# Bringing Quantum Mechanics to Coarse-Grained Soft Materials Modeling

Chun-I Wang and Nicholas E. Jackson*

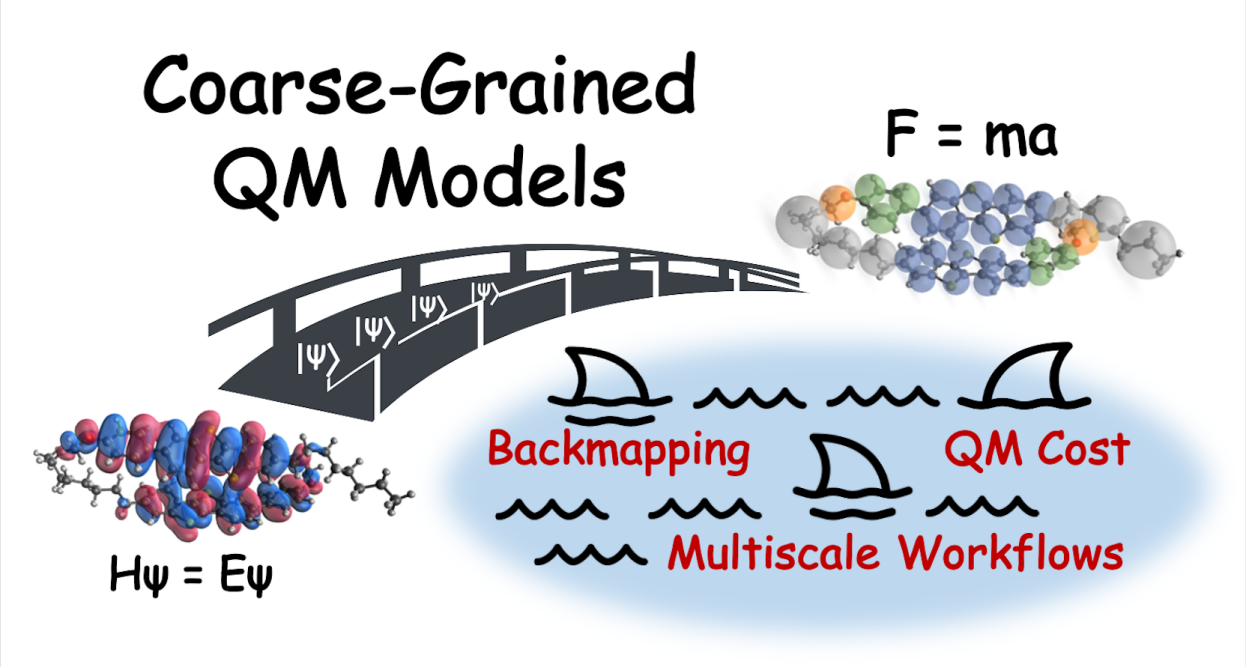
*Department of Chemistry, University of Illinois at Urbana-Champaign, 505 S Mathews Avenue, Urbana, Illinois, 61801, USA*

E-mail: jacksonn@illinois.edu

**Abstract**

Fundamental knowledge gaps are endemic in our understanding of how emergent soft materials properties are linked to the quantum mechanical (QM) world. The limitations of current QM modeling paradigms inhibit the understanding and design of soft materials classes for which QM phenomenology is critical. At its root, these limitations derive from the seemingly innocuous premise of requiring all atomic positions to solve the molecular Schrodinger equation, which necessitates supercomputing resources to incorporate even simple QM phenomenology into small (∼nm) soft materials systems. Here, we review emerging efforts to overcome these challenges through the development of electronic prediction models that operate at the coarse-grained resolution. We motivate the origins of this new computational paradigm, denoted Electronic Coarse-Graining (ECG), discuss its relationship to existing molecular modeling frameworks, and describe recent successes of ECG and related models for soft materials. Importantly, we highlight the classes of soft materials where ECG models can be potentially transformative.

## For Table of Contents use only




## Biography Sketches

Prof. Nick Jackson obtained his B.A. in Physics from Wesleyan University and his Ph.D. in Chemistry from Northwestern University. He then performed postdoctoral work jointly between Argonne National Laboratory and the University of Chicago as a Maria Goeppert Mayer Fellow. In January 2021 he started his independent laboratory as an Assistant Professor at UIUC. His lab develops and applies computational methods that enable the design of electronic functionality in soft materials.

Dr. Chun-I Wang received his PhD degree of Chemical Engineering at National Chung Cheng University under the supervision of Prof. Chi-Chung Hua in 2017. After working as a postdoctoral researcher with Prof. Chao-Ping (Cherri) Hsu at Academica Sinca from 2017 to 2021, he joined Prof. Nicholas Jackson's group at University of Illinois Urbana-Champaign. His research interest is investigating structure-property relationships in soft materials by integrating quantum mechainics, classical multiscale simulation, and machine learning.

# Introduction

To date, the dramatic progress of soft materials has been driven by the efficacy of models utilizing classical physics. From the first polymer model of Flory, to the most systematic coarse-grained (CG) methodologies of the present, many of the physical properties of soft materials can be understood without in-depth knowledge of the quantum mechanical (QM) phenomena operating at the molecular scale. In many such models, the relevant classical physics can be imbued with chemical specificity by fitting smalls sets of chemically dependent parameters. In more recent efforts to design structure and thermodynamics in soft materials from the "bottom-up", molecular QM calculations can be employed to parameterize atomic and CG potential energy surfaces, but the partition functions sampled by these force-fields are still entirely classical. The success of classical modeling for soft materials design is not surprising, as the structural, thermodynamic and mechanical properties of soft materials relevant to traditional applications (e.g. adhesives, containers, thermosets, mechanical composites) are often dictated by length scales far exceeding the Angstrom scale. In these scenarios, CG models that renormalize the intrinsically complex QM physics into effective tuning parameters are a logical and preferred strategy to fully QM simulations. These facts have led to a broad array of theoretical and computational models based in classical physics that are the workhorses of the soft materials communities.

While the successes of classical phenomenology in soft materials are beyond reproach, we here argue that some of the most pressing societal challenges necessitate the incorporation of non-trivial QM phenomena into soft materials design. When considering the unknowns at the frontier of soft materials, the following questions commonly emerge: How can we recycle polymers and/or design their degradability? How can we architect flexible electronics for transistors, solar cells, bioelectronics, or neuromorphic computing? How do we model reactivity to control functionality in emerging soft materials of relevance to energy generation and storage? What are the molecular mechanism(s) of contact electrification in polymers? Can we predict the fracture behavior and mechanochemistry of polymers with chemical speci-

ficity? How does biology efficiently move QM photoexcitations over mesoscopic distances, or transduce adenosine triphosphate (ATP) to mechanical motion? In this small sampling of soft materials applications, one arrives at the need to understand QM properties over heterogeneous morphologies extending well-beyond the Angstrom scale, and often to recouple these changes in electronic structure back into the structural and thermodynamic prediction. This nontrivial task requires that classical physics must be supplemented with more nuanced information about molecular electronic structure and its coupling to morphology at multiple length scales.

The modeling of QM properties in soft materials rests at the interface of two mature modeling paradigms between which little communication traditionally occurs: CG modeling and quantum chemistry (QC). At face value, each field is antithetical to the other: CG modeling treats length scales beyond the atomic scale by averaging multiple specific atoms into effective pseudo particles — this results in dramatically more efficient CG structural prediction models, but also the inability to perform chemically specific QC predictions. On the other hand, while QC can approach an exact treatment of the full electronic structure problem, the necessity to track all (or most) electronic degrees of freedom severely limits its scalability. To this end, accounting for all atomic and electronic degrees of freedom for even a 10 nm cube of polymer ($\sim 10^5$ atoms/$\sim 10^6$electrons) surpasses the capability of most modern computational resources. These considerations necessitate financially and environmentally taxing supercomputing calculations to incorporate even simple QM phenomenology into soft materials using traditional molecular modeling paradigms. It is our assertion that existing molecular modeling tools for soft materials are insufficient, and significant innovation is required to make QC practical for the spatiotemporal scales of soft materials modeling.

Here, we argue that to efficiently incorporate QM phenomenology in soft materials design, a new molecular modeling paradigm must be introduced that bridges the disparate fields of CG modeling and QC chemistry in a systematic, "bottom-up" fashion. Specifically, this article argues three points. First, that pressing societal challenges with soft materials

solutions cannot be addressed by purely classical models and require the incorporation of QM phenomenology. Second, that the traditional molecular modeling paradigms of CG modeling and QC, taken separately, are ill-equipped to tackle the aforementioned challenge of integrating QM phenomenology into soft materials design. Third, that there is a significant prospect for developing systematic CG models of electronic structure that renormalize QC-accurate electronic predictions over collective nuclear degrees of freedom. To justify these assertions, this perspective provides a broad overview of the current state of soft materials design in the most pressing fields, reviews existing computational paradigms for QM calculations in soft materials, and discusses emerging methodologies for systematically incorporating QM phenomena into soft materials modeling. The article concludes with a perspective on the future of CG electronic prediction models, and posits potentially fruitful future directions for researchers to pursue.

# Quantum Mechanical Challenges in Soft Materials

We begin by outlining a subset of soft materials design challenges for which purely classical modeling efforts are insufficient, and QM phenomenology must be incorporated (Figure 1). These topics do not encompass all possible examples of non-trivial QM phenomena in soft materials, but only those most familiar to the authors.

## Controlling Degradation for Sustainable Polymers

With 6.3 billion metric tons of plastic waste now contaminating the Earth,[1] the pursuit of sustainable polymers is a task that can brook no delay. To achieve sustainability, we need robust end-of-life management, which includes plastic waste collection, reuse, and recycling. Research in plastics degradation pathways can help to develop strategies for end-of-life management.[2,3] For instance, advanced chemical recycling transforms petroleum-based plastics into value-added materials via so-called upcycling, valorization, or circular chemistry.[4–6] The

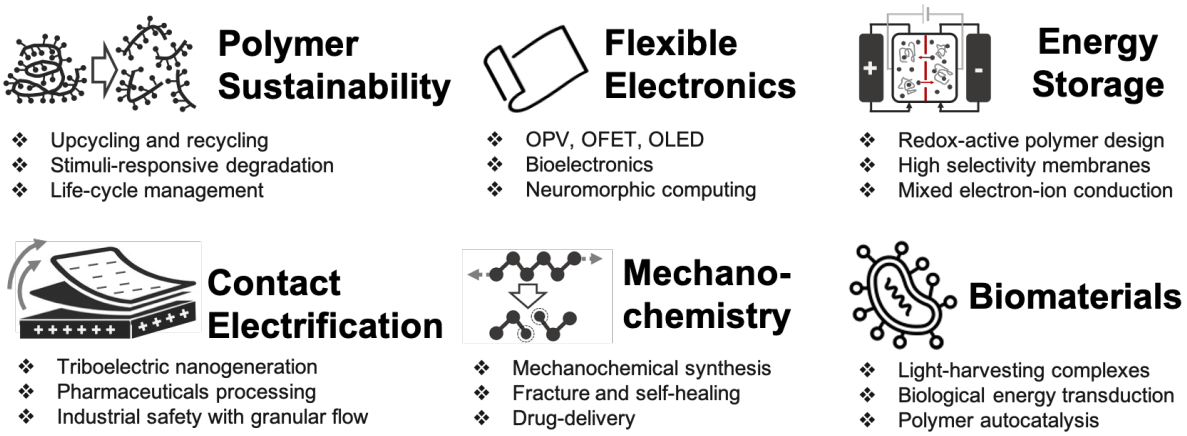



Figure 1: Frontiers of soft materials design with significant QM design components.

most successful example of upcycling is depolymerizing polyethylene terephthalate (PET) into its pristine resin.[1,6] The progress with polyethylene (PE), polypropylene (PP), polyvinyl chloride (PVC), and polystyrene (PS) has been slower because of their high ceiling temperatures.[1,6] Recent efforts are already underway to retain the polymer backbone and repurpose the pendant functional groups to exploit novel degradation pathways[1,3] so that recycling plastics can be competitive with commodity polymers in terms of cost and performance.[2,3] On the other hand, for the reuse of biomass-derived polymers, it has been found that the degradation rate varies widely and is dependent on temperature, humidity, morphology and oxygen availability.[3] For example, polylactide (PLA) requires industrial composting conditions such as high humidity and high temperature.[7] Scalable computational QM methods applicable at the mesocopic length scales offer insightful opportunities to reveal the mechanisms of chain scission, oxidation, cross-linking or catalytic activation during polymer degradation, as well as their relationships to multiscale morphology.

Understanding the roles of polymer crystallinity and ambient conditions on polymer degradation remains an experimentally critical, but computationally unexplored, avenue of sustainable polymer science.[4,5,8] In particular, there is a near complete absence of computational work on this topic due to the severe challenge of performing scalable QM calculations, with sufficient statistical sampling, across diverse polymer morphologies and chemistries.

Experimentally, the presence of crystallinity is known to drastically slow catalytic and biotic degradation pathways.[4,5,8] Even for amorphous polymers, it has been found that ordered chain segments with crystal-like properties are embedded in the polymer matrix, and that the degradation rate is controlled by semi-crystallinity.[5,8] Little is also understood about the nature and role of the antagonists that induce such degradation in polymers, such as UV radiation and singlet oxygen, which are hypothesized to induce bond breaking that deteriorates mechanical and optical properties.[9] Additives like excited-state quenchers, peroxide decomposers, and free radical scavengers can effectively inhibit undesirable photooxidative degradation,[9] but many additives are known to poison catalysts or deactivate catalyst sites during recycling. The search for suitable additives needs to consider excited state energy transfer between polymers and additive stabilizers during the exposure to UV radiation and oxygen, and to investigate the effects of these additives on the multiscale morphologies. The power of QM molecular modeling to elucidate degradation pathways in soft materials represents a currently untapped means for advancing soft materials design.

## Improving Redox Activity and Energy Storage in Soft Materials

The pressing need for clean, renewable energy necessitates the creation of alternative materials capable of energy storage. Redox flow batteries (RFBs) have become one of the most promising soft materials options due to their excellent safety, efficiency, flexibility, and scalability.[10,11] RFBs process electrochemical energy using two redox-active components dissolved in liquids that are stored and pumped through external tanks outside of the electrochemical cell. Inside the cell, two electrolyte circuits are separated by an ion-selective membrane where ion exchange occurs but the redox-active species are excluded. In the past decade, small organic compounds including $\pi$-conjugated molecules with carbonyl groups, fused ring aromatic molecules, stable radicals and organometallic complexes have served as potential redox-active centers.[10,11] However, these small molecule based RFBs suffer from the crossover of redox-active molecules, which reduces the Coulombic efficiency of the cell

and leads to capacity degradation during long-term cycling.[12] Additionally, the commonly used ion-exchange membranes necessary for species separation, such as Nafion, Fumasep, and Selemion, account for almost 40% of the cost of the reaction cell, impeding the scalability of RFBs.[13] Recently, the development of redox-active polymers has become a promising alternative for the RFB community due to their low cost and ability to suppress the crossover of the redox components.[10–13] However, there are still several QM challenges that limit polymeric RFB technologies, such as sluggish redox kinetics and unclear charge transport mechanisms.

QM methods are needed to accelerate the design of RFB materials, but are limited by the high cost of QC associated with the length scales of polymeric materials. Specific RFB topics demanding a QM treatment include (i) the sensitivity of redox-active molecules to the multiscale morphology, with side reactions being triggered under flow conditions or after several charging cycles;[14,15] and (ii) increasing the molecular weight of redox-active polymers dramatically varies the morphology of suspensions,[12–14] for instance, from suspending oligomers to single-chain aggregates, multi-chains aggregates or colloids. The general Dahms-Ruff formalism[16] is unable to describe the charge transport mechanisms for these complicated molecular environments, where electrons can migrate via intra and inter chain mechanisms, inside the polymer aggregate, or on the colloid surface. Only a few works have attempted to model the charge transport of nonconjugated redox-active polymers in electrolytes,[17–20] with none employing a QM level of fidelity at the required length scales.[21] The ability to predict the QM coupling between electronic states and the resulting electrostatic environments that control polymer structure and mechanics is currently absent from the literature and is another means by which QM modeling could transform soft materials design.

## Solving the Mystery of Contact Electrification in Soft Materials

Contact electrification is a ubiquitous phenomenon in which transfer of electrostatic charge occurs when two surfaces are brought into contact and then separated. Contact electrification is a double edged sword because it can cause billion-dollar losses or safety issues for indus-

try,[22,23] but also presents a promising pathway for the development of self-powered sensors and energy harvesting technologies.[24–26] From a destructive perspective, contact charging occurs when the manufacturing processes of polymers and pharmaceuticals involve flowing granular materials, and the resulting electrostatic charge leads to non-uniform blending in the products or undesirable adhesion at the reactor wall. More severely, processing or transportation of materials in powdered form with large surface-to-volume ratios can accumulate intensive contact charge, triggering explosion hazards. Also, unwanted electrical discharges caused by contact charging can inflict damage on the electronics components in semiconductors that incurs a massive financial loss to the electronics industry. Although the omnipresent contact charge could be detrimental, it has been productively employed in the application of self-powered sensors and energy harvesters.[24–26] Specifically, triboelectric nanogenerators harvest wasted mechanical energy such as human motion, sound vibration, mechanical triggering, rotating tires, wind, and flowing water. Understanding the molecular and QM foundations of contact electrification can promote the design of triboelectric nanogenerators, and reduce industrial costs, loss of life, and property damage.

Despite its long history, the physical origin of contact electrification in soft materials is an unresolved issue due to its non-equilibrium nature and the complex mechanisms which occur over a wide range of length and time scales.[23,27] It has been found that the formation of surface electrons is dependent on the atomic structure at the material surface (1-100 nm), surface roughness at the mesoscale (>1 $\mu$m), and particle size at the macroscale (10-1000 $\mu$m).[23] Also, it has been reported that various mechanisms, such as electron transfer, ion exchange, material transfer, and mechanochemical fracture (*i.e.*, altering or breaking of chemical bonds) can trigger contact certification.[27] However, explanations for the formation of surface charges on various materials are constantly under debate. Conventional theoretical efforts to incorporate multiple mechanisms throughout the whole range of length and time scales are futile, with few attempting to model contact electrification with QC accuracy. To make progress in this difficult area of soft materials modeling, scalable QM methods are

needed that can provide QM-informed characterizations of electron tunnelling, radical formation, and bond-breaking at the mesoscopic length scales underpinning interfacial contact.

## Managing Polymer Mechanochemistry

Since the mechanical degradation of polymers was observed in the 1930s, mechanically induced covalent bond scission has been regarded as a degradation pathway that limits soft materials lifetimes at both molecular and macroscopic scales.[28,29] In the past decade, efforts have been devoted to redirecting mechanical energy (*e.g.* by compression, shear, or friction) to productive chemical transformations.[28–31] The science and technology utilizing the influence of mechanical actions on chemical reactions is denoted "mechanochemistry",[28,29,31] and has sparked unprecedented developments in the fields of synthesis,[29,32] drug delivery,[30,33] and self-healing materials.[34,35] Mechanochemical synthesis is often solvent-free and can allow access to products unavailable by traditional reaction pathways since potential energy surfaces can be distorted by external forces.[36–38] Mechanophores (force sensitive functional molecular motifs)[28,29] can be also used to transduce external force through attached polymer chains into the mechanically selective scission of weak bonds, and then trigger the release of drug molecules.[30,33] Much practical interest in the force-induced degradation of polymers also arises in the field of self-healing materials and their mechanically-induced fracture and/or degradation.[34,35]

To address these challenges it is vital to understand the electronic structure of mechanophores and soft materials under the influence of an external mechanical force. Although traditional *ab initio* approaches have connected the relationship between molecular structure and reactivity,[36–38] mechanophores and self-healing polymers are sensitive to their molecular environment and the heterogeneous interface where the substrates transfer kinetic energy to reacting particles. Furthermore, while many studies on these materials have occurred in the solution-phase, the knowledge gained from these efforts is not always valid for solid-state mechanochemical reactions, and current theories are not able to thoroughly describe reac-

tion mechanisms.[32] With regard to modeling mechanical response in the elastic, plastic flow, and strain hardening regimes, significant work has been done using classical models,[39,40] but works incorporating QC-consistent reactive potentials are generally absent. Notably, the molecular determinants of mechanical response for polymer nanocomposites undergoing extreme deformation are still poorly understood because modeling cannot meet the requirements of long timescales and QM reactivity simultaneously. The mechanical behavior of polymers represents a crucial application in which existing modeling paradigms are lacking, and new QM methods capable of dissecting reactivity within nano to mesoscale soft material heterogeneities are needed.

## Architecting the Next Generation of Flexible Electronics

Over the past decades, the research and development of flexible electronics using organic semiconductors (OSCs) has rapidly expanded due to the merits of process simplicity, reduced fabrication cost, and mechanical pliability. Remarkable accomplishments have occurred in flexible electronic devices such as organic light-emitting diodes (OLEDs),[41,42] organic photovoltaics (OPVs),[43,44] organic field-effect transistors (OFETs),[45,46] as well as biomedical applications.[47,48] Recently, state-of-the-art OLEDs were theoretically estimated to achieve 100% internal quantum efficiencies.[41,42] For OPVs, the power conversion efficiencies have surpassed 18%,[43,44] a value that was believed to be unreachable some years ago. Both small-molecule and polymer based OFETs have achieved high charge mobilities exceeding 10 $cm^2V^{-1}s^{-1}$, which approaches the mobility regime of multi-crystalline silicon.[45,46] As the device performance of flexible electronics reaches the level of commercialization, the major challenge is the development of effective manufacturing processes that transition OSCs from laboratory-scale devices into large area devices without sacrificing their performance.[44,46,49] The most desirable manufacturing approach should be solution-based printing processes (*e.g.* roll-to-roll printing, doctor-blade coating, slot die coating, screening, gravure printing, and inkjet printing),[44,49] which exhibit low cost, high feasibility of customization, and little dif-

ferentiation from the current mature printing technologies. Moreover, emerging applications in bioelectronic and neuromorphic computing are beginning to utilize ionic and/or biological media in the presence of OSCs, further complicating the underlying multiscale design principles at play. As the length scales of processing in these applications concern the mesoscopic and macroscopic, efficiently coupling QM and classical mechanical simulations to design materials becomes ever more critical.

It could be reasonably argued that the field of OSCs represents the most successful integration of QM theory into soft materials development to-date.[50] The success of OSC-based devices can be attributed to the accumulated understanding in molecular design principles, structure-performance relationships, optoelectronic mechanisms, and device architecture optimizations, which involve knowledge spanning the micro-scale to macro-scale, and quantum to classical mechanics. In particular, multiscale simulations that combine bottom-up CG simulation, backmapping techniques, and/or QC calculations have revealed numerous fundamental insights about OSCs.[51–54] Although existing modeling methodologies will contribute to the progress of upscaling OSCs utilizing advanced processing techniques, they are extremely time-consuming to employ when both morphological evolution under various flow conditions and the optoelectronic properties of the final state need to be accounted for. These factors become even more pressing in the emerging neuromorphic and bioelectronic applications that utilize ionic currents to modulate electronic behavior.[55] Scalable CG models method that can simultaneously model bulk morphology, electronic structure, and mechanical response under various processing conditions would be transformative to the field.[44,45,47,49]

## Mimicking Biological Materials

Coincident with the discussed applications of predominantly synthetic soft materials is the emerging significance of understanding QM effects in biomaterials. Studies of electron transfer in proteins are going on nearly fifty years.[56] Understanding QM transport in photosyn-

thetic and light-harvesting complexes has occupied an enormous fraction of modern chemical research agendas.[57–59] Biological energy transduction in active materials typically derives from multi-electron state changes in ATP, for which QM effects are fundamental.[60] More recently, researchers are trying to understand how to manipulate autocatalytic processes to induce biology-inspired self-assembly pathways in synthetic materials.[61] As time goes on, it becomes clear that biology has effectively integrated QM phenomena over many length scales of intrinsically heterogeneous and disordered soft materials morphologies, making biological materials a fascinating grounds for integrating both QM and classical mechanical modeling.

Significant previous method development efforts have gone into addressing many of these challenging biological design problems. Arguably, biological materials have been the motivating force in the development of multiple distinct sub-fields of molecular modeling (*e.g.* CG theory, all-atom (AA) force-field development, quantum dynamics, and non-equilibrium statistical mechanics). In the realm of interfacing QM with classical simulations of biologically relevant length scales, QM/MM has been the workhorse for including QM effects into biological systems, but these efforts are often limited to small system sizes.[58,59] Recently, Mironenko and Voth have introduced the QM/CG-MM approach to the accelerate QC calculations in complex solvation environments, extending the accessible length and time scales for QM/MM calculations.[62] CG models have been instrumental in understanding the folding pathways of proteins, but can only include arguments about QM effects in an *ad hoc* fashion. The field of electron transfer pathways in proteins represents one of the most fascinating fields in which QM and classical modeling have interacted, but incorporating chemically-specific models in a computationally efficient manner still proves enormously challenging. Just as the study of biomaterials has benefitted from efforts in systematic "bottom-up" and "top-down" coarse-graining efforts, one can envision a systematic CG formalism applied to QM predictions in biological materials could potentially have enormous impact.

## Limitations of Current Molecular Modeling Paradigms

To motivate the introduction of a new paradigm for incorporating QM modeling into soft materials research, we briefly review the state-of-the-art for scalable QC, as well as systematic CG techniques for both QM and classical systems. The multiscale integration of QC, molecular dynamics (MD) simulation, and CG modeling is shown in Figure 2.

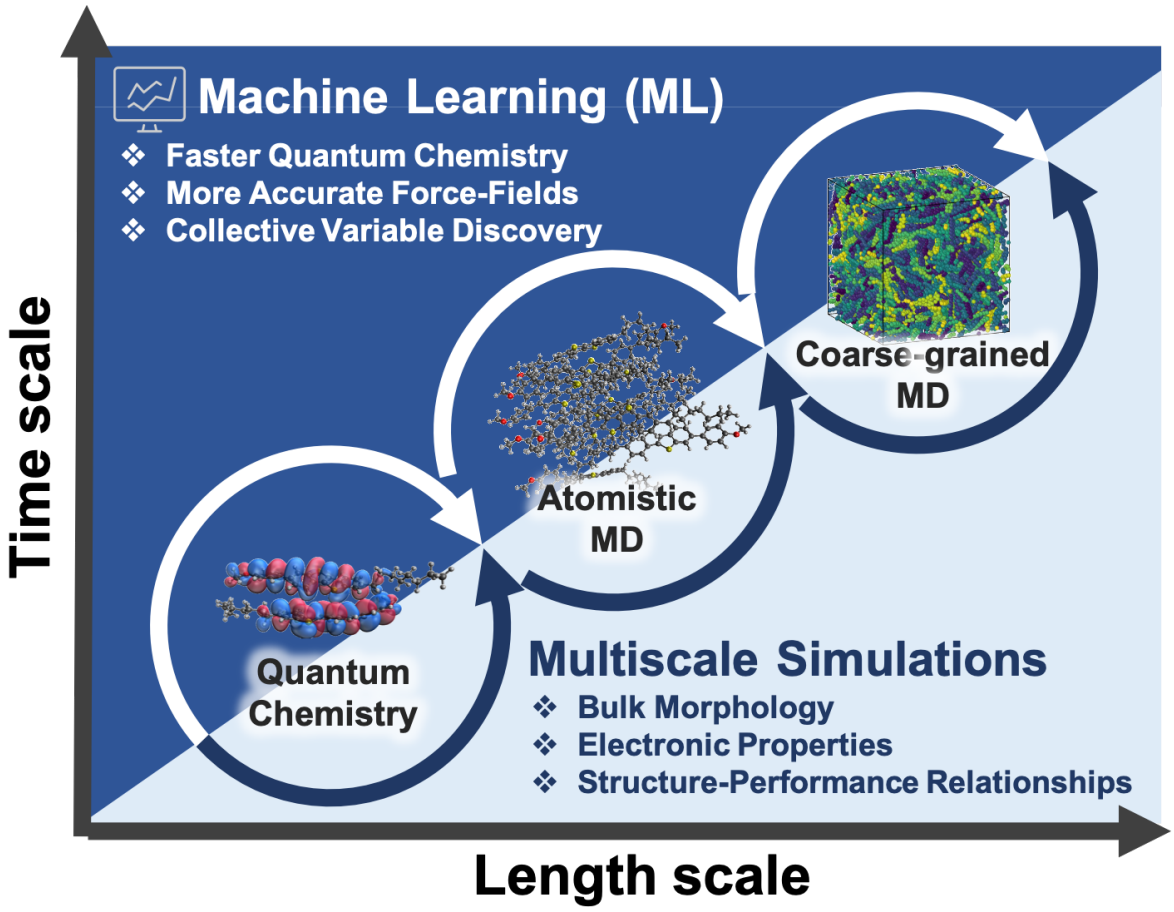



Figure 2: Existing paradigms in multiscale molecular modeling including QC, MD simulation, and CG techniques, as well as the aspects to which ML has contributed.

## Quantum Chemistry

In the past fifty years, QC methods have advanced rapidly with simulations on the ∼1 nm length scale becoming routine.[63,64] Two of the most prominent efforts for reaching multi-nm length scales have been the density functional theory tight binding (DFTB)[65,66] and fragment molecular orbital (FMO)[67,68] approaches. In FMO, approximately $O(N)$ scaling in the number of basis functions is possible, with single-point calculations of $10^4$ atoms being increasingly common. Similarly, DFTB calculations can permit geometry optimizations of $10^3 - 10^4$ atoms. In notable efforts, combinations of DFTB and FMO with efficient parallelization have facilitated geometry optimizations of one million atom systems[69]—a scale unthinkable only a decade ago.

Despite these impressive advancements, atomistic QC remains far removed from the required spatiotemporal scales of soft materials. While a one million atom DFTB+FMO calculation will provide a single minimum energy geometry for a ∼20 nm cubic box, soft materials chemistry applications necessitate QC characterizations at similar or larger length scales, averaged over the thermodynamic ensemble, and for multiple distinct state points and chemistries. These computational requirements render even the best scaling QC methods insufficient to address the ubiquitous molecular design question in soft materials: "*How do chemistry, morphology, and thermodynamic conditions dictate electronic processes in non-crystalline soft materials?*"

Notwithstanding this lack of scalability, arguably the most significant deficiency of existing QC methods involves the fact that, even if the excessive computational cost required to electronically characterize soft materials at mesoscopic scales could be paid, the effort would represent an extreme inefficiency in the utilization of high-performance computing (HPC) resources. When performing QC calculations at mesoscopic scales, a researcher typically separates the total system (10-100 nm) into subsystems (1-5 nm) for local QC characterization. These local calculations, and couplings between them, are then utilized to piece together an approximation of the original QM Hamiltonian. However, the nature of Boltzmann sampling

will lead to repeatedly performing nearly redundant QC calculations on subsystems throughout the original system. In many cases, only a few collective variables (*e.g.* dihedrals, kink formation, Huang-Rhys active normal modes) within subsystems induce large changes in the relevant electronic structure, leading to dramatic expenditures of HPC time to compute electronic structure variations that could in principle be projected onto a reduced set of collective coordinates.

An important consideration relates to the potential use of periodic boundary conditions in QC, as this is a standard way to include bulk effects on the electronic structure of crystalline materials. However, as soft materials are characterized by liquid-like morphological heterogeneity over nanoscopic and mesoscopic length scales, computational QM models exploiting periodicity via Bloch's theorem on small subsystems cannot be reasonably applied to model the electronic structure of disordered soft materials. Such approaches will dramatically underestimate the influence of structural disorder on the resulting electronic structure, resulting in artificially ordered systems. Similarly, a large enough subsystem that captures the requisite morphological disorder is likely to exceed the capabilities of modern QC computational power in the same way as for methods using localized basis sets.

An interesting path forward for QC concerns the systematic development of effective Hamiltonians in which portions of Hilbert space are averaged out, in what is effectively CG for electronic degrees of freedom. While the scaling of molecular QM calculations with such methods does not eliminate the scaling in the number of nuclear degrees of freedom, the reduction in electronic degrees of freedom can have a dramatic effect in reducing computational cost; in essence this strategy underlies the highly successful DFTB and FMO approaches. To these ends, a number of works have attempted to construct rigorous and systematic approaches for performing dimensionality reduction of electronic Hamiltonians.[70–73] A classic summary of these efforts is provided by Durand and Malrieu.[74] While the theoretical formalisms for constructing such models are generally well-defined, in practice they are rarely attempted due to limitations in accuracy and computational complexity. An adjacent field

of electronic structure theory involves the development of pseudopotentials[75] which parameterize effective nuclear potentials that incorporate the shielding from low energy electrons so that simpler models involving only higher energy electrons exhibit reduced computational cost. While these classes of methods are incredibly powerful across a broad range of materials science, they are still extremely limited in their application to the spatiotemporal scales necessary to characterize QM properties in soft materials. Moreover, when it comes to potentially integrating these calculations with CG structural prediction models that proliferate soft materials modeling, it is crucial to note that renormalized electronic structure models do not CG over nuclear degrees of freedom, and thus are still incompatible with the CG simulation models ubiquitous in soft materials modeling.

## Coarse-Graining and Backmapping

While QC calculations at mesoscopic scales are inhibited by the retention of all atomic degrees of freedom, morphology simulations at mesoscopic scales commonly benefit from the use of CG representations.[52,76] Classical simulations using AA force-fields are typically limited to 10's of nm and 100's of ns with the accuracy of the predictions depending strongly on parameterization. To overcome these limitations, researchers employ CG modeling paradigms in which the least important AA degrees of freedom are averaged into effective CG pseudoatoms. By tracking fewer degrees of freedom, as well as employing softer intermolecular potentials, CG modeling permits access to mesoscopic spatiotemporal scales while maintaining thermodynamically rigorous links between Hamiltonians at different resolutions.[77] Furthermore, the act of decimating degrees of freedom results in models that exhibit improved conceptual clarity relative to fully-detailed AA simulations.

In recent years, advances in CG methodologies have improved both the quality and availability of CG modeling, profoundly impacting soft materials design. Multiple groups have worked to understand both the representability and information errors associated with CG representations.[78–80] Many modifications and improvements to CG parameterization proto-

cols (*e.g.* Iterative Boltzmann Inverse (IBI),[81] Force-Matching,[82] and Relative Entropy[83]) have been developed with improved transferability and accuracy. The chemically generalizable Martini model has evolved to significant levels of complexity and prediction power across a broad sampling of chemical space.[84] In conjunction with these developments, considerable software efforts have been invested to make CG modeling strategies available to the broader community.[85–87] As a result, CG modeling forms the core computational tool set for soft materials simulations traditionally focused on structural and thermodynamic predictions.

As QC calculations require specification of all atomic degrees of freedom, and CG modeling utilizes averages over atomic degrees of freedom, additional methodologies are required for bridging simulations at these different length scales. At present, the state-of-the-art (Figure 3) involves (i) equilibrating morphologies using CG structural prediction models, (ii) backmapping AA coordinates onto CG configurations (a one-to-many problem), followed by (iii) *ad nauseam* QC calculations to characterize local electronic structure.[88] Solutions to the underdetermined backmapping problem continue to be posed,[89–96] but often rely on heuristics, ignore proper ensemble averaging, or limit chemical transferability. Moreover, backmapping strategies significantly complicate computational workflows by requiring interfacing the inputs and outputs of multiple independent simulation codes. Taken together, these considerations leave the field in the regrettable state of lacking computationally scalable and thermodynamically rigorous strategies for deriving electronic structure at the requisite spatiotemporal scales for soft materials.

Orthogonal to the standard paradigm for connecting QC and CG modeling methods are efforts to develop phenomenological models that make electronic predictions using CG descriptions of collective degrees of freedom. In their simplest and most common form, these models utilize normal mode representations of configurational changes and linear electron-phonon couplings parameterized to experiment or QC calculations. Attempts to go beyond linear descriptions of configurational changes and electronic couplings are generally only permitted when one possesses extensive pre-existing knowledge and physical intuition that can

inform appropriate mathematical forms of approximate Hamiltonians. Despite the crudeness of such CG electronic Hamiltonians, these approaches have been enormously useful[97–100] in understanding the electronic properties of organic materials despite their phenomenological nature and distinct lack of generalizability to systems for which physical intuition is absent. It is also useful to note that such models ignore the degeneracy of decimated degrees of freedom within the averaged CG model, and represent a mean value approximation of the electronic structure at a single thermodynamic state point. The lack of systematic CG methodologies for bottom-up CG of electronic predictions in soft materials systems represents a significant opportunity for advancing the *in silico* understanding and design of QM phenomena in soft materials.

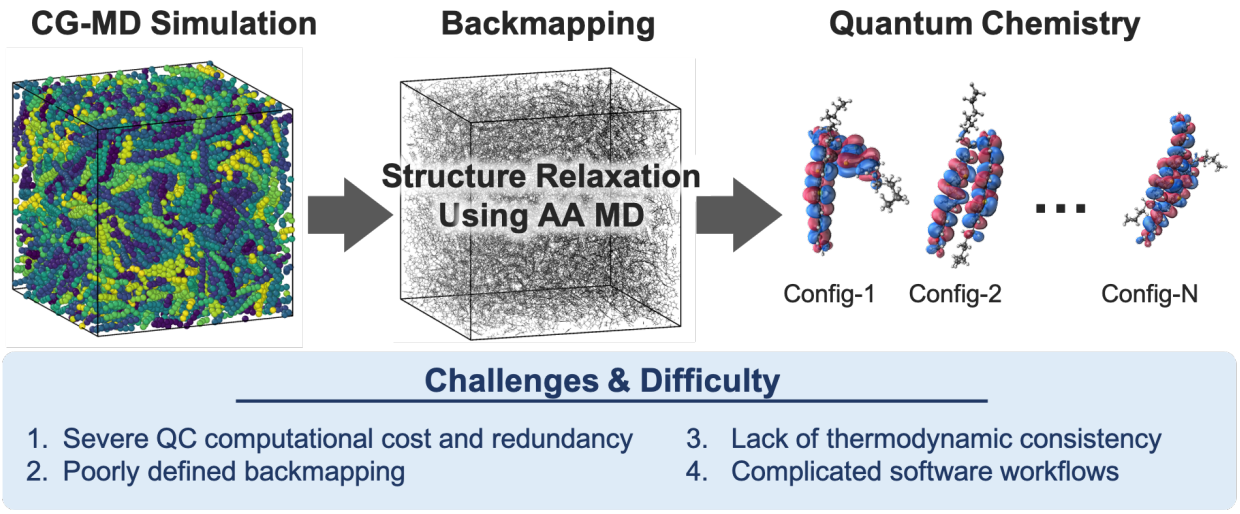



Figure 3: Current strategy for incorporating QM effects into soft materials predictions and its challenges.

## Coarse-Grained Electronic Models

The preceding discussion motivates the need for molecular models capable of accessing significantly longer spatiotemporal scales than are reachable using current QC methods. In the remainder of this article we introduce recent work in our and other research groups on the development of electronic prediction models that operate at the CG model resolution.

Within our research group, we refer to these models as Electronic Coarse-Graining (ECG) models. ECG models represent a family of techniques introduced by Jackson[101] targeted at the systematic, bottom-up CG of configurationally dependent electronic property predictions in soft materials.

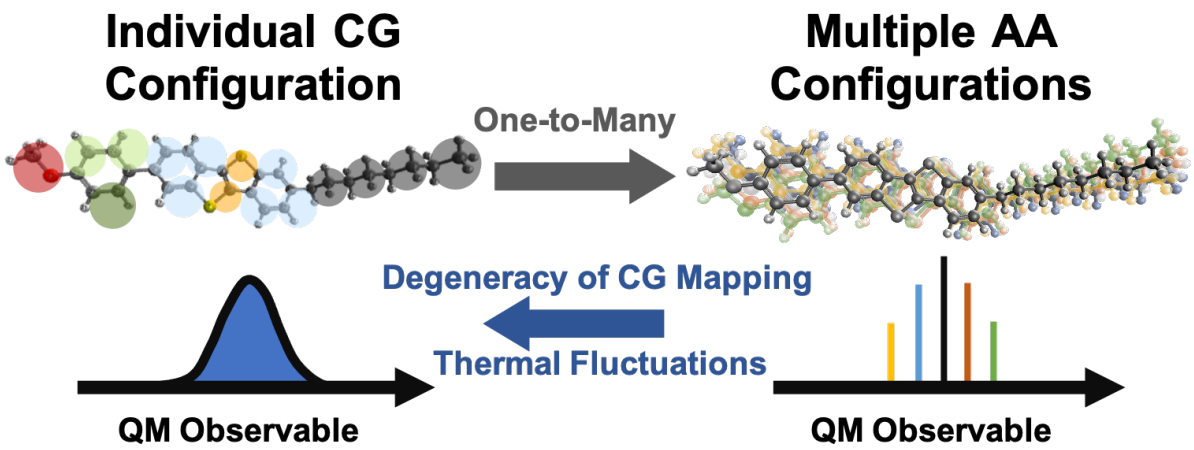



Figure 4: Basic goal of ECG methods: effectively mapping targeted electronic property distributions (derived from AA configuration space) directly to CG configuration space.

## The Development of ECG Methods

Critical to the development of CG electronic structure prediction models is the realization that a single CG molecular configuration corresponds to a collection of AA configurations. As depicted in Figure 4, this so-called "one-to-many" mapping between CG and AA configurations means that any AA-derived property can be regarded as an intrinsically noisy probability distribution of the property at the CG resolution. The noise on this distribution can be rigorously related to (i) the degeneracy of the CG mapping operator (how the atoms are combined into CG beads) and (ii) the thermodynamic state of the system (*e.g.* how strong the thermal fluctuations are in the AA model). For electronic properties of interest to QM modeling in soft materials, analytical approaches are, at present, incapable of directly computing the structure of this noisy distribution function from the "bottom-up". In these contexts, machine learning (ML) has proven a powerful method for mapping AA

electronic structure directly to the CG representation, and forms the basis of most recent work. It is important to note that by mapping the distribution of AA properties directly to the CG representation, one bypasses the need to bridge QC and CG simulations via the use of backmapping methodologies and *ad naeseam* QC. If one can exactly learn this noisy distribution as a function of CG representation, then the CG model predictions will correspond exactly to those derived from a computationally intensive backmapping approach.

The development of ECG models in our group employs the scheme depicted in Figure 5. This scheme consists of the following steps: (i) a sampling of AA configuration space at a specific thermodynamic state point (using Monte Carlo or MD simulations) which is used to parameterize a CG structural prediction model (*e.g.* IBI, force-matching, relative entropy), (ii) the labeling of the sampled AA configurations using QC calculations of the targeted electronic properties of interest, (iii) the projection of the AA configurations into a CG representation, (iv) the transformation of this dataset into a ML-readable format and (v) the training of a supervised ML regression model that maps CG configurations to AA QC. To date, a variety of ML architectures (Figure 6) have been implemented within ECG models. The result of this workflow is a model that provides electronic property predictions directly from the CG model resolution that are consistent with AA structural (*e.g.* Optimized Potentials for Liquid Simulations, OPLS) and QC (*e.g.* DFT) prediction models. While these methods primarily employ standard AA force-fields (*e.g.* OPLS) with classical MD to sample configuration spaces that are then characterized with QC, they could straightforwardly be replaced with *ab initio* potential energy surfaces to incorporate complex electronic nuclear couplings back into the configurational sampling. Due to the additional cost of such a procedure, especially at the length scales (>10 nm) of present interest, we have not pursued these approaches. However, in the context of incorporating bond-breaking in chemical reactivity, or excitation induced rigidity in polymers, such adaptations of the sampling procedure will be critical.

The first demonstration of ECG modeling was performed by Jackson *et al.* in which

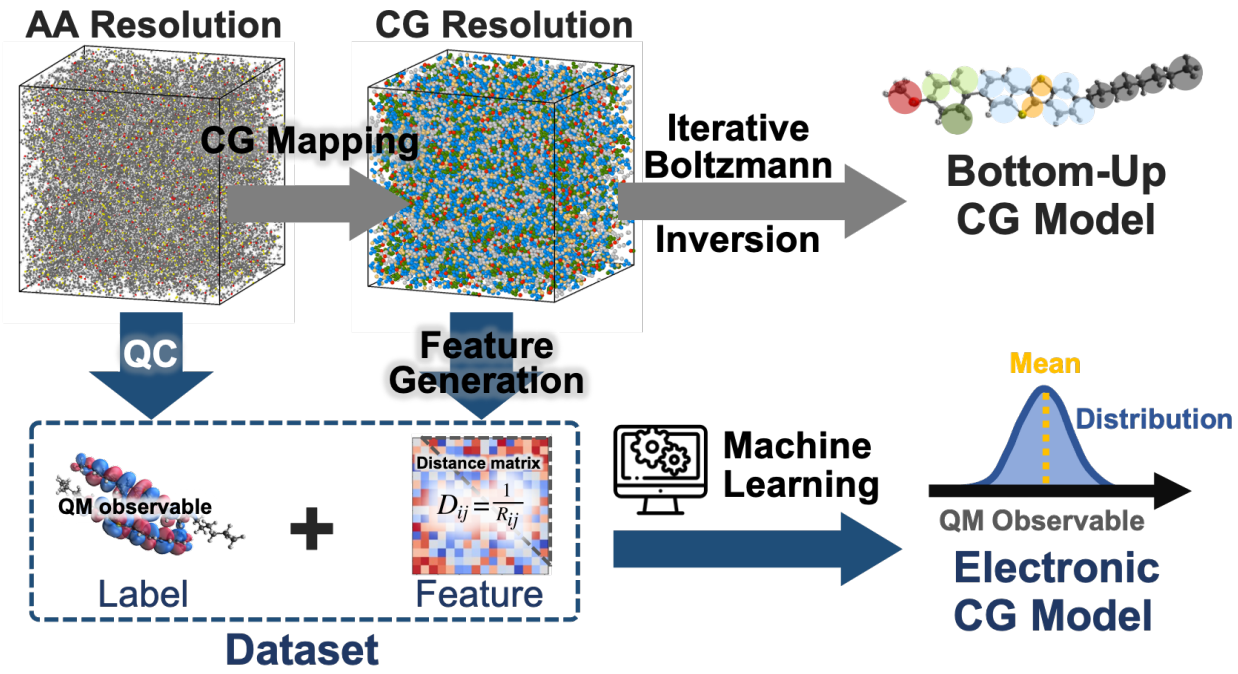



Figure 5: Workflow of ECG methods. The grey arrows sketch the processes of establishing bottom-up CG models. The blue arrows describe the workflow from building a dataset, to training a ML model, to predicting electronic observables for a given CG configuration.

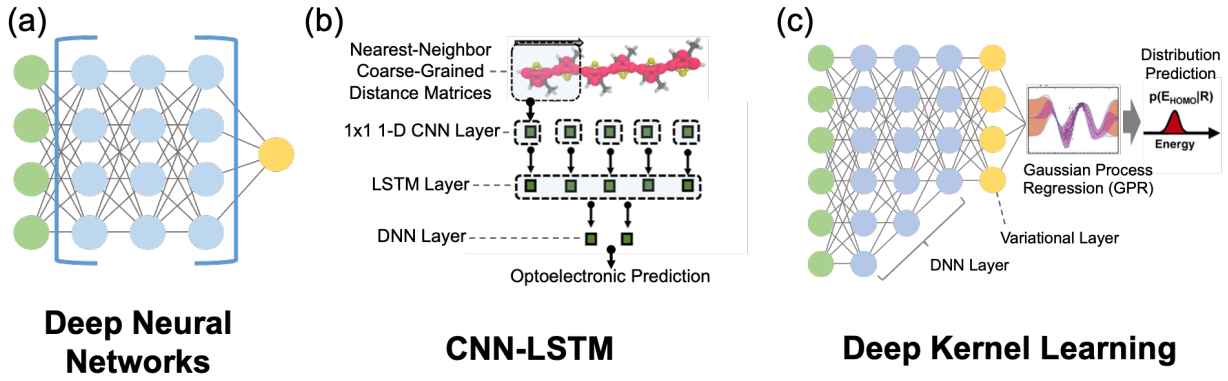



Figure 6: Various machine learning architectures employed in ECG works: a) feed-forward deep neural networks, b) 1D convolutional and long short-term memory networks, c) deep kernel learning with Gaussian process regression. Reproduced with permission from the American Chemical Society[102] and the American Institute of Physics.[103]

feed-forward deep neural networks (DNN, see Figure 6a) were employed to learn the mean values of the six highest occupied molecular orbital (HOMO) energies and intermolecular electronic couplings of thiophene-based chemistries directly from the CG resolution (Figure 7a).[101] In all models the CG configuration was featurized using the inverse distance matrix associated with the CG configuration. Models were trained at multiple temperatures, intramolecular rigidity constraints, and CG resolutions to assess the impact of thermodynamics and CG mapping operator selection on model performance. In nearly all cases below a critical threshold of CG model resolution, prediction accuracies of $\sim$ 10 meV's (Figure 7a,b) were achievable directly from the CG resolution. In the context of incorporating these predictions into *e.g.* semi-classical rate theories for charge transport, these errors are comparable to those derived from variations in QC method or basis set,[104] with the important distinction that all prediction errors in this model are "intrinsic" in that they derive from the degeneracy of the CG mapping operator. Moreover, a computational acceleration of $10^6$ was reported relative to the conventional backmapping and *ad nauseam* QC paradigm, along with a dramatic improvement in the simplicity of the computational workflow. This proof-of-principle demonstrated that QM prediction models can operate at the same resolution as the underlying CG structural prediction model, dramatically improving computational efficiency.

Building on this work, Jackson introduced a modification to the DNN architecture for ECG adapted to polymeric materials such that molecular weight transferable QM predictions could be achieved. This ECG model employed 1D convolutional and long short-term memory networks (CNN-LSTM) to learn a molecular weight transferable representation for CG polythiophene configurations transferable from a single repeat unit to 20 repeating units (Figure 6b).[102] Featurization was achieved by concatenating nearest-neighbor distance matrices between CG beads that entered into a convolutional layer prior to a LSTM layer. This model was trained over subsets of the underlying molecular weights, with testing data sets applied at molecular weights not in the training sets. As demonstrated in Figure 7c-d, this

CNN-LSTM version of ECG was able to successfully reproduce the absorption spectra, molecular orbital energies, and charge densities of polythiophenes obtained at the ZINDO/S level of QC. Contemporaneously, independent work by Simine and Rossky[105] introduced a CG electronic prediction model for the optical spectra of polythiophenes obtained at the Pariser-Parr-Pople level of theory with Configuration Interaction derived excited state energies. CG polythiophene chains were featurized as sequences of inter-thiophene dihedral angles which were subsequently processed by an LSTM. With the dramatically reduced computational cost of this model, simulations of 14,000 30mers of polythiophenes were trivially accessible, though results did not exhibit quantitative agreement with an employed backmapping algorithm. These works clearly demonstrate that CG electronic prediction models present a powerful means for bypassing existing QM modeling paradigms for conjugated polymers, enabling the production of QM properties at extended length scales crucial to soft materials.

With early works on ECG-style models studying predominantly thiophene-containing chemistries, recent work by Alessandri has extended ECG modeling to predict the energies and couplings of radical containing polymer chemistries directly from the CG resolution.[106] In this work, similar quantitative agreement was observed in the prediction of orbital energies and electronic couplings as for thiophene containing chemistries, though important insights regarding selection of the CG mapping operator were derived. Importantly, Alessandri compared the performance of an ECG model employing Martini-based structural predictions against a state-of-the-art backmapping protocol followed by *ad nauseam* QC. Good agreement between the backmapping derived and ECG models was obtained, but significant discrepancies were observed in the tails of the electronic density of states, which is particularly important as trap formation in the tails of bands typically limits charge carrier transport. The origin of this disagreement was attributed to the approximation of the noisy QM observable at the CG resolution as simply the mean value of the observable distribution. Of additional note was the important role of the accuracy of the structural predictions in the CG model, with IBI corrections to the Martini intermolecular potentials being required to ensure

quantitative agreement with the underlying AA model; models that did not employ this correction exhibited significant disagreement between the backmapped and ECG models. This implies that alternative CG parameterization algorithms such as force-matching or relative entropy, in which the structural distribution function is not the on-target property for reproduction, may exhibit inferior performance relative to IBI-parameterized models when uesd with ECG models. Finally, this work provided a quantitative comparison of the computational cost improvement for ECG relative to backmapping, for which the ECG model became dramatically more efficient when compared to the traditional backmapping with QC scheme after analyzing only 13 configurational snapshots. This acceleration is consistent with the previously posited $10^6$ speedup of ECG, and shows that ECG-style models can enable the computation of QM properties over broad samplings of thermodynamic and configurational state space relevant to designing soft materials.

Recently, our group has advanced supervised ML techniques for ECG that enable learning the CG electronic structure prediction beyond the common mean value approximation, treating the noisy distribution function within a Gaussian approximation. Sivaraman has introduced the deep kernel learning (DKL) framework employing Gaussian Process Regression to model the electronic property distribution throughout CG configuration space as a homoscedastic Gaussian distribution.[107] This framework allowed for the noisy observable at the CG resolution to be treated as a Gaussian distribution with variable mean and constant width. Moreover, the incorporation of a Gaussian distribution-derived variance facilitated the integration of ECG models with active learning strategies to achieve same level of accuracy with fewer training data.[107] ECG models using DKL could be efficiently trained over multiple state points using this combination of methods. While this effort represented a significant advance of ECG methods in moving from a mean value approximation to a Gaussian distribution, fundamentally there is *a priori* no reason to expect the variance of the distribution to be constant as a function of CG configuration (*i.e.* there is no expectation that the distribution be homoscedastic).

Recently, Maier has extended the ECG framework using DKL with approximate Gaussian processes to incorporate heteroscedasticity into ECG predictions of QM observables at the CG resolution (Figure 6c).[103] In this approach, the noisy observable is treated within a Gaussian approximation, the mean and width of which varies as a function of position in CG configuration space. Figure 7e indicates that the DKL-ECG method used in conjunction with IBI-derived CG structural prediction models could reproduce the exact QC valence electronic density of states for three different polymer chemistries relevant to charge transport. In comparing the performance of this model against the original ECG models that only learned the mean value of the distribution, significant improvement in reproducing the extremes of the observable distribution was observed. It was further shown that for the majority of CG mapping operators and resolutions the Gaussian approximation of the observable was suitable, though exceptions were found for extremely strong CG mappings. In the case of extremely fine CG mappings (*i.e.* the AA representation) this Gaussian distribution approaches a delta function, whereas for the case of very coarse mappings, multi-modal, non-Gaussian behavior was observed, which is anticipated based on recent work by Voth in the area of Ultra CG.[108] However, the heteroscedastic DKL ECG framework presents a manner of learning CG distributions of AA properties, notably QM properties, that are traditionally "lost" in CG model development. By knowing the structure of the learned distribution at all points in CG configuration space, stochastic samples of the AA observables can be generated that are consistent with the original ensemble. While this has been primarily shown in the context of molecular orbital energies and couplings, such a framework is straightforwardly transferable to a diverse set of AA properties including vibrational spectra, hydrogen bonding, optical spectra, NMR shifts, and charge density fluctuations. Importantly, DKL-based ECG provides a means of reproducing the statistically exact AA observable at the CG resolution without recourse to backmapping or *ad nauseam* QC, presenting an alternative modeling paradigm for the scalable analysis of QM properties in soft materials.

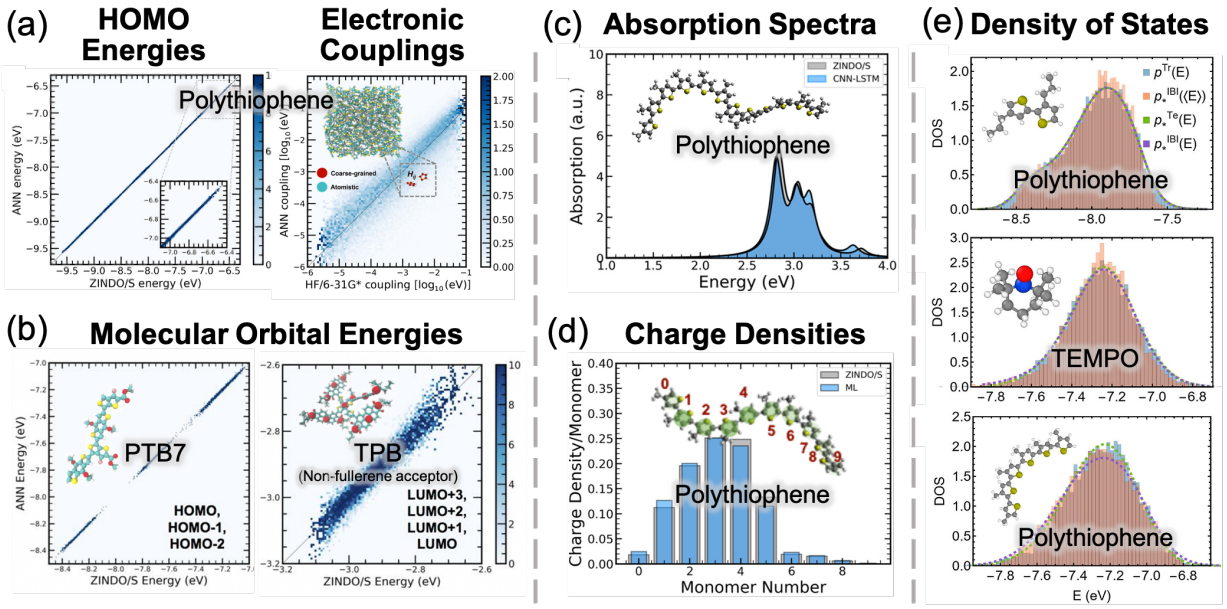



Figure 7: The ECG predictive performances for optoelectronic properties of various chemical compounds: a) six HOMO energies of polythiophene and electronic coupling of thiophene dimers, b) molecular orbital energies of PTB7 and TPB, c) absorption spectrum and d) charge density distribution of poly(3-methylthiophene), and e) electronic density of states of polytiophene and TEMPO. Reproduced with permission from the American Association for the Advancement of Science,[101] the American Institute of Physics,[103] and the American Chemical Society.[102]

## ECG Model Transferability

An aspect of soft materials design that typically confounds traditional CG modeling is the need to simulate structural and thermodynamic quantities over a broad range of thermal and processing conditions. Since application of the CG mapping operator destroys some of the entropy associated with the AA beads, CG models typically fail to properly capture the thermodynamic state dependence of property predictions, though emerging techniques are beginning to address this challenge.[109] These same considerations apply to the development of ECG models for QM property prediction, a point that has occupied considerable concern in our recent work.[107] Sivaraman examined the temperature transferability of DKL-ECG models trained over a 1000K temperature range for thiophene chemistries. In Figure 8a, it was observed that ECG models at high system temperatures displayed prediction accuracies at low temperatures that were comparable with the performance of DKL-ECG models trained at low temperature. However, the reverse was not true — models trained at low temperatures failed catastrophically when applied to configurations derived from high temperature simulations. These results agree with the physical intuition that the temperature should dictate the amount of configuration space that is explored during the sampling of AA configurations when the ECG training sets are constructed. In principle, the underlying QM prediction function can be better learned if a more diverse configurational space is incorporated into the training set. Although this physical intuition may be taken for granted (*i.e.* the more comprehensive the configurations in the dataset the better the model's performance), it is worth commenting that ECG models so far appear to avoid overfitting of the configurations when trained at high temperatures. Despite this success, it is important to note that while the effective mean of the distribution can be learned in a generally temperature transferable way by focusing ECG model training at high temperatures, the width and shape of the QM observable distribution has a temperature dependence that, at present, cannot be accounted for in ECG models. This challenging feature amounts to reincorporating the missing entropy of the CG model, and represents an interesting target for future

ECG model development.

While thermodynamic state point transferability is of particular interest to ECG modeling, a fascinating frontier for ECG model development concerns the creation of ECG models exhibiting chemical transferability. This fact is particularly useful as the mechanical and optoelectronic properties of polymers often exhibit strong dependencies on their molecular weight. In a naive conception of CG modeling, it would be required to train individual ECG models for polymers at varying molecular weights, significantly increasing the computational cost and complexity of developing ECG models. In previous work, we have demonstrated that, for a single chemistry, molecular weight transferability in polymer chains can be imbued using supervised ML approaches derived from sequence analysis (Figure 8b).[102] Interpolative performance of ECG models was generally excellent across the molecular weight spectrum, though extrapolative performance degraded strongly as a function of distance of the test data from the training data sets. Molecular weight transferability represents a logical first step in the development of broader chemical transferability in ECG models, a topic we will discuss in a later section.

## Building Physics into ECG Models

A concern with many models based in supervised ML is that predictions are not constrained in any way by fundamental physical laws, and that such deficiencies can result in catastrophic predictions when models extend beyond their training sets. In our research group, we operate with the philosophy that if a computational task can be achieved *efficiently* using existing molecular modeling paradigms, then it should. If it cannot, then supervised ML is a powerful tool to be exploited.

In an effort to integrate physics-based constraints into ML-based ECG models, we have established a $\Delta$-ML approach[110] that can leverage pre-existing knowledge of CG electronic structure in the form of model Hamiltonians.[102] We have developed ECG models that employ canonical tight-binding (TB) Hamiltonians describing electronic ground states and excited

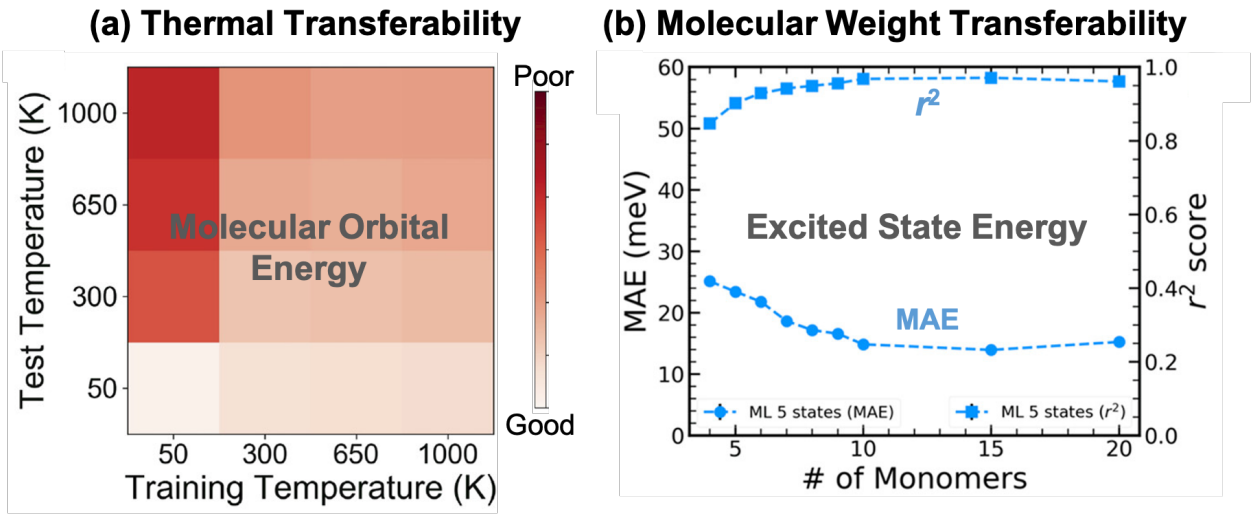



Figure 8: Demonstration of the ECG model capability in a) temperature transferability and b) molecular weight transferability. The heatmap shows the model performance (root-mean-squared error) for the prediction of HOMO energies of polythiophene. The line plot indicates the model accuracy for the lowest-lying excited-state energy of polythiophene at different molecular weights. Reproduced with permission from ref. 102 and 107. Copyright 2020 and 2022 American Chemical Society.

states at the CG resolution. In general, to parameterize the QM Hamiltonian for polymer systems, the diagonal elements (or the onsite energies) are often assigned to a constant value, and the off-diagonal elements (or the electronic coupling) are defined by phenomenological or approximate expressions, namely, a cosine function of the intermonomer dihedral angle for the ground state and terms incorporating transition dipole interactions for the excited states.[97–100] While the physics of the TB models can qualitatively capture the effects of the p-orbital overlap or point-dipole interactions, they cannot capture any details contained in the CG representation beyond known CG degrees of freedom, where human intuition is limited. The Δ-ML based ECG method uses the TB model to predict a first guess of the optoelectronic properties of interest, and then the difference between the guess and the exact QC result is learned by the ML. Δ-ML based ECG accounts for the additional local orientational degrees of freedom not captured by CG degrees of freedom included in the TB model. We believe that the concept of Δ-ML based ECG can be integrated with various phenomenological models so that any known physics can be incorporated in these models.

In our previous work on molecular weight transferability, significant improvements in both the interpolation and extrapolation across the studied molecular weight range was observed for models incorporating the simple known Hamiltonians.[102] An alternative strategy for constraining ECG models to known physics in future efforts is to build in the innate physical model as part of the loss function. Approaches similar to this have been developed in the physics informed ML models introduced by Karniadakis.[111] While this has not been attempted for ECG models, it forms an interesting avenue of pursuit for future efforts.

## Extracting Physical Insights from ECG Models

Another prevalent concern with the use of ML is the belief that by employing ML one effectively abandons all hope of physical insights and must treat the model as a "black-box". In our group, we consider the two broad goals of ECG models to be the following: to (i) dramatically decrease the computational cost of performing chemically specific QC characterizations of soft materials and (ii) develop physical insights into the collective variables that control changes in the electronic structure of soft materials. Whereas previous discussion has focused primarily on the utility of ECG models for achieving (i), here we describe how ECG models employing supervised ML can be repurposed to discover the dominant electronically active collective variables in soft materials for which no physical insights are known. This effort thematically parallels recent work utilizing ML to identify collective variables in soft materials.[112] This effort also overlaps significantly with the goal of identifying optimal CG mapping operators that capture maximal information at a given resolution.[78–80]

In general, CG representations of molecules and polymers are selected *ad hoc* via physical intuition, chemical intuition or convenience.[52,54] Although considerable efforts have been recently devoted to developing systematic CG mapping techniques, there is a lack of research on how to identify "optimal" mappings for the purpose of predicting observables.[78–80] For electronic predictions, it has been demonstrated that ECG models can be rapidly repurposed to identify optimal CG mapping operators and electronically active collective variables in

conjugated and radical containing polymers. As shown in Figure 9, ECG models can be trained at a variety of CG representations in order to identify the degradation of predictive performance as a function of CG mapping operator resolution.[101,106] Specifically, the drop of ECG performance in Figure 9a suggests the intermonomer dihedrals are a critical CG configurational degree of freedom in describing the electronic structure of conjugated molecules; this result is consistent with the previous four decades of purely theoretical research that employed approximate Hamiltonians using expressions incorporating the dihedral of the intermonomer torsional angle in thiophene chains. Similarly, in Figure 9b, work by Alessandri has shown that traditional CG mapping operators are not effective for reproducing electronic property predictions in CG models of radical transporting polymers, and that the N-O functional group must be maintained in CG resolution so the radical active nature could be captured. Consequently, in systems where the important collective variables that control electronic structure variations are not well-known, testing the performance of ECG models on a variety of CG mapping operators represents an interesting strategy for collective variable identification and the specification of optimal CG representations.

# Future Directions

## Multiscale Structure-Processing-Function Relationships for Organic Semiconductors

With the capability of predicting configurationally dependent electronic properties at the CG resolution using ECG, there is a significant promise for integrating existing ECG models with established CG simulation capabilities to explore the structure-processing-function relationships of OSCs with unprecedented scalability. Our group[113–116] along with others[117–121] have pioneered the use of anisotropic CG models that capture the multiscale morphologies of OSCs spanning $\pi$-stacking to domain formation. Since anisotropic $\pi$-conjugated features ($\pi$-stacking, inter-ring dihedral) are preserved at the CG resolution, the morphological

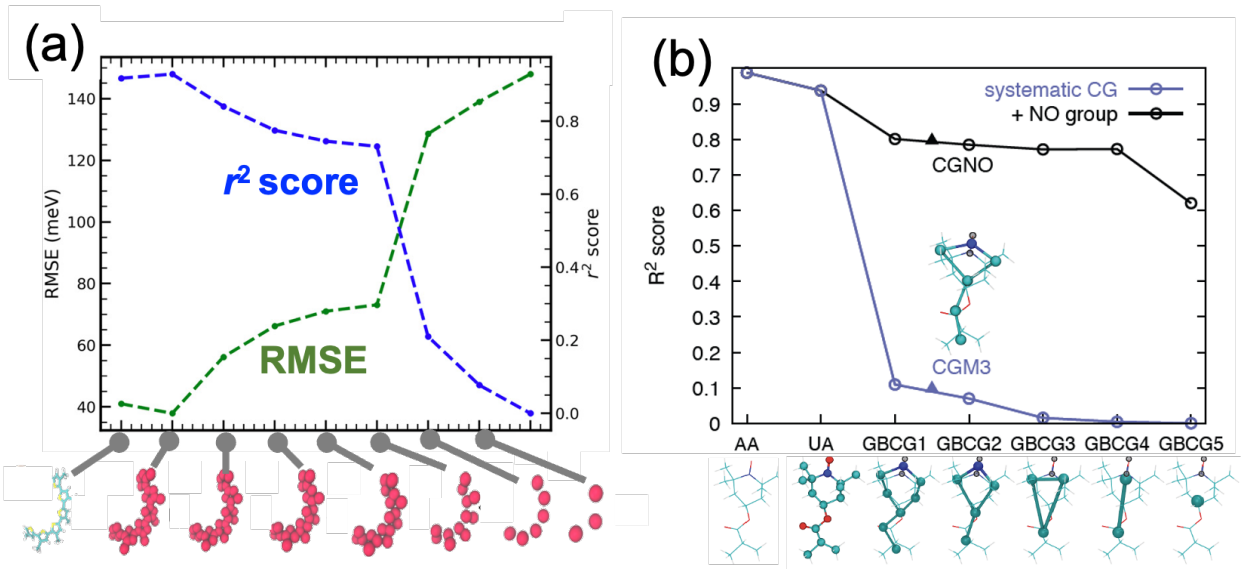



Figure 9: Examples of using ECG for collective variable discovery: the ECG model performances for predicting a) the HOMO energy of polythiophene as a function of CG resolution and b) the singly-occupied molecular orbital energy of a p-type radical polymer by using different CG representations. Reproduced with permission from the American Association for the Advancement of Science.[101]

dependence on their electronic properties can be systematically investigated by employing phenomenological analysis or TB Hamiltonians without implementing the backmapping process. The inclusion of these additional degrees of freedom in the CG structural prediction model allows for OSCs to potentially mimic the tremendous success of CG structural prediction models (*e.g.* Kremer-Grest)[122] in conventional polymer physics. The ability to bypass backmapping and QC calculations traditionally required to integrate multiscale structural predictions with electronic characterizations has the potential to be transformative for OSC science.

Recently, we have advanced this new paradigm of purely CG structural and electronic prediction models in OSCs to understand the behavior of OSCs in ionic media relevant to biosensing, bioelectronics, and neuromorphic devices.[113,114] In this work, thiophene-based conjugated polyelectrolytes were represented using the aforementioned anisotropic CG model, and single-chain conformations were modeled as a function of solvent, represented by both the solvent dielectric and solvent quality. A phenomenological Hamiltonian, in which the elec-

tronic couplings were parameterized by the inter-ring dihedrals and non-bonded monomer orientations, was then coupled to the sampled configurations to qualitatively address the correlation between chain conformations and charge transport properties.[97–100] Importantly, electrostatic potentials resulting from the ionic environment were directly incorporated into modulations in the electronic structure of the Hamiltonian to assess the impact of the electrostatic environment on the electronic wavefunction, which has never been attempted at these length scales for OSCs. Importantly, it was deduced that electrostatic potentials over the 10's of nm scale primarily influence variations in electronic structure through modulation of the thermodynamically stable conformations, and not through direct influence on the site energies of the electronic states. Current efforts in our research group are working to connect these anisotropic prediction models directly with ECG models to compute the exact valence Hamiltonian within a dimer approximation (see Figure 10a). These model Hamiltonians will not only possess the *ab initio* accuracy associated with the underlying QC, but will also stochastically include broadening of the local electronic structure by intermolecular vibrations averaged out during the CG model development. The ability to access these approximate QM Hamiltonians directly from the CG resolution has the potential to revolutionize the study of charge and exciton transport in OSCs at unprecedented length scales relevant to device function.

## QM Dynamics at the CG Resolution

As computational models for scalable morphological and electronic property predictions become available, obtaining QC-accurate descriptions of QM dynamics at CG resolutions presents itself as a high impact application of CG electronic prediction models. Conventionally, when the charge transport mechanism of soft materials is in the so-called hopping regime, a kinetic Monte Carlo method or iterative master equation solution can be parameterized using the simulated morphology, from which charge carrier mobility can be derived.[123–127] The probability of a charge carrier hopping between sites is derived from either Fermi's

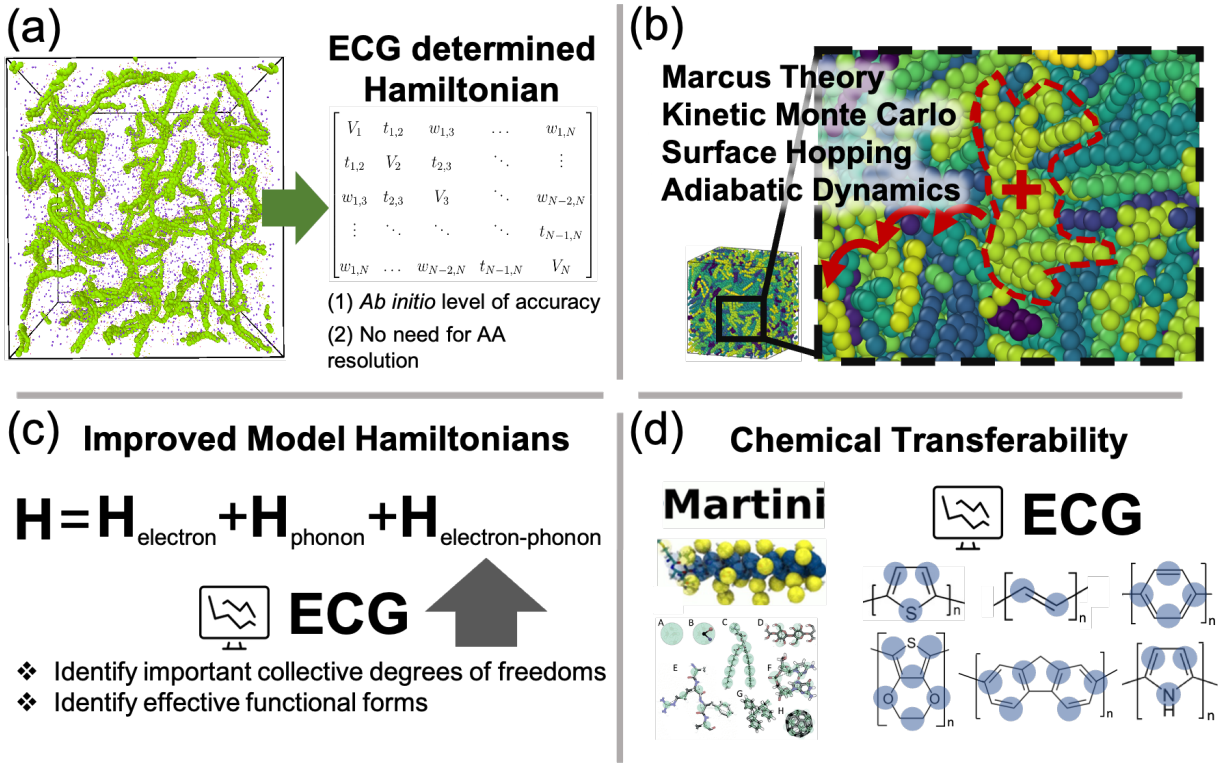



Figure 10: Future directions for ECG efforts: a) establishing multiscale structure-processing-function relationships for OSCs with *ab initio* accuracy, b) addressing QM dynamics at the CG resolution with configurationally dependent electronic structure, c) improving model Hamiltonians by identifying representative collective degree of freedoms, and d) developing a chemical transferable CG framework for conducting soft materials.

golden rule or Marcus theory in which the electronic coupling is arbitrarily specified and is assumed as a constant parameter.[98,125,126,128] However, in many conjugated systems the electronic coupling highly depends on the molecular alignment and packing structure, and its value could vary across several orders of magnitude.[129,130] By employing ECG frameworks, a deterministic kinetic Monte Carlo simulation can be established, where the configuration-dependent electronic coupling is evaluated. Moreover, in the case of the DKL-ECG, the electronic coupling "distribution" of a given CG configuration can be further determined so that the effect of the lost atomic degree of freedoms can be re-introduced in an exact stochastic manner at the CG resolution.

On the other hand, as illustrated in Figure 10b, for soft materials in which a charge carrier (or an exciton) is neither localized at a single molecule nor fully delocalized, polaronic hopping models and band theories cannot appropriately characterize the charge transport mechanism (or the exciton dynamics).[104,131–133] In this intermediate regime, the interactions between charge carriers and nuclear motions are highly correlated. The charge transport mechanisms in this regime are often characterized by employing mixed quantum-classical approaches.[104,132,134] Such approaches treat nuclear degrees of freedom classically by performing AA-MD simulation, with the charge carrier propagated in a fully QM fashion. To solve the time-dependent Schrodinger equation for the combined nuclear-electronic system, the two most common propagation techniques used in this context are Ehrenfest based mean field (MF) dynamics[135,136] and Tully's surface hopping (SH) method[137,138] along with electronic model Hamiltonians. The values of the on-site energies and electronic couplings of model Hamiltonians can be obtained using QC calculations, with the parameterization of spatial and temporal correlations from AA MD trajectories being a particularly efficient strategy.[139–141] The mixed quantum-classical approaches have provided significant insights,[104,132,133,142–147] but they are still constrained to limited spatiotemporal scales. An interesting path forward concerns the development of rigorous quantum dynamical approaches that operate solely at the CG model resolution. As ECG can straightforwardly obtain ap-

proximations to the full electronic Hamiltonian at every configurational snapshot, the charge carrier wave function can conceivably be evolved by implementing MF or SH propagation techniques. In many ways, this is not significantly different from a substantial literature on formal quantum dynamics approaches that utilize approximate Hamiltonians with normal mode expansions of collective variable representations, but the ability to connect directly to real chemistries beyond the standard linear approximations could be transformative. While much of this is speculative, the possibility to establish a connection between bottom-up CG models of QM phenomena and formal approaches to quantum dynamics remains a tantalizing prospect.

## Improved Model Hamiltonians from ECG

Connecting with the idea of performing CG quantum dynamical calculations parameterized directly from bottom-up AA representations is the idea of utilizing ECG-style models to construct more sophisticated model Hamiltonians which capture complex relationships between configurational and electronic degrees of freedom extending beyond the traditional linear approximations. To-date, existing model QM Hamiltonians employed in static and dynamical calculations typically employ one of two approaches to building configurational dependence into model Hamiltonians. First, normal mode expansions and their linear electron-phonon couplings are commonly used to induce configurational dependence in quantum dynamics calculations; this approximation has been omnipresent in nearly all studies of QM wavefunction propagation in photosynthetic complexes.[104,148] Second, in cases where specific collective variables are known, more complex dependencies can be specifically constructed.[149–151] Incorporating the inter-ring coupling in conjugated polymers as a function of the inter-ring dihedral degree of freedom represents a classic use of this approach. However, in most systems of interest, these non-linear dependencies of electronic structure on collective nuclear degrees of freedom are not known and require many years of careful study to elicit. The application of ECG models to discover electronically coupled collective variables (see Figure

9 and 10c) presents an avenue to (i) identify the important collective degrees of freedom that dictate variations in electronic structure and (ii) identify the functional forms of the Hamiltonian terms relating to these expressions. With the recent re-emergence of algorithms for symbolic regression, there is even the prospect of developing mathematical forms for these relationships that permit analytical approaches.

## Towards Chemical Transferability of ECG

To-date, all work on the development of CG electronic prediction models has been aligned with specific chemical systems from which rigorous "bottom-up" ECG-style models can be constructed. A particularly interesting path forward concerns to what degree chemical transferability of electronic predictions can be built into ECG model development. While at face value such a concept seems incredibly challenging, it is important to note that the chemically generalizable Martini model remains arguably the most used CG model to-date, with myriad offshoots to specific chemical subsets emerging. Whereas the Martini model employs a "bottom-up" style parameterization for intramolecular bonded terms and a more "top-down" style parameterization of intermolecular terms,[152,153] such a discrimination in ECG-style models is likely not achievable due to the difficulty of measuring relevant quantities such as the electronic density of states and the intrinsically non-additive behavior of many QM phenomena. Consequently, as shown in Figure 10d, including chemical transferability into ECG models will likely need to occupy a specific subset of relevant chemistries, similar to how Martini began with lipids and then progresses to broader sets of chemistries. Of immediate interest is the development of ML architectures for ECG model fitting that can include chemical transferability in a computationally efficient manner across the enumerable space of OSC chemistries. These concerns occupy significant present work within our research group, with the goal of constructing a drop-in replacement for conventional atomistic QC calculations on organic media at the CG resolution.

# Conclusions

Electronic structure models that operate at the CG nuclear resolution, while maintaining a rigorous statistical mechanical connection to the underlying *ab initio* QC of real molecules, represent a useful approach for tackling classes of soft materials modeling that evade current molecular modeling paradigms. Here, we have motivated the development of ECG-style prediction methods from existing deficiencies in current molecular modeling techniques, and have outlined the progress of these models to date. While the majority of efforts have concerned ECG model development in the context of charge transporting polymers, ECG model development is extensible to any soft materials systems for which QM phenomenology plays a non-trivial role. We have outlined a sampling of these soft materials applications for which ECG may play a critical role in future modeling efforts (Figure 1), as well as future directions that will make ECG models more transferable and of high impact to the soft materials chemistry community (Figure 10). The ability to perform QC-accurate electronic structure predictions without the need to reference all atomic degrees of freedom provides a promising path forward for the development of scalable QM prediction models in soft materials.

# Acknowledgement

The development of the coarse-grained electronic models described in this article was supported by the National Science Foundation Chemical Theory, Models, and Computation division under award CHE-2154916. We acknowledge support from the Dreyfus Program for Machine Learning in the Chemical Sciences and Engineering during this project.

# References


(1) Fagnani, D. E.; Tami, J. L.; Copley, G.; Clemons, M. N.; Getzler, Y. D. Y. L.; McNeil, A. J. 100th Anniversary of Macromolecular Science Viewpoint: Redefining Sustainable Polymers. *ACS Macro Lett.* **2021**, *10*, 41–53.

(2) Epps, T. H.; Korley, L. T. J.; Yan, T.; Beers, K. L.; Burt, T. M. Sustainability of Synthetic Plastics: Considerations in Materials Life-Cycle Management. *JACS Au* **2022**, *2*, 3–11.

(3) Haque, F. M. et al. Defining the Macromolecules of Tomorrow through Synergistic Sustainable Polymer Research. *Chem. Rev.* **2022**, *122*, 6322–6373.

(4) Keijer, T.; Bakker, V.; Slootweg, J. C. Circular chemistry to enable a circular economy. *Nat. Chem.* **2019**, *11*, 190–195.

(5) Ellis, L. D.; Rorrer, N. A.; Sullivan, K. P.; Otto, M.; McGeehan, J. E.; Román-Leshkov, Y.; Wierckx, N.; Beckham, G. T. Chemical and biological catalysis for plastics recycling and upcycling. *Nat. Catal.* **2021**, *4*, 539–556.

(6) Schneiderman, D. K.; Hillmyer, M. A. 50th Anniversary Perspective: There Is a Great Future in Sustainable Polymers. *Macromolecules* **2017**, *50*, 3733–3749.

(7) Nagarajan, V.; Mohanty, A. K.; Misra, M. Perspective on Polylactic Acid (PLA) based Sustainable Materials for Durable Applications: Focus on Toughness and Heat Resistance. *ACS Sustain. Chem. Eng.* **2016**, *4*, 2899–2916.

(8) Andrady, A. L. The plastic in microplastics: A review. *Mar. Pollut. Bull.* **2017**, *119*, 12–22.

(9) Yousif, E.; Haddad, R. Photodegradation and photostabilization of polymers, especially polystyrene: review. *SpringerPlus* **2013**, *2*, 1–32.

(10) Cao, J.; Tian, J.; Xu, J.; Wang, Y. Organic Flow Batteries: Recent Progress and Perspectives. *Energy Fuels* **2020**, *34*, 13384–13411.

(11) Lai, Y. Y.; Li, X.; Zhu, Y. Polymeric Active Materials for Redox Flow Battery Application. *ACS Appl. Polym. Mater.* **2020**, *2*, 113–128.

(12) Burgess, M.; Moore, J. S.; Rodríguez-López, J. Redox Active Polymers as Soluble Nanomaterials for Energy Storage. *Acc. Chem. Res.* **2016**, *49*, 2649–2657.

(13) Janoschka, T.; Martin, N.; Martin, U.; Friebe, C.; Morgenstern, S.; Hiller, H.; Hager, M. D.; Schubert, U. S. An aqueous, polymer-based redox-flow battery using non-corrosive, safe, and low-cost materials. *Nature* **2015**, *527*, 78–81.

(14) Ding, Y.; Zhang, C.; Zhang, L.; Zhou, Y.; Yu, G. Molecular engineering of organic electroactive materials for redox flow batteries. *Chem. Soc. Rev.* **2018**, *47*, 69–103.

(15) Zhang, L.; Feng, R.; Wang, W.; Yu, G. Emerging chemistries and molecular designs for flow batteries. *Nat. Rev. Chem.* **2022**, *6*, 524–543.

(16) Dahms, H. Electronic conduction in aqueous solution. *J. Phys. Chem.* **1968**, *72*, 362–364.

(17) Kemper, T. W.; Larsen, R. E.; Gennett, T. Relationship between Molecular Structure and Electron Transfer in a Polymeric Nitroxyl-Radical Energy Storage Material. *J. Phys. Chem. C* **2014**, *118*, 17213–17220.

(18) Sato, K.; Ichinoi, R.; Mizukami, R.; Serikawa, T.; Sasaki, Y.; Lutkenhaus, J.; Nishide, H.; Oyaizu, K. Diffusion-Cooperative Model for Charge Transport by Redox-Active Nonconjugated Polymers. *J. Am. Chem. Soc.* **2018**, *140*, 1049–1056.

(19) Tan, Y.; Casetti, N. C.; Boudouris, B. W.; Savoie, B. M. Molecular Design Features for Charge Transport in Nonconjugated Radical Polymers. *J. Am. Chem. Soc.* **2021**, *143*, 11994–12002.

(20) Abella, L.; Crassous, J.; Favereau, L.; Autschbach, J. Why is the Energy of the Singly Occupied Orbital in Some Radicals below the Highest Occupied Orbital Energy? *Chem. Mater.* **2021**, *33*, 3678–3691.

(21) Bello, L.; Sing, C. E. Mechanisms of Diffusive Charge Transport in Redox-Active Polymer Solutions. *Macromolecules* **2020**, *53*, 7658–7671.

(22) Galembeck, F.; Burgo, T. A. L.; Balestrin, L. B. S.; Gouveia, R. F.; Silva, C. A.; Galembeck, A. Friction, tribochemistry and triboelectricity: recent progress and perspectives. *RSC Adv.* **2014**, *4*, 64280–64298.

(23) Lacks, D. J.; Sankaran, R. M. Contact electrification of insulating materials. *J. Phys. D: Appl. Phys.* **2011**, *44*, 453001.

(24) Kim, W.-G.; Kim, D.-W.; Tcho, I.-W.; Kim, J.-K.; Kim, M.-S.; Choi, Y.-K. Triboelectric Nanogenerator: Structure, Mechanism, and Applications. *ACS Nano* **2021**, *15*, 258–287.

(25) Wang, Z. L.; Chen, J.; Lin, L. Progress in triboelectric nanogenerators as a new energy technology and self-powered sensors. *Energy Environ. Sci.* **2015**, *8*, 2250–2282.

(26) Lin Wang, Z.; Chen, J.; Lin, L. Progress in triboelectric nanogenerators as a new energy technology and self-powered sensors. *Energy Environ. Sci.* **2015**, *8*, 2250–2282.

(27) Lacks, D. J.; Shinbrot, T. Long-standing and unresolved issues in triboelectric charging. *Nat. Rev. Chem.* **2019**, *3*, 465–476.

(28) Li, J.; Nagamani, C.; Moore, J. S. Polymer Mechanochemistry: From Destructive to Productive. *Acc. Chem. Res.* **2015**, *48*, 2181–2190.

(29) O'Neill, R. T.; Boulatov, R. The many flavours of mechanochemistry and its plausible conceptual underpinnings. *Nat. Rev. Chem.* **2021**, *5*, 148–167.

(30) Klok, H.-A.; Herrmann, A.; Göstl, R. Force ahead: Emerging Applications and Opportunities of Polymer Mechanochemistry. *ACS Polym. Au* **2022**, *2*, 208–212.

(31) Ribas-Arino, J.; Shiga, M.; Marx, D. Understanding Covalent Mechanochemistry. *Angew. Chem. Int. Ed.* **2009**, *48*, 4190–4193.

(32) Krusenbaum, A.; Grätz, S.; Tigineh, G. T.; Borchardt, L.; Kim, J. G. The mechanochemical synthesis of polymers. *Chem. Soc. Rev.* **2022**, *51*, 2873–2905.

(33) Versaw, B. A.; Zeng, T.; Hu, X.; Robb, M. J. Harnessing the Power of Force: Development of Mechanophores for Molecular Release. *J. Am. Chem. Soc.* **2021**, *143*, 21461–21473.

(34) Dahlke, J.; Zechel, S.; Hager, M. D.; Schubert, U. S. How to Design a Self-Healing Polymer: General Concepts of Dynamic Covalent Bonds and Their Application for Intrinsic Healable Materials. *Adv. Mater. Interfaces* **2018**, *5*, 1800051.

(35) Thakur, V. K.; Kessler, M. R. Self-healing polymer nanocomposite materials: A review. *Polymer* **2015**, *69*, 369–383.

(36) Liu, Y.; Holm, S.; Meisner, J.; Jia, Y.; Wu, Q.; Woods, T. J.; Martinez, T. J.; Moore, J. S. Flyby reaction trajectories: Chemical dynamics under extrinsic force. *Science* **2021**, *373*, 208–212.

(37) Roessler, A.; Zimmerman, P. M. Examining the Ways To Bend and Break Reaction Pathways Using Mechanochemistry. *J. Phys. Chem. C* **2018**, *122*, 6996–7004.

(38) Stauch, T.; Dreuw, A. Advances in Quantum Mechanochemistry: Electronic Structure Methods and Force Analysis. *Chem. Rev.* **2016**, *116*, 14137–14180.

(39) Rottler, J. Fracture in glassy polymers: a molecular modeling perspective. *J. Phys.: Condens. Matter* **2009**, *21*, 463101.

(40) Tauber, J.; van der Gucht, J.; Dussi, S. Stretchy and disordered: Toward understanding fracture in soft network materials via mesoscopic computer simulations. *J. Chem. Phys.* **2022**, *156*, 160901.

(41) Hong, G.; Gan, X.; Leonhardt, C.; Zhang, Z.; Seibert, J.; Busch, J. M.; Bräse, S. A Brief History of OLEDs—Emitter Development and Industry Milestones. *Adv. Mater.* **2021**, *33*, 2005630.

(42) Shi, Y.-Z.; Wu, H.; Wang, K.; Yu, J.; Ou, X.-M.; Zhang, X.-H. Recent progress in thermally activated delayed fluorescence emitters for nondoped organic light-emitting diodes. *Chem. Sci.* **2022**, *13*, 3625–3651.

(43) Luo, D.; Jang, W.; Babu, D. D.; Kim, M. S.; Wang, D. H.; Kyaw, A. K. K. Recent progress in organic solar cells based on non-fullerene acceptors: materials to devices. *J. Mater. Chem. A* **2022**, *10*, 3255–3295.

(44) Bernardo, G.; Lopes, T.; Lidzey, D. G.; Mendes, A. Progress in Upscaling Organic Photovoltaic Devices. *Adv. Energy Mater.* **2021**, *11*, 2100342.

(45) Liu, K.; Ouyang, B.; Guo, X.; Guo, Y.; Liu, Y. Advances in flexible organic field-effect transistors and their applications for flexible electronics. *npj Flex. Electron.* **2022**, *6*, 1–19.

(46) Tang, W.; Huang, Y.; Han, L.; Liu, R.; Su, Y.; Guo, X.; Yan, F. Recent progress in printable organic field effect transistors. *J. Mater. Chem. C* **2019**, *7*, 790–808.

(47) Yokota, T.; Fukuda, K.; Someya, T. Recent Progress of Flexible Image Sensors for Biomedical Applications. *Adv. Mater.* **2021**, *33*, 2004416.

(48) Rivnay, J.; Owens, R. M.; Malliaras, G. G. The Rise of Organic Bioelectronics. *Chem. Mater.* **2014**, *26*, 679–685.

(49) Garlapati, S. K.; Divya, M.; Breitung, B.; Kruk, R.; Hahn, H.; Dasgupta, S. Printed Electronics Based on Inorganic Semiconductors: From Processes and Materials to Devices. *Adv. Mater.* **2018**, *30*, 1707600.

(50) Coropceanu, V.; Cornil, J.; da Silva Filho, D. A.; Olivier, Y.; Silbey, R.; Brédas, J.-L. Charge Transport in Organic Semiconductors. *Chemical Reviews* **2007**, *107*, 926–952.

(51) Jackson, N. E. Coarse-Graining Organic Semiconductors: The Path to Multiscale Design. *J. Phys. Chem. B* **2021**, *125*, 485–496.

(52) Jin, J.; Pak, A. J.; Durumeric, A. E. P.; Loose, T. D.; Voth, G. A. Bottom-up Coarse-Graining: Principles and Perspectives. *J. Chem. Theory Comput.* **2022**, *18*, 5759–5791.

(53) Kidder, K. M.; Szukalo, R. J.; Noid, W. G. Energetic and entropic considerations for coarse-graining. *Eur. Phys. J. B* **2021**, *94*, 153–29.

(54) Dhamankar, S.; Webb, M. A. Chemically specific coarse-graining of polymers: Methods and prospects. *J. Polym. Sci.* **2021**, *59*, 2613–2643.

(55) Gumyusenge, A.; Melianas, A.; Keene, S. T.; Salleo, A. Materials Strategies for Organic Neuromorphic Devices. *Annu. Rev. Mater. Res.* **2021**, *51*, 47–71.

(56) Beratan, D. N.; Betts, J. N.; Onuchic, J. N. Protein Electron Transfer Rates Set by the Bridging Secondary and Tertiary Structure. *Science* **1991**, *252*, 1285–1288.

(57) Scholes, G. D. Quantum-Coherent Electronic Energy Transfer: Did Nature Think of It First? *J. Phys. Chem. Lett.* **2010**, *1*, 2–8.

(58) Morzan, U.; Alonso de Armiño, D. J.; Foglia, N.; Ramírez, F.; González Lebrero, M. C.; Scherlis, D. A.; Estrin, D. A. Spectroscopy in Complex Environments from QM–MM Simulations. *Chem. Rev.* **2018**, *118*, 4071–4113.

(59) Bondanza, M.; Nottoli, M.; Cupellini, L.; Lipparini, F.; Mennucci, B. Polarizable embedding QM/MM: the future gold standard for complex (bio)systems? *Phys. Chem. Chem. Phys.* **2020**, *22*, 14433–14448.

(60) Keller, D.; Bustamante, C. The Mechanochemistry of Molecular Motors. *Biophys. J.* **2000**, *78*, 541–556.

(61) Kumar, R.; Liu, Z.; Lokitz, B.; Chen, J.; Carrillo, J.-M.; Jakowski, J.; Collier, C. P.; Retterer, S.; Advincula, R. Harnessing autocatalytic reactions in polymerization and depolymerization. *MRS Commun.* **2021**, *11*, 377–390.

(62) Mironenko, A. V.; Voth, G. A. Density Functional Theory-Based Quantum Mechanics/Coarse-Grained Molecular Mechanics: Theory and Implementation. *J. Chem. Theory Comput.* **2020**, *16*, 6329–6342.

(63) Ratcliff, L. E.; Mohr, S.; Huhs, G.; Deutsch, T.; Masella, M.; Genovese, L. Challenges in large scale quantum mechanical calculations. *WIREs Comput. Mol. Sci.* **2017**, *7*, e1290.

(64) Fabian, M. D.; Shpiro, B.; Rabani, E.; Neuhauser, D.; Baer, R. Stochastic density functional theory. *WIREs Comput. Mol. Sci.* **2019**, *9*, e1412.

(65) Elstner, M.; Porezag, D.; Jungnickel, G.; Elsner, J.; Haugk, M.; Frauenheim, Th.; Suhai, S.; Seifert, G. Self-consistent-charge density-functional tight-binding method for simulations of complex materials properties. *Phys. Rev. B* **1998**, *58*, 7260–7268.

(66) Bannwarth, C.; Ehlert, S.; Grimme, S. GFN2-xTB—An Accurate and Broadly Parametrized Self-Consistent Tight-Binding Quantum Chemical Method with Multipole Electrostatics and Density-Dependent Dispersion Contributions. *J. Chem. Theory Comput.* **2019**, *15*, 1652–1671.

(67) Fedorov, D. G.; Kitaura, K. Extending the Power of Quantum Chemistry to Large Systems with the Fragment Molecular Orbital Method. *J. Phys. Chem. A* **2007**, *111*, 6904–6914.

(68) Kitaura, K.; Ikeo, E.; Asada, T.; Nakano, T.; Uebayasi, M. Fragment molecular orbital method: an approximate computational method for large molecules. *Chem. Phys. Lett.* **1999**, *313*, 701–706.

(69) Nishimoto, Y.; Nakata, H.; Fedorov, D. G.; Irle, S. Large-Scale Quantum-Mechanical Molecular Dynamics Simulations Using Density-Functional Tight-Binding Combined with the Fragment Molecular Orbital Method. *J. Phys. Chem. Lett.* **2015**, *6*, 5034–5039.

(70) Freed, K. F. Is there a bridge between ab initio and semiempirical theories of valence? *Acc. Chem. Res.* **1983**, *16*, 137–144.

(71) Zheng, H.; Changlani, H. J.; Williams, K. T.; Busemeyer, B.; Wagner, L. K. From Real Materials to Model Hamiltonians With Density Matrix Downfolding. *Front. Phys.* **2018**, *6*, 43.

(72) Sawaya, R. C.; White, S. R. Constructing Hubbard models for the hydrogen chain using sliced-basis density matrix renormalization group. *Phys. Rev. B* **2022**, *105*, 045145.

(73) Zhang, F. C.; Rice, T. M. Effective Hamiltonian for the superconducting Cu oxides. *Phys. Rev. B* **1988**, *37*, 3759–3761.

(74) Durand, J.-P., Philippe; Malrieu *Ab Initio Methods in Quantum Chemistry*, 1st ed.; John Wiley & Sons, Ltd, 1987; Vol. 67; pp 321–407.

(75) Björnson, K.; Wills, J. M.; Alouani, M.; Grånäs, O.; Thunström, P.; Ong, C. S.;

Eriksson, O. In Situ Pseudopotentials for Electronic Structure Theory. *J. Phys. Chem. C* **2021**, *125*, 15103–15111.

(76) Ingólfsson, H. I.; Lopez, C. A.; Uusitalo, J. J.; de Jong, D. H.; Gopal, S. M.; Periole, X.; Marrink, S. J. The power of coarse graining in biomolecular simulations. *WIREs Comput. Mol. Sci.* **2014**, *4*, 225–248.

(77) Han, Y.; Jin, J.; Wagner, J. W.; Voth, G. A. Quantum theory of multiscale coarse-graining. *J. Chem. Phys.* **2018**, *148*, 102335.

(78) Foley, T. T.; Kidder, K. M.; Shell, M. S.; Noid, W. G. Exploring the landscape of model representations. *Proc. Natl. Acad. Sci. U.S.A.* **2020**, *117*, 24061–24068.

(79) Khot, A.; Shiring, S. B.; Savoie, B. M. Evidence of information limitations in coarse-grained models. *J. Chem. Phys.* **2019**, *151*, 244105.

(80) Chakraborty, M.; Xu, J.; White, A. D. Is preservation of symmetry necessary for coarse-graining? *Phys. Chem. Chem. Phys.* **2020**, *22*, 14998–15005.

(81) Müller-Plathe, F. Coarse-Graining in Polymer Simulation: From the Atomistic to the Mesoscopic Scale and Back. *ChemPhysChem* **2002**, *3*, 754–769.

(82) Izvekov, S.; Voth, G. A. A Multiscale Coarse-Graining Method for Biomolecular Systems. *J. Phys. Chem. B* **2005**, *109*, 2469–2473.

(83) Chaimovich, A.; Shell, M. S. Coarse-graining errors and numerical optimization using a relative entropy framework. *J. Chem. Phys.* **2011**, *134*, 094112.

(84) Souza, P. C. e. a. Martini 3: a general purpose force field for coarse-grained molecular dynamics. *Nat. Methods* **2021**, *18*, 382–388.

(85) Rühle, V.; Junghans, C.; Lukyanov, A.; Kremer, K.; Andrienko, D. Versatile Object-Oriented Toolkit for Coarse-Graining Applications. *J. Chem. Theory Comput.* **2009**, *5*, 3211–3223.

(86) Abraham, M. J.; Murtola, T.; Schulz, R.; Páll, S.; Smith, J. C.; Hess, B.; Lindahl, E. GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **2015**, *1-2*, 19–25.

(87) Plimpton, S. Fast Parallel Algorithms for Short-Range Molecular Dynamics. *J. Comput. Phys.* **1995**, *117*, 1–19.

(88) Alessandri, R.; Uusitalo, J. J.; de Vries, A. H.; Havenith, R. W. A.; Marrink, S. J. Bulk Heterojunction Morphologies with Atomistic Resolution from Coarse-Grain Solvent Evaporation Simulations. *J. Am. Chem. Soc.* **2017**, *139*, 3697–3705.

(89) Liu, P.; Voth, G. A. Smart resolution replica exchange: An efficient algorithm for exploring complex energy landscapes. *J. Chem. Phys.* **2007**, *126*, 045106.

(90) Liu, P.; Shi, Q.; Lyman, E.; Voth, G. A. Reconstructing atomistic detail for coarse-grained models with resolution exchange. *J. Chem. Phys.* **2008**, *129*, 114103.

(91) Wassenaar, T. A.; Pluhackova, K.; Böckmann, R. A.; Marrink, S. J.; Tieleman, D. P. Going Backward: A Flexible Geometric Approach to Reverse Transformation from Coarse Grained to Atomistic Models. *J. Chem. Theory Comput.* **2014**, *10*, 676–690.

(92) Chennakesavalu, S.; Toomer, D. J.; Rotskoff, G. M. Ensuring thermodynamic consistency with invertible coarse-graining. *arXiv* **2022**, 2210.07882 [cond–mat.stat–mech].

(93) Louison, K. A.; Dryden, I. L.; Laughton, C. A. GLIMPS: A Machine Learning Approach to Resolution Transformation for Multiscale Modeling. *J. Chem. Theory Comput.* **2021**, *17*, 7930–7937.

(94) Peng, J.; Yuan, C.; Ma, R.; Zhang, Z. Backmapping from Multiresolution Coarse-Grained Models to Atomic Structures of Large Biomolecules by Restrained Molecular Dynamics Simulations Using Bayesian Inference. *J. Chem. Theory Comput.* **2019**, *15*, 3344–3353.

(95) Wang, W.; Gómez-Bombarelli, R. Coarse-graining auto-encoders for molecular dynamics. *npj Comput. Mater.* **2019**, *5*, 1–9.

(96) An, Y.; Deshmukh, S. A. Machine learning approach for accurate backmapping of coarse-grained models to all-atom models. *Chem. Commun.* **2020**, *56*, 9312–9315.

(97) Yaliraki, S. N.; Silbey, R. J. Conformational disorder of conjugated polymers: Implications for optical properties. *J. Chem. Phys.* **1996**, *104*, 1245–1253.

(98) Grozema, F. C.; van Duijnen, P. Th.; Berlin, Y. A.; Ratner, M. A.; Siebbeles, L. D. A. Intramolecular Charge Transport along Isolated Chains of Conjugated Polymers: Effect of Torsional Disorder and Polymerization Defects. *J. Phys. Chem. B* **2002**, *106*, 7791–7795.

(99) Tipirneni, P.; Jindal, V.; Janik, M. J.; Milner, S. T. Tight binding models accurately predict band structures for copolymer semiconductors. *Phys. Chem. Chem. Phys.* **2020**, *22*, 19659–19671.

(100) Bombile, J. H.; Janik, M. J.; Milner, S. T. Tight binding model of conformational disorder effects on the optical absorption spectrum of polythiophenes. *Phys. Chem. Chem. Phys.* **2016**, *18*, 12521–12533.

(101) Jackson, N. E.; Bowen, A. S.; Antony, L. W.; Webb, M. A.; Vishwanath, V.; de Pablo, J. J. Electronic structure at coarse-grained resolutions from supervised machine learning. *Sci. Adv.* **2019**, *5*, eaav1190.

(102) Jackson, N. E.; Bowen, A. S.; de Pablo, J. J. Efficient Multiscale Optoelectronic Prediction for Conjugated Polymers. *Macromolecules* **2020**, *53*, 482–490.

(103) Maier, J. C.; Jackson, N. E. Bypassing backmapping: Coarse-grained electronic property distributions using heteroscedastic Gaussian processes. *J. Chem. Phys.* **2022**, *157*, 174102.

(104) Oberhofer, H.; Reuter, K.; Blumberger, J. Charge Transport in Molecular Materials: An Assessment of Computational Methods. *Chem. Rev.* **2017**, *117*, 10319–10357.

(105) Simine, L.; Allen, T. C.; Rossky, P. J. Predicting optical spectra for optoelectronic polymers using coarse-grained models and recurrent neural networks. *Proc. Natl. Acad. Sci.* **2020**, *117*, 13945–13948.

(106) Alessandri, R.; de Pablo, J. J. Prediction of Electronic Properties of Radical-Containing Polymers at Coarse-Grained Resolutions. *arXiv* **2022**, 2209.02072 [cond–mat.soft].

(107) Sivaraman, G.; Jackson, N. E. Coarse-Grained Density Functional Theory Predictions via Deep Kernel Learning. *J. Chem. Theory Comput.* **2022**, *18*, 1129–1141.

(108) Jin, J.; Voth, G. A. Ultra-Coarse-Grained Models Allow for an Accurate and Transferable Treatment of Interfacial Systems. *J. Chem. Theory Comput.* **2018**, *14*, 2180–2197.

(109) Pretti, E.; Shell, M. S. A microcanonical approach to temperature-transferable coarse-grained models using the relative entropy. *J. Chem. Phys.* **2021**, *155*, 094102.

(110) Ramakrishnan, R.; Dral, P. O.; Rupp, M.; von Lilienfeld, O. A. Big Data Meets Quantum Chemistry Approximations: The Δ-Machine Learning Approach. *J. Chem. Theory Comput.* **2015**, *11*, 2087–2096.

(111) Karniadakis, G. E.; Kevrekidis, I. G.; Lu, L.; Perdikaris, P.; Wang, S.; Yang, L. Physics-informed machine learning. *Nat. Rev. Phys.* **2021**, *3*, 422–440.

(112) Ferguson, A. L. Machine learning and data science in soft materials engineering. *J. Phys.: Condens. Matter* **2017**, *30*, 043002.

(113) Friday, D. M.; Jackson, N. E. Modeling the Interplay of Conformational and Electronic Structure in Conjugated Polyelectrolytes. *Macromolecules* **2022**, *55*, 1866–1877.

(114) Friday, D.; Jackson, N. Electron and Ion Transport in Semi-Dilute Conjugated. *ChemRxiv* **2023**, DOI: 0.26434/chemrxiv–2023–1pm15.

(115) Bowen, A. S.; Jackson, N. E.; Reid, D. R.; de Pablo, J. J. Structural Correlations and Percolation in Twisted Perylene Diimides Using a Simple Anisotropic Coarse-Grained Model. *J. Chem. Theory Comput.* **2018**, *14*, 6495–6504.

(116) Cohen, A. E.; Jackson, N. E.; de Pablo, J. J. Anisotropic Coarse-Grained Model for Conjugated Polymers: Investigations into Solution Morphologies. *Macromolecules* **2021**, *54*, 3780–3789.

(117) Gemünden, P.; Poelking, C.; Kremer, K.; Andrienko, D.; Daoulas, K. Ch. Nematic Ordering, Conjugation, and Density of States of Soluble Polymeric Semiconductors. *Macromolecules* **2013**, *46*, 5762–5774.

(118) Gemünden, P.; Poelking, C.; Kremer, K.; Daoulas, K.; Andrienko, D. Effect of Mesoscale Ordering on the Density of States of Polymeric Semiconductors. *Macromol. Rapid Commun.* **2015**, *36*, 1047–1053.

(119) Ricci, M.; Roscioni, O. M.; Querciagrossa, L.; Zannoni, C. MOLC. A reversible coarse grained approach using anisotropic beads for the modelling of organic functional materials. *Phys. Chem. Chem. Phys.* **2019**, *21*, 26195–26211.

(120) Khot, A.; Savoie, B. M. Top–Down Coarse-Grained Framework for Characterizing Mixed Conducting Polymers. *Macromolecules* **2021**, *54*, 4889–4901.

(121) Nguyen, H. T. L.; Huang, D. M. Systematic bottom-up molecular coarse-graining via force and torque matching using anisotropic particles. *J. Chem. Phys.* **2022**, *156*, 184118.

(122) Everaers, R.; Karimi-Varzaneh, H. A.; Fleck, F.; Hojdis, N.; Svaneborg, C. Kre-

mer–Grest Models for Commodity Polymer Melts: Linking Theory, Experiment, and Simulation at the Kuhn Scale. *Macromolecules* **2020**, *53*, 1901–1916.

(123) Gryn'ova, G.; Lin, K.-H.; Corminboeuf, C. Read between the Molecules: Computational Insights into Organic Semiconductors. *J. Am. Chem. Soc.* **2018**, *140*, 16370–16386.

(124) Yu, Z. G.; Smith, D. L.; Saxena, A.; Martin, R. L.; Bishop, A. R. Molecular geometry fluctuations and field-dependent mobility in conjugated polymers. *Phys. Rev. B* **2001**, *63*, 085202.

(125) Rühle, V.; Lukyanov, A.; May, F.; Schrader, M.; Vehoff, T.; Kirkpatrick, J.; Baumeier, B.; Andrienko, D. Microscopic Simulations of Charge Transport in Disordered Organic Semiconductors. *J. Chem. Theory Comput.* **2011**, *7*, 3335–3345.

(126) Nelson, J.; Kwiatkowski, J. J.; Kirkpatrick, J.; Frost, J. M. Modeling Charge Transport in Organic Photovoltaic Materials. *Acc. Chem. Res.* **2009**, *42*, 1768–1778.

(127) Balzer, D.; Smolders, T. J. A. M.; Blyth, D.; Hood, S. N.; Kassal, I. Delocalised kinetic Monte Carlo for simulating delocalisation-enhanced charge and exciton transport in disordered materials. *Chem. Sci.* **2021**, *12*, 2276–2285.

(128) Lu, N.; Li, L.; Geng, D.; Liu, M. A review for polaron dependent charge transport in organic semiconductor. *Org. Electron.* **2018**, *61*, 223–234.

(129) Wang, C.-I.; Braza, M. K. E.; Claudio, G. C.; Nellas, R. B.; Hsu, C.-P. Machine Learning for Predicting Electron Transfer Coupling. *J. Phys. Chem. A* **2019**, *123*, 7792–7802.

(130) Wang, C.-I.; Joanito, I.; Lan, C.-F.; Hsu, C.-P. Artificial neural networks for predicting charge transfer coupling. *J. Chem. Phys.* **2020**, *153*, 214113.

(131) Ortmann, F.; Bechstedt, F.; Hannewald, K. Charge transport in organic crystals: Theory and modelling. *Phys. Status Solidi B* **2011**, *248*, 511–525.

(132) Troisi, A. Charge transport in high mobility molecular semiconductors: classical models and new theories. *Chem. Soc. Rev.* **2011**, *40*, 2347–2358.

(133) Wang, L.; Beljonne, D. Flexible Surface Hopping Approach to Model the Crossover from Hopping to Band-like Transport in Organic Crystals. *J. Phys. Chem. Lett.* **2013**, *4*, 1888–1894.

(134) Berne, B. J.; Ciccotti, G.; Coker, D. F. *Classical and Quantum Dynamics in Condensed Phase Simulations*; World Scientific Publishing Company: Singapore, 1998; pp 515–538.

(135) Li, X.; Tully, J. C.; Schlegel, H. B.; Frisch, M. J. Ab initio Ehrenfest dynamics. *J. Chem. Phys.* **2005**, *123*, 084106.

(136) Castro, A.; Gross, E. K. U. Optimal control theory for quantum-classical systems: Ehrenfest molecular dynamics based on time-dependent density-functional theory. *J. Phys. A: Math. Theor.* **2013**, *47*, 025204.

(137) Tully, J. C. Molecular dynamics with electronic transitions. *J. Chem. Phys.* **1990**, *93*, 1061–1071.

(138) Tully, J. C. Perspective: Nonadiabatic dynamics theory. *J. Chem. Phys.* **2012**, *137*, 22A301.

(139) Lee, C. K.; Shi, L.; Willard, A. P. Modeling the Influence of Correlated Molecular Disorder on the Dynamics of Excitons in Organic Molecular Semiconductors. *J. Phys. Chem. C* **2019**, *123*, 306–314.

(140) Lee, C. K.; Willard, A. P. Representing the Molecular Signatures of Disordered Molecular Semiconductors in Size-Extendable Models of Exciton Dynamics. *J. Phys. Chem. B* **2020**, *124*, 5238–5245.

(141) Kordt, P.; Andrienko, D. Modeling of Spatially Correlated Energetic Disorder in Organic Semiconductors. *J. Chem. Theory Comput.* **2016**, *12*, 36–40.

(142) Wang, L.; Akimov, A.; Prezhdo, O. V. Recent Progress in Surface Hopping: 2011–2015. *J. Phys. Chem. Lett.* **2016**, *7*, 2100–2112.

(143) Troisi, A.; Cheung, D. L.; Andrienko, D. Charge Transport in Semiconductors with Multiscale Conformational Dynamics. *Phys. Rev. Lett.* **2009**, *102*, 116602.

(144) Heck, A.; Kranz, J. J.; Kubař, T.; Elstner, M. Multi-Scale Approach to Non-Adiabatic Charge Transport in High-Mobility Organic Semiconductors. *J. Chem. Theory Comput.* **2015**, *11*, 5068–5082.

(145) Heck, A.; Kranz, J. J.; Elstner, M. Simulation of Temperature-Dependent Charge Transport in Organic Semiconductors with Various Degrees of Disorder. *J. Chem. Theory Comput.* **2016**, *12*, 3087–3096.

(146) Giannini, S.; Carof, A.; Ellis, M.; Yang, H.; Ziogos, O. G.; Ghosh, S.; Blumberger, J. Quantum localization and delocalization of charge carriers in organic semiconducting crystals. *Nat. Commun.* **2019**, *10*, 1–12.

(147) Giannini, S.; Blumberger, J. Charge Transport in Organic Semiconductors: The Perspective from Nonadiabatic Molecular Dynamics. *Acc. Chem. Res.* **2022**, *55*, 819–830.

(148) Chenu, A.; Scholes, G. D. Coherence in Energy Transfer and Photosynthesis. *Annu. Rev. Phys. Chem.* **2015**, *66*, 69–96.

(149) Jackson, N. E.; Savoie, B. M.; Chen, L. X.; Ratner, M. A. A Simple Index for Characterizing Charge Transport in Molecular Materials. *J. Phys. Chem. Lett.* **2015**, *6*, 1018–1021.

(150) Jackson, N. E.; Kohlstedt, K. L.; Chen, L. X.; Ratner, M. A. A n-vector model for charge transport in molecular semiconductors. *J. Chem. Phys.* **2016**, *145*, 204102.

(151) Wang, C. I.; Hsu, C. H.; Hua, C. C. The correspondence between the conformational and chromophoric properties of amorphous conjugated polymers in mesoscale condensed systems. *Phys. Chem. Chem. Phys.* **2017**, *19*, 20818–20828.

(152) Souza, P. C. T. et al. Martini 3: a general purpose force field for coarse-grained molecular dynamics. *Nat. Methods* **2021**, *18*, 382–388.

(153) Marrink, S. J.; Monticelli, L.; Melo, M. N.; Alessandri, R.; Tieleman, D. P.; Souza, P. C. T. Two decades of Martini: Better beads, broader scope. *WIREs Comput. Mol. Sci.* **2022**, e1620.