's XtalOpt evolutionary algorithm will be used to predict new stable, superhard carbon-rich materials with stoichiometries such as BC_xN (x = 2,4,6,8), including calculation of properties and possible synthesis pathways. Kadkhodaei's group ML model will be expanded to predict formation energy and high-pressure synthesizability. Trenary's group will examine addition substrates for the formation of boride films. The Trenary and Zurek groups are collaborating on joint experimental/theoretical studies of diboride films. The experiments will also involve completing studies of B_4C (with samples from PARADIM, the JHU MIP). BAs and other thermal conductors of interest to AFRL, synthesis of the predicted "DAC-carbon", and searching for new ultrahard nitrides under pressure.

Broader Impacts and Workforce Development

Zurek gave a 4hr hands-on workshop on using the XtalOpt evolutionary algorithm to at the "High Pressure International School of Crystallography" in Erice, Italy, a 45 minute educational talk on computational crystal structure prediction, and 3 virtual seminars at PUIs on materials design. The high-throughput predictive capability for synthesis likelihood will impact materials discovery in various technological applications. Altogether, four graduate students, two undergraduates, and two postdocs are currently involved in the project and are being trained for workforce development.

Data Management and Open Access

The XtalOpt evolutionary algorithm, used for the crystal structure searches of superhard materials, is published under the BSD or GNU open-source licenses. The source code and compiled executables can be found on the XtalOpt website (https://xtalopt.github.io) and in the Computer Physics Communications Library (https://cpc/cs/qub.ac.uk). The outputs of the computations have been uploaded to the NOMAD repository (https://nomad-lab.edu). Kadkhodaei's lab has published the synthesizability prediction code on Github (https://github.com/kadkhodaei-research-group/XIE-SPP) under an open-access license to contribute to advancing scientific cyberinfrastructure. All experiments data are made available to the community as per NSF guidelines.

Advancing Along the Materials Development Continuum and Partnerships to Translation

The developed DFT/AI/ML models for synthesis and property predictions are providing new means to guide the discovery and synthesis of new functional superhard materials. The effort is demonstrating that a close coupling between simulation and experiment embodied in the MGI philosophy is essential. Our team is also partnering with the UF DMREF team, resulting in a publication,[3] and is collaborating with the JHU MIP for precursor boron-containing materials. The prospects for commercialization of various discoveries are being explored.

Publications and References

1. X. Wang, D. Proserpio, C. Oses, C. Toher, S. Curtarolo, and E. Zurek, *The microscopic diamond anvil cell: Stabilization of superhard, superconducting carbon allotropes at ambient pressure*, Angew. Chem. Int. Ed., in press, doi:10.1002/anie.202205129.
2. A. Davariashtiyani, Z. Kadkhodaie, and S. Kadkhodaei, *Predicting synthesizability of crystalline materials via deep learning.* Comm. Mater. **2**, 115. doi.org/10.1038/s43246-021-00219-x.
3. J. Lim, A. C. Hire, Y. Quan, J. S. Kim, S. R. Xie, R. S. Kumar, D. Popov, C. Park, R. J. Hemley, J. J. Hamlin, R. G. Hennig, P. J. Hirschfeld, and G. R. Stewart, *Creating superconductivity in WB_2 through pressure-induced metastable planar defects*, arXiv: 2109.11521, *Nature Comm*., submitted.

DMREF AI-Accelerated Design of Synthesis Routes for Metastable Materials

Lead Investigator: Richard G. Hennig, rhennig@ufl.edu
Participating Institutions: University of Florida
Website: https://super.mse.ufl.edu
Keywords: Metastable materials synthesis routes, machine learning, ultra-fast force fields, superconductivity, magnetism, hardness

Project Scope

The project will address the science of nonequilibrium processes by searching for and identifying the design rules for synthesis pathways of metastable materials. We will combine computational methods, including AI, with experimental approaches utilizing pressure, temperature, and magnetic fields to transform amorphous precursor materials into desired phases that are kinetically stable at ambient conditions and persist indefinitely at ambient conditions. We will build the fundamental knowledge base to discover and design materials that exhibit unusual compositions and structures, are inaccessible to synthesis processes at ambient pressure and provide desirable properties and functionality for applications such as superconductivity, magnetism, and superhardness.

Relevance to MGI

Our design approach is based on a closed loop combining AI for structure prediction and ab-initio methods for calculating properties with high pressure, magnetic field, and temperature synthesis and characterization. The work builds on proven approaches to synthesizing metastable materials by crystallizing amorphous precursor materials. We will develop computational methods that employ the ultra-fast force field machine learning of energy landscapes to model the finite temperature and pressure thermodynamics and kinetic stability of both crystalline and amorphous materials.

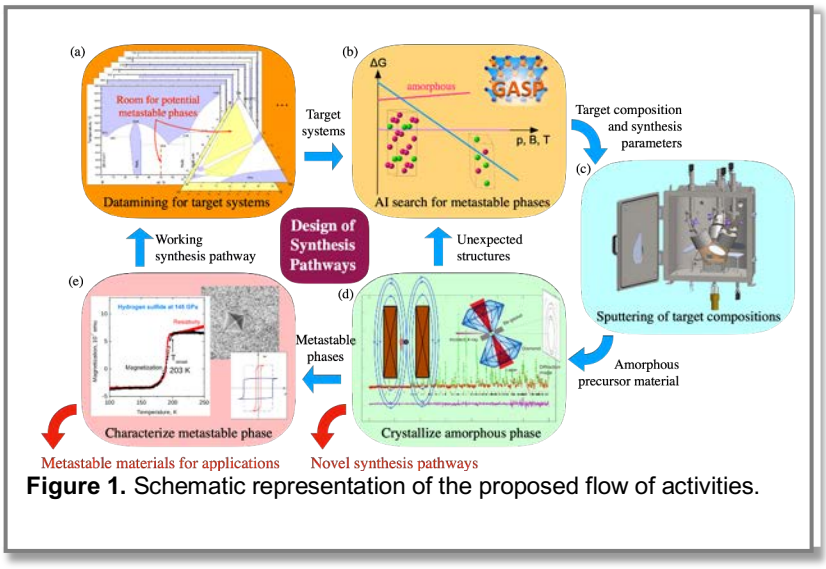


Figure 1. Schematic representation of the proposed flow of activities.

The design process will identify experimental conditions that enable the kinetic stabilization of desired high-pressure and temperature phases at ambient conditions in metastable form. Experimental synthesis and characterization will validate and guide the approach, leading to new synthesis pathways. Products will include new metastable synthesis methods, machine learning approaches for structure prediction and characterization of their thermodynamics and kinetics, and metastable materials with a wide variety of desirable properties such as superconductivity, magnetism, and superhardness.

Technical Progress

During the first nine months of the project:

1. We extended the ultra-fast force field (UF^3) method for energy landscapes.
2. We designed and assembled a tool for rapidly quenching liquids to synthesize amorphous materials.
3. We refined and tested the high-magnetic field furnace to anneal amorphous materials in high fields.
4. We applied an alloying approach to synthesize metastable superconducting MoB_2 at ambient pressure.

AI methodology. We extended the UF^3 machine learning method from two- and three-body interactions to include the effect of four-body interactions. We automated the hyperparameter selection to optimize the model complexity and simplify the parameter optimization. We also started coupling the UF^3 method to our genetic algorithm code GASP to accelerate the composition and structure predictions of metastable multinary compounds.

Synthesis of amorphous precursors. We designed and machined an anvil and hammer splat cooler to rapidly quench liquid samples. Applying a vacuum rapidly transfers the molten sample from a high-purity argon atmosphere onto a cooled copper anvil where a hammerhead hits the sample, rapidly increasing the sample's surface area while pressing it between two opposing pieces of water-cooled copper, thereby quick quenching the sample. We tested the equipment for materials with a tendency to form amorphous phases during rapid cooling. Fig. 2 shows the splat cooler.

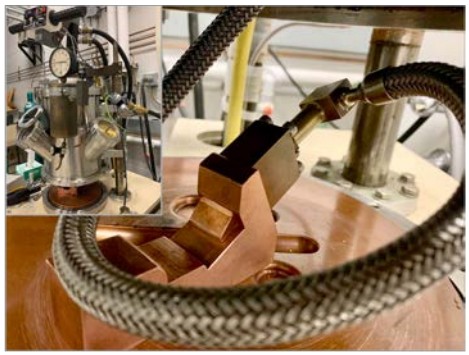


Figure 2. Anvil and hammer splat cooler to quench a liquid into an amorphous phase.

High-magnetic field furnace. We tested the capability of annealing samples in high magnetic fields of 9 T temperatures up to 1500 K. We expect that we can use higher fields up to 45 T available for shorter annealing up to 10 hours.

Alloying approach to extend pressure range of metastable defects. We had previously shown that the high-pressure synthesis of metastable stacking fault and twinning defects during compression of WB_2 is the likely mechanism for superconductivity. We extended our study to MoB_2 and showed that alloying sufficiently increases the stability of the metastable, superconducting AlB_2 phase at ambient pressure.

Future Plans

We will accelerate the genetic algorithm structure prediction with the UF^3 machine learning approach trained on density-functional relaxations and apply the method to various transition metal borides to identify metastable materials with the possibility of superconductivity or high hardness. In parallel, we will synthesize and anneal amorphous precursors in these systems to tune the experimental approach and validate our predictions. After this validation, we will extend the computational approach to include Monte Carlo and molecular dynamics simulations using the UF^3 model to predict the stability of stable and metastable crystalline and amorphous phases as a function of temperature, composition, and pressure. The results will be used to identify further promising materials systems and to help explain experimental observations.

Broader Impacts and Workforce Development

PI Hennig has developed and taught a new graduate course on *AI for Materials*. The next step is to co-teach this class as an undergraduate elective class and integrate the course into the MSE undergraduate and graduate curriculum. In addition to the software tools for machine learning, PI Hennig this year developed Jupyter notebooks for teaching fundamental thermodynamics concepts. We freely provide the education material and Jupyter notebooks to other teachers, students, and the interested general public through our Github organization for Materials Education at https://github.com/matscied. We plan to expand this effort and also provide additional teaching material. In the following years, we also will develop an experiment kit for K-12 use that demonstrates synthesis methods and make this kit available to teachers across the southeastern states, Puerto Rico, and the US Virgin Islands. Through our partnership with the education program of Prof. Ruzycki, we will train teachers in its use and refine the module based on feedback.

Data Management and Open Access

We freely provide all software and documentation on our Github page https://github.com/henniggroup. Furthermore, all datasets are made available on sites such as https://www.materialscloud.org, and synthesis recipes are published through peer-reviewed manuscripts.

Publications and References

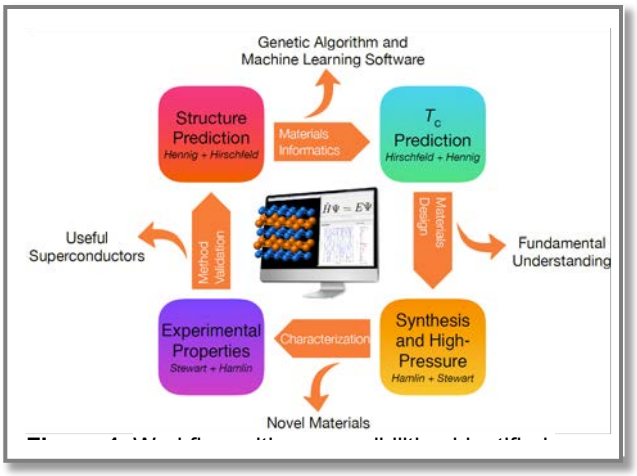
1. S. R. Xie, M. Rupp, R. G. Hennig, *Ultra-fast interpretable machine-learning potentials*, submitted to npj Computational Materials (2022), arXiv:2106.05235.
2. J. Lim, A. C. Hire, Y. Quan, J. S. Kim, S. R. Xie, S. Sinha, R. S. Kumar, D. Popov, C. Park, R. J. Hemley, Y. K. Vohra, J. J. Hamlin, R. G. Hennig, P. J. Hirschfeld, and G. R. Stewart, *Creating superconductivity in WB_2 through pressure-induced metastable planar defects*, submitted to Nat. Commun. (2022), https://arxiv.org/abs/2109.11521.
3. J. B. Gibson, A. C. Hire, and R. G. Hennig, *Data-Augmentation for Graph Neural Network Learning of the Relaxed Energies of Unrelaxed Structures*, submitted to npj Computational Materials (2022), https://arxiv.org/abs/2202.13947.
4. A. C. Hire, S. Sinha, J. Lim, J. S. Kim, L. Fanfarillo, P. M. Dee, J. J. Hamlin, R. G. Hennig, P. J. Hirschfeld, G. R. Stewart, *High Critical Field Superconductivity at Ambient Pressure in MoB_2 Stabilized in the P6/mmm Structure via Nb Substitution*, submitted to Phys. Rev. B (2022).

DSDCMS-Data Science Enabled Discovery of New Superconductors

Lead Investigator: Peter Hirschfeld, pjh@phys.ufl.edu
Participating Institutions: University of Florida
Website: https://super.mse.ufl.edu
Keywords: Superconductivity, materials discovery, machine learning, high-throughput

Project Scope

The aim of this project is to streamline the process of discovering new superconductors by developing algorithms that can learn the essential ingredients of a high-temperature electron-phonon superconductor and to use these algorithms to predict new superconductors, both at and above ambient pressure. These ingredients are to be calculated, using first-principles methods, based on descriptors relevant to T_c that can be used in a high-throughput database search, informed by physical principles known to be fundamental to superconductivity. The ultimate goal is a high-throughput scheme based on DFT-based descriptors for known superconducting and non-superconducting materials. Promising candidate materials will then be tested for superconductivity, under ambient conditions and under pressure, and subjected to more expensive *ab initio* electron-phonon and Eliashberg calculations of the superconducting properties, as well as of potential structural phase transformations under pressure using our genetic algorithm and machine-learning code GASP.



Relevance to MGI

For many years, theorists, even equipped with the tools of BCS and Migdal-Eliashberg (ME) theory, were unable to reliably predict new superconducting materials. The value of the critical temperature T_c depends on the electron-phonon coupling and the Coulomb pseudopotential via an essential singularity; thereforeT_c as a materials property is dependent in a very sensitive way on the details of the electronic structure and the competition of the electron-phonon interaction with the Coulomb interaction, such that quantitative calculations were considered impossible. Recently, advances in computational superconductivity using the tools of DFT and ME theory enabled accurate predictions of the superhydride high-pressure superconductors, spurring renewed interest in superconducting materials discovery through computation. This project aims to begin to put superconductor discovery on the same footing with other materials properties, such as thermoelectric coefficients, which proved susceptible to computational high-throughput calculation of a simple figure of merit. When new superconductors are predicted in this project, they are synthesized, characterized, and measured under pressure to gain new insights into the underpinnings of T_c, consistent with the MGI vision that the pace of new materials deployment could be accelerated through complementary efforts in theory, computation, and experiment.

Technical Progress

During the first 2.5 years of the project:

5. We applied symbolic regression to the calculation of T_c from the Eliashberg function α^2F, resulting in an improved "Allen-Dynes" formula that performs significantly better on superhydrides than the original [1].
6. We studied several binary and ternary beryllides under pressure [2].
7. We predicted several new diboride superconductors that would adopt the MgB_2 structure at moderate pressures, and discovered 17K superconductivity in WB_2 at 50 GPa.[3] .
8. We investigated kinetic barriers in superpressurized metastable Nb_3Si [4].
9. We investigated systematics of EPW[5] predictions for low-Tc superconductors and identified outliers that have T_c suppressed by spin fluctuations.

10. 11 Experiments performed at pressures greater than 100 GPa.
11. Co-edited and contributed to JPCM ``ROADMAP" issue on high-temperature superconductor discovery, 21 articles from the leading researchers in this rapidly advancing field [6].

AI methodology. Symbolic regression–based machine learning with SISSO code was used to discover a new Allen-Dynes type formula to predict T_c of strong-coupled superconductors. We extended the UF^3 ultra-fast force field method for energy landscapes to integrate with GASP structural relaxation algorithm.

High-throughput search for new hydrides. We have systematically searched for novel superconducting ternary hydrides, with two different strategies: i) focusing on structures in which a binary hydride high temperature superconductor has already been discovered, and replace atoms in existing hydrogen cages, and ii) partial atom substitution in stable ambient pressure hydrides to create new ternary hydrides.

Future Plans

DFT-based descriptors for high-throughput. In service of our long-term goals to construct a high-throughput scheme to discover new superconductors, we will identify using physical principles, and test, a series of DFT-based descriptors based on the chemical bonding, the electron localization functions and the charge density distributions. We aim to identify a set of DFT parameters that are either directly or indirectly related to w_{log} and λ which can be useful for future high throughput searches for high T_c superconductivity, utilizing machine-learning with analytic relations, random forest, and kernel-based methods to identify correlations between the electron-phonon interaction parameters and the DFT descriptors.

Spin fluctuation contributions to T_c. Eliashberg theory is quantitatively inaccurate for many systems with d- and f-electron states near the Fermi level, due to an approximate treatment of the Coulomb interaction, and in particular the neglect of spin fluctuations, which suppress conventional BCS pairing. Including such effects accurately in current codes for ab initio calculations of T_c is at the forefront of current research. For our proposed high-throughput approach, we do not need highly accurate calculations of T_c, but spin fluctuation suppressions of tens of percent can be important. We will include these and check improved accuracy of DFT+ME T_c calculations.

Broader Impacts and Workforce Development

PI Hennig has developed and taught a new graduate course on *AI for Materials*. The next step is to co-teach this class as an undergraduate elective class and integrate the course into the MSE undergraduate and graduate curriculum. We freely provide the education material and Jupyter notebooks to other teachers, students, and the interested general public through our Github organization for Materials Education at https://github.com/matscied. The project has helped train 4 graduate students, 5 undergrads, and 4 postdocs.

Data Management and Open Access

We freely provide all software and documentation on our Github page https://github.com/henniggroup. Furthermore, all dataset are made available on sites such as https://www.materialscloud.org and description of the experimental and computational methodology are published through the peer-reviewed manuscripts.

Publications and References

5. S. R. Xie, Y. Quan, A. C. Hire, B. Deng, J. DeStefano, I. Salinas, U. Shah, L.Fanfarillo, J. Lim, J. Kim, G. R. Stewart, J. J. Hamlin, P. J. Hirschfeld, and R. G. Hennig, *Machine learning of superconducting critical temperature from Eliashberg theory*, npj Computational Materials 8, 14 (2022).
6. L. Fanfarillo, J. Hamlin, R.G. Hennig, A. Hire, P. J. Hirschfeld, J. Kim, J. Lim, Y. Quan, G.R. Stewart, and S. Xie, Be_5Pt: *Properties of a novel pseudogapped semimetal,* Phys. Rev. B 102, 155206 (2020).
7. J. Lim, A. C. Hire, Y. Quan, J. S. Kim, S. R. Xie, S. Sinha, R. S. Kumar, D. Popov, C. Park, R. J. Hemley,4, 6 Y. K. Vohra, J. J. Hamlin, R. G. Hennig, P. J. Hirschfeld, and G. R. Stewart, *Creating superconductivity in WB_2 through pressure-induced metastable planar defects*, submitted to Nat. Commun. (2022), https://arxiv.org/abs/2109.11521.
8. J. Lim, J. Kim, A. Hire, Y. Quan, R. G. Hennig, P. J. Hirschfeld, J. J. Hamlin, and G. R. Stewart, *Probing the Origins of Superconductivity in Metastable A15 Nb_3Si,* J. Phys. Cond. Mat. 33,285705 (2021) .
9. S. Ponce, E. Margine, C. Verdi, and F. Giustino, *EPW: Electron–phonon coupling, transport and superconducting properties using maximally localized Wannier functions*, Computer Physics Communications 209, 116 (2016).
10. S. R. Xie, Y. Quan, A. Hire, L. Fanfarillo, G. R. Stewart, J. J. Hamlin, R. G. Hennig, and P. J. Hirschfeld, *The 2021 High-Temperature Superconductivity Roadmap*, J. Phys.: Condens. Matter 34183002 (2022).

GOALI: Salt Separation Membranes Based on Modifiable Two-Dimensional Covalent Organic Frameworks

Lead Investigator: John Hoberg, hoberg@uwyo.edu
Co-PI: Jonathan Brant, jbrant1@uwyo.edu
Co-PI: Laura Oliveira, Laura.deSousaOliveira@uwyo.edu
Co-PI: Bruce Parkinson, bparkin1@uwyo.edu
Co-PI: Zachary Gray, zack@wyonano.com
Participating Institutions: University of Wyoming and Wyonano
Website: none
Keywords: Covalent Organic Frameworks, nanofiltration, salt separation, membranes, computational screening

Project Scope

New materials based on covalent organic frameworks (COFs) with defined and modifiable pores make them naturally suited for semi-permeable membrane applications give their 2D nature. By putting charged functional groups in the pores and changing the pore sizes, ion selective membranes can reject anions and/or cations with a specific size threshold with a focus on desalinization. The team uses a feedback loop (Fig 1) between the synthesis of the membrane, membrane performance testing, and computational modeling with machine learning to guide the synthesis of the new membrane materials and optimize their performance for a particular separation.

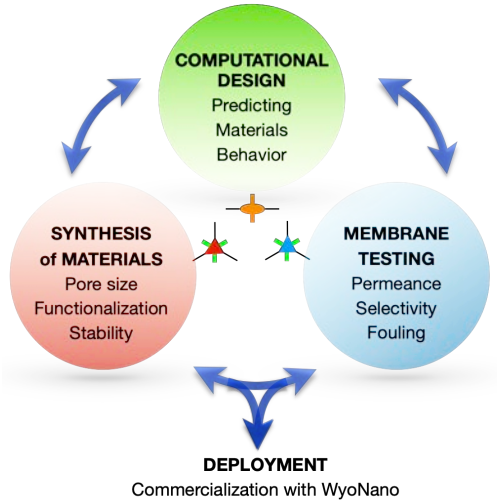


Figure 2. COF structures (top) and feedback loop.

Relevance to MGI

The research aligns with the goals of the DMREF Materials Genome Initiative to accelerate materials discovery and development by building the fundamental knowledge base needed to advance the design and development of COF materials with desirable properties and functionality. The modularity and corresponding extensive phase space of COFs makes them prime candidates for computer-aided, high-throughput, materials discovery and design. We use a machine learning-based approach to identify quantitative structure-property relationships (QSPR) to predict COF behavior. Through an automated computational screening of candidate frameworks, we will not just discover novel functional COFs, but learn how to reverse engineer frameworks for desired applications, starting with desalination. Our interwoven experimental and computational approach will guide the research towards solving some of the world's pressing issues related to clean water and energy.

Technical Progress

As this is a 2021 funded project, we have included the chart depicted in Figure 2 of the function-based design system that will be used to conceive, test and screen COF architectures that includes two independent levels of screening. Useful framework designs can be synthesized and consider empirical parameters, but in such a way as to not affect the set of QSPR functions that directly relate the descriptors to the COFs behavior and theoretical performance. Yellow and orange colors are used to indicate the computational and experimental dependence of the screening model, respectively. Full COFs and moieties will undergo the process of being attributed descriptors and evaluated for performance. Later, the direct link between the moieties and COFs will make use of moiety QSPR functions to intelligently produce frameworks with expected higher performance. Descriptor and performance metrics are shown as intersecting spheres to highlight the fluidity between them.

Future Plans

Computational studies will guide chemical synthesis and membrane fabrication through a collaborative feedback loop where new 2D COFs are either synthesized or predicted to be interesting by theory and then synthesized and tested for nanofiltration and separations.

Broader Impacts and Workforce Development

Since the start of this project in 2021, we have hired two post-doc research associates and two graduate students all of whom are mentors for four undergraduate STEM students who started in the first summer of the grant. Combined group meetings involve discussions and presentations of theorist, synthetic chemists and membrane engineers, which help to train the next generation in all aspects of this materials research. In each of these activities the MGI philosophies are accentuated to highlight the overlap and impact of both the theory and testing on guiding synthetic designs. Students involved in the project will be required to present their research to the outside community at science day presentations that are held at the end of the summer and again during the school year.

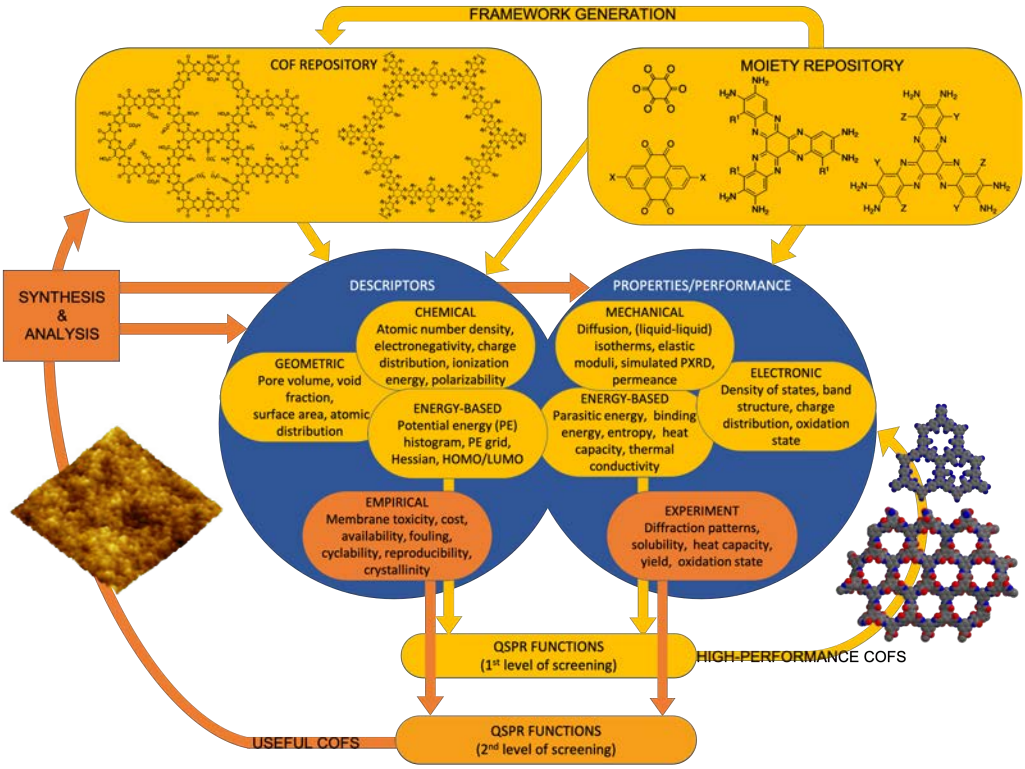


Figure 3.

Data Management and Open Access

We are committed to data management in accordance with the FAIR (findable, accessible, interoperable, and reusable) data principles. As such, our aim is to both use and contribute to existing databases. To date, we have performed calculations using COFs from the CoRE COF and CURATED COF databases, and modeled COFs synthesized by our team, which are not yet available in the literature or databases. The results from both calculations (e.g., band structures, charge distribution) and the geometries from the latter set of COF calculations, when complete and reported in the literature, will be included into existing databases, such as the CoRE COF database (for COF geometries) or the Novel Materials Discovery (NoMaD) database (for electronic structure properties), which is in keeping with the mission of the Materials Genome Initiative. Automation scripts for the machine learning component of the project have already been saved on GitHub but will also only be made public after the results obtained are published.

Advancing Along the Materials Development Continuum and Partnerships to Translation

The University of Wyoming has multiple patent applications on the technology that has been developed in the Hoberg lab. Along with the small local company, Wyonano, the PI's were collaborators on a recent NSF STTR Phase 1 grant focused on the development of membranes for salt separations. With this collaboration in hand, the prospect for commercialization is enhanced as several hurdles have already been overcome to make our technology viable in the commercial realm. This includes the purchase of scale-up equipment, a mind-set in the development of scalable processes and the development of a business model with Wyonano. Finally, a phase 2 STTR is planned in the future that should be enhanced with the inclusion of results from this research and results of MGI inclusion.

Publications and References

No publications have been submitted at this early stage but one student and two PIs have presented lectures at the MRS spring meeting acknowledging this DMREF grant.

AI Institute: Planning: Institute for AI-Enabled Materials Discovery, Design, and Synthesis

Lead Investigator: Vasant G. Honavar, vhonavar@psu.edu
Participating Institutions: Pennsylvania State University, University of Wisconsin, Massachusetts Institute of Technology, University at Buffalo, Northwestern University, Cornell University
Website: https://www.icds.psu.edu/institute-for-ai-enabled-materials-discovery-design-and-synthesis/
Keywords: Artificial Intelligence, Self-Driving Labs, Machine Learning, Potentials, Inverse Design, Optimization
Grant: DMR 2020243

Project Scope
Advances in materials design, discovery, and synthesis are increasingly by advances in instruments, data science and artificial intelligence methods, and computational modeling. The emergence of "big data", and advances in machine learning have dramatically accelerated some of the key steps in science, e.g., data acquisition and model fitting. However, other key elements of the scientific process, e.g., generating hypotheses, designing, prioritizing and executing experiments, integrating data, models, and simulations, drawing inferences and constructing explanations, reconciling scientific arguments, and communicating across disciplines, remain largely untouched by the advances in artificial intelligence (AI). Accelerating scientific progress, potentially by several orders of magnitude, by effectively addressing these bottlenecks presents a grand challenge for AI. This project brings together an interdisciplinary team of researchers with complementary expertise in AI and Material Science and Engineering (MSE) to launch a planning effort to lay the groundwork for an AI-Enabled Materials Discovery, Design, and Synthesis (AIMS) Institute.

Relevance to MGI
As described in the recently released MGI strategic plan, MGI requires an integration of experiment, computation, and theory through use of a closed, iterative feedback loop among project tasks to harness the power of data. We believe that realizing the MGI vision requires synergistic advances across multiple areas of AI, including: (i) knowledge representation frameworks for encoding, communicating, and reasoning with models or abstractions of scientific artifacts, e.g. data, experiments, hypotheses, models; (ii) optimizing scientific studies, experiments, etc.; (iii) machine learning and causal inference methods that can provide explanations of their results in the context of available knowledge, and recommend experiments to validate the predictions using the available experimental techniques); and (iv) algorithmic abstractions of AI-enabled human-machine, AI-enabled human-human, and machine-machine collaborations. Addressing this AI grand challenge requires fundamental advances across multiple areas of AI, and AI mediated human-machine systems that support collaborative team science. The AI advances would go hand-in-hand with use-inspired research driven by some of the most pressing MSE challenges. The planning effort has focused largely on (i) identifying opportunities for and catalyzing interdisciplinary research at the interface of AI and MSE; (ii) exploring scalable models for preparing diverse AI-savvy MSE students and professionals for diverse careers in academia, industry, national labs; (iii) identifying infrastructure needs; and (iv) identifying opportunities for partnerships across academia, national labs, and industry; and (v) developing partnerships for broadening participation of under-represented minorities and under-resourced institutions in research at the interface of AI and MSE.

Technical Progress
The deep interdisciplinary dialog between MSE and AI researchers initiated through workshops, seminars, and working groups has led to the realization that such interdisciplinary research may be organized into three broad categories, based on the types of problems considered: (i) Class 1 problems, e.g., prediction of materials properties from material composition, that are amenable to solution through careful application of existing state-of-the-art AI methods, where progress can be made by incorporating Data Sciences and AI methods into the training in MSE (ii) Class 2 problems e.g., training prediction on very small data sets with accurate domain and uncertainty quantification, which are readily formulated but will require significant modification of AI methods and close interaction between MSE and AI, and (ii) Class 3 problems that require deep sustained interdisciplinary communication. and collaboration between AI experts and material scientists and engineers to formulate MSE problems as. often novel AI problems, e.g., principled ways to infuse physical knowledge and constraints, e.g.,

symmetries, into machine learning algorithms. Some representative Class 2 problems include: Physics-guided learning of inter-atomic potential functions that can be applied to many components and atomic configurations without incurring impractically large computational or data requirements (a problem on which we have an exploratory project in progress). AI methods to accelerate materials simulations; AI methods for design and control of the microstructure of materials; AI methods for representing and traversing complex energy landscape. Some problems that begin to take on a Class 3 character include AI-guided design of materials that optimally meet the desired functional and non-functional requirements and AI-driven laboratories for materials discovery, design, and synthesis. Class 3 problems also include those which constitute many fundamental elements of the materials research scientific process, and still remain unarticulated as a computational problem and largely unimpacted by modern AI. The tackling of these constantly changing and difficult to define problems represents the major area of Class 3 problems. Another key observation is that addressing these MSE challenges will call for fundamental advances in several areas of AI, and the design of new AI tools for scientists that can: Create and share a knowledge base that summarizes what we know about a scientific question (annotated with uncertainty, provenance, and underlying assumptions); Summarize and prioritize questions that need to be answered to achieve an overall scientific objective (e.g., understanding how material composition and structure impacts material properties); Identify and rank alternative explanations of an observation based on the current state of scientific understanding in a given field; Construct computational models, e.g., of synthesizability of materials, that make experimentally testable predictions; Identify data that support or refute a given conjecture; Design and prioritize, orchestrate, and execute experiments; Create plans for replicating studies and validating their claims; Generate and rank alternative interpretations of the data; Document studies, communicate results; Infer causal relations from observations and data; solve inverse problems using data and simulations; Assemble scientific teams that are best equipped to answer specific questions; Track scientific progress, evolution of scientific disciplines, and scientific impact. Addressing the AI grand challenge of accelerating scientific progress requires synergistic advances across multiple areas of AI, data and computational infrastructure for collaborative science, including the appropriate governance structures and processes, standardized data and metadata representations, data sharing, data discovery, and data access protocols, integration of shared instruments with computing and data infrastructure, findable, accessible, reusable, and shareable data analytic pipelines, and a sustainable research ecosystem; and innovative modes of training a diverse cadre of scientists and engineers who are well-versed in MSE, and have a deep knowledge of AI to drive convergent advances across both disciplines.

Future Plans

We plan to organize (i) focused workshops to conduct a deep dive into a few key research themes identified in the report from our interdisciplinary workshop; (ii) pilot a set of education and training materials to introduce AI tools for MSE researchers and MSE research problems to AI researchers with diverse expertise and interests; (iii) articulate the desiderata and develop blueprints and infrastructure for collaborative, AI-enabled science.

Broader Impacts and Workforce Development

The project has established partnerships e.g., with the National Data Science Corps, an initiative of the NSF-funded Big Data Regional Innovation Hubs, to educate and train diverse cadre of MSE and AI students and professionals. We are also developing approaches to make Data Sciences and AI training available to MSE graduate and undergraduate students.

Data Management and Open Access

All of the project results are being freely disseminated to the community through the project webpage.

Advancing Along the Materials Development Continuum and Partnerships to Translation

We envision translational efforts through partnership with industry and national labs.

Publications and References

Honavar, V., Morgan, D., Van Duin A, Mahdavi, M. and Olivetti, E. Accelerating Materials Discovery, Design, and Synthesis: A Grand Challenge for Artificial Intelligence. Workshop report available through the project website. Under revision for journal publication.

Electrocatalysis Consortium (ElectroCat): A DOE Energy Materials Network (EMN) Consortium

Poster Presenter: McKenzie Hubert, Hydrogen and Fuel Cell Technologies Office (mckenzie.hubert@ee.doe.gov)
Lead Investigators: Deborah Myers (dmyers@anl.gov), Piotr Zelenay (zelenay@lanl.gov)
Participating Core Institutions: Argonne National Laboratory, Los Alamos National Laboratory, Oak Ridge National Laboratory, National Renewable Energy Laboratory
Website: https://www.electrocat.org/
Keywords: energy, hydrogen, fuel cell, electrolyzer, electrocatalyst

Project Scope

The ElectroCat (Electrocatalysis) Consortium is aimed at increasing U.S. competitiveness in developing and manufacturing fuel cell and water electrolyzer energy conversion devices by addressing primary challenges to the widespread implementation of these technologies. The platinum group metal (PGM) electrocatalysts that are the current standard in low-temperature fuel cell and electrolyzer systems are expensive and restrict the ability to be cost-competitive with traditional hydrocarbon-based technologies. In this sense, catalyst design represents the most pressing material barrier related to fuel cell and electrolyzer deployment. ElectroCat is addressing this barrier by accelerating the development and deployment of platinum group metal-free (PGM-free) electrocatalysts in low-temperature fuel cells and electrolyzers.

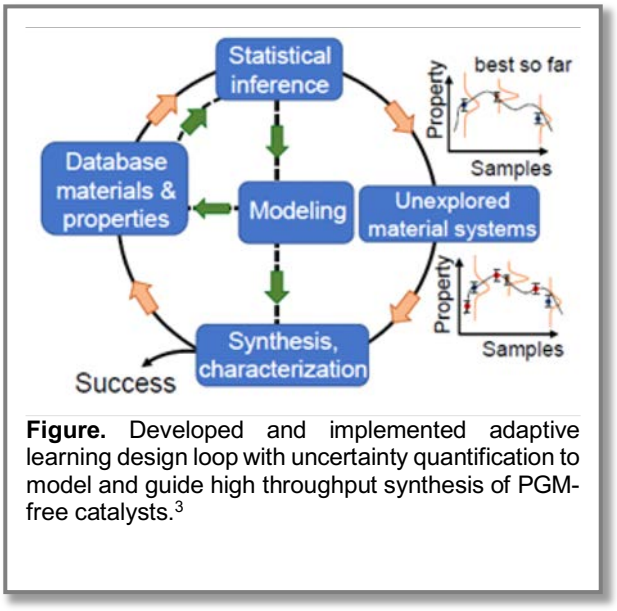


Figure. Developed and implemented adaptive learning design loop with uncertainty quantification to model and guide high throughput synthesis of PGM-free catalysts.[3]

Relevance to MGI

As part of the DOE's Energy Materials Network (EMN), ElectroCat employs a systematic approach in which catalysts are synthesized and analyzed rapidly and comprehensively using high-throughput, combinatorial methods. These in turn are guided by computational work, advanced characterization, and the fundamental electrocatalysis and materials knowledge housed across the national laboratory network. Streamlined data sharing with industry and academia partners is critical to the ElectroCat approach, rapidly building an understanding of PGM-free electrocatalysts across the field and, ultimately, enabling the incorporation of those materials into next-generation fuel cells and electrolyzers.

Technical Progress

Some key accomplishments that highlight the success of ElectroCat's systematic approach for PGM-free catalyst development include:

- *Identification of catalytic active sites:* Computational modeling using density functional theory and advanced characterization techniques such as in situ X-ray absorption spectroscopy were used to understand catalytic active sites in Fe-N-C catalysts for the oxygen reduction reaction.[1,2]
- *Machine learning-directed synthesis of novel, high-performing catalysts:* High-throughput catalyst synthesis and performance data were correlated in a machine learning model. Output from the model suggested "next-step" experiments that resulted in higher performing catalysts than were previously fabricated.[3,4]

- *Novel synthesis methods for improved catalyst activity:* Optimization of Ni-Fe composition coupled with novel, high surface area synthesis methods resulted in high activity for the oxygen evolution reaction in alkaline electrolyte.[3]

Future Plans

ElectroCat is continuing to apply its data science-guided high-throughput synthesis methods specifically to Fe-N-C catalysts for the oxygen reduction reaction in fuel cells, pushing the boundaries of catalytic activity. The consortium plans to implement best-in-class high-throughput-synthesized Fe-N-C catalysts in PEMFCs and expand their work in low-temperature electrolyzer catalyst development and leverage the successful approach they have employed for fuel cell catalyst development.

Broader Impacts and Workforce Development

In the coming year, ElectroCat plans to partner with Historically Black Colleges and Universities (HBCUs) on fuel cell and electrolyzer catalyst development projects. Students from HBCUs of selected projects will have an opportunity to work directly with the staff and post-docs at the core ElectroCat national laboratories. In addition, ElectroCat is committed to hiring and training new post-docs to conduct material synthesis, electrochemistry, and computational modeling work that is critical to the success of the consortium.

Data Management and Open Access

ElectroCat maintains a Data Hub where it stores data collected and models used. Industry and academic partners also share data through the Data Hub. Streamlined data sharing with these partners is critical to the ElectroCat approach, rapidly building an understanding of PGM-free electrocatalysts across the field and ultimately, enabling the incorporation of those materials into next-generation fuel cells and electrolyzers.

Advancing Along the Materials Development Continuum and Partnerships to Translation

By pooling their expertise, ElectroCat's national laboratory partners (co-led by Argonne National Laboratory and Los Alamos National Laboratory, with partners Oak Ridge National Laboratory and National Renewable Energy Laboratory) will advance the tools needed to model, synthesize, characterize, and optimize PGM-free catalysts and electrode structures to the point that they are easily applied to a broad range of catalyst systems and set a standard for rapid material analysis. The consortium approach allows for efficient collaboration between multiple lab partners and also provides a structure for outside stakeholders to leverage the expertise at the core national labs to accelerate innovation. Such capabilities include material synthesis and processing, catalyst and electrode characterization, and computational material modeling.

Publications and References

A list of ElectroCat publications can be found at https://www.electrocat.org/electrocat-publications/

1. L. Osmieri, R. K. Ahluwalia, X. Wang, H. T. Chung, X. Yin, A. J. Kropf, J. Park, D. A. Cullen, K. L. More, P. Zelenay, D. J. Myers, K.C. Neyerlin, *Appl. Catal. B: Environ.*, 257, 117929-117941 (2019).
2. H. Yu, M. Zachman, D. Myers, R. Mukundan, H. Zhang, P. Zelenay, K. Neyerlin, and D. Cullen, *Microsc. Microanal*. 27, 974-976 (2021).
3. DOE Hydrogen Program Annual Merit Review, June 6-8, 2022 (held remotely). Title: "ElectroCat 2.0 (Electrocatalysis Consortium);" P. Zelenay and D. Myers.
4. M. Karim, M. Ferrandon, S. Medina, E. Sture, N. Kariuki, D.J. Myers, E.F. Holby, P. Zelenay, and T. Ahmed, ACS *Appl. Energy Mater.,* 3, 9083-9088 (2020).

Data-driven integration of experiments and multi-scale modeling for accelerated development of aluminum alloys

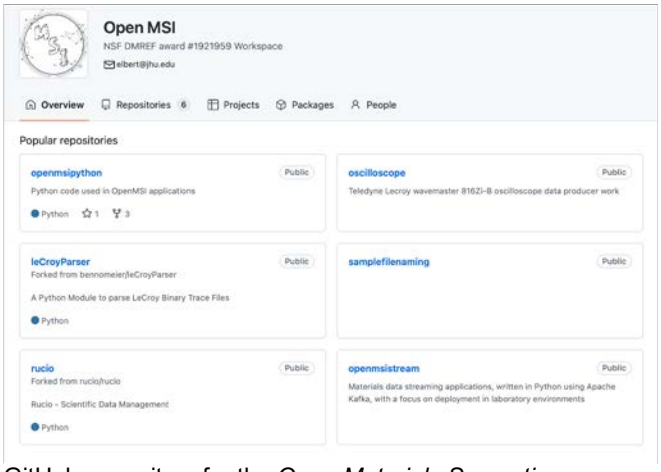
Lead Investigator: Todd Hufnagel, Johns Hopkins University, hufnagel@jhu.edu
Co-Principal Investigators: David Elbert, Johns Hopkins University, elbert@jhu.edu; Somnath Ghosh, Johns Hopkins University, sghosh20@jhu.edu; K.T. Ramesh, Johns Hopkins University, ramesh@jhu.edu
Website: https://github.com/openmsi
Keywords: Materials semantic infrastructure, OpenMSI, aluminum, spall, automated data streaming

Project Scope

We seek to establish a new paradigm for the materials design loop in which the flow of data, rather than individual modeling or experimental tasks, is viewed as central. We do this by leveraging modern data science tools to create the semantic framework and physical infrastructure necessary for seamless integration of experiments and models in complex materials development problems. We demonstrate this approach through a multi-scale modeling framework for the resistance of commercial aluminium alloys to spall failure, with automated connections to advanced microstructural characterization and high-throughput laser shock testing through a centralized data layer.

Relevance to MGI

Traditionally, materials design is an iterative process in which the properties of the material are gradually improved by adjusting the details how of the material is made. This process can be accelerated by introducing computational models into the loop and comparing the output with experimental measurements of the structure and properties of the material at each step. The cycle is slowed, however, by the need for humans to be involved at almost every step for data collection, analysis, and reduction. In this program we are leveraging modern data science tools to develop a data framework and infrastructure for seamless integration of experiments and models in complex materials development problems. In our paradigm the flow of data, rather than individual modeling or experimental tasks, is viewed as central, and the framework is generic, allowing the ideas to be applied across a wide range of application areas.



GitHub pepository for the *Open Materials Semantic Infrastructure (OpenMSI)*. Written in Python and implemented on Windows, macOS, and Linux to enable automated data flows, OpenMSI is built around two types of processes: *Data producers* run on laboratory computers and stream experimental data automatically to the data pool. *Data consumers* subscribe to specific data streams and may return results to the pool or direct new experimental or computational activities. Data streaming is automatic and processes interact only with the pool, not with each other.

Technical Progress

Our ideas are instantiated in the *Open Materials Semantic Infrastructure (OpenMSI)*. A key attribute of OpenMSI is the definition and description of classes of materials data, and connections among these classes, that are required to create an automated data flow in which each experimental and computational task in the design loop automatically pushes information to, or pulls information from, a data layer common to all tasks. At the heart of OpenMSI is a data layer, which can be implemented using either local or cloud-based compute and storage resources via Apache Kafka. Two kinds of processes communicate with this data pool. *Data producers* provide data to the pool; a typical example would be a process running on a laboratory computer connected to an instrument that produces experimental data. The data producer process runs in the background, automatically streaming data to the pool as it is collected without human intervention. The second kind of process, *data consumers*, receive data automatically streamed from the pool. A typical example would be a computational model or simulation, which uses experimental data as input. The data producers can specify the type(s) of data to

which they subscribe. Note that multiple data producers and data consumers can run on the same machine, allowing for a fully-automated closed loop in which data (experimental or simulation results) prompts subsequent actions by other processes. As shown in the box, OpenMSI is written in Python and has been implemented on Windows, macOS, and Linux operating systems. We have created a GitHub repository to enable distribution of OpenMSI to a community of potential users.

A specific demonstration of our approach is the implementation of a multi-scale modeling framework for the resistance of commercial aluminum alloys to spall failure. The first example of automation is a data producer which runs on a computer that collects photon doppler velocimetry (PDV) data during laser-driven microflyer spall tests. The PDV data, and the associated metadata, are automatically streamed to our data layer which is implemented as an Apache Kafka broker running on Confluent Cloud. Use of these data by consumers is facilitated by metadata captured in the lab and automatically organized as an instance of a Graphical Expression of Materials Data (GEMD) model data model. A corresponding data consumer process, running on a different computer, subscribes to this data stream, automatically receives the raw PDV data, extracts the associated metadata, and uses both to process the raw data to produce the final result.

Other progress has focused on other experimental and computational tasks necessary to demonstrate the OpenMSI concept on the aluminum spall problem. On the experimental side we continue to develop and refine the laser microflyer shock experiments described above, and have conducted a series of experiments to investigate the effects of heat treatment on aspects of microstructure relevant to the spall problem (size and spatial distribution of inclusions, and the aluminum grain structure). Modeling efforts are focused on the inverse microstructural design problem and developing a crystal plasticity model for deformation under dynamic loading.

Future Plans

- Continue development of the OpenMSI framework by writing new data producer/consumer processes for tasks relevant to the aluminum spall problem.
- Continue development the crystal plasticity framework to dynamic loading and implementation of a phase field model for spall void growth.
- Develop an inverse modeling framework for optimization of processing parameters for producing target microstructures identified in the micromechanical modeling.

Broader Impacts and Workforce Development

There are two PhD students and one post-doctoral scholar participating in the materials aspects of the project, either experimental (processing and characterization of samples), computational (micromechanical and process modeling), or both. In addition, two MS students have contributed to the code development. All of these are intimately involved in the data aspects, including definition of metadata and implementation of data producer/consumer processes.

Data Management and Open Access

We have established a GitHub repository for OpenMSI.

Advancing Along the Materials Development Continuum and Partnerships to Translation

The OpenMSI framework has the potential to dramatically accelerate the development of new materials by automating handling of data in the iterative materials design loop. Our code is open source and implemented on widely-used platforms (Windows, Linux, and macOS) so the barriers to implementing OpenMSI are minimal. Beyond our own examples cited above, OpenMSI has been implemented in two laboratories connected to the NSF PARADIM user facility, one for automated handling of x-ray diffraction data (at Johns Hopkins) and another for electron microscopy data on the Titan TEM at Cornell.A primary focus going forward will be advertising and disseminating the capabilities of OpenMSI to the community.

Publications and References

Nothing to report.

FLOSIC: Efficient Density Functional Calculations Without Self-Interaction

Lead Investigator: Koblar Jackson, jacks1ka@cmich.edu
Participating Institutions: Central Michigan University, the University of Texas at El Paso, Temple University, the University of Florida, and the University of Pittsburgh
Website: www.flosic.org
Keywords: Density functional theory, self-interaction correction, molecular magnetism, catalysis

Project Scope

The goal the FLOSIC Center is to enable efficient DFT calculations without self-interaction using the Fermi-Löwdin orbital (FLO) self-interaction correction (SIC) method, a computationally efficient approach to implementing the Perdew-Zunger SIC theory (PZSIC). The Center's research includes developing new theoretical approaches to SIC, creating efficient, user-friendly software implementing FLOSIC, and using FLOSIC in applications related to molecular magnetism and MOF-based catalysis.

Relevance to MGI

Density functional theory (DFT) calculations are at the core of the Materials Genome Initiative. DFT methods can make useful predictions of many materials properties at a relatively modest computational cost. Yet self-interaction errors (SIEs) are present in all DFT calculations and can skew the predictions, especially when bonds between atoms are stretched, but also for some equilibrium geometries when transition metal atoms are involved. Research in the FLOSIC Center is aimed at eliminating these SIEs.

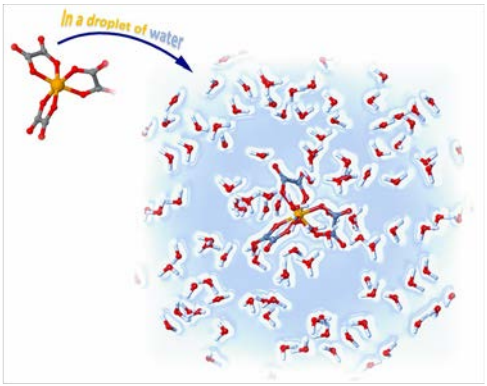


Figure 1. Self-interaction corrections can eliminate unphysical charge transfers than occur in DFT calculations, for example, when solvating a negatively charged transition metal complex in water. With SIC, the anion charge remains correctly localized on the complex, rather than delocalized into the water bath.

Technical Progress

Several advances were made in the theory area in the past year. First, the Zope group developed a numerical Poisson solver for scaled charge densities to enable self-consistent LSIC studies. The locally-scaled, self-interaction correction method LSIC improves over PZSIC in predicting many physical properties by using an iso-orbital indicator to reduce correction in regions where it is not needed. A second advance involves assessing the performance of FLOSIC-derived densities in reducing density errors. Perdew and collaborators found that density correction elevates the strongly-constrained and appropriately normed SCAN functional to coupled-cluster accuracy for water clusters, liquid water, and ice, and can replace the much more expensive coupled-cluster method to determine the many-body interactions needed for simulations of water in all its phases. The FLOSIC density was not as successful as the Hartree-Fock density in removing the density errors. Finally, the Pederson group generalized the FLOSIC method to incorporate complex FLOs (cFLOs). cFLO densities are smoother than real FLO densities and yield lower FLOSIC total energies. The cFLO/rFLO difference increases with increasing hybridization of the orbitals and is especially pronounced for molecules with multiple bonds.

Progress was also made in code/method development. An electrostatic model was developed by Peralta, Pederson, Perdew, and Jackson to determine approximate Fermi Orbital Descriptor (FOD) positions. Finding optimal FODs is a key, but computationally slow, step in FLOSIC calculations. The new "quick-FOD" method provides a very fast set of near-optimal FODs for a variety of chemical environments. A second approach to locating optimal FOD positions focuses on the topology of the electron density. The Peralta group found an analytical connection between FODs and the electron localization function (ELF) via the requirement of stationary Fermi orbitals. Initial results show a close similarity between ELF critical points and relaxed FODs for a number of small

molecules. Other code-related initiatives include the first FLOSIC calculations for rare-earth (f-electron) atoms (Pederson, Baruah), which show significant improvement in FLOSIC orbital energies compared to electron removal energies, and a scheme to obtain a physically correct starting charge density in situations where DFT calculations give unphysical charge transfers (Pederson). Finally, the computationally efficient r^{2}SCAN functional was recently implemented in the FLOSIC code (Zope).

In applications, a systematic study led by Baruah demonstrated that eliminating SIE results in significant improvements in calculated chemical reaction barriers. Christou and Peralta developed a protocol combining DFT computations, a magnetostructual correlation (MSC) for Fe^{3+}/O clusters, and experimental data to predict the magnetic properties of high nuclearity molecular metal-oxo clusters. Jackson, Peralta, Johnson and Perdew showed that strong over-binding of adsorbates on transition metal ions is eliminated in SIC calculations. Finally, Ruzsinszky found that both SIE and density errors play a role in the inaccurate DFT values of low-spin/high-spin energy difference in spin-crossover complexes.

Future Plans

The following briefly describes plans for new and on-going research in the Center. The Zope group will work to parallelize the self-consistent LSIC method and search for improved scaling schemes. The Perdew group will explore a new approach to removing SIEs from the SCAN meta-GGA. The aim is to preserve the SCAN description of weak and metallic bonds, but improve over SCAN for stretched covalent bonds. Ruzsinszky will apply FLOSIC and LSIC to absolute and relative core-level binding energies, as well as to polaronic effects. In code/method development, the Pederson group will enhance the cFLOSIC code and address formal issues connected to the use of complex orbitals. The Johnson group will explore the use of machine learning methods to determine near-optimal FOD positions. Additional initiatives include working on a periodic FLOSIC code, testing mixing schemes to improve FLOSIC SCF convergence, exploiting symmetry in FLOSIC calculations, reducing memory demands of FLOSIC calculations, and enhancing the usability of the code for non-experts. Plans for new applications include studying the adsorption of small molecules on transition metal binding sites in MOFs, barriers for catalytic reactions in MOFs, bond length alternation studies in polymer chains, studies of a Cr tri-anion complex in a large bath of explicit waters, and studies of magnetic properties of additional multi-nuclear Fe complexes, including a $Fe_{12}La_4$-oxo cluster.

Publications and References

1. S. Akter, J. A. Vargas, K. Sharkas, J. E. Peralta, T. Baruah, K. A. Jackson, and R. R. Zope, *How well do self-interaction corrections repair the over-estimation of static polarizabilities in density functional calculations?,* Physical Chemistry Chemical Physics **23**, 18678 (2021).
2. L. Li, C. S. Ewing, M. Abdelgaid, S. Banerjee, R. B. Garza, R. Hacku, N. D. HGess, S. Hong, M. Islamov, L. L. Luciani, S. Papadopoulos, C. Parker, R. A. Patel, R. Pawar, M. H. Poklar, M. Salem, F. Shabnam, P. B. Shukla, S. K. Vena, U. Wang, R. Wang, X. Fei, F. Yang, and J. K. Johnson,*Binding of CO and O on Low Symmetry Pt Clusters Supported on Amorphous Silica,* J. Phys. Chem. C, **125**, 13780 (2021).
3. D. B. Nguyen, M. R. Pederson, J. P. Perdew, K. A. Jackson, and J. E. Peralta, *Initial Fermi orbital descriptors for FLOSIC calculations: The quick FOD method* Chem. Phys. Lett. **780**, 138952 (2021).
4. C. M. Diaz, L. Basurto, S. Adhikari, Y. Yamamoto, A. Ruzsinszky, T. Baruah, and R. R. Zope, *Self-interaction-corrected Kohn-Sham effective potentials using the density-consistent effective potential method,* J. Chem. Phys. **155**, 064109 (2021).
5. S. Dasgupta, E. Lambros, J. P. Perdew and F. Paesani, *Elevating Density Functional Theory to Chemical Accuracy for Water Simulations through a Density-corrected Many-body Formalism,* Nat. Comm. **12**, 6359 (2021).
6. J. P. Perdew, *Perspective: Artificial intelligence "sees" split electrons,* Science **374**, 1322 (2021).
7. P. Mishra, Y. Yamamoto, J. K. Johnson, K. A. Jackson, R. R. Zope, and T. Baruah, Journal of Chemical Physics, *Study of self-interaction-errors in barrier heights using locally scaled and Perdew-Zunger self-interaction methods*, **156**, 014306 (2022).
8. K. P. K. Withanage, K. Sharkas, J. E. Peralta, J. K. Johnson, J. P. Perdew and K. A. Jackson, *Fermi-Löwdin Orbital Self-interaction correction of adsorption energies on transition metal ions,* J. Chem. Phys. **156**, 134102 (2022).
9. P. Mishra, Y. Yamamoto, P.-H. Chang, D. B. Nguyen, J. E. Peralta, T. Baruah, and R. R. Zope, *Study of Self-Interaction Errors in Density Functional Calculations of Magnetic Exchange Coupling Constants Using Three Self-Interaction Correction Methods*, J. Phys. Chem. A **126**, 1923 (2022).
10. K. P. K. Withanage, K. A. Jackson, and M. R. Pederson, *Complex Fermi- Löwdin orbital self-interaction correction*, J. Chem. Phys. **156**, 231103 (2022).

Rheostructurally-informed Neural Networks for geopolymer material design

Lead Investigator: Safa Jamali, s.jamali@northeastern.edu
Participating Institutions: Northeastern University (University of Delaware, and Georgetown University)
Website: none
Keywords: Geopolymers, Rheology, ML-Accelerated Additive Manufacturing, Data-Driven Process Design

Project Scope

Geopolymers are amorphous, solid matrices (typically from clays) that grow into flocs and aggregate into gels, yielding ceramic-like structures, with recoverable water as the reaction medium within pores of the gel. The possibility of using a range of clay minerals, and their durability make geopolymers among the most interesting candidates for sustainable new materials sources, with the potential to drastically reduce CO_2 emissions and water use. Rheology is key to controlling the flow and structure of highly heterogeneous geopolymers into desirable structures. The main objective of the project is to exploit the potential of geopolymer materials, by designing, from the initial chemistry, the final material properties through the gel development and its evolution during processing.

Relevance to MGI

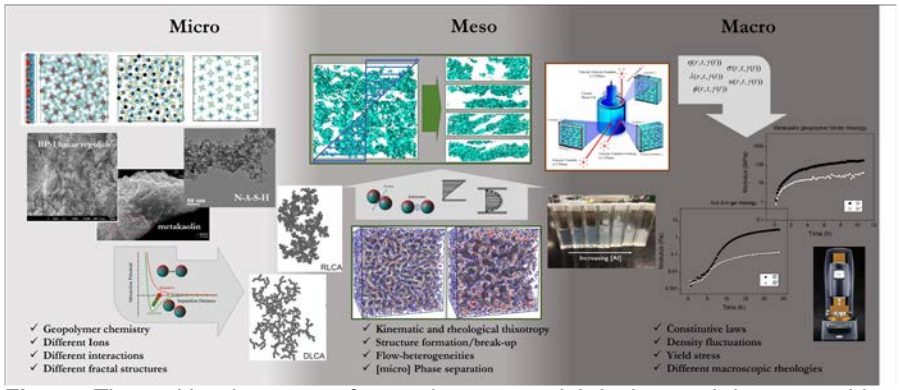


Figure. The multiscale nature of geopolymer material design, and the overarching plan for integration of simulation, experiments and machine learning.

Variety of different indegenous geopolymers and the intertwining of different processes in fig.1 makes their design cycle extremely long and difficult. In our project, we combine a range of advanced computational tools with experiments and develop neural networks that are informed by the micro-, meso-, and macro-scale. Employing physics-informed AI, we are developing the next generation of predictive meta-models for sustainable manufacturing of structures from geopolymers. Our approach consists of a concerted application of experiments, computation, and data-driven methods across multiple scales. This approach will bridge the gap between scales by identifying key targets for convergent research across scales. In the context of geopolymer processing and rheology, the mesoscale gel properties play a key role, and thus are identified as the check point for each tool as well as multiple opportunities for rigorous cross-validation, bench-marking, and iterative feedback between experimental and computational approaches. This truly multi-scale approach in which information from each scale is provided to next also fundamentally impacts the material design process in general. Here, we will focus on design and discovery of geopolymeric materials via data-driven algorithms that are informed by multiscale interrogation of these systems under different processing conditions.

Technical Progress

Although the project is still in its initial phase, we have made some progress towards developing the tools, and benchmarking different geopolymer systems against existing models/materials in the literature. Co-PI Wagner's team has made tremendous progress in understanding the rheology of two different geopolymer systems during structure formation and developed a constitutive model capable of predicting the moduli of the resulting gels. Co-PI Del Gado's team has developed detailed simulations of reactive gelation providing a fundamental understanding

of how different geopolymer chemistries will result in different porous gel structures and rheologies. PI Jamali's team has established a rigorous set of data-driven tools, referred to as digital rheometers, that leverage advances in physics-informed neural networks and can be used in place of a physical rheometer. These digital twins to a rheometer will enable fast and reliable prediction of rheological behavior [of an unknown or newly developed material] with a very few number of actual tests, accelerating the process of testing a material's mechanics.

Future Plans

The main objective of the project is to develop a set of rheostructurally-informed data driven algorithms that enable design and discovery of geopolymer materials and their processing into structural components. Thus, the future plans include: i) continuously testing different geopolymer systems and their mechanical, rheological, and structural characteristics to build a better understanding of their formation as well as a reliable data base for data-driven design, ii) performing large scale simulations on gelation of these nanoscale clusters into macroscopic ceramic-like structures to identify the most important features that must/can be manipulated during the processing, and iii) developing an integrated AI platform, informed of these processes as well as a series of experimental data sets, for acceleration of design cycle for geopolymer systems. The efforts in the year 3 will move onto actual additive manufacturing of these components into structures.

Broader Impacts and Workforce Development

Geopolymers are ubiquitous in nature and form a class of materials that promise exceptional mechanical properties at virtually no detriment to the environment. 8% of the world's CO_2 emissions come from cement production and this can be greatly reduced using geopolymers. Thus, societal broader impacts of the project are rather clear. This project, the tools that are being devised and used are: 1- versatile in nature, from experimentation to novel data-driven methodologies, and 2- of interest to fields of research well beyond geopolymer or rheology. Thus, a major focus of the PIs in this project has been to train graduate students and postdoctoral researchers that are familiar with all these different tools and algorithms. With an ever-growing role of AI and computational capabilities in science and engineering, it's imperative that the next generation of workforce in materials research is capable of using these tools and technologies. The group of PIs in the project are also actively engaged in educational activities beyond their home institutions, and through Society of Rheology's short courses and conferences, providing another platform for training the new generation of workforce in materials research.

Data Management and Open Access

Throughout the project, research data is being generated by all three groups involved, with existing platforms between the involved personnel. However, more importantly, the PI continuously releases the latest versions of the AI methodologies developed in the project through GitHub repositories accessible to public. These services and capabilities are accompanied by detailed tutorials as well as sample codes for other researchers in order to facilitate use and stimulate further feedback and collaboration between the users. This will promote a meaningful long-term impact to be made by the project's outcomes, with a sustainable and consistent improvement of the methodologies developed. More importantly the graduate student at Northeastern University will be collecting all literature data on geopolymer materials, their chemical characterization, resulting structures and rheological measures, contributing to a unique data collection for training of the neural networks.

Advancing Along the Materials Development Continuum and Partnerships to Translation

As described previously, the variety of different geopolymer sources and chemistries make it virtually impossible to develop a generic pathway to their design into structural materials. However, use of rheology and structure as guiding characteristics/features of these materials, and development of a logical relationship between their microscopic characteristics and macroscopic properties will enable this process development. This will be done, specifically, using our data-driven platforms and neural networks that are informed by different detailed interrogations of the system. Geopolymers are becoming increasingly more commercially viable options, and the results of this project can significantly impact their application as structural materials.

AI Institute: Planning: Novel Neural Architectures for 4D Materials Science

Lead Investigator: Yang Jiao (ASU, yang.jiao.2@asu.edu); Nik Chawla (Purdue, nikc@purdue.edu)
Participating Institutions: ASU, Purdue, Syracuse, U. Wyoming, U. Kentucky, UNR
Website: https://materialsai.github.io/
Keywords: 4D Materials Science, microstructure representation learning, structure-property mapping, dynamic neural graph for microstructural evolution, neural architecture optimization

Project Scope

The theme of the project involves the development of revolutionary approaches enabled by fundamental and use-inspired AI research to understanding structure-property relationships in vastly different materials systems for both predictively modeling and optimal material design, with complementary mechanical, electronic, thermal application areas. The goal is to accelerate converging research on new learning theories, experimentation methodologies, and validation protocols that will facilitate scientific modeling of the evolutionary and hierarchical structure-property mappings of complex materials systems. We mathematically formulate the ubiquitous challenges in modeling structure-property mappings across critical application domains (alloys, composites, porous materials, nuclear fuels, etc.), demonstrate the necessity and feasibility of AI in addressing these challenges, and identify a consortium of industry collaborators to transfer the knowledge into practical solutions and creating new workforce.

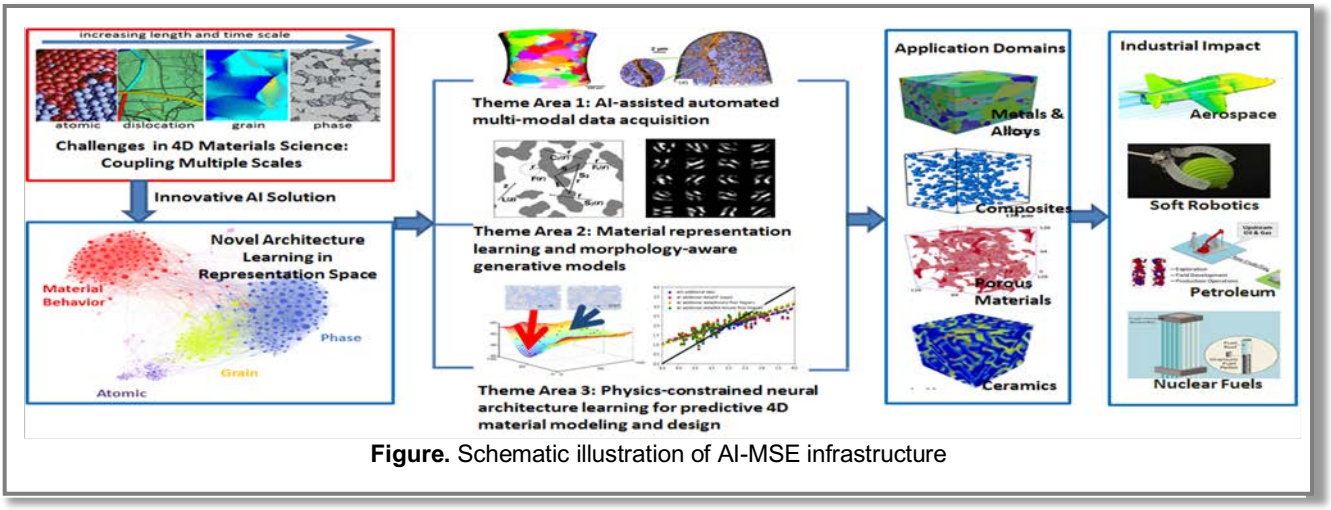


Figure. Schematic illustration of AI-MSE infrastructure

Relevance to MGI

We introduce the novel concept of AI-MSE that involves predictive neural network models that possess unique learning architectures capturing the physical causal relations across key microstructural features, and explicitly constrained by known physical laws and capable of incorporating newly discovered material physics. The project inspires one to re-think the utility of machine learning in materials science: From knowledge-agnostic feature learning to reasoning mechanisms adaptive to domain-specific knowledge. The methodologies and frameworks for constructing novel physics-based learning models developed can be readily applied to a variety of compelling problems in heterogeneous material systems including ceramics, alloys, composites, and porous materials. It will provide the key infrastructure for potential automated material characterization, research and discovery.

Technical Progress

Key technical progresses are made in the following areas: **(a)** AI-assisted acquisition of multi-modal 4D data characterizing evolution of features on multiple length scales, including the development of convolutional neural network (CNN)-based imaging processing methods for auto-data segmentation. **(b)** Development of quantitative microstructure representations (QMR) for (evolving) key structural features on multiple length (time) scales, through novel integration of physics-driven spatial correlation functions and data-driven feature learning methods. In particular, novel higher-order correlation-based microstructure representations and metrics for quantifying

evolving material structures are developed and tested on synthetic 4D data obtained via phase-field modeling. Physics-inspired network architecture for learning robust reduced dimension representations for complex evolving material systems are developed and validated in a variety of heterogeneous material systems. **(c)** Development of novel physics-driven architecture search methods to identify key QMR features and their causal relations for the prediction of material properties of interests and their validation via physics-based simulations and experiments. Dynamic neural graphs are combined with novel representation learning for informing emerging physics in evolving material systems with a hierarchy of structural complexity. This was employed to design a novel multi-functional smart composite exhibiting superior dynamic tuning performance that is ideal for soft robotics applications.

Future Plans

Our future plan includes **(i)** applying the AI-MSE tools developed in the project in specific material domains, relevant to aerospace and auto-industry, nuclear fuel technologies, soft robotics and petroleum engineering; **(ii)** obtaining insights in fundamental AI theories such as generalizability and expandability through domain-specific tools; and **(iii)** engaging professional societies for joint efforts of database sharing and educational activities.

Broader Impacts and Workforce Development

This project inspires the re-thinking of utility of AI in MSE, which requires the search and design of novel neural architectures that requires fundamental AI research. The AI-MSE infrastructure is suitable for handling a wide class of distinct material systems with applications in auto-industry, nuclear fuel technologies, soft robotics and petroleum engineering. The project strives to train and prepare next generation of AI-MSE scientists and engineers, by creating new courses (including both in-person and virtual sections) and research opportunities engaging over 30 graduate, undergraduate and high school students, with 10 female students. It also provides a myriad of research and educational opportunities for junior faculty, women, and other underrepresented groups. The central southwest location well positioned us to recruit students of Native American and Hispanic origin.

Data Management and Open Access

The project website (https://materialsai.github.io/) documents planning activities (14 seminar talks to date) and publications. For reproducibility and dissemination purposes, data and code repositories related to the publications are also provided. We plan to integrate the datasets and programs developed throughout this project into a documented software package and make it publicly available to the research community.

Advancing Along the Materials Development Continuum and Partnerships to Translation

The project has engaged with industrial partners (Uniformity Lab, ALPEMI, Intel, etc.) via research meetings, and workshops. These interactions direct our research to generate deliverables directly address industrial needs. Two common themes, i.e., (i) material processing control and optimization, (ii) predictive modeling under extreme conditions, are of interest to many industrial applications. We have accordingly been developing AI-based 4D MSE tools, combining multi-model representation learning and dynamic neural graphs to address these challenges, which will be packaged into ready-to-use software tools with GUI.

Selected Publications

1. H. Torbati-Sarraf, S. Niverty, R. Singh, D. Barboza, V. De Andrade, P. Turaga, and N. Chawla, *Machine Learning-based Algorithms for Automated Image Segmentation of Transmission X-ray Microscopy (TXM) Datasets*, JOM **73** 2173 (2021).
2. S. Cheng, Y. Jiao, and Y. Ren, *Data-Driven Learning of Three-Point Correlation Functions as Microstructure Representations*, Acta Materialia **229**, 117800 (2022)
3. P. E. Chen, R. Raghavan, Y. Zheng, K. Ankit, and Y. Jiao, *Quantifying Microstructural Evolution via Time-Dependent Reduced-Dimension Metrics Based on Hierarchical n-Point Polytope Functions*, Physical Review E **105**, 025306 (2022)
4. Y. Xu, P. E. Chen, H. Li, W. Xu, Y. Ren, W. Shan, and Y. Jiao, *Correlation-Function-Based Microstructure Design of Alloy-Polymer Composites for Dynamic Dry Adhesion Tuning in Soft Gripping*, Journal of Applied Physics **131**, 115104 (2022)
5. S. Sharifi, A. M. Nasab, P. E. Chen, Y. Liao, Y. Jiao, and W. Shan, *Robust Bi-continuous Metal-Elastomer Foam Composites with Highly Tunable Stiffness*, Advanced Engineering Materials, in press (2022)
6. T. Davydzenka, D. Sinclair, N. Chawla, P. Tahmasebi, *Deep-Layers-Assisted Machine Learning for Accurate Image Segmentation of Complex Materials*, Materials Characterization, (2021) [under revision]

Artificial intelligence and data science enabled predictive modeling of collective phenomena in strongly correlated quantum materials

Lead Investigator: Steven Johnston, University of Tennessee, Knoxville.
Participating Institutions: University of Tennessee, Knoxville; UC Davis; San Jose State University; Oak Ridge National Laboratory; Brookhaven National Laboratory; Los Alamos National Laboratory
Website: None.
Keywords: quantum materials, non-perturbative approaches, model extraction, artificial intelligence, smart experiments.

Project Scope

This project leverages machine learning (ML), artificial intelligence (AI), and data science to advance and accelerate the predictive modeling of correlated quantum materials. We aim to develop methods to 1) generate and confirm novel low-energy effective many-body models for quantum materials, 2) accelerate non-perturbative model solutions, and 3) create workflows that integrate experiment and theory to accelerate discovery at national user facilities. We will deploy this technology to study correlated systems hosting a mix of localized and delocalized excitations, which are challenging to model with conventional methods.

Relevance to MGI

Our project's ultimate goal is to provide the community with the necessary tools to develop "smart" experimental workflows at national user facilities. These methods will enable *in situ* analysis of data collected at the beamline, allowing theory and modeling to be fed back into the experiment in real-time to inform investigations using AI and ML methods. The tools we are developing can also be used to support high-throughput calculations for quantum materials that move beyond *ab initio* methods.

Technical Progress

Our research goals require fast solvers for the direct scattering problem; these methods are the foundation of any data-driven approach to solving the inverse scattering problem. This year we have worked to develop fast, scalable, and generalizable methods for solving two critical classes of models. The first is a nonperturbative hybrid (quantum) Monte Carlo (HMC) method for solving systems of itinerant electrons coupled to the lattice.[3,4] The second is a generalization of Landau-Lifshitz dynamics, a semiclassical approximation for computing dynamical correlation functions of quantum spin systems.[5] Crucially, both approaches scale linearly in the system size.

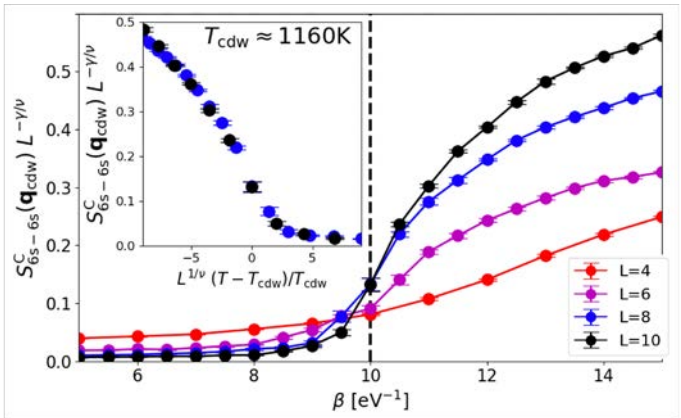


Figure. The charge structure factor for $BaBiO_3$, derived from numerically exact QMC simulations of a DFT-derived multiorbital sp-model defined on cubic clusters with $4L^3$ orbitals. The inset shows an Ising scaling analysis to determine the *thermodynamic* transition temperature.

Using our HMC approach, we successfully calculated the metal-insulator transition in $BaBiO_3$ in the *thermodynamic* limit (Fig. 1), using model parameters derived from first principles simulations. Using our semiclassical framework, we've found a general mechanism for realizing a new type of skyrmion in materials.[7] We have also used machine learning to derive an effective spin model with bilinear and biquadratic spin interactions that reproduce the phase diagram obtained with the full Kondo lattice model. A significant advantage of this procedure is that we can solve the effective model and compute its phase diagram in approximately two days. (It takes more than six months to obtain the phase diagram with the original Kondo Lattice Model.) The MGI can also leverage these capabilities in high-throughput

workflows to search for novel materials' properties that extend beyond ab initio methods.

Future Plans

We have several ongoing development efforts, including 1) a generalized Langevin dynamics code for computing dynamical correlation functions in thermal equilibrium and 2) a generalized determinant quantum Monte Carlo code for simulating arbitrary multi-orbital Hubbard models, with and without electron-phonon interactions. These codes will be crucial for supplying the fast solvers required to build an AI-powered workflow for solving the inverse scattering problem on correlated quantum materials. We will also complete combined experiments, machine learning, and simulation studies that we have begun to start a general integration of these methods.

Broader Impacts and Workforce Development

The students and post-docs in this project work in multidisciplinary research and receive hands-on training in experimental or theoretical condensed matter physics and its interface with artificial intelligence, machine learning, data science, and high-performance computing. We are also actively engaging with users at the national faculties to showcase the project's technology and train users in their usage.

Data Management and Open Access

All codes developed by this project are publicly available with complete documentation. For example, our HMC code for treating electron-lattice coupled models is available at https://github.com/cohensbw/elphdynamics, and our approach for learning phase transitions with symmetry enhanced AI autoencoders can be found at https://github.com/dagrawa2/ssb_detection_ising. We are also contributing to the development of the SUNNY code (https://github.com/SunnySuite), a fast solver for computing the dynamical spin structure factor of semi-classical spin models. We continue to work with members of the INS community to ensure that the implementation is easy to use for users at national facilities. We plan several more code releases over the project's lifetime.

Advancing Along the Materials Development Continuum and Partnerships to Translation

The technology we are developing can be utilized to develop smart experiments, where *in situ* data analysis can be conducted at national user facilitates, which is fed back into experiments in real time to accelerate discovery. It is not hard to imagine that similar technology could be developed for commercially available experimental devices for smaller scale laboratories. We plan to target such opportunities as our program matures.

Publications and References

1. Y.Shen *et al.*, *Role of Oxygen States in the Low Valence Nickelate $La_4Ni_3O_8$*. Physical Review X **12**, 011055 (2022).
2. Q. Parker, D. Perera, Y. W. Li, and T. Vogel, *Supervised and unsupervised machine learning of structural phases of polymers absorbed to nanowires*. Physical Review E **105**, 035304 (2022).
3. C. Miles, B. Cohen-Stead, O. Bradley, S. Johnston, R. Scalettar, and K. Barros, *Dynamical tuning of the chemical potential to achieve a target particle number in grand canonical Monte Carlo simulations*. Physical Review E **105**, 045311 (2022).
4. B .Cohen-Stead, O. Bradley, C. Miles, G. Batrouni, R. Scalettar, and K. Barros, *Fast and scalable quantum Monte Carlo simulations of electron-phonon models*. Physical Review E, in press (2022). arXiv:2203.01291.
5. D. Dahlbom, H. Zhang, C. Miles, X. Bai, C. D. Batista, and K. Barros, *Geometric integration of classical spin dynamics via a mean-field Schrödinger equation*. arXiv:2204.07563 (2022).
6. D. Agrawal, A. Del Maestro, S. Johnston, and J. Ostrowski, *A group-equivariant autoencoder for identifying spontaneously broken symmetries in the Ising model*. arXiv:2202.06319 (2022).
7. H. Zhang, Z. Wang, D. Dahlbom, K. Barros, and C. D. Batista, *CP^2 Skyrmions and Skyrmion Crystals in Realistic Quantum Magnets*. arXiv:2203.15248 (2022).
8. Z. Wang and C. D. Batista, *Skyrmion Crystals in the triangular Konto lattice model*. arXiv:2111.13976 (2021).
9. N. John, A. Del Maestro, B Rosenow, *Robustness of Helical Edge States Under Edge Reconstruction*, arXiv:2105.14763 (2021).
10. Seung-Hwan Do *et al.*, *Understanding temperature-dependent SU(3) spin dynamics in the S=1 antiferromagnet $Ba_2FeSi_2O_7$*. arXiv:2205.11770 (2022).

Center for Predictive Simulation of Functional Materials

Lead Investigator: Paul R. C. Kent, kentpr@ornl.gov
Participating Institutions: Oak Ridge National Laboratory, Argonne National Laboratory, Brown University, North Carolina State University, Sandia National Laboratories
Website: https://cpsfm.ornl.gov
Keywords: Quantum Monte Carlo, exascale computing, topological materials, vdW materials

Project Scope

The overarching goal of the Center for Predictive Simulation of Functional Materials is the development, validation, and distribution of external-parameter-free methods and open-source codes to predict and understand the properties of functional materials. We emphasize those with strong electronic correlations, van der Waals and spin-orbit interactions, where higher accuracy methodology is particularly required. The Center advances Quantum Monte Carlo (QMC) techniques and the open source QMCPACK code to reach this goal. By solving the Schrödinger equation stochastically, we avoid uncontrolled approximations. Recently, our methodologies, codes, and algorithms have advanced to tackle quantum materials in preparation for the era of Exascale supercomputing.

Relevance to MGI

Although databases of material properties predicted with quantum-mechanics based methods are becoming more common, the methods generally used introduce potentially non-trivial approximations and uncertainty. QMC methods are in principle significantly more accurate, but are not yet sufficiently general, restricting their use. Key improvements made in the Center over the last 2 years address all the fundamental approximations and limitations needed to study quantum materials: improvements in trial wavefunctions [2], inclusion of spin-orbit coupling (SOC), geometry optimization and forces. Significant effort has targeted the development and use of an expanded set of observables to gain improved insight into why particular properties are exhibited. These are also all required steps to a generally deployable methodology. These methods have then been used in various applied studies to gain insight into several classes of materials. This includes metal-insulator transitions in the delafossites and perovskites, and the role of correlation in non-magnetic and magnetic van der Waals systems, the latter requiring full geometry optimization and SOC within QMC to examine spin-phonon and spin-lattice couplings. These are first-of-a-kind demonstrations we will now expand on to construct more robust, automated workflows to begin creating accurate QMC-derived databases of materials properties.

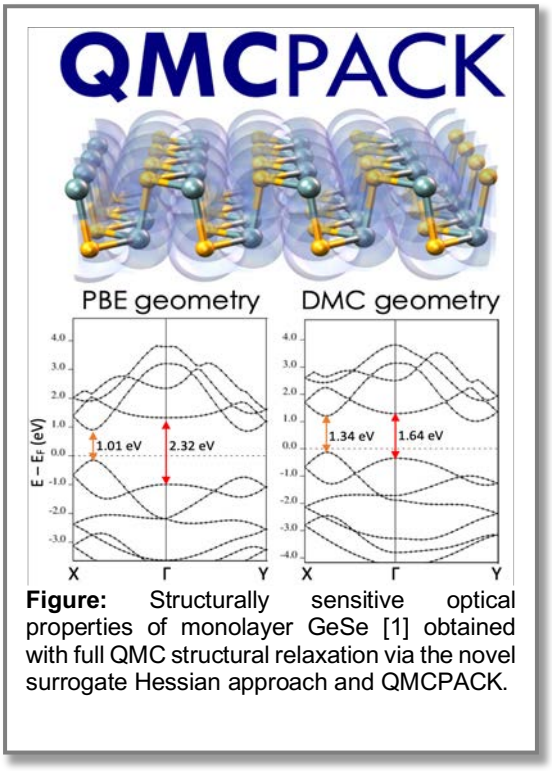


Figure: Structurally sensitive optical properties of monolayer GeSe [1] obtained with full QMC structural relaxation via the novel surrogate Hessian approach and QMCPACK.

Technical Progress

We have developed and implemented a formalism for spin-orbit effects in solids with QMC including a generalized fixed-phase method with efficient evaluation of spinor wavefunctions and a new generation of pseudopotentials [4] to capture many body effects and spin-orbit coupling for heavy atoms consistently. Many quantum materials and phenomena are now in-scope of QMC for the first time. e.g. we performed the first spin-orbit QMC calculation of a 3D periodic system, $RuCl_3$, a prominent candidate for a spin-liquid electronic phase.

We developed a QMC geometry relaxation method [3] that bypasses outstanding theoretical problems with forces. The method uses a surrogate energy Hessian of the atomic coordinates with subsequent structural optimization performed by QMC along favorable coordinate directions. The method scales linearly in the number of degrees of freedom and has the advantage of controlling biases within predefined tolerances, enabling us to obtain accurate geometries for 2D monolayers, such as GeSe [1] and CrI_3, fully within QMC for the first time.

These new methods, as well as previously established QMC methods have been applied to a range of materials including trihalides (CrI_3, $RuCl_3$), delafossites (ABO_2), and magnetic topological materials. For example, $MnBi_2Te_4$ (MBT) is a strong candidate magnetic topological material, but it is prone to be n-doped by intrinsic defects, impeding the observation of topological states. We examined magnetic anti-site defects with QMC [5] to provide guidance for synthesis efforts at reduced defect concentration. The QMC moments were used to construct estimation models for defect concentrations in MBT when combined with experimental magnetic data. The models, following literature validation, indicate current MBT samples are rich in Bi_{Mn} donor defects. Hence, further experimental growth optimization is desirable to access a currently unexplored, low defect region of the phase diagram.

Future Plans

Besides ongoing scientific application, we aim to significantly improve the real-world throughput and productivity of the method. This requires substantial improvement in the robustness of the density functional codes commonly used to generate trial wavefunctions through to automate this work as a focus for improvements in the entire QMC workflow. While QMCPACK exploits, at very high efficiency, current generation pre-exascale machines, more work is needed to establish fully robust, automated workflows. We will expand the functionality of the Nexus workflow system to perform tiered calculations encoding domain expert knowledge to enhance the robustness and reliability of the automated process.

Broader Impacts and Workforce Development

We organized a virtual QMC workshop with 244 registrants in late 2021, indicating very strong demand. We covered a full introduction to QMC and gave the first ever hands-on tutorial on spin-orbit QMC. We developed a new virtual machine that provides QMCPACK and all other needed software. This will be maintained on an ongoing basis as we believe this provides the easiest way to try a QMC calculation. https://github.com/qmcpack/qmc_workshop_2021 links to all the files and presentation materials. High quality recordings from the workshop were uploaded on the QMCPACK YouTube channel and have thousands of views.

Data Management and Open Access

We make regular releases of QMCPACK and the associated workflow software Nexus on GitHub, https://github.com/QMCPACK. Changes cover the full spectrum of new capabilities, expanded testing and documentation, and improved robustness. The surrogate Hessian relaxation method is published open source at https://github.com/QMCPACK/surrogate_hessian_relax. We offer an expanding database of QMC and many-body theory-appropriate, validated pseudopotentials at https://pseudopotentiallibrary.org. These potentials are now being used by the Simons Collaboration as well as new QMC groups. Data from the Center's publications is searchable and discoverable at the Materials Data Facility, https://www.materialsdatafacility.org/.

Advancing Along the Materials Development Continuum and Partnerships to Translation

This project is developing, applying, and validating a fundamental research tool. Today we can provide more accurate results than other electronic structure methods, potentially e.g., aiding the (down) selection of materials for experimental study. Wider deployment requires improving the robustness of all steps of the method to minimize hands-on steps and minimize the degree of expert knowledge required for typical calculations.

Publications and References

1. H. Shin, J. T. Krogel, K. Gasperich, P. R. C. Kent, A. Benali, and O. Heinonen, "*Optimized structure and electronic band gap of monolayer GeSe from quantum Monte Carlo methods"*, Physical Review Materials 5024002 (2021).
2. A. Annaberdiyev, G. Wang, C. A. Melton, M. C. Bennett, and L. Mitas, *"Cohesion and excitations of diamond structure silicon by quantum Monte Carlo: benchmarks and control of systematic biases"*, Phys. Rev. B **103**, 205206 (2021).
3. J. Tiihonen, P. Kent, and J. Krogel, *"Surrogate Hessian Accelerated Structural Optimization for Stochastic Electronic Structure Theories"*, J. Chem. Phys **156**, 054104 (2022).
4. G. Wang, B. Kincaid, H. Zhou, A. Annaberdiyev, C. Bennett, J. Krogel, L. Mitas, *"A new generation of effective core potentials from correlated and spin-orbit calculations: selected heavy elements"*, Submitted to J. Chem. Phys. (2022).
5. M. C. Bennett, J. Ahn, A. Pham, G. Wang, P. Ganesh, and J. T. Krogel. *"Magnetic measures of purity for MnBi2Te4"*, Submitted to Chemistry of Materials (2022).

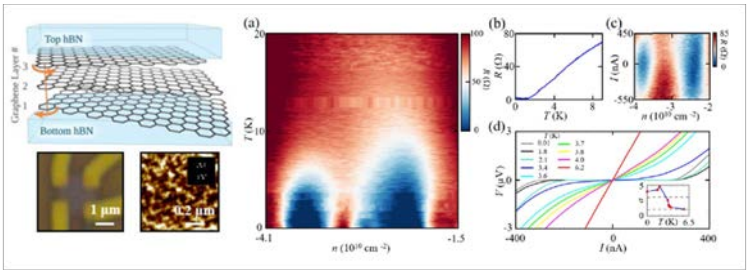
DMREF: Collaborative Research: The Search for Novel Superconductors in Moiré Flat Bands

Lead Investigator: Philip Kim (philipkim@g.harvard.edu), Mitchell Luskin (luskin@umn.edu)
Co PI: Efthimios Kaxiras (kaxiras@g.harvard.edu), Ke Wang (kewang@umn.edu, presenting in-person poster)
Participating Institutions: Harvard University, University of Minnesota **Website:** None
Keywords: Moiré Superlattice, 2D Material, Graphene, Superconductivity, Correlated Physics

Project Scope: The major goals of this project are to develop and investigate novel quantum behaviors in various twisted moiré systems based on graphene and transition metal dichalcogenides. The specific aims includes (1) Fabrication of mesoscopic crystalline superconductor samples in the 2D limit, with gate tunable superconductivity. (2) Developing theory and mathematical modeling of multi-scale electronic structure calculation for multi-layer van der Waals (vdW) systems. (3) Building a closed loop between experiments and theoretical modeling for a variety of choices for TMDs and targeted twist angle in these systems.

.

Relevance to MGI: PIs Luskin (University of Minnesota Math), Kaxiras (Harvard Physics), Kim (Harvard Physics), and Wang (University of Minnesota Physics) have been utilizing their combined expertise to develop and investigate novel superconductivity with unusual properties, such as gate tunable transition temperature and non-conventional pairing symmetries in moiré flat bands based on various vdW homo/hetero junctions. Kim and Wang fabricate state-of-the-art vdW samples, exploring a wide range of material combinations and experimental configurations guided by the theoretical modeling of Luskin and Kaxiras. Samples are comprehensively characterized via close collaboration and streamlining between Kim and Wang. Kim characterizes their structural and thermal properties via TEM and RF thermometry, while Wang characterizes electrical and magneto transport properties with atomically flat nano-gating in a dilution refrigerator. Comprehensive experimental data and analysis are frequently compiled, exchanged and discussed with Luskin and Kaxiras for benchmarking, which in-turn will provide critical materials parameters to further improve and complete the theoretical model.



Superconductivity in twist trilayer graphene Moiré of Moiré superlattices. Left panel: Schematic of the hBN-encapsulated tTLG device. Three individual pieces of monolayer graphene are transferred on top of each other, consecutively twisted by an angle with respect to the previous layer. **(a-d)** temperature and bias characteristics of superconducting-like zero-resistance states, at a record carrier density. The results were published in Phys. Rev. Lett as editors' suggestions in 2021.

Technical Progress: As an initial demonstration for tunable superconductivity in a moiré flat band, Kim started experimental study on twisted double bilayer graphene (tDBLG). The electronic band structure of Bernal stacked bilayer graphene (BBLG) can be tuned by an electric displacement field applied to the normal direction of the basal plane. Due to inversion symmetry breaking in bilayer graphene with vertical electric fields, an energy gap opens up at the charge neutrality point whose size can be tunable by applied electric fields. At the band edges of BBLG, a Mexican hat-type of dispersion relation can be realized due to the trigonal warping effect, leading to diverging density of states due to van Hove singularity associated with 1D rim of the Mexican hat potential. Upon applying displacement fields in the tDBG, the band gaps and shape of the band can be tuned, and even flatter bands can be realized by adjusting the applied displacement field. Wang has improved the fabrication recipe for twisted trilayer graphene (tTLG) samples, which allowed assembly of more advanced and specialized tTLG vdW structures that can be studied by AFM and TEM. Wang has imaged the topography of the tTLG moiré of moiré superlattice, to

supplement the low-temperature transport data that has demonstrated superconductivity in tTLG. Wang has developed a novel nano-device architecture and fabrication protocol, to allow gate-defined quantum contact for gapless materials such as graphene, with demonstration of continuously tunable transmission, reflection and hybridization rate of quantum Hall edge states at a strain and electrostatically engineered nano constriction. Kim and Wang have worked closely on transmission electron microscopy (TEM) investigation of the tTLG and tDBLG. We studied the change of the atomic structure and electronic properties of the twisted van der Waals (vdW) interface in engineered graphene heterostructures. In the past year, our team has successfully identified in tTLG both experimentally and theoretically, the emergence of unconventional superconductivity with correlated electrons, yielding several times higher transition field than previously reported tBLG systems.
Using first principles density functional theory (DFT) calculations, we have studied the effects of the twist angle and interface chemistry on the properties of various twisted bilayers and twisted trilayers. Due to the small bandwidth of these bands, correlations are likely to dominate, making these systems enticing platforms for studying strongly correlated physics. Using our continuum relaxation model for alternatively twisted trilayer graphene, we found that relaxation in the system forms three distinct types of domains, twistons, solitons and magic moirés, on the moiré of moiré scale by redistributing the local twist angle, which agrees with the Scanning Tunneling Microscopy data. Using our electronic structures model for multi-layered moiré of moiré van der Waals heterostructures, we show that the four-layer system of AB-bilayer graphene sandwiched by top and bottom Boron Nitride exhibits low-energy bands that are intrinsically layer polarized. In addition, we found that the layer polarization is reversible under electric field. This finding provides theoretical support for the observed novel material properties in this system.

Future Plans:We will (1) Continue to search for exotic Mott insulator and superconducting states in twisted bilayer/trilayer systems based on homo and hetero-structures of varieties of vdW materials. (2) Continue to develop and demonstrate gate-defined nanostructures for Moiré systems, towards gate-tunable coherent quantum electronic devices. (3) Continue to develop and improving new theoretical models for twisted bilayer and trilayer systems, guiding experimental design and interpreting new experimental findings.

Broader Impacts and Workforce Development:The PI tightly integrates the education and outreach program with the research in the frontier of low-dimensional materials systems. The PI/co-PI have been training graduate and undergraduate students in quantum matter research at the interface of experimental and theoretical science. This effort benefits students in their career development in both academia and industry. We also maintain a very tight outreach through the existing network setup via Harvard SEAS and University of Minnesota. DMREF students, postdocs and faculty will participate in the event to inform and engage the public through discussions of the advanced nanotechnologies developed by our team and its effect on device technologies.

Data Management and Open Access: The accumulation of data is computerized and the conditions under which the data was taken is recorded in electronic notebooks. The standards for data format and content are in accordance with the conventional scientific methodology where the mean measured values are reported and the uncertainties in these values either directly reported or discussed. All collected data is archived and described in various notebooks that are stored in the lab. The samples used during measurements are also stored in proper storage cabinets such as desiccators. The data generated in this research effort are published in research journals and presented at seminars, colloquia and conferences. Upon request, raw data may be available to anyone who desires it. The role and contribution of each author will be described in accordance with each specific journal guidelines and will be available upon request from the PI as well.

Advancing Along the Materials Development Continuum and Partnerships to Translation: The discovery of novel superconductors often requires quantum fluctuation, electron correlations, quasi-two dimensionality, and geometrical frustration of magnetic couplings. The development of the new electronic devices based on moiré structures thus opens important directions in both fundamental and applied research utilizing these new types of materials tuned by global and nano-patterned gating. The research exploits recent technological advances in the synthesis of low dimensional materials and nanodevice fabrications. The present study has features which are of interest to both the experimental and theoretical condensed matter communities and the applied mathematics community, and nanoscale transport and device research community. Our project presents a novel transport study

of an exotic Mott insulator and superconducting states, which is of fundamental importance to the condensed matter and nanoscience research communities, inviting further work. The gate-tunable quantum transport of electrons in these novel materials systems pave a path towards future nano-electronic devices.

Publications and References

Correlated Insulating States and Transport Signature of Superconductivity in Twisted Trilayer Graphene Superlattices, Xi Zhang, Kan-Ting Tsai, Ziyan Zhu, Wei Ren, Yujie Luo, Stephen Carr, Mitchell Luskin, Efthimios Kaxiras, and Ke Wang Phys. Rev. Lett. 127, 166802 (2021)

Lithium intercalation in bilayers and implications for moiré flat bands, Zhuyu Lu, Stephen Carr, Daniel T. Larson, Efithimios Kaxiras, PHYSICAL REVIEW B 102, 125424 (2020)

Moiré metrology of energy landscapes in van der Waals heterostructures, Dorri Halbertal, Nathan R. Finney, Sai S. Sunku, Alexander Kerelsky, Carmen Rubio-Verdú, Sara Shabani, Lede Xian, Stephen Carr, Shaowen Chen, Charles Zhang, Lei Wang, Derick Gonzalez-Acevedo, Alexander S. McLeod, Daniel Rhodes, Kenji Watanabe, Takashi Taniguchi, Efthimios Kaxiras, Cory R. Dean, James C. Hone, Abhay N. Pasupathy, Dante M. Kennes, Angel Rubio, and D. N. Basov, Nature Communications volume 12, Article number: 242 (2021)

Twisted Trilayer Graphene: A Precisely Tunable Platform for Correlated Electrons, Ziyan Zhu, Stephen Carr, Daniel Massatt, Mitchell Luskin, and Efthimios Kaxiras, PHYSICAL REVIEW LETTERS 125, 116404 (2020)

Electronic-structure methods for twisted moiré layers, Stephen Carr, Shiang Fang, Efthimios Kaxiras, Nature Review Materials 5, 748-763 (2020)

Modeling mechanical relaxation in incommensurate trilayer van derWaals heterostructures, Ziyan Zhu ,Paul Cazeaux, Mitchell Luskin, and Efthimios Kaxiras, PHYSICAL REVIEW B 101, 224107 (2020): DOI: 10.1103/PhysRevB.101.224107

Duality between atomic configurations and Bloch states in twistronic materials, , Stephen Carr, Daniel Massatt, Mitchell Luskin, and Efthimios Kaxiras, PHYSICAL REVIEW RESEARCH 2, 033162 (2020).

High-Bandwidth, Variable-Resistance Differential Noise Thermometry, A. V. Talanov, J. Waissman, T. Taniguchi, K. Watanabe, P. Kim, arXiv:2008.12739, submitted to Review of Scientific Instrument.

Comscope: Center for Computational Material Spectroscopy and Design

Lead Investigator: Gabriel Kotliar, Kotliar@physics.rutgers.edu.
Participating Institutions: Brookhaven National Laboratory, Rutgers University, LANL, University of Virginia, Ames National Laboratories
Website: https://www.bnl.gov/comscope
Keywords: strong correlations, quantum embedding methods, dynamical mean field theory, theoretical spectroscopy

Program Scope

Strongly correlated electron materials are at the core of important technologies, but until recently it was not possible to compute their properties from first principles. The aim of our DOE BES Computational Materials Science Center, Comscope, is to develop methodologies and computational platforms that will do for strongly correlated materials what DFT has done for less correlated systems. Our vision is to provide scientists tools based on quantum embedding methods such as DMFT or RISB combined with electronic strucutre methods such as GW and vertex corrected GW to accurately explore the space of correlated material with scant or no experimental input.

Relevance to MGI

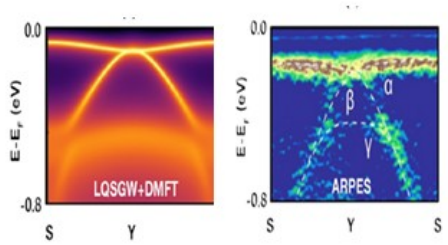


Fig. 1: Photoemission spectra of the high-performance low temperature thermoelectric FeSb2, theoretical prediction (left) and experimental verification (right) from Ref. 3.

The COMSCOPE software platform has been developed with the goal to make material design possible in the class of strongly correlated electron systems, as described in Ref. [1] and theory-experiment interactions following the MGI philosophy. We use FeSb2, a material which holds the world record of thermoelectric power at low temperature, as an example. Predictions of its electronic properties were carried out, with various methods. The predictions of qsGW+DMFT, our most accurate tool described well the optical conductivity of this material [2]. Theory then predicted that the momentum resolved electronic spectral function and revealed unique features such as incoherent spectral weight and high degeneracy near the Fermi level. Armed with these predictions the center carried out the first angle-resolved photoemission experiments on this material verifying key aspects of the theoretical predictions [3] as shown in Fig. 1. This study provides both a validation test for the qsGW+DMFT code module and a demonstration of why its development was necessary. Insights from theory and computation then led to further understanding and more experimental discoveries of large thermoelectric response in this class of systems [4][5].

Technical Progress

Comscope's packages, Comsuite [6][7] and FlapwMBPT [8], offer complementary approaches to tackling the problem of computing the behavior of correlated materials. The first of these, Comsuite, focuses on cases where correlations are primarily local, implementing three flavors of ab initio dynamical mean field theory in order of increasing computational complexity: LDA+RISB, LDA+DMFT, and GW+DMFT. Comsuite uses FlapwMBPT for its LDA and GW capabilities and ComCTQMC and ComRISB as a quantum impurity solver. FlapwMBPT is an all-electron LAPW basis first-principle code that permits GW+Vertex computations in order to treat non-local correlations beyond GW. As an important milestone, a fully self-consistent GW+EDMFT code which addresses a limitation of our earlier lqsGW+DMFT package will be presented in our poster together with tests on several materials.

Future Plans

A release of fully self-consistent GW+DMFT and further integration and simplification of all the center codes are planned to improve accessibility to a broader community of developers. To expand our community of users and increase the awareness of the new possibilities open by the quantitative treatment first-principle treatments of strongly correlated electron materials a new school is planned for the fall of 2022. We also plan accelerations of

key parts of the code using machine learning (ML) tools and more extensive use of GPU hardware. We will expand our functionalities to appeal to new users by adding new theoretical spectroscopies such as ARPES computations with full inclusion of matrix elements.

Broader Impacts and Workforce Development

Our center participates in training the next generation workforce in the possibilities to treat strongly correlated through many schools, the most recent of which took place in June 2022 in Jouvance Canada where COMSCOPE's software was presented including hands on training. The center contributes highly trained postdocs that continue their careers in National Laboratories such as BNL, Universities such as George Mason University, and technology companies such as Micron Technology Inc.

Data Management and Open Access

Our papers and preprint are broadly available in the arXiv and on the OSTI website. The center makes new state of the art open source codes available to the community in two ways: 1) We publish in dedicated peer reviewed scientific journals, such as Computer Physics Communications, as for example, references [6][7][8][9], and 2) our codes are then maintained and continuously supported with new downloadable versions from the center website https://www.bnl.gov/comscope/

Advancing Along the Materials Development Continuum and Partnerships to Translation

To translate very sophisticated theoretical and computational tools designed for complex problems such as strongly correlated electron materials to the broader material science community we use standard tools such as websites, publications, and schools. We have also used a direct collaborative process which we exemplify with our interactions with the experimental X ray spectroscopy community. EDRIX, was initially used in collaborative mode with experimentalists resulting in joint publications such as [9] [10]. It continued to be improved and made more user friendly so that it can be used by users without extensive theoretical or computational background and new versions added at the light source NSLS2 https://github.com/NSLS-II/edrixs. Recently it has been used in scientific efforts not funded by the center as for example the investigation of nickel analogs of cuprates [11]. We plan to follow this collaborative path to collaboratively reach the broad material science community.

Publications and References

1. Ran Adler, Chang-Jong Kang, Chuck-Hou Yee and Gabriel Kotliar, Correlated materials design: prospects and challenges Reports on Progress in Physics **82**, 1 012504 (2019).
2. C. C. Homes, Q. Du, C. Petrovic, W. H. Brito, S. Choi, and G. Kotliar, Unusual electronic and vibrational properties in the colossal thermopower material FeSb2, Scientific Reports **8**, 11692 (2018).
3. A. Chikina, J.-Z. Ma, W. H. Brito, S. Choi, P. Sémon, A. Kutepov, Q. Du, J. Jandke, H. Liu, N. C. Plumb, M. Shi, C. Petrovic, M. Radovic, and G. Kotliar, Correlated electronic structure of colossal thermopower FeSb2: An ARPES and ab initio study, Physical Review Research **2**, 023190 (2020).
4. Qianheng Du, Milinda Abeykoon, Yu Liu, G. Kotliar, C. Petrovic, Low-Temperature Thermopower in CoSbS, Phys. Rev. Lett. **123**, 076602 (2019).
5. Qianheng Du, Lijun Wu, Huibo Cao, Chang-Jong Kang, Christie Nelson, Gheorghe Lucian Pascut, Tiglet Besara, Kristjan Haule, Gabriel Kotliar, Igor Zaliznyak, Yimei Zhu, and Cedomir Petrovic, Vacancy defect control of colossal thermopower in $FeSb_2$, npj Quantum Materials **6**, 13 (2021).
6. S. Choi, P. Semon, B. Kang, A. Kutepov, and G. Kotliar, ComDMFT: a Massively Parallel Computer Package for the Electronic Structure of Correlated-Electron systems, Comp. Phys. Comm. **244**, 277 (2019).
7. Corey Melnick, Patrick Sémon, Kwangmin Yu, Nicholas D'Imperio, André-Marie Tremblay, Gabriel Kotliar, Accelerated impurity solver for DMFT and its diagrammatic extensions, Comp. Phys. Comm. **267**. 108075 (2021).
8. AL Kutepov, VS Oudovenko, G Kotliar, Linearized self-consistent quasiparticle GW method: Application to semiconductors and simple metals, Computer Physics Communications **219**, 407 (2017).
9. Y.L. Wang, G. Fabbris, M.P.M. Dean, and G. Kotliar, EDRIXS: An open source toolkit formulating spectra of resonant inelastic x-ray scattering, Comp. Phys. Comm. **243**, 151 (2019).
10. Y. Wang, Ruitang Wang, Jungho Kim, M. H. Upton, D. Casa, T. Gog, G. Cao, G. Kotliar, M. P. M. Dean, and X. Liu2, Direct Detection of Dimer Orbitals in n Ba5AlIr2O11, Physical Review Letters **122** (10), 106401 (2019).
11. Y. Shen. J. Sears, G. Fabbris, J. Li, J. Pelliciari, I. Jarrige, Xi He, I. Božović, M. Mitrano, Junjie Zhang, F. Mitchell, A. S. Botana, V. Bisogni, M. R. Norman, S. Johnston, and M. P. M. Dean, Role of Oxygen States in the Low Valence Nickelate $La_4Ni_3O_8$, Phys. Rev. X **12**, 011055 (2022).

Design and Optimization of Granular Metamaterials using Artificial Evolution

Lead Investigator: Rebecca Kramer-Bottiglio, rebecca.kramer@yale.edu
Participating Institutions: Yale University, University of Vermont
Website: https://www.eng.yale.edu/faboratory/
Keywords: Granular metamaterials, multifunctional materials, evolutionary algorithms, metamaterials

Future metamaterials will exhibit increased dynamic plasticity, enabling responses to different environmental inputs or task demands by reconfiguring their physical structure. Granular metamaterials offer an advantageous platform for such dynamic programmability, as particle properties can be widely tuned to achieve different responses. Granular metamaterial response is dependent on many variables, including the grain arrangement, grain mass, modulus, and shape, friction or other interactions between grains, and boundary conditions of the granular assembly. With such a vast number of possible combinations of micro-structural variables, the task of designing the complex relationship between microstructure and bulk properties is daunting. We are applying evolutionary algorithms to efficiently search the immense parameter space for granular metamaterial designs with specific material properties, as well as identify how a design can be perturbed to push it from one set of desired bulk properties to another.

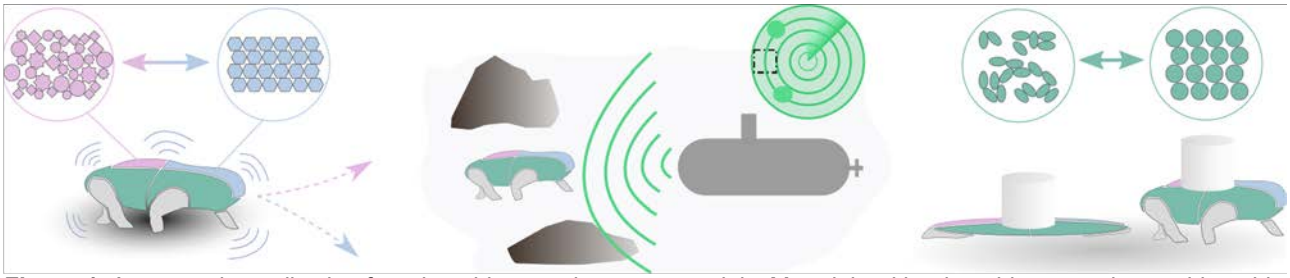


Figure 1. An example application for adaptable granular metamaterials. Materials with adaptable properties could enable a robot with (1) different responses to vibration, which could enable vibration-driven locomotion; (2) acoustic cloaking during covert missions; and (3) tunable bulk stiffness to adjust between compliance and load-bearing tasks.

Project Scope

Our effort seeks to establish a new artificial intelligence-driven approach to the design and optimization of granular metamaterials with adaptable properties. New knowledge and tools to be generated by this project include: (i) New evolutionary algorithms for granular metamaterials capable of editing their own configuration and properties; (ii) New physics-based DEM simulations that can predict the properties of granular materials with active grains; and (iii) New synthesis strategies for multifunctional grains and grain assemblies.

Relevance to MGI

Our project aims to answer two questions: (1) How can we automatically design granular assemblies with specific desired material properties? (*i.e.*, What should the grains' arrangement, moduli, shapes, etc. be to yield a given bulk property?); and (2) How can we automatically design a series of granular assemblies that allow continuous, time-ordered changes in material properties? To answer these questions, we are developing a closed-loop procedure wherein physics-based discrete element simulations, evolution-based optimization, and physical realizations are combined to produce granular metamaterials with desired, optimal, and adaptable bulk properties.

Technical Progress

Our progress to-date includes: 1. The evolution of acoustic logic gates in granular metamaterials [1-2] (Fig. 2). We investigated how to embed Boolean logic gates (AND and XOR) into granular assemblies by evolving configurations of high-mass and low-mass grains. Our results confirm the existence of gradients of increasing "AND-ness" and "XOR-ness" that can be optimized by evolutionary search. We evaluated the computational

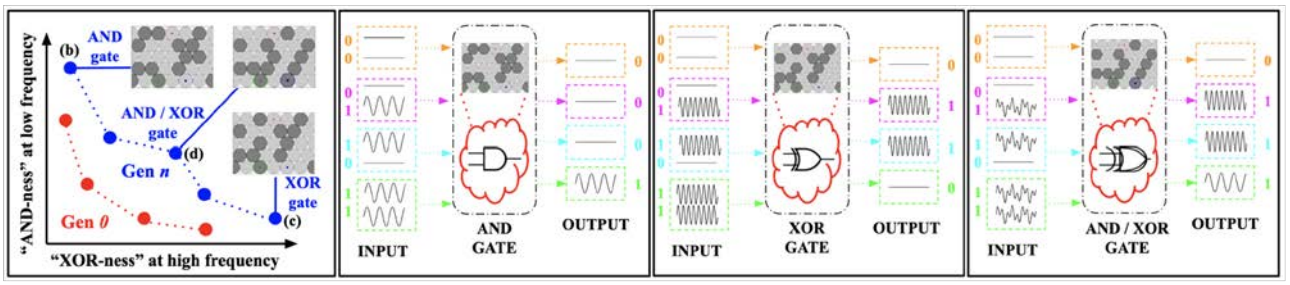


Figure 2. We used multi-objective evolutionary optimization to design basic logic gates embedded in a granular assemblies made from stiffness-contrasting grains. The designs can be "reprogrammed" via frequency modulation.

functionality of evolved assemblies by probing how they transform bits encoded as vibrations with zero or non-zero amplitude. We further compared the evolution of assemblies made of mass-contrasting particles and with assemblies made of stiffness-contrasting particles, and found that the latter were more evolvable.

2. The design of reprogrammable allosteric metamaterials from disordered networks generated from random configurations of jammed soft disks [3]. Prior works on disordered mechanical metamaterial networks—consisting of fixed nodes connected by discrete bonds—have shown that auxetic and allosteric responses can be achieved by pruning a specific set of the bonds from an originally random network. However, bond pruning is irreversible and yields a single bulk response. Using material stiffness as a tunable design parameter, we created metamaterial networks where allosteric responses are achieved without bond removal. We experimentally realized the networks through variable stiffness bonds that can strengthen and weaken on-demand. In a disordered mechanical network with variable stiffness bonds, different subsets of bonds can be strategically softened to achieve different bulk responses, enabling a multiplicity of reprogrammable input/output allosteric responses. We are currently expanding on this result to introduce variable stiffness particles that can adapt their stiffness on-demand, to enable granular metamaterial assemblies with adaptive force chains.

Future Plans

We expect to expand our particle capabilities to include variable stiffness (including addressable stiffness gradients rather than binary stiffness switching), variable size (particle growing and shrinking), and variable shape. As we increase the number of free variables, we expect the simulation environment, the multi-objective search algorithm, and fabrication of the evolved designs to become increasingly complex.

Broader Impacts and Workforce Development

Our primary activity to broaden participation in the design of next-generation adaptable materials is the creation a website, Twitch Plays Metamaterials, through which members of the public can compete or collaborate to make the best metamaterials. This new interactive Twitch channel will build upon the success of our team's prior introduction of Twitch Plays Robotics and Twitch Plays Soft Robotics.

Data Management and Open Access

Digital outputs are publicly available through publications posted on the PIs' websites.

Advancing Along the Materials Development Continuum and Partnerships to Translation

We expect the application of evolutionary algorithms to locally adaptable materials, such as granular assemblies, to enable increasingly sophisticated, programmable, and computationally dense metamaterials.

Publications

1. A. Parsa, D. Wang, C. O'Hern, M. Shattuck, R. Kramer-Bottiglio, J. Bongard. *Evolution of Acoustic Logic Gates in Granular Metamaterials*. EvoStar, Madrid, Spain, 2022.
2. A. Parsa, D. Wang, C. O'Hern, M. Shattuck, R. Kramer-Bottiglio, J. Bongard. *Evolving Programmable Computational Metamaterials*. The Genetic and Evolutionary Computation Conference (GECCO), Boston, USA, 2022.
3. N. Pashine, A. M. Nasab, R. Kramer-Bottiglio. *Reprogrammable allosteric metamaterials from disordered networks*.

Low-Cost Science Robot Kit

Lead Investigator: Aaron Gilad Kusne, aaron.kusne@nist.gov
Participating Institutions: NIST, UMD
Website: none
Keywords: autonomous, education, closed-loop

Project Scope

The next generation of physical science involves robot scientists – autonomous physical science systems capable of experimental design, execution, and analysis in a closed loop. Such systems have shown real-world success for scientific exploration and discovery, including the first discovery of a best-in-class material. To build and use these systems, the next generation workforce requires expertise in diverse areas including ML, control systems, measurement science, materials synthesis, decision theory, among others. However, educators need a low-cost, easy-to-use platform to teach the required skills. Industry can also use such a platform for developing and evaluating autonomous physical science methodologies. We are developing the next generation in science education, a low-cost robot science kit. The kit was used during two courses at the University of Maryland to teach undergraduate and graduate students autonomous physical science.

Relevance to MGI

MGI calls for the integration of experiment, computation, and theory as well as the education of the next generation workforce in the ability to integrate them. This allows students to investigate this integration through a closed-loop system for experiment design, execution, and analysis. Students used this system to solve challenges in 1) learning a relationship between experiment parameters and their results, 2) parameter determination when the model is known, 3) model determination when the model is unknown. The system can be used to develop novel autonomous and scientific machine learning methods.

Figure. A) Photo of the robot science kit, b) The kit's use a University of Maryland course.

Technical Progress

The students were asked to solve two challenges: 1) To write a code and operate the kit, so that it can autonomously identify the relationship between acid to base ratio and the resulting pH using the minimum number of experiments. 2) To modify the code so that the system identifies the acid to base ratio with a pH of 4.5 using the minimum number of experiments. They were also introduced to scientific challenges, such as noisy pH sensor measurements. Each student group was asked to write a Jupyter notebook to address the two challenges. Figure 3 shows a solution to the challenges, where off-the-shelf Gaussian process regression was fit to the data and a) paired with an exploratory active learning acquisition function for Challenge 1 and b) paired with an acquisition function that balances exploration and exploitation for Challenge 2. The two acquisition functions recommend different subsequent experiments.

At the end of the 3-week project, students presented their results and their autonomous strategies and discussed opportunities for future improvements. The students found that with autonomous control, they were able to solve the challenges with less than ten experiments, compared to an exhaustive study which takes dozens of experiments. Similarly, total experiment time fell from hours to minutes.

Three advanced challenges were provided to an undergraduate student to solve. In the first challenge, the student was provided the Henderson–Hasselbalch equation and asked to use the system to find the parameter values in the

minimum number of experiments. The student used Bayesian inference and active learning to guide subsequent experiments and focus in on the correct parameter values.

The second challenge provided to the student was one of hypothesis or model selection. The student was asked to use the system to figure out the underlying model and its parameters if a set of possible models was provided. Here again Bayesian inference and active learning were combined to guide subsequent experiments. Each model in the set was fit to the data and the distribution over parameter values was determined. By combining the distribution over the set of possible models, the student was able to identify the sample ratio at which the models most disagree – the ratio at which entropy is maximum. Bayesian information Criteria is used to select the best candidate model given model fit and complexity.

The final challenge given to the undergraduate was to use the system to determine the mechanistic function if no potential models are given, i.e., search over the space of all possible models (i.e., hypotheses) to determine the best model. Here the student combined symbolic regression with active learning in a closed loop. At each iteration, past data is fit using symbolic regression to identify a set of potential models. These models are quantified by fit quality using mean square error (MSE) and complexity. From these values a score is computed to rank the models. For each model, the MSE provides an estimate of model uncertainty, based on the validated assumption that the measurement noise is normally distributed and heteroskedastic. The next experiment is then selected using the same entropy selection criteria from the previous exercise. The model with the best score was an extremely close fit to the HH equation.

Future Plans

We are working on making the system modular and robust while maintaining low cost.

Broader Impacts and Workforce Development

This project focuses on next-generation workforce development.

Data Management and Open Access

Jupyter notebook developed for the system will be made open access.

Advancing Along the Materials Development Continuum and Partnerships to Translation

This system focuses on accelerating closed-loop autonomous experiment design, execution, and analysis. Universities are interested in the kit for their courses and industry partners are interested in the kit as a platform for demoing new machine learning techniques.

Publications and References

1. Saar, L., Liang, H., Wang, A., McDannald, A., Rodriguez, E., Takeuchi, I., & Kusne, A. G. (2022). A Low-Cost Robot Science Kit for Education with Symbolic Regression for Hypothesis Discovery and Validation. arXiv preprint arXiv:2204.04187.

OOF2 and OOF3D

Lead Investigator: Stephen Langer, stephen.langer@nist.gov
Co-investigators: Andrew Reid, Shahriyar Keshavarz
Participating Institutions: NIST
Website: https://www.ctcms.nist.gov/oof/.
Keywords: microstructure, finite element, image.

Project Scope

The OOF project is developing software for finite element analysis of material microstructures, based on real or simulated images of the microscopic structures, with the aim of predicting the effect of microstructure on material properties. OOF2 works with two dimensional micrographs, and OOF3D with three dimensional ones. A wide range of classical physical phenomena can be modeled, and non-linear properties are under development. The goal is to create a platform accessible to researchers without a background in computational methods, which they can easily extend to add the physics of interest.

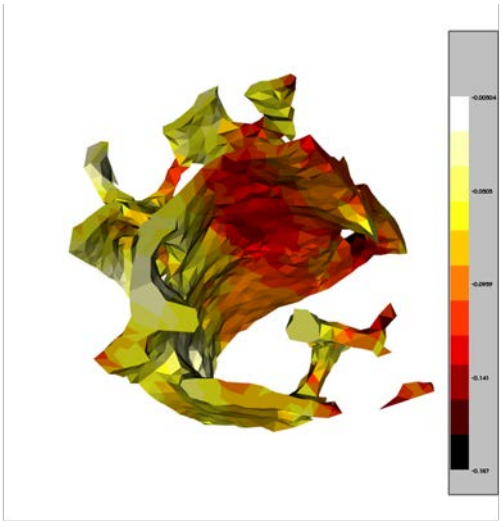


Figure. Heat flux magnitude computed in a geometry derived from a micrograph of bone. (The physical parameters were chosen arbitrarily for demonstration purposes, and aren't meaningful.)

Relevance to MGI

OOF2 and OOF3D facilitate one particular step of the materials design process: linking structure and properties via microstructure. Users can change physical and geometrical parameters and observe their effects on behavior. OOF occupies a spot in the toolset that is opposite and complementary to homogenization methods, since an OOF calculation applies to one particular realization of a microstructure. However, OOF is easily automated for application to a set of microstructures.

In a real-world example, OOF was used by a mechanical testing group at NIST, where it was able to meet the group's need for an unusual boundary condition to model scattering from polycrystalline steel under load. The calculations matched some previously unexplained experiments.

Technical Progress

Recent progress includes:

- A framework for crystal plasticity and large deformations was added.
- OOF2 and OOF3D depend on external libraries that are themselves evolving, such as Python, GTK, and VTK. Updating OOF2 and OOF3D for python3 and gtk3 is in progress.
- Several parts of OOF should be useful in other contexts, and either are or soon will be available separately:
    - A surprisingly tricky and error-prone part of the calculation of the intersection of a finite element and a set of pixels (or voxels in 3D) has been made robust, through a graph-based algorithm.
    - No substitute for the GTK 2 graphics canvas used in OOF2 is available in GTK 3, so a separate library, OOFCanvas was written for it.
    - The gui-testing infrastructure that was developed for OOF and GTK 2 is available now for GTK 3.

Future Plans

The crystal plasticity and large strain capabilities will be completed, including using machine learning to accelerate the expensive plasticity calculation. OOF2 and OOF3D will be moved from Python 2 to Python 3. New methods for meshing images will be investigated. The physical models will be expanded, including adding interface physics which is currently a glaring omission. Uncertainty quantification could be included. OOF can be used within Python scripts, but explicit interfaces with other MGI components should be developed.

Broader Impacts and Workforce Development

OOF2 has been used in educational settings. Both OOF2 and OOF3D are available at nanoHUB.

Data Management and Open Access

The OOF project does not produce any data itself. The programs are freely available at https://www.ctcms.nist.gov/oof. Specifically:

OOF2: https://www.ctcms.nist.gov/oof/oof2/ -— 2 dimensional microstructure analysis
OOF3D: https://www.ctcms.nist.gov/oof/oof3d/ — 3 dimensional microstructure analysis
OOFCanvas: https://www.ctcms.nist.gov/oof/oofcanvas/ — 2D GTK 3 compatible graphics library
VSB: https://www.ctcms.nist.gov/oof/vsb/ — Robust voxel/tetrahedron intersection calculation
gtklogger: https://www.ctcms.nist.gov/oof/gtklogger/ — GUI recording, replaying, and testing for GTK 3.

Advancing Along the Materials Development Continuum and Partnerships to Translation

OOF is often used to generate input meshes for commercial finite-element programs, and it would make sense for the developers of those codes to incorporate some of OOF's algorithms directly. Because of its ease of use, expandibility (open source, modular design), and cost ($0), OOF will continue to be valuable as a stand-alone product.

Publications and References

1. S. A. Langer and A.C.E Reid, *Calculating Voxel-Polyhedron Intersections for Meshing Images*, Proceedings of the 29th International Meshing Roundtable, p 14 (2021)
2. S. Keshavarz, Z. Molaeinia, A. C. E Reid, and S. A. Langer, *Morphology Dependent Flow Stress in Nickel-Based Superalloys in the Multi-Scale Crystal Plasticity Framework*, Crystals **7**, p 334 (2017)
3. J. L. Molaro, S. Byrne, S. A. Langer, *Grain-scale thermoelastic stresses and spatiotemporal temperature gradients on airless bodies, implications for rock breakdown: Thermoelastic stresses on airless bodies*, Journal of Geophysical Research: Planets **120**, p 255 (2015).
4. V. R. Coffman, A. C. E. Reid, S. A. Langer, and G. Doğan, *OOF3D: An image-based finite element solver for materials science*, Mathematics and Computers in Simulation **82**, p 2951 (2012).
5. A. C. E. Reid, R. C. Lua, E. García, V. R. Coffman, and S. A. Langer, *Modelling Microstructures with OOF2*, International Journal of Materials and Product Technology **35**, p 361 (2009).

Interface structure prediction from first-principles: Atomically thin interlayer enables defect-free incommensurate SnO_2/CdTe interface

Lead Investigator: Stephan Lany and Abhishek Sharan; National Renewable Laboratory, Golden, CO 80401

Once coined a “scandal”, the difficulty to predict crystal structures from first principles has been largely overcome in over the past 1-2 decades, and bulk crystal structure prediction has played a major role in research within the materials genome initiative. However, atomic structure prediction for general, incommensurate interface structures with an ambiguous atomic registry remains challenging, and explicit atomic structure models rarely exist. At the same time, this problem is of high importance because such models are prerequisite for first-principles predictions of interface properties. Here, we adopt our bulk crystal-structure prediction method to the interface geometry and apply it to the incommensurate SnO_2/CdTe heterojunction without and with addition of $CdCl_2$. Treatment with $CdCl_2$ is a beneficial and ubiquitous process step, which, however, lacks a mechanistic understanding. Employing our interface structure prediction approach, we discover a unique 2-dimensional $CdCl_2$ interphase, unrelated to the respective $CdCl_2$ bulk crystal structure, which facilitates the seamless transition from the rutile to zinc-blende lattices. Using band gap corrected electronic structure predictions for these large scale slab supercells containing more than 200 atoms and a vacuum layer, we calculate the interfacial density of states. Whereas the direct SnO2/CdTe interface is highly defective and causes a band gap reduction by almost 1 eV, the CdCl2 passivated interface is defect-free with an essentially ideal band alignment. Our work demonstrates progress in crystal structure prediction from first principles, leaping from bulk crystals to interfaces. Beyond the implications for photovoltaics and the specific materials system studied, our work has a broader impact by highlighting the exciting possibility to design atomically thin interlayers that enable defect-free incommensurate interfaces.

High-Precision Structural Measurements

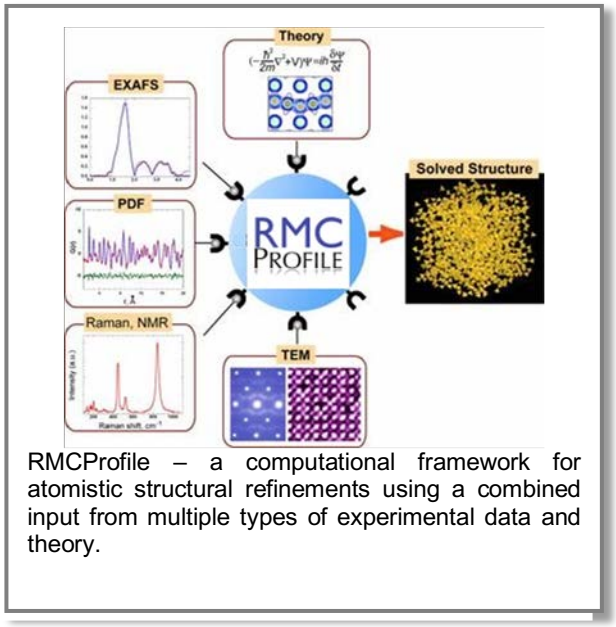
Lead Investigator: Igor Levin, igor.levin@nist.gov
Participating Institutions: NIST
Website: https://www.nist.gov/programs-projects/measurement-and-prediction-local-structure
Keywords: local structure, data fusion, reverse Monte Carlo

Project Scope

This project develops computational algorithms and computer software for determining high-fidelity models of atomic arrangements in complex materials using combined inputs from multiple experimental techniques. The new tools are implemented in the RMCProfile software (www.rmcprofile.org), which permits refinements of large-scale atomic configurations representing the instantaneous structure of a material. The atomic coordinates are determined by fitting various types of X-ray and neutron scattering data, including diffuse scattering, together with extended fine structure in X-ray absorption spectra. Such an approach provides internally consistent descriptions of atomic order over the scales ranging from sub-nanometer to tens nanometers directly from experimental data.

Relevance to MGI

Atomic arrangements underpin the properties of all solids and, as such, are fundamentally important for materials development. The power of computational materials science depends on effective models, which in turn require accurate data on the atomic order. This input is critical to the successful design and accelerated commercial deployment of new and improved materials and targeted functionalities, which is the primary goal of the Material Genome initiative. Practical materials tend to be complex, and detailed structural determination is often challenging, particularly with three-dimensional fidelity and precision across the sub-nanometer to macroscopic length scales, which are needed to correlate structure and function. Some of the existing limitations are associated with instrumental and data-processing deficiencies, whereas others are caused by the inability to integrate data from different experimental sources. Our project seeks to address these issues. In addition to their own value in supplying unprecedently detailed multiscale information on atomic order, the experimentally-derived models provide references for theoretical simulations.



RMCProfile – a computational framework for atomistic structural refinements using a combined input from multiple types of experimental data and theory.

Technical Progress

Recent highlights include several new capabilities for the RMCProfile software. First, we modified the implementation of the instrumental resolution correction for total neutron scattering, enabling the treatment of arbitrary-shape resolution functions. This capability is required for analyzing data from diffractometers at the Spallation Neutron Source, which are the only neutron instruments in the US optimized to characterize disordered materials. Second, we continued developing capabilities for structural refinements using simultaneous fitting of X-ray/neutron scattering from powders and single crystals. We performed this development while solving a highly complex structure of a classic uniaxial ferroelectric, strontium barium niobate (SBN), which remained debatable for decades. SBN contains assemblages of twin domains, up to ten nanometers in size, having atomic displacements modulated along orthogonal directions with a periodicity incommensurate with that of the underlying lattice. We implemented an algorithm for fitting 3D distributions of diffuse scattering from single crystals that contain nanoscale twins. The new capability helped uncover details of hitherto inaccessible correlated atomic displacements, providing new insights into structure-property relationships of these technologically relevant

systems. Our results also suggest ideas for designing new antiferroelectrics. Finally, we implemented restraints on the refined atomic coordinates, which help regularize the resulting atomic configurations by limiting values and symmetry of atomic displacement parameters (akin to atomic probability density distributions) in large-scale atomic configurations according to the average-structure data obtainable from traditional Bragg diffraction. The new capability proved critical in determining the low-temperature structure of a classic perovskite compound $BaZrO_3$, a promising candidate for fuel-cell and hydrogen-separation applications. Despite multiple theoretical studies, its ground state, as suggested by first-principles calculations, remained uncertain, with conflicting claims by leading theory groups. We settled this debate by combining atomistic structural refinements with electron diffraction at cryogenic temperatures to present conclusive evidence for a phase change in $BaZrO_3$ below 80 K associated with a structural distortion. This study, which calls for revising some of the theoretical approximations used to model such systems, illustrates the usefulness of our methods and tools for the development of simulations.

Future Plans

We plan to combine the advantages of the recently developed fast algorithms for calculations of single-crystal diffuse scattering in multiple-technique refinements implemented in RMCProfile. The new algorithms would need to be adapted to the RMCProfile philosophy and tested for their adequacy in systems with a complex displacive disorder. The outcome will be a much higher power version of RMCProfile taking full advantage of modern measurement and computing resources and capable of combining powder and single-crystal data from both X-ray and neutron sources to provide comprehensive high-fidelity structural models.

Broader Impacts and Workforce Development

We actively participate in the annual US Schools on Total Scattering, where we educate students and postdoctoral researchers on the approaches and software that we develop. These oversubscribed events, which showcase state of the art in local structure determination and provide hands-on tutorial sessions on several major software packages, including RMCProfile. The last school was virtual (https://conference.sns.gov/event/301/) with this format offering opportunities to engage leading experts from all over the world.

Data Management and Open Access

The RMCProfile package is developed jointly by several institutions, with the NIST and ORNL teams currently being the main contributors. The RMCProfile software and the auxiliary analysis tools that we develop are available to users via the www.rmcprofile.org website which is maintained by the ORNL group. Some tools that we developed for data reduction are available from the NIST website https://www.nist.gov/mml/materials-measurement-science-division/materials-structure-and-data-group (see the Software section).

Advancing Along the Materials Development Continuum and Partnerships to Translation

Establishing structure-property relations in complex systems often takes decades, with research communities going through multiple iterations partly because of the difficulties with determining all critical structural features. We envision our methods and software to shorten this path by providing comprehensive structural models that simultaneously reproduce several types of experimental data, eliminating ambiguities associated with single-technique analyses or qualitative comparison of results from several methods. As demonstrated by the $BaZrO_3$ example, our results also facilitate the development of accurate theorical simulations.

Publications and References

1. I. Levin, M. G. Han, H. Y. Playford, V. Krayzman, Y. Zhu, R. A. Maier, "Nanoscale-correlated octahedral rotations in $BaZrO_3$," Rhys. Rev. B., 104 (21) 214109 (2021)
2. I. Levin, V. Krayzman, H. Y. Playford, J. C. Woicik, R. A. Maier, Z. Lu, A. Bruma, M. Eremenko, M. G. Tucker, "The mediation of bond strain by vacancies and displacive disorder in A-site-deficient perovskites," Acta Mater., 2021/4/1. 116678 (2021)
3. Y. Zhang, M. Eremenko, V. Krayzman, M. G. Tucker, I. Levin, "New capabilities for enhancement of RMCProfile: instrumental profiles with arbitrary peak shapes for structural refinements using the reverse Monte Carlo method," J. Appl. Cryst., 53 [6] 1509-1518 (2020)

DMREF: Machine Learning Accelerated Design and Discovery of Rare-earth Phosphates as Next Generation Environmental Barrier Coatings

Lead Investigator: Jie Lian, Mechanical, Aerospace & Nuclear Engineering, Rensselaer Polytechnic Institute, lianj@rpi.edu; Liping Huang, Materials Science & Engineering, Rensselaer Polytechnic Institute, Huangl5@rpi.edu; Suvranu De, Mechanical, Aerospace & Nuclear Engineering, Rensselaer Polytechnic Institute, des@rpi.edu

Website: none

Keywords: Multicomponent rare-earth phosphate, environmental barrier coating (EBC), ceramic-matrix composite (CMC), steam oxidation and CMAS corrosion, microstructure design

Project Scope

A data-driven machine learning (ML) approach will be developed to inform materials design and guide materials performance evaluation to discover new rare-earth phosphates (Fig. 1). The configurational disorder and microstructure control of the designed multicomponent rare-earth phosphate can lead to enhanced performance against steam oxidation and CMAS interaction. Experimentally well-calibrated models from the atomic to the continuum level will be used to sample the large design parameter space to provide reliable inputs and microstructure descriptors for training the element-based and microstructure-based ML models. This data-driven materials design approach can accelerate design and discovery of novel rare-earth phosphates as next generation EBCs.

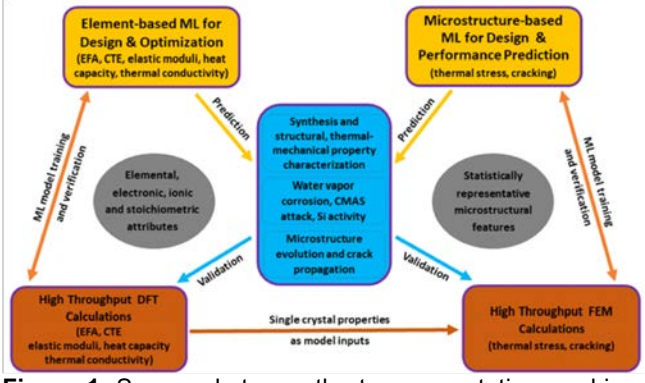


Figure 1. Synergy between the two computation-machine learning-experiment feedback loops for the accelerated design and discovery of rare-earth phosphate-based EBCs in this research program.

Relevance to MGI

This program will apply ML tools to guide extensive computational and experimental search strategies toward promising candidate materials and therefore accelerate the design and discovery of $REPO_4$ with cation complexity and microstructure control as the next generation EBCs for SiC-based ceramic matrix composites (CMCs). An element-based ML will be trained on high throughput density functional theory (DFT) calculations and will be used to guide the design and optimization of configurationally disordered rare-earth phosphates with key characteristics of EBCs that include high structural stability, desired thermal-mechanical properties, well-matched coefficient of thermal expansion with CMCs, excellent oxidation resistance against water vapor and molten calcium-magnesium aluminosilicate (CMAS), and chemical compatibility with the Si bond coat. A microstructure-based ML will be trained on high throughput finite element method (FEM) calculations and will be used to predict the optimal microstructure and performance of EBCs against CMAS corrosion at elevated temperatures. The optimization of the compositions in designing multicomponent rare-earth phosphates and the prediction of microstructure for optimized materials performance will be validated by experimentations. These iterative feedback loops among computation, machine learning and experiment will provide a pathway for success in accelerating the design and discovery of rare-earth phosphates as the next generation EBCs for CMCs.

Technical Progress

During the first year of this DRMEF program, our efforts are mainly focused on the methodology development of high throughput multiscale computation approaches in designing multicomponent phosphate ceramics and predicting their properties, and the experimental demonstration of materials synthesis and their performance evaluation. On experiments, multicomponent rare-earth phosphates based on empirical size disorder and thermal expansion coefficient matched with SiC have been synthesized and characterized. CMAS interaction with the multicomponent rare-earth phosphates is being investigated as functions of durations and temperatures, as compared to single component phosphates. Microstructures of the as-synthesized phosphates and upon CMAS interaction

including grain size, distribution and orientation, and CMAS penetration depths and interfacial reactions have been characterized as initial input data to build up 3D synthetic microstructure for the meso-scale FEM modeling. Thermal-mechanical properties of the model systems including thermal conductivity, thermal expansion coefficients and elastic properties are being measured that can be used to validate the macroscopic properties prediction by FEM. On multi-scale computations, algorithms are being developed to generate synthetic microstructures based on experimentally-obtained microstructures and simulation of generated microstructure for materials properties estimation. Experimental EBSD data is processed and fed into DREAM 3D software to generate 3D synthetic statically equivalent microstructure of the single component and multicomponent rare-earth phosphates. Abaqus files based on model output from DREAM 3D are generated to calculate the macroscopic properties of the various statistically representative synthetic microstructures.

Future Plans

We will continue to develop the integrated computational and experimental approaches based on machine-learning-guided materials design to discover new multicomponent rare-earth phosphates with optimized compositions and microstructures and transformational performances as next generation EBCs. In the upcoming year, we specifically plan for:

(1) High throughput DFT calculations to predict the single crystal properties as inputs for the FEM modeling and design of multicomponent phosphates with optimized thermal expansion coefficients that will be demonstrated by experimental synthesis.

(2) Extension of the developed FEM model to high entropy phosphates to simulate their macroscopic properties and predict their performance upon CMAS interaction which can be validated by experimental observations.

(3) Development of the machine learning models: (a) identification of a suitable model based on data set generated from high throughput FEM simulations; (2) training, testing and validation of the ML model; and (3) using the ML model to accelerate the material property estimation from microstructural descriptors.

Broader Impacts and Workforce Development

The program offers a unique opportunity for graduate students and young scientists to work in a dynamic and multidisciplinary environments with complementary expertise from data-driven ML material design and data curation, multiscale computation and modeling, and experimental demonstration and validation across different fields of materials science, aerospace, mechanical and nuclear engineering. Currently, the project supports two postdoc research fellows and two PhD students, where they are educated and trained as the next-generation workforce skilled in the comprehensively integrated experiment-simulation-data analytics approach. In addition, five undergraduate students participated in this project during the Spring 2022 semester and were trained on different aspects of experimentations, FEM modeling and image processing with numerical software and FEM tools such as DREAM 3D, ABAQUS, and Matlab. An integral part of the proposed research is to develop teaching modules that will be integrated into *MTLE 4500: Computational Materials Design*, which is a required upper level undergraduate course for Materials and related majors at RPI. The program will also engage middle and high school students to excite their interests in pursuing STEM through various outreaching programs such as *Engineering Summer Exploration Program at Rensselaer* and *High School Summer Research Program*.

Data Management and Open Access

Currently, the experimental and FEM data obtained are deposited at the RPI box: https://rpi.app.box.com/folder/153319373447 to promote the interaction and collaboration among the team members. Materials data and computational tools developed in this project will be curated and contributed to the MPContribs Portal and accessible via API (for queries) on the Materials Project platform by others to facilitate data-driven material design in general.

Advancing Along the Materials Development Continuum and Partnerships to Translation

This proposed research has the potential to transform the development of EBCs for CMCs by providing the knowledge base/design principles to optimize their structure and properties for reactive engine environments, and this data-driven materials design approach will lead to the accelerated design and discovery of new materials with compositional complexity and configurational disorder. Our collaboration with industry, GE Global Research, will further enhance the technological relevance of our EBCs and speed up their incorporation into potential new applications.

Publications and References

NA

Symmetry-guided Machine Learning for the Discovery of Topological Phononic Materials

Lead Investigator: Bolin Liao, bliao@ucsb.edu
Participating Institutions: University of California Santa Barbara, Massachusetts Institute of Technology
Website: None
Keywords: topological materials, phonon, thermal transport, machine learning

Project Scope

The overarching goal of this project is to use a Materials Genome Approach to discover, simulate, synthesize and characterize novel materials hosting phonon modes with nontrivial topology. These topological phononic materials are ideal testbeds for emerging new physics in topological bosonic systems. Practically, nontrivial phonon topology can provide a new degree of freedom to control phonon scattering and to manipulate thermal and thermoelectric transport properties. We will develop machine-learning based methods to systematically identify topological phononic materials, synthesize and characterize them in a closed-loop fashion. Our study will establish the foundation for understanding the physical consequences of topological phononics states.

Relevance to MGI

Our team will utilize a closed feedback loop approach to identify, simulate, synthesize and characterize novel topological phononic materials. We will first develop a machine-learning based method to accurately predict phonon properties of materials based on their crystal symmetry and chemical composition. For this purpose, we will also curate and establish a phonon property database for stoichiometric crystalline materials included in Materials Project. Machine learning predictions will be verified using first-principles phonon simulation and topological invariance analysis. Promising candidate materials will be synthesized as thin films and bulk single crystals and characterized using inelastic neutron and x-ray scattering, thermal transport, and surface-sensitive spectroscopy and scanning probe measurements. We will also incorporate machine-learning and data science methods to augment the sensitivity of the characterization methods. One recent example is shown in Figure 1, where machine learning is used to extract detailed phonon scattering information from an ultrafast electron diffraction experiment. An important feature of the proposed project are novel characterization techniques, such as those based on scanning ultrafast electron microscopy, which will directly image topological surface phonon modes that have not been experimentally probed so far. Simulation, synthesis and characterization results will inform and improve the machine-learning search algorithms, thus closing the materials discovery loop.

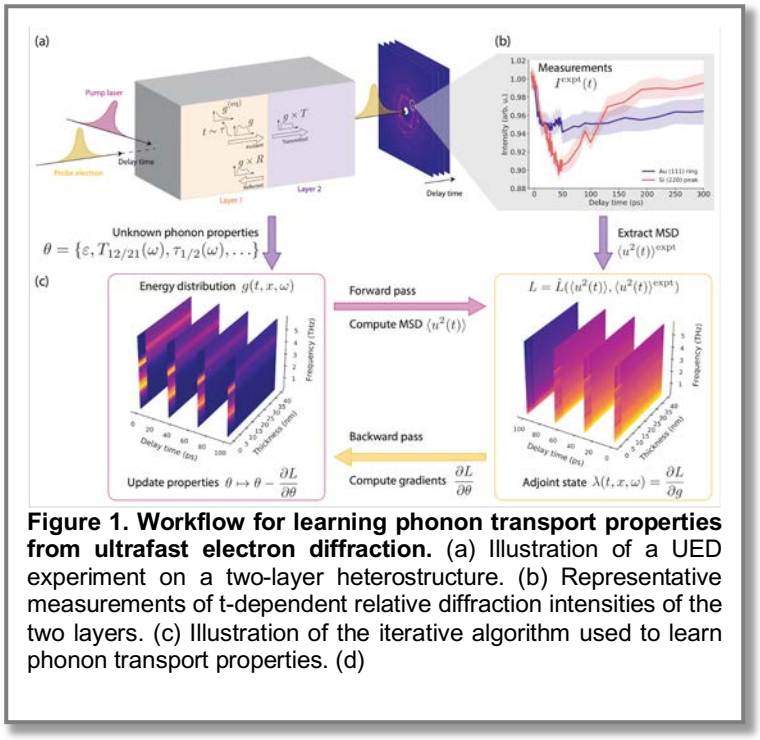


Figure 1. Workflow for learning phonon transport properties from ultrafast electron diffraction. (a) Illustration of a UED experiment on a two-layer heterostructure. (b) Representative measurements of t-dependent relative diffraction intensities of the two layers. (c) Illustration of the iterative algorithm used to learn phonon transport properties. (d)

Technical Progress

The project started in October 2021. In the past 8 months, we achieved the following technical progress: **(1) We demonstrated machine-learning enhanced phonon transport property extraction from ultrafast electron diffraction experiment.** One central challenge in understanding phonon thermal transport is a lack of experimental tools to investigate mode-based transport information. In this work, we established a framework that can reveal

microscopic phonon transport information in heterostructures, integrating state-of-the-art ultrafast electron diffraction with advanced scientific machine learning. With this novel machine-learning-augmented experimental framework, we are able to reliably recover the frequency-dependent interfacial transmittance with possible extension to frequency-dependent relaxation times of the heterostructure. This enables a direct reconstruction of real-space, real-time, frequency-resolved phonon dynamics across an interface. Our work provides a new pathway to experimentally probe phonon transport mechanisms with unprecedented details. This work is now under review and posted to open access arXiv[1]. **(2) We synthesized and characterized a candidate topological phononic material CoSn that hosts flat electronic and phononic bands.** We have successfully synthesized bulk crystals of CoSn, a Kagome compound hosting flat electronic and phononic bands, which can lead to strong and unusual electron-phonon interaction properties. We have conducted preliminary electrical and thermal transport and inelastic neutron scattering measurements and identified an unusual structural phase transition that can be attributed to higher-order electron-phonon interactions. This work represents a closed-loop collaboration by the three PIs: theoretical modeling by Liao, synthesis and scattering measurements by Li, and transport/magnetic measurement by Stemmer. **(3) We have fabricated and conducted preliminary transport measurements of a candidate material strained thin-film Cd_3As_2.** Stemmer group has fabricated thin films of topological Dirac semimetal Cd_3As_2 and we have started preliminary measurements to understand their phonon and thermal transport properties. We have obtained preliminary thermal and thermoelectric transport properties that signal contribution from the topological surface states. Inelastic x-ray scattering measurements have been planned to investigate the properties of the phonon modes. **(4) We have started curating a phonon property database incorporating phonon dispersions and related properties from both literature and our own high-throughput calculations.** We are still in the process of establishing the data infrastructure and hoping to increase the number of entries to ~1000 in the next year.

Future Plans

In the next reporting period, we hope to have a preliminary phonon database that we can use to train a neural network for predicting phonon properties and searching for new topological phononic materials. In the meantime, we will model, synthesize and characterize known topological phononic materials to understand the physical consequences of the topological phononic states.

Broader Impacts and Workforce Development

The project is currently supporting three graduate students, who will gain experience in both traditional material science and data science/machine learning. Liao is also hosting a summer intern undergraduate researcher from an underrepresented minority community. Beyond workforce development, we also aim to develop open access phonon database and open-source software for phonon topology analysis. The PIs also plan to develop new courses or incorporate data science concepts into existing courses on material science.

Data Management and Open Access

The project will generate the following digital objects that will be deposited onto online databases that are openly accessible to the public: (1) a machine-learning model for phonon properties; (2) a phonon property database; (3) a phonon topology analysis software and (4) educational course materials. All experimental and simulation data will also be properly annotated and published together with peer-reviewed journal articles. We will identify the DOIs of the deposited data once the accompanying paper is published online.

Advancing Along the Materials Development Continuum and Partnerships to Translation

In this project, we aim to accelerate the development of phononic materials by incorporating machine-learning methods and high-throughput computation into phononic materials searching and modeling. We hope to establish a phonon database that can significantly accelerate the development of materials with desirable thermal properties. Given the fundamental nature of the proposed project, we currently do not anticipation commercialization of the developed topological phononic materials in the near future.

Publications and References

1. Z. Chen, X. Shen, N. Andrejevic, T. Liu, D. Luo, T. Nguyen, N. Drucker, M. Kozina, Q. Song, C. Hua, G. Chen, X. Wang, J. Kong, and M. Li, *Panoramic mapping of phonon transport from ultrafast electron diffraction and machine learning*, arXiv: 2202.06199 (2022).

Collaborative Research: DMREF: Machine Learning-aided Discovery of Synthesizable, Active and Stable Heterogeneous Catalysts

Lead Investigator: Suljo Linic, linic@umich.edu
Participating Institutions: University of Michigan – Ann Arbor, Wayne State University
Website: none
Keywords: Machine learning; heterogeneous catalysis; alloys

Project Scope

The objective of our DMREF project is to bring together a team of experts in machine learning (ML), computational catalysis, and catalytic experimentation to augment the current computational paradigm for catalyst discovery by creating a fundamental framework that will use not only descriptors of catalytic activity but also the descriptors of materials synthesizability and stability under reaction conditions. Specifically, we will develop a workflow for predicting the synthesizability, activity and stability of bimetallic catalysts. In this early stage, we will focus on bimetallic catalysts for low temperature preferential CO oxidation in the presence of H_2 and partial oxidation of ethylene to for ethylene oxide. We describe our overall approach in Figure 1. Success will be measured by the ability of the ML to predict synthesizable catalysts as verified by rigorous experimentation.

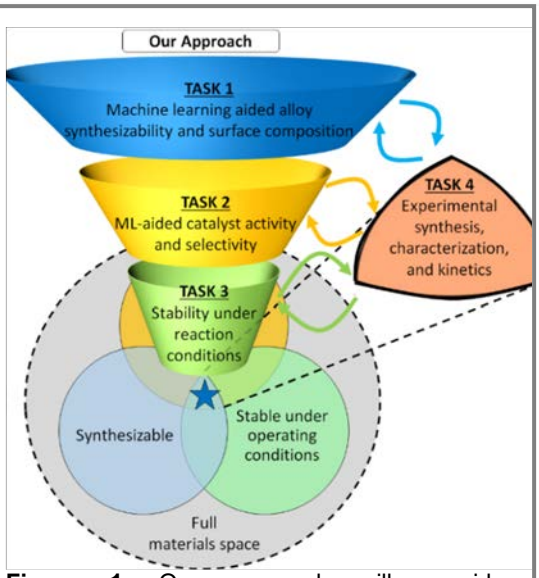


Figure 1. Our approach will consider synthesizability, surface composition of materials, activity (selectivity) and stability under operating conditions. This approach will yield a smaller phase space of potential target catalysts. Specifically, bimetallic alloys will be computationally screened based on synthesizability descriptors (TASK 1). The resulting stable alloy surfaces will subsequently be screened for the activity (selectivity) based on the activity (selectivity) descriptors using ML-aided DFT calculations (TASK 2). Finally, the top catalysts will be screened for stability under reactions conditions using ML-aided *ab initio* thermodynamics (TASK 3). Concurrent with the computational workflow, these catalysts will be synthesized and tested for catalytic performance and characterized under reaction conditions (TASK 4).

Relevance to MGI

Our proposed work addresses two NSF's *Big Ideas*, namely: (1) "Harnessing the Data Revolution" by combining ML and data analytics with computational and experimental catalysis research, and (2) "Grow Convergent Research" by addressing major challenges at the interface of the energy and material science by designing catalysts that enable sustainable chemical transformations. In line with DMREF's goals, our work will: (1) closely couple computation, theory, and experimentation in our catalytic materials design strategies, (2) help lead a paradigm shift in the way catalysis researchers perform research through an integrated ML-aided approach, (3) make all of our data and developed ML algorithms available to the broader community via dissemination in public repositories and GitHub, and (4) create a Data Science for Catalysis Training Program to educate and inspire the next generation of the catalytic materials workforce. Ultimately, by moving ML for catalysis beyond the studies of highly idealized systems and by incorporating the impact of stability and synthesizability into our analysis, we hope to serve as a go-to entity for practitioners of catalysis who seeks user-friendly inputs from fist principles calculations.

Technical Progress

Synthesis: Successfully synthesis of a library of alloys (PtM/support (M = Cu, Ru and Co); support = SiO_2 and high surface area carbon) have been achieved using two synthesis methods: borohydride reduction and strong electrostatic adsorption. Monometallic catalysts have been also synthesized as benchmarks. The reproducibility of the synthesis protocols has been validated at University of Michigan and Wayne State University.

Characterization: The synthesized alloys were rigorously characterized by using x-ray diffraction (XRD), in situ CO-Diffuse reflectance infrared Fourier transform spectroscopy (DRIFTS), X-ray absorption spectroscopy (XAS) and transmission electron microscopy with elemental mapping (STEM/EDS).
Database construction: Experimental synthesis procedures and all characterization data are being uploaded to our Metal Alloy Database (MAD) which will serve as a public repository for metal alloy catalyst data.[1] Currently we have several dozen materials synthesized and characterized. Graph-based procedures are generated and stored in MAD, accessible through a progress web app. Alloy characterization data will be available for download and visualization. We will leverage the MAD database to construct a computer-aided workflow and open-source tools for predicting the synthesizability, activity, and stability of catalysts.

Future Plans
The future plans involve the expansion of the compositional space of the alloys to obtain the necessary volume of data to train ML algorithms in predicting the synthesizability of these catalysts. We are currently working on developing high-throughput approaches for synthesizing controllably a large array of alloys. We will test these alloys in the two reactions of interest described above in this text.

Broader Impacts and Workforce Development
Our goal is to train and equip the next generation workforce in material science and catalysis as well as to educate the general public. Our outreach efforts focus on student professional development, broadening science participation, and informal science communication to help create a world-class scientific workforce. Thus far, we have been organizing cross-disciplinary training for graduate and undergraduate students at University of Michigan (U-M) and Wayne State University (WSU), providing them with a foundation to continue making scientific advances throughout their careers. For example, students funded by DMREF have received cross training on various characterization techniques for synthesis of alloys and ran experiments interchangeably at both institutions. PhD students have visited Stanford Linear Accelerator and Argonne National Lab's Advanced Photon Source to learn about x-ray absorption spectroscopy (XAS). We have made significant effort to recruit students from underrepresented minority groups to work on this project. We also are contributing to the scientific community by constructing an open-source database (publicly available repository) that will be accessible by the community to guide the synthesis and characterization of heterogeneous alloy catalysts using various techniques. This characterization data will be accessible as a training tool for new researchers. In the future we plan to create a Data Science for Catalysis Training Program, where promising undergraduates from WSU come to U-M during the summer to learn the basics of data science and catalysis. We will also engage underrepresented students and their parents in science outreach events from schools in Detroit, MI.

Data Management and Open Access
We are in the process of constructing a public database for all the alloy synthesis information and characterization data. This data is being loaded onto a searchable website where particular catalysts can be selected and their synthesis procedure and all the available characterization data can be displayed and downloaded. Currently the website is for the team only until several details are sorted out. In addition, we plan to make available validated results of electronic structure calculations via public repositories (e.g., Novel Materials Discovery Laboratory and Materials Project). This type of dissemination may occur after, before, or instead of publication in a scientific journal. Also, all high-fidelity experimental catalysis data for CO PROX and ethylene partial oxidation will also be stored in a MongoDB and will be made available to the public.
Advancing Along the Materials Development Continuum and Partnerships to Translation
We are pursuing several avenues to accelerate catalysis discovery and development. The first is to increase the reproducibility of synthesis and availability of synthesis information through the creation of the MAD database, which will avoid redundancies in material synthesis efforts, and is designed to accelerate the process of new researchers synthesizing desired materials. The second avenue is in using ML to predict new materials, with a much lower cost than experimentally synthesizing materials. The third avenue is in developing techniques to accelerate or automate materials synthesis to increase the throughput and reproducibility of materials discovery.

Publications and References
1. J. A. Esterhuizen, B. R. Goldsmith, and S. Linic. *Interpretable machine learning for knowledge generation in heterogeneous catalysis*, Nature Catalysis **5**, 175-184 (2022).

Elastomers Filled with Electro/Magneto-Active Fluid Inclusions: A New Paradigm for Soft Active Materials

Lead Investigator: Oscar Lopez-Pamies, pamies@illinois.edu.
Participating Institutions: University of Illinois Urbana-Champaign and The Pennsylvania State University
Website: https://publish.illinois.edu/liquidinclusions/team/
Keywords: Elastomers; Surface tension; Homogenization; Elasto-capillarity; Cloaking.

Project Scope

Elastomers filled with *fluid* inclusions — contrary to conventional *solid* fillers — have emerged over the past few years as a new class of materials with the potential to revolutionize a wide range of technologies. In this context, the main objectives of this research are to: **i)** derive and numerically implement the homogenized equations describing from the bottom up the macroscopic electro/magneto-mechanical response of elastomers filled with fluid inclusions, **ii)** deploy the derived equations to guide the design of elastomers filled with liquid-metal and ferrofluid inclusions with optimized macroscopic electro- and magneto-mechanical properties, and **iii)** fabricate and test prototypes of such new materials.

Relevance to MGI

The interaction between theory, computation, synthesis, and characterization has proved critical for the success of this project thus far. As an example, simulations of the mechanical response of suspensions of monodisperse liquid inclusions provided bounds on how small the inclusions ought to be and how soft the elastomer ought to be in order to have macroscopic behaviors that are dominated by the mechanics of the elastomer/liquid-inclusions interfaces. These bounds affected the protocol in which the materials were fabricated. On the other hand, the characterization of the mechanical response of the fabricated elastomers provided the means to identify via comparison with the theory how the surface tension of the elastomer/liquid interfaces depends on deformation.

Technical Progress

In a recent contribution [1, 2], we have formulated the homogenization problem that describes from the bottom up the mechanical response of elastomers filled with liquid inclusions under finite quasistatic deformations. Specifically, we have considered that the elastomer making up the matrix is a hyperelastic solid, that the liquid making up the inclusions is a hyperelastic fluid, that the interfaces separating the solid elastomer from the liquid inclusions are hyperelastic interfaces, and that the inclusions are spherical in shape in their ground state. We have shown that the resulting macroscopic behavior of such filled elastomers is that of a hyperelastic solid — distinctly, one that depends directly on the size of the inclusions and the constitutive behavior of the interfaces — and hence that it is characterized by an effective stored-energy function $W(\mathbf{F})$ of the macroscopic deformation gradient $\mathbf{F}$. For the case of filled elastomers with periodic microstructures, moreover, we have provided a formula for $W(\mathbf{F})$. The computation of $W(\mathbf{F})$ amounts to solving a super-cell problem that exhibits two non-standard features, both of which have profound implications not only on the mechanical

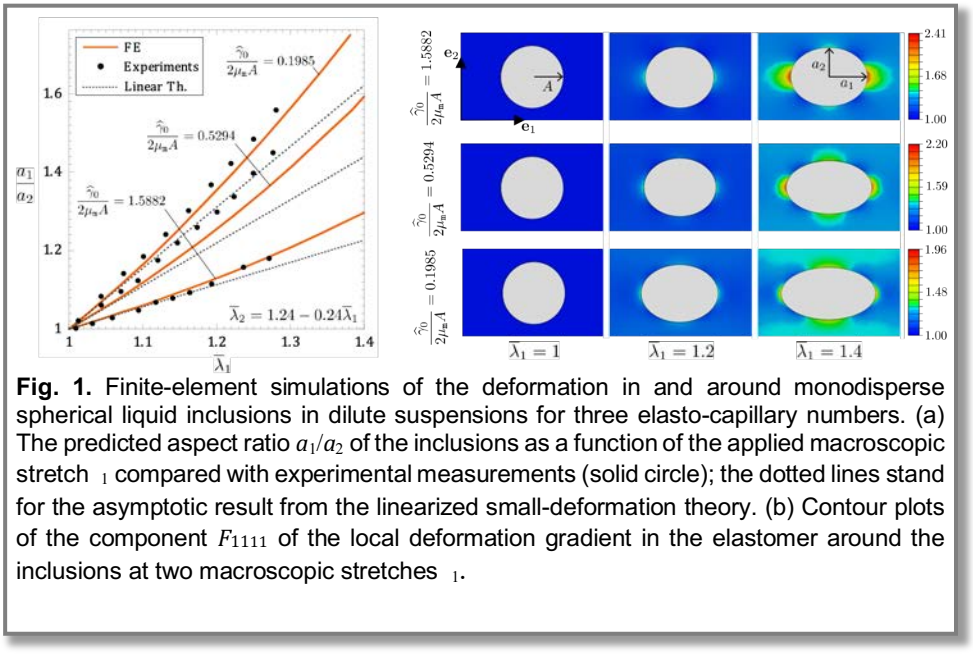


Fig. 1. Finite-element simulations of the deformation in and around monodisperse spherical liquid inclusions in dilute suspensions for three elasto-capillary numbers. (a) The predicted aspect ratio a_1/a_2 of the inclusions as a function of the applied macroscopic stretch $_1$ compared with experimental measurements (solid circle); the dotted lines stand for the asymptotic result from the linearized small-deformation theory. (b) Contour plots of the component F_{1111} of the local deformation gradient in the elastomer around the inclusions at two macroscopic stretches $_1$.

response of the material, but also on the mathematical analysis of the problem. The most remarkable for the latter is that in the limit of small deformations, when the governing equations are linearized, the otherwise powerful Lax-Milgram theorem does *not* apply to prove existence of solution.

To gain quantitative insight, we have also introduced and made use of a finite-element (FE) implementation of the proposed general formulation to work out numerical solutions for the macroscopic response of Gaussian elastomers filled with random isotropic distributions of incompressible liquid inclusions of monodisperse size and interfaces featuring a constant surface tension, which is arguably the most basic class of elastomers filled with liquid inclusions. The solutions have shown that the presence of inclusions with large (small) elasto-capillary numbers can lead to very significant stiffening (softening) of the macroscopic response, more so the larger the volume fraction of inclusions and the larger the macroscopic deformation. This behavior — which confirms previous experimental and theoretical results centered on small deformations — is due to the fact that inclusions with a relatively surface-tension-to-inclusions-size ratio pose significant resistance to deformation and hence *de facto* behave like stiff inclusions. The solutions have also shown that locally the deformations in and around the inclusions can be large, even when the macroscopic deformations are small, and hence that accounting for finite deformations is essential; see the Fig. 1.

Future Plans

In the remainder of the project, we will extend our investigation to the coupling between mechanical and electric/magnetic properties of elastomers filled with liquid-metal and ferrofluid inclusions.

Broader Impacts and Workforce Development

This project is training three PhD students for careers in industry and/or academia. This training is truly unique as our biweekly Zoom meetings and written exchanges provide all three students with complementary perspectives (mathematics, mechanics, computation, and materials science) in the design of an advanced class of materials.

Data Management and Open Access

We are in the process of finalizing our website for the project, where we plan to make accessible all the experimental data, codes, and publications that will be created. All 8 journal publications and 2 FE codes that have resulted from the project so far are also available from the PI's webpage and in GitHub. Moreover, we have publicized major outcomes from project through news outlets, see, e.g., https://cee.illinois.edu/news/researchers-derive-new-theory-behavior-new-class-materials.

Advancing Along the Materials Development Continuum and Partnerships to Translation

Nothing yet to report.

Publications and References

1. K. Ghosh, V. Lefèvre, O. Lopez-Pamies, *Homogenization of elastomers filled with liquid inclusions: The small-deformation limit*. SIAM Journal of Applied Mathematics. Under review (2022)
2. Ghosh, K., Lopez-Pamies, O. 2022. *Elastomers filled with liquid inclusions: Theory, numerical implementation, and some basic results*. Journal of the Mechanics and Physics of Solids **166**, 104930 (2022).
3. V. Lefèvre, O. Lopez-Pamies, *The effective shear modulus of a random isotropic suspension of monodisperse rigid n-spheres: From the dilute limit to the percolation threshold*. Extreme Mechanics Letters **55**, 101818 (2022).
4. V. Lefèvre, G.A. Francfort, O. Lopez-Pamies, *The curious case of 2D isotropic incompressible Neo-Hookean composites*. Journal of Elasticity, In press (2022).
5. G.A. Francfort, A. Gloria, O. Lopez-Pamies, *Enhancement of elasto-dielectrics by homogenization of active charges*. Journal de Mathématiques Pures et Appliquées **156**, 392–419 (2021).
6. K. Ghosh, B. Shrimali, A. Kumar, O. Lopez-Pamies, *The nonlinear viscoelastic response of suspensions of rigid inclusions in rubber: I — Gaussian rubber with constant viscosity*. Journal of the Mechanics and Physics of Solids **154**, 104544 (2021).
7. K. Ghosh, O. Lopez-Pamies, *On the two-potential constitutive modeling of dielectric elastomers*. Meccanica **56**, 1505–1521 (2021).
8. B. Shrimali, W.J. Parnell, O. Lopez-Pamies, *A simple explicit model constructed from a homogenization solution for the large-strain mechanical response of elastomeric syntactic foams.* International Journal of Non-Linear Mechanics 126, 103548 (2020).

Design of Organic-Inorganic Membranes for Extreme Chemical Environments

Lead Investigator: Mark Losego, losego@gatech.edu
Participating Institutions: Georgia Institute of Technology
Website: N/A
Keywords: membranes for separation, chemical extremes, vapor phase infiltration, density functional theory, machine learning

Project Scope

Chemical separations are critical for industrial production of clean water, chemicals, and pharmaceuticals. However, these separations are financially and environmentally costly, accounting for 10-15% of global energy consumption. Replacing thermally-based separation processes with engineered membrane materials could reduce this consumption by an order of magnitude. Polymer based membranes are of interest for their scalability, but advances in chemical stability are required for operation in aggressive organic solvents. This research program combines computational and experimental approaches to accelerate the design of organic-inorganic hybrid membranes with consistent performance in a broad range of solvent environments.

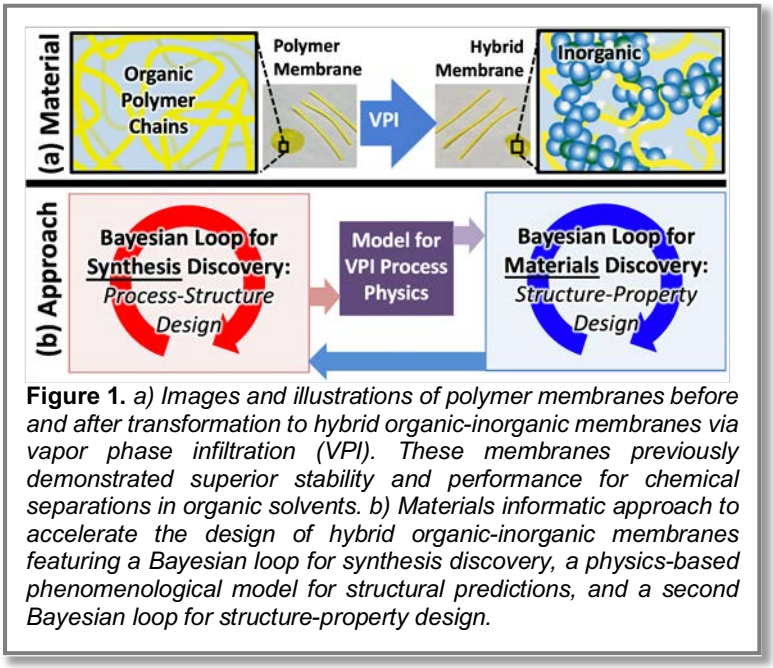


Figure 1. *a) Images and illustrations of polymer membranes before and after transformation to hybrid organic-inorganic membranes via vapor phase infiltration (VPI). These membranes previously demonstrated superior stability and performance for chemical separations in organic solvents. b) Materials informatic approach to accelerate the design of hybrid organic-inorganic membranes featuring a Bayesian loop for synthesis discovery, a physics-based phenomenological model for structural predictions, and a second Bayesian loop for structure-property design.*

Relevance to MGI

This project builds upon initial experimental success creating hybrid membranes via vapor phase infiltration (VPI, a gas-phase process that infuses polymers with inorganics at the atomic level) for chemical separations **(Figure 1a)**. However, the processing space for creating hybrid materials via VPI is vast, complex, and largely unexplored. To accelerate membrane discovery, this project develops powerful data-driven machine learning methods for both process-structure design for VPI and hybrid material structure-property design and which then are combined via a physics-based phenomenological model (**Figure 1b**). In the first Bayesian loop, multi-fidelity data from both experimental and computational (density functional theory) sources are fed into machine learning algorithms exploring fundamental parameters key to VPI (precursor vapor pressure, precursor solubility, binding energy between polymer and metalorganic, etc.). The outputs identify VPI systems of interest which are pursued experimentally for membrane development. Active learning approaches are also applied to identify areas of data sparsity which experimentalists then pursue to improve the learning process. A similar approach is taken for the second Bayesian loop focusing on how hybrid material chemical and physical structure influences separation performance with loops connected via a physics-based phenomenological model recently developed by the team.

Technical Progress

Significant progress has been made developing machine learning algorithms capable of handling multi-fidelity data (experimental and computed) to predict the vapor pressure of both organic and metalorganic precursors, a key component of both the VPI process (Loop #1) and in semi-conductor manufacturing (several manuscripts currently in preparation). Lessons from developing these algorithms will be applied to machine learning on data regarding the strength of chemical interactions between metalorganics and polymers which is currently being explored both through density functional theory and experimental measurements.[1] To connect these outputs to structure-property relationships (Loop #2), a new physics-based phenomenological model for vapor phase infiltration has been introduced and experimentally validated. The model has been designed to take input from the machine learning explorations and accurately captures and predicts how the convolution of reactions and diffusion during vapor phase

infiltration (VPI) shapes the evolution of the hybrid membrane structure in terms of inorganic loading both spatially and temporally.[6] However, the model relies on a complex set of differential equations with six distinct variables related to aspects of a given VPI process. To accelerate mapping of the model to experimental datasets via parameter identification, a new Bayesian optimization procedure for solving this inverse problem has been developed.[5] Importantly, the model accurately predicts structural information not only via *in situ* experimental measurements, but also via more accessible *ex situ* characterization which comprises a larger fraction of available VPI literature.[2] The capability of the model to produce both types of structural data is critical to the success of machine learning efforts on structure-property relationships (Loop #2) as it both expands data availability and utility of outputs. To better inform the data-driven approaches for structure-property relationships (Loop #2), experimental explorations have been conducted connecting the stability of hybrid materials (under thermal and heated aqueous treatments) to their chemical and physical structure as controlled via VPI processing parameters.[3,7] Further, the application spaces for the methodologies and knowledge developed within this project have been advanced to include structurally robust hybrid materials to serve as scaffolds for amine-based carbon capture.[4]

Future Plans

Further development of multi-fidelity machine learning methods for precursor vapor pressure and precursor-polymer binding energy is currently underway including the development of datasets via computational and experimental methods. Additional application of the phenomenological model to new precursor-polymer systems explored experimentally is being pursued to provide input for structure-property design (Loop #2). Further, connections between hybrid material structure and membrane performance (stability, permeance, selectivity) are also under active exploration to provide feedback on the output of structure-property design (Loop #2).

Broader Impacts and Workforce Development

To date this program has contributed to the multidisciplinary education of 6 graduate students (one URM, two women, and 1 international visitor), 2 post-doctoral trainees (both women), and 6 undergraduate students. The results of this program have been reported via 8 publications (5 collaborative), three public thesis defenses, as well as over 15 presentations at international scientific conferences and/or institutional seminars.

Data Management and Open Access

Where appropriate, data and code corresponding to publications have been made publicly available through publisher websites and the code for the VPI phenomenological model is available on GitHub.

Advancing Along the Materials Development Continuum and Partnerships to Translation

The results of this project are featured in two active patent applications and have been discussed with industry representatives including a Japanese membrane manufacturer. The significant application space for this work has further led to preliminary investigations regarding the scalability of these processes.

Publications and References

1. Y. Liu, E. K. McGuinness, B. C. Jean, Y. Ren, B. Gonzalez del Rio, R. P. Lively, M. D. Losego, R. Ramprasad, *Infiltration of Polymer of Intrinsic Microporosity 1 (PIM-1) with Trimethylaluminum (TMA) and Water: A combined computational and experimental study*, (under review)

2. B. C. Jean, Y. Ren, E. K. McGuinness, R. P. Lively, M. D. Losego*, Effects of Trimethylaluminum Vapor Pressure and Exposure Time on Inorganic Loading in Vapor Phase Infiltrated PIM-1 Polymer Membranes,* (under review)

3. K. Pyronneau, E. K. McGuinness, A. J. Gonzalez, B. C. Jean, H. V. Manno, N. R. McClelland, M. D. Losego, *Wash Fastness of Hybrid AlOx-PET Fabrics Created via Vapor-Phase Infiltration*, ACS Appl. Polym. Mater., **4** 3304 (2022).

4. F. Zhang, E. K. McGuinness, Y. Ma, Y. Ren, J. Leisen, M. D. Losego, R. P. Lively, *Vapor-phase-infiltrated AlO_x/PIM-1 "hybrid scaffolds" as solution-processable amine supports for CO_2 adsorption*, ACS Appl. Polym. Mater., **3** 4460 (2021).

5. C. Huang, Y. Ren, E. K. McGuinness, M. D. Losego, R. P. Lively, V. R. Joseph, *Bayesian optimization of functional output in inverse problems*, Optimization and Engineering **22**, 2553 (2021).

6. Y. Ren, E. K. McGuinness, C. Huang, V. R. Joseph, R. P. Lively, M. D. Losego, *Reaction-diffusion transport model to predict precursor uptake and spatial distribution in vapor-phase infiltration processes*, Chemistry of Materials **33**, 5210 (2021).

7. J. T. Bamford, R. A. Smith, C. Z. Leng, W. R. Gutekunst, M. D. Losego, *Measuring the glass transition temperature of vapor-phase infiltrated AlOx-PS-r-PHEMA organic-inorganic hybrid thin-film materials*, Macromolecules **54**, 6790 (2021).

8. J. Pilz, A. M. Coclite, M. D. Losego, *Vapor phase infiltration of zinc oxide into thin films of cis-polyisoprene rubber*, Mater. Adv. **1** 1695 (2020).

Center for Computational Study of Excited-State Phenomena in Energy Materials (C2SEPEM)

Lead Investigator: Steven G. Louie, Lawrence Berkeley National Laboratory (LBNL) and Physics Department, University of California at Berkeley (UC Berkeley), sglouie@berkeley.edu.
Participating Institutions: Lawrence Berkeley National Laboratory/University of California at Berkeley (*Jeffrey B. Neaton, Jack Deslippe, Naomi Ginsberg, Eran Rabani, Feng Wang, Chao Yang*), The University of Texas, Austin (*James R. Chelikowsky*), Stanford University (*Felipe da Jornada*), University of California, Los Angeles (*Daniel Neuhauser*), Yale University (*Diana Y. Qiu*)
Website: https://c2sepem.lbl.gov/
Keywords: Interacting-particle Green's functions, excited-state phenomena and spectroscopies, time-dependent phenomena, software for high-performance computing, BerkeleyGW

Project Scope

The mission of the Center for Computational Study of Excited-State Phenomena in Energy Materials (C2SEPEM) is to develop and implement new theories, methods, algorithms, and computer codes to explain and predict excited-state phenomena in materials. C2SEPEM performs research on first-principles many-body perturbation theory and advanced algorithms, as well as their experimental validation and efficient implementation for high-performance computers. Studies and software developments on quasiparticle excitations, optical spectra, trion and bi-exciton formation, exciton-exciton interactions, nonlinear optical processes, time-dependent phenomena, and more are carried out for bulk and reduced-dimensional systems.

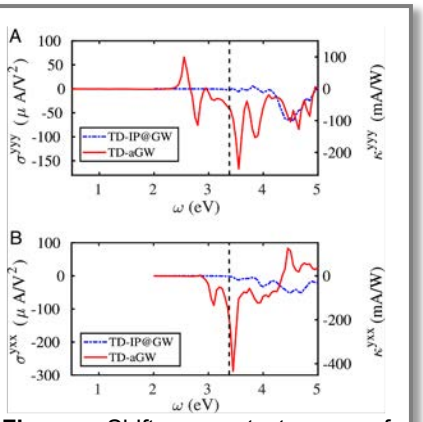


Figure. Shift current tensor of monolayer GeS along the (A) yyy and (B) yxx directions, calculated using the time-dependent adiabatic GW (TD-aGW) approach [7].

Relevance to MGI

Excited-state phenomena in a material typically give rise to its defining attributes and determine its usefulness. These phenomena are particularly important in processes of energy generation, transport and storage. However, *ab initio* methods for them, especially for multiple-particle correlated excitations and ultra-fast electron dynamics, have been under explored and hence limiting their studies in real materials. The methods and software developed are aiming to fill this gap, and are relevant to the predictive study of complex materials and validated through close collaboration with experimental groups. The end result will be integrated open-source software packages with capabilities to predict and understand a variety of excited-state phenomena from first principles.

Technical Progress

Technical progress of C2SEPEM covers development and release of community software packages, novel theories and methods for nonequilibrium dynamics, correlated multiparticle excitations, electron-phonon and exciton-phonon couplings, stochastic approaches, optical phenomena in two-dimensional (2D) materials, moiré heterostructures and defects, as well as advanced numerical algorithms and experimental studies. Forty (40) papers have been published or submitted since September 2020 acknowledging this Program (48% having 2+ co-PIs, 19 in Category A, 21 in Category B). These articles appeared in journals such as *Nature, Science, Nat. Materials, Nat. Commun., Proc. Natl. Acad. Sci., Phys. Rev. Letters, ACS Nano*, *Nano Letters, Comput. Phys. Commun.,* among others. The team members at C2SEPEM have been recognized with 7 significant awards. We released new software versions for BerkeleyGW (v3), StochasticGW (v2), and NanoGW. We organized the Annual BerkeleyGW Tutorial Workshops and the Annual Berkeley Excited State Conferences, each attracting over 300 participants.

Future Plans

Planned directions: 1) novel theories for excited-state phenomena; 2) new computational algorithms and their implementations; 3) porting of software packages to GPUs for exascale computing and maintaining active user and

developer communities; and 4) applying developed theories and methods to study of electronic, optical, and dynamics phenomena in novel materials, in collaboration with experimental efforts.

Broader Impacts and Workforce Development

C2SEPEM values and actively engages in diversity, equity, and inclusion (DEI) initiatives. The Center strives to support researchers from underrepresented minority (URM) groups, and many of our URM alumni moved on to prominent faculty positions (e.g., Yale, Stanford, Weizmann Institute of Science, Oxford). We actively involve in and search for funding sources for local outreach programs. C2SEPEM provides ample career development opportunities for young researchers to mature into future leaders. These include: rigorously trained in their respective disciplines; opportunities of collaborating with peers and scientists across several different disciplines on a project; trained in and exposed to the science and art of software development; asked to take important applied math and computer science courses, and instructed in best practices in writing inter-portable codes.

Data Management and Open Access

All software developed as part of C2SEPEM are made available to the community using open-source distribution licenses (the Open-Source Initiative (OSI)). The BerkeleyGW code is distributed with an open-source BSD license. Community input to the codes is solicited through workshops and meetings, and through the users' forum on the BerkeleyGW website. Many of the methods developed and the resulting codes and libraries will be general and relevant for the chemistry, condensed matter physics, materials science and engineering communities. C2SEPEM's goal is to deliver robust software, separate well-defined portable libraries and software tools in an open-source manner such that it allows the components to be integrated into codes outside of the Center, thereby maximizing the impact of the development efforts. The Center's data management framework supports the complete data lifecycle: requirements, collection, processing and validation, archiving and sharing, access and dissemination.

Advancing Along the Materials Development Continuum and Partnerships to Translation

C2SEPEM develops *ab initio* theories, methods and codes for excited-state properties and phenomena that are not adequately addressed for real materials from first principles currently. Such predictive methodologies and studies fill an important gap in the spectrum of tools for the theoretical and computational study of materials. The codes developed, as mentioned above, will be made available freely to all researchers.

Publications and References (10 recent ones)

1. K. H. Liou, C. Yang, and J. R. Chelikowsky, *Scalable Implementation of Polynomial Filtering for Density Functional Theory Calculation in PARSEC*, Comp. Phys. Commun. **254**, 107330 (2020).

2. M. Del Ben, C. Yang, Z. Li, F. H. da Jornada, S. G. Louie, and J. Deslippe, *Accelerating Large-Scale Excited-State GW Calculations on Leadership HPC Systems*, SC20: International Conference for High Performance Computing, Networking, Storage and Analysis **4**, 36 (2020).

3. D. Y. Qiu, F.H. da Jornada, and S. G. Louie, *Solving the Bethe-Salpeter equation on a subspace: Approximations and consequences for low-dimensional materials*, Phys. Rev. B **103**, 045117 (2021).

4. S. G. Louie, Y.-H. Chan, F. H. da Jornada, Z. Li, and D. Y. Qiu, *Discovering and understanding materials through computation*, Nat. Mater. **20**, 728 (2021).

5. Z. L. Li, M. Wu, Y. H. Chan, and S. G. Louie, *Unmasking the Origin of Kinks in the Photoemission Spectra of Cuprate Superconductors*, Phys. Rev. Lett. **126**, 146401 (2021).

6. H. Bretscher, Z. J. Li, J. Xiao, D. Y. Qiu, S. Refaely-Abramson, J. A. Alexander-Webber, A. Tanoh, Y. Fan, G. Delport, C. A. Williams, S. D. Stranks, S. Hofmann, J. B. Neaton, S. G. Louie, and A. Rao, *Rational passivation of sulfur vacancy defects in two-dimensional transition metal dichalcogenides*, *ACS Nano* **15**, 8780 (2021).

7. Y.-H. Chan, D. Y. Qiu, F. H. da Jornada, and S. G. Louie, *Giant exciton-enhanced shift currents and direct current conduction with subbandgap photo excitations produced by many-electron interactions*, Proc. Natl. Acad. Sci. U.S.A. **118,** e1906938118 (2021).

8. M. Nguyen, W. Li, Y. Li, E. Rabani, R. Baer, and D. Neuhauser, *Tempering Stochastic Density Functional Theory*, J. Chem. Phys. **155**, 204105 (2021).

9. W. Gao, F. H. da Jornada, M. Del Ben, J. Deslippe, S. G. Louie, and J. R. Chelikowsky, *Quasiparticle energies and optical excitations of 3C-SiC divacancy from GW and GW plus Bethe-Salpeter equation calculations*, Phys. Rev. Mat. **6**, 036201 (2022).

10. M. H. Naik, E. C. Regan, Z. Zhang, Y.-H. Chan, Z. Li, D. Wang, Y. Yoon, C. S. Ong, W. Zhao, S. Zhao, M. Iqbal Bakti Utama, B. Gao, X. Wei, M. Sayya, K. Yumigeta, K. Watanabe, T. Taniguchi, S. Tongay, F. H. da Jornada, F. Wang, and S. G. Louie, *Intralayer charge-transfer moiré excitons in van der Waals superlattices*, *Nature*, in press (2022).

Manipulating Optical Properties of van der Waals Materials Through Band Nesting Effects

Lead Investigator: Tony Low, tlow@umn.edu
Participating Institutions: University of Minnesota, University of Notre Dame
Website: none
Keywords: Band nesting, Light absorption, Transition metal dichalcogenides, van der Waals heterostructure, interlayer interaction, hyperbolicity, non-linear response, machine learning

Project Scope

We provide a theoretical design framework for achieving perfect light absorption (>99%) in 1-nm-thick van der Waals (vdW) materials. We show that this goal can be achieved with various transition metal dichalcogenides (TMD) materials covering a wide spectrum of frequencies across the infrared to visible range. Figure 1 illustrates the requirement on the optical conductivity of the ultrathin TMD layers in achieving perfect absorption in a cavity structure. Key findings of our work include the engineering of the ultrathin TMD layers in achieving the desired optical conductivity through band nesting, strains, and other approaches. We also experimentally realize these structures through mechanical exfoliation and MBE growth. Theoretically, we have also extended these band nesting concepts to engineer new novel optical effects, such as giant non-linear second harmonic response, and extreme optical birefringence in the form of in-plane hyperbolicity. On the machine learning front, we have developed a general framework for extracting physical rules of vdW materials using global and local machine learning methodologies.

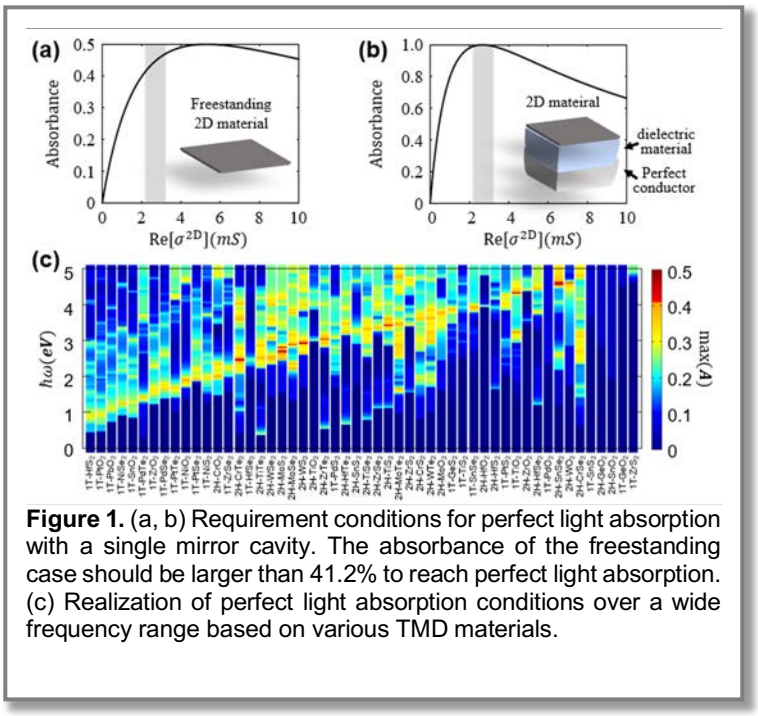


Figure 1. (a, b) Requirement conditions for perfect light absorption with a single mirror cavity. The absorbance of the freestanding case should be larger than 41.2% to reach perfect light absorption. (c) Realization of perfect light absorption conditions over a wide frequency range based on various TMD materials.

Relevance to MGI

The closed and iterative feedback loop between theory and experiment, in conjunction with machine learning, was established in this project and played a pertinent role to the success of this work. The combination of first-principles calculations (**Low**, UMN) and machine learning (**Cherkassky**, UMN) enables efficient theoretical calculations of vast material space. The theoretical prediction informed the experimental team for 2D heterostructure growth (**Hinkle**, ND), and fabrication and characterization (**Koester**, UMN).

Technical Progress

Due to the low light absorbance of two-dimensional materials, practical applications of photodetectors and optoelectronics based on two-dimensional materials have been greatly limited. Although there are several reports on perfect light absorption using two-dimensional materials, they typically employ complicated optical structures such as patterned structures or multi-layered supercell structures, which are not desirable in device fabrication. We seek 2D heterostructure materials based on TMDs to engineer heterostructures with perfect light absorption. TMDs are unique due to their band nesting behavior which allows extremely-high absorption at specific wavelengths due to the broad region of k-space over which absorption can occur. **Low** has successfully identified the heterostructure stacks that realize such perfect light absorption. **Hinkle** has successfully grown these heterostructure materials using MBE and demonstrate their feasibility to grow them *in-situ* within the same chamber. **Koester** has also made these heterostructure using large-area exfoliation techniques and utilizing the MBE samples and performed optical measurements. Our collective studies have demonstrated the viability of our designed ultrathin heterostructure as a candidate for perfect light absorption. This work is currently being prepared for a journal publication, and has also been submitted as a

conference abstract.

We also seek to demonstrate that band nesting effects in TMDs can be exploited to generate novel optical responses. We have proposed a new mechanism to generate giant anisotropic second harmonic nonlinear response *via* double resonance effects, achieved through electronic bandstructure engineering. The ideal band setup would be a triplet of nested bands separated by the fundamental resonance. **Low** had demonstrated theoretically that the proposed phenomenon can be realized in bilayer SnS by band tuning with perpendicular electrical bias, which maximizes the second harmonic susceptibility by several orders of magnitude. **Low** has also explored the tunability of hyperbolicity in a naturally in-plane hyperbolic material monolayer WTe_2 as a function of external perpendicular electric field and electronic doping. New interband optical transitions can be triggered which modifies the interband optical conductivity. By combining these two stimuli, it is possible to dynamically tune both the type and frequency range of hyperbolicity.

Future Plans

Further experiments are necessary to realize a perfect light absorber with a single mirror cavity structure. New predicted phenomena on nonlinear and hyperbolicity can also be explored experimentally. Machine learning methods are being developed by **Cherkassky** that allow for heterostructure design of 2D materials, particularly in identifying the technologically interesting type II and III band alignments.

Broader Impacts and Workforce Development

Low has co-organized a workshop on 2D materials nanophononics, with Luis Martin Moreno and Alexey Nikitin via Zoom with about 100 participants. S. R. Biswas, a Ph.D. student supported by this grant has graduated and joined Lumiled, US, and will perform design and theory on light emitting devices in this company, a skill he learned during the support of this grant. **Hinkle** gave the keynote presentation at the NSF Future of Semiconductors Workshop in March 2021. In the talk, our 2D materials research was presented as a potential part of the future of monolithic 3D integrated circuit technologies. Our team's collaborative results have been used as examples in **Hinkle** Electrical Engineering undergraduate course at Notre Dame (Electronic and Optoelectronic Devices). The recent findings of current NSF-supported research has worked exceedingly well at inspiring young students to see that the fundamental concepts taught in the class have real-world, modern applications.

Data Management and Open Access

All data generated by theoretical and experimental works is stored locally in the PI's universities and securely backed up through automated services. All data is available upon personal request to the PIs.

Advancing Along the Materials Development Continuum and Partnerships to Translation

We are in the process of filing for IP protection between our universities, for the methods, materials and device design that realizes the perfect absorption.

Publications and References

1. E. H. Lee, W. Jiang, H. Alsalman, T. Low, and V. Cherkassky. "Methodological framework for materials discovery using machine learning." *Physical Review Materials* 6, 043802 (2022)
2. S. R. Biswas, J. Yu, Z. Wang, D. R. da Costa, C. Zhao, S. Yuan, and T. Low. "Double Resonant Tunable Second Harmonic Generation in Two-dimensional Layered Materials through Band Nesting." *arXiv preprint arXiv:2108.06900* (2021).
3. G. Zhou, T. Sun, R. Younas, and C. L. Hinkle, "Materials and Device Strategies for Nanoelectronic 3D Heterogeneous Integration," IEEE 2021 International Conference on Simulation of Semiconductor Processes and Devices (SISPAD), pp. 163-166, (2021).
4. A. Chaves, J G. Azadani, H Alsalman, D. R. Da Costa, R. Frisenda, A. J. Chaves, S H Song et al. "Bandgap engineering of two-dimensional semiconductor materials." *npj 2D Materials and Applications* 4, no. 1 (2020): 1-21.
5. D. Seo, S. Lee, N. Izquierdo, S. H. Park, E. H. Lee, R. Younas, G. Zhou, C. Hinkle, T. Low, and S. J. Koester1, "A path toward achieving perfect absorption in 2D heterostructures," submitted to 2022 Fall Materials Research Society Meeting.

DMREF: Data Driven Discovery of Conjugated Polyelectrolytes for Neuromorphic Computing

Lead Investigator: Gang Lu, gang.lu@csun.edu
Participating Institutions: California State University Northridge and University of California Santa Barbara
Website: https://www.cpegenome.com/
Keywords: CPE database, neuromorphic computing, machine learning, organic electrochemical transistors.

Project Scope

The overarching objective of the project is to establish a collaborative, multidisciplinary and data-centric research program which accelerates the discovery of novel conjugated polyelectrolytes (CPEs) with chemical structures tailored for the demands of neuromorphic engineering. We will deliver the first comprehensive database dedicated to CPEs by combined high-throughput first-principles computation, machine learning and experimental validations. We will also formulate a set of design rules to guide future synthesis of CPEs with privileged properties for neuromorphic and more broadly, emerging optoelectronic applications.

Relevance to MGI

The research effort integrated high-throughput computation, machine learning, chemical synthesis, and materials and device characterization in a "closed loop" fashion. We have developed the first CPE database, based on which we have formulated a set of design rules for molecular structures with prescribed properties. More specifically, we predicted several promising CPEs for neuromorphic and optoelectronic applications. Experimental validation of these predictions is being carried out. There has also been iterative feedback between theory and experiment on the synthesis of CPEs. First-principles calculations were performed to understand the experimental observations in terms of doping levels on the side-chain length in the CPEs. Simultaneously, the experiments have motivated the development and refinement of theoretical models on the nature of self-doping and formation of polarons in the *n*- or *p*-doped CPEs. The PIs, students and Postdocs have benefited from the multidisciplinary, integrated experimental and computational approach in this project.

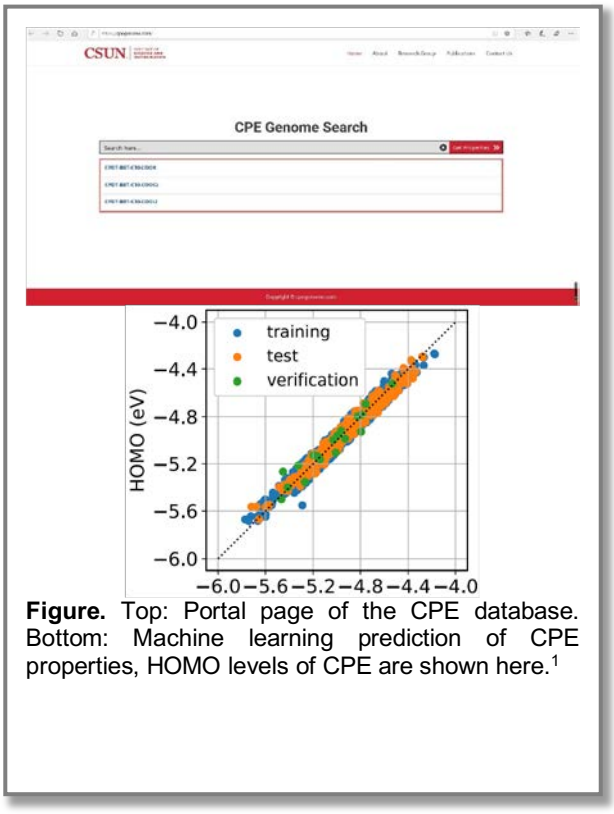


Figure. Top: Portal page of the CPE database. Bottom: Machine learning prediction of CPE properties, HOMO levels of CPE are shown here.[1]

Technical Progress

We have constructed a website for the CPE database, which is freely available to the research community. Recently, a Python code has been developed to perform machine learning predictions of CPE materials properties. We are working on the update of the web portal to incorporate the machine learning functionality. Based on the CPE database, we carried out a comprehensive study on the stability, structural deformation, electronic structure and optical absorption of positive/negative polarons and bi-polarons in CPEs. We shed light on the absence of negative polarons in CPEs and provide a strategy to improve the electrochemical stability of *n*-doped CPEs.[2] In addition, we have synthesized four different CPEs and investigate chemical structure effects on self-doping efficiency. In terms of self-doping efficiency, PCPDTPh and PCPDTTh exhibit much lower polaron due to deeper HOMO levels and disordered structures from high dihedral angles. From the simulation work, we revealed that the Gibbs free energy for polaron formation in PCPDTBZ is much higher than that of PCPDTBT. This higher Gibbs free energy restricts the polaron formation and is therefore undesirable for high self-doping efficiency. We have also discovered dual ionic-electronic interactions of CPEs when the CPE thin films are immersed in an electrolyte.[3,4] Upon applying voltages of different polarities to draw ions from the electrolyte into a thin film of PCPDTBT-SO_3K, we find that anion injection can dope the CPE while cation injection

de-dopes it. This discovery leads to the demonstration of OECTs using CPE-K as the active materials. OECTs have been the main organic artificial synapses for neuromorphic applications. CPE-K-based OECTs exhibit high performance and are the first dual mode transistors made with organic active materials.[4] We demonstrate neuromorphic functionalities of OECTs using biomaterial gels as the solid electrolyte and also develop a novel and efficient fabrication method of OECTs that has been applied to CPE-K-based OECTs.[5,6]

Future Plans

We will continue to expand the CPE database. More CPEs and additional properties will be included to the database. We will also improve the machine learning method to predict more CPE properties. A number of screened *n*-type CPEs will be verified experimentally. Based on the established platform charactering OECT performance, we will design novel CPEs with tunable electronic and ionic conductivities for applications in OECTs.

Broader Impacts and Workforce Development

Exposing students and postdoctoral scholars to data-centric research and education is an essential component of our objectives. Two graduate students and one undergraduate student have been supervised to learn basic physical principles, program skills to perform high-throughput first-principles calculations, and Python programming for machine learning modeling. Some of the CSUN students in this project have participated in the REU summer program (virtual) at Princeton MRSEC. The students at UCSB are also trained on the synthesis and characterization of targeted CPEs, some of them are screened from the database. We plan to incorporate data science and materials informatics in the future summer camps which host local high school students and science teachers.

Data Management and Open Access

We have created the first database dedicated to CPEs. The database includes CPE properties obtained by combined high-throughput first-principles computation and machine learning prediction. The CPE database is made readily available to the research community via a web portal https://www.cpegenome.com/. The experimental data includes the details of sample and device preparation, characterization data for selected CPEs. Some experimental data will be included in the database. The other means of providing access to these research data is through the publication of the results in peer-reviewed journals.

Advancing Along the Materials Development Continuum and Partnerships to Translation

Based on the CPE database, we can have a fundamental understanding on how the backbone structure and adjacent electrostatic forces control CPE properties and develop a set of design rules to screen the most promising CPEs to be synthesized and characterized. The closed feedback loops between theory, computation, synthesis and characterization could accelerate discovery of novel CPEs comparing to traditional trial-and-error materials discovery strategy. CPEs have been used as interfacial layers in organic solar cells, perovskite solar cells, and transistors. Thus, the material development from this project has impacted other research areas.

Publications and References

1. Y. Wan, F. Ramirez, X. Zhang, T. Q. Nguyen, G. C. Bazan, and G. Lu, *Data Driven Discovery of Conjugated Polyelectrolytes for Optoelectronic and Photocatalytic Applications*, npj Comput. Mater. **7**, 69 (2021).
2. Y. Wan, X. Zhang, G. C. Bazan, T. Q. Nguyen, and G. Lu, *Polarons in Conjugated Polyelectrolytes: A First-Principles Perspective*, submitted.
3. A. T. Lill, D. X. Cao, M. Schrock, J. Vollbrecht, J. Huang, T. Nguyen-Dang, V. V. Brus, B. Yurash, D. Leifert, G. C. Bazan, and T.-Q. Nguyen, *Organic Electrochemical Transistors Based on the Conjugated Polyelectrolyte PCPDTBT-SO3K (CPE-K)*, Adv. Mater. **32**, 1908120 (2020).
4. T.Nguyen-Dang, S. Chae, J. Chatsirisupachai, H. Wakidi, V. Promarak, Y. Visell, T. Q. Nguyen, *Dual-Mode Organic Electrochemical Transistors Based on Self-Doped Conjugated Polyelectrolytes for Reconfigurable Electronics.* Adv. Mater. **34**, 2200274 (2022).
5. T. Nguyen-Dang, K. Harrison, A. Lill, A. Dixon, E. Lewis, J. Vollbrecht, T. Hachisu, S. Biswas, Y. Visell, and T.-Q. Nguyen, *Biomaterial-based Solid-Electrolyte Organic Electrochemical Transistors for Electronic and Neuromorphic Applications*, Adv. Electron. Mater. **7**, 2100519 (2021).
6. T. Nguyen-dang, S. Chae, K. Harrison, L. C. Llanes, A. Yi, H. J. Kim, S. Biswas, Y. Visell, G. C. Bazan, T. Q. Nguyen, *Efficient Fabrication of Organic Electrochemical Transistors via Wet Chemical Processing.* ACS Appl. Mater. Interfaces **14**, 12469-12478 (2022).
7. X. Zhang, G. Lu, R. Baer, E. Rabani, and D. Neuhauser, *Linear Response Time-Dependent Density Functional Theory with Stochastic Range-Separated Hybrids*, J. Chem. Theory Comput. **16**, 1064-1072 (2020).

Accelerated Design of Redox-Active Polymers for Metal-Free Batteries

Lead Investigator: Jodie Lutkenhaus, jodie.lutkenhaus@tamu.edu
Participating Institutions: Texas A&M Engineering Experiment Station, University of Chicago
Website: none
Keywords: Organic Batteries, Organic Radical Polymers, Inverse Design, Closed-loop Design

Project Scope

Radical-containing polymers are promising as redox-active materials, but their current performance remains inferior to current Li-ion battery materials. This project addresses the specific challenges of improving the electrochemical performance of the radical-containing polymers as both cathodes and anodes. Specifically, the new knowledge to be gained regarding the cathode and anode include: (1) which chemical modifications adjust the redox potential of the redox active group; (2) which co-monomer patterns promote charge transport; and (3) which electrolyte promotes electrochemical stability of the polymer.

Relevance to MGI

In alignment with the Materials Genome Initiative (MGI), this project will culminate in the production and dissemination of a searchable polymer property database on redox-active polymers, polymer-specific inverse design machine learning tools, multiscale models for the redox kinetics, and a proof-of-concept all-organic battery.

Technical Progress

Our project overview, workflows, and tools are summarized in Figure 2 (see below). As we are approaching the end of our first year, we are currently in **Phase 0** of the project. We are populating a database with existing polymer/small molecule descriptors and target properties taken from the literature, including experimental data from **Lutkenhaus** and **Rowan.** Moreover, we are integrating our computational databases, including the molecular databases developed by the **Tabor** group and bulk polymer simulation protocols from the **de Pablo** group. **Lutkenhaus** and **Tabor** are doing first simulations on a specific polymer cathode chemistry as a test case for polymer-ion-solvent interactions. The team has identified its first leads for high-throughput synthesis for anode materials and meets every other week via Zoom.

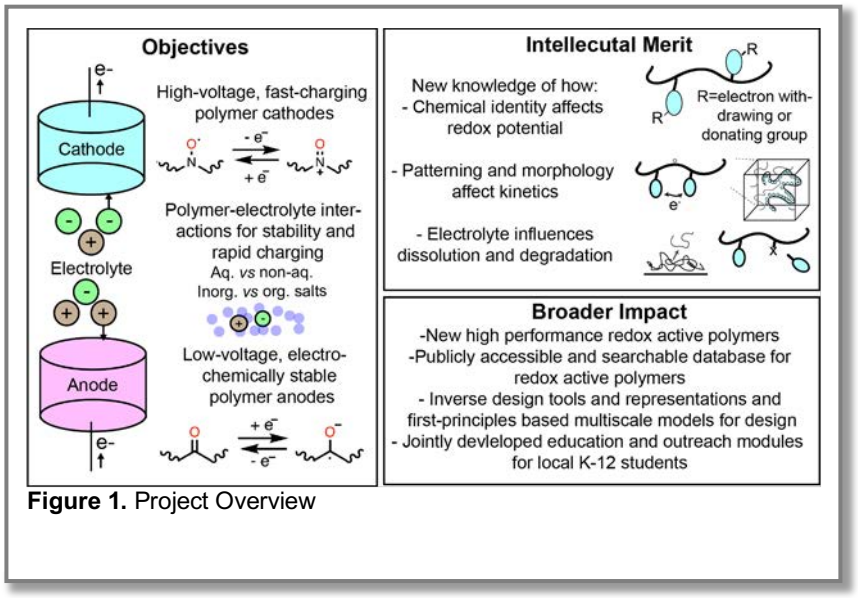


Figure 1. Project Overview

Future Plans

We are approaching the end of **Phase 0,** marked by the first syntheses and characterizations of targeted polymers, and by the first models of the same.

In **Phase 1**, we will explore targeted classes of polymers with bench-scale and high-throughput polymer synthesis/characterization to determine practical limits on both our synthetic space and design space. We will also conduct integration tests of data-driven Bayesian optimization algorithms with the robotic systems and identify both

critical and potential variables for further optimization. In parallel, molecular and multiscale simulation will be used to validate our models, to estimate the degree of further potential property optimization from the initial experiments, and to identify high-risk/high-reward sectors of chemical space. At the end of **Phase 1**, trends regarding cathode and anode polymers, as well as the electrolytes, will be elucidated.

In **Phase 2**, we will engage in an iterative design process of new high-performance polymers and electrolytes with a closed-loop integration of high-throughput experimental tools and computational molecular model predictions. The completion of **Phase 2** will be marked by the dissemination of the following Final Products to the community: (1) New high-performance redox-active polymers for all-organic batteries; (2) A comprehensive open and living database of computational and experimental properties of polymers; (3) Inverse design tools and representations tailored for polymer design, and (4) First-principles based multiscale models for radical polymer charge transport, stability, solubility, and kinetics.

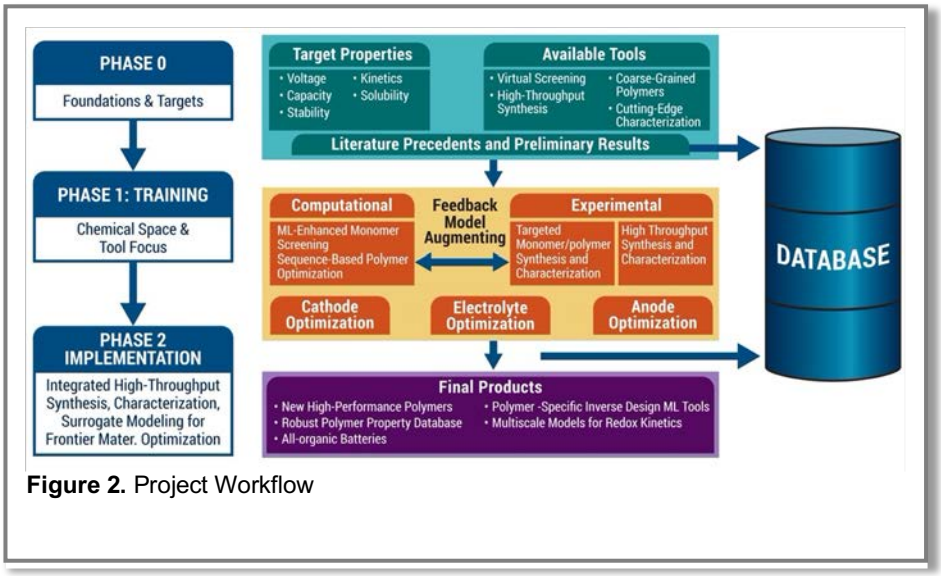


Figure 2. Project Workflow

Broader Impacts and Workforce Development

We are beginning to launch our plans for broader impacts and workforce development. This project will provide educational training opportunities in synthetic chemistry, polymer science, electrochemistry, and computational chemistry and physics specific to organic batteries. Education and outreach activities will be jointly developed and deployed at The University Chicago's No Small Matter Molecular Engineering Fair, Texas A&M University's Physics & Engineering Festival, and other venues. Workforce development is planned through participant mentoring and workshops targeted to industrial, academic, and government researchers.

Data Management and Open Access

We are currently using an internal GitHub repository for exchanging data and synchronizing the computational efforts across the team. In the future, both our experimental and computational data will be released via the Materials Data Facility, a platform that provides public access to its contents and issues a DOI for its entries, thereby facilitating use, exchange, and discovery of datasets. Moreover, the computational and machine learning tools that we will develop will be made available via public GitHub repositories, while code releases will be also archived in the MDF. The public GitHub repositories will allow the community to more easily re-use and build on our code.

Advancing Along the Materials Development Continuum and Partnerships to Translation

The MGI approach where synthesis, characterization, and simulation are tightly integrated with close feedback loops is key to designing promising organic radical polymer battery materials on a less than decade timescale. Without this tight integration, we would be limited to selected fundamental theoretical and experimental studies on only a few model systems, and the design process would evolve too slowly. In terms of commercialization prospects, metal-free materials are necessary to reduce the risks associated with current supply chains for metal-based battery systems. We intend to explore translation for certain chemistries that meet our design goals for open circuit voltage, conductivity, and stability.

Publications and References

None to report yet, as this project is still in its first year.

Designing Plasmonic Nanoparticle Assemblies For Active Nanoscale Temperature Control By Exploiting Near- And Far-Field Coupling

Lead Investigator: David J. Masiello, masiello@uw.edu
Participating Institutions: University of Washington, Temple University, Rice University
Website: https://faculty.washington.edu/masiello/DMREF/
Keywords: Thermoplasmonics, lattice plasmons, surface lattice resonances, photothermal imaging, super-resolution imaging

Project Scope

The ability to spatially localize heat flow and thus temperature at both nanoscale (<100 nm) and micron-scale (~1-1000 μ m) dimensions has important implications for applications ranging from big data to nanomedicine. The goal of this DMREF-supported work is to overcome thermal diffusion and exercise precise, spatially non-uniform temperature control using only light as a heat source. By arranging plasmonic metal nanoparticles into well-defined configurations and using specific wavelengths/polarizations of light, the local thermal profiles can be actively manipulated, offering long-range global control over spatially-nonuniform thermal gradients spanning from nano- to micron-scales

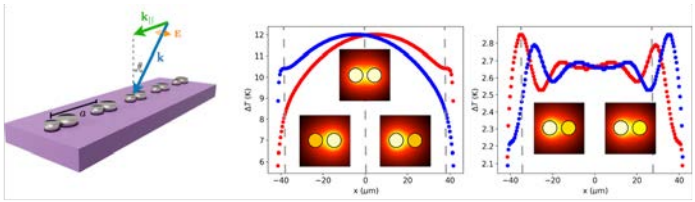


Figure. Diffractively-coupled plasmonic metal nanoparticle lattices (left panel) are predicted to offer strong potential for the creation of actively-tunable thermal gradients spanning from 10s of nanometers to 100s of microns using only continuous-wave light. For example, with each unit cell composed of a strongly-coupled plasmonic nanoparticle dimer (each nanoparticle having ~50 nm diameter), the spatial profiles of both intra- (middle/right insets) and inter- (middle/right blue and red traces) unit-cell thermal gradients ($\Delta T(x)$) are predicted to be optically controllable despite the deleterious effects of heat diffusion. The middle and right panels show numerical calculations of the spatially-dependent temperature increases spanning the lattice when it is optically excited at the Γ-point at the in-phase (middle) and out-of-phase (right) surface lattice resonances of the 80 micron dimer lattice. These numerical results indicate the surprising ability to create actively-tunable, tailored thermal profiles in steady-state with both nanoscale and micronscale features using light.

Relevance to MGI

To design thermal metamaterials, theory takes a lead role in devising structures that show the ability to produce localized, spatially non-uniform temperature profiles using only light. Because light is used to manipulate temperature profiles, a single structure allows us to probe many different temperature configurations simply by changing the properties of the excitation source, allowing theoretical predictions to be tested and refined without the need for expensive, time-consuming fabrication and letting the team quickly converge on optimized structures.

Technical Progress

In the previous funding cycle, the team demonstrated the ability to localize and control temperature on the nanoscale using dimers and trimers of gold nanorods, arranged in strategic patterns in order to create wavelength-dependent non-uniform temperature profiles. Photothermal imaging was used to validate predictions from theory, showing perturbations in the point spread function of diffraction-limited images that were linked to asymmetric temperature profiles. New theoretical tools were developed to understand these results and refine our experimental understanding of temperature information encoded in photothermal images. The team also demonstrated a super-resolution imaging strategy to measure light-induced temperature changes with <20 nm spatial resolution.

As we move towards nanoparticle unit cells arranged in periodic lattices, we require nanoparticles that are large enough to scatter light, which allows the nanoparticles to couple to the lattice, but also small enough to absorb power from the light field, which is converted into both local and global temperature increases. We have combined theoretical modeling with experimental nanoparticle synthesis to elucidate the behavior of measured photothermal signals for nanoparticles in the size regime where both scattering and absorption are significant. To date, all past work in the photothermal imaging community has focused on the small (absorption only) nanoparticle limit where

scattering from the heated nanoparticle can be ignored. However, in our materials, scattering cannot be neglected. We have probed the nanoparticle size-dependence of the signal and discovered a surprising signal nonlinearity versus pump power in the larger nanoparticle regime relevant to our photothermal metamaterial systems. We are now incorporating this knowledge into the analysis of our team's first plasmon lattice formed via nanosphere lithography. In this spatially extended periodic system, we are finding that both the electromagnetic coupling and thermally conductive coupling between gold particles within the lattice is essential to describe the lattice's temperature rise under continuous wave optical illumination with the temperature read out via DNA melting nanothermometry techniques developed by our DMREF team.

Future Plans

Future plans include the fabrication of precision-engineered Bravais and non-Bravais plasmonic nanoparticle lattices designed to realize specific thermal profiles under optical illumination when excited at specific (k,ω) points in the Brillouin zone, leading to nanoscale temperature control across 100s of microns. Another future plan includes the design of plasmonic lattices composed of nanoparticle dimer unit cells that are tuned into a topologically nontrivial regime. Our theoretical predictions indicate that optically-induced thermal gradients can be strongly localized to the edges of the lattice without appreciable penetration into the lattice interior. Such a metamaterial will be highly unusual as it will localize increased temperature only to its surfaces and not to its interior (or bulk).

Broader Impacts and Workforce Development

Our team is training the next generation workforce not only through the specific research skills obtained by members of each of the participating labs, but also through the highly interdisciplinary and collaborative nature of the work. Our team meets every two weeks to share results over zoom, and we have recently been able to initiate travel between the participating institutions (previously hindered by the pandemic). Students and postdocs have presented DMREF-supported work at ACS National Meetings and Gordon Research Conferences, and the PIs are co-organizing a symposium at the Fall ACS National Meeting, providing multiple platforms for disseminating the MGI mission and supported work. The team is also planning to host a summer workshop on photothermal processes, bringing together team members with world leaders in the field. We are working with the University of Washington Clean Energy Institute staff to design a Nanotechnology summer camp module for middle school students with focus on photothermal materials, with the intention of transitioning this program to other participating institutions.

Data Management and Open Access

As part of our original DMREF project, we have established a team website (https://faculty.washington.edu/masiello/DMREF) that is maintained by the Masiello group. Our team has also created a profile at https://dmref.org/. The source codes and all relevant operating instructions for the numerical implementations of our thermal discrete dipole approximation (*T*-DDA) and photothermal imaging pipeline codes are now publicly accessible and freely downloadable from the Masiello group GitHub repositories: https://github.com/MasielloGroup/t-dda https://github.com/MasielloGroup/confocal_photothermal_pipeline. Our team website also includes all publications, team contact info, and press resulting from DMREF funding. This website will be augmented to incorporate all of the experimental protocols, procedures, and results originating from these research activities. Appropriate links to external data repositories housing any additional computer code as well as all raw and processed experimental data will be provided.

Advancing Along the Materials Development Continuum and Partnerships to Translation

We are still in the basic research stages of our project, and are not ready for commercialization. However, the MGI approach has drastically accelerated our ability to discover, develop, and deploy new materials with actively controllable thermal properties using light. Barriers to commercialization lie in finding the right industrial need and market for such thermal metamaterials.

Publications and Reference

1. Journal of Physical Chemistry B, **125**, *pp. 12197-12205* (2021).
2. Nano Letters, **21**, *pp. 5386-5393* (2021).
3. ACS Photonics, **8**, *pp. 1159–1168* (2021).

Computational Discovery of Polymeric Membranes for Polar Solvent Dehydration

Lead Investigator: Clare McCabe, c.mccabe@vanderbilt.edu
Participating Institutions: Vanderbilt University
Website: https://osf.io/eyt3j/
Keywords: polymer membranes, screening, molecular simulation, machine learning, data mining

Project Scope

The rational design of the next generation of membranes for solvent-solvent separations is a significant challenge, given the vast chemical and design space, but could be realized by Materials Genome Initiative (MGI)-inspired screening and has the potential to transform the membrane development paradigm. This project is developing functionality- and performance-driven screening with close coupling between simulations and experiment to tailor high-performance membranes for targeted separations. Specifically, we will focus on the dehydration of polar solvents by pervaporation where effective new materials can eliminate the need for high-cost and high-energy separations while enabling effective solvent reuse for sustainable manufacturing.

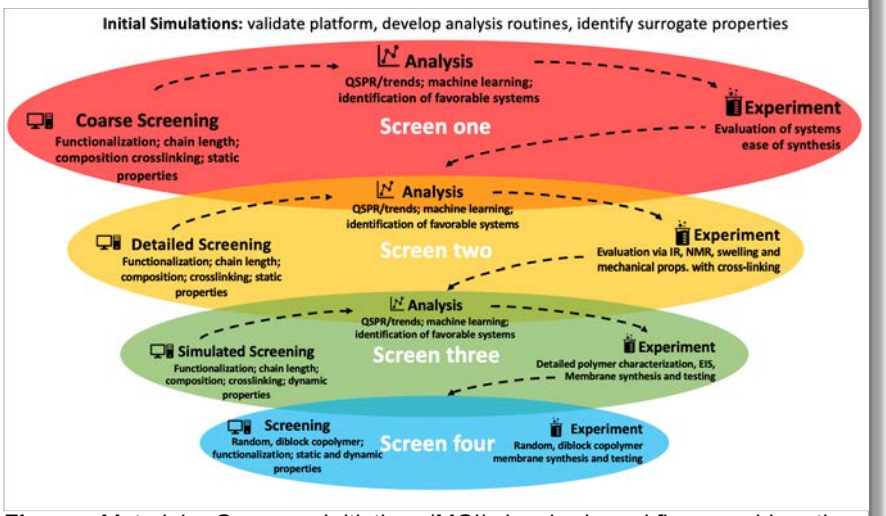


Figure. Materials Genome Initiative (MGI) inspired workflow enables the integration of computational and experimental data to predict high-performance membrane compositions for specific polar solvent dehydration.

Relevance to MGI

We are developing an integrated MGI-inspired computational and experimental screening platform, with the goal of accelerating the rational design of membranes for the dehydration of polar solvents. This will be achieved through the combination of extensive molecular simulations using the Molecular Simulation and Design Framework (MoSDeF), machine learning using DeepForge, and experiments based around the combination of ring-opening metathesis polymerization chemistry combined with spin coating. Our versatile approach will enable the synthesis of a wide array of polymer membrane compositions from a common central scaffold and the integral synergy between molecular-level screening simulations and experiment to molecularly design and identify new membrane compositions that are tailored to specific dehydration separations. The goals of this project are the identification, synthesis, and testing of new candidate polymers for polar organic solvent-water membrane separations, development of a robust library of polymer membrane properties, development of machine learning models that relate chemistry to measured properties of membrane films, and release of a generally applicable set of software tools that will enable rapid screening and machine learning studies on soft-matter systems.

Technical Progress

We have developed a method called spin coating ring opening metathesis polymerization (scROMP) that enables the fabrication of polymeric membranes simply and rapidly with control over membrane thickness by varying spin speed. The current "scaffold" polymer contains two acyl chlorides in each repeat unit for straightforward conversion by amines and alcohols to a wide variety of polymer film compositions. The initial

focus has been on modification of the scaffold polymer film with taurine and phenylenediamine to generate charged, cross-linked membranes that are designed to favor water over polar organics. Simulations are examining the structure and interactions of oligomers of this same composition in water versus polar organics. Results from the simulations will ultimately guide the choice of polymer compositions and the cross-linking levels to achieve highly selective transport of water from the polar organic.

Future Plans

Through the integrated screening approach shown in the figure, the objectives of the proposed research are:

1. Screen 1: Based on a polymer backbone that is compatible with the sc-ROMP process, use molecular simulations to perform a *coarse* screen of the design space to determine each candidate's potential to separate water via pervaporation and determine which of the potential systems have viable synthesis routes.
2. Screen 2: Computationally perform a detailed screen over the candidates selected from screen one. Experimentally synthesize the promising candidates as they are identified to verify expected compositions via IR spectroscopy and NMR.
3. Screen 3: Evaluate polymer membrane candidates identified from screen two. Computationally evaluate dynamic in addition to static properties. Experimentally characterize structure, swelling, and physicochemical properties. Examine the effect of cross-linking on structure and mechanical properties by tuning the ratio of monomer to cross-linker. Begin preliminary membrane testing.
4. Using molecular simulation data combined with experiment, implement machine learning algorithms to determine key relationships in the screening design space between water and candidate solvent molecules, using a broad array of film compositions to identify films with high potential selectivities (screens 1– 3).
5. Screen 4: Fabricate and characterize polymer membrane candidates identified in screen 3 and from ML studies. Test their ability to separate water via pervaporation from target polar organics with high selectivity and flux.

Broader Impacts and Workforce Development

This project, with its integration of computational modeling, machine learning, material synthesis, characterization, and performance evaluation in targeted applications, will serve as an excellent educational platform for participating graduate students and postdoctoral researchers to experience the full suite of interconnected components described in the MGI vision. By developing competency in the three foundational pillars of experiment, computation, and data science, the proposed project will help develop a workforce aligned with the MGI vision. Additionally, multiple integrated educational activities at the undergraduate and K-12 levels will highlight the potential of computational materials science and the need for close coupling with experiment and data science, inspiring the next generation of the MGI workforce.

Data Management and Open Access

To facilitate sharing and dissemination of data we have created a publicly accessible project on the Center for Open Science's OSF.io web portal (https://osf.io/eyt3j). OSF.io will provide a central location for linking to content, including anticipated datasets on zenodo.com, and will include direct integration with GitHub.com, where our software is being developed, and figshare.com where we will post preprints and relevant presentations/posters, including detailed experimental data.

Advancing Along the Materials Development Continuum and Partnerships to Translation

The integration of computational approach to identify optimal polymer compositions along with a new membrane fabrication process and highly diverse polymer chemistry will enable the rapid discovery of new materials for membrane separations. We have already filed a patent on the scROMP process: G. K. Jennings and L. Prozorovska, "Polymer Film Preparation Using Cyclic Olefin Monomers," U.S. Application Number 17/583,370 filed on Jan. 25, 2022.

Publications and References

None yet.

Quasi-Direct Semiconductors

Lead Investigator: José Menéndez, Arizona State University
Participating Institutions: Arizona State University, University of Texas at Austin.
Website: None.
Keywords: quasi-direct, semiconductors, electron-phonon, optical properties

Project Scope

Quasi-direct semiconductors have a direct band gap slightly above the fundamental indirect band gap. Over 200 such compounds appear in the Materials Project database. Surprisingly, there is no satisfactory theory of optical absorption in these materials: the standard perturbation theory approach breaks down when the photon energy reaches the direct band gap, causing an unphysical infinite absorption. Our program will develop a new *ab initio* theory and computational methods to calculate absorption and emission spectra from quasi-direct semiconductors and will validate the new theoretical tools by studying the optical properties of samples fabricated using novel chemical precursors.

Relevance to MGI

This project combines the expertise of an experimental team that recently carried out an in-depth characterization of the quasi-direct optical properties of Ge with the theoretical and computational know-how of a leading group in the study of the electron-phonon interaction and its impact on optical properties. The theoretical team will develop a new quasi-degenerate many-body perturbation theory framework to capture both direct and indirect phonon-assisted absorption on the same footing. The theoretical predictions will be validated by carrying out high precision measurements of the absorption, initially in Ge and subsequently on other materials within an improvement experiment-theory loop. Initial benchmarking of the Ge optical properties has already been performed by studying to what extent the standard theory of excitonic direct absorption reproduces the experimental dielectric function above the direct gap. Once an accurate model is established the theoretical team will perform high throughput screening of attractive quasi direct materials for synthesis by the experimental group, and both teams will combine their expertise to parametrize simplified models suitable for fitting experimental data. It is anticipated that the software module implementing the new *ab initio* methodology will be released as part of the open-source EPW software project (https://epw-code.org).

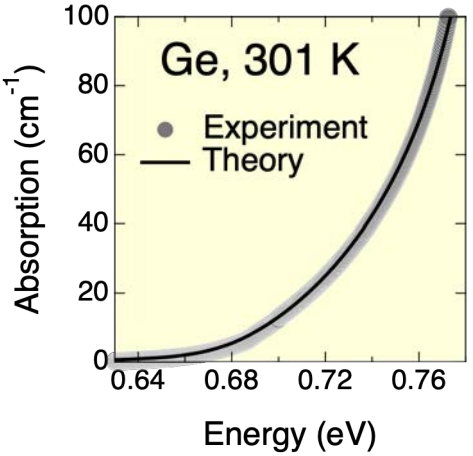


Figure. Indirect absorption in Ge is resonantly enhanced as the photon energy approaches the direct gap at 0.80 eV. The standard theoretical expression (solid line) diverges at the direct gap and beyond. A new theoretical approach is needed. Experimentally, it is very difficult to separate the contributions from resonantly enhanced indirect absorption from bona fide direct absorption. Furthermore, as the absorption coefficient exceeds 100 cm^{-1}, it becomes increasingly difficult to makes samples that are thin enough for accurate transmittance/reflectance measurements.

Technical Progress

Initial work included a study of the dielectric function of Ge near the direct band gap. Experimental data were fitted with a realistic expression which for the first time did not include adjustable "amplitude" or "phase" parameters. Progress was also made developing a new technology to grow Ge films on Si using chlorogermane as a CVD precursor. These led to two publications as listed below. Initial work on the theoretical side included the reformulation of the theory of indirect absorption in the language of many-body quantum field theory, and a preliminary design of its *ab initio* implementation in EPW.

Future Plans

In the near future the experimental work will focus on Ge, GeSi and GeSn alloys. Films with thicknesses exceeding 3 μm will be grown on Si, strain-relaxed by annealing, and their optical properties will be determined by transmittance/reflectance measurements as well as spectroscopic ellipsometry. In parallel with these efforts, *pin* diodes will be fabricated and their responsivity will be measured. For this purpose, we are developing simple by accurate codes to calculate the I-V characteristics of these diodes, including all important recombination mechanisms and optical interference effects between all layers in the structure. A "universal" model for the I-V characteristics of Ge diodes has already been completed and will be submitted for publication shortly. Surprisingly, no such model is currently available, despite the central historical importance of Ge in semiconductor science. During Y2 the theory team will focus on the implementation of the new theory of optical absorption in the EPW code, and benchmark this implementation against accurate experimental data available for silicon and germanium.

Broader Impacts and Workforce Development

At ASU, two graduate students are supported by the project. A course on the physics of p-n junctions was taught for the first time, with undergraduate student participation. One of the undergraduates has continued his involvement with the team and is participating in a pedagogical manuscript on diode physis that will be submitted to the American Journal of Physics. At UT Austin, one postdoc and one undergraduate student are supported by this project. During the week 13-19 June 2022, the UT Austin Co-PI hosted an in-person summer school on electron-phonon physics from first principles, involving both theory lectures from experts, hands-on tutorials on open-source codes, and a weekend hackathon to promote programming literacy and high-performance computing among students. This school included lectures and tutorials on phonon-assisted optical absorption. A completed set of slides, assignments, and videos can be found at https://epw2022.oden.utexas.edu/74-schedule.

Data Management and Open Access

The main outputs of this project will be (1) accurate measurements of the optical absorption coefficient in quasidirect semiconductors, (2) development of a new software module for performing *ab initio* calculations of optical absorption and emission in quasidirect semiconductors, (3) Optical absorption calculated from first principles for several quasidirect semiconductors using the new module. The new EPW module is currently being written, and will be released after publication of the corresponding manuscript, likely during Y2. The UT Austin Co-PI has a long-standing track record of development of open-source software, for example the EPW code for electron-phonon physics (https://epw-code.org/) and the SternheimerGW code for GW quasiparticle calculations (http://www.sternheimergw.org/). Both are open-source projects released under the most permissive GNU GPL license. They can be downloaded and used by anyone without limitation or registration. The team supports a growing community via the user forum (https://forum.epw-code.org/) which has over 650 registered users.

Advancing Along the Materials Development Continuum and Partnerships to Translation

The team has already developed new strategies for semiconductor growth as a byproduct of the effort to develop suitable samples for optical measurements. The further development of these materials for optoelectronic applications is being pursued via SBIR/STTR programs in collaboration with local companies in the Phoenix metropolitan area. New collaborations will be pursued once the project advances and the application potential of quasi-direct semiconductors becomes more apparent.

Publications and References

1. C. Emminger, N. S .Samarasingha, M. Rivero Arias, F. Abadizaman, J. Menéndez, S. Zollner, *Excitonic effects at the temperature-dependent direct bandgap of Ge*, J. Applied Physics **131**, 165701 (2022). DOI: 10.1063/5.0080158.
2. A. Zhang, M. Mircovich, S. Ringwala, C.D. Poweleit, M. Roldán, J. Menéndez, J. Kouvetakis, *Ultra-low temperature synthesis of Ge based optical materials and devices on Si using GeH_3Cl*, submitted to ACS Applied Electronic Materials. (2022).

DMREF: III-Nitride Monolayers and Extreme Quantum Dots

Lead Investigator: Zetian Mi, ztmi@umich.edu
Participating Institutions: University of Michigan, Ann Arbor
Website: none.
Keywords: GaN, two-dimensional semiconductor, quantum dot, exciton, molecular beam epitaxy.

Project Scope

This project will create a new family of nano-quantum materials (nQMs), including two-dimensional (2D) III-nitrides and extreme quantum dots (XQD). Our team predicted that the Coulomb interaction in 2D III-nitrides is so strong that it binds electrons and holes into excitons with a binding energy exceeding 1 eV. We will establish an nQM platform with unique quantum functionalities by using the most accurate quantum theory to design the quantum optoelectronic properties of nQMs. Our breakthroughs will transform conventional low-efficiency AlGaN into high-brightness deep ultraviolet (UV) emitters and establish the first controllably interacting semiconductor qubit platform for quantum information science application.

Relevance to MGI

We propose a focused theory-epitaxy-characterization-quantum optoelectronics collaboration. We will develop a systematic quantum theory that integrates three first-principles methods, including Density-Functional Theory, Many-Body Perturbation Theory, and Quantum-Dynamic Cluster Expansion, to precisely predict and determine the electronic, optical, excitonic, and entanglement properties of 2D III-nitrides and XQDs. Atomically thin III-nitrides and XQDs with superior quality will be realized by combining ultrahigh temperature molecular beam epitaxy and selective area growth with polarization, surface, polarity, and quantum engineering. The presence of disorder and spectral diffusion, deleterious effects that plague conventional QDs, will be addressed by i) deterministically growing and positioning XQDs on hBN whose atomically smooth surface eliminates interface interdiffusion and disorder, and ii) creating interface excitons with well-aligned dipole and with identical emission that is largely determined by the band alignment of AlN/hBN. We combine time-resolved deep-UV photoluminescence spectroscopy and differential transmission spectroscopy to demonstrate controllable XQD–XQD coupling and controlled formation of direct and indirect excitons, biexcitons, and dropletons at room temperature. The potential of the nQMs platform will be realized by taking key steps toward extraordinary quantum optoelectronic devices, including deep UV light emitters, entanglement detectors, and coupled XQD systems.

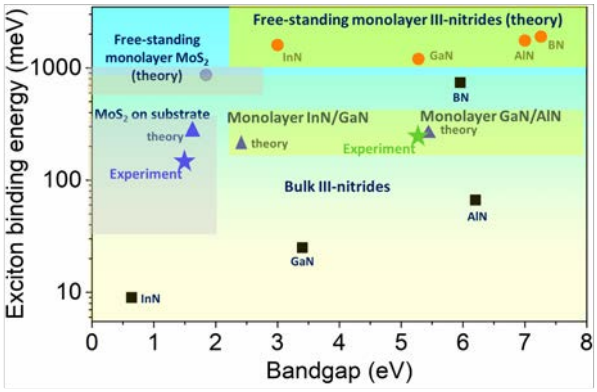


Figure. Diagram of exciton binding energy vs. energy bandgap for various 2D semiconductors including the emerging 2D III-nitrides as well as conventional 2D TMDs.

Technical Progress

Since the beginning of this project in 2021, our team has performed detailed studies of the theory, epitaxy, and characterization of monolayer III-nitrides and extreme quantum dot heterostructures. We have developed the first method for the epitaxy of pristine, wafer-scale monolayer hBN on graphene. We discover that the in-plane hBN/G interface can be precisely controlled, enabling the scalable epitaxy of unidirectional monolayer hBN on graphene, which exhibits a uniform moiré superlattice consistent with single-domain hBN, aligned to the underlying graphene lattice. Furthermore, we identify that the deep-ultraviolet emission at 6.12 eV stems from the 1s-exciton state of monolayer hBN with a giant renormalized direct bandgap on graphene.

Future Plans

Quantum Theory and Design. We will develop a systematic quantum theory to guide experiments via precise predictions on how the many-body dynamics of excitations drives UV emission as well as qubit-qubit interactions in them. We will compute the band structure and interaction matrix elements (Coulomb, dipole, and phonon) of nQMs by combining Density-Functional Theory (DFT) and Many-Body Perturbation Theory (MBPT). We will then use this input in Quantum-Dynamic Cluster Expansion (QDCE) to deliver extremely efficient and accurate predictions for the quantum kinetics of quasiparticles.

Epitaxy of III-Nitride Monolayers and XQDs. We will investigate the synthesis of atomically thin GaN, InN, and hBN, and the deterministic formation of XQDs. The presence of disorder and spectral diffusion, deleterious effects that plague conventional quantum dots, will be addressed by i) deterministically growing and positioning XQDs on hBN whose atomically smooth surface eliminates any interface interdiffusion and disorder, and ii) creating interface excitons with well-aligned dipole and with identical emission.

Optical Spectroscopy. Optical properties of III-nitride monolayers and QDs will be studied by using ultrafast spectroscopy to reveal the strong Coulombic many-body interactions. The formation of quantized direct and indirect excitons, biexcitons, and dropletons will be experimentally measured by combining time-resolved PL spectroscopy and differential transmission spectroscopy and by studying 2-photon quantum-light excitation.

Quantum Optoelectronics. Unique properties of 2D III-nitride nanostructures will be tested and verified through a few device demonstrations, including deep UV light emitters and superradiance and controlled qubit interaction.

Broader Impacts and Workforce Development

This project will pioneer 2D III-nitride nQMs with quantum optoelectronic properties enabling novel applications ranging from high-efficiency deep UV light emitters to controlled qubit interactions. The broader impacts also include the highly interdisciplinary nature of this project and outreach to undergraduates, underrepresented minorities, and K-12. The education and outreach program includes 1) encouraging underrepresented minorities and women in careers in science and engineering through the appealing potential social impacts of the research, 2) involving undergraduate students in frontier research, and 3) communicating the research to the general public, including the popular "Saturday Morning Physics" public lecture series at the University of Michigan and the recently established Midwest Quantum Collaboratory to reach out to the broader public, particularly underrepresented groups on quantum science and technology.

Data Management and Open Access

During this project, data from quantum theory, design and modeling, epitaxy, structural and optical characterization and analysis, and device processing and measurements will be generated. Data that are of high impact results will be made available to the general public through peer-reviewed journal and conference publications (including supplementary material), seminars, and presentations. More detailed data analysis is also published after the completion of each project, which may take the form of final project reports, master and doctoral thesis, books, and book chapters.

Advancing Along the Materials Development Continuum and Partnerships to Translation

Some IP related to this project has been licensed to NS Nanotech, Inc., which was a spin-off from the University of Michigan with Mi being a co-founder.

Publications and References

1. P. Wang, W. Lee, J.P. Corbett, W.H. Koll, N.M. Vu, D.A. Laleyan, Q. Wen, Y. Wu, A. Pandey, J. Gim, D. Wang, D.Y. Qiu, R. Hovden, M. Kira, J.T. Heron, J.A. Gupta, E. Kioupakis, Z. Mi, "Scalable Synthesis of Monolayer Hexagonal Boron Nitride on Graphene with Giant Bandgap Renormalization", Advanced Materials, 2201387, 2022.

Localized Phase Transformation (LPT) Strengthening for Next-Generation Superalloys

Lead Investigator: Michael Mills, mills.108@osu.edu
Participating Institutions: The Ohio State University, University of Michigan, GE Research, NASA Glenn Research Center, Air Force Research Laboratory
Website: https://github.com/mesoOSU.
Keywords: Alloy development, superalloys, phase transformations,

Project Scope

This project exploits a new design strategy that utilizes a localized phase transformation (LPT) phenomenon to disruptively improve the high-temperature creep performance of Ni-base superalloys. LPTs occur only at extended defects, limiting deformation along these defects, and are confined locally at these defects, avoiding the formation of detrimental phases in the bulk of the alloys. This project integrates sophisticated computational models, at multiple scales, advanced materials characterization techniques, and combinatorial and accelerated methods of materials processing and property evaluation to exploit the connection between defect type and the nature of LPTs for development of new high-performance superalloys.

Relevance to MGI

In collaboration with GOALI partner GE Research and supplemental funding partners NASA Glenn Research Center and Air Force Research Labs (AFRL), we are (a) searching the alloy composition space for possible LPT by classical thermodynamic analysis, high-throughput DFT, Monte Carlo calculations, using CALPHAD modeling and machine learning. The computational predictions are validated by advanced electron microscopy and atom probe tomography; (b) establishing the connection between defect type and the nature of LPTs using phase field simulations with the DFT calculations as inputs; and (c) quantitatively determining the effect of LPTs on the operative deformation mechanisms and develop physics-based deformation models that capture these effects. Our partners will respond to recommendations by producing alloys in single crystal and polycrystal forms. Such an integrated research effort will raise significantly the state-of-the-art in the discovery and development of new superalloys and will result in new science in alloy design.

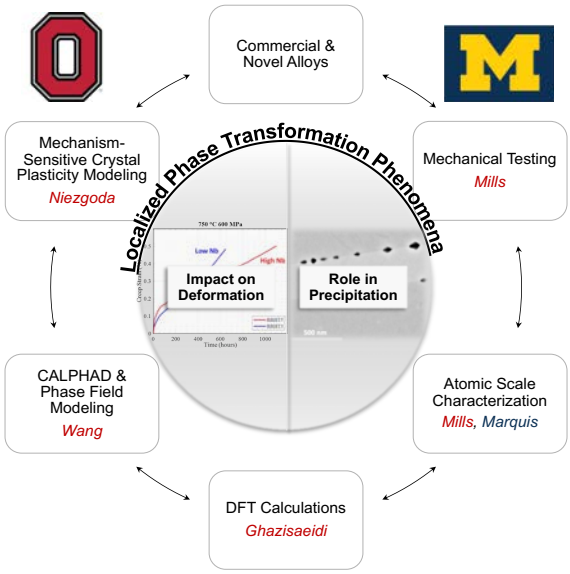


Figure 4 OSU X UMich Discovery Loop: Outer flowchart delineates all the experiments being undertaken to investigate localized phase transformations in commercial and novel alloys. In addition, the inner circle contains the two overarching themes guiding these experiments.

Technical Progress

A Monte Carlo based method has been developed for fast prediction of coexisting phases in multicomponent systems and have been applied to investigate the formation of both η and δ phases in bulk regions. Examination of LPT phenomena is ongoing.

First principles (DFT) calculations have been performed on the interaction energies between solutes and the stacking faults with the presence of other solute atoms, addressing the synergistic effects between the alloying elements. DFT calculations have also been utilized to assist the characterization of the phases forming at the stacking fault regions, providing insight into the impact of LPT and deformation mechanisms.

3-dimentional atom-probe tomography characterization of atomic-scale structure and quantitative analysis of elemental segregation to stacking faults and twin boundaries have been carried out on different alloys.

A new alloy composition was selected based on our thermodynamic parameter (ordering tendency) calculated through CALPHAD database. This alloy should have even more enhanced η-hardening than the current, best LPT-strengthened alloy. The selected alloy is being manufactured by our collaborators at NASA Glenn Research Center. Collaboration with GE has also revealed the remarkable preferential precipitation of γ' precipitates that occurs on faults after processing, specifically annealing twin boundary in alloy 718 and 718-variant alloys. Preliminary

evidence indicates that formation of χ-phase LPT at annealing twin boundaries leads to the preferential nucleation of γ' particles and consequent intense local shearing adjacent to these boundaries. A thermodynamic model supported by CALPHAD calculations has been postulated to explain this effect and is now being evaluated relative to a wide range of commercial and experimental alloys.

Future Plans

Evaluation of creep response has begun on single crystals grown and provided by GE Global along with commercial alloy ME501 and recently developed NA1. Differences in creep behavior should relate directly to propensity for LPT. In addition, two monomodal versions (large vs small precipitates) of polycrystalline ME3 have been jointly developed which seek to evaluate the effect of precipitate size of degree of segregation to faults during γ LPT. To expand the understanding of LPT strengthening in Ni-base superalloys, UM will further study the planar faults via STEM and APT to examine the local structure and chemistry at stacking faults to explain their high-temperature creep properties difference. The practical limits of LPT strengthening is also being explored in alloys for which thicker (i.e. non-LPT) χ-phase forms during creep deformation and is deleterious to ductility. A mechanistic model for this limiting behavior is being developed with input from DFT calculations and phase field modeling.

Broader Impacts and Workforce Development

Faculty and students working on the project are engaging with GE Global Research, NASA Glenn Research Center and AFRL staff and facilities in the program. PIs are providing opportunities for undergraduate students to participate in interdisciplinary senior capstone projects that involve integrated materials modeling and experimentation. In the past, the undergraduates involved in these projects were not only gaining valuable research experience, but they were also helping us (including the graduate students working on our previous DMREF project) in outreach activities for K-12 students from local high schools who attended tours of the Center for Electron Microscopy and Analysis (CEMAS) and introductions to engaging materials science topics. DMREF PIs continue to host high school interns in their labs on DMREF materials simulation projects.

Data Management and Open Access

Data and software is currently available from Co-PI Niezgoda's GitHub repository https://github.com/mesoOSU. As the project enters its final years, data will be moved to community repositories for archiving.

Advancing Along the Materials Development Continuum and Partnerships to Translation

The discovery of LPT strengthening by the OSU-DMREF program has opened the door to new avenues for alloy development and property improvement. This advancement has energized GE Research and NASA Glenn researchers and has stimulated new directions for alloy development. DMREF and supplemental funding will continue to support experiments and modeling that will enable assessment of the behavior of new alloys that have been recently postulated and produced by NASA Glenn, but without the computational and experimental support from the MGI process would not be able to move towards commercialization in a timely manner.

Publications and References

1) Feng, L., Mills, M. J., & Wang, Y. (2021). Generalized stacking fault energy surface mismatch and dislocation transformation. *npj Computational Materials*, *7*(1), 1-8.
2) Zenk, C. H., Feng, L., McAllister, D., Wang, Y., & Mills, M. J. (2021). Shearing mechanisms of co-precipitates in IN718. *Acta Materialia*, *220*, 117305.
3)Qiu, D., Zhao, P., & Wang, Y. (2021). A general phase-field framework for predicting the structures and micromechanical properties of crystalline defects. *Materials & Design*, *209*, 109959.
4) Qiu, D., Zhao, P., Trinkle, D. R., & Wang, Y. (2021). Stress-dependent dislocation core structures leading to non-Schmid behavior. *Materials Research Letters*, *9*(3), 134-140.
5) Sun, X., Lu, S., Xie, R., An, X., Li, W., Zhang, T., ... & Vitos, L. (2021). Can experiment determine the stacking fault energy of metastable alloys?. *Materials & design*, *199*, 109396.
6) Mukhopadhyay, S., Sriram, H., Zenk, C. H., DiDomizio, R., Detor, A. J., Hayes, R. W., ... & Mills, M. J. (2021). Creep Behavior of Compact γ′-γ ″Coprecipitation Strengthened IN718-Variant Superalloy. *Metals*, *11*(12), 1897.

Accelerated Data-Driven Discovery of Ion-Conducting Materials

Lead Investigator: Dr. Yifei Mo, yfmo@umd.edu
Participating Institutions: University of Maryland, Northwestern University, Lehigh University
Website: none
Keywords: *ion-conducting oxides, microstructure design, closed-loop, high-throughput materials discovery*

Project Scope

The object of this DMREF project is to integrate the rapid synthesis of ceramics with first-principles data-driven computation, high-throughput measurements and characterization, and microstructural modeling into a closed-loop framework to significantly accelerate the discovery and design of new ceramic oxide materials using Na-ion conductors as model systems. Project scope includes creating 1) an integrated closed-loop framework for accelerated discovery and rational design of ion-conducting materials by simultaneously considering composition, microstructure, and sintering conditions to achieve optimal ion-conducting properties; 2) novel ion-conducting materials systems; and 3) new knowledge and a publicly shared database encompassing a wide range of oxide materials, including composition-sintering-microstructure-property relationships, design rules for ion-conducting materials, and microstructure sintering models. Project hypothesis: if the conventionally sluggish and time-consuming synthesis and sintering methods of ceramic oxides is replaced with the rapid, novel Ultrafast High-temperature Sintering (UHS) method, predictive material design, and high-throughput characterization, then can we accelerate the discovery of novel ion-conducting oxide materials. Success will be measured in terms of closed-loop efficiency and fidelity, *i.e.,* how quickly and accurately all three collaborating institutions can predict, fabricate, and characterize Na-ion conductor compounds, including interactive steps for optimization. A tangible goal will be the outcome of new Na-ion conducting oxides with new compositions and microstructures achieved by rationally guided sintering conditions of UHS.

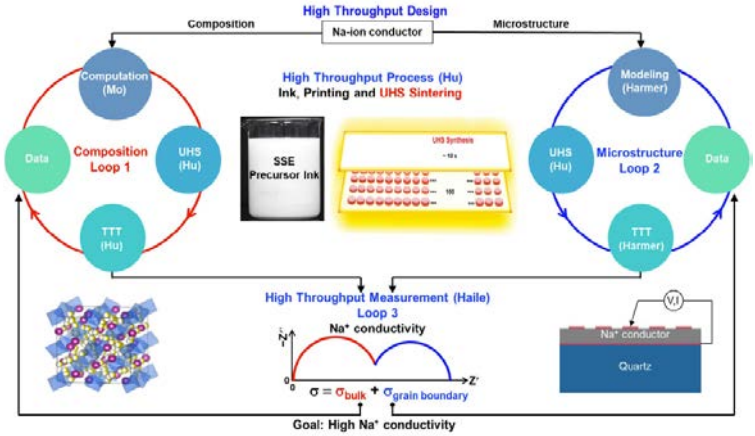


Figure 5. Closed-loop approach for materials discovery and design integrating high-throughput rapid synthesis, measurements, computation, and modeling.

Relevance to MGI

The proposed integrated framework of data will be used to significantly accelerate the discovery and rational design of new materials with improved performance in a fraction of the time. **Loop 1** (Composition discovery loop), using first principles computation, in combination with multi-level high-throughput computation screening and a data-driven machine learning approach, we will predict previously unexplored compounds that exhibit fast ion conduction. The computational predictions of new Na-ion conducting compositions will be rapidly validated by syntheses and characterization of UHS-prepared materials. The experimental results (either validated or unsuccessful) will be used as feedback for the computation. **Loop 2** (Microstructure loop) will guide the rational design of oxide microstructures under different UHS sintering conditions. Through an iterative feedback loop between UHS, characterization, and modeling, we will achieve greater knowledge of the microstructural evolution in UHS and establish Time-Temperature-Transformation (TTT) diagrams of the materials and microstructures. **Loop 3** (Combined loop of composition, microstructure, and sintering) will establish a unified closed loop for designing new ion-conducting materials systems with simultaneous consideration of composition, microstructure, and synthesis conditions.

Technical Progress

Since the issuance of this award, scientific progress has been made in fulfilling the goals of closed-Loop 1. Ionic conductivity data is currently being collected for Na-ion conducting materials modified from the Inorganic Crystal Structure Database (ICSD), while characterization of fabricated samples is used as validation data back to machine learning. A key improvement of Loop 1 has been the automation of the UHS sintering voltage profiles in

which ceramic oxides are fabricated under computer-controlled sintering parameters for improved sample quality and consistency.

Future Plans

Immediate plans include continued optimization of Loop 1, as well as advancement into Loop 2. The main aspects of loop 2 are described in three tasks: 1) characterize bulk microstructures and grain boundary complexions, 2) employ 3D Monte Carlo grain growth and grain boundary transformation models, and 3) design/optimize/evaluate new materials via UHS. In addition, as some of the predicted compositions from Loop 1 have never been synthesized, the first iteration of Loop 2 will study known model systems (b/b"-Al_2O_3 or NASICON) to ensure that Loop 2 can understand/determine synthesis science, microstructure evolution, and grain boundary transformations. Lastly, in Loop 3, the large set of composition-sintering-microstructure-conductivity data, obtained over the course of the project, will be used in an established autonomous experimental framework *Closed-loop Autonomous system for Materials Exploration and Optimization (CAMEO)* established at NIST and UMD in prior MGI work to guide the next experiments parameters (compositions and sintering conditions) and will be iterated as a closed-loop to discover Na-ion conductors with optimal ion conductivity.

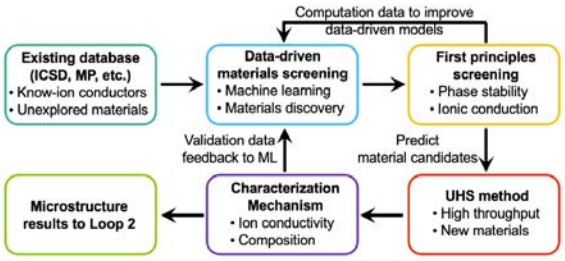


Figure 2. Scheme of Loop 1. Materials composition discovery by data-driven and first-principles

Broader Impacts and Workforce Development

The Na-ion conducting materials newly discovered in this work can be used as key components for Na-ion and Na-metal batteries, which are economic, environmental-friendly, and sustainable alternatives to Li-ion battery technology, thus promoting renewable energy use for transportation and grid-energy storage. In addition, all PIs will utilize their interdisciplinary fields to create unique educational opportunities for a diverse group of graduate, undergraduate, and K-12 students, with particular emphases on under-represented minorities. Education (*e.g.* hands-on training of summer high school students) and outreach activities (*e.g.* student oriented conference presentations and local activities such as an educational booth at "Maryland Day" to educate parents and K-12 students from the community) will be developed and employed in conjunction with the proposed research activities. In the spirit of MGI, our education and outreach efforts at all collaborating institutions will emphasize the unique components of data-driven closed-loop materials design as essential training for the next-generation MGI workforce.

Data Management and Open Access

Our current backbone database has been constructed using the open-source code *pymatgen-db (https://github.com/materialsproject/pymatgen)* based on *MongoDB*, which is developed under the MGI effort of the *Materials Project*. We are further expanding the capability of this database to handle the large amount of diverse data from high-throughput UHS experiments, conductivity measurements, characterization, microstructure modeling, and TTT diagrams from the many materials to be generated in this project. To facilitate the sharing of data within the team and to the public, a web portal will be established to allow direct access to the database. To meet the FAIR standard of data usage, we will add query features to the database for an effective search of material structures, compositions and properties. In addition to the web portal, a python-based API will be developed for the fast, easy, and scalable dissemination of a large amount of data.

Advancing Along the Materials Development Continuum and Partnerships to Translation

R&D of ion-conducting ceramic materials has largely been conducted in a slow, trial-and-error fashion impeded by three challenges: 1) long synthesis and processing times (~10–100 hours), leading to undesired side reactions and volatile elemental loss; 2) a limited number of available compounds that exhibit fast ion-conduction with mechanical and chemical stability; and 3) a lack of understanding and rational control of microstructures during conventional sintering. This project leverages the unique UHS method to rapidly sinter dense layers of oxide ceramics from precursors in ~10 seconds by simply sandwiching ceramic pellets between two Joule-heated carbon films. This simple yet effective proximity heating strategy, with a temperature up to 2000-3000K enables a 10^2–10^4-fold acceleration in ceramic synthesis and sintering time; and provides a route for rapid verification and feedback to first-principles computational predictions, enabling a truly computation-guided, closed-loop materials discovery process. The UHS method is the fundamental technology behind the entrepreneurial start-up HighT-Tech, LLC, founded to make UHS a commercially relevant technology.

DMREF: Computationally Driven-Genetically Engineered Materials (CD-GEM)

Lead Investigator: Jin Kim Montclare, montclare@nyu.edu
Participating Institutions: NYU Tandon School of Engineering, NYU, NYU School of Medicine
Website: https://wp.nyu.edu/tandonschoolofengineering-cdgem/
Keywords: MGI, biosynthesis, computation, high-throughput microrheology, gelation

Project Scope

We have explored the mechanism and dynamics of our supramolecular assembling coiled-coil protein, dubbed Q, for its ability to form nanofibers, encapsulate a small molecule, and form a hydrogel. We have hypothesized that: 1) these supramolecular assemblies can be used as scaffolds for drug delivery and imaging and 2) quantifying the gelation will allow us to home in on the sequence-structure-assembly relationship to allow controlled design of fibers and hydrogels. Our initial goals are to use Q to create an imageable drug delivery agent and measure the tunability of Q gelation kinetics in high throughput.

Relevance to MGI

Self-assembling protein biomaterials have become increasingly popular to develop scaffolds and hydrogels for tissue engineering, therapeutics, and diagnostic agents. However, predicting design to self-assembly function is still a challenge. Where we have previously reported developing Q, we have not yet tested its ability to be functionalized or explored its tunability. Our recently published projects have shown that we are able to functionalize Q for drug delivery and imageability and characterize morphological differences in Q, as well as quantify the impact of ionic strength and pH on the gelation of Q by developing a high throughput microrheological assay.

Using these learnings, we have hypothesized the impact of design changes on the self-assembly of Q variants. For example, in prepared manuscripts, using Rosetta and electrostatics computation, we have been able to directly correlate sequence to fiber assembly and rationally design hydrogels with faster sol-gel transitions, an example of our iterative feedback loop for optimized coiled-coil designs. The result of this work will be the first foray into quantifiable coiled-coil nanofiber and hydrogel design-to-function prediction.

Technical Progress

We have completed a series of important and fundamental studies to elucidate the mechanism and dynamics of Q hydrogel formation. We have studied the impact of the chemical environment to impart different gelation kinetics and material properties exhibiting stimuli responsiveness to not only temperature but to ionic strength and pH (**Figure 1a**). We have followed up these studies with a close examination of the gelation kinetics across a compositional space of pH values and ionic strengths by developing a high-throughput microrheological assay (**Figure 1b**). Here, we employed two independent methods of passive microrheology – MPT and DDM and added our own independent analysis by fitting relaxation exponents to sigmoidal curves to create a more facile analysis system (**Figure 1c**). These results have confirmed the electrostatic dependence of physical crosslinking for the Q protein and will assist in the optimization of variants for preferred gelation kinetics and UCST behavior.

We have also leveraged the functional properties of Q for its ability to form nanofibers by conjugation to ultrasmall superparamagnetic iron oxide (USPIO) nanoparticles. The resulting hybrid material is capable of doxorubicin encapsulation as well as

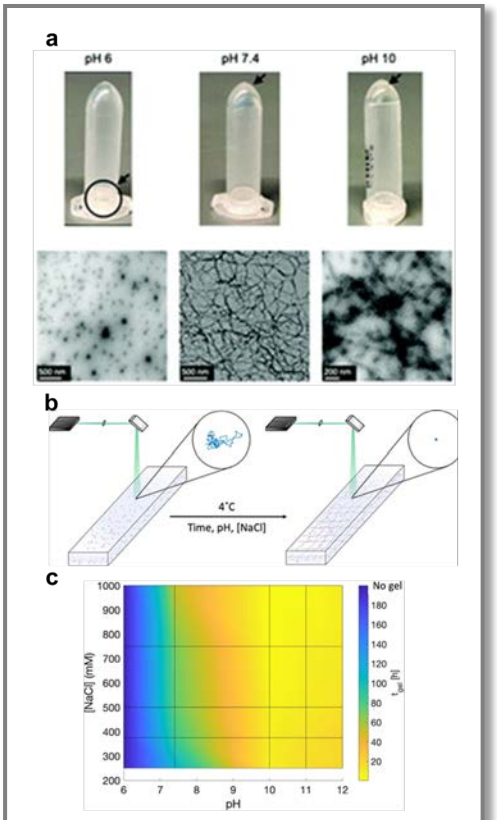


Figure 1 a. Tube inversion and transmission electron micrographs of Q at pH 6, pH 7.4, and pH 10. **b.** Scheme of high-throughput microrheology assay using a capillary tube **c.** Surface plot of gelation times determined by sigmoidal curve-fitting of MPT-determined mean-square displacements for the entire compositional space studied.

sensitive T_2*-weighted MRI darkening for strong imaging capability. This work illustrates the potential to use coiled-coil protein fibers in diagnostic and therapeutic applications.

In current projects, we have established the relationship between computational quantification of Q variant design electrostatics and stability and corresponding gelation kinetics and fiber size. To this extent we have created hydrogels with faster gelation properties and predicted and synthesized larger protein fibers with > 3x nanofiber diameters. We have also established ^{19}F MRI imageability of trifluoroleucine (TFL) incorporated Q variants.

Future Plans

Where we have been able to synthesize and predict fiber size from coiled-coil electrostatics, we now look to use computational design and electrostatics to predict Q variant hydrogel gelation kinetics. In this work, we utilize the aforementioned high-throughput microrheological assay to determine the sol-gel transition times for a variety of Q designs. To this extent, we can pinpoint precise differences and determine gelation from sequence and electrostatics with the assistance of machine learning packages. The ongoing project looks to fill out the compositional space of these parameters to home in on a reliable sequence-gelation relationship.

In our efforts to create imageable drug delivery materials from Q variants, current work involves computationally designed fluorine functionalized hydrogels for ^{19}F MRI imageability. Using the established designs of Q variants, we have used Monte-Carlo simulations in Rosetta to stabilize the hydrophobic core of Q and optimize for incorporation of TFL to create a hydrogel capable of ^{19}F MRI visualization.

Broader Impacts and Workforce Development

We have used these projects and its resources to focus on technology-centered outreach to K-12. We have recruited upper class undergraduate students in Biomolecular Engineering and Computer Science as paid fellows to develop a sequence of modules that exposes the students in grade 6-12 to technology enable modules on a variety of science topics. This summer, we will be integrating the research of ongoing projects with the ARISE program where we will use the research as a vehicle for our PhD students to teach and introduce research experience to New York City students lacking access to high quality STEM education experiences, especially to those from demographic groups underrepresented in STEM disciplines and careers.

Data Management and Open Access

This proposal is a collaborative effort between the groups of **Montclare**, **Bonneau** and **Wadghiri** across NYU schools of engineering, science, and medicine as well as Dr. **Renfrew** from Flatiron Institute and Dr. **Gupta** from AFRL. Together, the group is comprised of a multidisciplinary group of researchers including those from Materials Chemistry, Chemical & Biological Engineering, Chemistry, Computational Biology, Biomedical Engineering, Medicine, Biotechnology and Bioinformatics. This offers a unique collaborative environment and excellent training environment for students and postdocs. All the preliminary data presented here is from this multidisciplinary collaboration. Students and postdocs in the groups work fluidly together. We make all publications available and the links are on our website (https://wp.nyu.edu/tandonschoolofengineering-cdgem/).

Advancing Along the Materials Development Continuum and Partnerships to Translation

As a direct result of fundamental understanding of the Q variant and chemical environment design space, we are able to mitigate misdesign and have created many new variants capable of fiber assembly and or gelation. Variants possessing different assembly abilities match different applications depending on their needs. For example, faster gelling variants are critical for *in situ* gels and slower gelling variants may be ideal for tissue engineering. We have previously attained patents for work related to the Q design and look to do so with current and future designed variants that are capable of imageability and drug delivery which will have commercialization potential.

Publications and References

1. M. Meleties, P. Katyal, B. Lin, D. Britton, J. K. Montclare, *Self-assembly of Stimuli Responsive Coiled-coil Fibrous Hydrogels*, Soft Matter, **17**, 6470-6476 (2021)

2. M. Meleties, D. Britton, P. Katyal, B. Lin, R. L. Martineau, M. K. Gupta, J. K. Montclare, *High-Throughput Microrheology for Assessment of Protein Gelation Kinetics*, Macromolecules, **55**, 4, 1239-1247 (2022)

3. L. K. Hill, D. Britton, T. Jihad, K. Punia, X. Xie, E. Delgado-Fukushima, C. F. Liu, O. Mishkit, C. Liu, C. T. Hu, M. Meleties, P. D. Renfrew, R. Bonneau, Y. Z. Wadghiri, J. K. Montclare *Engineered Protein-Iron Oxide Hybrid Biomaterial for MRI-traceable Drug Delivery*, Mol. Syst. Des. Eng., Advanced Article (2022)

4. A. Qu, M. Nicolas, E. M. Leung, S. M. Jones, P. Katyal, K. Punia, M. Maxfield, J. K. Montclare *Exploring the Viability and Role of Virtual Laboratories in Chemistry Education Using Two Original Modules*, J. Chem. Educ. 2022, **99**, 4, 1596-1603

5. M. Meleties, R. L. Martineau, M. K. Gupta, J. K. Montclare, *Particle-Based Microrheology As a Tool for Characterizing Protein-Based Materials*, ACS Biomater. Sci. and Eng., Advanced Article (2022)

Materials Architected by Adaptive Processing

Lead Investigator: Thao (Vicky) Nguyen[1], vicky.nguyen@jhu.edu
Gretar Tryggvason[1], David Elbert[1], Peter Olmsted[2], David Kazmer[3],
Kalman Migler[4], Jonathan Seppala[4], Hilmar Koerner[5]
Participating Institutions: Johns Hopkins University[1], Georgetown[2], University of Massachusetts Lowell[3],
NIST[4], AFRL[5]
Website: none.
Keywords: polymer processing, recycled polymer, immiscible blends, flow-induced crystallization.

Project Scope

Materials manufacturing is subject to uncertainty in raw materials and processes that can drastically alter the resulting properties. To address these challenges, this project will develop a data-centric approach for integrated materials design and manufacturing, called Materials Architected by Adaptive Processing (MAAP). For example, we will create MAAP to design and produce polymer blends with superior properties from recycled polyethylene and polypropylene. The superior properties will be realized by designing novel shape-multiplying elements that manipulate the melt streams to create domain architectures and crystalline morphologies optimized for strength and toughness. We will develop an event-driven, microservices data layer, MAAP-DESc, to automate the contextualized data flow between the processing, characterization, and modeling tasks of the project. An instrumented coextrusion process with modular shape-multiplying elements will be designed from modeling studies to investigate the controllability of the architected blends. Multiscale modeling will be used to study the flow-induced crystallization of the multi-component melt, the stability of the melt streams, and the formation of phase domains during processing to determine how the measurable processing parameters control the crystalline morphologies and domain architecture.

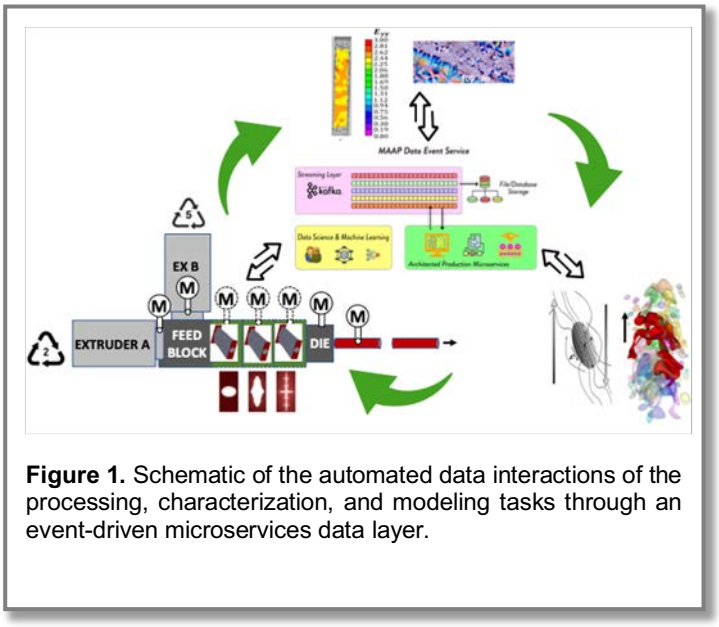


Figure 1. Schematic of the automated data interactions of the processing, characterization, and modeling tasks through an event-driven microservices data layer.

Relevance to MGI

An integrated methodology for concurrent materials design and manufacturing will create an innovative data architecture that maps research tasks onto microservices loosely coupled through a streaming data backbone. This approach maximizes human productivity and collaboration and integrates artificial intelligence and machine learning advances while reducing uncertainty from design to production. The MAAP framework will enable a new paradigm for the design of in situ process monitoring and control for the next generation of materials processing methods. In MAAP, the same data and data infrastructure applied for materials design will be harnessed for process monitoring and control to ensure consistent production of materials with targeted properties. The integrative MAAP approach will make the materials-by-design process more robust and provide adaptive processing systems for consistent materials production.

Technical Progress

The first 6 months of the project was devoted to developing equipment, experimental protocols, and modeling methods. The Kazmer group commissioned and characterized a cast film system comprising five extruders, producing nine layers using layer multiplying elements. We performed initial studies of virgin and

recycled polypropylene (PP) and found that recycled PP with higher viscosity drives instabilities & defective film. However, layer multiplying elements significantly reduced instabilities. The Nguyen lab performed initial uniaxial tension tests of the virgin, recycled, and alternating virgin and recycled PP multilayered sheets. As expected, the inclusion of recycled PP in the composite sheet degraded the strength and elongation to failure of the PP sheets. The Tryggvason group is developing fluid dynamics simulations of the formation and evolution of polymer blends of immiscible polymers.

Future Plans

For the next 12 months, we will continue to develop the film cast system and study layered virgin and recycled polyethylene (PE) systems and PP-PE layered composite systems. Through collaborations with NIST and AFRL, we will develop polarized light microscopy and X-ray scattering methods to characterize the polymer crystalline morphologies. We expect that multiphase fluid dynamics simulations will be able to study flow instabilities observed in experiments and the resulting morphology. We have started and will continue to develop the MAAP-DESc interface to automatically pull data from in-line temperature and pressure measurements of the film cast system.

Broader Impacts and Workforce Development

This project also advances the fundamental understanding of mixing and flow-induced crystallization in multi-polymer melts and applies that understanding to produce architected blends from recycled polymers. The ability to upcycle plastic waste will improve sustainability and reduce the environmental impact of plastics production. This project will provide access for students at all levels to training in the three key competency areas of data, computation, and experimentation that are called out in the TMS MGI Workforce Study. Graduate students and postdocs will be exposed through collaboration to research in a national lab environment (AFRL and NIST). Outreach will include meaningful research internships for high-school students and training seminars for practicing engineers, technicians, and managers in recycling, melt blending, coextrusion, and quality control. We will contribute to building a pipeline for STEM by recruiting high-school interns from all-girls and minority-majority schools.

Data Management and Open Access

Our project plan includes use of a single, unified semantic description of data. This allows use of a single database of metadata to describe and provide access to the entire project. Structured data will be hosted in SQL databases, while unstructured data (e.g., image files) will be curated in cloud-based storage. During the lifetime of the project, data will be maintained and hosted at Johns Hopkins University (JHU) by the Institute for Data Intensive Engineering and Science (IDIES). In addition, data will also be stored and archived at their institute of origin. MAAP-DESc dramatically simplifies the management of our digital data products. In MAAP-DESc, files will be processed by an asynchronous data broker triggering metadata harvesting, the population of our sample description database, creation of DOIs, and movement of files to SciServer's petabyte-capable cloud storage. The automatic population of our metadata database makes data findable, accessible, and reusable to the project immediately.

Advancing Along the Materials Development Continuum and Partnerships to Translation

A core strength of the project is the integration between designing the material and process control for the constituent production of the material. This will facilitate the translation.

Publications and References

None

Conductive Protein Nanowires as Next Generation Polymer Nanocomposite Fillers

Lead Investigator: Stephen S. Nonnenmann, Department of Mechanical & Industrial Engineering, University of Massachusetts-Amherst, ssn@umass.edu.; Todd Emrick, Department of Polymer Science & Engineering, University of Massachusetts-Amherst, tsemrick@umass.edu.; Derek R. Lovley, Department of Microbiology, University of Massachusetts-Amherst, dlovley@umass.edu.; Jessica D. Schiffman, Department of Chemical Engineering, University of Massachusetts-Amherst, schiffman@umass.edu.; Arthi Jayaraman, Department of Chemical & Biomolecular Engineering, University of Delaware, arthij@udel.edu.
Website: https://github.com/arthijayaraman-lab/self-supervised_learning_microscopy_images
Keywords: nanocomposites, protein nanowires, nanoelectronics, molecular modeling, machine learning

Project Scope

Composite flexible electronics typically feature soft, polymer matrices embedded with hard, brittle conductors that fail due mechanical mismatch. This DMREF project focuses on designing and establishing conductive protein nanowires (CPNs) as a revolutionary design paradigm as novel fillers for flexible electronics. Our DMREF plan integrates molecular modeling, simulations, machine learning (ML), synthetic biology, polymer chemistry, materials synthesis and engineering, and nanoelectronics in a synergistic manner to produce optimal CPN nanowires that exploit their conductive inner cores with surface-exposed amino acids that promote tunable interactions with polymer matrices for enhanced transport and processability.

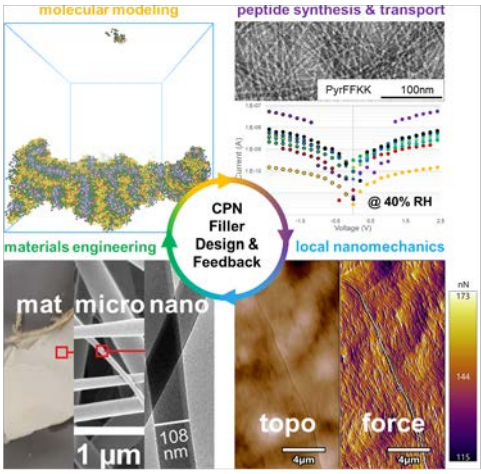


Figure 1. Convergence of molecular simulations of self-assembly peptides into CPNs, synthesis and fabrication of CPN based composites, and characterization of structure (aided by machine learning), electronic transport, and mechanical properties.

Relevance to MGI

Flexible electronics typically comprise soft, polymer matrices embedded with hard, brittle conductors and often fail due mechanical mismatch. This project introduces conductive protein nanowires (CPNs) as a revolutionary design paradigm for novel fillers in elastomer matrices. CPNs exhibit metallic-like conductivity yet possess lower elastic moduli compared to metallic nanowires. Contrasting common fillers like carbon nanotubes, CPNs disperse in water and agglomerate controllably in solvents. Our project integrates molecular modeling, simulation, & ML (Jayaraman), synthetic biology (Lovley), polymer chemistry (Emrick), materials engineering (Schiffman), and nanoelectronics (Nonnenmann) in a synergistic manner (Fig. 1). The interplay of amino acid sequences that form the CPNs and matrix polymers creates a complex system for which predictive modeling is required to identify key design choices and interaction parameters to achieve dispersed CPN-fillers in the matrix. These predictions inform peptide synthesis, structural characterization, and electronic transport studies to accelerate the production of CPN-enabled nanocomposites and devices.

Technical Progress

Robust CPNs & CPN superstructures. Our team translated anaerobic microbial CPN synthesis to a unique *E. coli* strain under aerobic conditions, greatly expanding potential of large-scale fabrication and production while preserving high conductivity.[1,2] We also demonstrated controlled agglomeration of CPNs into superstructured bundles with high aspect ratios (>10^5) using common organic solvents.[3] ***Pili-inspired conductive oligopeptides***. Our pili-based nanocomposites and bundles inspired the development of aromatic-containing, end-capped oligopeptide nanowires. Hydrophobicity, p-p interactions, and conductivity served as design principles to guide solid-phase peptide synthesis across a broad library (> 20) of oligopeptides. Oligopeptide hydrophobicity/size were determined to dictate the extent of solution assembly. Oligopeptide nanowire assembly was observed in aqueous solutions with 1 wt.% peptides, with structures tens of oligopeptide units thick and thousands of units

long. Current increased by two orders of magnitude with increasing size and hydrophobicity of the end-capping group. The current response of nanowire films displayed a strong sensitivity to relative humidity (RH), varying by three orders of magnitude in the 40-70% RH range. ***Electrospun oligopeptide nanofibers.*** Electrospinning represents the most cost-effective, controllable, scalable, and facile means to produce nanofiber mats boasting high mechanical strength, flexibility, and porosity.[4] Entanglement is generally required to obtain smooth, uniform fibers, posing challenges for small oligopeptide molecules. We overcame such challenges and successfully electrospun oligopeptides from a polar fluoroalcohol solvent. ***Nanomechanical mapping***. High-rate force mapping (~ 1 kHz) was performed of aromatic, end-capped peptide nanowire assemblies and observed preferential alignment and localized moduli variations. on microelectrode surfaces. ***Computational materials design:*** Using coarse-grained (CG) molecular dynamics (MD) simulations we developed rules linking chemical and physical heterogeneity in nanowires and nanorods to dispersion & percolation in polymer nanocomposites. ***Open-source ML workflow development for TEM image analysis****:* We have developed a new open-source self-supervised transfer learning based ML approach using small sets of TEM images for training; this method accurately analyzes TEM images for nanowire morphology classification, and segmentation as effectively as methods trained on significantly larger datasets.

Future Plans

New peptide sequences and capping group designs will yield assemblies with stronger aromatic interactions to further enhance electronic output. Electrical transport and tensile properties from the full peptide nanofiber mats to single nanofiber scales will be investigated using microelectrodes, atomic force microscopy, and rheological load cells. Polymer matrices such as polyvinyl alcohol (PVA) will be used to increase the cohesive strength of the peptide nanofibers. Nanomechanics measurements will adopt Sneddon solutions to further calibrate and quantify modulus imaging. Ongoing CGMD simulations of peptide assembly linking solvent selection, end-capping size, and charge density influence on the assembled nanowires will be verified using synthesis and characterization.

Broader Impacts and Workforce Development

Our project cross-trains graduate researchers in distinct disciplines – computational materials, polymer-nanoparticle synthesis, nanocharacterization, and microbiology. The students lead our collaborative meetings to solicit inter-disciplinary feedback and drive discussions of progress. The open source developed ML workflow for TEM image analysis trained using few labeled images and small data sets will be useful broadly in the soft materials community. Nanomaterials and entrepreneurship demos involving materials synthesis, hands-on characterization, and team building are planned for post-pandemic high school engineering summer camp.

Data Management and Open Access

We adopted the use of online repository figshare to supplement publications with raw data. Our approach aims to create optimal data descriptors for files output from polymer/peptide simulations, characterization during chemical synthesis, materials processing, biosynthesis, and electrical/rheological characterization. We are generating two groups of descriptors: 1) elemental and structural information and 2) physical properties (i.e., conductivity, elasticity, percolation thresholds) of our CPN library.

Advancing Along the Materials Development Continuum and Partnerships to Translation

While CPNs represent soft, conductive fillers as next-generation composite fillers, issues such as scalable production and geometries typically required of devices hinder broader adoption. Our efforts in aerobic synthesis, superstructured bundling, and manufacturing of fiber mats are advancing CPNs towards commercial pathways and efforts to align with industrial partners will be pursued in the second half of the award period.

Publications and References

● T. Ueki, D.J. Walker, T.L. Woodard, K.P. Nevin, S.S. Nonnenmann, D.R. Lovley, *An Escherichia coli Chassis for Production of Electrically Conductive Protein Nanowires*, ACS Synthetic Biology **9**, 647 (2020). ● X. Liu, D.J. Walker, S.S. Nonnenmann, D. Sun, D.R. Lovley, *Direct Observation of Electrically Conductive Pili Emanating from Geobacter Sulfurreducens*, mBio **12** e02209 (2021). ● Y.-L. Sun, B.J. Montz, R. Selhorst, H.-Y. Tang, J. Zhu, K.P. Nevin, T.L. Woodard, A.E. Ribbe, T.P. Russell, S.S. Nonnenmann, D.R. Lovley, T. Emrick, *Solvent-Induced Assembly of Microbial Protein Nanowires into Superstructured Bundles*, Biomacromolecules **22**, 1305 (2021). ● P. Rathore, J.D. Schiffman, *Beyond the Single-Nozzle: Coaxial Electrospinning Enables Innovative Nanofiber*

Chemistries, Geometries, and Applications, ACS Applied Materials & Interfaces **13**, 48 (2021). ● S. Lu, A. Jayaraman*, Effect of nanorod physical roughness on the aggregation and percolation of nanorods in polymer nanocomposites*, ACS Macro Letters **10**, 1416–1422 (2021) ● S. Lu, Z. Wu, A. Jayaraman, *Molecular Modeling and Simulation of Polymer Nanocomposites with Nanorod Fillers* J. Phys Chem B. **125**, 2435-2449 (2021)

Paired ionic-electronic conductivity in self assembling conjugated rod-ionic coil segmented copolymers and mesogens with ionic liquid units

Lead Investigator: Christopher Kemper Ober, cko3@cornell.edu
Participating Institutions: Cornell U, U of Chicago, U of Washington Seattle
Website: http://dmref.coecis.cornell.edu/index.html
Keywords: Computational modeling; mixed ionic/electronic conductor; charge carrier transport; self-assembly.

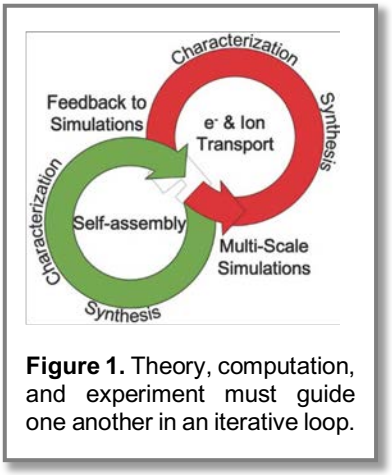


Figure 1. Theory, computation, and experiment must guide one another in an iterative loop.

Project Scope

The scope of our project is to develop mixed ion/electron conducting mixed ionic electronic conductors (MIECs) that may provide new materials for energy conversion and storage and be key components in organic electrochemical transistors. Our MIECs consist of two classes of materials: 1) conjugated liquid crystals (oligomers with thiophene and ethylene oxide units) and 2) conjugated polymers made of a high molecular weight polythiophene backbone and short PEG side-chains. Our objective is to fundamentally understand the structure-property relationships for ion and electron transport in MIECs through an iterative cycle (Fig.1) of modeling, synthesis, and characterization.

Relevance to MGI

Iterative steps involving predictive modeling and state-of-the-art synthesis, processing and characterization enhance our understanding of the structure-property-function relationships present in MIECs and facilitates new materials discovery. Multiscale simulations are used to explain and predict the self-assembly behavior of chemical species consisting of ion-conducting and electron-conducting components. Moieties with favorable structures are then synthesized and processed into a thin film morphology by the experimental teams. Both transport characterization and simulation are then integrated into a machine learning (ML) approach to accelerate materials discovery.

Technical Progress

I. *Designing new MIEC liquid crystals (LCs).* We explored the complex relationship between the self-assembled mesophase and mixed ionic/electronic conductivity in LCs. For example, we have designed a synthetic pathway to produce new MIEC LCs containing 2,5-bis(thien-2-yl)thieno[3,2-b]thiophene (BTTT) as mesogen attached to tetra(ethylene oxide), BTTT/PEO4. Second, we examined its thermal behavior, and identified a mesophase with a lamellar structure. We adopted both grazing incidence wide-angle X-ray scattering (GIWAXS) and molecular dynamics (MD) simulations to reveal solid state herringbone packing in the self-assembled structures. Finally, we measured the ionic and electronic conductivities of BTTT/PEO4 and compared them with those of prior MIECs.

II. *Designing new polythiophene-based MIECs.* Our previous studies explored alternative side chain chemistries on the polythiophene backbone. The side chain was constrained to have 3 "units", either an alkyl (propyl) "A" group or an EO (or PEO) group "P". Seven new side groups were simulated, namely: -AAP, -PAP, -APA, -PAA, -APP, -PPA and -PPP. Molecular dynamics simulations were used to understand how changing the oxygen position and density would change ionic conductivity in both crystalline and amorphous systems. A selection of these polymers was synthesized and characterized [1] and observed influence of LiTFSI on structure and conductivity followed the MD simulations. Increased oxygen atom distance resulted in improved crystallinity and conductivity.

Future Plans

We will continue to examine both low molar mass LC (model materials) and polymeric MIECs using the iterative process we have now established. We have previously shown higher polymer film crystallinity and improved mixed-conduction results for P3APPT compared to P3MEEMT [1,2]. We will now expand our molecular toolset and include charged "ionic liquid" functional groups, to provide better undoped self-assembly due to stronger electrostatic forces. Our MD simulations and ML studies also suggest faster Li ion conduction happens with nitrile and sulfide functional groups compared to PEG segments. And we thus expect improved mixed-

conduction performance of our new polymer materials.

Broader Impacts and Workforce Development

As an example of educational outreach, the Nealey and Patel groups led participation by DMREF team members from all schools in a six-week Molecular Engineering course to teach the principles in the design of soft materials. For summer 2021, the Chicago Collegiate Scholars Program (CSP) with 18 selected students received a module on *Electric Circuits* from our DMREF team. Additionally, Ober and a group member were the speakers of the 2022 Turner J. Alfrey Professorship Lecture Series held by Michigan State University. Ober and his student described recent DMREF work on MIECs, and potential applications to students, faculty, industrial scientists and the public.

Data Management and Open Access

Our project website (http://dmref.coecis.cornell.edu/index.html) is intended to collect and make available our DMREF information and data. The site is continuously updated to reflect the activities of the group. We provide links to different resources including: (i) A DMREF-specific public Github repository which contains many of our simulation codes and scripts, (ii) a public Github repository of codes from the Escobedo group members, some of which are of relevance to the simulation of polymers and oligomers in general, (iii) a link to a repository called "Cornell Box" which contains the force field parameters of the candidate chemical species, (iv) links to a "Cornell Box" repository which contains software being used for machine learning applications in the project and other supplementary information, and (v) links to repositories used by various group members for research data storage.

Advancing Along the Materials Development Continuum and Partnerships to Translation

To date no patents have been filed but the work has shown the suitability of these materials in organic electrochemical transistors (OECT). David Ginger (U Washington) has been involved in device design and we expect to send materials to George Malliaras (Cambridge, UK) to better understand their potential for bioelectronics.

Publications

1. T. Ma, B. X. Dong, J. W. Onorato, J. Niklas, O. Poluektov, C. K. Luscombe, S. N. Patel, Correlating conductivity and Seebeck coefficient to doping within crystalline and amorphous domains in poly(3-(methoxyethoxyethoxy)thiophene), Journal of Polymer Science **59** 2797-2808 (2021).
2. B. X. Dong, Z. Liu, J. W. Onorato, T. Ma, J. Strzalka, P. Bennington, C. K. Luscombe, C. K. Ober, P. F. Nealey, S. N. Patel, *Ionic Dopant-Induced Ordering Enhances the Thermoelectric Properties of a Polythiophene-Based Block Copolymer*. Advanced Functional Materials, **31**, 2106991 (2021).
3. J. W. Onorato, Z. Wang, Y. Sun, C. Nowak, L. Q. Flagg, R. Li, B. X. Dong, L. J. Richter, F. A. Escobedo, P. F. Nealey, S. N. Patel, C. K. Luscombe, *Side chain engineering control of mixed conduction in oligoethylene glycol-substituted polythiophenes*, Journal of Materials Chemistry A **9**, 21410-21423 (2021).
4. R. Gupta, M. Misra, W. Zhang, A. Mukhtyar, S. P. Gido, A. Ribbe, F. A. Escobedo, E. B. Coughlin, *Topological Frustration as a New Parameter to Tune Morphology Revealed through Exploring the Continuum between A-B-C 3-Arm Star and Linear Triblock Polymers,* Macromolecules **54**, 4401-4411 **(**2021).
5. B. X. Dong, C. Nowak, J. W. Onorato, T. Ma, J. Niklas, O. G. Poluektov, G. Grocke, M. F. DiTusa, F. A. Escobedo, C. K. Luscombe, P. F. Nealey, S. N. Patel, *Complex Relationship between Side-Chain Polarity, Conductivity, and Thermal Stability in Molecularly Doped Conjugated Polymers*, Chemistry of Materials, **33**, 741-753 (2021).
6. P. J. W. Sommerville, Y. Li, B. X. Dong, Y. Zhang, J. W. Onorato, W. K. Tatum, A. H. Balzer, N. Stingelin, S. N. Patel, P. F. Nealey, C. K. Luscombe, *Elucidating the Influence of Side-Chain Circular Distribution on the Crack Onset Strain and Hole Mobility of Near-Amorphous Indacenodithiophene Copolymers*, Macromolecules **53**, 7511-7518 (2020).
7. T. Ma, B. X. Dong, G. L. Grocke, J. Strzalka, S. N. Patel, *Leveraging Sequential Doping of Semiconducting Polymers to Enable Functionally Graded Materials for Organic Thermoelectrics*, Macromolecules **53**, 2882–2892 (2020).
8. C. G. Bischak, L. Q. Flagg, K. Yan, T. Rehman, D. W. Davies, R. J. Quezada, J. W. Onorato, C. K. Luscombe, Y. Diao, C. Li, D. S. Ginger, *A Reversible Structural Phase Transition by Electrochemically-Driven Ion Injection into a Conjugated Polymer*, Journal of the American Chemical Society **142**, 7434–7442 (2020).
9. M. Misra, Z. Liu, B. X. Dong, S. N. Patel, P. F. Nealey, C. K. Ober, F. A. Escobedo, *Thermal Stability of π-Conjugated n-Ethylene-Glycol-Terminated Quaterthiophene Oligomers: A Computational and Experimental Study*, ACS Macro Letters **9**, 295–300 (2020).
10. L. Q. Flagg, C. G. Bischak, R. J. Quezada, J. W. Onorato, C. K. Luscombe, D. S. Ginger, *P-Type Electrochemical Doping Can Occur by Cation Expulsion in a High-Performing Polymer for Organic Electrochemical Transistors*. ACS Materials Letter **2**, 254–260 (2020).

References

1. J. W. Onorato, Z. Wang, Y. Sun, C. Nowak, L. Q. Flagg, R. Li, B. X. Dong, L. J. Richter, F. A. Escobedo, P. F. Nealey, S. N. Patel, C. K. Luscombe, *Side chain engineering control of mixed conduction in oligoethylene glycol-substituted polythiophenes*, Journal of Materials Chemistry A **9**, 21410-21423 (2021)
2. B. X. Dong, C. Nowak, J. W. Onorato, J. Strzalka, F. A. Escobedo, C. K. Luscombe, P. F. Nealey, S. N. Patel, *Influence of Side-Chain Chemistry on Structure, and Ionic Conduction Characteristics of Polythiophene Derivatives: A Computational*

and Experimental Study, Chemistry of Materials **31**, 1418-1429 (2019)

Ultra-Strong Composites by Computational Design

Lead Investigator: Gregory M. Odegard, gmodegar@mtu.edu
Participating Institutions: Michigan Technological University, Florida State University, University of Utah, Massachusetts Institute of Technology, Virginia Commonwealth University, Georgia Tech, Johns Hopkins University, Pennsylvania State University, University of Colorado at Boulder, University of Minnesota, Florida A&M University, Nanocomp Technologies, Solvay, NASA, Air Force Research Laboratory
Website: http://www.us-comp.com/
Keywords: Carbon nanotubes, Composites, Mechanical testing, Topology optimization

Project Scope

The goal of this project is to use computational modeling to drive the development of the next generation of composite materials for crewed missions to deep space. The composite material needs to have 3× specific strength and stiffness relative to state-of-the-art carbon fiber composites and be fabricated and tested as a panel-level proof-of-concept (TRL-3). The project is focused on carbon nanotube (CNT) reinforced polymer matrix composites. Computational modeling tools include density functional theory (DFT), molecular dynamics (MD), mesoscale modeling, finite element analysis (FEA), peridynamics, and topology optimization. Novel mechanical testing techniques are being developed to test these ultra-strong composites.

Relevance to MGI

This project involves 11 universities, 2 companies, NASA, and AFRL; and thus is the largest MGI effort focused on achieving the common quantitative goals for the next generation composites in the U.S. New computational and experimental tools are being developed that are specifically targeted to develop these ultra-strong composites. These new tools are being utilized to generate new data that is used to train machine learning (ML) models that are focused on the nanometer and microscale length scales. In turn, the ML models provide information on the effect of nano- and micro-scale features on composite mechanical properties. This knowledge supports the material manufacturer (Nanocomp) in understanding the effect of processing parameters on nano/micro-scale feature development, and thus on CNT material performance. Ultimately, this knowledge leads to improvements in CNT manufacturing and thus CNT-based composite material strength and stiffness. Figure 1 shows a diagram that demonstrates the relationships between modeling, manufacturing, and testing.

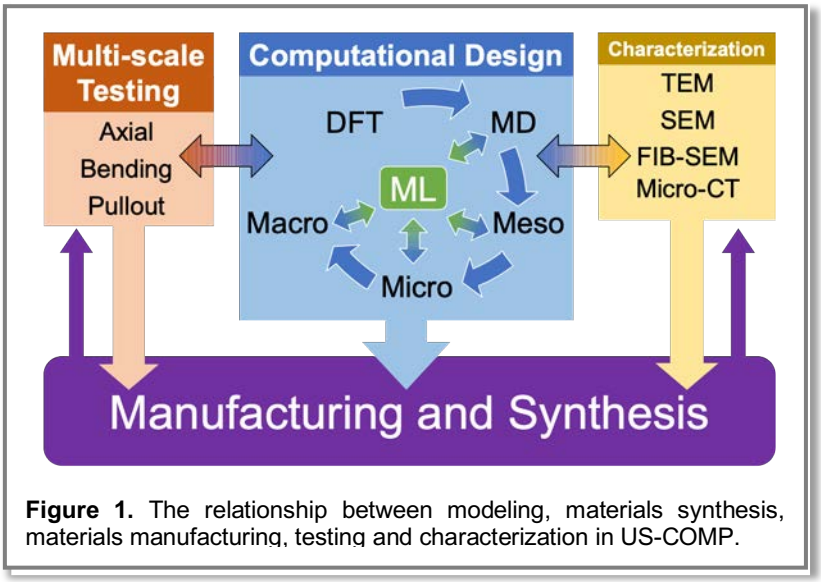


Figure 1. The relationship between modeling, materials synthesis, materials manufacturing, testing and characterization in US-COMP.

The US-COMP working group is sub-divided into a series of collaborative teams. Each team includes experts in simulation and experiments, and is diverse in terms of student/faculty/industry/government lab participation. In this manner, each team has a specific goal that relates to the common quantitative objectives, and has input from researchers with different perspectives, research experience levels, and areas of expertise. The original US-COMP teams were built on discipline-specific focus areas (modeling, materials synthesis, materials manufacturing, testing and characterization), but this structure was not effective in achieving the project goals, so the collaborative team structure was adopted. Since the switch to the new team structure, we have made substantial progress to our goals.

Technical Progress

Since the start of US-COMP in 2017, substantial progress has been made toward achieving the institute goals. Our modeling and characterization efforts have established the overall structure of our materials, as well as the critical features that control the material performance [1-4]. Our machine learning models have revealed the role

that critical features (e.g. CNT defects, CNT size distribution, CNT/polymer interfacial functionalization) play in the strength and failure characteristics of the CNT composites [5]. Also, novel force fields have been developed that enable the accurate prediction of the physical, mechanical, and thermal properties of CNT composites [6, 7].

Future Plans

In the next year, US-COMP plans to continue efforts on CNT composite laminate fabrication and testing, machine learning for improved CNT composite design, CNT surface treatment techniques, and topology optimization for CNT/carbon fiber hybrid composite preforms. All of these efforts will continue to rely on structure-property relationships established with computational modeling.

Broader Impacts and Workforce Development

To date, US-COMP has supported 56 graduate students, 43 undergraduate students, and 19 postdocs. Most of these students have participated in one or more of our weekly meetings, annual poster sessions, annual review meetings, and/or annual working meetings. The graduated PhD students have taken jobs in academia, industry, and federal research labs.

Data Management and Open Access

To date, US-COMP has supported the publication of 73 journal/conference papers (only a small fraction are included in the references herein), 5 licenses, 2 force field parameter sets, and numerous open-source software codes. It is important to note that much of our research is protected by non-disclosure agreements and export control restrictions.

Advancing Along the Materials Development Continuum and Partnerships to Translation

US-COMP researchers are working closely with our two industry members, Solvay and Nanocomp. These two companies are invited to participate in our weekly meetings, and their input is continuously sought. We have non-disclosure agreements with both companies, which allows us to open communicate important information regarding the manufacturing and use of polymer resins and raw CNT materials. Without this interaction, we would not have made the substantial progress that we have achieved thus far.

Publications and References

1. Xu, H., G. Drozdov, J.G. Park, B.D. Jensen, K.E. Wise, Z. Liang, G.M. Odegard, E.J. Siochi, and T. Dumitrică, *Computationally Guided Design of Large-Diameter Carbon Nanotube Bundles for High-Strength Materials.* ACS Applied Nano Materials, 2021. **4**(10): p. 11115-11125.
2. Patil, S.U., M.S. Radue, W.A. Pisani, P. Deshpande, H. Xu, H. Al Mahmud, T. Dumitrică, and G.M. Odegard, *Interfacial characteristics between flattened CNT stacks and polyimides: A molecular dynamics study.* Computational Materials Science, 2020. **185**: p. 109970.
3. Pisani, W.A., M.S. Radue, S.U. Patil, and G.M. Odegard, *Interfacial modeling of flattened CNT composites with cyanate ester and PEEK polymers.* Composites Part B: Engineering, 2021. **211**: p. 108672.
4. Deshpande, P.P., M.S. Radue, P. Gaikwad, S. Bamane, S.U. Patil, W.A. Pisani, and G.M. Odegard, *Prediction of the Interfacial Properties of High-Performance Polymers and Flattened CNT-Reinforced Composites Using Molecular Dynamics.* Langmuir, 2021. **37**(39): p. 11526-11534.
5. Rahman, A., P. Deshpande, M.S. Radue, G.M. Odegard, S. Gowtham, S. Ghosh, and A.D. Spear, *A machine learning framework for predicting the shear strength of carbon nanotube-polymer interfaces based on molecular dynamics simulation data.* Composites Science and Technology, 2021. **207**: p. 108627.
6. Damirchi, B., M. Radue, K. Kanhaiya, H. Heinz, G.M. Odegard, and A.C.T. van Duin, *ReaxFF Reactive Force Field Study of Polymerization of a Polymer Matrix in a Carbon Nanotube-Composite System.* The Journal of Physical Chemistry C, 2020. **124**(37): p. 20488-20497.
7. Odegard, G.M., S.U. Patil, P.P. Deshpande, K. Kanhaiya, J.J. Winetrout, H. Heinz, S.P. Shah, and M. Maiaru, *Molecular Dynamics Modeling of Epoxy Resins Using the Reactive Interface Force Field.* Macromolecules, 2021. **54**(21): p. 9815-9824.

Nonperturbative Studies of Functional Materials Under Nonequilibrium Conditions (NPNEQ)

Lead Investigator: T. Ogitsu, ogitsu1@llnl.gov.
Participating Institutions: Lawrence Livermore National Laboratory, Lawrence Berkeley National Laboratory, SLAC National Accelerator Laboratory, Stanford University
Website: https://sc-programs.llnl.gov/basic-energy-science-at-llnl/npneq
Keywords: RT-TDDFT, coupled spin-electron-ion dynamics, opensource software, high performance computing, functional materials

Project Scope

NPNEQ provides software tool for materials science research community to explore emerging novel materials properties that are relevant for MGI. We develop GPU ready opensource real-time time-dependent density-functional-theory (RT-TDDFT) software, which is validated through theory-experiments collaborative research taking advantage of the ultrafast experiment capability at SLAC. The software, named *INQ* consisting of highly modular programmable library set, enables us to directly investigate novel properties emerging from quantum mechanically coupled spin-electron-ion dynamics. The NPNEQ project offers tutorials of *INQ* software at LLNL Computational Chemistry and Materials Science (CCMS) Summer School Program.

Relevance to MGI

For opensource computer software tool to be effective in materials science research, following factors become crucial: readability, programmability, transferability, and excellent performance on HPC systems. For this reason, *INQ* software consists of highly modular library that have simple interface, well defined functionality. This helps scientists to customize the code for their own scientific research. RT-TDDFT is still immature methodology in need of further theory/algorithm development, which will require comprehensive validation based on computer simulations and experiments. Our software aims to reduce burden for such activities by offering programmable library that works on wide variety of computer architectures. Our project assists scientific community to explore emerging quantum material behaviors by providing user/programmer friendly software tool.

Technical Progress

In Q3 2021, *INQ* code became functional and available at https://gitlab.com/npneq/inq The code demonstrated excellent TDDFT simulation scalability over 1000+ GPU on Lassen supercomputer at LLNL. The *INQ* simulation results are compared with existing code such as Quantum-Espresso (QE) and Octopus for various physical quantities such at total energy of reference systems, dipole fluctuation and optical absorption profile and published in Ref [1]. See Figure 1. We published 3 papers led by NPNEQ project (Refs. [1-3]) and 7 papers published as collaborations with external funded projects (Refs. [4-10]).

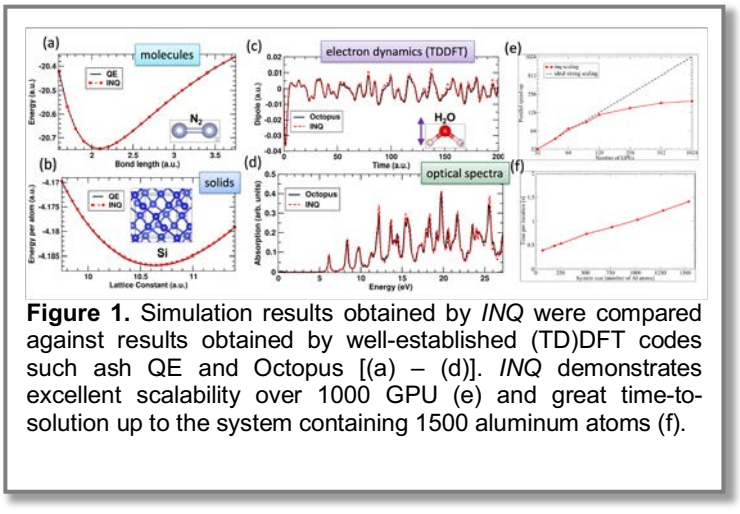


Figure 1. Simulation results obtained by *INQ* were compared against results obtained by well-established (TD)DFT codes such ash QE and Octopus [(a) – (d)]. *INQ* demonstrates excellent scalability over 1000 GPU (e) and great time-to-solution up to the system containing 1500 aluminum atoms (f).

Future Plans

We plan on making *INQ* software available on Perlmutter at NERSC (FY22), Frontier at Oak Ridge National Laboratory (FY23) and El Capitan at LLNL (FY24). These HPC systems have distinct novel hardware designs, which will let us re-evaluate our software design approach and improve the portability. We will also assist compilation and execution of *INQ* on CPU based HPC systems available at NSF and/or DOE supercomputer centers to facilitate broader use of our code. Several commonly used/relevant functionalities in materials science research, such as fast hybrid exchange-correlation functional, spin-orbit coupling that does not rely on orbital projection

(Breit-Pauli Hamiltonian), spiral boundary condition, will be implemented in the next few years. As the validation of new software implementation, we will perform theory-experiments collaborative research on spin dynamics in layered anti-ferromagnetic systems, demagnetization in optically excited nickelates, non-linear, non-equilibrium response in topological materials.

Broader Impacts and Workforce Development

We will continue to provide *INQ* tutorials leveraging summer/winter schools such as LLNL CCMS summer school. The students will learn how to develop software tools such as *INQ* and to perform research using the code. (TD)DFT is becoming commodity in materials science research, therefore, our tutorial will become great platform for the next generation MGI workforces to lean about basic research skills. Conference and workshop presentations as well as publications in peer reviewed journals will also be used to share know-how regarding software development, evaluation and insights obtained from research using our software.

Data Management and Open Access

INQ software, its users/programmers guide, reference inputs/outputs for validation, publications and presentations related to our software as well as information about scientific research results using our software will be available at https://gitlab.com/npneq and https://sc-programs.llnl.gov/basic-energy-science-at-llnl/npneq

Advancing Along the Materials Development Continuum and Partnerships to Translation

NPNEQ provides opensource RT-TDDFT software *INQ* to materials science community for accelerating materials design and discovery. We are being engaged in discussions with many other software development efforts for materials science research such as NSF funded MolSSI (https://molssi.org), and the other BES funded software center projects such as Center for Predictive Simulation of Functional Materials (PI: Paul Kent) and Berkeley GW (PI: Steven Louis) to facilitate broader use of our software. Recently, we initiated discussion with IBM about new software sharing effort that will accelerate use of software for materials science research such as *INQ* in industry led research programs.

Publications and References

1. X. Andrade, C. D. Pemmaraju, A. Kartsev, L. Z. Tan, S. Rajpurohit, J. Xiao, A. Lindenberg, T. Ogitsu, and A. Correa, *INQ, a modern GPU-accelerated computational framework for (time-dependent) density functional theory*, J. Chem Theo. Comp. **17**, 7447 (2021).
2. S. Rajpurohit, C. D. Pemmaraju, T. Ogitsu, and L. Z. Tan, *Nonperturbative study of bulk photovoltaic effect enhanced by an optically induced phase transition*, Phys. Rev. B **10**, 094307 (2022).
3. T. J. Smart, H. Takenaka, T. A. Pham, L. Z. Tan, J. Z. Zhang, T. Ogitsu, and Y. Ping, *Enhancing Defect Tolerance with Ligands at the Surface of Lead Halide Perovskites*, J. Phys. Chem. Lett. **12**, 6299 (2021).
4. Jorge Kohanoff, Alfredo Correa, Gleb Gribakin, Conrad Johnston, Andres Saul, *Radioactive decay of ^{60}Sr in cement: a non-equilibrium first-principles investigation*, Euro. Phys. J. D **75**, 426 (2021).
5. K. M. Siddiqui1, D. B. Durham, F. Cropp, C. Ophus, S. Rajpurohit, Y. Zhu, J. D. Carlstrom, C. Stavrakas, Z. Mao, A. Raja, P. Musumeci, L. Z. Tan, A. M. Minor, D. Filippetto, and R. A. Kaindl, *Ultrafast optical melting of trimer superstructure in layered 1T'-TaTe2*, Nat. Comm. **4**, 152 (2021).
6. M.-F. Lin, N. Singh, S. Liang, M. Mo, P. Nunes, K. Ledbetter, J. Yang, Michael Kozina, S. Weathersby, X. Shen, A. Cordones-Hahn, T. Wolf, K. Gaffney, C. D. Pemmaraju, M. Ihme, X. Wang, *Imaging the solvated hydronium cation and hydroxyl radical in the ionized liquid water*, Science **374**, 92 (2021).
7. S. Rajpurohit, L. Z. Tan, C. Jooss, and P. E. Blochl, , *Ultrafast spin-nematic and ferroelectric phase transitions induced by femtosecond light pulses*, Phys. Rev. B **102**, 174430 (2020).
8. E. E. Quashie, R. Ullah, X. Andrade, and A. A. Correa, *Effect of chemical disorder on the electronic stopping of solid solution alloys*, Acta Mater. **196**, 576 (2020).
9. J. Xiao, Y. Wang, H. Wang, C. D. Pemmaraju, S. Wang, P. Muscher, E. J. Sie, C. M. Nyby, T. P. Devereaux, X. Qian, X. Zhang and A. M. Lindenberg, *Berry curvature memory through electrically driven stacking transitions*, Nature Physics **16**, 1028 (2020).
10. L. McClintock, R. Xiao, Yasen Hou, C. Gibson, H. Clark T., D. Abramovitch, L. Z. Tan, R. T. Senger, Y. Fu, S. Jin, and D. Yu, *Temperature and Gate Dependence of Carrier Diffusion in Single Crystal Methylammonium Lead Iodide Perovskite Microstructures*, J. Phys. Chem. Lett. **11**, 1000 (2020).

DMREF: Collaborative: Accelerating the Adoption of Sintering-Assisted Additive Manufacturing Using Integrated Experiments, Theory, Simulation and Data Science

Lead Investigator: Eugene A. Olevsky, San Diego State University, eolevsky@sdsu.edu.
Lead Investigator: Rajendra K. Bordia, Clemson University, rbordia@clemson.edu
Lead Investigator: Lisa Rueschhoff, AFRL, lisa.rueschhoff.1@afresearchlab.com
Participating Institutions: San Diego State University, Clemson University, and AFRL

Keywords: Additive manufacturing, sintering, multi-scale simulations, machine learning.

Project Scope

The overarching goal of the proposed research is the development of a new type of experimentally guided and validated multi-scale direct and inverse sintering model taking into account specifics of micro- and macro structure

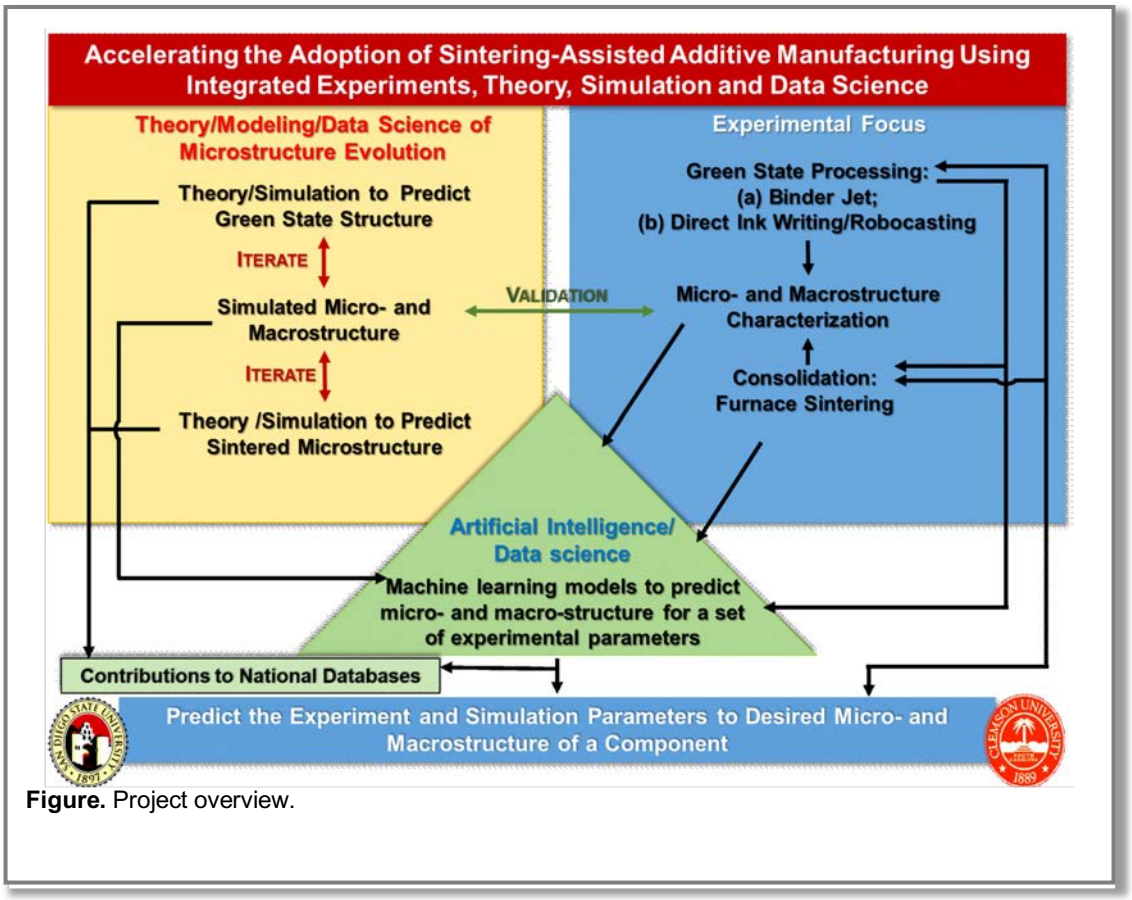


Figure. Project overview.

in the SAAM processes. The project overview is presented in the figure below.

The solution to this fundamental inverse sintering problem enables the determination of the optimal green state processing conditions; pre-sintering components' shape, and micro- and macro-structure; and sintering conditions required to obtain the desired shape and microstructure at the end of sintering. This is the ultimate objective of this integrated collaborative project.

Relevance to MGI

The proposed integration of theory and multiscale simulations with experiments in a data-driven predictive framework addresses the complex interplay between green-state processing conditions and sintered microstructure.

The project will provide fundamental, basic knowledge and a novel practical approach to design and optimize the manufacturing of advanced ceramic systems with programmable macroscopic characteristics and microstructure and hence properties and performance. Due to significant reduction in trial-and-error experiments, the solution of the "inverse SAAM problem" will accelerate the adoption of SAAM for advanced ceramics – an important goal of the Materials Genome Initiative.

Technical Progress

Different Machine Learning algorithms were compared for image-based porosity classification from a diverse and complex porosity image set. Within them Deep Convolutional Neural Network (DCNN) performed with the highest accuracy, shown simplicity of the model, and a variety of materials and defects can be easily included in the model by simply adding a new training dataset [1].

A Convolutional Neural Network (CNN) and a Faster Region-Based CNN (R-CNN) for image-based detection of porosities in printed parts by SAAM have been compared. It was found that the R-CNN outperformed the standard CNN with region proposal algorithms in both F1 Score (57%-73% for CNN and 84% for R-CNN) and image processing time (15s for CNN and > 1s R-CNN) [2].

Future Plans

- The development of a new constitutive model of sintering of AM ceramics based on the mutual calibration of the meso-scale sintering simulation code, experiments, and the machine learning of the specifics of the green and sintered microstructure of the SAAM produced parts.
- Development of full multi-scale, experimentally validated and machine learning guided, virtual experiments for sintering and microstructure development in SAAM processes which fully account for the specifics of the microscopic and macroscopic structures produced using SAAM.
- The solution of the "inverse" problems of sintering enabling the determination of the optimal AM conditions, and green components' shape and micro- and macro-structure required to obtain the desired shape and micro-structure of the sintered component.
- Validation implementation of developed simulation platforms for SAAM produced components of interest to AFRL, GE, Saint-Gobain, and Robocasting Enterprises.

Broader Impacts and Workforce Development

Five Ph.D. students are currently involved in the project, including two female, and one Hispanic. We plan to recruit additional students and postdoctoral research associates with diverse backgrounds to work on the project.

Data Management and Open Access

The primary analyzed data will be published in the form of peer-reviewed journal articles, theses, and other print or electronic publishing formats (e.g., pdf format). In addition, the associated metadata that describes the theoretical models and data analysis methods, experimental setups, fabrication procedures, will be made available in materials databases and repositories for re-use by the scientific community. We will publish our data in materials databases such as the Materials Data Facility (MDF) and the Open Science Facility (OSF). Accompanying workflow descriptions will be released as educational materials through the NSF sponsored Nanohub platform. To aid citation of software and data, the PIs will secure digital object identifiers (DOI) where possible.

Advancing Along the Materials Development Continuum and Partnerships to Translation

We have established strong partnership with two companies. Both of them have been very enthusiastic about this project. We are in the process of establishing IP agreements with them and once these are in place, we will request a GOALI supplement to apply the approaches and fundamental scientific discoveries to problems of direct interest to our industrial partners.

Publications and References

1. N. Satterlee, E. Torresani, E. Olevsky, J. Kang, *Comparison of machine learning methods for automatic classification of porosities in powder-based additive manufactured metal parts*, Int. J. Adv. Manuf. Technol., **120**, 6761–6776 (2022).
2. N. Satterlee, E. Torresani, E. Olevsky, J. Kang, *Using image-based R-CNN for automatic detection and characterization of porosities in sintering-assisted additive manufactured metal parts*, J. Manuf. Syst. (submitted)

The Synthesis Genome: Data Mining for Synthesis of New Materials

Lead Investigator: Elsa Olivetti, elsao@mit.edu; Yuriy Roman, Massachusetts Institute of Technology; Gerbrand Ceder, University of California, Berkeley; Andrew McCallum, University of Massachusetts, Amherst
Participating Institutions: Massachusetts Institute of Technology; University of California, Berkeley; University of Massachusetts, Amherst
Website: https://www.synthesisproject.org/
Keywords: Natural Language Processing, Materials Synthesis, Solid state synthesis, Sol-gel synthesis, Zeolites

Project Scope
The objective of the research is to build predictive tools for materials synthesis leveraging natural language processing techniques coupled with thermochemical information to identify patterns in the parameters we extract from existing literature. In this way we aim to accelerate the development and testing of new materials for a variety of applications.

Relevance to MGI
Our project benefits from computational automation and tools. We have established data links between extracted compounds and high-throughput computed ab initio data in the Materials Project, thereby further building out a connected ecosystem of data. This work helps to accelerate materials discovery is by linking high-throughput atomistic simulations, literature mining, human-computer interaction, synthesis and characterization. This data-driven approach linked with experiment has let us develop new theories of electrostatic interactions between the frameworks for zeolites and the structure-directing agents that influence the stability and catalytic properties of the desired zeolites. Our dataset of synthesis "recipes" facilitates the synthesis of known materials by providing an initial clue on the choice of precursors, conditions or synthesis type for a particular material. The fundamental insights related to synthesis described above also enable broader acceleration of materials discovery outside of the cases we explore.

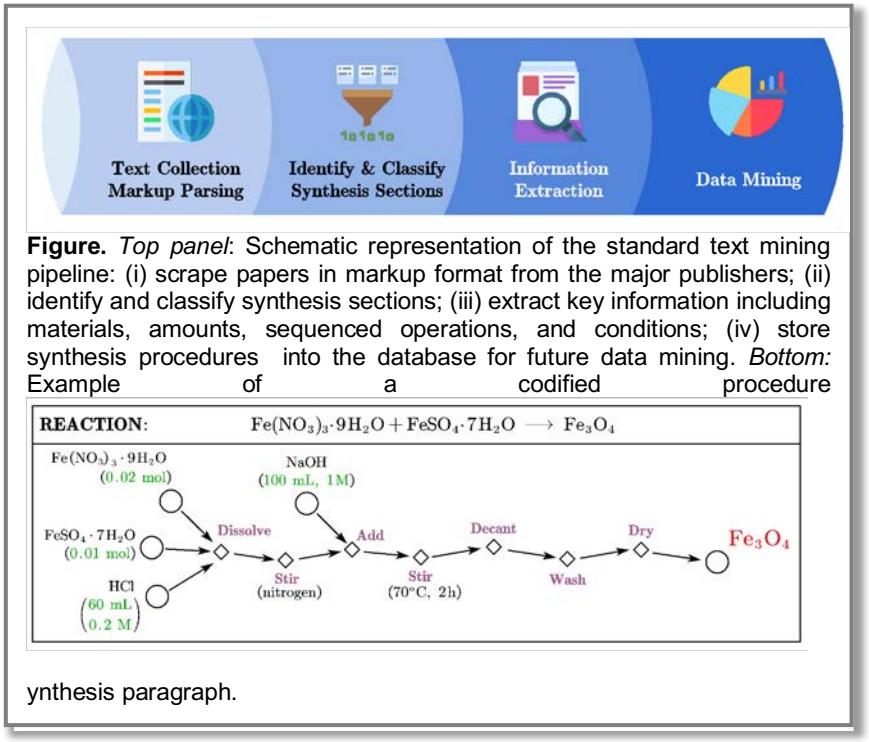


Figure. *Top panel*: Schematic representation of the standard text mining pipeline: (i) scrape papers in markup format from the major publishers; (ii) identify and classify synthesis sections; (iii) extract key information including materials, amounts, sequenced operations, and conditions; (iv) store synthesis procedures into the database for future data mining. *Bottom:* Example of a codified procedure ynthesis paragraph.

Technical Progress and Future plans
Our technical progress has been in: natural language processing pipeline development, development of datasets in solid-state and solution-based synthesis procedures, machine learning model development to inform zeolite synthesis, and experimental synthesis of zeolites based on above data. We have improved our token classification (i.e., identifying specific synthesis aspects such as targets, materials, and operations) starting with a new dataset of 595 annotated paragraphs using new token labels that better capture important synthesis components and a new annotation procedure that increases both the efficiency and consistency of the dataset. We have also optimized our token classification model examining state-of-the-art modeling techniques including word embeddings (BERT, ElMO) and model architectures such as transformer models and attention mechanisms. For the dataset development,

for example, we applied advanced machine learning and natural language processing techniques to construct a dataset of 35,675 solution-based synthesis procedures extracted from the scientific literature. Each procedure contains essential synthesis information including the precursors and target materials, their quantities, and the synthesis actions and corresponding attributes. Every procedure is also augmented with the reaction formula. We have made freely available the first large dataset of solution-based inorganic materials synthesis procedures. We have now linked the experimental with all the calculations and database created from text extraction for the case of zeolites. This work will continue to develop out data sets on synthesis and link those to experimental inquiry, particularly for the vast domain of zeolites, but also energy relevant materials.

Broader Impacts and Workforce Development

By developing content for tutorials in machine learning for materials and natural language processing, team members focus on formalizing the methods we are using in the research and put them in context for the larger domain of data driven materials informatics. These include development of computer notebooks for tutorial participants. There have been several conferences that are specific to machine learning methods and knowledge graph development outside of the materials domain that the students have participated in.

Data Management and Open Access

We have made code available through GitHub as it is refined for public consumption. We have content available through the Materials Project and synthesisproject.org as well as https://github.com/CederGroupHub/text-mined-synthesis_public. A recent goal have been to make these tools more user friendly as we have regular inquiries about the datasets and models through our synthesisproject email lists. The research team has been in contact with other groups doing NLP on material science domain. Work in the new dataset generating in the reporting period has brought data to size parity with other datasets in the related domains (organic synthesis procedures).

Advancing Along the Materials Development Continuum and Partnerships to Translation

The datasets in solid state synthesis and sol-gel synthesis provided by the team at Berkeley have fostered data collaborations throughout the project. Our work on zeolite synthesis provides a specific example of the iterative work of the project from simulation to materials informatics to synthesis. In this case, we use simulation to calculate the binding affinity for an organic structure directing agent that forms the pore structure within a zeolite. This simulation provides a binding matrix for close to 600K zeolite-OSDA pairs using a computational pipeline. The basis set for this simulation was created by extracting around 550 OSDAs from the literature through automated extraction methods developed between UMA, Berkeley and MIT. The simulation data was also benchmarked on its ability to reproduce 1100 zeolite-OSDA pairs extracted from the literature. These data extracted from literature were used to refine the computational calculations to use more than just binding energies. This capability has let the experimentalists within the work propose new OSDAs for synthesis of specific zeolites. The experimental data from synthesizing these newly proposed candidates can then feed back into the simulation work.

Publications and References

Wang, Z., Kononova, O., Cruse, K., He, T., Huo, H., Fei, Y., Zeng, Y., Sun, Y., Cai, Z., Sun, W. and Ceder, G., *Dataset of solution-based inorganic materials synthesis procedures extracted from the scientific literature*. Scientific Data, 9, 1-11. (2022)

H. Uvegi, Z. Jensen, T.N. Hoang, B. Traynor, T. Aytas, R. Goodwin, E.A. Olivetti, *Literature mining for alternative cementitious precursors and dissolution rate modeling of glassy phases* (2021) , Journal of the American Ceramic Society)

Z Jensen, S Kwon, D. Schwalbe-Koda, C. Paris, R. Gomez-Bombarelli, Y. Román-Leshkov, A. Corma, M. Moliner, E. Olivetti, *Discovering Relationships between OSDAs and Zeolites through Data Mining and Generative Neural Networks*. ACS Central Science, (2021)

R. Mahbub, K Huang, Z. Jensen, Z. Hood, J Rupp, EA. Olivetti, *Text mining for processing conditions of solid-state battery electrolytes* Electrochemistry Communications (2020)

Schwalbe-Koda, D., Kwon, S., Paris, C., Bello-Jurado, E., Jensen, Z., Olivetti, E., Willhammar, T., Corma, A., Román-Leshkov, Y., Moliner, M. & Gómez-Bombarelli, R. *A priori control of zeolite phase competition and intergrowth with high-throughput simulations*. Science, 374, 308-315. (2021).

Olivetti, E. A., Cole, J. M., Kim, E., Kononova, O., Ceder, G., Han, T. Y. J., & Hiszpanski, A. M. *Data-driven materials research enabled by natural language processing and information extraction*. Applied Physics Reviews, 7, 041317. (2020).

O Kononova, T. He, Haoyan Huo, Amalie Trewartha, Elsa A. Olivetti, and Gerbrand Ceder. *Opportunities and challenges of text mining in materials research*. Iscience 24, (2021): 102155.

Accelerating Discovery of Mixed Rare Earth Silicates for Extreme Environments

Lead Investigator: Elizabeth Opila, opila@virginia.edu.
Participating Institutions: University of Virginia; University of Texas, Dallas; Rolls-Royce Inc.
Website: none.
Keywords: multicomponent rare-earth silicates, thermochemical stability, thermal conductivity, elastic modulus, thermal expansion

Project Scope

The discovery of new materials that enable higher temperature and more efficient turbine operation will be accelerated through computational methods validated with experimental results. Materials discovery of multicomponent rare earth silicates suitable for turbine component coatings will be conducted using AFLOW: high throughput property prediction. Controlled combinations of the seventeen rare earth elements in mono- and di-silicates provide opportunities to optimize desirable coating properties, including low thermal conductivity, thermal expansion, modulus, and high stability in reactive environments.

Relevance to MGI

The computation-experiment-feedback loop coupled with machine learning and high throughput computation will enable optimization of coating composition and properties for extreme environment application.

Technical Progress

Recent progress includes systematic DFT-informed calculations of rare earth disilicate thermal conductivity (**Figure 1**), thermal expansion coefficient and elastic moduli to identify property trends as a function of cation radius in comparison to experimental measurements. Melting temperature of a series of increasingly complex multicomponent rare earth di-silicates and mono-silicates have been measured by laser heating and thermal arrest (latent heat) on cooling. Comparison will be made to results from lattice parameter and crystal structures determined at the Argonne National Lab synchrotron with levitated laser heated specimens.

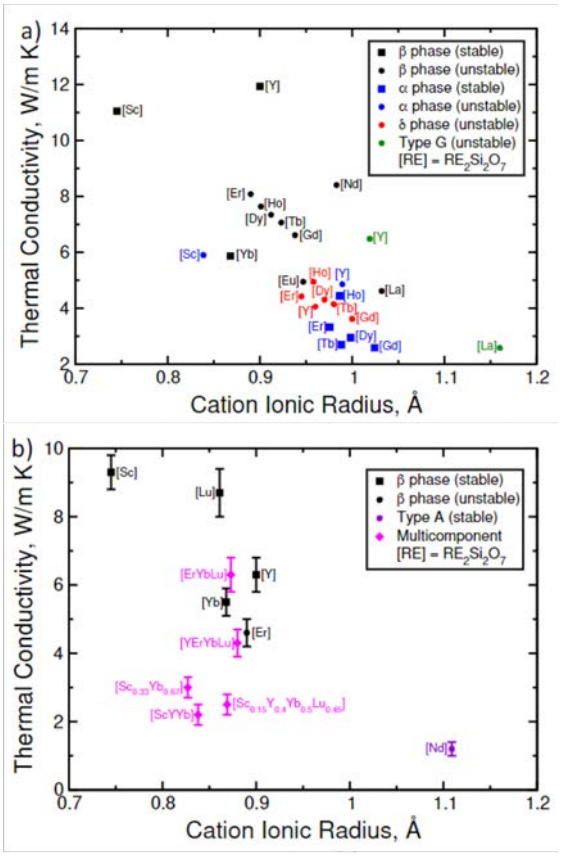


Figure 1. Thermal conductivity as a function of cation radius for rare earth silicates obtained from (a) DFT calculations and (b) experimental measurements.

Future Plans

Future plans include extending this approach to thermochemical stability of rare earth silicates in high temperature reactive environments and ionic transport rates to further develop computational methods and property prediction essential for application in extreme environments.

Broader Impacts and Workforce Development

This project involves seven graduate students who are exposed to the MGI philosophy, important for work force development. In addition, a full-day online AFLOW workshop was organized by the PIs of this project at University

of Virginia on August 17, 2021. This workshop focused on online tools for database access, structure analysis, and machine-learning, with a major objective being to introduce experimentalists to materials genomics concepts, tools and resources and was attended by at least 25 students and faculty. Opila also presented a webinar February 17, 2021 for the American Ceramic Society that was attend by over 70 participants, entitled, "High entropy environmental barrier coatings" that described the approach and results from this effort.

Data Management and Open Access

Results for this effort are publicly available in journal publications, partially listed below. In addition, data are made available through the AFLOW database https://aflowlib.org/ and the UVA LibraData Repository https://doi.org/10.18130/V3/XI7QGF.

Advancing Along the Materials Development Continuum and Partnerships to Translation

The project is accelerating materials discovery of rare earth silicates by systematically exploring thermodynamic stability, crystal structure and bond strength as a function of rare earth cation size, mass, and oxygen coordination. This enables trends in properties such as thermal conductivity, thermal expansion, and modulus to be predicted for complex mixtures of silicates. This is a GOALI award in collaboration with Rolls-Royce, Inc. who is interested in commercializing these compositions as high temperature coatings for turbine engine environments. A patent application has been filed in collaboration with Rolls-Royce: US Patent App. 16/723,675, 2020 Thermal and/or environmental barrier coating system.

Publications and References

1, R. Friedrich, M. Esters, C. Oses, S. Ki, M.J. Brenner, D. Hicks, M.J. Mehl, C. Toher, and S. Curtarolo, *Automated coordination corrected enthalpies with AFLOW-CCE*, Physical Review Materials **5,** 043803 (2021).
2. D. Hicks, M.J. Mehl, M. Esters, C.Oses, O. Levy, G.L.W. Hart, C. Toher, and S. Curtarolo, *The AFLOW Library of Crystallographic Prototypes: Part 3*, Computational Materials Science **199,** 110450 (2021).
3. G.L.W. Hart, T. Mueller, C. Toher, and S. Curtarolo, *Machine learning for alloys*, Nature Reviews Materials **6,** 730–755 (2021).
4. C.Toher, C. Oses, M. Esters, D. Hicks, G.N. Kotsonis, C. Rost, D.W. Brenner, J.-P. Maria, and Curtarolo, *High-entropy ceramics: Propelling applications through disorder*, MRS Bulletin **47,** 194–202 (2022).
5. H. Kulik, T. Hammerschmidt, J. Schmidt, S. Botti, M.A.L. Marques, M. Boley, M. Scheffler, M. Todorović, P. Rinke, C. Oses, A. Smolyanyuk, S. Curtarolo, A. Tkatchenko, A Bartok, S. Manzhos, M. Ihara, T. Carrington, J. Behler, O. Isayev, M. Veit, A. Grisafi, J. Nigam, M. Ceriotti, K.T. Schütt, J. Westermayr, M. Gastegger, R. Maurer, B. Kalita, K. Burke, R. Nagai, R. Akashi, O. Sugino, J. Hermann, F. Noé, S. Pilati, C. Draxl, M. Kuban, S. Rigamonti, M. Scheidgen, M. Esters, D. Hicks, C. Toher, P. Balachandran, I. Tamblyn, S. Whitelam, C. Bellinger, and L.M. Ghiringhelli, *Roadmap on Machine Learning in Electronic Structure*, Electronic Structure, in press (2022).
6. K. Quiambao-Tomko, R.R.White, J.A. Tomko, C.M. Rost, L. Backman, E.J. Opila, and P.E. Hopkins, *Detection of sub-micrometer thermomechanical and thermochemical failure mechanisms in titanium with a laser-based thermoreflectance technique*, Journal of Applied Physics **131,** 055104 (2022).
7. P.E. Hopkins, J.A. Tomko, and A. Giri, *Quasi-harmonic theory for phonon thermal boundary conductance at high temperatures*, Journal of Applied Physics **131,** 015101 (2022).
8. D.H. Olson, J.A. Deijkers, K. Quiambao-Tomko, J.T. Gaskins, B.T. Richards, E.J. Opila, P.E. Hopkins, and H.N.G. Wadley, *Evolution of microstructure and thermal conductivity of multifunctional environmental barrier coating systems*, Materials Today Physics **17,** 100304 (2021).
9. M. Ridley and E. Opila, *Variable thermochemical stability of $RE_2Si_2O_7$ (RE = Sc, Nd, Er, Yb, or Lu) in high-temperature high-velocity steam*, Journal of the American Ceramic Society **105,** 1330–1342 (2022).
10. M.J. Ridley and E.J. Opila, *Quantitative evaluation of (0001) sapphire recession in high-temperature high-velocity steamjet exposures,* J Eur Ceram Soc **42,** 631–637 (2022).

Chemical Reactivity and Spectroscopy Through Adaptive Quantum Mechanics/Many-Body Representations: Theoretical Development, Software Implementation, and Applications

Lead Investigator: Francesco Paesani, fpaesani@ucsd.edu.
Participating Institutions: University of California, San Diego
Website: https://paesanigroup.ucsd.edu/software/mbx.html.
Keywords: many-body interactions, data-driven models, machine learning, molecular dynamics, molecular fluids.

Project Scope

The scope of this project is the development and application of a new theoretical/computational framework, called many-body energy (MB-nrg), for predictive molecular dynamics (MD) simulations of chemical and physical transformations across different phases of relevance for water–energy systems.

Relevance to MGI

As discussed in "The Water–Energy Nexus: Challenges and Opportunities", several challenges must be addressed in order to advance water–energy systems, including the characterization of complex fluids, the control of chemical reactions, and the ability to make predictions about physical and chemical transformations. The ability to reliably predict and control the emergent behavior in these solutions requires fundamental knowledge of the underlying physical mechanisms which, in turn, demands a molecular-level understanding of the interactions between all the components of the solutions. This knowledge is key to developing engineering processes that can control the behavior of complex solutions at the microscopic level which, in turn, will enable a more efficient use of natural resources. Our project addresses this challenge by integrating advanced theoretical methods and computational techniques into efficient software for HPC platforms which enables predictive MD simulations, with unprecedented accuracy.

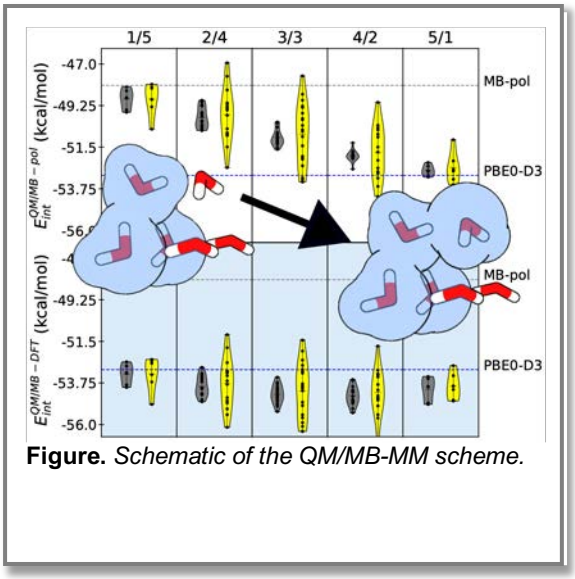


Figure. *Schematic of the QM/MB-MM scheme.*

Technical Progress

We demonstrated that our theoretical/computational MB-nrg framework for data-driven many-body potential energy functions (MB PEFs) is applicable to generic molecular fluids (e.g., CO_2/H_2O and CH_4/H_2O mixtures) and enable predictive MD simulations, with unprecedented accuracy. We developed and implemented our hybrid QM/MB-MM method which replaces conventional MM models with data-driven MB PEFs. Our QM/MB-MM scheme builds upon mutually polarizable QM/MM schemes for polarizable force fields with inducible dipoles and uses permutationally invariant polynomials to effectively account for quantum-mechanical contributions (e.g., exchange-repulsion, and charge transfer and penetration) that are difficult to describe by classical expressions adopted by conventional MM models. We demonstrated that our QM/MB-MM scheme enables a smooth energetic transition as molecules are transferred between QM and MB-MM regions, without the need of a transition layer. By effectively elevating the accuracy of both the MM region and the QM/MM interface to that of the QM region, we showed that our new QM/MB-MM approach achieves an accuracy comparable to that obtained with a fully QM treatment of the entire system. We extended our MB-nrg framework to the development of MB PEFs with arbitrary QM accuracy, and introduced a general framework for the development of data-driven MB PEFs (MB-QM PEFs) that represent the interactions between molecules at an arbitrary quantum-mechanical (QM) level of theory. Through a systematic analysis of individual many-body contributions to the interaction energies, we demonstrated that all MB-QM PEFs preserve the same accuracy as the corresponding QM calculations. We demonstrated that

DFT limitations in representing aqueous systems can be overcome within our MB-nrg framework by using density-corrected functionals that provide a more consistent representation of individual many-body contributions.

Future Plans

Building upon the current MB-nrg theoretical/computational framework, we are planning to recast the Bethe–Goldstone-like equation for covalently-bonded systems into a data-driven many-body expansion that allows for calculating the energy of generic molecules as the sum of individual *n*-body contributions associated with distinct monomers. In parallel, we aim to achieve a large degree of portability for our MBX software by building new accelerator-based capabilities into the already highly modular design of MBX. We will achieve this goal by executing a strong, collaborative software development plan just as we have done so already in creating a new interface in MBX to enable large-scale parallel many-body molecular dynamics simulations with LAMMPS.

Broader Impacts and Workforce Development

Our team comprises several undergraduate and graduate students, and postdoctoral scholars, as well as high school students from underprivileged and underrepresented minorities in the transborder San Diego-Tijuana region. All members are trained in modern programming languages, computer simulations, and theoretical models.

Data Management and Open Access

All data generated as part of our project are stored on a central, dedicated primary file server. Automatic snapshots of this storage pool are taken daily, hourly, and frequently (every 15 minutes). Any data deemed unneeded or unnecessary for preservation are archived onto single-copy tape and removed from the server. Data requests from interested researchers are fulfilled upon request. Specific computational details necessary to reproduce the results of our publications are made available on GitHub: https://github.com/paesanilab/Data_Repository. Our MBX software is open and publicly available on GitHub: https://github.com/paesanilab/MBX.

Advancing Along the Materials Development Continuum and Partnerships to Translation

By integrating advanced theoretical methodologies and state-of-the-art computational techniques into efficient software for exascale HPC platforms, our MB-nrg framework enables predictive computer simulations of complex molecular systems, with unprecedented accuracy. We already demonstrate the predictive power of our MB-nrg framework in fundamental studies of aqueous solutions, molecular fluids, and metal-organic frameworks. We have recently started contacting VC firms potentially interested in commercializing our software.

Publications and References

1. M. Riera. E. Lambros, T.T. Nguyen, A.W. Götz, F. Paesani, *Low-order many-body interactions determine the local structure of liquid water*, Chem. Sci. **10**, 8211 (2019).
2. M. Riera, E.P. Yeh, F. Paesani, *Data-driven many-body models for molecular fluids: CO_2/H_2O mixtures as a case study*, J. Chem. Theory Comput. **16**, 2246 (2020).
3. E. Lambros, F. Lipparini, G.A. Cisneros, F. Paesani, *A many-body, fully polarizable approach to QM/MM simulations*, *J. Chem. Theory Comput.* **16**, 7462 (2020).
4. M. Riera, A. Hirales, F. Paesani, *Data-driven many-body models with chemical accuracy for CH_4/H_2O mixtures*, J. Phys. Chem. B **124**, 11207 (2020).
5. V.W.D. Cruzeiro, A.P. Wildman, X. Li, F. Paesani, *On the relationship between hydrogen-bonding motifs and the $1b_1$ splitting in the X-ray emission spectrum of liquid water*, J. Phys. Chem. Lett. **12**, 3996 (2021).
6. E. Lambros, J. Hie, F. Paesani, *Assessing the accuracy of the SCAN functional for water through a many-body analysis of the adiabatic connection formula*, J. Chem. Theory Comput. **17**, 3739 (2021).
7. E.F. Bull-Vulpe, M. Riera, A.W. Götz, F. Paesani, *MB-Fit: Software infrastructure for data-driven many-body potential energy functions*, J. Chem. Phys. **155**, 124801 (2021).
8. S. Dasgupta, E. Lambros, J.P. Perdew, F. Paesani, *Elevating density functional theory to chemical accuracy for water simulations through a density-corrected many-body formalism*, Nat. Commun.**12**, 6359 (2021).
9. S. Yue, M. Riera, R. Ghosh, A.Z. Panagiotopoulos, F. Paesani, *Transferability of data-driven, many-body models for CO_2 simulations in the vapor and liquid phases*, J. Chem. Phys. **156**, 104503 (2022).
10. E. Palos, E. Lambros, S. Swee, J. Hu, S. Dasgupta, F. Paesani, *Assessing the interplay between functional-driven and density-driven errors in DFT models of water*, J. Chem. Theory Comput. **18**, 3410 (2022).

The Materials Project

Lead Investigator: Kristin Persson, kapersson@lbl.gov
Participating Institutions: Lawrence Berkeley National Laboratory, UC San Diego, Dartmouth College
Website: https://www.materialsproject.org
Keywords: electronic structure, databases, open-source software

Project Scope

The vision of the Materials Project (https://www.materialsproject.org/) is to leverage the continuing developments in materials theory to compute many of the basic properties of all known materials and beyond, in order to a) rapidly screen for novel materials with interesting properties; b) accelerate the development of quantitative higher length scale models; and c) form a well-tested and curated dataset on which to develop and test machine learning and materials informatics methods to extract insights. In parallel, we demonstrate how such an information-driven approach serves the materials research community and accelerates materials design. Within the BES-MSE core program, we apply our approach on novel functional electronic compounds, alloy systems and interfaces exhibiting magnetic, elastic and/or electric coupling, and topological materials.

Relevance to MGI

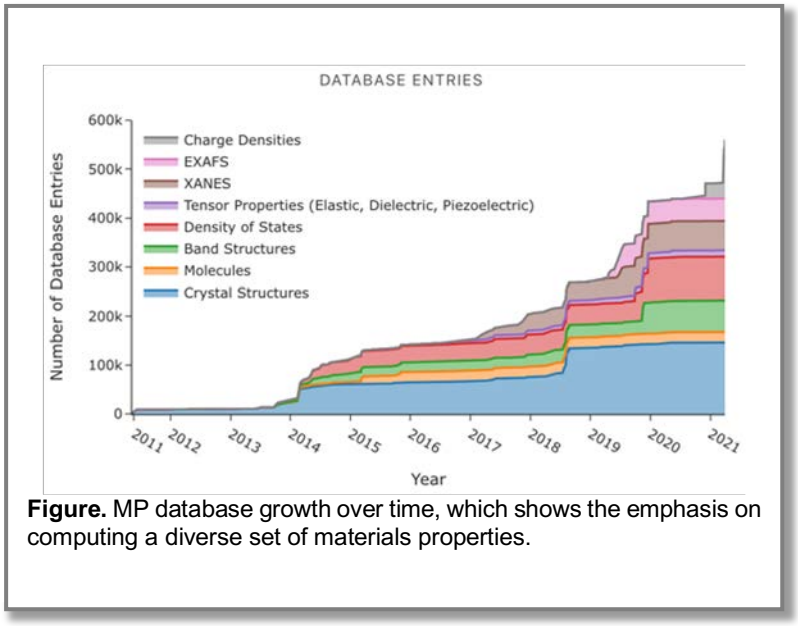


Figure. MP database growth over time, which shows the emphasis on computing a diverse set of materials properties.

Our program leverages accurate materials property and structure prediction with high-performance computing and data mining to realize an efficient materials data and design platform. Many property requirements transcend a particular application (e.g., elasticity, charge transport, stability, permittivity, etc), which means that gathering all properties in one place, combined with the software infrastructure and expertise already found at the Materials Project is maximizing the impact on the materials community. Our program leverages accurate materials property and structure prediction with high-performance computing and data mining to realize an efficient materials data and design platform. Many property requirements transcend a particular application (e.g., elasticity, charge transport, stability, permittivity, etc), which means that gathering all properties in one place, combined with the software infrastructure and expertise already found at **the Materials Project** is maximizing the impact on the materials community. Indeed, its largest impact is not coming from the people who calculate and organize the data, but from the hundreds of thousands of scientists in academia and industry we empower to perform data-driven design and development. This highlights the role of the Materials Project as a de facto "Materials Genome" for the research community.

Technical Progress

The Materials Project has made recent technical progress in a variety of areas, including development and application of new theoretical methods, new materials discovery efforts, a completely updated and revamped backend and frontend data infrastructure that can serve as a model for cloud/DOE hybrid platforms, new community

outreach initiatives, increased application of machine learning in the framework, and new data sets generated both internally and contributed form external users. Newly developed atomate (open-source automatic workflows for computational materials science) modules include workflows for amorphous, magnetic, grain boundary, work-function and charged point defect systems. We launched MPContribs-ML, a resource for machine learning data sets and benchmarking machine learning algorithms. The Matbench site, now fully integrated into MPContribs-ML, has comprehensive benchmarks for many of the state-of-the-art algorithms developed for structure-property and composition-property relationships and maintain a current "leaderboard" of best methods. MP has also developed several novel machine learning frameworks. For example, a new graph-based ML model, GUFF17, can perform structural relaxation, phonon calculations, molecular dynamics, elasticity calculations, etc., for all materials. We developed the AtomSets framework to perform transfer learning to help researchers use materials data of different sizes for training ML models. We also integrated materials synthesis information from text mining of 5 million research papers and integrated synthesis information into Materials Project.

Based on theoretical predictions, we synthesized $BiInO_3$, a new polar material predicted to exhibit exotic persistent spin texture, and solid solutions of $PbV_{1-x}Ti_xO_3$, exploring the phase boundary for enhanced piezoelectric response. We have also probed a promising ferroelectric system ($Pb_{1-x}Sr_xHfO_3$) derived previously through a high-throughput computational search. Of particular interest was $Pb_{0.5}Sr_{0.5}HfO_3$ which was found to exhibit the highest electric-breakdown strength (E_B) of 5.12 MV/cm (three-times the value for $PbHfO_3$), a 6% increase in the capacitive energy storage efficiency (from 91% for $PbHfO_3$ to 97%), a 2.7-times increase in the recoverable energy density (from 21 J/cm^3 for $PbHfO_3$ to 77 J/cm^3), and a 10^6-times improvement in fatigue endurance (cycling stability). This performance places $Pb_{0.5}Sr_{0.5}HfO_3$ among the most promising antiferroelectrics for such capactive energy storage applications and this work was published in *Adv. Mater. 2021*. Other investigations are currently in progress.

Future Plans

Future plans include increased high-throughput accuracy (R2SCAN metaGGA calculations, non-collinear magnetism, efficient calculation of vibrational properties including anharmonic effects), increasing compound coverage (machine learning predictions, non-stoichiometric compounds using cluster expansions), design of novel functional topological, ferroic and Heckmann materials, data-driven realization and validation of target solid-state inorganic materials, and various infrastructure upgrades to computing science and operations including a completely new web site, API updates, and further evolving the MPContribs data contribution platform.

Broader Impacts and Workforce Development

In addition to the core work in theory, experimentation, and data dissemination, MP regularly interacts with its community through several avenues: MP newsletter, monthly seminar series (2000 registrants, typically 250–350 live attendees), a three-day annual workshop (now in its seventh year), the Materials Science Community Discourse (discussion forum with 7000 unique visits/day), and a Github platform for discussions about code features and bug fixes with other scientific software developers and contributors.

Data Management and Open Access

The Materials Project makes data accessible through a public web site as well as an API. It was officially recognized as a Public Reusable Research (PuRE) data resource, a DOE initiative which recognizes that MP is an authorative resource in its community, and ensures its data management meets specific standards. Furthermore, MP now serves as a data repository for external data sets. The existing capabilities of MPContribs with respect to data discoverability, maintainability, and exposure have already led the Materials Science Division (MSD) at LBNL to recommend it as a central part of its data management plan.

Advancing Along the Materials Development Continuum and Partnerships to Translation

The Materials Project has been used both internally and by the community to predict several functional materials whose properties have been confirmed by experiment. Furthermore, a large fraction of users are affiliated with industry which serves as a testament to the fact that such a materials genome approach can be translated to commercial applications.

Publications and References

Recent publications:

1. M. Acharya, S. Mack, A. Fernandez, J. Kim, H. Wang, K. Eriguchi, D. Meyers, V. Gopalan, J. B. Neaton, and L. W. Martin, "Searching for New Ferroelectric Materials Using High-Throughput Databases: An Experimental Perspective on BiAlO3 andBiInO3", Chem. Mater. 32, 7274 (2020) DOI: 10.1021/acs.chemmater.0c01770

2. A.M. Ganose, J. Park, A. Faghaninia, R. Woods-Robinson, K.A. Persson, and A. Jain. "Efficient calculation of carrier scattering rates from first principles", Nat. Commun. 12, 2222 (2021). DOI: 10.1038/s41467-021-22440-

3. M. Markov, L. Alaerts, H. Pereira Coutada Miranda, G. Petretto, W. Chen, J. George, E. Bousquet, P. Ghosez, G.-M. Rignanese, and G. Hautier. "Ferroelectricity and multiferroicity in anti–Ruddlesden–Popper structures", Proc. Natl. Acad. Sci. U.S.A., 118, e2026020118 (2021). DOI: 10.1073/pnas.2026020118

4. Mattei, J. Dagdelen, M. Bianchini, A. Ganose, A. Jain, E. Suard, F. Fauth, C. Masquelier, L. Croguennec, G. Ceder, K. A. Persson, and P. Khalifah, "Enumeration as a tool for structure solution – a materials genomic approach to solving the cation-ordered structure of Na3V2(PO4)2F3", Chem. Mater., 32, 20, 8981–8992 (2020) DOI: 10.1021/acs.chemmater.0c03190

5. J. Kim, M. Acharya, S. Saremi, G. Velarde, E. Parsonnet, P. Donahue, A. Qualls, D. Garcia, L. W. Martin, "Ultrahigh capacitive energy density in ion-bombard relaxor ferroelectric films", Science 369, 81-84 (2020). DOI: 10.1126/science.abb0631

6. H. Zheng, X. Li, R. Tran, C. Chen, M. Horton, D. Winston, K.A. Persson, S.P. Ong. "Grain Boundary Properties of Elemental Metals", Acta Materialia, 186, 40-49 (2020) DOI: 10.1016/j.actamat.2019.12.030

7. T. E. Smidt, S. A. Mack, S. E. Reyes-Lillo, A. Jain, and J. B. Neaton, "An automatically curated first-principles database of ferroelectrics", Sci. Data 7, 72 (2020) DOI: 10.1038/s41597-020-0407-9

8. C. Chen, Y.X. Zuo, W.K. Ye, X.G. Li, and S.P. Ong. "Learning Properties of Ordered and Disordered Materials from Multi-Fidelity Data", Nat. Comput. Sci. 1, 46-53 (2021). DOI: 10.1038/s43588-020-00002-x

9. A.M. Ganose, A.J. Searle, A. Jain and S.M Griffin. "iFermi: A Fermi surface analysis and manipulation tool", J. Open Source Softw. , 6(59), 3089 (2021). DOI: 10.21105/joss.03089

10. A. Wang, R. Kingsbury, M. McDermott, A. Jain, S.P. Ong, S. Dwaraknath, and K.A. Persson. "A Framework for Quantifying Uncertainty in DFT Energy Corrections," Sci. Rep. 11 (1), 15496 (2021). DOI: 10.1038/s41598-021-94550-5

Coarse-grain models of polymer melts: Tools and methods for preserving structure and dynamics

Lead Investigator: Frederick R. Phelan Jr., Materials Science & Engineering Division, National Institute of Standards & Technology (NIST), frederick.phelan@nist.gov.
Investigator: Lilian C. Johnson, Materials Science & Engineering Division, National Institute of Standards & Technology, lilian.johnson@nist.gov.
Keywords: Molecular Dynamics (MD); Coarse-grained force-fields (CGFF); Iterative Boltzmann Inversion (IBI); Dissipative Potentials; Friction Parametrization; Polymer Dynamics.

Project Scope

We are developing software and data tools which parametrize coarse-grained force-fields in a systematic way while preserving the chemistry, thermodynamics, and dynamics of the underlying atomistic system. Our approach is based on statistical thermodynamics where we utilize conservative potentials to model the molecular forces and dissipative potentials to account for the smoother potential energy landscape arising from the loss of configurational entropy. For conservative potentials, we have developed a software code based on the Iterative Boltzmann Inversion (IBI) method. For dissipative potentials, we are developing tools and approaches to parametrize friction based the Langevin thermostat as a first and the most basic approach to approximate the lost degrees of freedom.

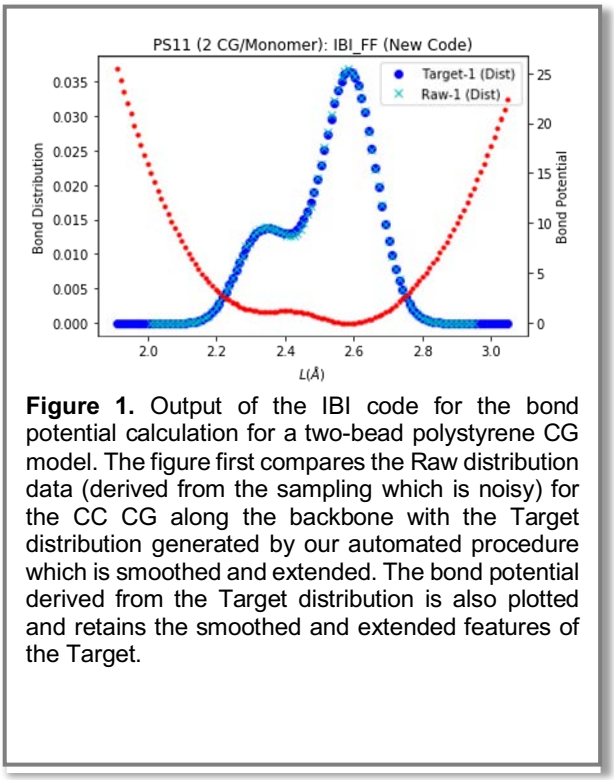


Figure 1. Output of the IBI code for the bond potential calculation for a two-bead polystyrene CG model. The figure first compares the Raw distribution data (derived from the sampling which is noisy) for the CC CG along the backbone with the Target distribution generated by our automated procedure which is smoothed and extended. The bond potential derived from the Target distribution is also plotted and retains the smoothed and extended features of the Target.

Relevance to MGI

Computation at mesoscopic length scales between the atomistic and continuum for polymeric and similar soft materials requires coarse-graining techniques. These techniques enable computational access to the greater length and time scales inherent to these materials by subsuming atoms associated with monomers or other chemical moieties into "coarse-grained" sites and then describing the physics by means of a potential parametrized for the coarse-grained description of the material. Bottom-up methods such as IBI, in which the CGFF is directly derived from atomistic simulation data, have the advantage of retaining a high degree of chemical specificity, and thus, can typically capture equilibrium structure quite well. However, bottom-up methods retain a Hamiltonian description of the force-field potential but with reduced degrees of freedom (or configurational entropy). This results in accelerated dynamics and a softer mechanical response due to a smoothed free-energy landscape arising from a reduction in topological constraints and fluctuations associated with the fine-grained atomistic details, which are averaged out. We employ an approach to correct the accelerated dynamics of CG models based on the statistical mechanics of coarse-graining by the addition and parameterization of a nonconservative (or dissipative) potential to the material description as represented by a Hamiltonian potential. The dissipative potential typically involves a friction parameter whose parametrization is fundamentally related to the integrals of autocorrelation functions of the atomistic system, making the entire approach derivative of atomistic reference simulations. Our approach combines these two methods and supports the NIST role in the MGI to create techniques and standards for the interoperation of modeling systems at multiple length and time scales.

Technical Progress

- **Iterative Boltzmann Inversion**: We have developed a new software code based on the Iterative Boltzmann Inversion (IBI) method which automates the development of conservative potentials for coarse-grained molecular dynamics (MD) calculations of soft materials.
- **Dissipative Potentials**: We are developing tools and approaches for parametrizing friction based on different measures of translational and rotational motion, and thermomechanical properties of molecular systems, as well as investigating the temperature dependence of the friction.
- **COMSOFT-Tools**: Our code is built on a series of reusable Python modules called COMSOFT Tools which are designed to enable data handling, data interoperability and molecular analysis functions for applications involving multiscale MD data.

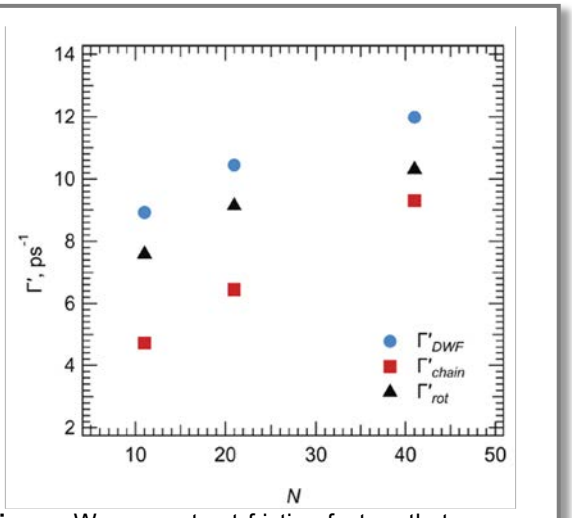


Figure. We can extract friction factors that recover specific measures of dynamics shown here for polystyrene oligomers with varying number of monomers, N. Dynamics here include the Debye-Waller factor, chain diffusivity, and chain rotational relaxation time. From reference 1, not subject to copyright in the US.

Future Plans

Our immediate plans are to publish the IBI code through the COMSOFT-Tools GitHub repository along with documentation on usage. The first step of future plans will be to integrate tools for parametrizing friction into the repo which we anticipate will be driven by a process of continuous improvement as we develop practical implementations of theoretical concepts. We also anticipate the development of better tools for mapping all-atom systems into coarse-grains and machine learning approaches based on COMSOFT-Tools for the better description of bonded potentials.

Broader Impacts and Workforce Development

Training in the usage of our software tools will be accomplished through online manuals, tutorials, videos and possibly workshops. We anticipate that as the tools mature and achieve greater usage, their knowledge will naturally be carried into the workforce by undergraduates and especially graduate students working in molecular simulation of soft materials.

Data Management and Open Access

As mentioned above, the IBI code and related tools will be published through an open access GitHub repository. Data derived from the software will be published through various auspices, including publications, the NIST Science Data portal, and through the companion Web Force-Field Project (WebFF) which will provide an online platform for publishing force-field data for soft materials.

Advancing Along the Materials Development Continuum and Partnerships to Translation

Accurate force-fields for coarse-grained MD are very important to future materials discovery but remain very difficult to develop for quantitative (rather than qualitative) materials calculations. We anticipate that the development and maturation of these tools will greatly accelerate the development of CGFF, which will directly aid the MGI approach to materials development wherever it is being employed. We have had some early discussion with a commercial software company about the implementation of the IBI code in their package.

Publications and References

1. Lilian C. Johnson and Frederick R. Phelan, Jr., "Dynamically consistent coarse-grain simulation model of chemically specific polymer melts via friction parameterization," J. Chem. Phys. 154 (8), 084114 (2021). DOI: 10.1063/5.0034910

Transforming Photonics and Electronics with Digital Alloy Materials

Lead Investigator: Viktor Podolskiy, viktor_podolskiy@uml.edu.
Participating Institutions: U Mass Lowell, UT Austin, Purdue U, U Virginia
Website: none
Keywords: Digital Alloys, Hyperbolic Materials, Metamaterials, Plasmonics, Strong light-matter interaction

Over the course of the project we aim to capitalize on the advances in digital alloy semiconductor growth technologies and in the fields of plasmonics and metamaterials, to simultaneously engineer electronic and photonic wavefunctions to explore and realize new avenues for light-matter interaction at mid-IR frequencies.

Project Scope

Our project aims to address the fundamental aspect of light-matter interactions, strong mismatch between the wavelength of light and the size of the atom by bringing together opportunities for light compression in hyperbolic metamaterials with electronic wavefunction engineering brought by digital alloy technologies. The team brings together experts from four Universities and ARFL.

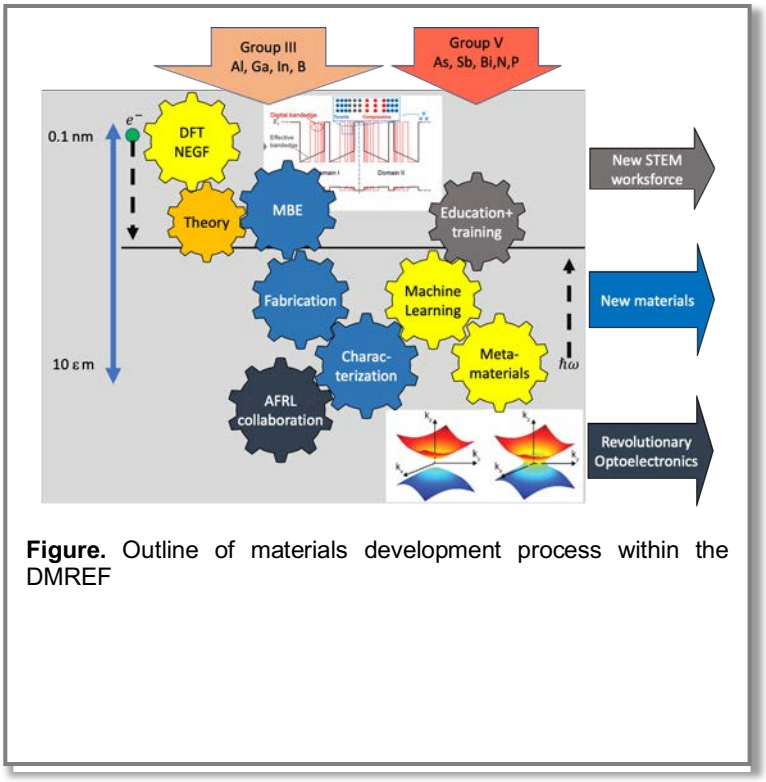


Figure. Outline of materials development process within the DMREF

Relevance to MGI

The members of the team are experts in molecular beam epitaxy, microfabrication, optical and electronic characterization, condensed matter theory, electromagnetic theory and simulations, first-principles simulations, quantum transport, and machine learning. By working together in a unified iterative loop, the team aims to comprehensively address materials discovery process (see Fig.). In order to facilitate the rapid materials discovery and development, the team has already instituted regular bi-weekly whole-team progress meetings that identify crucial bottlenecks in the progress, with additional sub-teams meetings on as-needed basis to address these bottle-necks.

The current team grows out of multiple former NSF programs (DMREF [UML, UT, Purdue], NSF-ECCS [UT, UVA]). We have used the first 9 months of current DMREF to develop a vision for the electronic and photonic control of our materials, and to evaluate initial analytical, computational, synthesis, and characterization efforts for the program, via meetings with all DMREF PIs and student/postdoc members of their research teams.

Technical Progress

As already mentioned above, we used the first nine months of the project to establish the regular progress meetings and the common terminology base for the members of the team. Leveraging the progress of our previous DMREF project, the analytical theory, developed by Purdue, and rigorous electromagnetic simulations, developed by UML were used to design the materials parameters and the geometry to strongly confine mid-IR light within

wavelength/15 spatial areas, while simultaneously generating strong gradients of electromagnetic fields. The growth/microfabrication run has been completed and the materials are currently being characterized by the UT team. UVA team is building computational models for understanding the minigap formation in strained digital alloy superlattices and their impact on AC conductivity and electron-photon coupling.

In parallel with the experimental efforts, UML and Purdue teams are continuing to develop the framework for the new classes of electromagnetic materials whose electromagnetic response would be dominated by the strong field gradients (as opposed to average electromagnetic fields). UVA, UML, and Purdue teams are working on incorporating these developments within the non-equilibrium Greens function formalism that would allow faster and more accurate development of the resulting materials platform.

Future Plans

In the upcoming year we aim to build upon these initial studies, realizing strong field gradients and using these designed materials to explore the electromagnetic response beyond the dipole approximation. We also expect to bring online new, first principles-based tools for calculating such electromagnetic response beyond the currently-available effective mass approximations.

Broader Impacts and Workforce Development

DMREF currently supports four graduate students at participating institutions. These students are learning the broader set of subjects that naturally come with interdisciplinary materials-science projects. Three students supported by previous DMREF/NSF projects have graduated with their Ph.D and are now employed at Applied Materials (UT student), PSI (UML student) and Intel (UVA student). During the proposal planning process, we anticipated student-exchange program to facilitate converging training. UT Ph.D. student is spending Summer 2022 at AFRL; while there, he will begin near-field optical measurements on structures designed by Purdue/UML and grown/fabricated at UT. We are now planning additional exchanges that we expect to take place in 2022 and 2023.

Data Management and Open Access

The members of the team are committed to FAIR data practices. In particular, our research codes will be available to the broader community. The first set of these, stemming from the recently completed DMREF project is available on GitHub (https://github.com/EvanSimmonsUML/BallisticMetamaterials), and is currently being used by the team to design the first generation of materials for the current DMREF. Graphics User Interface for this set of codes is under development.

Advancing Along the Materials Development Continuum and Partnerships to Translation

The team has partnered with ARFL to explore translation of fundamental materials science towards applications. Ph.D. student from UT Austin is spending the summer of 2022 with collaborators at AFRL, working on designs and simulations of novel photonic structures, as well as near field measurements of quantum structures in the mid-infrared.

NSF BioPACIFIC Materials Innovation Platform

Lead Investigator: Javier Read de Alaniz, javier@chem.ucsb.edu, Director of BioPACIFIC MIP. Eleni Papananou, hellen@ucsb.edu, User Coordinator of BioPACIFIC MIP, Arica Lubin, alubin@ucsb.edu, Education Lead of BioPACIFIC MIP
Participating Institutions: University of California Santa Barbara and University of California Los Angeles
Website: biopacificmip.org
Keywords: Materials Innovation Platform, Biopolymers, Data driven discovery, Automated polymerization, High-throughput bioreactor

Project Scope

BioPACIFIC MIP's mission is to enable the discovery, characterization, and development of advanced biomaterials through innovative research at the interface of synthetic biology and material science. This is accomplished by providing unique infrastructure and technology, engaging education, and the assemblage of shared knowledge through establishment and operation of a national user facility.

Relevance to MGI

The BioPACIFIC MIP serves as the hub for four interconnected research elements: (1) living bioreactors, (2) automated synthetic tools for the production of bio-derived materials, (3) hierarchical computational tools and (4) a state-of-the-art characterization facility to enable structure-property relationship determination. All in-house research projects are therefore driven by the MGI approach.

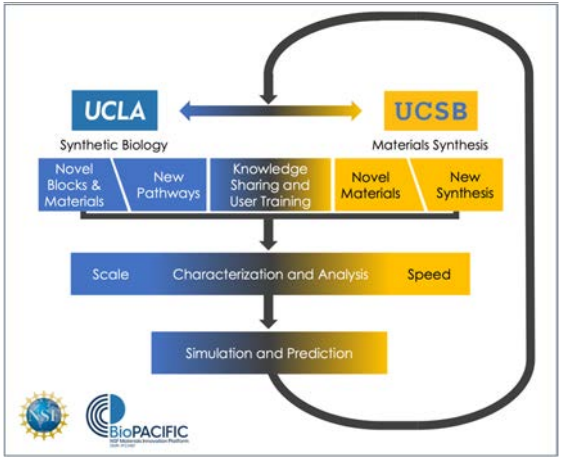


Figure. BioPACIFIC MIP Scope.

Technical Progress

BioPACIFIC MIP in-house research programs are organized into four MGI-based Synergistic Experimental Thrusts (SETs).

- SET 1 – Bioderived Materials – Engineering biological systems to discover, optimize and produce new monomers and to develop and apply new biological polymerization.
- SET 2 – Sequence-defined Materials – Development of sequence specific synthesis and the application of cryoEM and MicroED and multi-scale modeling to guide discovery.
- SET 3 – Functional Biomimics – Mimicking the functionality of nature to discover and produce biomaterials with superior properties and performance.
- SET 4 – Degradation-optimized Materials – Developing new bio-derived monomers and high-throughput methods to monitor polymer degradation.

Future Plans

Accelerating materials research by integrating high-throughput synthesis, robotics, characterization techniques, and artificial intelligence is increasingly recognized as the way of the future. Going forward, more research initiatives seeking to automate their discovery workflows will be supported and developed. This creates enormous potential to accelerate discovery by increasing safety, efficiency, reproducibility, and design space while also reducing error and research downtime. Given this expanding landscape, it will be important to overcome several challenges and capture the full value of these investments. These include: (1) creating and maintaining unified,

high-quality, open-access datasets; (2) building efficient models that connect the structure-property relationships required for inverse design; (3) lowering the cost of high-throughput robotics to expand access; (4) training next-generation leaders in high-throughput experimentation; and (5) developing remote capabilities to improve access.

Broader Impacts and Workforce Development

The BioPACIFIC MIP is dedicated to providing a clear connection between researcher training and their career aspirations to empower trainees from diverse backgrounds to become the next generation workforce. The BioPACIFIC MIP provides undergraduate and graduate students, postdoctoral fellows, and participating users training in how to set up and analyze experiments that leverage automation and large data sets, summer school workshops, industry networking opportunities, and transferable professional skill training.

Knowledge Sharing

A central pillar of the BioPACIFIC MIP is fostering new modalities of research and education, through sharing tools, codes, samples, data, and know-how. The BioPACIFIC MIP is managed and staffed with technical experts that impart knowledge by training researchers on the proper use of equipment and advise on how to properly set up, carry out, and analyze experiments, and moreover, expand and retain the BioPACIFIC MIP's institutional knowledge base through hands-on interaction with users. Sharing is not envisioned to be a one-way flow of information out of the BioPACIFIC MIP but rather a collaborative partnership between the BioPACIFIC MIP and its users.

Advancing Along the Materials Development Continuum and Partnerships to Translation

BioPACIFIC MIP provides an open, collaborative research ecosystem free from constraints that are common to topically-focused research centers and can thus accommodate biomaterials innovations stemming from any of its research pillars – synthetic biology, synthetic chemistry, additive manufacturing analytical techniques, or novel computation. This unique structure makes BioPACIFIC MIP an ideal organization to support translational research and train the next generation of innovators through a comprehensive program to develop and commercialize the most promising technologies within its research portfolio. However, accomplishing this goal is challenging and will require dedicated resources for technology development, funding for IP to lower the barrier to file patent applications, and better understanding from academia on the complete picture of technology adoption by industry.

Publications

1. M. Gu, Y. Luo, Y. He, M.E. Helgeson, M. T. Valentine MT *Uncertainty quantification and estimation in differential dynamic microscopy*, Physical Review E, **104**, 34610 (2021). https://doi.org/10.1103/PhysRevE.104.034610.
2. C, Choi, Y. Okayama, P. T. Morris, L. L. Robinson, M. Gerst, J. C. Speros, C. J. Hawker, J. Read de Alaniz, C. M. Bates *Digital Light Processing of Dynamic Bottlebrush Materials*, Advanced Functional Materials, **32**, 2200883 (2022). https://doi.org/10.1002/adfm.202200883.
3. J. Misa, J. M. Billingsley, K. Niwa, R. K. Yu, Y. Tang, *Engineered Production of Strictosidine and Analogues in Yeast*, ACS Synthetic Biology, **11**, 1639 (2022). https://doi.org/10.1021/acssynbio.2c00037.
4. L. L. Robinson, S. L. Self , A. D. Fusi, M. W. Bates, J. Read de Alaniz, C. J. Hawker, C. M. Bates, C. S. Sample, *Chemical and Mechanical Tunability of 3D-Printed Dynamic Covalent Networks Based on Boronate Esters*, ACS Macro Letters, **10**, 857 (2021). https://doi.org/10.1021/acsmacrolett.1c00257.
5. Y. Luo, M. Gu, C. E. R. Edwards, M. T. Valentine, M. E. Helgeson, *High-throughput microscopy to determine morphology, micro-rheology, and phase boundaries applied to phase separating coacervates*, Soft Matter, **18**, 3063 (2022). https://doi.org/10.1039/D1SM01763B.
6. R. C. Law, A. Lakhani, S. O'Keeffe, S. Erşan, J. O. Park, *Integrative metabolic flux analysis reveals an indispensable dimension of phenotypes*, Current Opinion in Biotechnology, **75**, 102701 (2022). https://doi.org/10.1016/j.copbio.2022.102701.
7. Y. Luo, M. Gu, C. E. R. Edwards, M. T. Valentine, M. E. Helgeson, *High-throughput microscopy to determine morphology, micro-rheology, and phase boundaries applied to phase separating coacervates*, Soft Matter, **18**, 3063 (2022). https://doi.org/10.1039/D1SM01763B.
8. S. J. Bailey, F. Stricker, E. Hopkins, M. Z. Wilson, J. Read de Alaniz, *Shining Light on Cyclopentadienone–Norbornadiene Diels–Alder Adducts to Enable Photoinduced Click Chemistry with Cyclopentadiene*, ACS Applied Materials & Interfaces, **13**, 35422 (2021). https://doi.org/10.1021/acsami.1c08670.

Lifetime Sample Tracking Integrated into a Materials Innovation Platform for 2D Materials Synthesis

Lead Investigator: Joan Redwing (Director of Two-dimensional Crystal Consortium, jmr31@psu.edu), Wesley Reinhart (reinhart@psu.edu), Kevin Dressler (kxd13@psu.edu), Vincent Crespi (vhc2@psu.edu) and lead LiST developer Konrad Hilse (rkh24@psu.edu)
Participating Institutions: Penn State University
Website: www.mri.psu.edu/mip
Keywords: Materials Innovation Platform, Knowledge Sharing, Machine Learning, Materials Synthesis

Project Scope

To tightly integrate the broad inorganic thin-film and bulk synthetic capabilities and tools of the Two-Dimensional Crystal Consortium (2DCC) with a comprehensive data acquisition and curation tool to capture synthetic and property data and metadata of sufficient breadth and depth to enable machine learning enabled investigations that enhance materials discovery and application-enabling materials quality.

Relevance to MGI

Although theory has seen predictive synthetic success in systems amenable to near-equilibrium synthetic techniques, the broader impact of 2D materials is predicated on scalable synthesis of high-quality thin films and heterostructures on technologically relevant substrates using growth techniques whose outcomes can be dominated by non-equilibrium events rare in space and time. Data-driven material design often centers on equilibrium, whereas a comprehensive treatment of kinetic intermediates, substrate interactions, defect geometries and key impurities for a single material system can rival an equilibrium database across the entire periodic table. This information richness of growth is a re-expression of the fact that key properties of materials are typically dominated by how they were grown. To make transformative advances under kinetically dominated growth conditions thus requires a close integration of theory, data, and experiment that captures comprehensive growth information that can be correlated to theory, simulation, and characterization data, while advancing tools capable of atomistic simulation across time and spatial scales relevant to experiment. 2DCC has developed a Lifetime Sample Tracking (LiST) database which is being extended to knowledge graph capabilities and coupled to both experiment and reactive force field simulation tools that cover conditions relevant to 2D materials growth.

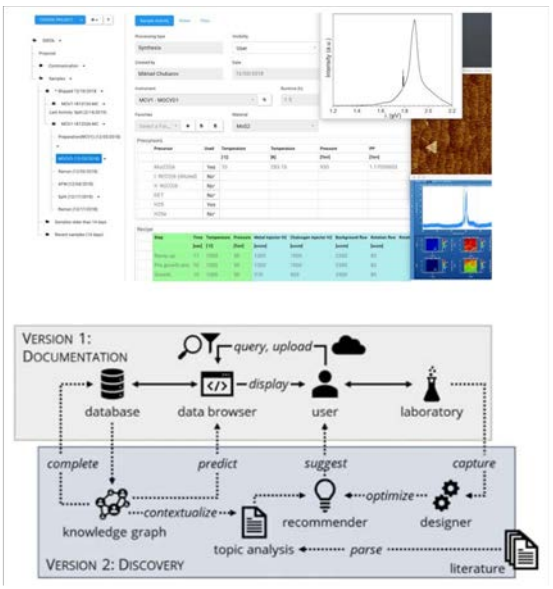


Figure 1. LiST 1.0 captures and curates growth protocol and characterization data and metadata on samples produced in the 2DCC, both user and in-house. The LiST 2.0 knowledge graph will transform the LiST repository from a documentation tool to a discovery tool.

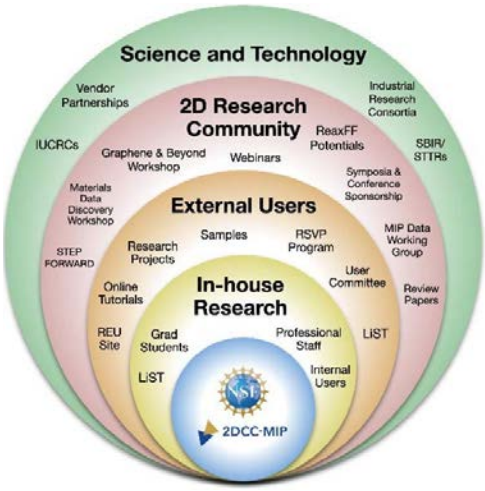
Technical Progress

To meet this challenge requires comprehensive datasets that embody the history of each sample: how it was grown; its defects, impurities, etc. This metadata describing the context of a target property measurement can be larger than the data comprising the measurement itself, while also being key input for knowledge-graph-based analyses. Over the past ~4 years the 2DCC has designed, developed and refined LiST (**Li**fetime **S**ample **T**racking), a database whose interface is depicted in the Figure 1. LiST has customized interfaces and automated data ingestion and has captured growth and characterization data on thousands of samples to date. This sets the stage for the implementation of knowledge-graph capabilities into a future LiST 2.0, an effort that has launched recently with

the hiring of personnel in 2022. Note that the "tool" being developed here is essentially the entirety of the 2DCC, as the existing LiST 1.0 is tightly integrated into experimental synthetic workflows across the 2DCC, which operates as an integrated platform.

Future Plans

Whereas LiST 1.0 uses a traditional database whose consistent schema makes the materials data straightforward to query, synthesize, and maintain since users or programs interacting with this structured database know what to expect, the addition of knowledge graph capabilities through integration with an ontology describing the relationships between entities will enable users of LiST 2.0 to infer relationships not explicitly provided by the initial data design. The resulting context-aware, graph-structured data is intended to enable more natural interaction such as question-and-answer and recommendation systems. The LiST 2.0 knowledge graph is also planned to be able to complete certain forms of sparsity in data by finding alternate paths between data using pattern recognition. As the knowledge graph accrues implicit and explicit relationships, data browser software will be designed to make informed guesses on the context and intent of user interactions in terms of both optimizing synthesis and interpreting characterization results, following methods of graph completion, where context described by a sub-graph is compared against statistical models of the overall knowledge graph to predict the most relevant entities. Such algorithms are intended to automatically adapt the software interface to common workflows and to also enable discovery by informing users of hidden relationships that might be unknown to human experts.



Broader Impacts and Workforce Development

The 2DCC has a multipronged strategy towards workforce development that includes a comprehensive proposal system for user-driven materials synthesis research projects, an annual Graphene and Beyond Workshop for the broader community, extended in-depth training of on-site Resident Scholar Visitors Program (RSVP) fellows across experimental and theory capabilities, and research experiences for undergraduates.

Knowledge Sharing

The LiST 1.0 database has been configured to enable creation of "data packages" consisting of specified subsets of synthesis, characterization and simulation data that can be shared as a DOI (envisioned for association to user publications, for example). Efforts are ongoing to more fully align the metadata formats in this package with community standards, as are efforts to release an API for direct remote access to the LiST data store. As ease-of-use and facility of data ingestion across a range of systems mature, LiST is planned to be offered to curate selected third-party data, likely starting with current and prior users who have storage, curation and analysis needs that extend beyond the data generated directly at the 2DCC. As development proceeds, List 2.0 will also serve as a data source for users interested in performing their own data-driven materials research using our data (with safeguards for data integrity and confidentiality). The 2DCC also receives and fulfills User Data Requests, as the most recent category of user proposals. These proposals, for example, request growth and characterization for thin-film TMDs grown by MOCVD, including Raman/PL, AFM, XRD, and SEM data and detailed growth protocols. As LiST grows in capability and size, the complexity and depth of user data requests that we can fulfill will expand.

Advancing Along the Materials Development Continuum and Partnerships to Translation

Interest has been expressed in the LiST data curation tool by diverse users across the community and there may be prospects for sharing of the underlying technology to commercialization partners in the longer term. The existing LiST 1.0 provides a foundational data-source towards accelerated materials discovery and development that makes use of the knowledge-graph capabilities of List 2.0.

Publications and References

As a user facility, the 2DCC has a broad range of publications that is difficult to summarize in a truncated list. A comprehensive listing organized by platform science drivers is available at:

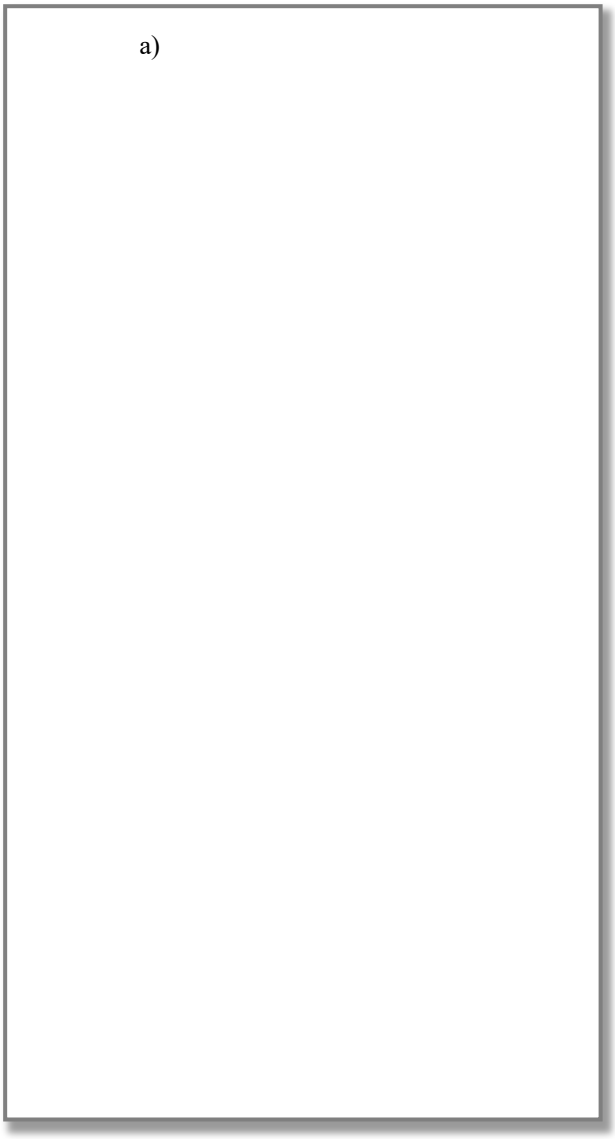
http://www.mri.psu.edu/2d-crystal-consortium/research/publications

Three-Dimensional Mechanical Metamaterials from Disordered Networks by Global Node Optimization

Lead Investigator: Marcos A. Reyes-Martinez, Meng Shen, Edwin P. Chan, Christopher L. Soles, Sidney R. Nagel, Heinrich M. Jaeger, Juan J. de Pablo
Participating Institutions: National Institute of Standards and Technology (NIST), University of Chicago
Website: None
Keywords: Auxetic Metamaterials, Disordered Networks, Impact Mitigation, Inverse design

Project Scope

Current strategies for the design of auxetic metamaterials are largely Edisonian and it is challenging to realize isotropic structures with precise control of their mechanical behavior. Here, we introduce a global node optimization strategy to design three-dimensional, isotropic auxetic metamaterials from disordered networks. After specifying the target Poisson's ratio for a disordered structure, an inverse design algorithm is used to adjust the positions of all nodes in a disordered network structure until the desired mechanical response is achieved. The design algorithm is validated through experimental realizations of the computationally-designed structures.



Relevance to MGI

Control over the Poisson's ratios is of interest for a variety of applications including impact mitigation, membrane separations and biomedical engineering. The design and fabrication of disordered network mechanical metamaterials brings together theory, computational design of structures with the desired mechanical properties, fabrication, characterization and validation.

The performance of disordered network mechanical metamaterials is intrinsically related to the local elastic properties of the network, namely local angle bending (k_θ) and bond stretching resistance (k_b). Data from compression experiments on disordered networks samples, including tracking of specific nodes during compression, is crucial in understanding the role of k_θ and k_b in the global mechanical behavior of the network. After initial testing of 3D printed sample of the computationally designed structures, it was discovered that fine control of the angle bending resistance in samples is needed. This experimental data feeds our simulation efforts to make our model more precise. The dual material printing strategy adopted in this presentation corresponds to a low k_θ / k_b ratio and is a direct result from the loop between theory, simulations, fabrication adopted in this project.

Technical Progress

Reid *et al.*[1] showed the effectiveness of local bond pruning to design and fabricate two-dimensional auxetic materials. While they achieved good agreement between simulation and experimental measurements, the process was not as effective in achieving auxetic behavior in three-dimensional networks. Additionally, given that the pruning method relies on removing the bonds that contribute the most to the bulk modulus, it results in a reduction

of the density of the overall structure. The resulting density reduction from pruning is advantageous for mitigating the magnitude of the transmitted stress upon impact[2] but the reduced mechanical integrity can be detrimental for other applications.

We demonstrate that the proposed algorithm allows independent control of the shear and bulk moduli, while preserving the density and connectivity of the networks. When the angle bending stiffness in the network is kept low, it is possible to realize Poisson's ratios as low as −0.6. During the optimization, the bulk modulus of these networks decreases by almost two decades, but the shear modulus remains largely unaltered. At the same time, the overall density of the structure remains nearly constant. The materials designed in this manner are fabricated by dual-material 3D-printing, and are found to exhibit the mechanical responses that were originally encoded in the computational design.

Future Plans

Figure 1b shows how three different values of k_θ/k_b affect the ultimate Poisson's ratio achievable through the global Node Optimization method. While the present work focuses on experimentally demonstrating the capabilities of our inverse design method for a specific k_θ/k_b value, a wider range remains to be explored. In addition, and specifically for impact mitigation applications, bond to bond interactions need to be added to the model and is an ongoing area of work. This is important to capture accurate predictions of mechanical behavior beyond the linear elastic regime of deformation.

The approach presented here for the design of three-dimensional mechanical metamaterials from disordered networks leads to facile development of isotropic, auxetic structures. Nevertheless, the level of control achieved by simultaneous adjustment of all node positions could be useful for the development of materials that exhibit multiple designer properties including mechanical, optical or thermal characteristics. These are fertile areas of research that should be explored.

Broader Impacts and Workforce Development

One goal of MGI is to train and equip the next generation workforce in materials research and the organization of summer/winter schools, and (for the NSF) education of students and the general public. How is the MGI philosophy and project outcomes being conveyed to, and impacting, those outside the immediate community? For existing projects, please describe how this is being accomplished. For new projects, describe plans to achieve this objective. Please limit this section to 250 words.

Data Management and Open Access

An objective of MGI is to make digital outputs publicly accessible and useful to the community to harness the power of materials data and unify the materials innovation infrastructure. How are digital data, software, and program outputs being integrated into the community? How are data from experiment and computation being combined into a searchable materials data infrastructure? Please describe how this project is addressing this objective. How has the project provided the community with access to its outputs? Please identify URLs and DOIs (counted as 1 word each) where data and code are available. Please limit this section to 150 words.

Advancing Along the Materials Development Continuum and Partnerships to Translation

The inverse design approach presented here has clear applications in the design of new metamaterials in the area of impact mitigation for protective equipment and shoe soles, among others. A U.S. Patent Continuation Application (# 17/731,998) has been filed under the title "INVERSE DESIGN MACHINE AND MAKING A DESIGNER IMPACT-MITIGATING ARCHITECTURED ISOTROPIC STRUCTURE". This will facilitate collaborations and has already captured the attention of industrial entities

Publications and References

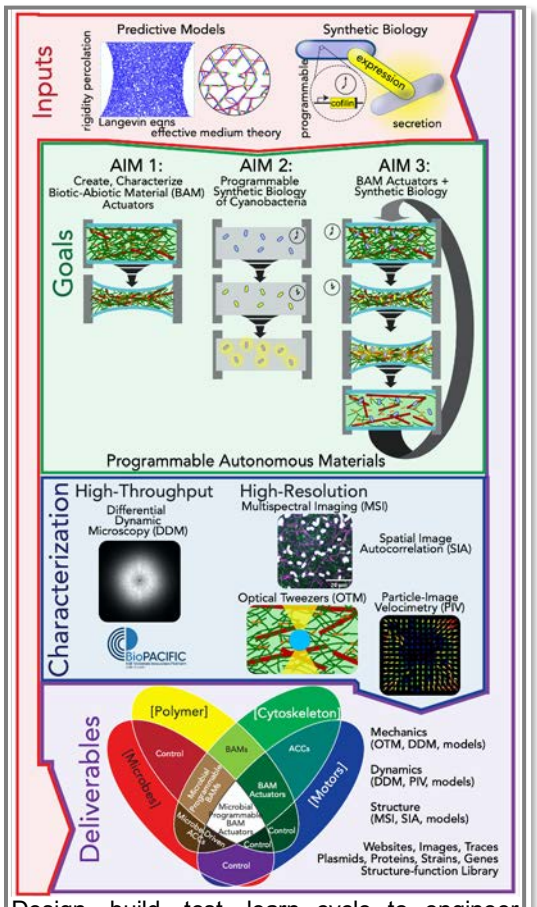
1. Reid, D.R., Pashine, N., Wozniak, J.M., Jaeger, H.M., Liu, A.J., Nagel, S.R., de Pablo, J.J. "Auxetic metamaterials from disordered networks", PNAS, 2018, E1384-E1390, 115(7).
2. Reyes-Martinez, M.A., Chan, E.P., Soles, C.L., Han, E., Murphy, K.A., Jaeger, H.M., Reid, D.R., de Pablo, J.J. "Tuning the mechanical impedance of disordered networks for impact mitigation", *Soft Matter*, 2022, **18**, 2039

Collaborative Research: DMREF: Living biotic-abiotic materials with temporally programmable actuation

Lead Investigator: Rae Robertson-Anderson, randerson@sandiego.edu
Participating Institutions: University of San Diego, Rochester Institute of Technology, Syracuse University, University of Chicago, University of California – Santa Barbara
Website: none
Keywords: Active Matter, Synthetic Biology, Biomaterials, Cytoskeleton, Programmable, Soft Matter

Project Scope: We will develop foundational technologies, predictive models, and formulation libraries to pioneer a new class of autonomous reconfigurable materials with self-generated spatiotemporal control. Guided by multi-scale modeling, and leveraging advances in synthetic biology and active matter, we will integrate biological circuits into biotic-abiotic composites to engineer materials that self-actuate programmable work cycles. Our design paradigm couples abiotic hydrogels to living cytoskeleton layers infused with bacteria that will secrete cytoskeleton-modifying proteins on a programmable schedule. Our proof-of-concept will be a gap-closing micro-actuator that photo-responsively closes and autonomously re-opens at times and locations programmed into the cells.

Relevance to MGI: We will capitalize on our team's unique expertise and strong collaborative track record to engineer living biotic-abiotic materials (BAMs) that unite microbes, proteins, biopolymers, and hydrogels, for in situ bioproduction and self-regulation to autonomously drive actuation and perform work. The proof-of-concept photo-responsive gap-closing actuator we will create will serve as a powerful testbed material that can be leveraged by our team and the broader MGI community to manufacture and deploy autonomous biotic-abiotic materials. We will use iterative design, build, test, learn (DBTL) cycles to accelerate discovery–linking theory, fabrication, computation, and characterization to establish a broad phase space of structure-mechanics-function relationships of our bio-fueled materials platform. We will partner with NSF BioPACIFIC Material Innovation Platform (MIP) to develop high-throughput screening schema and formulation libraries to rapidly optimize and publicly share our material blueprints and technologies. Our strong team work and dissemination, along with robust training opportunities for the next generation of MGI researchers–from high school students to postdocs–will empower the MGI community to capitalize on our discoveries to engineer programmable dynamic prosthetics, self-sensing PPE, self-regulated micro-robotic devices, and autonomous adaptive optics.



Design, build, test, learn cycle to engineer autonomous programmable biotic-abiotic materials. We will use multi-scale modeling and synthetic biology to drive the creation of self-actuating materials that we will robustly characterize using precision, yet high-throughput methods, and we will broadly disseminate design rules through a multidimensional property-function database.

Technical Progress: Progress under each Aim is described below.
Aim 1. Curate and disseminate a comprehensive platform to design, build, test, and learn from active biotic-abiotic material (BAM) actuators optimized for stiffness, strength, and contractility. We have started building a library of active cytoskeleton composites (ACCs) that are tunable, modular and reconfigurable; and developing a robust characterization suite to map formulation to structure and mechanics. We have incorporated kinesin and myosin into actin-microtubule composites to drive non-equilibrium restructuring and dynamics including contractility, acceleration, phase separation, bundling and flow. We have developed models to predict mechanical properties, and Fourier analysis methods to characterize spatiotemporally varying ACC dynamics and structure. Finally, we have incorporated spherical hydrogels into ACCs to begin tuning abiotic-biotic

interactions and couplings in BAMs.

Aim 2. Develop cyanobacteria as a programmable synthetic biology platform able to deliver molecules according to a user-defined schedule. We have identified an initial target molecule to engineer into cyanobacteria. The small peptide, thymosin, depolymerizes actin upon secretion and is small enough (<10 kD) to allow for efficient secretion.

Aim 3. Integrate biotic-abiotic actuators and living bioproduction elements to create programmable autonomous materials. We have purchased equipment needed to grow cyanobacteria cells within ACCs.

Future Plans

1. We will curate and disseminate a comprehensive platform to design, build, test, and learn from active BAM actuators optimized for stiffness, strength, and contractility. We will continue to build the formulation-mechanics library of our ACCs and work towards optimal coupling to abiotic hydrogels. We will incorporate high-throughput algorithms and approaches, fueled by predictive modeling, into our characterization suite to accelerate materials design and formulation database development. We will test and iteratively optimize predictions from multi-scale mechanistic models and new active matter theories. We will partner with the BioPACIFIC MIP to disseminate design rules via a searchable database of structure-property relationships.

2. We will develop cyanobacteria as a programmable synthetic biology platform to deliver molecules within materials according to a user-defined schedule. We will pioneer foundational advances in the synthetic biology of cyanobacteria to enable living cells to generate programmable waves of gene expression and secretion of cytoskeleton-modifying proteins, driven by their circadian clock, to empower living biotic-abiotic materials.

3. We will integrate biotic-abiotic actuators and living bioproduction elements to create programmable autonomous materials. Employing the design strategies and tools developed in Aims 1 and 2, we will engineer materials that can execute designed yet autonomous actions on a programmable schedule through in situ production of material-modifying proteins. We will incorporate bacteria with genetically-encoded timing circuits into BAM actuators (to program ACC disassembly via timed secretion of the depolymerization factors (e.g., thymosin), providing a programmable off-switch to BAM actuators. Through iterative DBTL cycles, and leveraging multiple engineered cells with different oscillator frequencies, we will construct complex temporal programs of self-actuated material restructuring and force-generation and decode and encode design rules for programmable living materials.

Broader Impacts and Workforce Development: We have engaged a diverse mix of undergraduates, postbaccalaureates, graduate students, and postdocs in this project, with a focus on broadening participation of URM, women and first-generation researchers, providing rich professional development opportunities including virtual and in-person project meetings and presentations at national conferences. The undergraduate summer research exchange program that we developed, in which a cohort of peer-mentored students across the five campuses collaborate to pursue interdisciplinary materials research also includes professional development programs to arm students with skills to lead a world-class materials workforce.

Data Management and Open Access: Dissemination and accessibility of material designs and characterization algorithms is a major goal of Aim 1. We will develop the shared database this year. Characterization code is currently available via GitHub and results disseminated via conference presentations, publications and PI websites.

Advancing Along the Materials Development Continuum and Partnerships to Translation: Our materials platform will result in novel technologies and predictive design paradigms for temporally-encoded autonomous materials with applications in self-propulsive materials, programmable prosthetics, micro-robotics, and self-healing infrastructure. The synthetic biology tools, high-throughput screening schema, and formulation-property database we will develop and disseminate via the BioPACIFIC MIP will transform future studies of active soft materials, enable use of engineered cyanobacteria for material modulation, and provide critical inputs to machine learning algorithms for material design. We will partner with local industry and foundation sponsors to host annual symposia where we share our results with the academic and industrial active matter and synthetic biology communities.

Publications and References

1. DH Achiriloaie, CJ Currie, J Michel, KA Lindsay, G Lee, M Sasanspour, MJ Rust, JY Sheung, M Das, JL Ross, RJ McGorty, RM Robertson-Anderson. Kinesin and Myosin Motors Compete to Drive Rich Multi-Phase Dynamics in Programmable Cytoskeletal Composites. Under review at *Nature Communications* (2022)
2. HN Verwei, G Lee, G Leech, II Petitjean, GH Koenderink, RM Robertson-Anderson, R McGorty. Quantifying Cytoskeleton Dynamics Using Differential Dynamic Microscopy. *JoVE*, in press (2022).

Uncovering Mechanisms of Grain Boundary Migration in Polycrystals for Predictive Simulations of Grain Growth

Lead Investigator: Gregory S. Rohrer, gr20@andrew.cmu.edu
Participating Institutions: Carnegie Mellon University
Website: none.
Keywords: Microstructure evolution, grain boundary properties, X-ray microscopy

Project Scope

Grain boundary network structures are determined by grain boundary migration when the material is processed at high temperature. Therefore, controlling materials properties is predicated on understanding and controlling grain boundary migration. The two prevailing models for grain boundary migration are diffusive migration and defect-controlled migration. To accurately simulate microstructure evolution, it is necessary to know if, and under what conditions, these two models provide an accurate description of grain boundary migration. X-ray microscopy will be used to measure the structure of the grain boundary networks in ferritic iron, nickel, and strontium titanate, and how they evolve with time. The results will be compared to atomistic simulations of grain boundary migration and to predictions from the two theories for grain boundary migration to determine which one provides a superior description of the temperature dependence. Understanding the mechanism of interface migration will make it possible to better predict microstructure evolution, a necessary step in accelerating the development of polycrystalline materials.

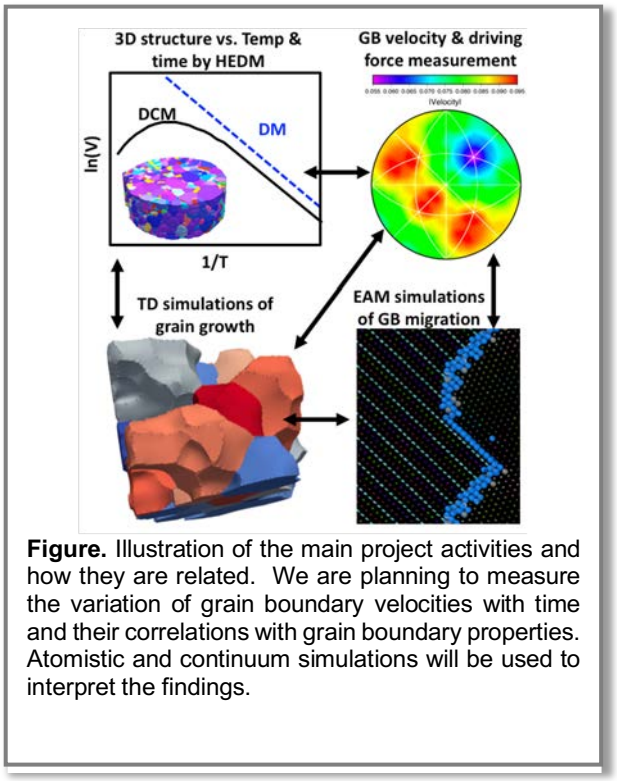


Figure. Illustration of the main project activities and how they are related. We are planning to measure the variation of grain boundary velocities with time and their correlations with grain boundary properties. Atomistic and continuum simulations will be used to interpret the findings.

Relevance to MGI

It has been common practice for designers to incorporate metallic and ceramic components into structures and devices assuming uniform and fixed bulk properties. However, a material's properties are sensitive to the types of interfaces within the material and this is a result of materials processing. To make materials part of the design process, it is necessary not only to know the range properties available, but also how to achieve them though processing. This project's aim of developing accurate predictive simulations for microstructure evolution will accelerate the incorporation of polycrystalline components by defining processing conditions to achieve specific microstructures and properties. This is integral to the broader goals of the Materials Genome Initiative.

Technical Progress

In the first year, we have focused on analyzing and simulating HEDM data from grain growth experiments in Ni, α-Fe, and $SrTiO_3$. In 2021, we made the surprising discovery that in Ni, the grain boundary migration velocity was uncorrelated to the grain boundary curvature.[1] This was surprising because it is commonly accepted that the velocity and curvature are positively correlated. Further analysis of the data from α-Fe and $SrTiO_3$ confirmed the findings in these disparate materials. To understand the origin of this result, we are exploring the hypothesis that grain boundaries move in such a way as to eliminate higher energy boundaries preferentially, not only reducing the overall grain boundary area (as would happen with curvature driven growth), but also replacing existing grain

boundary area with lower energy boundaries. It is possible that this unrecognized driving force is responsible for moving grain boundaries in directions not correlated with their curvature. We have developed a procedure to assign anisotropic energies to the boundaries based on experimental and computed findings. Using these assignments for Ni, we have found that as grain growth proceeds, the grain boundary network is made up of an increased fraction of lower energy interfaces, consistent with our hypothesis. We are attempting to repeat this for the α-Fe data and to show that simulations[3,5] with the same assigned properties lower the energy by the same mechanism.

Future Plans

Our plans for the future include new experimental campaigns to collect temperature dependent 4D grain growth data at the synchrotron (APS) in later 2022 and 2023. These measurements are necessary to determine if grain boundary migration occurs by a diffusive or defect mediated mechanism. We will also compare simulations of grain growth with accurate grain boundary energies to determine if this is sufficient to explain the experimental observations.

Broader Impacts and Workforce Development

We have restarted (post pandemic) the 3D microstructure workshop, which is scheduled to take place August 17-19, 2022. The workshop consisted of three types of presentations. Roughly one third of the presentations are about state of the art 3D materials research. Another third are tutorials illustrating the capabilities and uses of Dream.3D. The final third of the workshop consists of practical sessions in which participants work on individual projects with the guidance of the speakers. About 50 participants each year learn about 3D microstructures investigation tools, including those being developed as part of this project. The professionals who attend are from industry and government labs. The graduate students and post docs who attended come from universities in the United States and Europe. See: http://mimp.materials.cmu.edu/rohrer/3DMS_workshop_22/

Data Management and Open Access

We have made the data, analysis code, and simulation code used in this project available through internet repositories, listed below:
Data: http://mimp.materials.cmu.edu/~gr20/Grain_Boundary_Data_Archive/
Analysis code: https://github.com/gr20cmu/gbXstallography
Examples of using analysis code and data: http://mimp.materials.cmu.edu/rohrer/gbXstallography/
Grain boundary energy reconstruction code: https://github.com/Yufeng-shen/TJ2GBE
Simulation code: https://github.com/JadeXiaoyaoPeng/GrainGrowth_TD_iso
Simulation code: https://github.com/Kiana-Naghibzadeh/GrainGrowth_TD_aniso

Advancing Along the Materials Development Continuum and Partnerships to Translation

We have collaborated with NIST on developing accelerated EBSD data collection strategies. This led to some follow on funding under the umbrella of the AFRL center for data-driven discovery of optimized multifunctional materials systems. This development project is ongoing and has not yet been deployed for use.

Publications and References[1-5]

[1]A. Bhattacharya, Y. F. Shen, C. M. Hefferan, S. F. Li, J. Lind, R. M. Suter, C. E. Krill, and G. S. Rohrer, "*Grain boundary velocity and curvature are not correlated in Ni polycrystals*," Science, **374** 189 (2021). https://doi.org/10.1126/science.abj3210

[2]J. Chua, V. Agrawal, T. Breitzman, G. Gazonas, and K. Dayal, "*Phase-field modeling and peridynamics for defect dynamics, and an augmented phase-field model with viscous stresses*," Journal of the Mechanics and Physics of Solids, **159** 104716 (2022). https://doi.org/10.1016/j.jmps.2021.104716

[3]S. Florez, K. Alvarado, B. Murgas, N. Bozzolo, D. Chatain, C. E. Krill, M. Y. Wang, G. S. Rohrer, and M. Bernacki, "*Statistical behaviour of interfaces subjected to curvature flow and torque effects applied to microstructural evolution*s," Acta Materialia, **222** 117459 (2022). https://doi.org/10.1016/j.actamat.2021.117459

[4]S. K. Naghibzadeh, N. Walkington, and K. Dayal, "*Accretion and ablation in deformable solids with an Eulerian description: examples using the method of characteristics*," Mathematics and Mechanics of Solids, **27** [6] 989 (2022). https://doi.org/10.1177/10812865211054573

[5]X. Y. Peng, A. Bhattacharya, S. K. Naghibzadeh, D. Kinderlehrer, R. Suter, K. Dayal, and G. S. Rohrer, "*Comparison of simulated and measured grain volume changes during grain growth*," Physical Review Materials, **6** 033402 (2022). https://doi.org/10.1103/PhysRevMaterials.6.033402

GlycoMIP, an NSF Materials Innovation Platform

Lead Investigator: Maren Roman, maren.roman@vt.edu (Director of GlycoMIP)
Co-Principal Investigator: Robert J. Woods, rwoods@ccrc.uga.edu
Participating Institutions: Virginia Tech, University of Georgia
Website: https://glycomip.org
Keywords: glycomaterials; carbohydrates; rational molecular design; Materials Innovation Platform.

Project Scope: Anchored at Virginia Tech and the University of Georgia, GlycoMIP, an NSF Materials Innovation Platform, accelerates discovery in glycomaterials science and technology through a unique national user facility and leading-edge in-house research, and advances the implementation of the Materials Genome Initiative paradigm shift in glycomaterials development. Strengthened by world-leading expertise at Brandeis University, Rensselaer Polytechnic Institute, and the University of North Carolina at Chapel Hill, the in-house research of GlycoMIP addresses the grand challenge of rational molecular design of glycomaterials through efficient convergence of physical sciences, engineering, computation, and life sciences to enable computer-guided design of primary structure and molecular architecture.

Figure 1. The GlycoMIP user facility offers users access to two automated glycan assembly systems.

Relevance to MGI: GlycoMIP's In-House Research program comprises four closed, iterative feedback loops, each integrating experiment, computation, and theory: Loop 1 integrates chemical synthesis, automated sensing, and mesoscale modeling for the development of novel glycomaterials and high throughput characterization techniques; Loop 2 integrates enzyme-based glycan synthesis with computationally supported vibrational optical activity spectroscopy to advance solution-state molecular analysis and the automation of biocatalytic methods; Loop 3 integrates measurement and computational prediction of molecular interactions and binding properties of glycomaterials to determine structure–property relationships; and Loop 4 applies machine learning methods to process and spectroscopy data for the development of predictive glycan synthesis models and the automation of data analysis. Synergies between the Loops have also fostered inter-Loop collaborations. The long-term goal of GlycoMIP's In-House Research program is to enable and provide tools for the rational design of glycomaterials with specific, desired molecular and material properties as well as the chemical or chemoenzymatic synthesis of these glycomaterials.

Technical Progress: Glycomaterials are materials that contain or consist of carbohydrates. Because of their complex molecular structures, glycomaterials are more difficult to design, create, and characterize than other (bio)polymers. Open to external researchers from across the nation, the GlycoMIP user facility at Virginia Tech and the University of Georgia provides access to unique experimental and computational tools for glycomaterials synthesis, characterization, and modeling and facilitates Materials Genome Initiative approaches to materials research and development. These include automated glycan synthesizers for de novo glycan synthesis, a mass spectrometry (MS) imaging system for in-situ glycan analysis, high-resolution, multi-stage MS systems for in-depth glycan structural analysis, surface plasmon resonance and biolayer interferometry systems for high-throughput quantitation of binding events, two vibrational optical activity spectrometers for analysis of solution-state conformations, and a rheometer for the characterization of glycan-based solutions and gels.

Future Plans: GlycoMIP periodically offers online and in-person educational, training, and outreach events,

including town halls, open houses, summer schools, short courses, Scientists-on-Screen events, and undergraduate research opportunities. Upcoming events will be announced at https://glycomip.org/events/.

Broader Impacts and Workforce Development: GlycoMIP employs multiple approaches to train the next generation of glycomaterials researchers in accelerated glycomaterials development. Undergraduate students are trained through remote and in-person lectures and participation in GlycoMIP research. GlycoMIP-affiliated graduate students are integrated into multidisciplinary research teams and have responsibilities within the GlycoMIP user facility. GlycoMIP-funded postdocs are given leadership roles within the in-house research, education and outreach, and user programs, as well as the cyber resources development activities. Broadening the participation of underrepresented groups is an emphasis in all GlycoMIP activities. GlycoMIP offers grant programs to support user projects from non-R1 and HBCU/MSI institutions as well as audience-targeted education and outreach events.

Knowledge Sharing: GlycoMIP is currently creating a knowledge sharing platform, designed to facilitate the sharing of data, tools, and know-how both within the glycomaterials research community and among its users. Features of the platform include project pages, focused on individual research projects; several open-access databases that share glycomaterials-related data in a manner that ensures findability, accessibility, interoperability, and reusability; and online tools for data generation (*e.g.* computer simulation tools), data analysis (*e.g.* spectroscopy signal assignment tools), and data conversion (*e.g.* chemical notation or file format conversion tools).

Advancing Along the Materials Development Continuum and Partnerships to Translation: GlycoMIP is currently establishing an industry stakeholder group to engage in collaborative research and tech-transfer efforts.

Publications and References (https://glycomip.org/publications**)**

1. E. Stancanelli, W. Liu, R. Wander, J. Li, Z. Wang, K. Arnold, G. Su, A. Kanack, T. Q. Pham, V. Pagadala, A. Padmanabhan, Y. Xu, J. Liu, *Chemoenzymatic Synthesis of Homogeneous Heparan Sulfate and Chondroitin Sulfate Chimeras*, ACS Chemical Biology **17**, 1207 (2022). DOI: 10.1021/acschembio.2c00146
2. Y. Yu, L. Fu, P. He, K. Xia, S. Varghese, H. Wang, F. Zhang, J. Dordick, R. J. Linhardt, Chemobiocatalytic Synthesis of a Low-Molecular-Weight Heparin, ACS Chemical Biology **17**, 637 (2022). DOI: 10.1021/acschembio.1c00928
3. S. Y. Joshi, S. Singh, S. A. Deshmukh, *Coarse-grained molecular dynamics integrated with convolutional neural network for comparing shapes of temperature sensitive bottlebrushes*, npj Computational Materials **8**, 45 (2022). DOI: 10.1038/s41524-022-00725-7
4. F. Zhang, P. He, A. L. Rodrigues, W. Jeske, R. Tandon, J. T. Bates, M. A. Bierdeman, J. Fareed, J. Dordick, R. J. Linhardt, *Potential Anti-SARS-CoV-2 Activity of Pentosan Polysulfate and Mucopolysaccharide Polysulfate*, Pharmaceuticals 15, 258 (2022). DOI: 10.3390/ph15020258
5. E. Ahat, Y. Song, K. Xia, W. Reid, J. Li, S. Bui, F. Zhang, R. J. Linhardt, Y. Wang, *GRASP depletion-mediated Golgi fragmentation impairs glycosaminoglycan synthesis, sulfation, and secretion*, Cellular and Molecular Life Sciences 79(4) (2022) 199. DOI: 10.1007/s00018-022-04223-3
6. R. Wander, A. M. Kaminski, Z. Wang, E. Stancanelli, Y. Xu, V. Pagadala, J. Li, J. M. Krahn, T. Q. Pham, J. Liu, L. C. Pedersen, *Structural and Substrate Specificity Analysis of 3-O-Sulfotransferase Isoform 5 to Synthesize Heparan Sulfate,* ACS Catalysis **11**, 14956 (2021). DOI: 10.1021/acscatal.1c04520
7. Y. Song, P. He, A. L. Rodrigues, P. Datta, R. Tandon, J. T. Bates, M. A. Bierdeman, C. Chen, J. Dordick, F. Zhang, R. J. Linhardt, *Anti-SARS-CoV-2 Activity of Rhamnan Sulfate from Monostroma nitidum*, Marine Drugs **19**, 685 (2021). DOI: 10.3390/md19120685
8. T. G. DeHart, M. R. Kushelman, S. B. Hildreth, R. F. Helm, B. L. Jutras, *The unusual cell wall of the Lyme disease spirochaete Borrelia burgdorferi is shaped by a tick sugar*, Nature Microbiology **6**, 1583 (2021). DOI: 10.1038/s41564-021-01003-w
9. R. Dwivedi, P. Samanta, P. Sharma, F. Zhang, S. K. Mishra, P. Kucheryavy, S. B. Kim, A. O. Aderibigbe, R. J. Linhardt, R. Tandon, R. J. Doerksen, V. H. Pomin, *Structural and kinetic analyses of holothurian sulfated glycans suggest potential treatment for SARS-CoV-2 infection*, Journal of Biological Chemistry **297**, 101207 (2021). DOI: 10.1016/j.jbc.2021.101207
10. E. Mohammadi, S. Y. Joshi, S. A. Deshmukh, *A review of computational studies of bottlebrush polymers*, Computational Materials Science **199**, 110720 (2021). DOI: 10.1016/j.commatsci.2021.110720

DOE H-Mat Consortium

Poster Presenter: Neha Rustagi, neha.rustagi@ee.doe.gov
Lead Investigators: Chris San Marchi, cwsanma@sandia.gov; Kevin L. Simmons, kl.simmons@pnnl.gov
Participating Core Institutions: Sandia National Laboratories, Pacific Northwest National Laboratory, Savannah River National Laboratory, Oak Ridge National Laboratory, Argonne National Laboratory
Website: https://h-mat.org/
Keywords: energy, hydrogen storage, materials compatibility, polymer, metal, pipelines

Project Scope

The H-Mat consortium conducts materials compatibility R&D to increase the life, lower the cost, and understand the performance of polymers and metals in hydrogen service. H-Mat labs are currently working with over thirty partners in industry and academia on projects aiming to develop high strength steels compatible with hydrogen service, engineer polymers to improve their durability in hydrogen, develop accelerated test methods for metals in hydrogen, evaluate the performance of non-destructive evaluation technologies in pressure vessels, and understand the performance of conventional pipeline materials in hydrogen blends as part of DOE's HyBlend initiative.

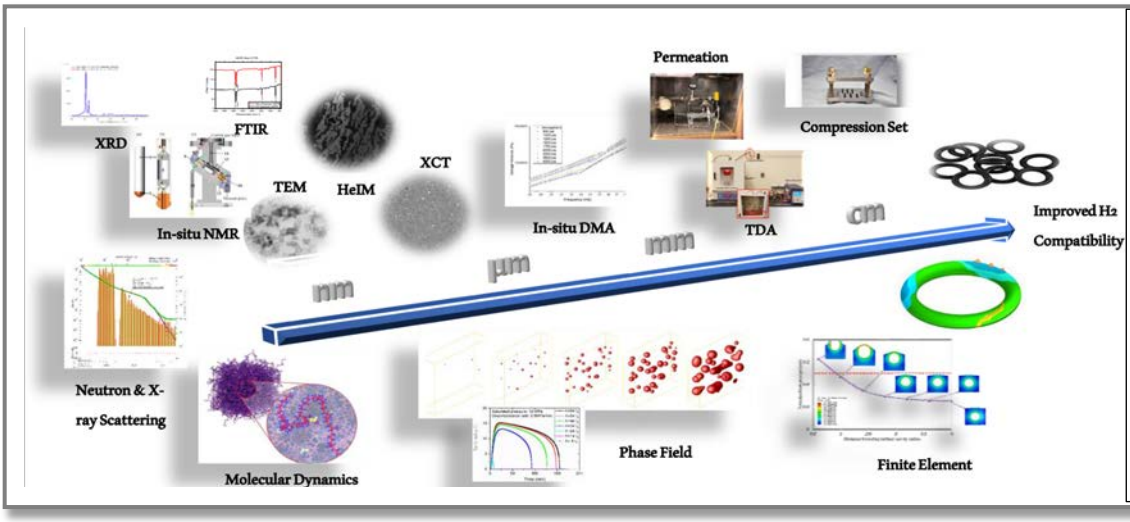


Figure 6. Examples of experimental and theoretical techniques used to compatibility of materials with hydrogen.

Relevance to MGI

As part of the DOE's Energy Materials Network (EMN), H-Mat aims to understand mechanisms of hydrogen-induced degradation of materials. This foundational understanding can inform science-based strategies to design the microstructure of materials for improved resistance to degradation in high-pressure hydrogen through integration of advanced computational materials science and innovative experimental capabilities across microstructural length scales. Ultimately, the development of novel material alloys and chemistries can reduce the cost and improve performance of components in hydrogen service.

Technical Progress

Prior accomplishments of H-Mat labs include:

- Developing an ASME code case describing the life of common pressure vessel materials in hydrogen, which reduced test requirements for industry and increasing the service life of pressure vessels by up to 3X.
- Developing test methods for polymers in hydrogen, which were ultimately adopted by CSA/ANSI and enable consistency in data generation across research stakeholders.
- Supporting industry-led teams in engineering components (e.g. dispensing hoses) for extended life in hydrogen.

- Developing an ASME code case that allows for the use of fiber reinforced polymer (FRP) piping in high-pressure hydrogen service, and informing the current ASME code for hydrogen piping and pipelines.

Future Plans

Additional ongoing work within H-Mat includes:

- Developing a tool that characterizes the risk of hydrogen blending in natural gas pipelines, as a function of materials used, vintage, and operating conditions.
- Developing a framework to account for crack nucleation in structural design, to enable 50% longer design life.
- Conducting systematic testing of polymer additive and filler materials to inform the development of novel materials that reduce swelling of seals in hydrogen by 50%.

Broader Impacts and Workforce Development

H-Mat is supporting over 30 stakeholders across universities and industry on materials compatibility R&D, through projects funded by solicitations issued by DOE and NSF, as well as through cooperative research and development agreements (CRADAs). The member labs also work closely with the international community, and recently established a Memorandum of Understanding with Kyushu University in Japan to enhance data sharing, enable researcher exchange, and coordinate future R&D.

Stakeholder interested in partnering with H-Mat are encouraged to contact the consortium at Contact Us - - H-Mat DataHUB for more information.

Data Management and Open Access

H-Mat hosts a public datahub that compiles the publications, presentations, and datasets for all projects under the H-Mat consortium. This Hub is available at: https://data.h-mat.org/

Advancing Along the Materials Development Continuum and Partnerships to Translation

R&D conducted by the H-Mat consortium informs codes and standards that ultimately guide the design of components used in hydrogen service. Recent examples include code cases within the ASME Boiler and Pressure Vessel Code, ASME Code for Hydrogen Piping and Pipelines, and the CSA/ANSI CHMC 2. H-Mat labs also work regularly with industry stakeholders to inform material design and component design, and to ensure that foundational R&D is targeting priority challenges within industry.

Publications and References

Publications from the H-Mat consortium can ben found on their website:
https://data.h-mat.org/dataset/?q=publications%7Cpresentations

Discovery and Design of Additives for Novel Polymer Morphology and Performance

Lead Investigator: Gregory C. Rutledge, Chemical Engineering, Massachusetts Institute of Technology, rutledge@mit.edu; Richard D. Braatz, Chemical Engineering, Massachusetts Institute of Technology, braatz@mit.edu; George Rodriguez, ExxonMobil Chemical Company, george.rodriguez@exxonmobil.com; Carlos Lopez-Barron, ExxonMobil Chemical Company, carlos.r.lopez-barron @exxonmobil.com
Website: http://web.mit.edu/rutledgegroup/, http://web.mit.edu/braatzgroup/
Keywords: nucleation kinetics, polyolefin, molecular dynamics, optimization, morphology

Project Scope

The vision of this project is the development of a paradigm for accelerated materials discovery and design through the synergistic application of molecular simulation, experimental evaluation, and advanced optimization. The instantiation of this paradigm is the discovery of additives that alter crystallization behavior of polyolefins, and thereby the semicrystalline morphology and properties of the material. The anticipated outcome of such effort is a systematic approach to new additives, and semicrystalline materials with exceptional thermal, mechanical, and optical properties.

Relevance to MGI

This project is a collaboration between academia and industry to create a new method of materials research, involving the application of (i) molecular level simulation, (ii) experimental validation, (iii) materials design and optimization, and (iv) industrial application, to identify essential relationships between molecular structure, morphology, and performance (see Fig. 1). This paradigm is realized here through the identification of additives that alter the nucleation of crystallites in polyethylenes (PE), and thereby their properties as well. Molecular dynamic (MD) simulations are used to conduct broad screenings of additive classes, which are validated experimentally in select cases by measurements of heterogeneous nucleation. Bayesian optimization (BO) is used to design and optimize candidates within and across additive classes. Best-in-class candidates are being singled out for subsequent synthesis, development, and characterization of morphology and properties within the laboratories of ExxonMobil, the industrial collaborator.

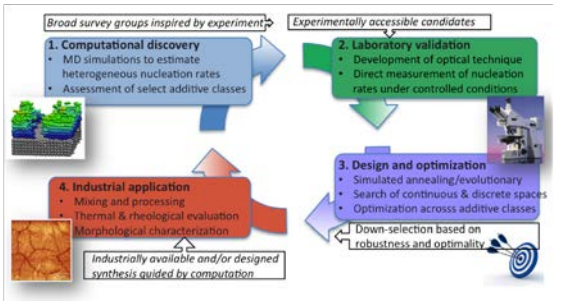


Figure 1. Schematic illustration of iterative process involving discovery through computational assessments, validation of computational indicators, design and optimization for rapid improvement, and industrial application.

Technical Progress

The measurement of heterogeneous nucleation rates with MD simulations of a PE oligomer (pentacontane, C50) on surfaces of diamond-like and graphene-like materials was demonstrated previously [1,2]. We used this approach to measure the nucleation rate of C50 with graphene as a function of crystallization temperature. For the case of graphene, no significant difference in nucleation rate was seen for the three temperatures investigated. However, graphene nucleates C50 so rapidly that temperature effects are difficult to quantify. Characterizing the temperature dependence of nucleation for other, less effective nucleating agents may still be valuable for comparison to experimental results.

Recently, a novel experimental approach was introduced to study heterogeneous nucleation in polypropylene [3]. Following this method, we created samples consisting of micro-droplets of PE in an immiscible polystyrene (PS) matrix. By melt blending various nucleating agents with PE prior to the final blending with PS, we measured crystallization rates for eight different nucleating agents. The dependence of crystallization rate on PE domain size, however, does not suggest purely heterogeneous nucleation measurement, but rather a combination of nucleation and growth rates. Quantifying nucleation rates separate from growth in our samples remains a challenge.

Classes of tetrahedrally coordinated (diamond-like) and hexagonally coordinated (graphene-like) nucleating agents were explored using BO to discover optimal additive candidates in a parametrized (molecular genome-like) space [4]. Silicon was found to be close to optimal within the tetrahedral class of agents, while a material similar to

graphene but somewhat softer (e.g. boron nitride) is predicted to be optimal within the hexagonal class. A nearest neighbor search of crystalline candidates from the open quantum materials database suggests a material like CdS may be near-optimal.

Future Plans

In order to take advantage of the capabilities of the BO algorithm, we are expanding the class of materials under investigation. The MD simulation approach used to study diamond-like monatomic materials is now being applied to exemplary materials within the broader class of diamond-like binary compounds. Initial data for heterogeneous nucleation rates of C50 on these materials will be determined by MD and fed to BO to predict new candidates for additional simulations. This feedback loop between simulations and BO will lead to a faster convergence to the optimal nucleating agent within this class of binary materials (i.e. fewer simulations compared to a grid search).

To close the loop between simulations and experiments, the experimental approach will be modified to measure nucleation rates separate from crystal growth. A method using optical microscopy is being developed which would allow nucleation and growth to be visually differentiated so that accurate nucleation rates can be measured. This experimental validation will also be used to further refine the models used for BO to improve recommendations and modeling techniques.

Broader Impacts and Workforce Development

This project serves society through the education and professional development of young engineers and scientists, and through the dissemination of results at national and international venues [5,6]. To date, the project has trained two postdoctoral associates, and three students (one PhD, one masters, and one undergraduate). The undergraduate was trained through MIT's Undergraduate Research Opportunities Program (UROP), which promotes the participation of underrepresented minority students. Through this project, this student had the opportunity to use advanced characterization tools, such as an electron microscope, that would not typically be part of the undergraduate experience.

Data Management and Open Access

Digital data from this project is retained in three forms during the lifetime of the project. All non-confidential data is made publicly available at the earliest reasonable time, with the exception of publications and presentations that are subject to a period of review by the industrial partner for removal of proprietary information. All publicly released data is made available upon request to the MIT PI. After the conclusion of the program, all data will be retained on MIT's archive DSpace@MIT (http://libraries.mit.edu/dspace-mit/) for the maximum period, which is currently permanent. Source code used in publications [4] is available on GitHub.

Advancing Along the Materials Development Continuum and Partnerships to Translation

At present, the industrial design and selection of additives are still largely a matter of empiricism. However, by using experimental observations to guide computational investigations of new additives, and by using the results of computational studies to inform an optimization strategy, subsequent experimental efforts can be focused on the most promising materials. Bayesian optimization approaches with noise are efficient, broad, and high-throughput, and capable of finding the optimal additive(s). The industrial partner in this GOALI project meets monthly with the MIT team, and has the resources to develop and commercialize best-in-class additives that result in nanocomposites that meet or exceed its design requirements. The challenge of verifying nucleation rates experimentally and the computational expense of determining them by simulation remain the primary barriers to implementation of the MGI approach for this class of problems.

Publications and References

1. A. J. Bourque, C. R. Locker, and G. C. Rutledge, *Heterogeneous Nucleation of an n-Alkane on Tetrahedrally Coordinated Crystals*, J. Phys. Chem. B **121**, 904 (2017).
2. A. J. Bourque and G. C. Rutledge, *Heterogeneous Nucleation of an n-Alkane on Graphene-like Materials*, Euro. Poly. J. **104**, 64 (2018).
3. B. Wang, R. Utzeri, M. Castellano, P. Stagnaro, A. Müller, D. Cavallo, *Heterogeneous Nucleation and Self-Nucleation of Isotactic Polypropylene Microdroplets in Immiscible Blends: From Nucleation to Growth-Dominated Crystallization*, Macromolecules **53**, page (2020).
4. S. Diwale, M.K. Eisner, C. Carpenter, W. Sun, G. Rutledge, R. Braatz, *Bayesian Optimization for Material Discovery Processes with Noise,* Molecular Systems Design & Engineering, 2022, 7, 622-636.
5. N. Volchko, G. Rutledge, R. Braatz, *Heterogeneous Nucleation Mechanisms in Polyolefins: Experiments Linked with Molecular Simulations*, AIChE Annual Meeting (2020).
6. N. Volchko, G. Rutledge, R. Braatz, *Heterogeneous Nucleation in Polyethylene: Experiments and Molecular Simulations*, AIChE Annual Meeting (2021).

Chemical Catalysis for Bioenergy (ChemCatBio): A DOE Energy Materials Network (EMN) Consortium

Poster Presenter: Kurt Van Allsburg, National Renewable Energy Laboratory (kurt.vanallsburg@nrel.gov)
Lead Investigator: Josh Schaidle (joshua.schaidle@nrel.gov), Dan Ruddy (dan.ruddy@nrel.gov)
Participating Institutions: National Renewable Energy Laboratory, Pacific Northwest National Laboratory, Oak Ridge National Laboratory, National Energy Technology Laboratory, Los Alamos National Laboratory, Argonne National Laboratory
Website: chemcatbio.org
Keywords: bioenergy, catalysis, biofuels, bioproducts, thermochemical

Project Scope

The Chemical Catalysis for Bioenergy (ChemCatBio) consortium is accelerating the catalyst and process development cycle for catalytic technologies that convert biomass and waste resources into renewable fuels, such as sustainable aviation fuel and biofuels for other difficult-to-electrify sectors, and chemicals. ChemCatBio is part of the DOE Energy Materials Network (EMN) and is funded by DOE's Bioenergy Technologies Office.

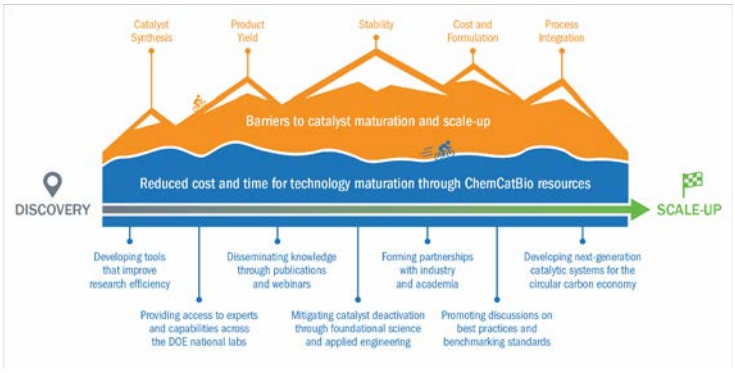


Figure. The path to catalyst deployment is slow and difficult. ChemCatBio is accelerating the catalyst and process development cycle.

Relevance to MGI

As part of the DOE's EMN, ChemCatBio employs a systematic approach leveraging close partnership between experiment and modeling, while also building tools for accelerated improvement. Catalysts under realistic biomass processing conditions are characterized for active sites using *in situ* and *operando* spectroscopy and characterized for atomic structure using state-of-the-art microscopy. The multi-scale modeling team simulates the performance of multiphase and catalytic processes under realistic conditions in which chemical kinetics and transport processes are all relevant factors. ChemCatBio has developed a Catalyst Property Database to accelerate discovery of catalyst-property correlations, a catalyst cost estimation tool (CatCost™) to allow inclusion of catalyst economics in R&D decisions, and a Surface Phase Explorer to understand coadsorption of two species on a surface. As the consortium prepares for its next three years, a major focus area will be the development of predictive design tools that use advanced data analytics, artificial intelligence/machine learning, and the combined insight of performance, structure, and cost insight to support faster discovery of promising materials.

Technical Progress

ChemCatBio leverages national laboratory capabilities in catalyst synthesis, catalyst characterization, catalyst performance evaluation, multi-scale modeling, and technoeconomic analysis / lifecycle analysis to accelerate the development of catalytic technologies that convert biomass and waste resources into renewable fuels and chemicals. The multi-scale modeling team uses atomic-scale experimental data from the catalyst characterization team to elucidate ways to improve catalyst performance through manipulating operating conditions or modifying the catalyst. The catalyst synthesis team develops modified catalysts. In one example, ChemCatBio's experimental and modeling teams closely partnered to baseline a particular conversion process, then identified the active sites to maximize desired hydrogenation and minimize undesired dehydration, then determined sites to improve selectivity, then developed and tested the next-gen catalyst. In under 2 years, the outcome was a next-gen catalyst with

improvement in selectivity from 62% to 89%. In another example, the modeling team investigated how catalyst pore size predicted catalyst lifetime. The experimental team used an improved catalyst with wider mix of micro/nano/mesopores and improved a yield from 10% to 40%.

Future Plans

ChemCatBio has worked to be a central hub of knowledge for the bioenergy community and seeks to continue that plan. In line with the priorities of the DOE Bioenergy Technologies Office, ChemCatBio is increasing its focus on catalysis for biofuels in difficult-to-electrify transportation sectors. For example, ChemCatBio supports DOE's effort in the Sustainable Aviation Fuel (SAF) Grand Challenge for the U.S. to supply 3 billion gallons SAF per year by 2030 and produce up to 35 billion gallons SAF per year by 2050 to meet 100% of aviation fuel demand.

Broader Impacts and Workforce Development

ChemCatBio publishes The Accelerator, a newsletter covering ChemCatBio events, funding opportunities and solicitations, publications, new technologies and research projects, research teams, and more. ChemCatBio has hosted webinars to inform the public on research capabilities.

Data Management and Open Access

ChemCatBio developed the Catalyst Property Database (CPD) at https://cpd.chemcatbio.org/ to make it easier to find an apply theoretical simulation data in catalyst development research. The CPD contains density functional theory computed adsorption energies for surface species in catalyst reactions. ChemCatBio also developed the CatCost tool at https://catcost.chemcatbio.org/ to combine industry-standard cost estimation methods and resources into an intuitive suite of tools bring actionable cost insight to every step of catalyst research and development. CatCost enables rapid development of comprehensive catalyst cost estimates. ChemCatBio also developed the Surface Phase Explorer (SPE) at https://spe.nrel.gov/ to create surface ab initio phase diagrams for adsorption and coadsorption. ChemCatBio also maintains its Data Hub at https://datahub.chemcatbio.org/ for internal data management and for external open access of the data used in the CPD, CatCost, and SPE tools.

Advancing Along the Materials Development Continuum and Partnerships to Translation

ChemCatBio held a Directed Funding Opportunity to work directly with industry partners and help solve their catalyst-related barriers to commercialization. Phase 1 of the opportunity featured nine industry collaboration projects. Phase 2 of the opportunity is ongoing and includes 3 of the Phase 1 projects. Additionally, ChemCatBio has cooperative research and development agreements with several other industry partners.

Publications and References

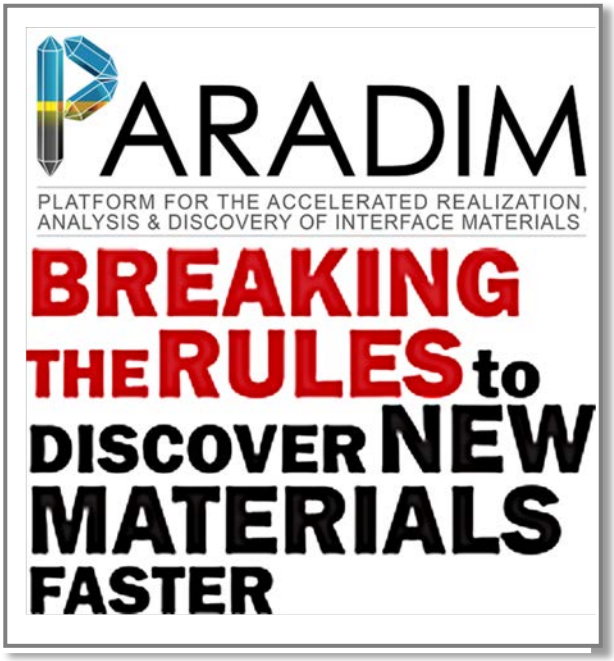
ChemCatBio has over 149 publications, 3 technology licenses, 6 software inventions, and 29 issued patents and patent applications. A full index of publications is at https://www.chemcatbio.org/publications.html

Platform for the Accelerated Realization, Analysis, and Discovery of Interface Materials (PARADIM): An NSF Materials Innovation Platform

Lead Investigator: Darrell G. Schlom, Department of Materials Science and Engineering, Cornell University, schlom@cornell.edu.
Co-Principal Investigator: Lena Kourkoutis, Applied and Engineering Physics, Cornell University, lena.f.kourkoutis@cornell.edu, Tyrel M. McQueen, Departments of Chemistry/Physics and Astronomy/Materials Science and Engineering, Johns Hopkins University, mcqueen@jhu.edu, Kyle Shen, Department of Physics, Cornell University, kmshen@cornell.edu.
Website: https://paradim.org
Keywords: Electronic materials, interface materials, quantum materials, materials-by-design, quantum fabrics.

Project Scope

Creating new interface materials with unprecedented properties, by design rather than by serendipity, is accomplished in PARADIM through a synergistic set of user facilities dedicated to theory (figuring out where to put the atoms for useful behavior), synthesis (putting the atoms in the targeted positions), and characterization (seeing that the atoms are indeed in the desired positions). Each of these world-class user facilities is equipped with the latest tools, techniques, and expertise to realize this materials-by-design dream. Users from throughout the nation are using PARADIM to discover and create interface materials for the next generation of electronics and optoelectronics. These new materials are enabling novel ways for electrons to carry information in solid-state devices and efficiently interact with magnetic, electrical, and optical stimuli.



Relevance to MGI

PARADIM's vision is to democratize materials discovery in the U.S.A. and to enable a more effective way of pursuing materials research, one that accelerates materials discovery by establishing a materials discovery ecosystem—a national community of practitioners—and equipping them with theoretical and experimental methods that enable them to reduce to practice the inorganic materials of which they dream. Discovering electronic materials—by design rather than by serendipity—is accomplished through a synergistic set of state-of-the-art facilities dedicated to theory, synthesis, and characterization. PARADIM's bulk crystal growth facility at Johns Hopkins is the only location within the United States where all major optical floating-zone techniques are available at a single site, including the only location in the world where floating zone growths in supercritical fluids are possible. In PARADIM's thin film facility users can select between a record 62 elements to synthesize the materials they envision by MBE, with integrated ARPES for precise electronic structure determination in quantum fabrics. PARADIM's electron microscopy facility has enabled the highest spatial resolution scanning tunneling electron microscopy (0.39 Å) ever achieved on 2D materials, and is continuing to push the frontier with advanced pixel array detectors and state of the art data analysis techniques. PARADIM's theory facility draws on expertise and computing capabilities across institutions to provide ground-breaking computational models and predictions, including through the PARADIM-developed tool specifically for interface materials: Mismatched INterface Theory (MINT). All PARADIM facilities electronically capture as much data as possible and is developing the big-data tools and techniques to extract insights from these data in real-time. PARADIM's in-house research team leverages these facilities to develop quantum fabrics for the next generation of electronics.

Technical Progress

During its sixth year of operation, PARADIM has made several key accomplishments. Research enabled by the Platform led to a total of 189 peer-reviewed journal publications—including in *Nature*, *Science*, and *Science Advances*. Among all PARADIM works several have been recognized by Web of Science as highly cited papers: 9 publications acknowledging PARADIM, plus 3 additional papers acknowledging use of the Helios FIB, for which PARADIM was the largest contributor beyond the funds coming from Cornell University. 57 publications in peer-reviewed journals have been published, including 21 journal publications based on External User Projects (non-in-house research, non-local users).

Exciting discoveries made by PARADIM users include: (i) a new material that is free of rare-earth elements yet has the highest magnetostrictive coefficient ever achieved at room temperature (10x higher than bulk FeGa and nearly twice as high as Terfenol-D); (ii) the most efficient transverse thermoelectric ever discovered, with performance comparable to commercial longitudinal thermoelectrics, but with drastically reduced energy loss at the contacts; (iii) a new layered nickelate superconductor—predicted by theory and experimentally realized; (iv) the first example of using biaxial strain to transmute a normal metal into a superconductor; (v) a method to peel single-crystal oxide films off their substrates to permit wafer-scale interfacing between high-quality films of virtually any complex oxide, free of epitaxial constraints; (vi) Reduction to practice of a new theoretical prediction—utilizing colinear antiferromagnetic metals with low symmetry, rather than spin-orbit coupling—to achieve efficient spin-transfer torque switching; (vii) synthesis of a visible-light photocathode with unusually high efficiency (quantum efficiencies exceeding 2% at 532 nm) at film thicknesses as low as 4 nm. These epitaxial Cs_3Sb films have an order of magnitude higher quantum efficiency at 650 nm than polycrystalline Cs_3Sb films.

Future Plans

PARADIM will continue to update and enhance its world unique capabilities in synthesis, characterization, and theory to accelerate materials by design, and provide a nexus for the discovery and use of interfacial quantum materials.

Broader impact (Only required for NSF projects)

PARADIM ran three Summer Schools in 2021 that had to adapt to the ongoing COVID-19 pandemic and were held in various formats: purely remote, hybrid, and in-person, for a total of 89 selected participants. Thanks to webinar capabilities, all lectures of the 3 Summer Schools were live streamed to the web and reached a wide audience. For the case of the DFT Summer School, this live streaming reached up to 260 additional attendees! All recorded materials including from past years are available as part of PARADIM's Materials-by-Design Toolbox (https://www.paradim.org/toolbox) to inform potential users about relevant capabilities available through PARADIM user facilities. PARADIM ran an REU program at Cornell and Johns Hopkins for a total of 13 participants, who developed new protocols and systems to improve the MIP-experience for future users of PARADIM. The program concluded with an in-person convocation at Cornell University with all participants (including from JHU) and many of the mentors. In addition, PARADIM with 2DCC-MIP co-organized NSF Community Forums on *Envisioning Pathways to Accelerated Materials Discovery for two-dimensional, interfacial, and layered materials (2DILM)* in January and March 2022 (a draft report is available at https://www.materialsarchive.org/2dilm). PARADIM together with 2DCC-MIP organized the *1st NSF-DMR Materials Innovation Platform Forum* at the Fall MRS 2021. 60 onsite participants + 74 virtually via Zoom learned about the current Platforms, including GlycoMIP and BioPACIFIC MIP, several user success stories and 25 flash talks highlighted the breadth of research pursued at 2DCC and PARADIM.

Data Management and Open Access

All PARADIM facilities are open to all users via a proposal process, facilitating the use of world-unique capabilities by a broad range of scientists. In addition, significant material products from PARDAIM successes will be produced in bulk by REU interns and made available to other researchers via a proposal process. All data from PARADIM facilities is recorded and stored for future use. After a period of inactivity or completion of scientific publications by the primary users, all data associated with user projects is made publically available.

Advancing Along the Materials Development Continuum

Multiple commercial entities currently utilize PARADIM facilities for research directly related to commercialization and the entry into new markets. This is in addition to efforts designed to enhance the deployment of intellectual property (IP) being generated by the in-house research team.

DMREF: Collaborative Research: GOALI: Multiscale Design of Zeolite Sites for Precise Catalytic Transformations

Lead Investigator: William F. Schneider, wschneider@nd.edu; Rajamani Gounder, rgounder@purdue.edu.
Participating Institutions: University of Notre Dame, Purdue University, BASF
Website: none
Keywords: zeolite, heteroatom, catalysis, NO_x reduction

Project Scope

This project explores the hypothesis that the density and locations of heteroatoms within a zeolite framework can be predicted and controlled, and that by exercising this control, the catalytic function and durability of zeolite catalysts can be controlled. We hypothesize that the interaction energies between charged structure directing agents and a heteroatom-decorated framework reports on the likelihood of that specific heteroatom distribution. We test this hypothesis through simulations, synthesis, *ex situ* characterization, and catalyst testing.

Relevance to MGI

Zeolites are versatile materials used in a wide range of adsorption, separation and catalytic applications, including chemical, petrochemical, and emissions control catalysis. Their regular crystalline structure, together with their diverse topologies and heteroatom compositions and arrangements, make them excellent candidates for materials design and optimization. However, current design strategies rely on empirical and anecdotal evidence, with unknown consequences for the density and arrangement of heteroatom sites (typically Al) that determine ion-exchange, adsorption and separation behavior, catalytic performance in acid and redox reactions, and structural durability. Zeolites crystallize in the presence of structure directing agents (SDAs), which serve to organize precursors and to charge-compensate heteroatoms. Here we use first-principles and classical simulations to predict heteroatom distributions in the presence of SDAs, develop titrants to interrogate the locations of those heteroatoms, and use probe catalytic reactions to assess the ability to control reactivity. The project thus advances theory-guided design of zeolite catalysts with tailored chemical functionality.

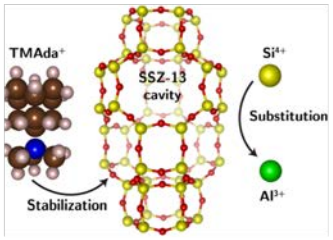


Figure. Interaction energies between OSDA and Al-substituted zeolite framework report on distribution of Al in the framework, with consequences for zeolite properties.

Technical Progress

We focus our efforts on chabazite (CHA), a zeolite framework with important catalytic applications and that that possesses only one symmetry-distinct framework site, so that the only degrees-of-freedom in material design (for a defect-free lattice) from a molecular perspective are the density and distribution of heteroatoms.
(1) **Materials design and synthesis**: We have explored the influence of N,N,N-trimethyl-1-admantylammonium ($TMAda^+$), a common organic structure-directing agent, on Al distributions in CHA. As illustrated in the Figure, $TMAda^+$ can fill the CHA cage in two orientations. We use density functional theory (DFT) to parameterize a classical forcefield and then apply that forcefield to 1000's of unique combinations of $TMAda^+$ orientations and Al distributions. We fingerprint those distributions based on local Al configurational features and find that certain features are energetically biased against or for by the OSDA. Predictions are compared against samples synthesized with $TMAda^+$ and titrated for the same configurations. Materials are observed to evolve away from the predicted unfavorable Al pair features with increasing crystallization time, highlighting the contributions of preorganization

and reorganizational kinetics, and thermodynamics of the zeolite product. To test the tunability of design, we contrast behavior in the presence of mixtures of organic and inorganic SDAs. Using *ab initio* molecular dynamics (AIMD), we contrast energies of key features in the absence and presence of the additional inorganic (Na^+) and find that its presence significantly alters Al siting preferences. Comparisons with experiment confirm models; further, inorganics appear to "lock in" certain features earlier in the crystallization process, revealing that SDA and crystallization time are both design parameters for Al siting control. Finally, genetic algorithm searches are used to contrast the siting behavior of SDA alternatives to $TMAda^+$ to identify candidates for alternative feature generation.
(2) **Feature interrogation**: To expand the toolkit for interrogating Al distributions, we use computations to explore the energetics of alkaline earth cation exchange vs Al location. We find that Ba^{2+} has distinctive and intriguing exchange energetics that suggest that it can titrate a different ensemble of Al features than are other ions.
(3) **Catalytic function**: To test the influence of Al distribution on catalytic function, we focus primarily on Cu-exchanged CHA, a material space known to be active for the selective catalytic reduction (SCR) of NO_x with NH_3. We exploit materials design knowledge from (1) to prepare materials with a wide range of Al content and distribution and load materials with Cu to create a composition space described by Si/Al and Cu/Al ratios. Experiments reveal that per Cu catalytic rates depend on both composition parameters. We employ kinetic Monte Carlo (kMC) simulations and Bayesian optimization to extract composition-dependent kinetic parameters, revealing that rates can be mapped onto composition-dependent Cu kinetic "footprints." Analysis of these footprints reveal correlations with local Al pair features, pointing the way towards optimization of catalytic activity.

Future Plans

Future priority directions include: (1) elaboration of $TMAda^+$/CHA strategy to identify OSDAs with different Al feature biases; (2) development and incorporation of vacancy defect models to enable exploration of factors that influence accessible Si/Al ratios; (3) catalyst syntheses to test computational predictions, including base material composition, influence on speciation and spatial distribution of Cu ions, and consequences for catalytic activity.

Broader Impacts and Workforce Development

This program supports the research activities of experiment- and computation-focused students across two institutions, ensures integration and cross-training with biweekly meetings of the academic team and monthly meetings with the industrial partner. Results are broadly disseminated at specialty and general conferences.

Data Management and Open Access

All research data are fully archived in the literature. We take advantage of Zenodo repositories to archive all computational results and associated codes, for example https://doi.org/10.5281/zenodo.6422186.

Advancing Along the Materials Development Continuum and Partnerships to Translation

Close interactions with BASF partners informs and guides research questions and provides a pathway for rapid application of new insights and materials.

Publications and References

1. Krishna, S. H.; Jones, C. B.; Gounder, R. Temperature Dependence of Cu(I) Oxidation and Cu(II) Reduction Kinetics in the Selective Catalytic Reduction of NOx with NH3 on Cu-Chabazite Zeolites. *Journal of Catalysis* **2021**, *404*, 873–882. https://doi.org/10.1016/j.jcat.2021.08.042.
2. Hoffman, A. J.; Bates, J. S.; Di Iorio, J. R.; Nystrom, S. V.; Nimlos, C. T.; Gounder, R.; Hibbitts, D. Rigid Arrangements of Ionic Charge in Zeolite Frameworks Conferred by Specific Aluminum Distributions Preferentially Stabilize Alkanol Dehydration Transition States. *Angew. Chem. Int. Ed.* **2020**, *59*, 18686–18694. https://doi.org/10.1002/anie.202007790.
3. Wang, X., Wang, Y., Moini, A., Gounder, R., Maginn, E. J., Schneider, W. F., Influence of N,N,N-trimethyl-1-adamantyl ammonium (TMAda+) Structure Directing Agent on Al Distributions and Pair Features in Chabazite Zeolite, *Chem. Mater.* Submitted.

DMREF/Collaborative Research: Designing Mutable Metamaterials with Photo-Adaptive Meta-Atoms

Lead Investigator: M. R. Shankar, ravishm@pitt.edu.
Participating Institutions: M. R. Shankar (U. Pitt), M. Brongersma (Stanford), H. Noh (Stanford), K. Dayal (Carnegie Mellon), R. Lipton (Louisiana State)
Website: none.
Keywords: Active Metastructures, Machine Learning, Liquid Crystal Elastomers, Inverse Design, Plasmonics

Project Scope

The goal of this research is to discover topologically optimized polymer-composite plasmonic metastructures, which can be structurally modulated via photochromic switching of the macromolecular network. Translating photomechanical adaptivity into changes in the scattering amplitude and phase of incident light will unlock active metasurfaces capable of device-level functionality. It is hypothesized that plasmonic structures can couple photons into molecularly ordered, photomechanically active polymers on which they are resident, which drives conformational changes and actuation. This photomechanical response back-couples to the plasmonic structures to generate mutable optical responses that can be harnessed in active metasurfaces for beam steering, and wavefront shaping.

Relevance to MGI

This research integrates analytical models by Lipton, computational methods from Brongersma to design structures that couple actinic photons with photochromic switches that inhabit gaps between plasmonic antenna and a metallic substrate. These forward models also offer a library of optical responses that emerge as a function of the (adaptive) geometry of the nanostructures. Inverse models (Noh and Dayal) utilize machine learning to build on this library to unlock the ability to design nanostructures for targeted optical responses. These forward and inverse models inform an experimental campaign to fabricate plasmonic metastructures that are resident upon photoactive liquid crystal elastomers (LCE). Magnetically assisted blueprinting of molecular order in LCE that contain photochromic (azobenzene) switches directs the mechanical actuation that is back coupled to the resident plasmonic structures. Characterization of the optical properties validate the computational methods and refine the machine learning based approaches for inverse designs. The iterative coupling between theory and experiments (Figure 1) accelerates the discovery of active plasmonic structures, catalyzes the development of new experimental methods to integrate organic and inorganic materials and unlocks novel approaches that couple analytical, computational and machine learning algorithms.

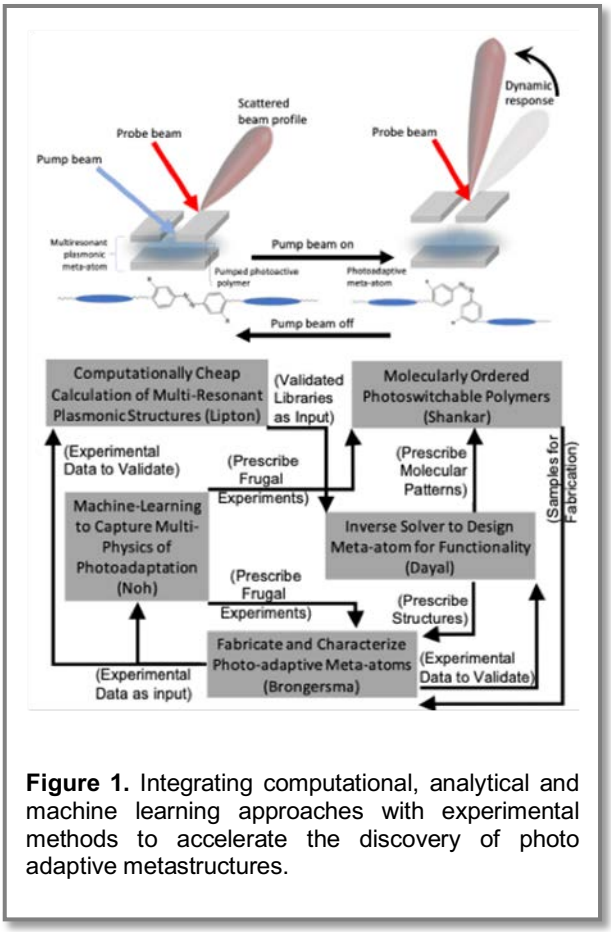


Figure 1. Integrating computational, analytical and machine learning approaches with experimental methods to accelerate the discovery of photo adaptive metastructures.

Technical Progress

Analytical models are being developed to characterize the effect of geometry and material properties on the interaction of multicomponent impedance surfaces with incident electromagnetic radiation. A new analytic model for plasmonic particles patterned on very thin film coatings is established. It models the metal on film layer as an effective impedance layer sensitive to periodic localized plasmon resonances. It is obtained by matched asymptotic expansions perpendicular to the layer and periodic two scale expansions along the surface. This will be used in conjunction with rigorous coupled wave analysis for beam steering. Also, homogenization theory was used to replace the Maxwell's system of equations using a transfer matrix method. These are complemented by computational models of the interaction of light with nanostructures

with a focus on intensification of optical fields in gaps between stripe cavities and metallic substrates. The forward models were used to train a neural network to capture the essential features of the optical response as a function of geometry and material properties. The optimized networks showed 99.7% accuracy in capturing the forward models and 96% accuracy for inverse modeling. 99% accuracy was achieved for reconstruction of the optical response. The physics informed deep learning model developed alleviates the issue of non-interpretable result that is often encountered in case of predictions with deep learning models. The predictions are scientifically consistent and thus interpretable. We also use a tandem architecture to solve the non-unique prediction problem due to one – many mapping between response and design of metasurface.

In parallel, a soft lithography-based approach was developed for creating nanometer scale LCE films on gold substrates. The alignment was enforced using an externally applied 2T magnetic field to overcome the effect of the metallic substrate. The samples were characterized for their photochromism, where their structural evolution was tracked using absorption and Raman spectroscopy as a function of the irradiation dose. The samples were decorated with nanoparticles and the scattering response was characterized as a function of the structural modulation with light. Nanostencil lithography was developed for solvent free patterning of nanostructures on polymer films. The idea is to fabricate a hard mask that can be directly placed in contact with the polymer. Evaporation of metal layers through the mask creates the desired nanostructures. To create the nanostencils, silicon nitride films were deposited on double polished silicon wafers. Patterning and etching the wafers resulted in nanostencils with ultrafine features.

Future Plans

Our team has developed a framework for modeling nanostructure-light interaction, discovering plasmonic structures using machine learning and a platform for fabricating the composite metasurfaces. The ongoing efforts will build on these capabilities to realize photoadaptive functionalities with plasmonic metastructures, which allow for encoding device-level functionalities in metasurfaces.

Broader Impacts and Workforce Development

This project supports two PhD students at the University of Pittsburgh, who are developing methods for creating nanostructured thin films with optimized photoactive compositions. Carnegie Mellon University has hired one PhD student to work on the machine learning. This student is co-advised by Prof. Noh and Dayal. Stanford has hired one PhD student to work on the nanofabrication of the plasmonic structures and characterization of the photomechanical responses. Lousiana State University has recruited one PhD student to develop the analytical models.

Data Management and Open Access

Our team is committed to sharing the codes on open-source platforms and disseminating the results via publications and conference presentations.

Advancing Along the Materials Development Continuum and Partnerships to Translation

Our program that integrates experiments and theory will accelerate the development of adaptive metasurfaces that can generate device-level functionality using light itself as the actinic power source. We are also pursuing follow-on projects that build on this effort to encode autonomy within the material design itself.

Publications and References

The following papers and presentations acknowledge support from this grant:

1. Gao, Junfeng, et al. "Molecularly Directed, Geometrically Latched, Impulsive Actuation Powers Sub-Gram Scale Motility." Advanced Materials Technologies 7.5 (2022): 2100979.
2. Clement, Arul, et al. "Complexity from simplicity: Confinement directs morphogenesis and motility in nematic polymers." Extreme Mechanics Letters 47 (2021): 101362.
3. Wang, Zhijie, and M. Ravi Shankar. "Spatially selective, solid state etching of diamond using lithographically patterned FeCoB." Diamond and Related Materials 121 (2022): 108763.
4. Robert Lipton and R. Perera. “Bloch spectra for high contrast elastic media.” Journal of Differential Equations 332 (2022): 1-49.
5. Li, Wei, Robert Lipton, and Matthias Maier. "Lorentz resonance in the homogenization of plasmonic crystals." Proceedings of the Royal Society A 477.2256 (2021): 20210609.
6. Abiti Adili, et.al. “Controlling the Dispersion of Metamaterials in 3-d,” Proceedings of the Royal Society A (2022). Accepted for Publication.
7. Sarkar, Sulagna, et al. “Physics-Guided Machine Learning Framework for Simultaneous Inverse Design and Optical Response Prediction of Photonic Metamaterials”, SIAM Annual Meeting (2022).
8. Sarkar, Sulagna, et al. “Physics based Machine Learning for inverse design of metasurfaces”, Mechanistic Machine Learning and Digital Twins for Computational Science, Eng. & Tech.: An IACM Conference (2021).

DMREF – Strain-adaptive materials

Lead Investigator: Sergei S. Sheiko, sergei@email.unc.edu.
Participating Institutions: UNC at Chapel Hill and Carnegie Mellon University
Keywords: Polymers, networks, strain-adaptive, tissue-mimetic

Project Scope

The goal is to design strain-adaptive materials that emulate living tissue's response to deformation–vital for soft robotics, wearable electronics, and biomedical devices. This is achieved by developing a new materials design platform that harnesses polymer network architecture to program grand variation of mechanical phenotypes ranging from soft fat tissue to firm skin. This platform will create a materials design search engine that will automatically provide *molecular codes* for the predictive design of tissue-mimetic materials with tailored properties.

Relevance to MGI

Our project leverages individual expertise with *daily* iterative feedback between theoretical predictions of structure-property correlations (Dobrynin), precise polymer synthesis (Matyjaszewski), and quantitative measurements of physical properties (Sheiko). Unlike many phenomenological theories currently used in the materials design, our theoretical models establish *quantitative* correlations between specific architectural parameters and specific physical properties. The developed design rules are directly applied by Matyjaszewski and Sheiko groups to the synthesis of brush-like macromolecules with precise molecular architecture which self-assemble into polymer networks with predetermined mechanical properties. Furthermore, we have developed a unique quality control methodology which allows non-invasive forensics of polymer networks to verify the correlation between the prescribed codes and the actual structure of synthesized materials. This *design-by-architecture* approach equipped with the *forensics* tool significantly enhances the materials design efficiency in contrast to the conventional trail-and-error approach based on explorative mixing of polymers, solvents, and particles leading to ill-defined materials with unknown internal organization.

Technical Progress

The key outcomes include 1) A methodology for the synthesis of graft block-copolymers that self-assemble to allow accurately control both brush dimensions and network topology; 2) A bridge between soft-and-compliant gels and soft-yet-firm tissues by tuning modulus and firmness independently of one another; 3) A new class of magnetoactive elastomers that allow for external regulation of both elastic and visco-elastic properties. 4) A forensics methodology that enables non-invasively deciphering the internal organization of polymer networks.

Independently Tuning Softness and Firmness. Softness and firmness are opposing traits that synergistically define the elastic response of biological systems. Currently, no single class of synthetic materials can provide independent control of these mechanical characteristics, particularly without altering chemical composition. To address this challenge, we explored hierarchical self-assembly of bottlebrush mesoblocks into reversible networks by judiciously incorporating side chains of different lengths at an encoded grafting density delivering thermoplastic elastomers with unprecedented softness-firmness combinations (**Fig. 1a**).

Forensics of Polymer Networks. Polymer networks are a prominent class of soft matter essential in applications from packaging to soft robotics. Despite a century of research, understanding the properties of elastomers and gels remains a grand challenge in polymer science because the network architecture remains a "black box" dominated by the convolution of a topological hierarchy. By analyzing the non-linear deformation of elastomers, we exposed their interior structure, including crosslink density, strand persistence length,

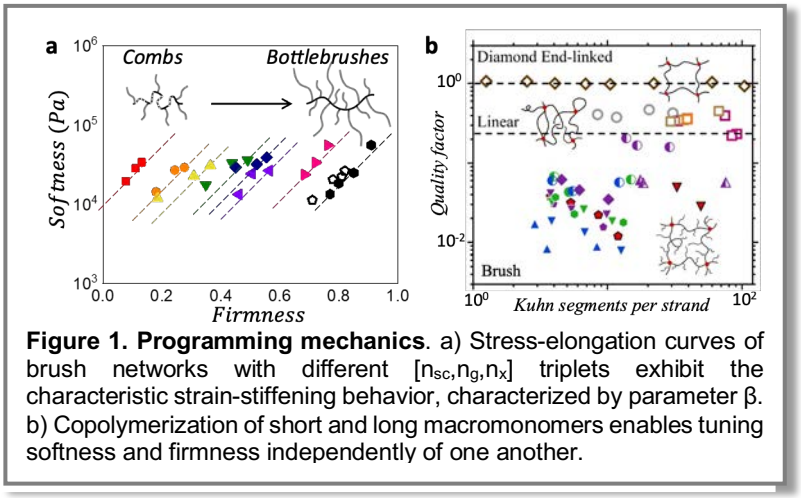


Figure 1. Programming mechanics. a) Stress-elongation curves of brush networks with different [n_{sc},n_g,n_x] triplets exhibit the characteristic strain-stiffening behavior, characterized by parameter β. b) Copolymerization of short and long macromonomers enables tuning softness and firmness independently of one another.

and fraction of the load-bearing strands. The decoded structures enabled classification of various polymer networks according to the effectiveness of stress distribution characterized by a topological quality factor (**Fig. 1b**).

Future Plans

The next step is to begin training Machine Learning algorithms in implementing Human Intelligence, *aka* theoretical predictions, towards the creation of a platform for Artificial Intelligence (AI)-governed manufacturing of polymer networks with precisely targeted combinations of properties. This will be achieved by integrating the molecular code-guided synthesis and reverse forensics of as-prepared networks.

Broader Impacts and Workforce Development

This project builds a foundation for an open-source Big Data search engine that will generate protocols for predictive fabrication of personalized soft materials. The interdisciplinary nature of the project provides maximum opportunity for (i) integrating science, technology, and education through modular teaching, multiexpertise training, and collaborations with industry and medical schools, (ii) enhancing diversity by broadening participation of underrepresented groups (2 female and 2 hispanic students), and (iii) fostering infrastructure for collaborative research through students exchange and cross-mentoring. Each student in our group is trained to do synthesis, measure physical properties, and conduct theoretical analysis of experimental data. This interdisciplinary training is unique and vital for building the next generation workforce in materials research.

Data Management and Open Access

The research outcomes are published in scientific (including open access) journals that are accompanied by the supporting information containing all data sets. These data sets are available for free download from journals' websites. We are currently working on creating a Data Bank of architectural codes that emulate the mechanical response of a highly diverse range of biological tissues from ultra-soft fat tissue to highly resilient skin. All primary data have been made publicly available by the time of publication through a designated URL link: https://tarheels.live/dobryningroup/data-request/.The available data include computer simulation trajectories, synthetic protocols, molecular characterizations, and physical properties.

Advancing Along the Materials Development Continuum and Partnerships to Translation

Tissue-mimetic materials are vital for high-impact technologies such as soft robotics, wearable electronics, and biomedical devices. In addition to industrial projects and collaborations with medical schools, several companies have expressed their interest in our materials. Currently, one company is considering licensing one of our IPs (one US patent and three non-provisional patent applications).

Publications and References

1. M. Maw, B.J. Morgan, Y. Tian, K. Matyjaszewski, A.V. Dobrynin, and S.S. Sheiko, "*Brush Architecture and Network Elasticity: Path to the Design of Mechanically Diverse Elastomers*" Macromolecules **55**, 2940 (2022).
2. K. Cushman, A.N. Keith, J. Tanaka, S.S. Sheiko, W. You, "*Investigating the Stress-Strain Behavior in ROMP-based Brush Elastomers*" Macromolecules **54**, 8365 (2021).
3. S.A. Kostrov, E. Dashtimoghadam, A.N Keith, S.S. Sheiko, E.Y. Kramarenko, "*Regulating Tissue-Mimetic Mechanical Properties of Bottlebrush Elastomers by Magnetic Field.*" ACS Appl. Mater. Interf. **13**, 38783 (2021).
4. M. Jacobs, H. Liang, A.V. Dobrynin, "*Theory and Simulations of Hybrid Networks*" Macromolecules **54** , 7337 (2021).
5. E. Dashtimoghadam, A.N. Keith F. Vashahi, M. Vatankhah, S.S. Sheiko "*Injectable Non-leaching Tissue-mimetic Bottlebrush Elastomers: A New Platform for Advancing Reconstructive Surgery*" Nat. Comm. **12**, 1 (2021).
6. I. Yu, E.R. Newton, D.C. Gillis, K. Sun, B.C. Cooley, A.N. Keith, S.S. Sheiko, N.D. Tsihlis, M.R. Kibbe, "*Coating small-diameter ePTFE vascular grafts with tunable poly (diol-co-citrate-co-ascorbate) elastomers to reduce neointimal hyperplasia*." Biomater. Sci. **9,** 5160 (2021).
7. H. Starvaggi, Y. Tian, H. Liang, A.V. Dobrynin, "*Bottlebrushes and Combs with Bimodal Distribution of the Side Chains: Diagram of States and Scattering Function*.", Macromolecules **54**, 1818 (2021).
8. A.N. Keith, C. Clair, Y. Tian, A.V. Dobrynin, S.S. Sheiko, "*Independently Tuning Elastomer Softness and Firmness by Incorporating Side Chain Mixtures into Bottlebrush Network Strands*.", Macromolecules **53**, 9306 (2020).
9. D. Zhang, E. Dashtimoghadam, F. Fahimipour, X. Hu, M. Vatankhah, S.S. Sheiko, "*Tissue-adaptive materials with independently regulated modulus and transition temperature*." Adv. Mater. **32**, 2005314 (2020).
10. Y. Cong, A.N. Keith, F.A. Leibfarth, M.R. Martinez, K. Matyjaszewski, S.S. Sheiko, "*Understanding the Synthesis of LBL Block Copolymers: Plastomers with Well-Defined Mechanics*." Macromolecules **53**, 8324 (2020).

DMREF: Engineering the On-the-Fly Control of 3-D Printed Block Bottlebrush Assemblies via Dynamic Bonds and Materials Processing

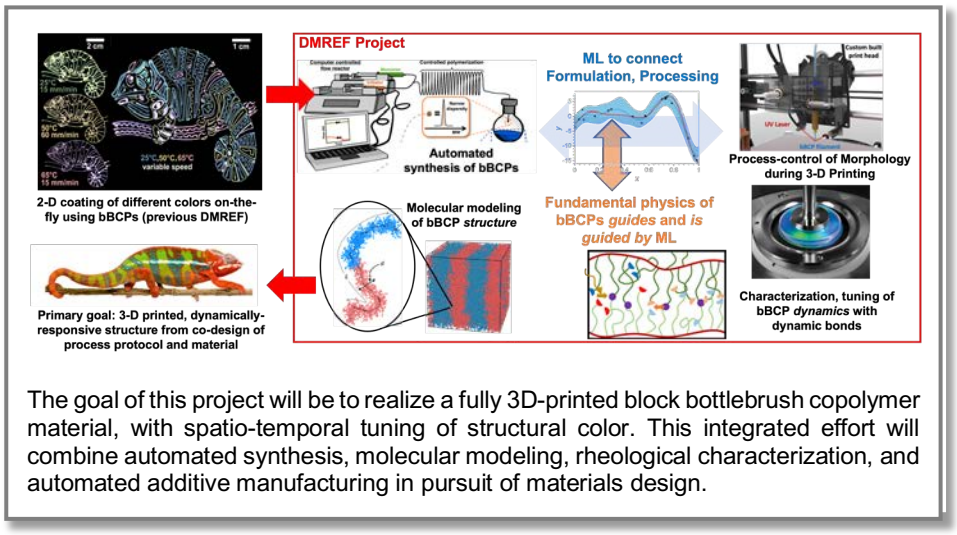
Lead Investigator: Charles E. Sing, cesing@illinois.edu.
Participating Institutions: University of Illinois at Urbana-Champaign
Website: none.
Keywords: Polymer architecture, self-assembly, additive manufacturing, dynamic bonds, structural color.

Project Scope

This DMREF program will establish the fundamental science and design rules for harnessing the versatility of process-directed, 3-D printed bottlebrush block copolymer (bBCP) materials. Processing variables such as flow rate can manipulate and pattern self-assembled morphologies at the molecular level. Bottlebrush block copolymers are an ideal platform because they self-assemble rapidly, and they can be synthesized with a wide range of architectures and dynamic interactions that provide a large parameter space. We will efficiently explore this parameter space via an experimental/computational approach that is informed by machine learning; we seek to engineer processing-dependent assembly by controlling molecular shape and interactions.

Relevance to MGI

This research program will use a combined approach of synthesis, characterization, modeling, and machine learning to realize material-processing approaches for advanced additive manufacturing. This is intended as an integrated approach to materials design, utilizing the combined expertise of the PIs. The overall goal will be to co-design a bottlebrush block copolymer material, chemical functionality, and a 3-D printing process to exhibit on-the-fly, spatially-varying morphologies. To this end, Guironnet will synthesize molecules to be used and characterized by Diao and Rogers. Rogers will develop insights into rheology and structure, accompanied by molecular simulation from Sing, that will provide insight into printing performed by Diao. Automated synthesis and formulation from Guironnet, along with printing and characterization of assembled structures by Diao will provide the foundations for automated printing schemes developed in collaboration with Maruyama at AFRL. This will be used to identify interesting parameter regimes in synthesis, and focus the attention of fundamental studies by Rogers and Sing; these physical insights will inform automated printing to make it generally adaptable to 3-D printing processes. Over half of the 19 papers published in the previous iteration of this DMREF project were multi-PI efforts, evidence of a highly-integrated research team.



Technical Progress

This project has resulted in several key research accomplishments: (1) Libraries of bottlebrush and bBCP polymers were synthesized, and synthesis was automated to rapidly make bottlebrush polymers of different architectures, including shape-defined structures. (2) Characterization using SAXS and rheology showed the mechanistic connection between bBCP assembled morphology, deformation, and structural color. (3) These synthesized bBCPs were incorporated into solution coating printing processes, with a proof-of-principle of on-the-fly printing of different colors. (4) A molecular model of bottlebrush structure was developed, along with a coarse-

grained counterpart. Both models agree with experiment, both dilute and assembled bottlebrushes. We successfully established a comprehensive picture of bBCP solution assembly, and demonstrated that processing could be used to modulate structural color on-the-fly.

Recent progress has centered around crosslinking and solvent-control of bBCP assembly in and out of equilibrium. This includes (1) developing synthetic methodologies to quantitatively functionalize the tips of bottlebrush polymer side-chains, (2) using Monte Carlo models to simulate the synthesis of shaped bottlebrushes, (3) elucidating the nature of bottlebrush polymers in solution using simulation and scaling arguments, (4) combining coarse-grained models and SAXS characterization to explore the origins of structural color, including the effect of evaporation-driven assembly, and (5) characterizing the yielding rheology of bottlebrush assemblies, important for 3-D printing. We have begun to address the coupled challenges of understanding assembly versus processing versus crosslinking timescales, and how these relate to material properties and molecular structure.

Future Plans

The next steps for this project are to systematically understand the role of triggerable/dynamic bonds on molecular self-assembly and material properties. Intermediate efforts will focus on (1) extending coarse-grained modeling approaches to contend with semidilute bBCP interactions and the dynamic molecular interactions induced by crosslinking, (2) using recovery rheology techniques to understand the emergence of yield-stress behavior and quantify printability in self-assembled bBCPs, (3) exploring the use of solvent to trigger and manipulate large transitions in molecular morphology of bBCPs, (4) establishing automated and versatile syntheses to functionalize bBCPs with a variety of dynamic or permanent crosslinkable moieties, and (5) quantifying the connection between structural color, morphology, and processing conditions. Longer-term efforts will be centered around integrating fundamental studies of bBCPs into an automated materials discovery scheme, combining automated synthesis and printing, as well as fundamental physical modes and insights, to design a versatile set of bBCPs that can exhibit dramatic and controllable structural color changes.

Broader Impacts and Workforce Development

The educational impact of this DMREF project comes in part from the collaboration between four PIs with complementary backgrounds. Cross-disciplinary knowledge was gained by students, who will benefit from a broader education that will prepare them for STEM careers. This program has played a role in training 9 graduate students, 1 postdoctoral researcher, and several undergraduate students. We also participate in outreach efforts, developing programming for the Girls' Adventures in Mathematics, Engineering, and Science (GAMES) camp, the St. Elmo Brady STEM Academy, and using expertise in 3-D printing for a 'CCCR-Champaign County COVID Relief projects to make 3-D printed face masks and respirators.

Data Management and Open Access

This project will generate a large amount of data composed of computational code, outputs from simulations, synthesis characterizations, and rheological and scattering data. We facilitate long-term stewardship of our data by adhering to the FAIR (findable, accessible, interoperable, and reusable) guiding principles. Both theoretical/computational and experimental data will be archived and hosted on data repositories when possible, using standard file formats that will be made available upon publication.

Publications and References

1. T.P. Pan, S. Dutta, Y. Kamble, B.B. Patel, M.A. Wade, S.A. Rogers, Y. Diao, D. Guironnet, C.E. Sing, *Materials Design of Highly Branched Bottlebrush Polymers at the Intersection of Modeling, Synthesis, Processing, and Characterization*, Chem. Mater. **34**, 1990 (2022).

2. B.B. Patel, T. Pan, Y. Chang, D.J. Walsh, J.J. Kwok, K.S. Park, K. Patel, D. Guironnet, C.E. Sing, Y. Diao, *Concentration-Driven Self-Assembly of PS-b-PLA Bottlebrush Diblock Copolymers in Solution*, ACS Polymers Au. DOI: 10.1021/acspoolymersau.1c00057 (2022).

3. D.J. Walsh, M.A. Wade, S.A. Rogers, D. Guironnet, *Challenges of Size-Exclusion Chromatography for the Analysis of Bottlebrush Polymers,* Macromolecules, **53,** 8610 (2020).

4. D.J. Walsh, S. Dutta, C.E. Sing, D. Guironnet, *Engineering of Molecular Geometry in Bottlebrush Polymers,* Macromolecules, **52,** 4847 (2019).

5. B.B. Patel, D.J. Walsh, D.H. Kim, J. Kwok, B. Lee, D. Guironnet, Y. Diao, *Tunable Structural Color of Bottlebrush Block Copolymers Through Direct-Write 3D Printing from Solution,* Sci. Adv. **6,** eaaz7202 (2020).

Collaborative Research: DMREF: GOALI: Discovering Materials for CO_2 Capture in the Presence of Water via Integrated Experiment, Modeling, and Theory

Lead Investigator: Randall Q. Snurr, snurr@northwestern.edu
Participating Institutions: Northwestern University, University of South Alabama, NuMat Technologies
Website: none
Keywords: Carbon capture, nanoporous materials, metal-organic frameworks, separations, adsorption

Project Scope

The concentration of carbon dioxide in the atmosphere has risen rapidly over the past century, creating significant concerns. Carbon capture and sequestration is widely viewed as an essential tool, along with other technologies, for keeping atmospheric CO_2 levels from rising further. This project focuses on developing new materials for selective adsorption of CO_2 versus N_2. Metal-organic frameworks (MOFs) are a promising class of materials for carbon capture due to their modular building-block synthesis, which allows for targeted tuning of material properties and for an almost unlimited number of potential MOFs. A primary emphasis of this project is the effect of water on CO_2/N_2 selectivity and CO_2 capacity. The main goal is to develop integrated simulation, theoretical, and experimental methods for understanding the effect of water on CO_2/N_2 separations in MOFs and to use these tools to speed up the discovery of new adsorbents for CO_2 capture and related separations.

Relevance to MGI

In line with the Materials Genome Initiative, this project will contribute to the development of new strategies for creating adsorbent materials with programmable structure, by precisely combining preassembled building blocks (i.e., MOF nodes, organic linkers, ions, and functional groups). In particular, we focus on a few MOF platforms that can be systematically tuned by changing the organic linkers and introducing extra-framework anions and cations, and restructured MOF nodes. These "platform" MOFs are chosen from families of MOFs known to exhibit excellent stability. Optimization of MOF synthesis will be accelerated by use of robotic synthesis tools coupled with machine learning algorithms. Molecular simulation will be used to test new proposed material variations and provide molecular-level insights into observed behavior. The simulation models will be validated against adsorption data collected by our team. All simulated and experimental adsorption data will be placed in publicly accessible databases and will be used to assess the accuracy of methods for predicting mixture behavior from single-component adsorption isotherms. For the most promising materials, we will perform single-crystal X-ray studies of molecular siting and arrangements in the pores.

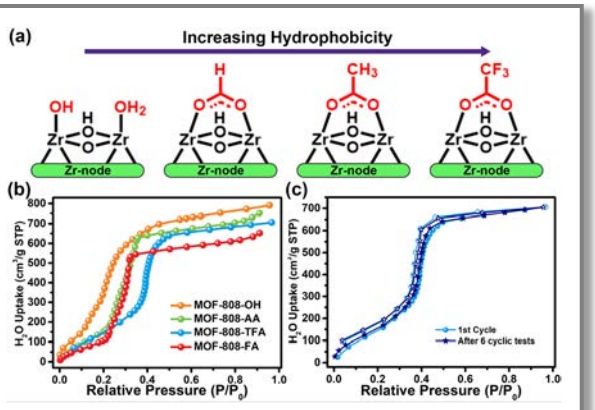


Figure 1. One concept explored in this project is how MOF hydrophobicity affects CO_2/N_2 selectivity and CO_2 working capacity for carbon capture. Water is an ever-present component in flue gas streams, and dehumidification of the flue gas stream is cost-prohibitive. Therefore, development of MOFs that can selectively take up CO_2 in the presence of water would greatly advance the field of solid adsorbent-based CO_2 capture technologies.

Technical Progress

The stability of MOFs in water directly affects their capacity as adsorbents for carbon capture and other applications. Recently, we examined the water stability of four node-modified variants of the mesoporous MOF, NU-1000, namely formate-, Acac-, TFacac-, and Facac-NU-1000, comparing these with node-accessible NU-1000. These ligands graft to NU-1000's hexa-Zr(IV)-oxy nodes by displacing terminal aqua and hydroxo ligands. The

node-modifying ligands show different degrees of hydrophobicity. Facac-NU-1000, containing the most hydrophobic ligands, showed the greatest water stability, being able to undergo at least 20 water adsorption/desorption cycles without loss of water uptake capacity. The measurement of multicomponent adsorption isotherms has begun using our existing, state-of-the-art, volumetric adsorption apparatus. We have validated our volumetric adsorption apparatus by measuring the binary adsorption isotherm of a refence adsorbent and have begun collecting mixed-gas adsorption isotherms on one of the platform MOFs.

Future Plans

The main goal of the proposed work is to develop integrated simulation, theoretical, and experimental methods for understanding the effect of water on CO_2/N_2 separations in MOFs and use these tools to speed up the process of finding new adsorbents for CO_2 capture and related separations in a systematic fashion. Selectively capturing CO_2 over N_2 in the presence of water is challenging, because most "sites" or functional groups that one might design to bind CO_2 also bind H_2O. We will explore families of materials generated from two related concepts.

Another objective is to further develop some essential tools for this research. Given the novelty of the concepts for introducing extra-framework cations and anions, we must develop and test appropriate force fields for molecular simulation of these systems. The simulation models will be validated against adsorption data collected by our team, including multicomponent adsorption measurements, which are extremely scarce in the literature. We will implement and test robotic synthesis tools – coupled with machine learning – to accelerate the optimization of MOF synthesis procedures, and we will test and further develop a method recently developed by our team for performing single-crystal X-ray studies of molecular siting that does not require synchrotron access. Such a tool should significantly expand access to single-crystal X-ray crystallography methods for a variety of applications.

Broader Impacts and Workforce Development

The project contributes to the education of 4 PhD students in an interdisciplinary project, as well as a post-doc and undergraduates. Summer undergraduate students will be recruited through a program that brings minorities and disadvantaged undergraduates to our campuses. Web-based education activities will reach a wider audience.

Data Management and Open Access

Simulation results will be placed in a searchable, public database (https://mof.tech.northwestern.edu/). We include relevant metadata such as details about the force field parameters used to generate the data, so that all calculations are fully reproducible. The database is easily searchable by MOF properties, and the adsorption data are stored in a standardized JSON format that is interoperable with the NIST adsorption database. By making all this data more easily accessible, we hope to facilitate new data mining and machine learning studies on these materials, leading to new insights on adsorption in MOFs.

Advancing Along the Materials Development Continuum and Partnerships to Translation

The project aims to discover new materials for carbon capture by using "platform" materials that can be tuned in a systematic way. Importantly, these materials can be explored computationally to accelerate the discovery process. The proposed research has the potential to contribute to the mitigation of CO_2-driven climate change. This work may also lead to nanoporous materials that may be useful in applications such as chemical separations, gas storage, sorption-based cooling, and water purification. The involvement of a co-PI from NuMat Technologies will help identify "show stoppers" and ensure that the team is working on industrially relevant problems.

Publications

1. K.M. Fahy, M.R. Mian, M.C. Wasson, F.A. Son, T. Islamoglu, O.K. Farha, *Exchange of coordinated carboxylates with azolates as a route to obtain a microporous zinc–azolate framework*, Chem. Commun. 2022, **58**, 4028-4031.
2. J. Liu, R. Anderson, K.M. Schmalbach, T.R. Sheridan, Z. Wang, N.M. Schweitzer, A. Stein, N.A. Mara, D. Gomez Gualdron, J.T. Hupp, *Dual-functional modification to enhance water stability of mesoporous zirconium metal–organic frameworks*, Journal of Materials Chemistry A, submitted.

Collaborative Research: DMREF: GOALI: Physics-Informed Artificial Intelligence for Parallel Design of Metal Matrix Composites and their Additive Manufacturing

Lead Investigator: Aaron Stebner, aaron.stebner@gatech.edu
Participating Institutions: Georgia Tech, University of Utah, University of Michigan, Elementum 3D, AFRL
Website: 'none'
Keywords: Artificial Intelligence (AI), Machine Learning (ML), Metal Matrix Composites (MMC), Additive Manufacturing (AM), Data Imputation

Project Scope

To enable AI-driven parallel design of metal matrix composites and their additive manufacturing, five AI-enabling algorithmic gaps will be filled: 1) automatic data analysis-interpretation-curation from process-structure-property (PSP) data sources; 2) automatic data cleaning and concatenation of databases so that AI can modify and append the data spaces when new data sources or data features are incorporated into a research problem; 3) automatic data feature mapping across multiple length and time scales to complete PSP data ontologies; 4) data feature engineering; 5) PSP machine learning models of global relationships across multiple nested PSP submodels. MGI success will be measured against a techno-economic baseline study of commercializing a new MMC for AM by GOALI partner Elemenutm 3D.

Relevance to MGI

Basic research to inform algorithm development will enable AI discovery and optimization of materials and their manufacturing toward reducing deployment times and costs by half, to meet the MGI vision. Such AI can be visualized as a network of nested PSP relationships expressed as $y = f(\mu, \varphi)$ (**Fig. 1**) where μ represents the input PSP variables, y represents the output PSP variables, and φ denotes variables describing the physics controlling the material + manufacturing phenomenon of interest.

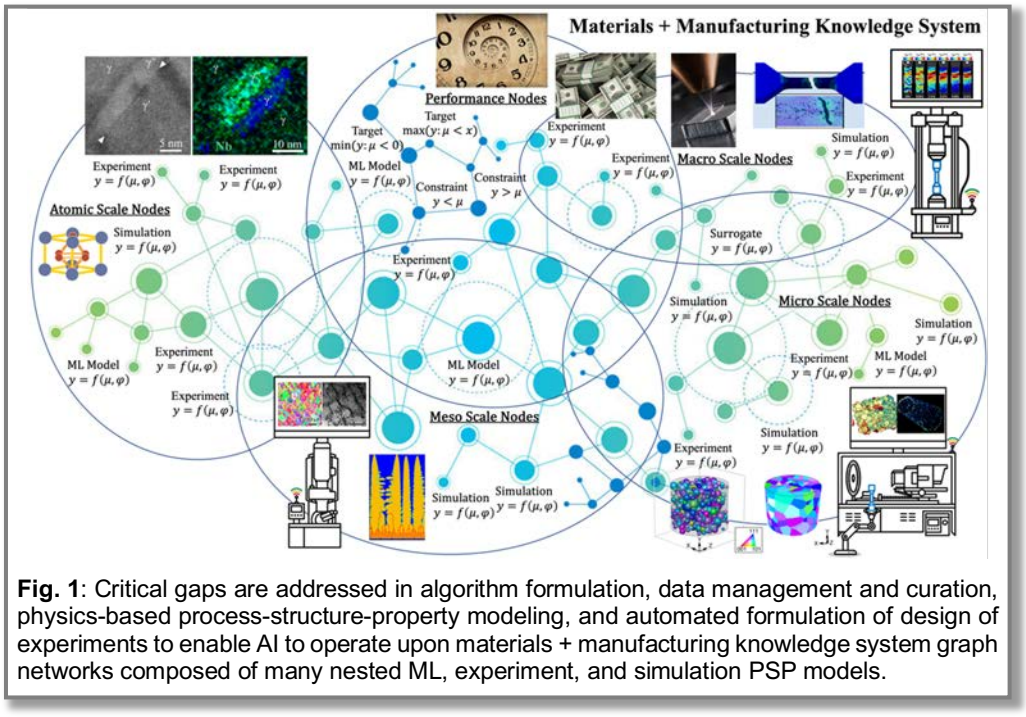


Fig. 1: Critical gaps are addressed in algorithm formulation, data management and curation, physics-based process-structure-property modeling, and automated formulation of design of experiments to enable AI to operate upon materials + manufacturing knowledge system graph networks composed of many nested ML, experiment, and simulation PSP models.

We address a material + manufacturing grand challenge in meeting these algorithmic needs: innovation of an ICME framework to discover and optimize MMCs and their AM. Specifically, AM technologies have created new abilities for the manufacture of MMCs that exhibit superior performances relative to their alloy counterparts and, complementarily, MMCs enable AM of alloys that could not otherwise be processed to high quality. Still, physics-based advancements such as the incorporation of alloy grain inoculation incited by ceramic particles into process-structure models and an ability to correlate thermodynamic phase transformation calculations with AM results must be made to enable data sources to generate meaningful results for AM-MMC discovery and optimization.

Technical Progress

GOALI partner Elementum 3D has sent a historical sample library and the students are developing characterization and data management procedures using that library. Most of the Co-I's met together in February at

Georgia Tech and exchanged initial progress/ideas. The physics-informed search metric for the ML-CALPHAD framework was updated and used to predict 3 new 5000 series aluminum MMC's for laser powder bed fusion AM and the experimental validation is underway, to be completed this summer. An initial screening design of experiments framework has been developed for laser powder blown directed energy deposition, and an accompanying test artifact is being designed, to be completed this summer.

Future Plans

Initial focus is on Al-based MMCs that are commercially available today from our GOALI partner Elementum 3D (E3D). We are demonstrating and verifying the utility of the AI-driven ICME framework in its ability to discover and optimize first within the Al-based class, then move to Ti-based. We focus on Laser Powder Bed Fusion (LPBF) and Laser powder-blown Directed Energy Deposition (LDED) AM technologies. In total, this chosen scope addresses today's state-of-the-art, while still providing both the data and physical challenges required to develop, demonstrate, and assess the proposed advancements in enabling AI-driven parallel design of MMCs and their AM. To ensure pedigree materials performance data for AM-MMCs, test artifacts coupled to statistical designs of experiments are being developed to enable a holistic view of PSP relationships by enabling repeatable analysis of microstructure, residual stress, chemical composition, surface integrity, distortion, internal channel integrity, powder removal testing, geometric accuracy, and thermomechanical properties. Furthermore, the usual relationships mapping $\{y, \mu, \varphi\}$ in collaborative, multiscale materials are established for this program to facilitate human-driven research activities, e.g., CALPHAD and electron microscopy will inform phase field models, phase field simulation and X-ray characterizations will inform FE, etc.

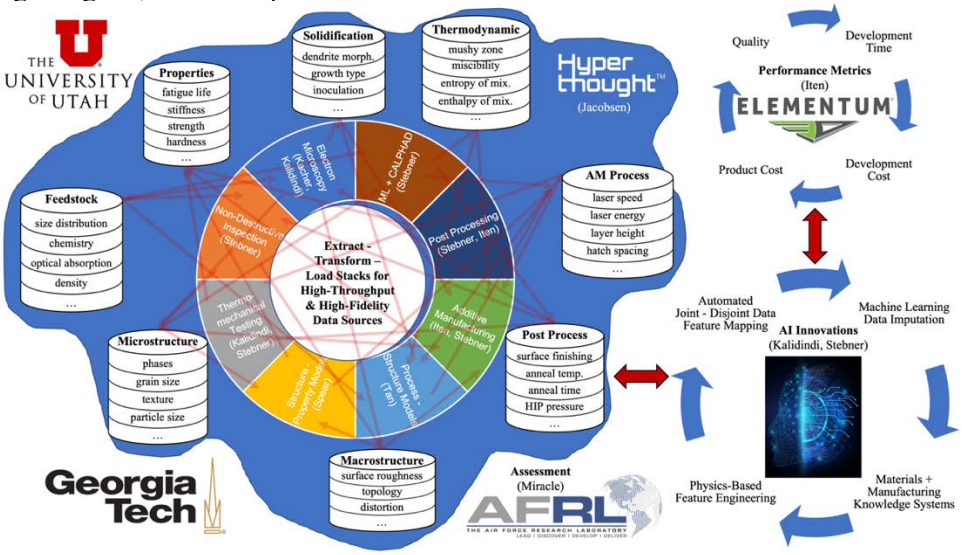


Fig. 2: A work breakdown structure schematic. The transparent arrows mark the myriad of data pipelines that we will integrate into the AFRL HyperThought data lake in developing extract-transform-load (ETL) Stacks for the indicated data sources. AI algorithms innovations will incorporate the performance metrics set by GOALI partner Elementum 3D as optimization constraints. AFRL researchers will assess progress annually.

Broader Impacts and Workforce Development

Four of the five graduate RA positions have been filled, with three of underrepresented gender/race. A new 4-course ML for scientists and engineers course sequence was piloted last year, with the DMREF students (in addition to more than 60 other students) taking the courses and providing feedback, and DMREF PI's developing and teaching 3 of the 4 courses. The Graduate RA's have completed their first year of courses and research training, including a 2-day virtual workshop with the AFRL HyperThought team in January. Students used DMREF data and tasks in their course projects.

Data Management and Open Access

We have deployed a Hyperthought instance for this project and students are working to establish data pipelines this year. Software will be distributed through GitHub repos. All forthcoming and in development.

Advancing Along the Materials Development Continuum and Partnerships to Translation

The work done to tie test artifact design with statistical designs of experiments (DOX) is reducing the number of samples needed to elucidate a PSP design space sufficient for statistical/ML modeling by 30-40%. The work done in high throughput characterization is reducing the time and costs of generating experimental data by more than 60%. Commercial adoption of the test artifact/DOX methodology is being pursued through participation in ASTM standards committees. The GT-Elementum 3D collaboration was awarded an ARO STTR grant for extending these findings to cold spray AM, GT was awarded an AFRL grant for qualification of wire arc AM of aluminum, and the entire DMREF team forms the nucleus of a larger collaborative team proposing to the NASA Space Technology Research Institute for certification of metals additive manufacturing.

Discovery of high-temperature, oxidation-resistant RCCAs

Lead Investigator: Alejandro Strachan, strachan@purdue.edu
Participating Institutions: Purdue University
Website: none.
Keywords: complex concentrated alloys, active learning, high-temperature applications, oxidation resistance

Project Scope

Motivated by the need to metallic alloys for high-temperature applications, we seek to discover refractory complex concentrated alloys (RCCAs) with unsurpassed combination of high-temperature strength and oxidation resistance, Fig. 1. We target: i) a mass gain per area of 1 mg/cm2 after 100 h at 1,000 ºC (an order of magnitude improvement), ii) specific strength 70 MPa.cm^3/g. We propose to explore N-component alloys with Ti, V, Cr, Zr, Nb, Mo, Ru, Hf, Ta, W, and Re as main elements (with N between 3 and 6), and with Al, Si, Sc, and Y as additives to enhance the formation of a protective oxide scale.

Relevance to MGI

We designed a tightly coupled, iterative approach to materials design that combines physics-based simulations, machine learning, and experiments, see Fig. 2. Importantly, the quantities of interest high-temperaturee oxidation resistance and strength) require costly and time-consuming experiments. To address this challenge, we use surrogate properties during our optimization: we restrict our search Al-containing alloys (for oxidation resistance) and test for hardness as a surrogate for strength. Our approach starts with modeling existing experimental results (approximately 300 values) and a design space with ~30,000 alloys. To minimize the number of possible experiments, we used CALPHAD simulations using Thermo-Calc to restrict our search to alloys expected to be single-phase and BCC. This brings down the number of alloys to ~5,000. The existing data from the literature is modeled using gaussian processes regression and we use the maximum expected improvement (MEI) measure to select the next alloys to test experimentally. MEI combines both the average predicted value and its uncertainty to balance exploitation with exploration. In addition, since our experimental capabilities alloy for the processing of batches of 5 alloys in parallel, we use an iterative approach to select 5 alloys that are not close in composition.

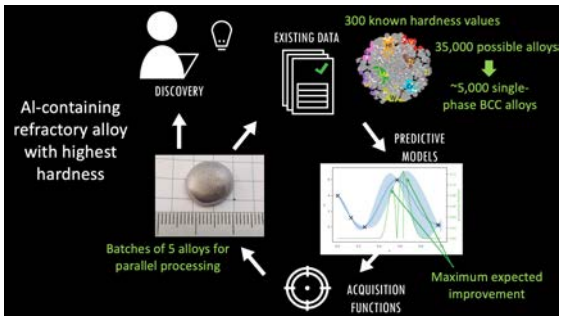


Figure 2. Schematic representation of the active learning loop used for alloy design.

Technical Progress

The results of the active learning approach for alloy design are summarized in Figure 3. We show the measured hardness of each alloy in the five iterative batches. The red horizontal line represents the value of the hardest Al-containing RCCA reported to date, AlMoNbTi, with a value of 510 HV. We can see that even in batch 1, our approach produces alloys harder than the state of the art. The model prediction for each alloy, shown as blue stars, improves in accuracy with batch number, as more data is added to the database. In batch 5, we find an alloy with a hardness of 700 HV, a 30 percent improvement. Importantly, this alloy, $Al_{31}V_{32}Nb_{27}W_{10}$, is the hardest 4-component RCCA regardless of Al content, as can be seen by comparing the blue dashed line that represents the value for a CrNbTiW alloy.

Prior efforts on active learning can also be coupled with physics-based simulations to optimize materials before running experiments. However, prior efforts have been mostly limited to the use of electronic structure calculations and properties that can be obtained at the unit cell level and with negligible noise. We coupled active learning with molecular dynamics simulations to identify multiple principal component alloys (MPCAs) with high melting

temperatures. Building on cloud computing services through nanoHUB, we developed a fully autonomous workflow for the efficient exploration of the high dimensional compositional space of MPCAs. We characterized how uncertainties arising from the stochastic nature of the simulations and the acquisition functions used to select simulations affect the convergence of the approach. Interestingly, we found that relatively short simulations with significant uncertainties can be used to efficiently find the desired alloys as the random forest models used for AL average out fluctuations.

Future Plans

Future efforts will focus on i) further mechanical testing of the alloys discovered via active learning, ii) further developments of synthesis methods that enable grain control, and iii) modeling oxidation resistance.

Mechanical testing. The active learning loop shown in Fig. 2 is close to convergence. We will complete one additional batch and test the strength of selected samples. Due to their high hardness these, single-phase alloys, are expected to be brittle and they will be tested in compression.

Synthesis. We have been using a powder metallurgy method combining Pechini method to XX and magnesiothermic reduction to form RCCAs with controlled grain size. These efforts will be continued and the alloys will be compared with those obtained using arc melting.

Oxidation modeling. We performed an extensive literature search and collected published mass gain data for a range of alloys. As will be discussed below, this data has been analyzed and made FAIR using nanoHUB's Sim2Ls. Continuing efforts will use machine learning tools to develop models that can classify alloys based on their regime of oxidation and regression to predict oxidation resistance.

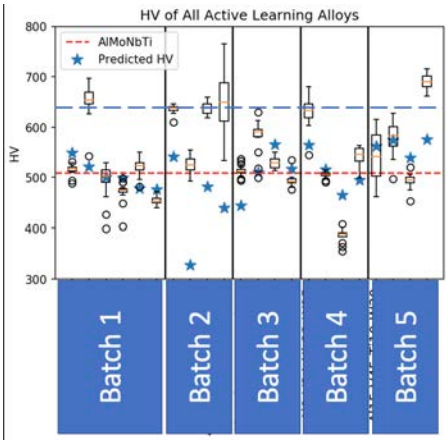


Figure 3. Alloy hardness during active learning loop. Box and whiskers represent experimental results, blue stars represent model predictions.

Broader Impacts and Workforce Development. Our team builds on NSF's nanoHUB to enhance the impact of the research and to share educational material. PI Strachan organizes a series of workshops to introduce machine learning to researchers that has attracted thousands of users. The series can be found in: https://nanohub.org/groups/ml/handsontraining. Co-PI Titus is co-organizing a similar series on CALPHAD methods: https://nanohub.org/groups/tcacademic/thermocalc_workshops. These efforts are in addition to the education of graduate and undergraduate students at Purdue University.

Data Management and Open Access

Our team builds on NSF's nanoHUB to make data and models FAIR:

- Oxidation mass uptake database: https://nanohub.org/resources/refoxdb (DOI: 10.21981/EA6T-BE60)
- Sim2L for melting temperature prediction: https://nanohub.org/resources/meltheas , (DOI: 10.21981/W5VD-T039)
- Active learning workflow: https://nanohub.org/resources/activemeltheas, DOI: 10.21981/NK7E-HA16

Advancing Along the Materials Development Continuum and Partnerships to Translation. A patent was filed on fabrication methods for RCCA and our team is pursuing several interactions with companies and defense labs.

Publications and References

1. K. H. Sandhage, "Methods for fabricating refractory complex concentrated alloys and composites containing such alloys, and bodies containing the same." U.S. Patent Application No. 17/208,424.
2. D. E. Farache, JC Verduzco, ZD McClure, S Desai, A Strachan "Active learning and molecular dynamics simulations to find high melting temperature alloys." Computational Materials Science **209** 111386, (2022).
3. M. S. Farnell, Z. D. McClure, S. Tripathi, A. Strachan, Modeling environment-dependent atomic-level properties in complex-concentrated alloys. The Journal of Chemical Physics, **156**(11), 114102, (2022).
4. Z. D. McClure, A. Strachan, A. Expanding materials selection via transfer learning for high-temperature oxide selection. JOM, 73(1), 103-115, (2021).
5. O.N. Senkov, T.I. Daboiku, T.M. Butler, M.S. Titus, N.R. Philips, and E.J. Payton, "High-Temperature Mechanical Properties and Oxidation Behavior of Hf-27Ta and Hf-21Ta-21X (X is Nb, Mo or W) Alloys," International Journal of Refractory Metals and Hard Materials 96 105467 (2021).

Beyond-DFT Electrochemistry with Accelerated and Solvated Techniques (BEAST)

Lead Investigator: Ravishankar Sundararaman, sundar@rpi.edu
Participating Institutions: Rensselaer Polytechnic Institute, University of Colorado Boulder, National Renewable Energy Laboratory, University of South Carolina, and Lawrence Berkeley National Laboratory.
Website: https://beast-echem.org
Keywords: Electrocatalysis, Solvation, Many-body perturbation theory, Delta learning.

Project Scope

The project aims to address the two key challenges in theoretical understanding of electrochemical reactions: (1) the lack of a universal framework that efficiently treats arbitrary electrolytes, solvents and applied potentials with sufficient detail and fidelity to realistically and accurately model electrochemical systems, and (2) the deficiencies of density functional theory (DFT) in describing charge transfer and reaction barriers. We aim to develop accurate and efficient exascale-ready solvated beyond-DFT methods by (1) advancing electrolyte solvation models, (2) combining them with the random phase approximation (RPA) for beyond-DFT electrochemical calculations and (3) accelerating these techniques using 'Δ-learning' approaches to make RPA-quality predictions at DFT cost, made possible by developing a beyond-DFT electrochemical database.

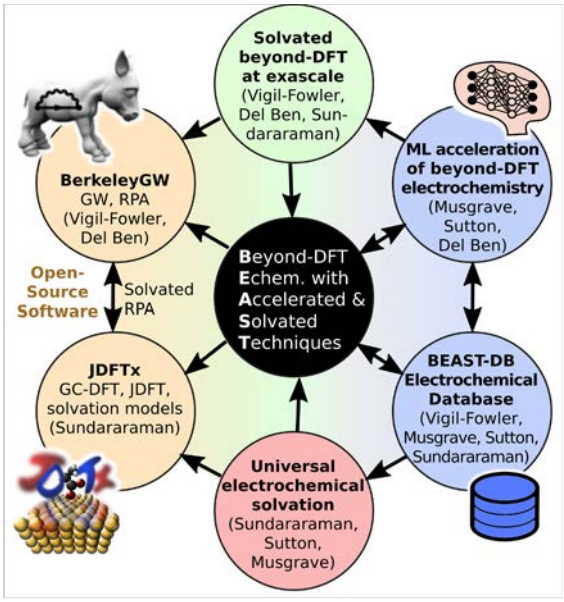


Figure. Universal electrochemical solvation models and beyond-DFT methods, optimized for exascale and accelerated by machine learning, will facilitate accurate beyond-DFT electrochemistry using open-source software, JDFTx and BerkeleyGW.

Relevance to MGI

BEAST integrates theory, computation and data to develop advanced models for accurately describing chemical processes at electrified interfaces and then applies these methods to generate an extensive data base for electrocatalytic reactions. While MGI approaches have successfully accelerated the discovery new materials for many applications, its application to identifying new materials in electrochemical systems has been hindered by the lack of methods for describing processes at electrified interfaces. However, the lack of a universal first-principles computational framework that efficiently treats solvents and applied potentials to realistically and accurately model electrochemical systems and DFT errors in describing charge transfer and reaction barriers have hindered the application of the MGI approach to electrochemical materials. In this project, we focus on materials for electrocatalysis, but the methods we develop will also be applicable to the discovery of materials for other electrochemical systems. Feedback from experiment is obtained from ongoing collaborations and interactions facilitated by our comprehensive data base of electrocatalytic reactions.

Technical Progress

1. Towards our goal of automatically generating classical DFT solvation for arbitrary electrolytes, we have advanced on: (1) automatically generating machine-learned (ML) force fields from ab initio molecular dynamics (AIMD) simulations, and (2) extracting free energies and density response of inhomogeneous fluids from force fields, needed to train ML classical DFTs.

2. To enable beyond-DFT solvated calculations at exascale, we have focused on (1) optimizing the JDFTx-BGW interface, and (2) developing implementation strategies of cubic-scaling RPA and RPA forces, with initial testing where possible. Additionally, to enable readiness on leadership-class HPC architectures, we have targeted computational resources with a strong focus on GPU capabilities.
3. Towards our goal of RPA-quality electrochemical predictions at DFT cost by Δ learning, we have: (1) generated GC-DFT data on a range of catalysts and chemistries, (2) developed electronic descriptors for learning single-particle energies beyond semilocal DFT, and (3) performed initial design and detailed implementation plan for the BEAST-DB.

Future Plans

- Using ML force fields, we will generate inhomogeneous electrolyte response data to train ML classical DFT solvation models. We will then automate our active learning scheme to generate ML force fields for complex electrolytes, eventually feeding into a universal ML classical DFT solvation workflow.
- We will expand the first-principles electrochemical capabilities to take advantage of GPU supercomputers, efficiently and at a larger scale. We will optimize the prefactor in cubic-scaling RPA and implement RPA forces based on initial theoretical formulations, creating a path to practical RPA calculations of electrochemical systems.
- We will rapidly expand our initial grand-canonical DFT dataset to include additional bulk materials, facets and reactant species for greater coverage of electrochemical reactions in BEAST-DB. We will generalize electronic descriptors and Δ-learning models from initial bulk prototypes to surface+adsorbate calculations typical for electrochemistry. We will deploy these data sets and algorithms into BEAST-DB with initial demonstrations targeted for the second BEAST workshop in 2023.

Broader Impacts and Workforce Development

The project involves close interdisciplinary collaboration between five organizations, facilitating close interactions between 10 graduate students and 3 postdocs in fields spanning chemistry, chemical engineering, materials science, physics, electrical engineering and high-performance computing, providing mutual training in fields they would not have encountered in core research-group projects. Planning is underway for the first annual BEAST workshop on August 15-16 2022 (virtual format), which will provide training in first-principles electrochemistry and beyond DFT methods for around 100 early career researchers. The BEAST project will organize in-person workshops for around 50 early-career researchers in summer 2023 and 2025, in addition to another virtual workshop with 100 participants in summer 2024.

Data Management and Open Access

We have created https://beast-echem.org with a research overview, publications and software resources as a central repository for disseminating all results of this program. This will include BEAST-DB containing all our grand-canonical DFT and RPA simulations of electrochemical systems, searchable by reaction mechanism, catalyst material/facet/adsorption site and applied bias, once it becomes ready for public access starting in 2023. All computational and methodological developments are being continuously deployed through the open-source JDFTx (http://jdftx.org) and BerkeleyGW (http://berkeleygw.org) codes. The usage of both the database and the codes will be disseminated to the community through our annual BEAST workshops.

Publications and References

1. R. Sundararaman, D. Vigil-Fowler and K. Schwarz, "Improving the Accuracy of Atomistic Simulations of the Electrochemical Interface", *Chem. Rev.* **122**, 10651 (2022)
2. A. Shandilya, K. Schwarz and R. Sundararaman, "Interfacial water asymmetry at ideal electrochemical interfaces", *J. Chem. Phys.* **156**, 014705 (2022)
3. Y. A. Alsunni, A. W. Alherz, and C. B. Musgrave, "Electrocatalytic Reduction of CO2 to CO over Ag(110) and Cu(211) Modeled by Grand-Canonical Density Functional Theory", *J. Phys. Chem. C* **125**, 23773 (2021)

SPARC-X: Quantum simulations at extreme scale — reactive dynamics from first principles

Lead Investigator: Phanish Suryanarayana, phanish.suryanarayana@ce.gatech.edu
Participating Institutions: Georgia Institute of Technology, Lawrence Livermore National Laboratory
Website: http://www.sparc-x.com
Keywords: Kohn-Sham, Density functional theory, Electronic structure, Real-space, Software

Project Scope

The objective of this project is to develop SPARC-X: an open source computational framework for performing Kohn-Sham Density Functional Theory (DFT) calculations, including those that scale linearly with system size, which can leverage petascale/exascale parallel computers to study chemical phenomena at length and time scales previously accessible only by empirical approaches — e.g., 100,000 atoms for a few picoseconds using semilocal functionals or 1,000 atoms for a few nanoseconds using hybrid functionals.

Relevance to MGI

Over the course of the past few decades, quantum mechanical calculations based on Kohn–Sham DFT have become a cornerstone of materials research by virtue of the predictive power and fundamental insights they provide. The tremendous popularity of DFT can be attributed to its generality, simplicity, and highly favorable accuracy-to-cost ratio relative to other such *ab initio* theories. However, Kohn-Sham calculations are associated with large computational expense, which severely limits the range and type of physical systems that can be investigated. In particular, the computational design process for materials is severely restricted by the large time to solution associated with such simulations. The objective of this project is to develop a computational framework that accelerates Kohn-Sham DFT calculations by orders of magnitude (Fig. I), which would make it an attractive choice for use in frameworks targeted towards the design of new materials with tailored properties.

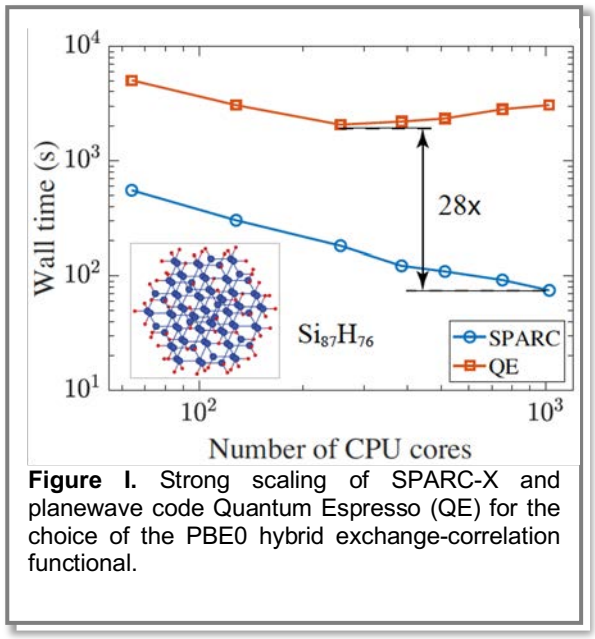


Figure I. Strong scaling of SPARC-X and planewave code Quantum Espresso (QE) for the choice of the PBE0 hybrid exchange-correlation functional.

Technical Progress

In this project, an open-source, real-space density functional theory (DFT) code named SPARC [1, 5] has been developed. It accommodates both Dirichlet and Bloch-periodic boundary conditions, enabling the treatment of finite, semi-infinite, and charged systems, as well as bulk 3D systems. Current features of SPARC include: (i) applicable to isolated systems such as molecules as well as extended systems such as crystals, surfaces, and wires; (ii) local, semilocal, and hybrid exchange-correlation functionals; (iii) dispersion interactions through DFT-D3, vdW-DF1, and vdW-DF2; (iv) norm conserving pseudopotentials, including nonlinear core corrections; (v) spin polarized and unpolarized calculations; (vi) spin-orbit coupling; (vii) calculation of ground state energy, atomic forces, and stress tensor; (viii) structural relaxation and ab initio molecular dynamics (NVE, NVT, and NPT), (ix) linear-scaling, i.e., O(N), Spectral Quadrature (SQ) method; (x) Discrete Discontinuous Basis Projection (DDBP) method; and (xi) MATLAB version available for rapid prototyping: M-SPARC [7].

SPARC is not only portable, but is also straightforward to install, use, and modify, with external dependencies limited to industry standard MPI, BLAS, and LAPACK/ScaLAPACK. It has been extensively benchmarked and validated against well-established planewave codes, where SPARC has shown to be an order of magnitude faster in time to solution, with increasing advantages as the number of processors is increased (Fig. I). It can efficiently utilize modest as well as large-scale computational resources, with parallel scaling bringing solution times to a few seconds for O(100-500) atoms and about a minute for systems with O(500-1000) atoms. Using the O(N) SQ method, SPARC has been scaled to system sizes containing more than a million atoms (Fig. II). SPARC is freely available at: https://github.com/SPARC-X/SPARC.

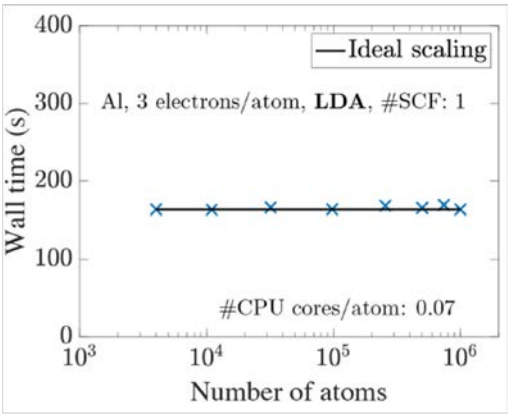


Figure II. Weak scaling of the SQ method in SPARC for a metallic system on Quartz computer at LLNL [1].

Future Plans

Key developments targeted in SPARC include: (i) O(N) DFT with substantially reduced prefactor using the DDBP+SQ method; (ii) reduced scaling density functional perturbation theory (DFPT); (iii) reduced scaling many-body random phase approximation (RPA) for the correlation energy; and (iv) on-the-fly machine learned force fields (MLFF) to further increase the length and time scales accessible.

Broader Impacts and Workforce Development

All codes and data are made freely available at https://github.com/SPARC-X. The project provides a unique opportunity to conduct research and mentor/train postdoctoral scholars as well as graduate students in an area which is at the interface of chemistry, applied mathematics, physics, materials science, and high performance computing.

Data Management and Open Access

All codes and data made freely available at https://github.com/SPARC-X.

Advancing Along the Materials Development Continuum and Partnerships to Translation

No plans for commercialization.

Publications and References

1. P. Suryanarayana, Q. Xu, and J. E. Pask, *Roadmap on Electronic Structure Codes in the Exascale Era*, Modelling Simul. Mater. Sci. Eng. Submitted (2022)
2. S. Shah, P. Suryanarayana, and E. Chow, *Pseudodiagonalization Method for Accelerating Nonlinear Subspace Diagonalization in Density Functional Theory*, J. Chem. Theory and Comput., **18**, 6, 3474–3482 (2022)
3. Q. Xu, X. Jing, B. Zhang, J. E. Pask, and P. Suryanarayana, *Real-space density kernel method for Kohn–Sham density functional theory calculations at high temperature*, J. Chem. Phys., **156**, 094105 (2022)
4. S. J. Sahoo, X. Jing, P. Suryanarayana, and A. J. Medford, *Ab-Initio Investigation of Finite Size Effects in Rutile Titania Nanoparticles with Semilocal and Nonlocal Density Functionals*, J. Phys. Chem. C, 126, **4**, 2121(2022)
5. Q. Xu, A. Sharma, B. Comer, H. Huang, E. Chow, A. J. Medford, J. E. Pask, and P. Suryanarayana, *SPARC: Simulation Package for Ab-initio Real-space Calculations*, SoftwareX, **15**, 100709 (2021)
6. C. M. Diaz, P. Suryanarayana, Q. Xu, T. Baruah, J. E. Pask, and R. R. Zope, *Implementation of Perdew--Zunger self-interaction correction in real space using Fermi--Lowdin orbitals*, J. Chem. Phys., **154**, 084112 (2021)
7. Q. Xu, A. Sharma, and P. Suryanarayana, *M-SPARC: Matlab-Simulation Package for Ab-initio Real-space Calculations*, Software X, **11**, 100423 (2020)

DMREF: Collaborative Research: Fundamentals of short-range order-assisted alloy design: Thermodynamics, kinetics, mechanics

Lead Investigator: Cem Tasan, tasan@mit.edu
Participating Institutions: Massachusetts Institute of Technology, Oak-Ridge National Laboratory
Website: None
Keywords: Short-range order, alloy, mechanical property, martensitic transformation, stacking faults.

Project Scope

The goal of the project is to contribute to the fundamental understanding of short-range ordering (SRO) in metallic materials, and to identify the design rules for the effective utilization of SRO. SROs commonly develop in complex-concentrated alloys (CCAs) owing to the complex compositions, and cause detrimental effects on the mechanical performances of alloys, by increasing slip planarity. We have shown that SROs can lead to new deformation mechanisms and unexpected improvements in mechanical properties, in conjunction with the stacking-fault energy (SFE)-related effects. We aim to further address the complex interplay between SRO and SFE systematically, through integrated theoretical-experimental investigations.

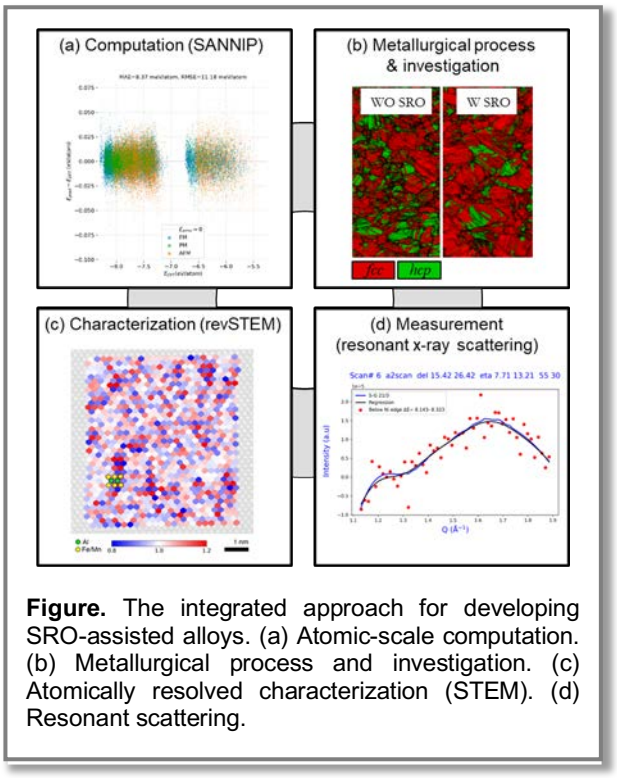


Figure. The integrated approach for developing SRO-assisted alloys. (a) Atomic-scale computation. (b) Metallurgical process and investigation. (c) Atomically resolved characterization (STEM). (d) Resonant scattering.

Relevance to MGI

The small dimensions of SROs make them difficult to detect by conventional transmission electron microscopy or atom probe tomography methods. This results in the lack of understanding of SROs, as well as, in the absence of guidelines to design materials that benefit from SRO-induced effects. This project seeks to address these gaps by following an integrated approach combining theoretical modeling, metallurgical processing (fabrication & testing), and atomically-resolved revolving scanning transmission electron microscopy (revSTEM) techniques. To this end, the candidate CCAs are designed using computational methods with benchmarked criteria. The alloys are processed, and the mechanical properties are measured by conventional metallurgy methods. The deformation behaviors are analyzed by integrated scanning electron microscopy techniques such as electron channeling contrast imaging and micro-digital image correlation, and in-situ synchrotron tensile tests. The corresponding atomistic mechanisms are investigated by revSTEM experiments and atomistic simulations.

Technical Progress

To achieve these goals, several collaborative activities have been initiated, including:

Spin-aware neural network interatomic potential for predicting SROs: We have been developing a robust framework for spin-aware neural network interatomic potential (SANNIP), to efficiently and accurately model both the thermodynamics and kinetics of SRO in CCAs. A workflow has been developed to carry out Monte Carlo and molecular dynamics simulations with the new potential, and the potential is validated on Fe on conventionally challenging problems such as Curie temperature.

Beneficial effects of SRO in low SFE CCAs: We have designed new CCAs with low SFEs by thermodynamic calculations and showed that SRO can effectively manipulate the transformation-induced plasticity behaviors therein. SRO induces discrete stacking faults and suppressed twinning and martensitic transformations that are

prone to cracking, while preserving tensile properties. The effect of SRO on dislocation is discussed by measuring dislocation activation volume.

Atomically resolved SRO characterization by STEM: We have successfully observed SRO in FeMnAlC and NiCoCr CCAs by revSTEM. Progress in analysis techniques has been made to provide evidence regarding planar shear of SRO theory through cluster size measurement and comparison. 4D-STEM imaging and Cepstral analysis at low magnification showed significant and correlated strains with SRO in a deformed sample. Additionally, analysis of integrated differential phase contrast data and STEM image simulations shows the ability for quantifying carbon interstitial content based on linear potential dependency that can be correlated with substitutional SRO.

SRO-dislocation interaction: We have measured thermal activation volumes of dislocations in various CCAs by performing stress-relaxation tests. The measured activation volumes are correlated with the average distance between SROs characterized by revSTEM. The origin of the correlation between activation volume and SRO effects have been investigated by atomic simulations. The atomic configurations for the simulations are structured considering the atomically resolved revSTEM results.

Resonant X-ray scattering for quantitative measurement of SROs: It is difficult to measure SROs by scattering because of weak x-ray contrast of constituent elements and diffuse nature of the scattering signal. We have carried out x-ray resonance near the x-ray absorption edge to enhance the contrast and observe SROs. As a test case we studied a NiV alloy at the Ni edge and found broad chemical SRO peaks ½(1 1 1) at 1.5 $Å^{-1}$ and (1 0 0) at 1.7 $Å^{-1}$.

Future Plans

In the fourth year, we are planning to focus on 1) extension of the first-principles calculations and neural network interatomic potential training to model CCAs, to enable large-scale, quantum-mechanical accurate molecular dynamics simulations on SRO effects, 2) investigation of SRO effects on CCAs with low SFE in other environments and find further alloy design criteria, 3) application of STEM methodologies for other CCAs, 4) application of resonant X-ray scattering for CCAs with more than three components.

Broader Impacts and Workforce Development

The project has helped bridging different fields, i.e. physical metallurgy, solid mechanics, electron microscopy, through its common goal of designing novel multi-component alloys. The overlapping findings have been communicated to a broader audience in various conferences. Moreover, the project has provided a number of training opportunities to researchers of all stages, on alloy design, thermomechanical processing, characterization, artificial neural network learning, first principle and molecular dynamics simulations, and STEM imaging.

Data Management and Open Access

We are committed to sharing data and software expeditiously and efficiently with the broader community through publications and presentations, and by making the data leading to those publications and presentations available via the internet for verification and reproduction. This project generates various data from experiments to simulations. These data include microstructure images, strain maps, diffraction, scattering, crystallographic data, first-principles calculations, molecular dynamics simulations, and neural network training. All digital data are stored on university servers at the host institutions and backed up following standard practices, which provides full access to data, metadata, and software associated with published research results to potentially interested researchers and educators. Finally, a permanent URL forwarding address (http://alum.mit.edu/www/liju99) has been established that will collect and preserve these resource links for simulation data.

Advancing Along the Materials Development Continuum and Partnerships to Translation

The proposed concept regarding metastability engineering by SRO has potential for commercialization. This, however, would be difficult without following the MGI approach, given SRO's atomic-scale dimensions. We are planning to initiate a collaboration with Army Research Laboratory to find further applications of the developed alloys or to apply the design concept for existing steels for better wear and damage resistance.

Publications and References

1. H.S. Oh, K. Odbadrakh, Y. Ikeda, S. Mu, F. Körmann, C-J. Sun, H. S. Ahn, K. N. Yoon, D. Ma, C. C. Tasan, T. Egami, E. S. Park, "Element-resolved local lattice distortion in complex concentrated alloys: An observable signature of electronic effects", *Acta Materialia*, 216, 117135, 2021.

Accelerating Thermoelectric Materials Discovery via Dopability Prediction

Lead Investigator: Eric Toberer, Colorado School of Mines, Golden, CO, etoberer@mines.edu.
Co-Principal Investigator: G. Jeffrey Snyder, Northwestern University, Evanston, IL
Co-Principal Investigator: Elif Ertekin, University of Illinois at Urbana-Champaign, Urbana, IL
Co-Principal Investigator: Vladan Stevanovic, Colorado School of Mines, Golden, CO
Co-Principal Investigator: Michael Toney, University of Colorado Boulder, Boulder, CO
Website: None
Keywords: thermoelectrics, semiconductors, dopability, transport, alloys

Project Scope

The efficiency of thermoelectric materials peaks at an optimal carrier concentration, thus it is critical to effectively control carrier concentration in these semiconductors through doping. In fact, the dopability of semiconductors is often a bottleneck to the realization of high-performing thermoelectric materials, even when intrinsic properties are optimized (Figure 1). The goal of this project is to discover a new set of high-performing thermoelectric materials by computationally and experimentally evaluating dopability in semiconductor materials, in addition to evaluating thermal and electronic transport properties. Specifically, we aim to (1) develop a computational dataset with experimentally validated dopability assessments, (2) develop new synthesis techniques to experimentally assess dopability, and (3) develop predictive models of dopability and thermoelectric transport.

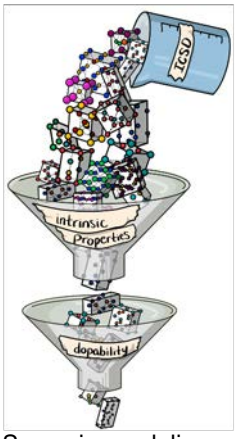


Figure 1. Screening and discovery of new thermoelectric materials by evaluating intrinsic transport properties and dopability.

Relevance to MGI

This project combines computation, theory, and experiment for more effective prediction and engineering of high-performance thermoelectric materials. In many cases, theory has played a critical role in identifying effective dopants. For instance, theoretical approaches helped eliminate the formation of compensating defects and identified extrinsic dopants that would not have been otherwise considered (or could occupy more than one lattice site). Such theoretical analysis has been applied to Mg_3Sb_2 and diamond-like semiconductors. Moreover, experimental results have guided theoretical and computational efforts, leading to the improvement of computational models. Experimental studies in diamond-like semiconductors showed the need to improve phase stability predictions in these compounds. Furthermore, experimental findings have motivated new paths for computational study. For example, the discovery of the ordered vacancy compound, Hg_2GeTe_4, motivated the computational study of high dopability in ordered vacancy compounds.

Technical Progress

In terms of materials search and screening, numerous new thermoelectric compounds were discovered/studied. For instance, the chemical replacements in structural prototype (CRISP) approach was used to predict 5 n-type ABX Zintl compounds and 3 n-type ABX_4 compounds with promising thermoelectric properties. Additionally, we developed an expanded set of dopability predictions in complex diamond-like semiconductors and undertook a systematic analysis of dopability in ternary diamond-like materials. Finally, by extending the half-Heusler compositional space to include quaternary compounds (instead of only ternary compounds), over 100 new half-Heusler thermoelectric compounds were predicted with reduced lattice thermal conductivity compared to the ternary compounds.

Significant progress has been made towards developing predictive models for dopability and transport. We

developed a comprehensive model for native defects and dopability limits using a tight-binding description of a binary ionic semiconductor. Furthermore, we used analytic models to reproduce experimental trends in the thermal conductivity of alloys with respect to point-defect scattering. This work gives insight into methods to reduce the lattice thermal conductivity of materials via doping and alloying. We also applied the tight-binding model to understand band engineering strategies in PbTe (and other IV-VI rock salt semiconductors) and how alloying/doping can be used to optimize their performance. This work led to the prediction of possible Fermi surface topologies in these IV-VI compounds, which could lead to greater thermoelectric performance if experimentally realized. Lastly, we used analytical models for band-gap prediction to better understand thermoelectric transport in VFe_2Al. Accurate band gap estimates are crucial for defect engineering.

We also developed new synthetic techniques, especially for phase boundary mapping, and experimentally assessed dopability. For instance, a study of Mg_2Si-Mg_2Sn alloys, an important system for high-temperature n-type thermoelectrics, has elucidated miscibility conditions in this system. Additionally, phase boundary mapping in the Pb-Te-I system has shown that to successfully n-type dope PbTe with iodine, it is necessary to synthesize PbTe in Pb-rich conditions. We also experimentally studied the dopability of a variety of compounds. For example, the extrinsic dopability of 1-1-4 ($KGaSb_4$ prototype) Zintl compounds was assessed to find optimal n-type dopants and evaluate the predictions from defect calculations.

Future Plans

In the future, we will continue publishing our project findings and applying our new understanding from these studies to new projects. Phase boundary mapping can be applied to understand the complex phase space of 14-1-11 Zintl compounds. Additionally, the computational analysis of the band structures of IV-VI semiconductors (PbTe, SnTe, GeTe, etc.) is being applied to understand the chemistry and physics of topological insulators and to understand how changes to the crystal symmetry (e.g., rhombohedral distortion of cubic phases) influences the electronic band structure of these compounds. Moreover, the CRISP approach is being used to discover promising new materials in BiCuSeO-related systems that can be successfully.

Broader Impacts and Workforce Development

Throughout this project, there have been numerous efforts to engage the broader scientific community and to train the next generation of scientists and engineers. In 2020, Toberer was a co-organizer of the first Virtual Conference on Thermoelectrics (VCT2020) after the International Conference on Thermoelectrics was cancelled due to the COVID-19 pandemic. There were over 900 registered participants and 300 oral presentations delivered by students and postdocs. Toberer was also the lead organizer of the 2019 North American Solid State Chemistry Conference, which included a separate summer school, led by Stevanovic, which focused on defects and dopability in materials. Ertekin gave a tutorial at the Fall 2019 annual workshop of the Materials Computation Center (MCC) at the University of Illinois. This workshop taught about 60 participants (primarily graduate students from other institutions) about the MGI approach to computationally driven discovery of functional materials.

Data Management and Open Access

Our team has created a data format that is constructed such that useful defect properties can be calculated. This data format is compatible with an open access Python toolkit (VTAnDeM) that can perform the necessary calculations (https://github.com/ertekin-research-group/VTAnDeM).

Advancing Along the Materials Development Continuum and Partnerships to Translation

Radioisotope thermoelectric generators have been used as power sources for NASA space missions for several decades, but these systems are still limited by the efficiency of the thermoelectric materials. This work can have direct impacts improving the performance of these systems, and there is ongoing collaboration between the PI's of this project and NASA-JPL to implement these improvements. The findings of this project also have the potential to be employed in commercial thermoelectric applications for industrial waste-heat recovery and solid-state refrigeration. For instance, half-Heusler thermoelectrics are desirable for commercial applications because they are typically made of low-cost, non-toxic materials. Therefore, the computational prediction of over 100 potential quaternary half-Heusler could lead to advances in industrial-scale application of thermoelectrics.

Publications and References

1. M. Y. Toriyama, M. K. Brod, L. C. Gomes, F. A. Bipasha, B. A. Assaf, E. Ertekin, G. J. Snyder, Tuning valley degeneracy with band inversion. *J. Mater. Chem. A* ***10***, 1588-1595 (2022).
2. M. Y. Toriyama, A. M. Ganose, M. Dylla, S. Anand, J. Park, M. K. Brod, J. M. Munro, K. A. Persson, A. Jain, G. J. Snyder, How to analyse a density of states. *Materials Today Electronics* ***1***, 100002 (2022).
3. S. Guo, S. Anand, M. K. Brod, Y. Zhang, G. J. Snyder, Conduction Band Engineering of Half-Heusler Thermoelectrics Using Orbital Chemistry. *J. Mater. Chem. A* ***10***, 3051-3057 (2022).
4. S. Anand, C. Wolverton, G. J. Snyder, Thermodynamic Guidelines for Maximum Solubility. *Chem. Mater.* ***34***, 1638-1648 (2022).
5. L. Borgsmiller, M. T. Agne, J. P. Male, S. Anand, G. Li, S. I. Morozov, G. J. Snyder, Estimating the lower-limit of fracture toughness from ideal-strength calculations. *Mater. Horiz.* ***9***, 825-834 (2022).
6. G. J. Snyder, A. Pereyra, R. Gurunathan, Effective mass from Seebeck coefficient. *Adv. Funct. Mater.* ***32***, 2112772 (2022).
7. B. L. Levy-Wendt, B. R. Ortiz, L. C. Gomes, K. H. Stone, D. Passarello, E. Ertekin, E. S. Toberer, M. F. Toney, D. Collaboration, Understanding Cu incorporation in the Cu 2 x Hg 2− x GeTe 4 structure using resonant x-ray diffraction. *Physical Review Materials* ***5***, 015402 (2021).
9. X. Zhang, Z. Wang, B. Zou, M. K. Brod, J. Zhu, T. Jia, G. Tang, G. J. Snyder, Y. Zhang, Band Engineering SnTe via Trivalent Substitutions for Enhanced Thermoelectric Performance. *Chem. Mater.* ***33***, 9624-9637 (2021).
10. X. Zhang, M. Y. Toriyama, J. P. Male, Z. Feng, S. Guo, T. Jia, Z. Ti, G. J. Snyder, Y. Zhang, First principles investigation of intrinsic and Na defects in XTe (X=Ca, Sr, Ba) nanostructured PbTe. *Materials Today Physics* ***19***, 100415 (2021).

DMREF: Accelerated discovery of metastable but persistent contact insecticide crystal polymorphs for enhanced activity and sustainability

Lead Investigator: Mark E. Tuckerman, mark.tuckerman@nyu.edu
Participating Institution: New York University
Website: crystalmath.io (Under Construction)
Keywords: molecular crystals, phase transformations, metastability, crystal engineering, infectious diseases

Project Scope

Contact insecticides are an essential component in the fight against vector-borne infectious diseases, such as malaria. When mosquitos contact small crystalline particles of an insecticide they absorb the active substance through their feet to lethal effect at the target site. The NYU investigators recently discovered that insecticides readily form crystal polymorphs with efficacies that are inversely correlated with their thermodynamic stability. Their use as insecticides, however, requires kinetic stability against transformation into their thermodynamically more stable but less active forms. This project aims to develop an integrated experimental-theoretical framework for the accelerated discovery of crystalline forms that meet these criteria.

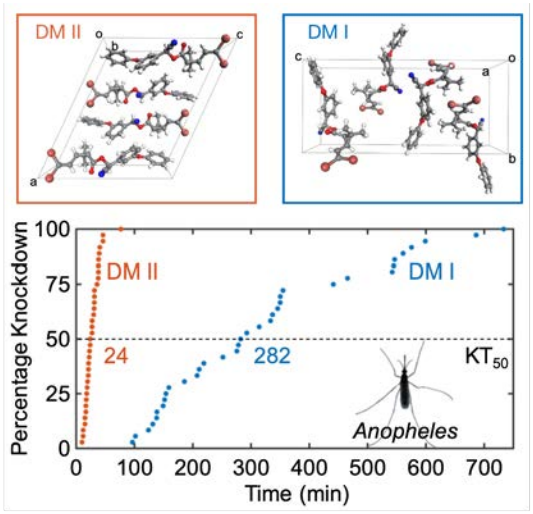


Figure. *Top*: Crystal structures of two polymorphs of deltamethrin (DM, one of the foremost pyrethroid insecticides). *Bottom*: Knockdown times using dusts containing DM I and II. Our newly engineered DM II is roughly twelve times faster acting than the commercially available DM I. (Yang *et al.*, PNAS 117, 26633 (2020))

Relevance to MGI

The project relies on a continuous and mutual feedback among theory, computation and experiment in three key areas: (*i*) discovery of thermodynamically accessible polymorphs, (*ii*) assessment of kinetic stability and transformation pathways, and (*iii*) evaluation and control of surface properties. Initially, computational results will be validated by existing experimental data and the theoretical framework subsequently improved. Theoretical predictions will then inform the experimental search for new polymorphs and crystallization protocols. New findings will reinforce the tandem theoretical and experimental approaches. The contact insecticides investigated in this project comprise different classes of molecules, with common structural features within each class. Through insights from theory, computation and experiment the project aims to identify correlations between these features to identify thermodynamically accessible polymorphs and their relative kinetic stabilities against phase transformations. The integrated experimental and theoretical framework will be developed, validated and refined, initially for a subset of molecules, followed by an optimized workflow for the accelerated exploration of a larger library of target compounds.

Technical Progress

The project start date was April 1, 2022. The project team has initiated the computational approach for evaluating and refining empirical force fields needed for large-scale simulations of polymorphic transitions and crystallization. A particular focus is the accurate description of relative energies between different polymorphs, the energy landscape of conformers of each molecule, and the energetic barrier for transitions between conformers, which is regarded as a critical property with respect to phase transformations. The team is computationally and experimentally investigating the effect of doping crystals of imidacloprid, a widely used neonicotinoid contact

insecticide with nine polymorphs, with its olefin metabolite. The metabolite is a structural mimic of imidacloprid, differing only by replacement of the methylene groups in the imidazoline ring by an olefinic C=C bond. The crystallographic lattice constants for imidacloprid Form IV and this "IMI-olefin" are nearly identical, suggesting the likelihood of solid solution that may favor the occurrence of the metastable Form IV. Preliminary DFT calculations of the torsional profile for IMI-olefin reveal that the energy penalty for the highest energy conformation in Form IX (one of four crystallographically independent molecules) is less for IMI-olefin than for IMI itself, suggesting facile incorporation of the IMI-olefin. It is reasonable to suggest that the resulting lower free energy would stabilize Form IX against transformation to the thermodynamically favored Form I.

Future Plans

The initial stage of the project will involve development and testing of the computational framework accompanied by experimental screening of contact insecticide polymorphs to address (*i*) control of the crystallization of specific metastable polymorphs and (*ii*) stabilizing metastable polymorphs against transformation into less active forms. Theory and computation will be used to chart the microscopic mechanisms of crystallization with respect to the competition between different crystalline polymorphs and amorphous phases as well as their relative stabilities, which will inform the experimental search for new polymorphs. Specifically, the research team will use simulations to evaluate the structural and dynamical properties of molecules in the melt, which is central in understanding the nucleation of different crystalline phases. This will include the effect of temperature and pressure as well as the presence of solid-liquid interfaces, including small seeds of different polymorphs and amorphous structures. The resulting growth simulations will be compared with crystallization of specific polymorphs under various conditions, ranging from growth in solution to melt growth, including seeded growth. Prediction of the relative kinetic stabilities of different polymorphs will rely on characterization of the phase transformation mechanisms and dynamics. A computational approach to sample solid-solid transformations in molecular crystals will be developed, initially through exploration of the structure of interfaces between different polymorphs. Transformation pathways will be sampled by establishing a combined machine-learning/molecular simulation framework to drive structural transformations through changes in local structural motifs. Analysis of transformation paths will be combined with experimental observations of phase behavior for polymorphs of a particular compound to fully characterize dynamics.

Broader Impacts and Workforce Development

Undergraduates, graduate students, and postdocs will benefit from an interdisciplinary training that will create reciprocal awareness in the tools and language of materials modeling, theory, synthesis and characterization. This is facilitated through biweekly meetings fostering the continuous exchange between researchers on all levels and in different fields. In partnership with the Scientific Frontiers Program at NYU the team will design K-12 educational initiatives, hosting students and teacher on site for hands-on workshops including both computation and experiments revolving around crystal growth. Summer research experiences for Black, Latino, and Native American students enrolling in the Collegiate Science Technology Entry Program (CSTEP) at NYU will be designed to cultivate interest in crystal growth and materials science in general. The team also will host female high school juniors enrolled in the NYU GSTEM program for real-world projects that highlight machine learning, molecular simulations and materials discovery.

Data Management and Open Access

The Data Management Plan for this project is designed to ensure the accurate recording and the integrity of data during the award period and beyond and adhering as closely as possible to the FAIR data principles. All code developed within the project will be made available to the community as open-source software, and all simulation and experimental data will be archived and made available in public repositories.

Advancing Along the Materials Development Continuum and Partnerships to Translation

The development of improved contact insecticides by introducing new active ingredients is a lengthy and complex process as it requires a careful assessment of the pharmacology and toxicology for every new compound. The project aims to leverage its strength in crystal engineering to increase the efficacy of contact insecticides through control of crystal structure, obviating the need to create new compounds. In a broader context, the design of metastable crystalline forms has potential applications in various technologically relevant areas including organic electronics, pharmaceuticals, and energetic materials.

Publications and References

The start date for this project was April 1, 2022. Publications acknowledging the project are not available at this time.

Numerically-Exact Relativistic Many-Body Electronic Structure of Heavy Elements

Lead Investigator: Edward F. Valeev, Virginia Polytechnic Institute and State University
Participating Institutions: Stony Brook University, University of Tennessee
Website: none
Keywords: Heavy elements, many-body methods, relativistic electronic structure

Project Scope

The vast importance of heavy elements for energy production and storage, classical and quantum information processing, and enabling chemical transformations, as well as the environmental impacts thereof, underscores the critical need for understanding and controlling the electrons in heavy-element-containing molecules and materials. Unfortunately, our ability to predict the behavior of electrons in heavy-element matter is severely hampered by the limitations of the traditional numerical representations used to describe relativistic quantum states of electrons in heavy elements. This project will remove the bottlenecks in modeling the electrons in such systems predictively by employing non-traditional real-space numerical representations.

To achieve these goals this project's team will build on their recent breakthroughs in the real-space numerical simulation of electrons that demonstrated its unrivaled ability to model electrons with better numerical precision than achievable by fixed atomic orbital basis sets that are typically employed in simulating electrons in molecules. This project will extend these efforts to the simulation of correlated electronic motion in heavy-element compounds. For the first time these developments will permit to approach a nearly exact numerical description of relativistic electrons in heavy-element-containing molecules. This will provide crucial benchmarks for more approximate methods, as well as enable predictive simulation of chemistry and spectroscopy of heavy-element compounds, such as the proposed model compounds for separating mixtures of lanthanides and actinides that will be studied in this project.

To fulfill the promise of the novel simulation methods developed in this project, they will be implemented in publicly-available open-source software that can be deployed to modern high- performance computing platforms, such as the leadership-class computing facilities of the Department of Energy (DOE), including the exascale supercomputers soon to be deployed by the DOE's Exascale Computing Initiative . The ability to execute on modern computing platforms will be essential to accommodate the extra cost of the relativistic description of the quantum electrons in molecules and to allow simulation of realistic (i.e., sufficiently large) molecular models of experimental separations platforms.

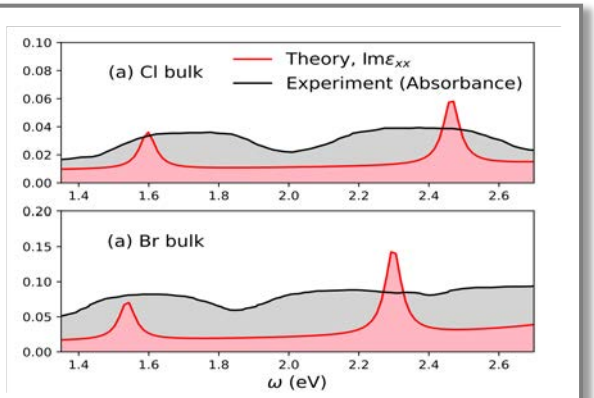
High-fidelity Green's functions in correlated materials

Lead Investigator: Mark van Schilfgaarde, mark.vanschilfgaarde@nrel.gov.
Participating Institutions: Mark van Schilfgaarde
Website: https://www.questaal.org/
Keywords: Electronic Structure, Green's functions, Strong Correlations, Unconventional superconductivity.

Green's function methods lie between quantum chemical methods (QC) favored by chemists and density functional (DFT) methods used ubiquitously. They are less efficient than the latter but more efficient than the former; further they have access to spectroscopic properties much more naturally than QC or DFT methods. GF methods have traditionally divided into two tracks: low-order many-body perturbation theory (MBPT), applicable to systems with weak or moderate correlations, and nonperturbative Dynamical Mean Field Theory (DMFT) when the independent particle picture is no longer adequate. Traditionally in *ab initio* formulations, both are added as a correction to density functional theory, taken as a standard reference one-body hamiltonian H^0. However, uncontrolled approximations in H^0 propagate to the higher level theory, which obscures and also limits the range of validity, in either track. By using MBPT itself to construct an *optimal* H^0, we show that the approximations become much better controlled. The fidelity of the theory significantly improves, and discrepancies with experiments become more systematic, making it possible to both clarify the limits of either path.

The optimization scheme, called Quasiparticle Self-Consistent *GW* (QS*GW*) approximation, can be systematically improved by adding higher order diagrams (QSGW^{++}).

These techniques are implemented in Questaal. A primary thrust area of this development is to adapt Questaal to the chemistry environment. Our aim is to make it possible to study low-dimensional systems with high fidelity (one- and two-particle spectral functions). We believe that with the addition of new magnetic diagrams we should be able to approach quantum chemical accuracy much more efficiently than traditional quantum chemical methods do. New algorithms are planned should enable efficient scaling with system size, adaptable to large HPC machines. We have already adapted existing algorithms for a highly parallel, GPU-enabled environment, which makes it possible to study 2D materials and interfaces.



Benchmarking against experimental optical absorption in bulk $CrCl_3$, $CrBr_3$. Comparison between the adopted optical absorbance data from the recent experimental work on $CrCl_{3-x}Br_x$[4] against our theoretical results. The excitonic peaks are remarkably deep --- several eV. The theoretical spectra for Imε_{xx} is rigidly blue-shifted by ~0.5 eV compared to experiments. The 2D case looks similar but the peak intensities are smaller, so the 3D case is shown.

Project Scope

The prime thrust area of this development is to adapt Questaal to the chemistry environment. Our aim is to make it possible to study low-dimensional systems with high fidelity. New algorithms are planned that should enable efficient scaling with system size, and adaptable to large HPC machines. Our hypothesis is that this theory can well characterize strongly correlated materials, esspecially spectroscopic properties in systems too complex to be accessible to of methods available today. Another focus area is to provide theoretical support for two key NREL programs: the solar photochemistry program and the CHOISE EFRC.

Relevance to MGI

The main focus of this project is theoretical/computational. Understanding the character complex and/or strongly correlated system a kind of holy grail in chemistry and physics. The current frontier of research in this area is to find approximations that are efficient yet retain sufficiently high fidelity. Another key focus area of our project is the ability to compute both 1-body and 2-body spectroscopic information. QC and DFT methods mostly focus on one-body properties. These are important as the total energy is a one-body property (as are derived quantities). But

two-body properties determine how a system responds to perturbations, which is at least as important. By exploiting two key advantages inherent in Green's function methods (better scaling with system size, and two-particle response functions) this project fills a need that is much less studied by traditional chemists.

Technical Progress

We have a large challenge to redesign Questaal for a chemistry environment, with new approaches to computing the response functions and self-energy. As this will take some time to get even initial results, we are pursuing a multi-track process. First there is a short-term track to redesign existing algorithms to make them scale efficiently on large GPU based machines. This was successfully adapted the 3D periodic code to such machines to make it possible to carry out QSGW^{++} calculations on much larger systems, including 2D materials and interfaces. We have used it to investigate the deep-lying excitons in CrX_3, a rare 2D ferromagnetic insulator. In another track, we have made our first attempt at embedding, showing that we can describe with high fidelity a point defect in a semiconductor host without adjustable parameters and have published an initial paper on it. In another track we are pursuing two materials topics of importance both to NREL and MGI (see Future Plans).

Future Plans

The prime focus is on method developments. The key technical advances we plan or are currently underway are : (1) to formulate a new basis of Jigsaw Puzzle Orbitals; (2) implement the theory in real space and use methods to achieve efficient scaling with system size; (3) Add diagrams not considered in an ab initio context to surmount key limitations of the existing theory; (4) augment the core diagrammatic method with dynamical mean field theory; (5) construct efficient computation of RPA total energy; (6) design specifically for next-generation machines.

We plan to apply the enhanced theory to low-dimensional systems: (1) chiral organic molecules on halide perovskites semiconductors, to support NREL's CHOISE EFRC. It has been found experimentally that this provides a path to yield spin-polarized light without application of a magnetic field. (2) photoexcitations in nanoscale particles used in catalysis, e.g. organic molecules on noble metal surfaces As a first step we will generate a spatially resolved dielectric function to determine where the light gets absorbed. Our ultimate aim is to determine how light enhances catalysis.

Broader Impacts and Workforce Development

Questaal is an open-access code. We held a workshop in May with 20 in person registrants and 60 virtual registrants. Several people a week request access to the code, from all over the world.

Data Management and Open Access

For data sharing we are in the process of setting up an interface to Aiida centered at EPFL.

Advancing Along the Materials Development Continuum and Partnerships to Translation

Having tools that reliably predict materials properties is the prime focus of MGI. Questaal has the potential to provide next-generation tools for this important effort.

Publications and References

1. S. Acharya, D. Pashov, A. Rudenko, M. Rosner, M van Schilfgaarde, M Katsnelson, “Real- and Momentum-Space Description of the Excitons in Bulk and Monolayer Chromium Tri-Halides,” npj 2D Mater App **6**, 33 (2022).
2. O. Dernek, D. Skachkov, W. R. L. Lambrecht, M van Schilfgaarde, “Real space representation of the quasiparticle self-consistent GW self-energy and its application to defect calculations,” Phys. Rev. B 105, 205136 (2022).
3. A. N. Rudenko, S. Acharya, F. Tasnádi, d. Pashov, A. V. Ponomareva, M van Schilfgaarde, KI A. Abrikosov, M I. Katsnelson, “Electronic and optical properties of crystalline nitrogen versus black phosphorus: A comparative first-principles study,” Phys. Rev. B 105, 205135 (2022)

DMREF: Collaborative Research: Developing Damage Resistant Materials for Hydrogen Storage and Large-scale Transport

Lead Investigator: T. A. Venkatesh, t.venkatesh@stonybrook.edu
Participating Institutions: Stony Brook University, Stanford University, Massachusetts Institute of Technology
Website: none
Keywords: Hydrogen embrittlement, fatigue, dislocation dynamics, atomistic simulation, machine learning.

Project Scope

This project aims to: (1) develop a physics-based modeling platform based on hydrogen interaction with crystalline defects at both atomistic and microstructural scales, informed by and integrated with experiments, to accelerate the pace of discovering controlling mechanisms of hydrogen embrittlement under fatigue; (2) enable faster development of hydrogen resistant materials in the energy transportation sector as it transitions from the transport of fossil fuels such as natural gas to hydrogen based renewables.

Relevance to MGI

This project is designed to advance our knowledge of crack tip processes that control damage accumulation and propagation under fatigue loading, and the role of hydrogen in making the material more brittle (Figure 1). We hypothesize that the controlling mechanisms occur in the plastic zone around the crack tip, over a length scale of about 1 to 10 microns, which is too small for continuum theory to be predictive and too large for atomistic simulations to handle by brute-force. We close this knowledge gap at the mesoscale, through a tightly coupled experimental-computational program. Our computational efforts build upon the recent advances made in atomistic simulations (at Sandia National Laboratory/Stanford), dislocation dynamics (DD) simulations (at Stanford), with insights on crystal plasticity (from MIT) and continuum level modeling (from Stony Brook). Our experimental efforts leverage improved and unique capabilities, such as Nanoindentation (at MIT/Stony Brook), X-Ray Tomography (at Brookhaven National Laboratory), and *in situ* testing in hydrogen environment (at Sandia National Laboratory).

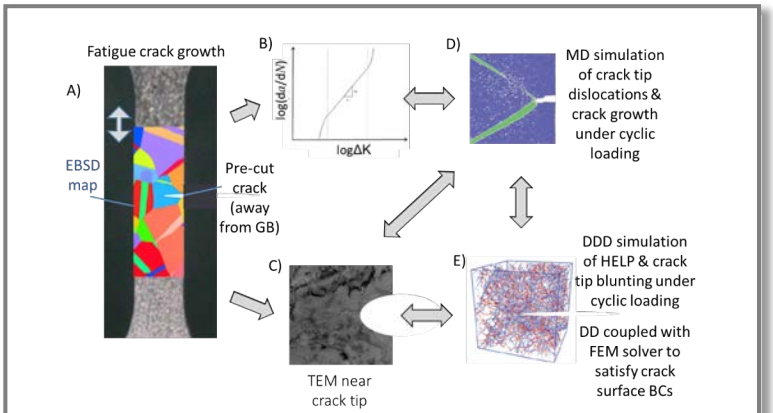


Figure 1. Schematic illustrating: A) An oligo-crystalline pipeline steel sample with a pre-cut crack being tested under cyclic fatigue load for crack growth; the overlapped EBSD map showing orientations of individual grains. B) The crack growth rate (d*a*/d*N*) observed in Stages I, II and III. C) TEM examination of the dislocations near the crack tip. D) MD simulation to study crack tip plasticity and crack growth under cyclic loading with influence of hydrogen. E) DD simulation of hydrogen induced shear localization and crack tip deformation.

Technical Progress

The successful realization of the hydrogen economy will require a coordinated effort from governments, research institutions, industries, policy makers, investors, and community leaders across the world. With a view towards enabling such a wide set of stakeholders to make informed decisions about the roll-out of the hydrogen economy, a study was conducted to review the current status and implementation of hydrogen blending in natural gas pipeline networks [1]. The primary emphasis of this review was to highlight top-level state-of-knowledge across

various stages in the lifecycle of a hydrogen economy—from methods for hydrogen production, to transportation of hydrogen blends and end-use applications. A computational fluid dynamic model has been developed to quantify frictional losses and energy efficiency of transport of methane-hydrogen blends across straight pipe sections [2]. It is observed that, in general, an increase in the energy costs is expected when hydrogen, with its lower density, is transported along with methane (which has higher density) in various blend ratios. Molecular dynamic (MD) simulations have been invoked to study hydrogen (H) adsorption as well as its effects on the strength of pearlitic steels [3]. In particular, the effects of H on the strength of the cementite (Fe_3C)/ferrite interphase and the ferrite/ferrite grain boundaries, have been investigated.

Future Plans

We will use high-throughput nano-indentation experiments under different hydrogen environments to calibrate our atomistic-DD model. We will use experiments and simulations of cyclic tensile tests to clarify the dislocation mechanisms in the limit of very low stress intensity factor amplitude (ΔK), where the hydrogen effects have not been conclusively determined or understood. We will examine the fatigue crack growth rate (da/dN), with the emphasis on the question of why hydrogen effects require a critical value of ΔK to be fully developed. We will develop physics-based continuum level models for hydrogen effects on fatigue crack growth to create reliable engineering roadmaps for life prediction and risk assessment for hydrogen storage and transport structures.

Broader Impacts and Workforce Development

This project will lead to broad impact at several levels: (i) By advancing our current understanding of the influence of hydrogen interactions with dislocations on the mechanisms associated with fatigue crack growth in pipeline steels, the proposed research activity will provide scientific and technological impact in the hydrogen storage and transport industry; (ii) This project will educate, train and mentor graduate and undergraduate students and postdoctoral associate and will provide a platform for promoting science to a broader audience, including under-represented minorities and high school students; (iii) The simulation tools, database and interactive visualization tools will be made widely available through online portals to facilitate dissemination and education; (iv) A major 6-day symposium with over 150 presentations was co-organized with hydrogen program leaders from NREL and Sandia National Laboratory to facilitate the dissemination of knowledge in the areas of advanced materials for hydrogen and fuel cell technologies to a world-wide community.

Data Management and Open Access

The source code of the atomistic and dislocation dynamics simulation programs developed in the project will be released to the public: http://micro.stanford.edu/wiki/MD++_Manuals. We will also develop wiki pages that fully describe the steps needed to reproduce the computation results in our publications. The wiki pages include all the input files necessary for reproducing all the numerical results/data. We will deposit experimental data and Abaqus input (.inp) and results (.fil) files in ascii format in DSpace@MIT and Stony Brook's Academic Commons, which will be accessible to the general public.

Advancing Along the Materials Development Continuum and Partnerships to Translation

We are working closely with our industry collaborators such as National Grid in this project who will help identify the commercialization potential for some of the fundamental insights obtained in this project.

Publications and References

1. D. Mahajan, K. Tan, T. A. Venkatesh, P. Kileti, and C. Clayton, *Hydrogen Blending in Gas Pipeline Networks – A Review*, ***Energies***, (2022).
2. K. Tan, D. Mahajan and T. A. Venkatesh, *Computational fluid dynamic modeling of methane-hydrogen mixture transportation in pipelines: estimating energy costs*, ***MRS Advances***, (2022).
3. X. Wang, Y. Zhao, G. Cheng, Y Zhang, T. A. Venkatesh, *Hydrogen adsorption in phase and grain boundaries of pearlitic steels and its effects on tensile strength*, ***MRS Advances***, (2022).

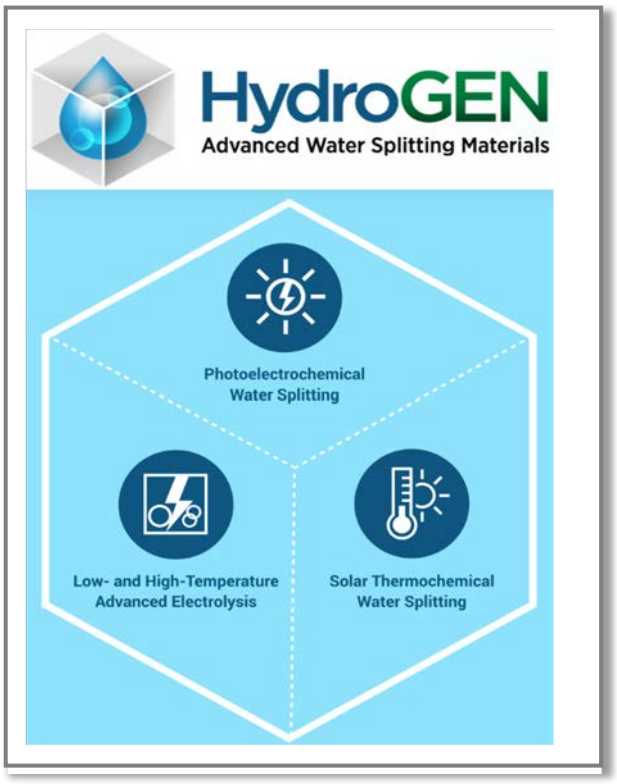
DOE HydroGEN EMN Consortium

Presenter: James Vickers, HFTO, james.vickers@ee.doe.gov
Lead Investigator: Huyen Dinh, (NREL) huyen.dinh@nrel.gov
Participating Institutions: NREL, INL, LBNL, SNL, LLNL
Website: https://h2awsm.org/
Keywords: Hydrogen, Solar Fuel, Electrolysis, Thermochemical, Solar Fuel, Photoelectrochemical

Project Scope

HydroGEN is a consortium of six U.S. Department of Energy (DOE) national laboratories focused on addressing advanced water splitting materials challenges by making unique, world-class national lab capabilities in photoelectrochemical (PEC), solar thermochemical (STCH), and low- and high-temperature electrolytic (LTE, HTE) water splitting more accessible to academia, industry, and other national labs. HydroGEN is part of the DOE Energy Materials Network (EMN) and is funded by DOE's Hydrogen and Fuel Cell Technologies Office.

HydroGEN aims to facilitate collaborations between federal laboratories, academia and industry. The consortium is guided by a steering committee with representatives from each member lab and DOE. The steering committee is available to clarify the capabilities offered by the consortium and to help interested users identify capabilities relevant to a given research project.



Relevance to MGI

Accelerating advanced materials development, from discovery through scale-up, has the potential to revolutionize whole industries and is critical for the United States to compete globally in manufacturing in the 21st century. However, today only a small fraction of materials innovations makes it to widespread adoption and commercialization. The goal of HydroGEN as part of the (EMN) is to dramatically decrease the time-to-market for advanced materials that are critical to manufacturing many clean energy technologies, enabling manufacturers of all sizes to develop and deliver innovative, made-in-America products to the world market.

Technical Progress

Given the breadth of HydroGEN, many accomplishments have been made with a few below. Theory-based computational tools developed by seedling projects are rapidly identifying new STCH materials. Projects have provided mechanistic understanding of PEC device degradation guided by theory and in operando characterization. Developed models to explain and predict Ni migration degradation in solid oxide electrolysis Ni-YSZ fuel electrodes. Significantly improved alkaline exchange membrane electrode assembly kinetics with water-only feeds.

Future Plans

HydroGEN 2.0 has specific goals set for each of the advanced water splitting (AWS) pathways:

- LTE: improve alkaline exchange membrane electrolysis performance and durability by determining the role of supporting electrolyte and the limiting factors behind DI water operation
- HTE:

    - Metal-Supported Solid Oxide Electrolysis: improve performance and durability with a scale-up cell
    - Proton Conducting Solid Oxide Electrolysis: understand the proton conduction and electronic leakage mechanisms of electrolyte materials in proton-conducting SOEC
- PEC: materials stability and device durability
- STCH: identify and understand how structural features, composition, and defect dynamics engender high capacity–high yield behavior in materials

Broader Impacts and Workforce Development

HydroGEN is vastly collaborative and has produced many high value products that it disseminates to the R&D community. The consortia facilitates strong community engagement and participation, nationally and internationally. They have hosted 4 Annual AWS community-wide benchmarking workshops and promote STEM work force development.

Data Management and Open Access

A goal of HydroGEN is to capture data, tools, and expertise developed at each resource node so that they can be shared and leveraged throughout the EMN and in future programs. This includes establishing data repositories and, where appropriate, distributing data to the scientific community and public.

HydroGEN's data team, led by the National Renewable Energy Laboratory (NREL), has developed a Data Hub that combines experimental and computational data into a searchable materials data infrastructure and encourages researchers to make their data available to others. HydroGEN partners are expected to support the Data Hub with any non-proprietary data generated. The Data Hub has over 300 registered users and over 400 datasets to date. Visit the HydroGEN Data Hub.

Advancing Along the Materials Development Continuum and Partnerships to Translation

Materials innovations in clean energy technologies are key to enhancing the performance and durability of hydrogen production technologies and reducing the cost of renewable hydrogen generation. Consistent with the MGI and EMN vision, moving transformational energy technologies forward requires an integrated approach spanning materials design and synthesis through process scale-up. Through HydroGEN, experts at the national laboratories support projects which have included 39 universities, 11 labs and 10 companies resulting in numerous publications and several patents. Funding form the DOE has totaled $43.8M since the consortia was kicked-off in 2016. Current consortium efforts seek to discover and design new materials, increase the efficiency and durability of critical components and systems, and advance manufacturing and scale-up efforts.

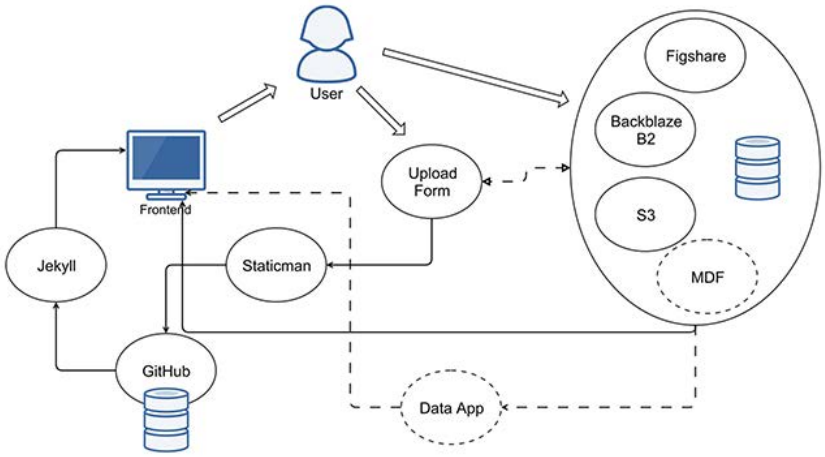
Publications and References

Hydrogen has had 118 Publications, with 2,783 citations, including 436 authors. For more information see website.

PFHub: The Phase-Field Community Hub

Lead Investigator: Daniel Wheeler, daniel.wheeler@nist.gov.
Participating Institutions: NIST, CHiMaD
Website: https://pages.nist.gov/pfhub
Keywords: phase-field, materials-science, jekyll-website, reproducible-science, scientific-portal

Scientific communities struggle with the challenge of effectively and efficiently sharing content and data. An online portal provides a valuable space for scientific communities to discuss challenges and collate scientific results. Examples of such portals include the Micromagnetic Modeling Group (μMAG), the Interatomic Potentials Repository (IPR) and on a larger scale the NIH Genetic Sequence Database (GenBank). In this work, we present a description of a generic web portal that leverages existing online services to provide a framework that may be adopted by other small scientific communities. The first deployment of the PFHub framework supports phase-field practitioners and code developers participating in an effort to improve quality assurance for phase-field codes.



Schematic overview of the PFHub framework for building scientific research portals, simply.

Project Scope

The phase-field method (PFM) describes material interfaces at the mesoscopic scale between atomic scale models and macroscale models. The PFM is well established and there are an assortment of code frameworks (e.g., FiPy, MMSP, MOOSE, PRISMS-PF) available for solving the wide variety of phenomena associated with phase-field (e.g. dendritic growth, spinodal decomposition, grain growth). However, it is difficult for novice as well as seasoned phase-field practitioners to assess the capability of codes for different phenomena. PFHub aims to provide a low barrier for comparing code output data using a standard set of metrics.

Relevance to MGI

PFHub is a community effort spearheaded by the Center for Hierarchical Materials Design at Northwestern University and the National Institute of Standards and Technology in support of phase-field code development. The current PFHub deployment focuses on improving cross-collaboration between phase-field code developers and practitioners by providing a standardized set of benchmark problems and a workflow for uploading and comparing benchmark results from different phase-field codes.

Community based scientific efforts often require web services to share and display data in unique ways between groups and institutions. These services are difficult to implement due to the groundwork required to investigate and prototype the many data-sharing and CMS tools available. The PFHub framework provides a template for other scientific projects beyond the phase-field community.

Technical Progress

The CHiMaD phase field working group has released two new benchmarks to validate nucleation phenomena in phase field codes. The first benchmark focuses on homogeneous nucleation for a single seed and multiple seeds. The second benchmark focuses on athermal heterogeneous nucleation. The first benchmark specification is now on the PFHub website and the second benchmark specification is in progress.

A new benchmark result upload mechanism is now available on the PFHub website. The mechanism allows a submitter to simply fill out a GitHub issue to submit results. GitHub actions are then used to check, validate and display the data for approval by the submitter via a GitHub pull-request.

A Python module has been released to gather and display PFHub data. This allows interested parties to use PFHub data independently for the website. Furthermore, the Python module removes the requirement for a JavaScript frontend as the gathering, refining and displaying of data is automated via GitHub actions and the resulting displays stored stacially on the website. Using Python in place of JavaScript makes the infrastructure more accessible and adaptable.

The project has received uploads from 4 new codes, which include AMPE (adaptive mesh phase field evolution, LLNL), MEMPHIS (mesoscale multiphysics phase field simulator, Sandia), MEUMAPPS (microstructure evolution using massively parallel phase field simulations, ORNL), SymPhas (general Purpose Software for phase field, phase field crystal, and reaction-diffusion simulations, Univ. of Western Ontario). The developers from these codes have all become involved with the CHiMaD phase field working group and have attended workshops and have stated that the benchmarks are important in the development of each code.

Future Plans

The current work underway involves a complete FAIR re-implementation of the upload process. The new process will require the submitter to first archive benchmark result data to a research repository such as Figshare, Zenodo or the Materials Data Facility. The submitter will then share the resulting DOI with PFHub via an automated mechanism such as a GitHub issue. The PFHub CI system can then extract relevant metadata and data directly from the research repository and compile the data into comparison views that include all the PFHub data. The submission is then checked via an automated pull-request and a temporary data view is supplied for human inspection before final submission.

Broader Impacts and Workforce Development

The phase field benchmarks are used for instructing researchers at materials summer schools and are included in the computational materials curriculum at a number of institutions. The benchmarks are in fact designed with the goal of being a pedagogical resource for teaching phase field methods. In addition, the PFHub website archives many examples of solving the phase field benchmarks providing a valuable resource for learning the practical implementation of phase field in common code bases. This exposes researchers to a variety of different phase field code choices and prevents siloed research and development.

Publications and References

1. A.M. Jokisaari, P.W. Voorhees, J.E. Guyer, J. Warren, O.G. Heinonen, *Benchmark problems for numerical implementations of phase field models*, Computational Materials Science **126**, 139-151 (2017)
2. A.M. Jokisaari, P.W. Voorhees, J.E. Guyer, J.A. Warren, O.G. Heinonen, *Phase field benchmark problems for dendritic growth and linear elasticity,* Computational Materials Science **149**, 336-347 (2018)
3. A.M. Jokisaari, W. Wu, P.W. Voorhees, J.E. Guyer, J.A. Warren, O.G. Heinonen, *Phase field benchmark problems targeting fluid flow and electrochemistry,* Computational Materials Science, **176**, 109548 (2020)
4. W. Wu, D. Montiel, J.E. Guyer, P.W. Voorhees, J.A. Warren, D. Wheeler, L. Gránásy, T. Pusztai, O.G. Heinonen, *Phase field benchmark problems for nucleation*, Computational Materials Science, **193**, 110371 (2021)
5. D. Wheeler, T. Keller, S.J. DeWitt, A.M. Jokisaari, D. Schwen, J.E. Guyer, L.K. Aagesen, O.G. Heinonen, M.R. Tonks, P.W. Voorhees, J.A. Warren, *PFHub: The Phase-Field Community Hub*, Journal of Open Research Software **7(1)**, 29 (2019)

Collaborative Research: DMREF: Accelerated Discovery of Artificial Multiferroics with Enhanced Magnetoelectric Coupling

Lead Investigator: Li Yang, lyang@physics.wustl.edu
Participating Institutions: Washington University in St Louis, University of Texas, Austin, University of Nebraska-Lincoln
Website: none
Keywords: multiferroic, magnetoelectric coupling, *ab initio* simulation, van der Waals material, moiré superlattices

Project Scope

This project is to combine computational search and experimental fabrications and characterizations to search for and realize artificial multiferroics exhibiting enhanced magnetoelectric (ME) coupling by integrating two-dimensional (2D) van der Waals (vdW) magnets with ferroelectrics in two types of heterostructures: vdW/vdW and vdW/oxide heterostructures. Taking advantage of the nonvolatile polarization and piezoelectric effect in ferroelectrics, tunable magnetism in layered vdW materials, and substantial interface-to-volume ratio, we propose to achieve electrical control of the magnetic order, magnetic anisotropic energy (MAE), magnetic Curie or Néel temperature (T_C or T_N), and magnetic moiré pattern through interfacial coupling mechanisms.

Relevance to MGI

Artificial multiferroics offer a promising solution to achieve enhanced coupling between the ferroelectric (FE) and magnetic orders that are not available in single phase materials. Given that a large number of 2D magnetic and FE materials have already been predicted and grown, the possible combinations of these building blocks can reach the order of 10^4-10^5. It is thus imperative to launch a computation-led program to search for optimal artificial multiferroics, as shown in the right figure. A heterostructure needs to be computed as a whole system before re-evaluating the ferroic orders and their coupling. Experimental studies of prototypical heterostructures will be used to validate computational methods and guide improved designs. We will develop new high-throughput computational methods to identify about a few thousand suitable multiferroics. The interpretation of unexpected experimental observations will be elucidated by computations and complementary measurements. Exploratory materials synthesis and characterizations will reduce the candidate multiferroics. Ultimately, we aim to identify and realize ~10 new artificial multiferroics with figures of merit better than currently available ones in the following aspects: high ferromagnetic T_C or antiferromagnetic T_N (~100 K or higher), soft coercive field (100 Oe or lower), tunable MAE, and modulation of magnetism at low switching voltage (~1 V or lower).

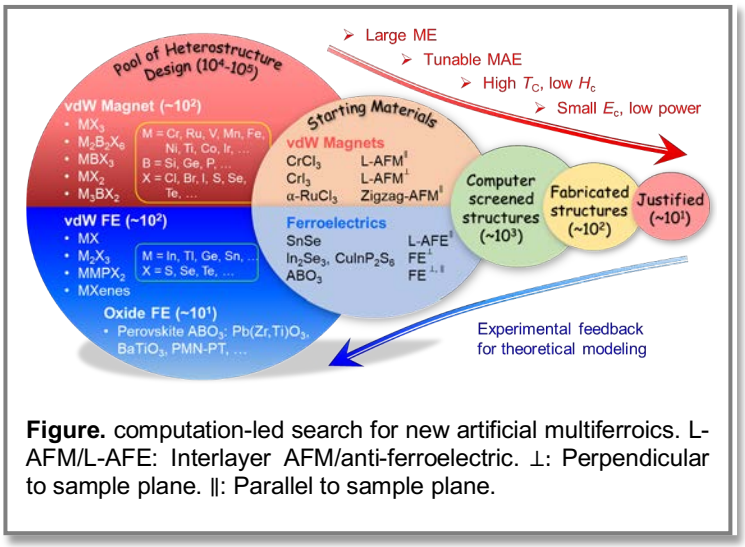


Figure. computation-led search for new artificial multiferroics. L-AFM/L-AFE: Interlayer AFM/anti-ferroelectric. ⊥: Perpendicular to sample plane. ∥: Parallel to sample plane.

Technical Progress

Several vdW materials have been identified as promising building blocks of heterostructures for realizing the proposed magnetoelectric properties. For the magnetic component, using physical vapor transport (PVT), we have successfully fabricated high-quality 2D antiferromagnetic (AFM) $CrCl_3$ flakes down to monolayer thickness. The structure, stoichiometry, morphology, and room temperature stability are investigated. Tunneling magnetotransport studies of graphite/few-layer $CrCl_3$/graphite tunnel junctions reveal in-plane magnetic anisotropy and the Neel temperature T_N of 17 K. For the FE component, we have successfully grown a promising piezoelectric/FE vdW material, SnS, by chemical vapor deposition (CVD). In-plane polarization of each SnS layer and anti-ferroelectric ordering between adjacent layers are confirmed by piezo-force microscopy (PFM).

In addition to progress in materials growth, we have established an optical spectroscopy method to identify the Néel vector in 2D AFM $NiPS_3$. Obtaining information on AFM orders remains a difficult task as the vanishing net moments bring significant challenges in both optical and electrical readout methods. We show that polarization-resolved photoluminescence (PL) of sharp exciton resonances in $NiPS_3$ can reveal the direction of local spin chains. The PL signal is maximal along the direction perpendicular to the local spin chains. With the help of first-principles simulations, we suggest that the sharp exciton resonances may originate from states bound to S-vacancies. The unique correlation between anisotropic excitons and local spin-chain directions makes $NiPS_3$ a compelling AFM material for developing future spin-photon transduction devices.

Beyond vdW magnets, we have predicted novel magnetoelectric coupling in molecular magnets. Using first-principles calculations and Heisenberg Hamiltonian, we demonstrate this effect in a family of magnetic molecules, transition metallic Porphyrins (TM-Pcs). This spin-electric coupling may provide a new approach for designing and controlling molecular spintronics for realizing the smallest functional magnets.

Future Plans

With the progress of growth and characterizing vdW magnetic and ferroelectric layers, we will aim to realize new artificial multiferroics with enhanced ME coupling by exploiting their unique interfacial interactions. More heterostructures made by vdW/vdW and vdW/oxides will be proposed by simulations to search for magnetoelectric properties. Meanwhile, experimental groups will further improve the PVT and CVD growth techniques to fabricate more vdW magnets (such as CrX_3 and $MnBi_2Te_4$) and ferroelectrics/piezoelectrics (such as SnSe and SnTe), as well as exploiting mechanical exfoliation of bulk samples. We will further make heterostructures of these materials to explore the strain and charge transfer in the interface and the magnetoelectric properties. Developed optical tools, including linear and nonlinear optical responses, will be combined with magnetotransport studies and scanning probe microscopy techniques, including PFM and magnetic force microscopy, to detect the magnetic and polarization orders in these heterostructures.

Broader Impacts and Workforce Development

The project has provided educational and research opportunities for training the next-generation workforce in material science. Two graduate students and two postdoc have been supported by this program, including a female Ph.D. student, a Hispanic postdoc, and a female postdoc. A graduate-level course, “First-principles studies of quantum materials”, have been developed at Washington University in the spring semester of 2022. We plan to further develop the curriculum by including new topics such as data mining and machine learning. Experimentalists on the team will be invited to give guest lectures to provide a comprehensive perspective.

Data Management and Open Access

The simulation data are collected and stored at Texas Advanced Computing Center. The experimental data are stored in lab computers and backed up to an external hard drive and cloud-based servers such as Dropbox, One Drive, or Box monthly. Our institutes (Washinton University in St Louis, University of Nebraska-Lincoln and University of Texas-Austin) provide unlimited data storage space on one or more cloud-based servers. These servers are managed centrally on a long-term basis. The generated data will be published and/or deposited for public access.

Advancing Along the Materials Development Continuum and Partnerships to Translation

The project accelerates materials discovery and development through its combined theoretical and experimental approach. The computational search to identify candidate magnetic and ferroelectric materials and interfaces are continuously improved by feedback from material growth and characterizations. The developed optical characterization not only provides an efficient tool to characterize magnetic orders but also helps theoretical simulations to reveal the fundamental physics mechanism behind the magnetoelectric coupling.

Publications and References

1. Jia Wang, *et al.*, “Physical Vapor Transport Growth of Antiferromagnetic $CrCl_3$ Flakes Down to Monolayer Thickness”, submitted.
2. Dong Seob Kim, *et. al.*, “Strongly bound excitons reveal local spin chain directions in a van der Waals antiferromagnet” under review at Nature Nanotechnology
3. Yan Lu, *et al.*, "Electric Field Tuning of Magnetic States in Single Magnetic Molecules", submitted.
4. Wenqi Yang, *et al.*, "Point Defects in Two-Dimensional $RuCl_3$", submitted.

Inverse Design of Architected Materials with Prescribed Behaviors via Graph Based Networks and Additive Manufacturing

Lead Investigator: Xiaoyu (Rayne) Zheng
Participating Institutions: UCLA and Penn State University
Website: none
Keywords: Graph neural network, architected metamaterials, dynamic and structural behaviors

Project Scope

This project will extend graph-based generative machine learning modeling techniques to identify the underlying motifs within architected materials to understand their dynamic behaviors as well as provide an inverse design framework for optimized functional responses. This extensive and experimentally validated framework will be used to discover fundamental knowledge pertaining to structural and dynamic characteristics, which will then be leveraged to inversely design materials with prescribed dynamic fingerprint.

Relevance to MGI

Our project entails a strategic integration and closed-loop combination of simulation, experimental, manufacturing, and machine learning activities that mutually inform and reinforce each other, and therefore is highly relevant to the MGI strategic plan.

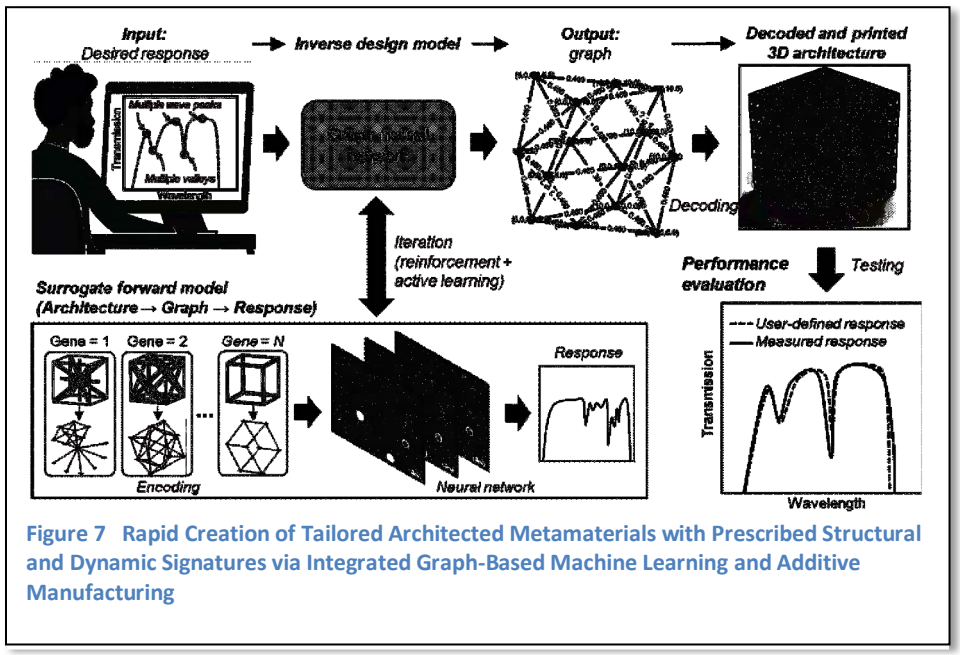


Figure 7 Rapid Creation of Tailored Architected Metamaterials with Prescribed Structural and Dynamic Signatures via Integrated Graph-Based Machine Learning and Additive Manufacturing

By seamlessly integrating new advances in graph network theory, artificial intelligence, numerical simulations, and high-speed additive manufacturing approaches, this project will accelerate the understanding, inverse design, and synthesis of architected materials with tailorable dynamic fingerprints, where materials with inversely designed three-dimensional micro-architectures will respond with prescribed impact shielding and wave transmission behaviors and replicated via desktop additive manufacturing. This project also bridges the critical gap between numerical and experimental training data, the latter of which has been largely neglected in previous machine learning models, but are critical in compensating complex manufacturing errors.

Technical Progress

We are developing a graphic neural network model using a graph representation as input to represent topology genes of metamaterials. Deep graph-based models (e.g., GNN) can directly capture the topology of a given architected metamaterial and can leverage the power of deep networks for automated feature discovery. Unlike other deep learning techniques such as convolutional neural networks (CNNs), GNNs can handle graphs with arbitrary topology and extend the deep learning models from regular structures (e.g., lattice and sequence) to

irregular structures (e.g., graph).

Future Plans

To construct GNN gene pool for architected metamaterials. We will identify proper strategies for constructing representative training topologies and graphs, from which approximately 10,000 unit cell designs can be generated to study their dynamic and structural features by combining simulation and experimental data. We will start from stress-strain curve responses to establish the bounds and then move to dynamic responses where we will focus on elastic wave transmission curve and sound absorption curve, though other dynamic response curves (e.g., acoustic wave transmission curve and elastic wave absorption curve) can be readily considered using the methodologies developed here. For elastic waves, band gaps (which are manifested as valleys in the transmission spectrum) are of particular interest.

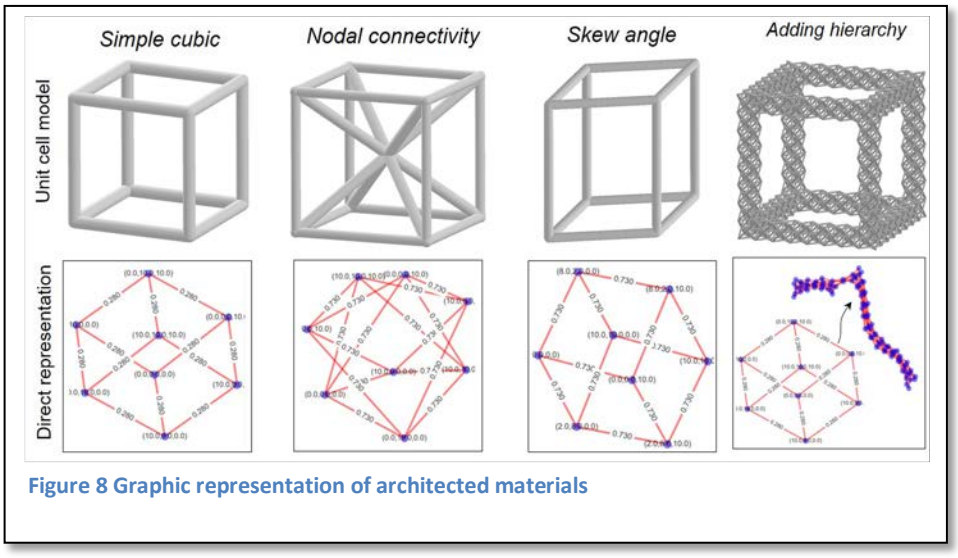


Figure 8 Graphic representation of architected materials

To develop a multi-fidelity model for training data We will establish a novel multi-fidelity graph-based model that can seamlessly learn from experimental and simulation data within the same learning pipeline. The model will seamlessly fuse and combine high-fidelity experimental data and low-fidelity synthetic data within the same training pipeline so that low- and high-fidelity data can mutually augment, complement, and inform each other. Additionally, we plan to develop a molecular dynamic representation of lattice structures to accelerate the training data.

To inverse design, measure, and demonstrate novel metamaterials possessing extraordinary dynamic and/or structural properties using proposed metamaterial design methodology. It will serve as a final validation of the graph- and RL-based inverse design scheme. Another important goal of this task is to expand graph explanation methods to study sub-graph features, in order to identify influential structural features that contribute to the coupling between different material properties. This investigation could provide crucial insight towards the creation of metamaterials with prescribed structural and dynamic behaviors that have been long considered mutually exclusive.

Broader Impacts and Workforce Development

The project will train graduate and undergraduate students in the new paradigm of autonomous inverse design and additive manufacturing based on desired behaviors. Moreover, demonstration modules, design games, and additive printing activities will be used for outreach to K-12 students.

Data Management and Open Access

All the machine learning codes will be implemented in Python and will rely on open-source packages and libraries, including Scikit-learn, NumPy, and SciPy. All publications, conference proceedings, and technical reports will be hosted on eScholarship—UCLA's open-access institutional repository. Publications relevant to the field of condensed matter will also be deposited on arXiv. Aspects relevant to educational advancement, e.g., webinar contents, presentations, Ph.D. theses, etc., will also be made available via electronic/web-based media.

Advancing Along the Materials Development Continuum and Partnerships to Translation

Current paradigms in engineering material design and manufacturing rely on searching database or handbooks that record the measured mechanical properties of existing materials. However, few approaches, if any, are capable of inversely creating materials with prescribed behaviors due to challenges in constraints of natural materials, designing nonlinear behaviors, and accounting for process-dependent manufacturing errors. Leveraging recent foundational advances in machine learning and additive manufacturing, this project will extend graph-based generative machine learning models to inversely synthesize 3D architected materials with prescribed behaviors.

From Wavefunctions to Exchange Correlation for Large-scale Electronic Structure Calculations

Lead Investigator: Paul Zimmerman, paulzim@umich.edu.
Participating Institutions: University of Michigan, Ann Arbor
Website: None.
Keywords: Configuration Interaction, Inverse DFT, Data-driven exchange correlation approximation, Wavefunction theory, Density functional theory

Project Scope

This project seeks to increase the accuracy and application scope of wave function theory (WFT) and density function theory (DFT) electronic methods through new scalable, highly accurate computational algorithms, and data-driven approaches. To provide a step function increase in DFT accuracy, density inversion will map the full configuration interaction (FCI) electron density to the exchange-correlation potential, which is the quantum mechanical component of DFT. Alongside the use of WFT to create precise electron densities, this development will allow training of new density functionals to reproduce the exact 3D exchange-correlation potential. The new functionals—constructed via system identification and machine learning through deep neural networks—will push DFT towards the ability to capture strong correlation, with the goal of achieving WFT-like accuracy.

Relevance to MGI

The proposed work uses the integration of recent advances in the algorithms, scalable implementations on HPC architectures and data-driven approaches to improve the exchange-correlation description in density functional theory. In particular, we use incremental full configuration interaction (iFCI) calculation to compute accurate ground-state densities and energies on polyatomic molecules, and subsequently compute the exchange correlation (XC) potential corresponding to these ab-initio correlated densities via recent advances in the solution of the inverse DFT problem. We use the data on the exact XC potentials to not only evaluate the existing XC functional models [1], but also use data-driven approaches to improve the XC functional description. We anticipate that these developments will enable accurate electronic structure calculations using density functional theory for materials systems with both weak and strong correlations, and thus accelerate our understanding of materials behavior and materials discovery.

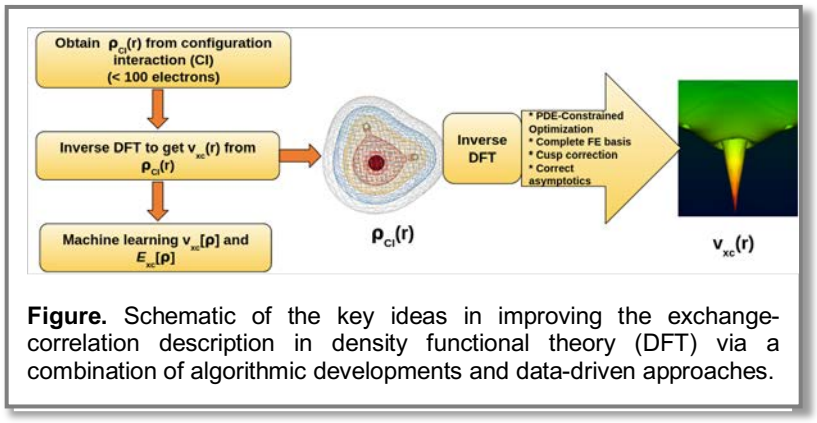


Figure. Schematic of the key ideas in improving the exchange-correlation description in density functional theory (DFT) via a combination of algorithmic developments and data-driven approaches.

Technical Progress

In the first year of this project, we have extended our inverse DFT formalism [2] to spin polarized (open shell) systems, and we are generating the exact XC potentials from ab-initio correlation electron densities for both closed shell and open shell systems obtained from accurate iFCI calculations. We presently have the data on a range of atomic and molecular systems and are continuing to generate more data. We have also made progress on improving the accuracy and scalability of configuration interaction calculations to generate the ground-state densities for larger molecular systems, thus enabling us to compute the exact XC potentials for larger systems. Further, we have developed the framework to use the data on the ground-state densities, energies, and the corresponding exact XC potentials to develop better XC models using machine-learning ideas. As a first step towards learning an XC functional, we have created a neural network (NN) based local spin density approximation (LSDA), utilizing the

exact XC potentials from our inverse DFT calculations. Our training samples include H_2, Li, LiH, N, O, and Ne. Our preliminary studies show an error of < 25 mHa in total energy per atom predictions on test systems, marking a significant improvement over the > 100 mHa errors in existing LSDA models (e.g., PW92). Similarly, we attain errors 15-25 mHa in the atomization energies on the test systems, which in turn are substantial improvements over existing LSDA models. We have also developed the general framework for NN based models for GGA and meta-GGA functionals, and are working towards developing NN models for these classes of functionals.

Future Plans

In the next years of this project, we will continue generating exact XC potentials for a wider range of molecules to augment the training data for our NN based XC models. Additionally, given the initial promise of the NN based LSDA models, we intend to develop models with increasing expressivity: (a) GGA, (b) meta-GGA, and (c) generalized derivative approximation (GDA), which will include both integral and fractional derivatives of the density. For each of the models, we will extensively benchmark their accuracy against thermochemical datasets. As the accuracy of these models will depend on the available training data, we will develop highly accurate electron densities using a parallel implementation of the heatbath configuration interaction (HBCI) method. Further, we also will port the inverse DFT code to GPUs to accelerate the inverse DFT calculations and the evaluation of the exact XC potentials.

Broader Impacts and Workforce Development

The junior researchers on this project are being trained by the PIs on a variety of aspects of scientific computation, ranging from theoretical and mathematical aspects of electronic structure to good programming practices and parallel computing. The project outcomes are being shared with the broader community via open-source codes and data-sharing, besides publications and presentations at meetings.

Data Management and Open Access

The open-source aspect is a big part of this project, where the developed codes, the data obtained on the XC potentials, and the developed XC models will all be made public over the course of the project. Some aspects of the codes are already open source: (a) SlaterGPU library (https://github.com/ZimmermanGroup/SlaterGPU); (b) DFT-FE codebase (https://github.com/dftfeDevelopers/dftfe) over which the public version of inverse DFT algorithm will be made available.

Advancing Along the Materials Development Continuum and Partnerships to Translation

We anticipate that the data on the exact XC potentials will substantially accelerate the development of new XC functionals with increasing accuracy for both weakly- and strongly-correlated systems. Thus, this effort will overall improve the accuracy of DFT calculations, which is the workhorse of electronic structure calculations in MGI. These developments coupled with the advances towards scalable large-scale DFT calculations on exascale computing architectures [3] has the significant potential of accelerating materials discovery.

Publications and References

1. Kanungo, B., Zimmerman, P., Gavini, V., A comparison of exact and model exchange-correlation potentials for molecules, *J. Phys. Chem. Lett.* 12, 12012-12019 (2021).
2. Kanungo, B., Zimmerman, P., Gavini, V., Exact exchange-correlation potentials from ground-state electron densities, *Nature Communications* 10, 4497 (2019).
3. Motamarri, P., Das, S., Rudraraju, S., Ghosh, K., Davydov, D., Gavini, V., DFT-FE: A massively parallel adaptive finite-element code for large-scale density functional theory calculations, *Comput. Phys. Commun.* 246, 106853 (2020).