Macromolecules

pubs.acs.org/Macromolecules Article

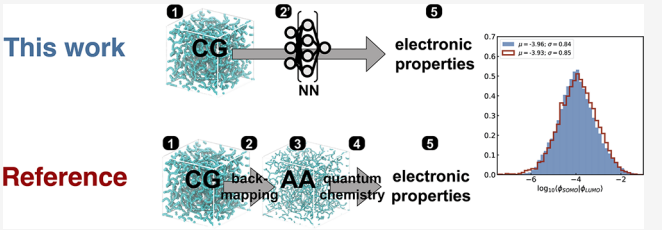
Prediction of Electronic Properties of Radical-Containing Polymers at Coarse-Grained Resolutions

Riccardo Alessandri and Juan J. de Pablo*

Cite This: *Macromolecules* 2023, 56, 3574−3584 Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: The properties of soft electronic materials depend on the coupling of electronic and conformational degrees of freedom over a wide range of spatiotemporal scales. The description of such properties requires multiscale approaches capable of, at the same time, accessing electronic properties and sampling the conformational space of soft materials. This could in principle be realized by connecting the coarse-grained (CG) methodologies required for adequate conformational sampling to conformationally averaged electronic property distributions via backmapping to atomistic-resolution level models and repeated quantum-chemical calculations. Computational demands of such approaches, however, have hindered their application in high-throughput computer-aided soft materials discovery. Here, we present a method that, combining machine learning and CG techniques, can replace traditional backmapping-based approaches without sacrificing accuracy. We illustrate the method for an emerging class of soft electronic materials, namely, nonconjugated, radical-containing polymers, promising materials for all-organic energy storage. Supervised machine learning models are trained to learn the dependence of electronic properties on polymer conformation at CG resolutions. We then parametrize CG models that retain electronic structure information, simulate CG condensed phases, and predict the electronic properties of such phases solely from the CG degrees of freedom. We validate our method by comparing it against a full backmapping-based approach and find good agreement between both methods. This work demonstrates the potential of the proposed method to accelerate multiscale workflows and provides a framework for the development of CG models that retain electronic structure information.

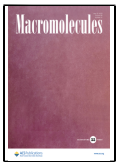


INTRODUCTION

Radical-containing polymers, also known as open-shell macromolecules, macromolecular radicals, or simply radical polymers, possess intriguing redox, optoelectronic, and magnetic characteristics that make them appealing for applications ranging from energy storage and optoelectronics to spintronics and memory storage.[1−6] Nonconjugated, radical-containing polymers are organic polymers that have a nonconjugated backbone bearing pendant stable radical sites. They constitute a class of charge-carrying polymers that do not rely on π-conjugation to transport charges successfully. The rational design of radical polymers with enhanced characteristics could be greatly advanced by deriving relationships that connect their molecular structure, morphology, and electronic properties. These relationships are inherently multiscale, involving the coupling of electronic and conformational degrees of freedom over a wide range of spatiotemporal scales. New modeling approaches capable of describing such coupling are needed.

Recent work has shown that coarse-grained (CG) models can be used to probe polymeric material length and time scales reaching the mesoscale.[7−10] In contrast, explicit quantum-chemical calculations, usually using density functional theory (DFT), which are necessary to access electronic properties,[11−13] are extraordinarily demanding and can only capture picosecond and Angstrom-level processes. To bridge these two scales, that is, to take into account large-scale morphological features generated via self-assembly processes (possibly as a function of processing conditions) when computing electronic properties, recent efforts have sought to introduce new multiscale modeling approaches.[12,14] In such approaches, the soft material morphology generated via CG simulations is backmapped to the atomistic resolution required for the quantum-chemical calculations, and such calculations are then performed on conformations drawn from the backmapped morphologies. These multiscale approaches have been primarily developed in the context of organic semiconductors[12,14−17] and provide a means to access the conformational dependence of electronic properties. However, considerable computational demands and workflow complexity of such have limited their applicability for high-throughput computer-aided materials discovery.

Received: January 26, 2023
Revised: April 12, 2023
Published: May 11, 2023



https://doi.org/10.1021/acs.macromol.3c00141
Macromolecules 2023, 56, 3574−3584

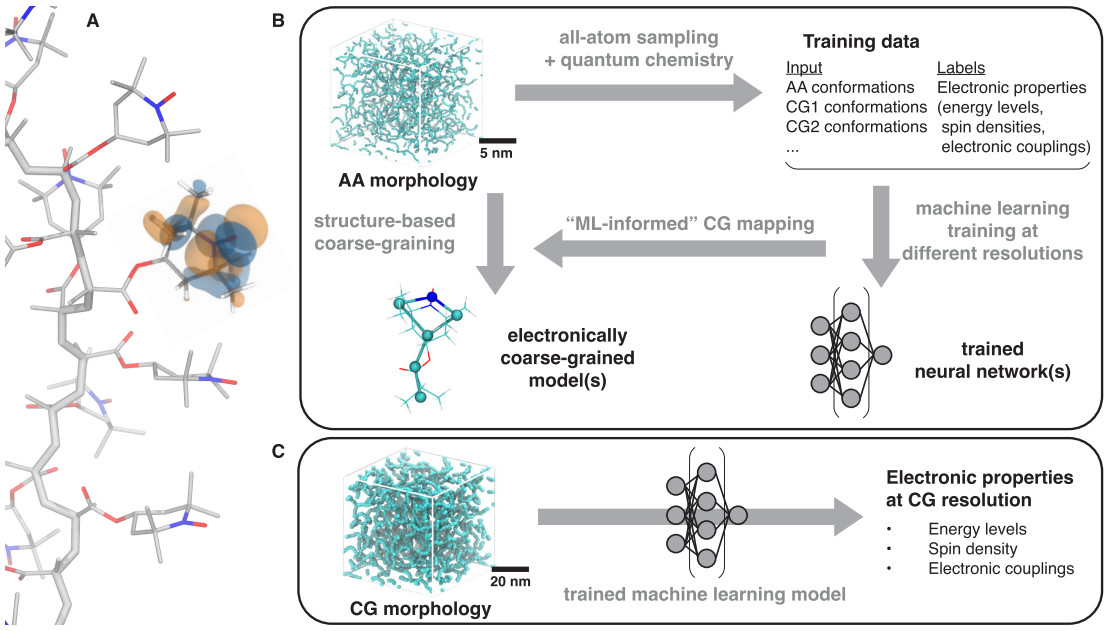


Figure 1. Schematic representation of the proposed method for nonconjugated radical polymers. (A) PTMA radical polymer: its nonconjugated backbone localizes electronic structure features, such as the SOMO orbital represented, to the pendant radical groups. Hydrogen atoms are only shown on the radical group on which the SOMO orbital is localized. (B) Development: starting from an all-atom (AA) model, conformations are sampled via MD simulations. Quantum-chemical calculations are performed on those conformations to generate the training data set; the data set serves to train machine learning models, in this case NNs, at different levels of CG resolution. The information gained during training is used to inform the development of CG models that are built on reference AA data via structure-based coarse-graining techniques. (C) Application: the CG models permit study of structures over larger spatiotemporal scales; the trained machine learning models are used to predict electronic properties based on the conformations taken from the CG morphology. The renderings of the morphologies show only polymer backbones for clarity.

Computational studies of radical-containing polymers have been limited.[1,2,13,18−24] Kemper et al. performed what appear to be the first[18] and one of the few molecular dynamics (MD) investigations of radical polymers. They simulated the widely used p-type radical polymer poly(2,2,6,6-tetramethylpiperidinyloxy-4-yl methacrylate) (PTMA), a polymethacrylate with pendant nitroxide radical TEMPO (2,2,6,6-tetramethylpiperidin-1-oxyl) groups, and provided early insights into some of PTMA's molecular-scale dynamics and its coupling to electronic properties.[18−20] More recent efforts that relied on coarser models have sought to predict the conductivity of radical polymers.[22,23] Such models require as input molecular-level information about the specific redox-active unit in order to be predictive and to differentiate between the performance of different polymers. A different type of effort has focused on molecule-specific characteristics by exploring sets of different redox-active units with DFT-based methods.[13,24] A limitation of these studies has been the fact that they do not account for condensed-phase conditions, which may affect the resulting electronic properties.

In this work, we rely on machine learning (ML) to connect conformational properties to electronic structure properties in condensed phases. The use of ML techniques to study polymers has been limited by a variety of inherent challenges.[25−28] These include the complexity of data representation in systems that are disordered, the dependence of properties on an ensemble of configurations rather than on a single one, and the dependence of processes on interactions that span multiple length and time scales. To the best of our knowledge, previous work aimed at predicting the *electronic properties of polymers* using ML has focused exclusively on polythiophene-based conjugated polymers.[29−33] Jackson et al., in particular, introduced the concept of electronic coarse-graining (ECG),[29] a ML-based methodology aimed at the prediction of electronic properties based on a molecule's CG representation, and applied it mostly to prototypical conjugated polymers and single molecules in vacuum.[29,31,34,35] However, no actual CG models, that require not only a CG mapping but also bonded and nonbonded interaction parameters, have been developed so far in the ECG framework; therefore, no validation against traditional backmapping-based approaches has been made. A comparison against backmapping-based approaches has been made by Simine et al., where a similar approach to ECG is used to infer the absorption spectra of a polythiophene system from CG degrees of freedom. However, significant discrepancies between the two approaches were reported and ascribed to the backmapping protocol.[32] Building on the ECG work, in what follows, we present an efficient ML-enabled method that can replace state-of-the-art but demanding backmapping-based multiscale approaches. We illustrate the method in the context of nonconjugated, radical-containing polymers, which represent an emerging class of conducting materials. We begin by (1) training supervised ML models to learn the dependence of several electronic properties on molecular conformations at CG resolution and identifying CG mappings that retain electronic structure information. For this first part, we take inspiration from the ECG methodology.[29] Next, we (2) parametrize CG models that retain electronic structure information using established[10,36] coarse-graining techniques. Finally, we (3) use such CG models to simulate condensed phases and the trained ML models to predict electronic properties solely from the CG degrees of freedom. We validate our method by comparing its results to those obtained via a standard backmapping-based approach and show that the proposed method offers a 10^6 speedup with respect to

https://doi.org/10.1021/acs.macromol.3c00141
Macromolecules 2023, 56, 3574−3584

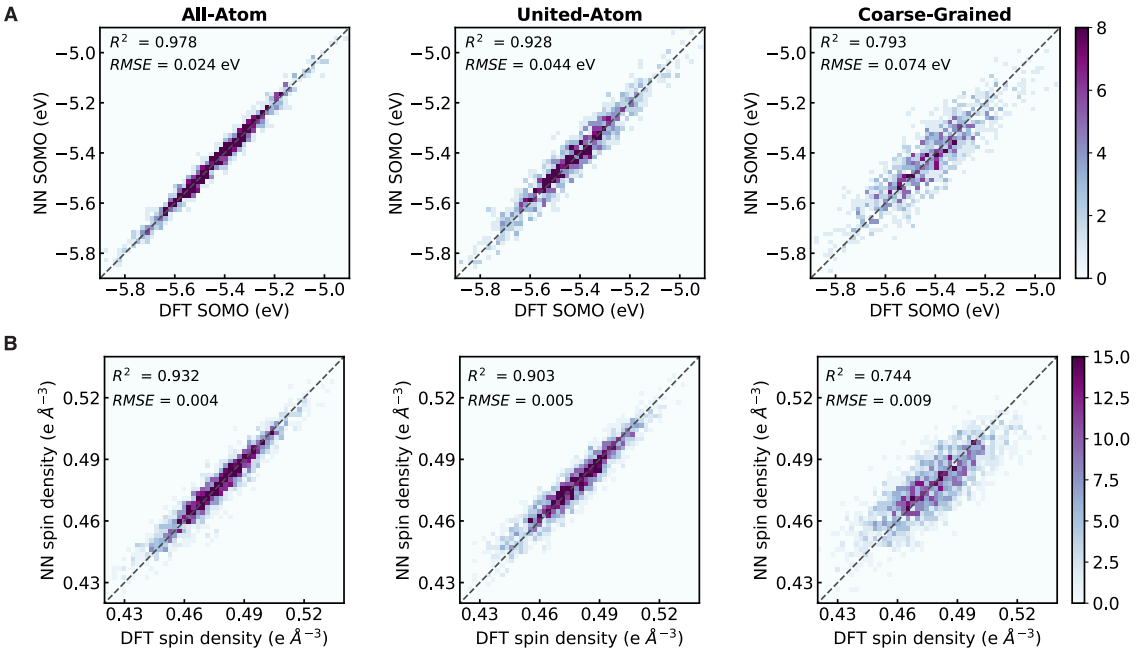


Figure 2. Predictive performance of NN models for monomer conformation-dependent electronic properties at all-atom and coarse-grained resolutions. Two-dimensional histograms showing the best accuracy achieved by the NNs for different resolutions, the reference all-atom resolution, united-atom, and CG, when learning (A) SOMO energy levels and (B) spin densities. In the CG case, the mapping used is GBNO1 (see Figure 3). The R^2 scores and RMSE are indicated on each plot and represent the results on the test set. All correlations fit to a linear model with slope = 1. The NN hyperparameters are optimized for each resolution.

backmapping-based approaches without loss of accuracy. By exploring different coarse-graining strategies, we are able to delineate general guidelines for development of CG models that retain electronic structure information and outline a viable pathway to accelerate multiscale workflows aimed at predicting electronic properties while encompassing CG-level spatiotemporal scales.

RESULTS

Method Overview. Figure 1 shows a schematic of the proposed method applied to nonconjugated, radical polymers. Only the main aspects of the method are outlined here, and readers are referred to the Materials and Methods for additional details. To illustrate the method, we use the well-studied, p-type radical polymer PTMA, which bears pendant TEMPO units (Figure 1A). An all-atom (AA) model is used to generate condensed-phase conformations of PTMA using MD simulations. The electronic properties of the resulting configurations are computed via DFT calculations (Figure 1B). Given the nonconjugated backbone of this class of radical-containing polymers, the monomers can be treated as electronically independent (see, e.g., the singly occupied molecular orbital (SOMO) distribution in Figure 1A). The data generated in this way, all-atom molecular conformations with associated electronic properties, are used as training data to develop supervised ML models. More specifically, we train feed-forward artificial neural networks (NNs) using the conformations (in the form of a reciprocal distance matrix) as input, and the electronic properties as labels. The all-atom molecular conformations can be mapped onto the corresponding CG configurations and NNs can therefore be trained *at different CG levels of resolution* (Figure 1B). We build two different data sets: one containing single monomer conformations and associated monomer conformation-dependent properties (SOMO energy level, spin density) and a second containing dimer conformations and associated dimer conformation-dependent properties (electronic couplings). After training, we generate two main outputs: (1) trained NNs for the different electronic properties (one for each different resolution) and (2) information on the CG mapping required for the CG model to retain sufficient information about specific electronic properties. The latter information, together with established coarse-graining techniques, are used to develop CG models for simulation of polymer condensed phases over larger spatial and temporal scales; the trained NNs allow one to retrieve electronic properties at the CG level with only negligible computational demands (Figure 1C), thereby permitting the investigation of electronic properties over large ensembles of CG morphologies.

Monomer Conformation-Dependent Electronic Properties. We first examine the performance of the NN models for prediction of electronic properties that depend on the conformation of a single PTMA *monomer*, using the *SOMO energy level* and the *spin density* as examples of the target properties to be learned. The SOMO energy level determines the position of the transport energy level and hence the charge transport type (p-type vs n-type), while knowledge of the spin density distribution gives access to the degree of (de)-localization of the unpaired electron(s). The latter property contributes to the stability of the radical center. Both energy levels and spin densities are conformation dependent; the extent to which they depend on conformation contributes to the degree of energetic disorder in the system.

Figure 2 shows that SOMO energies and spin densities can be predicted at CG resolution. Figure 2 shows correlation plots between the DFT reference SOMO energies and spin densities, computed at the B3LYP/6-311G(d,p) level of DFT (Materials and Methods), and the same quantities as predicted

https://doi.org/10.1021/acs.macromol.3c00141
Macromolecules 2023, 56, 3574−3584

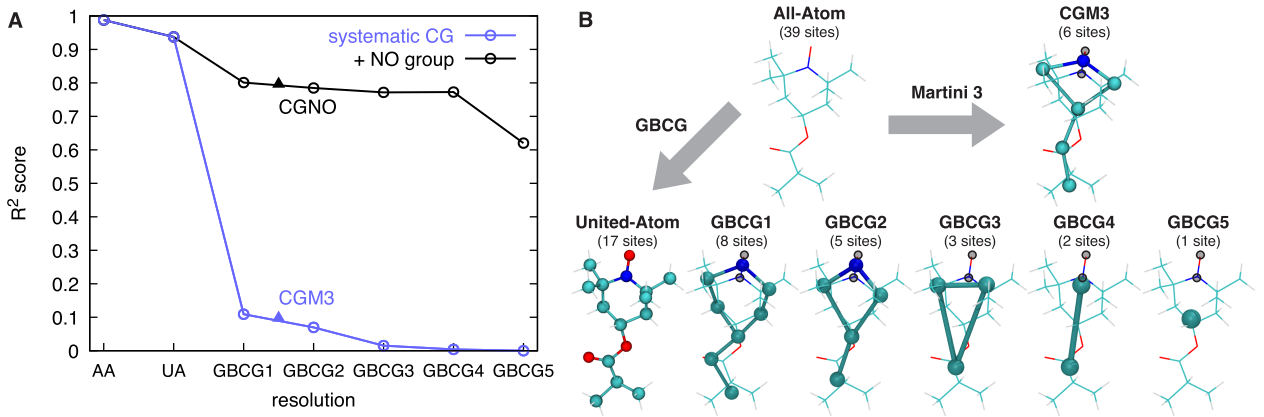


Figure 3. Dependence of the predictive performance of NN models on CG resolution. (A) R^2 score on the test set as a function of CG resolution for prediction of the SOMO energy level: all-atom (AA), united-atom (UA), and CG mappings obtained with the systematic graph-based coarse-graining approach (GBCGX, with X = [1.5]; blue line, hollow circles) and CG Martini 3 mapping (CGM3, blue triangle). The black data points (black line, hollow black circles for the GBCG-based mappings, black triangle for the CGM3-based mapping, CGNO) represent mappings where the nitrogen and oxygen atoms of the nitroxide group are described by an explicit CG site each (represented by semitransparent black circles in the renderings of (B)). (B) Renderings of the AA, UA, and CG mappings and the underlying atomistic structure. The number of sites that each resolution entails is also reported. For the detailed atom-to-bead correspondence, see Figures S10 and S12. The NN's hyperparameters are optimized for each resolution.

by the NN model at AA, united-atom (UA), and CG resolutions. The performance of the NN on the AA configurations (R^2 = 0.978, RMSE = 0.024 eV for SOMO energies; R^2 = 0.932, RMSE = 0.004 for spin densities) represents the maximum achievable performance for a given molecular representation, given that all other resolutions involve a loss of information for the NN model to perform the regression task on. At UA resolution, the NN also achieves very good predictive performance (R^2 = 0.928, RMSE = 0.044 eV for SOMO energies; R^2 = 0.903, RMSE = 0.005 for spin densities), just below that obtained at AA resolution. At CG resolution, there is a noticeable drop of performance with respect to the UA and AA cases, but the NN still provides R^2 = 0.793, RMSE = 0.074 eV for SOMO energies and R^2 = 0.744, RMSE = 0.009 for spin densities. To which degree these increased errors would eventually impact the calculation of charge transport rates[22] is difficult to estimate and will require a separate study. The CG mapping used here is "GBNO1" (see Figure 3B); a detailed investigation of the NN model accuracy for different CG mapping choices and resolutions is presented in the next section.

It is useful to investigate the minimum data requirements needed by the NN to learn monomer conformation-dependent electronic properties. Figure S1 shows that, after ≈4,000 data points, the performance of the NN starts to plateau at all resolutions. This represents a relatively small number of data points, and calculations at higher levels of theory could therefore be easily used to train such NNs. We also note that the required NNs are rather small (e.g., 4 hidden layers with 3 neurons each already maximize the performance at CG resolutions), and the same performance is achieved for a relatively wide range of hyperparameters (see Figure S8).

Impact of CG Resolution and Specific CG Mapping. The CG level of resolution and the specific CG mapping have considerable influence on the ability of the NN to predict electronic properties. In this section, we provide an analysis of resolution that is useful for identifying design principles for CG models that represent the best compromise between computational efficiency and accuracy.

To design CG mappings, we resort to two different strategies. First, we apply the Graph-Based Coarse-Graining (GBCG) algorithm of Webb et al. to arrive at a series of increasingly coarser mappings in a systematic and automated manner with minimal human intervention.[37] Second, we devise a mapping according to the Martini 3 building-block coarse-graining strategy.[10] All the mappings are shown in Figure 3B; the atom-to-bead correspondence is given in detail in Figures S10 and S12. Additional details are provided in the Materials and Methods and the Supporting Information.

Figure 3A shows how the predictive performance of the NN model, as quantified by the R^2 coefficient, changes as a function of the resolution of the molecular representation for the prediction of the SOMO energy level. We see that GBCG mappings lead to poor predictive performance ($R^2 < 0.2$), even at the finest CG resolution level, GBCG1. The same is true for the Martini 3 CG mapping (CGM3, the resolution of which is between GBCG1 and GBCG2). Given the high predictive performance of the UA resolution, these results indicate that these "standard" CG mappings are missing key degrees of freedom. Notably, they do not include the nitroxide group explicitly. As can be seen from Figure 3A (black line), the explicit inclusion of the nitroxide group (i.e., the positions of the nitrogen and oxygen atoms are included in the distance matrix used as the NN input vector) considerably improves the performance across the different CG mappings, with R^2 remaining around 0.8 for GBCG2−GBCG4. Even for resolution GBCG5, which represents the PTMA monomer with only 1 site, when expanded to include the explicit description of the nitroxide group, one can achieve a R^2 above 0.6. This result is expected and very recently observed also in ref 35: the SOMO wave function is localized around the nitroxide group (see Figure 1A) and, consequently, the conformation of this group relative to the other atoms in the TEMPO unit governs the SOMO wave function coefficients and associated energy.

Further confirmation of the importance of the nitroxide group for prediction of SOMO energies comes from a complementary test, where we exclude either the nitrogen or the oxygen atom of this group from the UA and AA resolutions

https://doi.org/10.1021/acs.macromol.3c00141
Macromolecules 2023, 56, 3574−3584

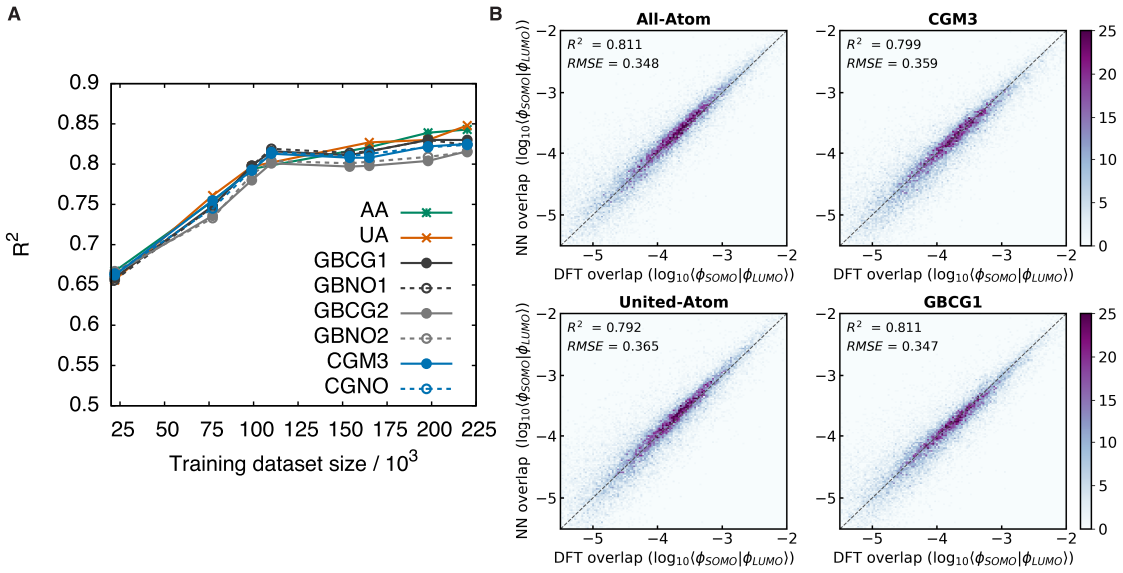


Figure 4. Predictive performance of NN models for dimer conformation-dependent electronic properties at all-atom and coarse-grained resolutions. (A) NN performance as a function of training data set size for different resolutions: all-atom (AA), united-atom (UA), GBCG1−2, GBNO1−2, and CG Martini (CGM3 and CGNO). See Figure 3B for a schematic representation of the mappings. (B) Two-dimensional histograms showing the best accuracy achieved by the NNs for the different resolutions. The R^2 scores and RMSE are indicated on each plot and represent the results on the test set. The NN's hyperparameters are optimized for each resolution and data set size.

(Figure S2). Such exclusions lead to a dramatic decrease of predictive performance at the AA and UA levels, with R^2 values <0.5 (Figure S2). Figure S2 shows that excluding the nitrogen from the UA and AA representations leads to R^2 values of ≈0.42 while excluding the oxygen from the same representations degrades R^2 down to ≈0.1, indicating that the oxygen position is more informative than that of the nitrogen for SOMO energy predictions. In the case of strongly localized radical groups such as the nitroxide in PTMA, the inclusion of the nitroxide atoms is essential for the NN model to be able to accurately infer the SOMO energy.

Dimer Conformation-Dependent Electronic Properties. We now examine the accuracy of NNs for the prediction of electronic properties that depend on the conformation of a *dimer*, using the *electronic couplings* as the target property to be learned. Electronic couplings are a key molecular-scale property that determines charge transport, the latter being in turn critical for many applications of radical-containing polymers.[1,2] Charge transport in these materials is generally believed[1,2] to occur via a hopping mechanism, where a charge jumps between a radical site and an adjacent ionized site, in the case of PTMA, and of p-type transport in general, a cation site. We approximate the electronic coupling as the orbital overlap between the SOMO of the neutral radical and the lowest unoccupied molecular orbital (LUMO) of the cation, $\langle\phi_{\mathrm{SOMO}}|\phi_{\mathrm{LUMO}}\rangle$ (Materials and Methods), and we train NNs to learn its base 10 logarithm, $\log_{10}(\langle\phi_{\mathrm{SOMO}}|\phi_{\mathrm{LUMO}}\rangle)$.

As shown in Figure 4, the logarithm of the electronic coupling can be learned with good and comparable accuracy at AA, UA, GBCG1-GBCG2, and CGM3 resolutions. The predictive performances (on the test set), as quantified by the R^2 coefficient, are around 0.8−0.81 at AA, UA, GBCG1, and CGM3 resolutions. The performance gap between AA and UA resolutions and the best CG resolutions are small, particularly when compared to the monomer conformation-dependent properties depicted in Figure 2. Electronic couplings, however, are significantly more complex, i.e., data-intensive and requiring larger NNs, and difficult to learn than energy levels or spin densities (compare Figures 4A and S1), across all resolutions. In particular, if a training data set of ≈4,000 data points was sufficient to train good-performing NNs for the prediction of SOMO energies (Figure S1), at least ≈100,000 data points are necessary for the analogous task on electronic couplings (Figure 4A). After ≈100,000 data points, the predictive performance starts to plateau across the different CG resolutions. Moreover, NNs with 400 neurons per layer are required (see Table S1). This complexity, both in the data set size required and NN architecture, is consistent with previous reports.[29,38] Overall, while more demanding, our results confirm that the task of predicting electronic couplings at CG resolutions can be accomplished. This is particularly important given their role in determining charge transport.

Another aspect that differentiates electronic couplings from the monomer-dependent electronic properties is the degree of chemical detail required by the CG mapping resolution. In particular, resolving the nitroxide group is not critical in the case of electronic couplings: adding an explicit description of the nitroxide group does not improve the results within the statistical uncertainty of our predictions (compare GBCG1−2 (solid) to GBNO1−2 (dashed) and CGM3 (solid) to CGNO (dashed) in Figure 4A). This observation can be understood in terms of the electronic coupling being dominated by the relative position between the two monomers. The nitroxide groups do not contribute to the definition of this relative position, and hence, their explicit representation does not improve predictions of electronic couplings significantly.

Electronic Properties of Condensed-Phase CG Simulations. We now develop CG models that retain electronic structure information and apply the trained NNs to predict electronic properties from condensed-phase CG simulations. We validate the NN predictions by comparing them against the electronic properties obtained via the reference, state-of-the-art approach that requires backmapping and explicit quantum-chemical calculations.

https://doi.org/10.1021/acs.macromol.3c00141
Macromolecules 2023, 56, 3574−3584

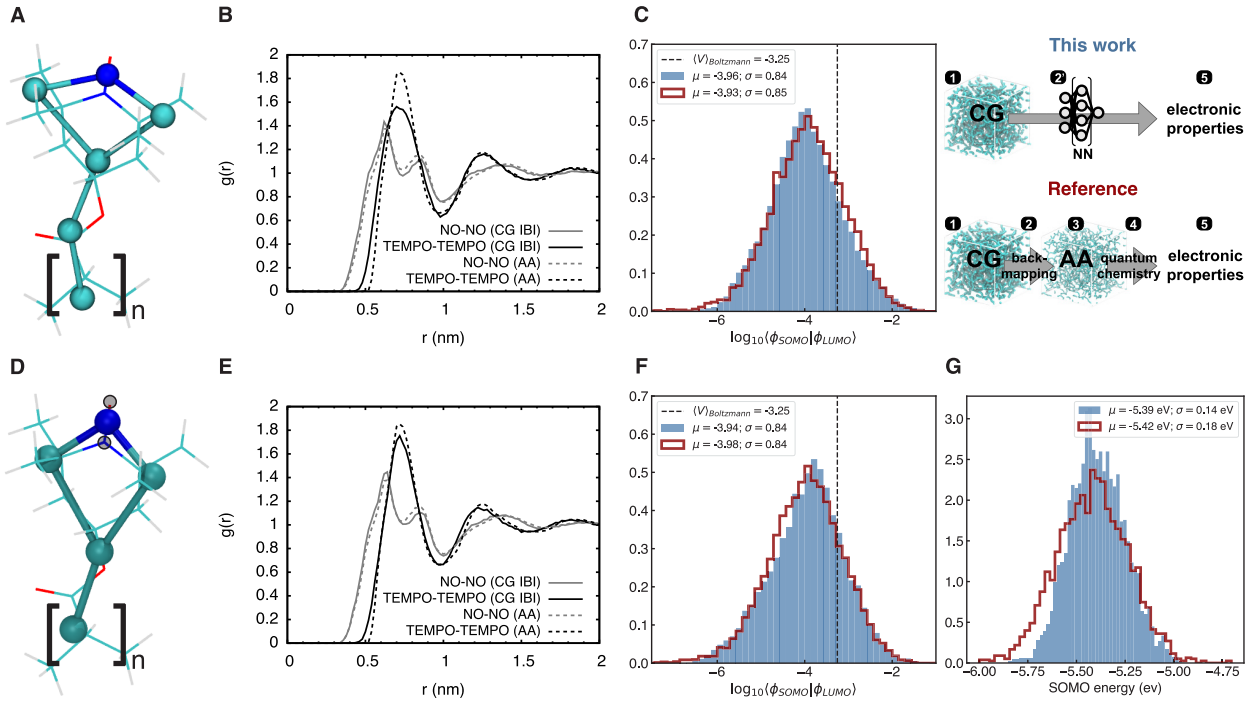


Figure 5. Application and validation of the proposed ML-based method to PTMA, a prototypical radical polymer. (A) Schematic representation of the CG model (CGM3 mapping) and the underlying AA structure. (B) Representative AA vs CG RDFs for the CGM3 model with nonbonded interactions obtained via IBI. (C) Electronic couplings computed with the reference backmapping-based approach (red) and the proposed ML-based method (blue) starting from a CG morphology generated via CG MD using the CGM3 model. (D) Schematic representation of the CG model (GBNO2 mapping) and the underlying AA structure; the nitroxide atoms are described by two virtual dummy sites (transparent gray circles; see text for details). (E) Representative AA vs CG RDFs for the GBNO2 model. (F) Electronic couplings and (G) SOMO energies computed with the reference backmapping-based approach (red) and the proposed ML-based method (blue) starting from a CG morphology generated via CG MD using the GBNO2 model. Mean (μ) and standard deviation (σ) of the electronic property distributions are reported in the legends.

To investigate what the most suitable coarse-graining strategies are when developing CG models that retain electronic structure information, we first build two CG models with the same mapping and bonded interactions but different nonbonded interactions. To realize this, we choose the CGM3 mapping that, as we have seen, allows for good accuracy in predicting electronic couplings. At the same time, the CGM3 mapping allows us to test two sets of nonbonded interactions derived via two different CG philosophies, namely, the building-block CG Martini 3 approach[10] and the structure-based CG IBI method.[36] For details on the CG models, see Materials and Methods and the Supporting Information.

We first consider the results generated with the CGM3 IBI model (Figure 5A–C). The model is used to produce a morphology via CG MD, and we then use two methods to retrieve the electronic couplings of such a CG morphology: the reference approach that involves backmapping to atomistic resolution and explicit DFT calculations (taken to be the ground truth; in red in Figure 5C) and the proposed ML-based method (in blue in Figure 5C). The agreement between the two electronic coupling distributions is excellent, with both the NN-predicted mean (μ) and standard deviation (σ) being around 1% of the corresponding reference values. We stress that the trained NN is able to predict electronic couplings of a CG simulation with an accuracy comparable to that of the reference approach but at a fraction of the computational cost (see the next section for a discussion on the computational efficiency). Moreover, the NN was trained on CG conformations mapped from an AA MD simulation and had not seen conformations drawn from a CG simulation before. However, the structural accuracy of the CG model (Figure 5B) is such that the CG model spans a conformational space that is consistent with the underlying AA structure, and therefore, the NN is able to make predictions on conformations drawn from previously unseen CG morphology realizations with excellent accuracy.

In contrast to the CGM3 IBI model, the CGM3 Martini 3 model leads to unsatisfactory electronic property predictions (Figure S4). In fact, not only does the electronic coupling distribution obtained with the Martini model exhibit a considerably larger standard deviation (+15%, Figure S4C) than the reference distribution, but also that distribution shows qualitative discrepancies with respect to the reference distribution (i.e., it is not a unimodal distribution). These discrepancies can be rationalized by looking at the structure of the Martini CG morphology, for example, by analyzing TEMPO–TEMPO and nitroxide–nitroxide radial distribution functions (RDFs) (Figure S4B). We see that the RDFs are qualitatively different, showing, for example, an extra peak at ≈0.4 nm in the case of the nitroxide–nitroxide RDF. These discrepancies are indicative of molecular conformations occurring in the Martini CG morphology that have no AA counterpart and are hence nonphysical. In contrast, structural accuracy is the parametrization target of structure-based coarse-graining techniques such as IBI and, therefore, as shown in the previous paragraph, the resulting CG models represent the underlying AA structure with high fidelity (Figure 5B). This structural accuracy appears to be an essential requirement when developing CG models that retain electronic structure information.

https://doi.org/10.1021/acs.macromol.3c00141
Macromolecules 2023, 56, 3574–3584

Having settled on IBI as the strategy of choice for developing CG models that retain electronic structure information, we now turn to the development of a CG model that is able to retain information on not only electronic couplings but also energy levels. Informed by the results of the previous sections, we know that an explicit description of the nitroxide group is required for energy level prediction and we therefore develop a corresponding CG model. We choose the GBNO2 mapping (Figure 5D), as it represents a good compromise between NN predictive accuracy, showing accurate electronic coupling (Figure 3) and energy level (Figure 4) predictions, and coarse-graining degree. We derive nonbonded interactions via IBI for all the GBCG2 beads (Supporting Information). The nitrogen and oxygen atoms of the nitroxide group are instead described by two virtual dummy sites that do not interact via nonbonded interactions with any of the other CG sites. As such, the GBNO2 model is expected to allow for predictions of both electronic couplings and energy levels.

As done for CGM3-based CG models, we generate a morphology via CG MD using the GBNO2 model just described and predict this time both energy levels and electronic couplings using the trained NNs (Figure 5D−F). Again, we compare the NN predictions to the ground truth taken to be the energy levels and electronic couplings obtained with the reference, backmapping-based approach (Figure 5F,G). The agreement between the two electronic coupling distributions is also excellent in this case, with the NN-predicted μ being within 1% of the reference one and σ matching the reference value. In contrast, the agreement between the two SOMO energy distributions is only satisfactory, with a very good agreement between the NN-predicted and reference μ (<1%) but a NN-predicted σ that is 22% lower than the reference. We note that, in this case, we are looking at an electronic property that depends strongly on the intramonomer degrees of freedom. The degeneracy of the CG representation, i.e., the fact that multiple AA conformations correspond to the same CG conformation, is likely responsible[34] for the narrowing of the SOMO distribution predicted at CG resolution. As discussed later, a recently proposed[34] Deep Kernel Learning approach may remedy this shortcoming and improve the prediction. Albeit with this limitation to keep in mind for the SOMO energy predictions, the GBNO2 model retains information on both energy levels and couplings and can be used to explore large ensembles of morphologies at CG resolution and retrieve electronic structure information with negligible computational costs.

Comparison to Quantum Chemistry Approaches. To highlight the importance of considering condensed-phase effects in soft materials when predicting electronic properties, including electronic couplings, we compare our results to some of the Boltzmann-averaged electronic couplings that have been used in recent, state-of-the-art quantum-chemical modeling studies[13,24] aimed at understanding transport in radical polymers. In such studies, the electronic coupling for a given molecular species is evaluated by first generating gas-phase dimer structures either randomly[24] or with more elaborate dimer surface sampling algorithms.[13] For each of the generated dimer structures, the electronic coupling is computed with the level of theory of choice, along with either a total or binding energy for that structure. The energies, which are indicative of the relative stability of each structure *in the gas phase*, are then used to obtain a Boltzmann-averaged electronic coupling according to

$$\langle V \rangle = \frac{\sum_{i=1}^{N} V_i \exp\left(\frac{-E_i}{k_B T}\right)}{\sum_{i=1}^{N} \exp\left(\frac{-E_i}{k_B T}\right)} \quad (1)$$

where V_i is the electronic coupling of the i-th dimer, E_i is its binding energy (calculated as the single point energy difference between the dimer (E_i^{dimer}) and the isolated species (E_i^{radical}, E_i^{cation}), according to $E_i = E_i^{\text{dimer}} - E_i^{\text{radical}} - E_i^{\text{cation}}$), k_B is the Boltzmann constant, T is the temperature (300 K), and N is the total number of dimers.

For a comparison to the method introduced here, we must compute the Boltzmann-averaged electronic coupling at the same level of theory used in the present work. Hence, we randomly generate 10,000 dimer structures following the procedure by Li and Tabor[24] (see the Supporting Information for details) and compute the $\langle \phi_{\text{SOMO}} | \phi_{\text{LUMO}} \rangle$ orbital overlap for each structure. For the same dimer structure, we also compute the binding energy (at the ωB97X-D/6-311G(d,p) level of DFT) and calculate the Boltzmann-averaged electronic coupling according to eq 1. The resulting coupling is shown with a vertical dashed line in Figure 5C,F. We can see that the Boltzmann-averaged coupling overestimates the magnitude of the mean electronic coupling that arises in condensed phases. The discrepancy between the two is not surprising, given that the gas-phase sampling of the dimer structures neglects any solid-state packing preference dictated by (1) steric constraints (including the fact that the monomers are attached to a polymer backbone) and (2) (un)favorable intra- and interchain interactions. A second intrinsic limitation of computing Boltzmann-averaged electronic couplings is the lack of an estimate of the so-called structural[39] disorder present in the system, i.e., the width of the electronic coupling distribution. This disorder may affect the charge transport landscape and may therefore impact the conductivity of the material. In contrast, the proposed method, whose accuracy depends on the structural accuracy of the CG model, gives access to not only (1) electronic couplings that account for the packing in the condensed phase but also (2) the full distributions of electronic couplings, thereby allowing for quantification of the structural disorder in the system.

DISCUSSION

Having demonstrated the feasibility of designing and developing CG models that retain electronic structure information and using such models as a more efficient substitute for current multiscale approaches that involve backmapping procedures, we conclude by briefly outlining a few opportunities for future extensions of the presented method.

The computational efficiency is a key advantage of the proposed method with respect to backmapping-based approaches and is expected to enable innovative high-throughput investigations of electronic properties over ensembles of soft material morphologies (see below). The upfront cost of the data set generation and ML model training is (considering both electronic couplings and SOMO energies): $\approx 375 \times 10^3$ CPU hours for training data generation and $\approx 4.2 \times 10^3$ CPU hours for training. Note that this estimation takes into account both the exploration of suitable

https://doi.org/10.1021/acs.macromol.3c00141
Macromolecules 2023, 56, 3574−3584

mappings and the hyperparameter grid search performed for all the NNs. The training data generation dominates the upfront cost. Computing electronic properties with the backmapping-based approach requires a total of $\approx 27 \times 10^3$ CPU hours for a single snapshot ($\approx 2.5 \times 10^3$ hours for backmapping and relaxation and $\approx 24.5 \times 10^3$ hours for the quantum-chemical calculations). Instead, the NN predictions take $\approx 1.5 \times 10^{-1}$ CPU hours per snapshot. See Table S4 for a detailed breakdown. Therefore, if electronic properties are predicted for 13 CG frames, we break even with the upfront training cost (see Figure S25). For any further frame, the proposed method will provide a speedup of 10^6 times. Further efficiency gains may be realized if more generic ML models, i.e., models trained on a larger fraction of the (CG) chemical space, will be developed; see also the discussion below.

The fact that neither the GBCG nor the Martini CG mappings describe the nitroxide group explicitly makes these mappings unable to predict the SOMO energy level. Although it is unsurprising that such fine detail is being averaged out by the CG mappings, this result highlights the limitation of current methods to devise CG mappings that retain specific electronic structure information (e.g., SOMO energy). While less strongly localized radical units might be less affected by this (i.e., we speculate that, if a radical is more delocalized, a coarser representation may still be able to capture the conformational dependence of the SOMO energy), a systematic way to identify mappings that preserve specific electronic properties would be greatly beneficial for automating the CG mapping step of the proposed method.

We have shown that a structurally accurate CG model is essential for the proposed method. Hence, in general, structure-based coarse-graining techniques, such as IBI,[36] should be used when deriving interactions for CG models that retain electronic structure information. However, other kinds of CG models such as chemically specific, building-block CG force fields (e.g., Martini[10]) that do not use the structure as a target for parametrization may still be valuable. First, while structural inaccuracies cannot be excluded for Martini-like CG approaches, they are also not necessarily present in every CG model developed within such frameworks. In fact, the latest Martini force field update (version 3)[10] has taken structural accuracy more into account via bonded parameter optimization that targets the AA molecular volume and shape and a wider range of systematically precalibrated nonbonded interactions; for example, structural accuracy is expected to improve for conjugated structures.[10,40] Hence, more generally, regardless of how the CG model potentials are obtained, as long as structural accuracy is preserved, which can be inspected, e.g., by comparing AA and CG RDFs, we expect a CG model to be suitable for the proposed method. Second, given the computational efficiency of Lennard-Jones-based models such as Martini as compared to models that use (IBI-derived) custom nonbonded potentials (Table S5) and the transferability of CG models developed within building-block CG frameworks, it is still valuable to investigate how such CG strategies can contribute to the development of CG models that retain electronic structure information.

With respect to the ML model and featurization schemes used here, there is ample room for improvement. Here, we used NNs due to their simplicity of implementation and application flexibility. However, recently Gaussian process regression has been combined with NNs to realize a Deep Kernel Learning (DKL) approach to ECG.[34,35] The approach allows one to incorporate the distributional nature of the electronic property predictions resulting from the CG mapping degeneracy, which, as we have seen in the present work, may be needed for electronic properties that strongly depends on detailed intramolecular features (Figure 5G). DKL moreover provides prediction uncertainties, and it is hence suitable for active learning strategies. Regarding the input featurization, the (reciprocal) distance matrices used here constitute one of the simplest molecular representations that guarantees rotational and translation invariance. However, a host of more elaborated featurization schemes, such as representations based on symmetry functions or graph convolutions, are available.[41] In particular, input representations containing chemical identity and possibly other physical descriptors are expected to play a key role in order to explore the question of whether ML models able to predict electronic properties at CG resolution and that are *transferable* across CG chemical space can be trained. Finally, applications to conjugated polymers, in contrast to the nonconjugated polymers studied here, need to consider that electronic states can delocalize along the polymer backbone. To tackle this problem, long short-term memory networks combined with a Δ-ML approach have been shown[31] to be a promising strategy to achieve the molecular-weight transferability necessary to consider conjugation along polymer backbones.

The proposed method opens up interesting avenues for applications that require efficient modeling of radical-containing polymers and, more generally, in soft electronic materials. For example, electronic properties could be studied as a function of morphology processing conditions in a high-throughput fashion. In particular, charge transport networks and fluctuations of electronic couplings over CG spatiotemporal scales can be studied by the proposed approach. Electronic couplings thus gathered could be fed into, for example, kinetic Monte Carlo simulations aimed at computing conductivities. As discussed before, such high-throughput studies are made possible by the 10^6 increase in efficiency of the proposed method with respect to current backmapping-based approaches.

In conclusion, an efficient ML-enabled method has been presented that connects electronic properties of soft materials with the CG spatiotemporal scales required to sample their phase space. The method is able to replace current backmapping-based multiscale approaches and directly connect electronic structure information to CG degrees of freedom. The method has been demonstrated for nonconjugated, radical-containing polymers, but it is suitable for any soft material with electronic properties. Overall, the findings reported underscore the potential of the proposed strategy to propel the investigation of soft materials' electronic properties that depend on conformational degrees of freedom over a wide range of spatiotemporal scales, a critical step toward the bottom-up design of soft electronic materials.

MATERIALS AND METHODS

AA Models. Initial parameters were obtained for a OPLS-AA/CM1A force field from the LigParGen server.[42,43] QUBEKit[44,45] was subsequently used to derive parameters for bond and angle potentials from DFT calculations (B3LYP/6-311++G(d,p)) that are tailored to PTMA. Charges were obtained with Gaussian via the CHELPG method.[46] The resulting force field led to a density for methyl methacrylate and TEMPO, molecular fragments that make up PTMA and for which experimental reference densities are available, in good agreement (within 2.5%) with experiments (Table S2).

https://doi.org/10.1021/acs.macromol.3c00141
Macromolecules 2023, 56, 3574−3584

Data Sets. Two data sets were built: one with single PTMA monomer conformations and associated monomer conformation-dependent electronic properties (namely, SOMO energy level and spin density) and a second one with dimer conformations and associated dimer conformation-dependent electronic properties (electronic couplings). We refer to these as "monomer dataset" and "dimer dataset", respectively. The following protocol was used to ensure good conformational sampling for the training data generation. We ran a condensed-phase AA MD simulation at 680 K (i.e., $\approx 1.3 \cdot T_g$; see Figure S9 for the estimation of the glass transition temperature, T_g, of the AA model) and, after 30 ns of equilibration, we gathered snapshots every 10 ns. Each snapshot was relaxed at 300 K for at least 20 ns. The relaxed snapshots were then used to extract the conformations for which quantum-chemical calculations were performed. Two different procedures were used to extract conformations for the two data sets, both leveraging the MDAnalysis python library[47,48] and Open Babel.[49] For the monomer data set, 4 snapshots containing 3,000 conformations each were used. For each conformation, coordinates were extracted up to the first carbon of the backbone ("C01" of Figure S10) and the carbon atom was capped with 3 hydrogen atoms. For each such conformation, reciprocal distance matrices at the different resolutions were computed and stored and an input file for a single point calculation (B3LYP/6-311++G(d,p)) was written. The (4·3,000) 12,000 conformations were filtered to increase the diversity of the data set by removing all conformations that had a RMSD of less than 0.8 Å with any other conformation. Gaussian[46] was used to obtain the SOMO energy and Mulliken spin density (B3LYP/6-311++G(d,p)). A final data set of 10,778 data points was obtained. For the dimer data set, 24 snapshots containing about 9,200 dimer conformations each were used. A dimer was selected if the distance between the respective nitrogen atoms was within a cutoff of 10 Å. For each selected dimer, the DFT-optimized geometry of the monomer was aligned (by minimizing the RMSD[50,51]) to each MD monomer structure. In this way, intramonomer vibrations were neglected. Intermonomer vibrations are much faster than the intermonomer motions and negligibly affect intermonomer couplings.[52] The electronic couplings were approximated as the orbital overlap[52] between the SOMO of the neutral radical and the LUMO of the cation, $\langle \phi_{\mathrm{SOMO}} | \phi_{\mathrm{LUMO}} \rangle$. For each dimer, Gaussian[46] was used to compute (1) the orbitals of monomer i and (2) the orbitals of monomer j of the of i–j dimer both by performing a single point calculation using as starting point the converged orbitals of the DFT-optimized geometry; (3) the overlap matrix by performing a calculation that produces only the overlap matrix in the MO basis set of the dimer; finally, Multiwfn[53] was used to evaluate $\langle \phi_{\mathrm{SOMO}} | \phi_{\mathrm{LUMO}} \rangle$ by providing the Gaussian checkpoint files from the three calculations described above. Note that the average overlap between the $\langle \phi_{\mathrm{SOMO},i} | \phi_{\mathrm{LUMO},j} \rangle$ and $\langle \phi_{\mathrm{LUMO},i} | \phi_{\mathrm{SOMO},j} \rangle$ values was taken, as the two values are in general different (although, at least for the present system, they are very similar). A final data set of 221,406 data points was obtained.

Coarse-Graining. GBCG mappings were generated using the spectral grouping variant of the GBCG algorithm,[37] as implemented at https://github.com/xmwebb/GBCG. The Martini CG mapping was devised following the Martini 3 guidelines.[10,40] Bonded parameters for the CG models were obtained by fitting standard bond, angle, and dihedral potentials to reference distributions obtained from mapped atomistic simulation (Supporting Information). Nonbonded parameters for the structure-based CG models were derived by iterative Boltzmann inversion (IBI)[36] as implemented in the VOTCA package.[54] 15 and 10 pair interactions were parametrized for the CGM3 and GBNO2 CG models, respectively. For more details, see the Supporting Information. Nonbonded parameters for the Martini model, that is, Martini bead types, were assigned based on Martini 3 guidelines[10,40] (and validated by computing octanol/water free energies of transfer; see Table S3) and available models for similar polymers[55] (Supporting Information).

Backmapping. Backmapping was performed with Backward,[56] which relaxes the AA structure with the atomistic force field after an initial geometrical reconstruction of the AA structure based on the CG particle positions. The geometrical reconstruction is specified via CG-to-AA mapping files.

Molecular Dynamics Simulations. Starting configurations for all polymer simulations were set up with Polyply,[55] and MD simulations were run with Gromacs versions 2021.x or more recent.[57] AA simulations used a time step of 1 fs and the Verlet scheme with a nonbonded cutoff of 1.1 nm, dispersion correction, and the particle mesh Ewald (PME) method for long-range electrostatic interactions. Bonds involving hydrogen atoms were constrained. Temperature and pressure were controlled by a Nosé–Hoover thermostat (coupling parameter, τ_T, of 1.0 ps) and a Parrinello–Rahman barostat (coupling parameter, τ_P, of 5.0 ps), respectively. The Berendsen barostat was used for equilibration purposes (τ_P = 0.5 ps). For the Martini CG simulations, the Verlet scheme had a straight nonbonded cutoff of 1.1 nm. Temperature and pressure were controlled by a velocity-rescaling thermostat (τ_T = 1.0 ps) and a Parrinello–Rahman barostat (τ_P = 12 ps), respectively. For the IBI-based CG simulations, the group cutoff scheme was used in order to use tabulated potentials with a nonbonded cutoff of 1.5 nm. Note that CG simulations with IBI-derived nonbonded potentials were run with Gromacs version 2019.5 because tabulated potentials are not available in more recent Gromacs versions.

ML Model Details. For each monomer or dimer conformation, a reciprocal distance matrix $\mathbf{D}$ between all particles of a particular resolution was computed. Its elements are $D_{kl}^{(i,j)} = |\mathbf{r}_k^{(i)} - \mathbf{r}_l^{(j)}|^{-1}$ where $\mathbf{r}$ is the position vector, i and j are the monomer indices, and k and l are the atom indices. For the monomer data set, $j = i$ and hence $D_{kl}^{(i,i)} = |\mathbf{r}_k^{(i)} - \mathbf{r}_l^{(i)}|^{-1}$. Each matrix was flattened, and the resulting one-dimensional vector (of dimension N^2) was used as the input feature for the NN. The electronic properties (SOMO energy level, spin density, electronic coupling) are the labels associated with the input vectors for the supervised ML regression task. The SOMO energy and spin density (only the values on the nitroxide nitrogen and oxygen atoms were considered, given that all other atoms have a spin density <0.07; hence, two values were passed to the NN) were used as obtained from the DFT calculations. The base 10 logarithm of $\langle \phi_{\mathrm{SOMO}} | \phi_{\mathrm{LUMO}} \rangle$ was used in the case of the couplings. A fully connected, feed-forward NN with a M-dimensional input layer followed by 4 batch-normalized hidden layers with the same number of neurons was used so as to have a NN flexible enough. M is the dimension of the (flattened) input vector, and it therefore depends on the molecular resolution ($M = N^2$). Hyperparameters that were optimized include: the number of neurons in the hidden layers, the batch size, and the number of training epochs. The default learning rate of the NAdam optimizer (0.001) was used for training. Visualizations of the hyperparameter grid search for SOMO energy prediction for some of the resolutions can be found in Figures S5–S8. Standard scaling was applied to the input and output features. 10% of the data sets (1,077 data points for the monomer data set; 22,140 for the dimer data set) was held out and used as the test set. The remaining data points were used to build training and validation sets by using 5-fold cross-validation. Hyperparameters were optimized by grid search based on the 5-fold cross-validated performance. For the best performing NN models for each property and molecular resolution, see Table S1. The final model performance was measured by applying the best model chosen based on the 5-fold cross-validation to the held-out test set. When predicting electronic couplings, as done when generating the dimer data set: (1) couplings were inferred for dimers for which the distance between the nitrogen virtual sites (or the beads representing the nitroxide group in the case of the CGM3 mapping) was within a 10 Å cutoff; (2) the CG-mapped DFT-optimized structures were aligned to each monomer of the CG MD simulation snapshot structure, and electronic couplings were inferred based on the reciprocal distance matrix computed for the dimer conformation after alignment. All the ML methods were implemented using the Keras[58] and scikit-learn[59] libraries.

Data Availability. Data and code to reproduce the findings of this work are available at https://github.com/ricalessandri/ECG-

https://doi.org/10.1021/acs.macromol.3c00141
Macromolecules 2023, 56, 3574–3584

RedoxPoly-Cathode. The developed AA and CG polymer models are also implemented in the Polyply[55] library (https://github.com/marrink-lab/polyply_1.0). The TEMPO and PTMA monomer Martini models are also available on the Martini 3 small molecule library[40] (https://github.com/ricalessandri/Martini3-small-molecules).

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.macromol.3c00141.

Additional results: NN performance vs data set size for SOMO energy prediction, extension of Figure 3, extension of Figure 4, and extension of Figure 5 (CGM3Martini model); ML model hyperparameters; density and T_g data of the AA model; details on the development of the CG models: mappings, bonded parameters, IBI nonbonded parameters, and Martini nonbonded parameters; random dimer structure generation; computational performance data (PDF)

■ AUTHOR INFORMATION

Corresponding Author

Juan J. de Pablo − *Pritzker School of Molecular Engineering, University of Chicago, Chicago, Illinois 60637, United States;* orcid.org/0000-0002-3526-516X; Email: depablo@uchicago.edu

Author

Riccardo Alessandri − *Pritzker School of Molecular Engineering, University of Chicago, Chicago, Illinois 60637, United States;* orcid.org/0000-0003-1948-5311

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.macromol.3c00141

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

R.A. is supported by the Dutch Research Council (NWO Rubicon 019.202EN.028). The work was partly supported by grant NSF-DMR-2119673 funded by the National Science Foundation. The authors acknowledge the Research Computing Center of the University of Chicago for computational resources.

■ REFERENCES

(1) Wang, S.; Easley, A. D.; Lutkenhaus, J. L. 100th Anniversary of Macromolecular Science Viewpoint: Fundamentals for the Future of Macromolecular Nitroxide Radicals. *ACS Macro Lett.* **2020**, *9*, 358−370.

(2) Tan, Y.; Hsu, S.-N.; Tahir, H.; Dou, L.; Savoie, B. M.; Boudouris, B. W. Electronic and Spintronic Open-Shell Macromolecules, Quo Vadis? *J. Am. Chem. Soc.* **2022**, *144*, 626−647.

(3) Kim, J.; Kim, J. H.; Ariga, K. Redox-Active Polymers for Energy Storage Nanoarchitectonics. *Joule* **2017**, *1*, 739−768.

(4) Janoschka, T.; Hager, M. D.; Schubert, U. S. Powering up the Future: Radical Polymers for Battery Applications. *Adv. Mater.* **2012**, *24*, 6397−6409.

(5) Oyaizu, K.; Nishide, H. Radical Polymers for Organic Electronic Devices: A Radical Departure from Conjugated Polymers? *Adv. Mater.* **2009**, *21*, 2339−2344.

(6) Wilcox, D. A.; Agarkar, V.; Mukherjee, S.; Boudouris, B. W. Stable Radical Materials for Energy Applications. *Annu. Rev. Chem. Biomol. Eng.* **2018**, *9*, 83−103.

(7) de Pablo, J. J. Coarse-Grained Simulations of Macromolecules: From DNA to Nanocomposites. *Annu. Rev. Phys. Chem.* **2011**, *62*, 555−574.

(8) Saunders, M. G.; Voth, G. A. Coarse-Graining Methods for Computational Biology. *Annu. Rev. Biophys.* **2013**, *42*, 73−93.

(9) 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.

(10) Souza, P. C. T.; Alessandri, R.; Barnoud, J.; Thallmair, S.; Faustino, I.; Grünewald, F.; Patmanidis, I.; Abdizadeh, H.; Bruininks, B. M. H.; Wassenaar, T. A.; Kroon, P. C.; Melcr, J.; Nieto, V.; Corradi, V.; Khan, H. M.; Domański, J.; Javanainen, M.; Martinez-Seara, H.; Reuter, N.; Best, R. B.; Vattulainen, I.; Monticelli, L.; Periole, X.; Tieleman, D. P.; de Vries, A. H.; Marrink, S. J. Martini 3: a General Purpose Force Field for Coarse-Grained Molecular Dynamics. *Nat. Methods* **2021**, *18*, 382−388.

(11) Olivier, Y.; Sancho-Garcia, J.-C.; Muccioli, L.; D'Avino, G.; Beljonne, D. Computational Design of Thermally Activated Delayed Fluorescence Materials: the Challenges Ahead. *J. Phys. Chem. Lett.* **2018**, *9*, 6149−6163.

(12) Friederich, P.; Fediai, A.; Kaiser, S.; Konrad, M.; Jung, N.; Wenzel, W. Toward Design of Novel Materials for Organic Electronics. *Adv. Mater.* **2019**, *31*, 1808256.

(13) 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.

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

(15) 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.

(16) Alessandri, R.; Sami, S.; Barnoud, J.; de Vries, A. H.; Marrink, S. J.; Havenith, R. W. A. Resolving Donor−Acceptor Interfaces and Charge Carrier Energy Levels of Organic Semiconductors with Polar Side Chains. *Adv. Funct. Mater.* **2020**, *30*, 2004799.

(17) Rolland, N.; Modarresi, M.; Franco-Gonzalez, J. F.; Zozoulenko, I. Large scale mobility calculations in PEDOT (Poly(3,4-ethylenedioxythiophene)): backmapping the coarse-grained Martini morphology. *Comput. Mater. Sci.* **2020**, *179*, 109678.

(18) 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.

(19) Kemper, T. W.; Larsen, R. E.; Gennett, T. Density of States and the Role of Energetic Disorder in Charge Transport in an Organic Radical Polymer in the Solid State. *J. Phys. Chem. C* **2015**, *119*, 21369−21375.

(20) Kemper, T. W.; Gennett, T.; Larsen, R. E. Molecular Dynamics Simulation Study of Solvent and State of Charge Effects on Solid-Phase Structure and Counterion Binding in a Nitroxide Radical Containing Polymer Energy Storage Material. *J. Phys. Chem. C* **2016**, *120*, 25639−25646.

(21) Joo, Y.; Agarkar, V.; Sung, S. H.; Savoie, B. M.; Boudouris, B. W. A nonconjugated radical polymer glass with high electrical conductivity. *Science* **2018**, *359*, 1391−1395.

(22) 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.

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

(24) Li, C.-H.; Tabor, D. P. Discovery of lead low-potential radical candidates for organic radical polymer batteries with machine-

learning-assisted virtual screening. *J. Mater. Chem. A* **2022**, *10*, 8273−8282.
(25) Bereau, T.; Andrienko, D.; Kremer, K. Research Update: Computational materials discovery in soft matter. *APL Mater.* **2016**, *4*, 053101.
(26) Ferguson, A. L. Machine learning and data science in soft materials engineering. *J. Phys.: Condens. Matter* **2018**, *30*, 043002.
(27) Jackson, N. E.; Webb, M. A.; de Pablo, J. J. Recent advances in machine learning towards multiscale soft materials design. *Curr. Opin. Chem. Eng.* **2019**, *23*, 106−114.
(28) Chen, L.; Pilania, G.; Batra, R.; Huan, T. D.; Kim, C.; Kuenneth, C.; Ramprasad, R. Polymer informatics: Current status and critical next steps. *Mater. Sci. Eng. R Rep.* **2021**, *144*, 100595.
(29) 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*, No. eaav1190.
(30) Miller, E. D.; Jones, M. L.; Henry, M. M.; Stanfill, B.; Jankowski, E. Machine learning predictions of electronic couplings for charge transport calculations of P3HT. *AIChE J.* **2019**, *65*, No. e16760.
(31) Jackson, N. E.; Bowen, A. S.; de Pablo, J. J. Efficient Multiscale Optoelectronic Prediction for Conjugated Polymers. *Macromolecules* **2020**, *53*, 482−490.
(32) 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. U.S.A.* **2020**, *117*, 13945−13948.
(33) Lee, C.-K.; Lu, C.; Yu, Y.; Sun, Q.; Hsieh, C.-Y.; Zhang, S.; Liu, Q.; Shi, L. Transfer learning with graph neural networks for optoelectronic properties of conjugated oligomers. *J. Chem. Phys.* **2021**, *154*, 024906.
(34) Sivaraman, G.; Jackson, N. E. Coarse-Grained Density Functional Theory Predictions via Deep Kernel Learning. *J. Chem. Theory Comput.* **2022**, *18*, 1129−1141.
(35) Maier, C. J.; Jackson, N. Bypassing Backmapping: Coarse-Grained Electronic Property Distributions Using Heteroscedastic Gaussian Processes. *J. Chem. Phys.* **2022**, *157*, 174102.
(36) Reith, D.; Pütz, M.; Müller-Plathe, F. Deriving effective mesoscale potentials from atomistic simulations. *J. Comput. Chem.* **2003**, *24*, 1624−1636.
(37) Webb, M. A.; Delannoy, J.-Y.; de Pablo, J. J. Graph-Based Approach to Systematic Molecular Coarse-Graining. *J. Chem. Theory Comput.* **2019**, *15*, 1199−1208.
(38) 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.
(39) Bondarenko, A. S.; Patmanidis, I.; Alessandri, R.; Souza, P. C. T.; Jansen, T. L. C.; de Vries, A. H.; Marrink, S. J.; Knoester, J. Multiscale modeling of molecular structure and optical properties of complex supramolecular aggregates. *Chem. Sci.* **2020**, *11*, 11514−11524.
(40) Alessandri, R.; Barnoud, J.; Gertsen, A. S.; Patmanidis, I.; de Vries, A. H.; Souza, P. C. T.; Marrink, S. J. Martini 3 Coarse-Grained Force Field: Small Molecules. *Adv. Theory Simul.* **2022**, *5*, 2100391.
(41) Wu, Z.; Ramsundar, B.; Feinberg, E. N.; Gomes, J.; Geniesse, C.; Pappu, A. S.; Leswing, K.; Pande, V. MoleculeNet: a benchmark for molecular machine learning. *Chem. Sci.* **2018**, *9*, 513−530.
(42) Jorgensen, W. L.; Tirado-Rives, J. Potential energy functions for atomic-level simulations of water and organic and biomolecular systems. *Proc. Natl. Acad. Sci. U.S.A* **2005**, *102*, 6665−6670.
(43) Dodda, L. S.; Cabeza de Vaca, I.; Tirado-Rives, J.; Jorgensen, W. L. LigParGen web server: an automatic OPLS-AA parameter generator for organic ligands. *Nucleic Acids Res.* **2017**, *45*, W331−W336.
(44) Allen, A. E. A.; Payne, M. C.; Cole, D. J. Harmonic Force Constants for Molecular Mechanics Force Fields via Hessian Matrix Projection. *J. Chem. Theory and Comput.* **2018**, *14*, 274−281.
(45) Horton, J. T.; Allen, A. E. A.; Dodda, L. S.; Cole, D. J. QUBEKit: automating the Derivation of Force Field Parameters from Quantum Mechanics. *J. Chem. Inf. Model.* **2019**, *59*, 1366−1381.
(46) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Petersson, G. A.; Nakatsuji, H.; Li, X.; Caricato, M.; Marenich, A. V.; Bloino, J.; Janesko, B. G.; Gomperts, R.; Mennucci, B.; Hratchian, H. P.; Ortiz, J. V.; Izmaylov, A. F.; Sonnenberg, J. L.; Williams-Young, D.; Ding, F.; Lipparini, F.; Egidi, F.; Goings, J.; Peng, B.; Petrone, A.; Henderson, T.; Ranasinghe, D.; Zakrzewski, V. G.; Gao, J.; Rega, N.; Zheng, G.; Liang, W.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Throssell, K.; Montgomery, J. A., Jr.; Peralta, J. E.; Ogliaro, F.; Bearpark, M. J.; Heyd, J. J.; Brothers, E. N.; Kudin, K. N.; Staroverov, V. N.; Keith, T. A.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A. P.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Millam, J. M.; Klene, M.; Adamo, C.; Cammi, R.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Farkas, O.; Foresman, J. B.; Fox, D. J. *Gaussian 16, Revision A.03*; Gaussian Inc.: Wallingford, CT, 2016.
(47) Michaud-Agrawal, N.; Denning, E. J.; Woolf, T. B.; Beckstein, O. MDAnalysis: a toolkit for the analysis of molecular dynamics simulations. *J. Comput. Chem.* **2011**, *32*, 2319−2327.
(48) Gowers, R. J.; Linke, M.; Barnoud, J.; Reddy, T. J. E.; Melo, M. N.; Seyler, S. L.; Dotson, D. L.; Domanski, J.; Buchoux, S.; Kenney, I. M.; Beckstein, O. MDAnalysis: a Python package for the rapid analysis of molecular dynamics simulations. In *Proceedings of the 15th Python in Science Conference (Scipy)*, Austin, TX, 2016; pp 102−109.
(49) O'Boyle, N. M.; Banck, M.; James, C. A.; Morley, C.; Vandermeersch, T.; Hutchison, G. R. Open Babel: An open chemical toolbox. *J. Cheminformatics* **2011**, *3*, 1−14.
(50) Theobald, D. L. Rapid calculation of RMSDs using a quaternion-based characteristic polynomial. *Acta Crystallogr.* **2005**, *61*, 478−480.
(51) Liu, P.; Agrafiotis, D. K.; Theobald, D. L. Fast determination of the optimal rotational matrix for macromolecular superpositions. *J. Comput. Chem.* **2010**, *31*, 1561−1563.
(52) Troisi, A.; Orlandi, G. Hole Migration in DNA: a Theoretical Analysis of the Role of Structural Fluctuations. *J. Phys. Chem. B* **2002**, *106*, 2093−2101.
(53) Lu, T.; Chen, F. Multiwfn: A multifunctional wavefunction analyzer. *J. Comput. Chem.* **2012**, *33*, 580−592.
(54) 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.
(55) Grünewald, F.; Alessandri, R.; Kroon, P. C.; Monticelli, L.; Souza, P. C. T.; Marrink, S. J. Polyply; a python suite for facilitating simulations of macromolecules and nanomaterials. *Nat. Commun.* **2022**, *13*, 68.
(56) 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.
(57) 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*, 19−25.
(58) Chollet, F. *Keras*; 2015; https://github.com/fchollet/keras.
(59) Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V.; Vanderplas, J.; Passos, A.; Cournapeau, D.; Brucher, M.; Perrot, M.; Duchesnay, É. Scikit-learn: Machine Learning in Python. *J. Mach. Learn. Res.* **2011**, *12*, 2825−2830.

https://doi.org/10.1021/acs.macromol.3c00141
Macromolecules 2023, 56, 3574−3584