UNDERSTANDING SELF-ASSEMBLY AND CHARGE TRANSPORT IN ORGANIC SOLAR CELLS THROUGH EFFICIENT COMPUTATION

by

Evan Miller

A dissertation

submitted in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy in Materials Science and Engineering

Boise State University

August 2019

BOISE STATE UNIVERSITY GRADUATE COLLEGE

DEFENSE COMMITTEE AND FINAL READING APPROVALS

of the dissertation submitted by

Evan Miller

Dissertation Title: Understanding Self-Assembly and Charge Transport in Organic Solar Cells Through Efficient Computation

Date of Final Oral Examination: 17th August 2019

The following individuals read and discussed the dissertation submitted by student Evan Miller, and they evaluated the student's presentation and response to questions during the final oral examination. They found that the student passed the final oral examination.

Eric Jankowski, Ph.D.	Chair, Supervisory Committee
Jeunghoon Lee, Ph.D.	Member, Supervisory Committee
Lan Li, Ph.D.	Member, Supervisory Committee
Carla Reynolds, Ph.D.	Member, Supervisory Committee

The final reading approval of the dissertation was granted by Eric Jankowski, Ph.D., Chair of the Supervisory Committee. The dissertation was approved by the Graduate College.

Dedicated to Denise and Thea

ACKNOWLEDGMENTS

Firstly, I would like to thank my advisor Dr. Eric Jankowski. I believe I was extraordinarily blessed in finding an advisor who was so supportive of my growth as a scientist and enabling me to find my research interests. I would also like to thank Dr. Matthew Jones, who was a constant source of guidance for various projects and a compatriot in exploring data science and machine learning. Thank you to all the other members of the CME lab, both past and present. I am incredibly grateful to work with such an amazing group of researchers. It would have been impossible for me to do this research without this group and other groups' effort so in Appendix A I list more explicitly the contributions to and the tools used this work.

I would also like to thank my family members: my parents who were great examples of teaching and learning and my brothers with whom I spent many hours discussing scientific topics around the dinner table. Lastly, I would like to thank my wife Denise. I am eternally grateful for her sacrifice in coming to Boise and her constant support over the years.

ABSTRACT

Organic solar cells capable of sustainably generating electricity are possible if: (1) The structures assembled by photoactive molecules can be controlled, and (2) The structures favorable for charge transport can be determined. In this dissertation we conduct computational studies to understand relationships between organic solar cell compounds, processing, structure and charge transport. We advance tools for encapsulating computational workflows so that simulations are more reproducible and transferable. We find that molecular dynamic simulations using simplified models efficiently predict experimental structures. We find that the mobilities of charges through these structures—as determined by kinetic Monte Carlo simulations—match qualitative trends expected with molecular ordering and in some cases agree quantitatively with experimental measurements. We identify percolating clusters with overlapping pi-orbitals as vital for fast charge transport, which are achieved through polymer "tie-chains" and extended molecular stacking. We find that machine learning predictions of electronic couplings from quantum chemical calculations gives a two-order-of-magnitude speed improvement relating structure to charge transport versus repeating the quantum calculations. We identify limitations of our structural and charge transport predictions, and provide recommendations for advancing future investigations of organic solar cells. In sum, the computational tools developed and employed herein enable the most broad and experimentally-validated sampling of self-assembled structure as a function of chemistry and processing to date. The fundamental understanding gained from these simulations informs the self-assembly

and structure-transport relationships needed to advance organic solar cell engineering.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AA – All-Atom

ANN – Artificial Neural Network

BHJ – Bulk Heterojunction

COM – Center-of-Mass

CG – Coarse-Grained

CH – Cahn-Hilliard

CHARM – Chemistry at Harvard Macromolecular Mechanics

DBP – Dibenzo-tetraphenyl Periflanthene

GAFF – Generalized Amber Force-Field

GIXS – Grazing Incidence X-ray Scattering

GPU – Graphical Processing Unit

HPC – High Performance Computing

HOMO – Highest Occupied Molecular Orbital

HOOMD – Highly Optimized Object-oriented Many-particle Dynamics

ITIC – 3,9-bis(2-methylene-(3-(1,1-dicyanomethylene)-indanone))- 5,5,11,11- tetrakis (4-hexylphenyl)-dithieno[2,3-d:2',3'-d']-s-indaceno [1,2-b:5,6-b'] dithiophene

KMC – Kinetic Monte Carlo

KNN – K-Nearest Neighbors

LUMO – Lowest Unoccupied Molecular Orbital

MC – Monte Carlo

MD – Molecular Dynamics

MSD – Mean-Squared Displacement

ML – Machine Learning

MD – Molecular Dynamics

NVT – Canonical ensemble: constant number of particles, volume and temperature

OLS – Ordinary Lease Squares

OPLS – Optimized Potentials for Liquid Simulations

OPV – Organic Photovoltaic

P3HT – Poly-(3-hexylthiophene)

PAH – Polycyclic Aromatic Hydrocarbon

Pe – Perylene

Pt – Perylothiophene

QCC – Quantum Chemical Calculations

RF – Random Forest

SVM – Support Vector Machine

TraPPE – Transferable Potential for Phase Equilibria

TRUE – Transferable, reproducible, understandable and extensible

UA – United-Atom

VRH – Variable Range Hopping

ZINDO/S – Zerner's Intermediate Neglect of Differential Overlap

LIST OF SYMBOLS

α Tuning parameter

CA   Aromatic Carbon

CT   Aliphatic Carbon

D Diffusivity

$\Delta E_{i,j}$ Energy difference

dt Timestep

ε Depth of the Lennard-Jones potential well

ϵ Electrical Permittivity

ε_s Scaling factor for ε, represents inverse solvent strength

η Power conversion efficiency

F Force

FF Fill factor

$g(r)$ Radial distribution function

$\hbar$ Planck's reduce constant

$J_{i,j}$ Electronic coupling or transfer integral

$J - V$ Current-Voltage curve

J_{SC} Short circuit current

k_{b,θ,d_n} Spring constant for bonds, angles or dihedrals

k_B Boltzmann constant

$k_{i,j}$ Rate coefficient

λ Reorganization energy

μ Charge mobility in cm^2/Vs

μ_0 Zero-field charge mobility in cm^2/Vs

O   Oxygen

N Number of simulation elements

P_{IN} Input power spectrum

P_{MAX} Maximum power point in $J - V$ curve

ϕ Packing fraction

φ The dihedral angle describing four simulation elements

ψ Order parameter describing the number of "large" clusters

$\psi\prime$ The ψ order parameter normalized by the relative standard deviations of bond lengths

S   Sulfur

s_q Structure factor

σ_{LJ} Van der Waals radius

σ Standard deviation

r Separation

ρ Mass density, normally in units of g/cm^3

θ The angle defining three simulation elements

t Time

t_a Decorrelation time

t_r Relaxation time

T Temperature

T_{KMC} Temperature in a Kinetic Monte Carlo simulation

τ KMC Time unit

τ_s MD time unit

TPS Timesteps per second

U Energy, the subscript signifies what kind of energy

v velocity

V_{OC} Open-circuit voltage

x_r Random number over $[0, 1)$

ξ Order parameter describing if stacked molecules are in register

CHAPTER 1

INTRODUCTION

1.1 Motivation

Limiting global climate change while providing power for a growing human population necessitates developing renewable energy that is scalable and inexpensive[15]. The sun provides more than three orders of magnitude the amount of energy needed by humans and can meet current and projected demand well beyond the year 2100[24] making it an ideal candidate to meet these growing needs. Traditional inorganic devices have limitations which prevent their viability to meet this need, namely high production costs and competition for the materials with the microelectronics industry,[15] which keeps material demand and prices high. Organic photovoltaics (OPVs)—which use carbon-based compounds to convert light into electricity—are advantageous over traditional silicon devices due to low-cost, scalable manufacturing[2,20,21,23]. OPVs can also have energy pay-back times as little as one day if estimated 15% power conversion efficiencies (η) are achieved[8]—a threshold that has recently been surpassed by laboratory devices[16,28].

However, it typically takes $\sim$4 years to scale up production from prototype, and mass-produced devices are typically $\sim$5% less efficient than champion laboratory devices[4]. The challenge to mass-producing efficient devices is reliably controlling active layer morphology—the spontaneously forming microstructure of organic electron

donor and acceptor—that determines charge generation and collection, and therefore η[5]. Thermodynamic self-assembly can provide the needed morphological control in principle. However, answering "What self-assembled morphologies permit high η?" and "What ingredient-processing combinations will obtain these structures?" is difficult due to the complex interplay between materials[14], solvents[7], annealing protocols[19,25,26] and temperature[13]. That is, it is infeasible to enumerate all the infinite combinations of ingredients and processing protocols that could be used to manufacture OPVs. To identify the most promising candidates we use computer simulations, which are faster to perform and can provide atomistic insight unavailable in laboratory experiments. We apply high-performance computing (HPC) to investigate thousands of candidates, looking for patterns in the structures that facilitate charge transport and the ingredient-processing conditions that result in these structures.

Addressing the issues of active layer structure and assembly with HPC still has challenges due to the time- and length- scales that must be considered and their associated trade-offs to throughput[11]. For example, to traverse a ~100 nm active layer, a single charge may undergo billions of femtosecond-long angstrom-scale hops between molecules. A complete picture of charge transport therefore requires molecular resolution at the device scale. Molecular dynamics simulations, especially with modern hardware accelerators and new models, can help to bridge these scales[3]. However, computationally equilibrating morphologies is hindered by long relaxation times that increase with the number of simulation elements[9]. Therefore, selecting reasonable system sizes and applying computation-reducing assumptions such as coarse-graining[10], "rigid" bodies[18], implicit electrostatics[17] and implicit solvents[22] are necessary.

In addition to utilizing efficient models to increase performance, it is also impor-

tant to use simulation "best-practices" to improve researcher efficiency[12]. These best-practices include using, refactoring and contributing to peer-reviewed software, which will save time in writing and de-bugging prototype code[1]. Utilizing these best-practices also has the benefit of making simulations more transferable, reusable, usable and extensible or *TRUE*. As such, these best practices make scientific investigations, including investigating OPV self-assembly and charge-transport, more efficient and reproducible[6,27]

1.2   Outline

This dissertation is structured as follows: In Chapter 2 we review the basics of OPVs, including the physics behind light-to-electricity conversion and quantifying device performance. We then explain the computational methods employed in this study and put our simulation work in context. In Chapter 3, we present our self-assembly studies of two polyaromatic hydrocarbons, showing how simplified models enable thousands of thermodynamic statepoints to be sampled. In Chapter 4, we develop an optimized force-field for poly-(3-hexylthiophene) (P3HT), apply it to oligomers, validate against experiments and conclude with a "recipe" for simulating OPV self-assembly. In Chapter 5 we pivot to calculating charge transport, first in blends of dibenzo-tetraphenyl periflanthene (DBP) with fullerene derivatives. In Chapter 6 we apply our charge transport techniques to P3HT and incorporate chain polydispersity. We generate a mobility "heatmap" that we use to correlate structural features with charge mobility and confirm how "tie-chains" are vital to charge movement between P3HT crystallites. In Chapter 7 we ameliorate a bottleneck in linking P3HT structure with charge transport by using machine learning to reduce transfer integral determination time

by $\sim 400\times$. Lastly, in Chapter 8 we describe our tool `Planckon` for making OPV simulations more *TRUE* and easy to peform.

BIBLIOGRAPHY

[1] Carl S. Adorf, Joshua A. Anderson, and Sharon C. Glotzer. How to Professionally Develop Reusable Scientific Software – And When Not To. *Computing in Science & Engineering*, 21:66–79, 2018.

[2] Ana Claudia Arias, J Devin MacKenzie, Iain McCulloch, Jonathan Rivnay, and Alberto Salleo. Materials and Applications for Large Area Electronics: Solution-Based Approaches. *Chemical Reviews*, 110(1):3–24, 2010.

[3] Efrem Braun, Justin Gilmer, Heather B. Mayes, David L. Mobley, Jacob I. Monroe, Samarjeet Prasad, and Daniel M. Zuckerman. Best Practices for Foundations in Molecular Simulations [Article v1.0]. *Living Journal of Computational Molecular Science*, 1(1), 2019.

[4] Jon E. Carlé, Martin Helgesen, Ole Hagemann, Markus Hösel, Ilona M. Heckler, Eva Bundgaard, Suren A. Gevorgyan, Roar R. Søndergaard, Mikkel Jørgensen, Rafael García-Valverde, Samir Chaouki-Almagro, José A. Villarejo, and Frederik C. Krebs. Overcoming the Scaling Lag for Polymer Solar Cells. *Joule*, 1(2):274–289, 2017.

[5] Wei Chen, Maxim P. Nikiforov, and Seth B. Darling. Morphology Characterization in Organic and Hybrid Solar Cells. *Energy & Environmental Science*, 5(8):8045–8074, 2012.

[6] Tom Crick, Benjamin A. Hall, and Samin Ishtiaq. Reproducibility in Research: Systems, Infrastructure, Culture. *arXiv*, pages 1–12, 2015.

[7] Minh Trung Dang, Guillaume Wantz, Habiba Bejbouji, Mathieu Urien, Olivier J. Dautel, Laurence Vignau, and Lionel Hirsch. Polymeric Solar Cells Based on P3HT:PCBM: Role of the Casting Solvent. *Solar Energy Materials and Solar Cells*, 95(12):3408–3418, 2011.

[8] Nieves Espinosa, Markus Hösel, Dechan Angmo, and Frederik C. Krebs. Solar Cells with One-Day Energy Payback for the Factories of the Future. *Energy & Environmental Science*, 5(1):5117, 2012.

[9] D. Frenkel. Simulations: The Dark Side. *The European Physical Journal Plus*, 128(1):10, 2013.

[10] Thomas E. Gartner and Arthi Jayaraman. Modeling and Simulations of Polymers: A Roadmap. *Macromolecules*, 52:755–786, 2019.

[11] Thomas F. Harrelson, Adam J. Moulé, and Roland Faller. Modeling Organic Electronic Materials: Bridging Length and Time Scales. *Molecular Simulation*, 43(10-11):730–742, 2017.

[12] Eric Jankowski, Neale Ellyson, Jenny W. Fothergill, Michael M. Henry, Mitchell H. Leibowitz, Evan D. Miller, Mone't Alberts, Samantha Chesser, Jaime D. Guevara, Chris D. Jones, Mia Klopfenstein, Kendra K. Noneman, Rachel Singleton, Ramon A. Uriarte-Mendoza, Stephen Thomas, Carla E. Estridge, and Matthew L. Jones. Perspective on Coarse-Graining, Cognitive Load, and Materials Simulation. *Computational Materials Science*, Pending Reveiw, 2019.

[13] F.C. Krebs, Sabine Bertho, Griet Janssen, Thomas J. Cleij, Bert Conings, Wouter Moons, Abay Gadisa, Jan D'Haen, Etienne Goovaerts, Laurence Lutsen, Jean Manca, Dirk Vanderzande, Jan D'Haen, Etienne Goovaerts, Laurence Lutsen, Jean Manca, and Dirk Vanderzande. Effect of Temperature on the Morphological and Photovoltaic Stability of Bulk Heterojunction Polymer:Fullerene Solar Cells. *Solar Energy Materials and Solar Cells*, 92(7):753–760, 2008.

[14] Hansol Lee, Chaneui Park, Dong Hun Sin, Jong Hwan Park, and Kilwon Cho. Recent Advances in Morphology Optimization for Organic Photovoltaics. *Advanced Materials*, 30(34):1–39, 2018.

[15] Katherine A. Mazzio and Christine K Luscombe. The Future of Organic Photovoltaics. *Chemical Society Reviews*, 44(1):78–90, 2015.

[16] Lingxian Meng, Yamin Zhang, Xiangjian Wan, Chenxi Li, Xin Zhang, Yanbo Wang, Xin Ke, Zuo Xiao, Liming Ding, Ruoxi Xia, Hin-Lap Yip, Yong Cao, and Yongsheng Chen. Organic and Solution-Processed Tandem Solar Cells With 17.3% Efficiency. *Science*, 2612:1094–1098, 2018.

[17] Evan Miller, Matthew Jones, Michael Henry, Paul Chery, Kyle Miller, and Eric Jankowski. Optimization and Validation of Efficient Models for Predicting Polythiophene Self-Assembly. *Polymers*, 10(12):1305, 2018.

[18] Evan D. Miller, Matthew L. Jones, and Eric Jankowski. Enhanced Computational Sampling of Perylene and Perylothiophene Packing with Rigid-Body Models. *ACS Omega*, 2(1):353–362, 2017.

[19] Jong Hwan Park, Jong Soo Kim, Ji Hwang Lee, Wi Hyoung Lee, and Kilwon Cho. Effect of Annealing Solvent Solubility on the Performance of Poly(3-hexylthiophene)/Methanofullerene Solar Cells. *The Journal of Physical Chemistry C*, 113:17579–17584, 2009.

[20] Sean E. Shaheen, David S Ginley, and Ghassan E Jabbour. Organic-Based Photovoltaics: Toward Low-Cost Power Generation. *MRS Bulletin*, 30(01):10–19, 2005.

[21] Sean E. Shaheen, Rachel Radspinner, Nasser Peyghambarian, and Ghassan E. Jabbour. Fabrication of Bulk Heterojunction Plastic Solar Cells by Screen Printing. *Applied Physics Letters*, 79(18):2996–2998, 2001.

[22] Hyeyoung Shin, Tod A. Pascal, William A. Goddard, and Hyungjun Kim. Scaled Effective Solvent Method for Predicting the Equilibrium Ensemble of Structures with Analysis of Thermodynamic Properties of Amorphous Polyethylene Glycol-Water Mixtures. *Journal of Physical Chemistry B*, 117(3):916–927, 2013.

[23] Roar R. Søndergaard, Markus Hösel, and Frederik C. Krebs. Roll-to-Roll Fabrication of Large Area Functional Organic Materials. *Journal of Polymer Science, Part B: Polymer Physics*, 51(1):16–34, 2013.

[24] United Nations Development Programme. *World Energy Assessment. Energy and the Challenge of Sustainability*. 2000.

[25] Eric Verploegen, Chad E. Miller, Kristin Schmidt, Zhenan Bao, and Michael F. Toney. Manipulating the Morphology of P3HT-PCBM Bulk Heterojunction Blends with Solvent Vapor Annealing. *Chemistry of Materials*, 24(20):3923–3931, 2012.

[26] Eric Verploegen, Rajib Mondal, Christopher J Bettinger, Seihout Sok, Michael F Toney, and Zhenan Bao. Effects of Thermal Annealing Upon the Morphology of Polymer-Fullerene Blends. *Advanced Functional Materials*, 20(20):3519–3529, 2010.

[27] Greg Wilson, D. A. Aruliah, C. Titus Brown, Neil P. Chue Hong, Matt Davis, Richard T. Guy, Steven H. D. Haddock, Kathryn D. Huff, Ian M. Mitchell, Mark D. Plumbley, Ben Waugh, Ethan P. White, Paul Wilson, and N. P. C. Hong. Best Practices for Scientific Computing. *PLoS biology*, 340(May):e1001745, 2014.

[28] Jun Yuan, Yunqiang Zhang, Liuyang Zhou, Guichuan Zhang, Hin-Lap Yip, Tsz-Ki Lau, Xinhui Lu, Can Zhu, Hongjian Peng, Paul A. Johnson, Mario Leclerc, Yong Cao, Jacek Ulanski, Yongfang Li, and Yingping Zou. Single-Junction Organic Solar Cell with over 15% Efficiency Using Fused-Ring Acceptor with Electron-Deficient Core. *Joule*, pages 1–12, 2019.

CHAPTER 2

BACKGROUND

Organic photovoltaics (OPVs) have gained significant attention as a technology to provide low-cost, scalable energy. In this chapter, we aim to provide a background on OPVs, describe their current state and discuss how we aim to improve them. In Section 2.1, we discuss the physical properties of organic molecules that make the conversion of light into electricity possible and make comparisons to their inorganic counterparts. Section 2.1 also provides insight into how the self-assembled structure affects the conversion of light into electricity. In Section 2.2 we then discuss how computers can be used to predict structure and charge transport in OPVs. Lastly, in Section 2.2.3 we present the state of the field for using computer simulations to understand OPV self-assembly and charge transport.

2.1 Organic Photovoltaic Basics

2.1.1 Conjugated Organic Molecules

The defining feature of an organic compound is that it is primarily composed of carbon with other light, non-metallic atoms such as hydrogen, nitrogen, oxygen and sulfur. All atoms are composed of positively charged nuclei with negative electron orbital "clouds" surrounding their nuclei[9]. For the light elements frequently found in OPVs, there are four orbitals that are of interest: a spherical s and three "dumbbell" shaped
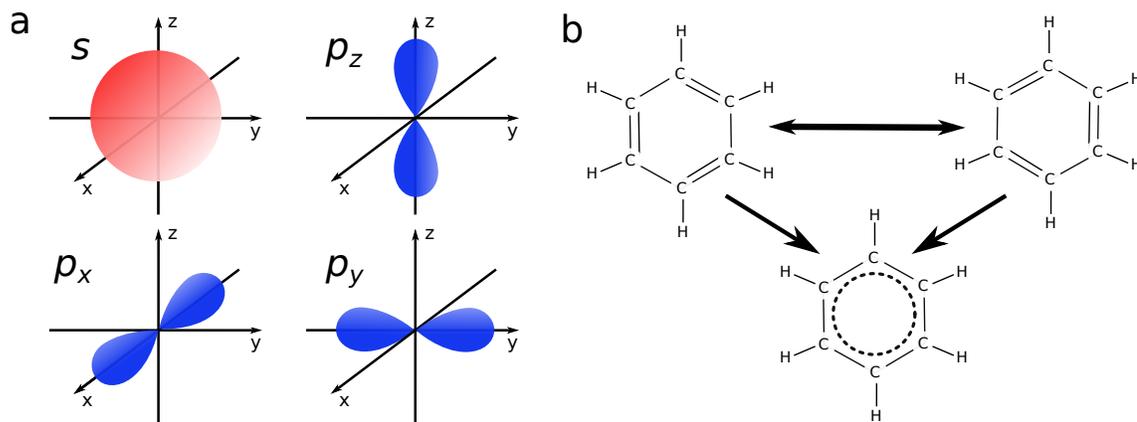
Figure 2.1: The important electron orbitals in organic compounds and the resonant structure of benzene. (a) The important orbitals in carbon are: s, p_x, p_y and p_z. (b) The molecule benzene is a simple example of an aromatic molecule, rather than having alternating single and double bonds, it has resonant bonds in which the electrons are diffused over the whole of the molecule.

p orbitals, where each p orbital is oriented along the three different axes (p_x, p_y and p_z see Figure 2.1a). An atom's lowest energy configuration can be achieved by filling the outer shell of electrons, often referred to as the "octet rule", by sharing electrons via overlapping electron orbitals with neighboring atoms (covalent bonding)[44]. In carbon, there are four electrons in its outermost shell with which it will typically form four bonds[9]. This bonding can be in the form of four single bonds where each of its four orbitals is overlapping with a different partner atom. Another way carbon can satisfy its octet rule is by forming two single bonds and one double bond in which two orbitals are shared between two atoms (termed sp^2 hybridization). When a double bond is formed, one sharing of the electrons exists directly between the atoms - a σ-bond, and the second sharing exists out of the plane between the two atoms - a π-bond[72]. The presence of the additional π-bond results in a reduction of the distance (1.54 Å) between the carbons to 1.34 Å[99]. Some molecular structures result in alternating single and double bonds. In these structures, the double and

single bonds can be thought as repeatedly switching to create a resonant or aromatic structure in which the π-bond is dispersed between all three atoms. This can be shown in the example of the molecule benzene, consisting of six carbon and six hydrogen atoms (see 2.1b). Each carbon in benzene is double bonded to one carbon, single bonded to a carbon and single bonded to the hydrogen. However, this double-single bond structure is not what is observed experimentally, rather the distances between carbons are constant (1.40 Å[99]), indicating that the double bond is shared with both neighbors simultaneously. This aromaticity over multiple atoms and bonds makes structures to more stable[16], and because the electrons are distributed over more atoms, the structure is more resilient to the addition/removal of charges.

Because these aromatic organic compounds are able to carry a surplus or absence of charges, they are electronically active and free charges are able to travel over these types of molecules in the form of quantized "hops"[48]. Because these compounds can transport charges they can be used in electronic devices, and these compounds are used as semi-conductors in the device[118]. The designation "semi-conductor" is assigned based on the energy gap between the electron filled and electron empty states for the material of < 2 eV[9]. In inorganic devices this is the energy gap between the conduction and valence bands. In organic compounds, the gap is described by the difference between the highest occupied molecular orbital (HOMO) and the lowest occupied molecular orbital (LUMO).

Another way organic compounds are analogous to, but distinct from their inorganic counterparts is in tuning their band-gaps to obtain desired electronic behavior. For example the Fermi-level (which marks the top of the valence band at 0 K in semiconductors) can be decreased or increased by doping with atomic species with a deficit of electrons ($p-type$) or a surplus ($n-type$)[107]. When a $p-type$ and $n-type$

semiconductor are placed together to create a $p-n$ junction, the energy difference between the two types creates a mechanism for controlling the direction of charges within device[108]. In organic compounds, this changing of the HOMO and LUMO levels can be accomplished by functionalizing, or bonding different chemical species to the molecule[71]. To help distinguish these electroactive molecules, we hereafter refer to these molecules as "chromophores". By tailoring the electronic levels of different chromophores, electrons have a driving force to move from the higher energy chromophore (donor) to the lower energy chromophore (acceptor)[8].

One difference between inorganic and organic semiconductors in how these charges move. In inorganic semiconductors, the deficit or surplus of electron states allows electrons to move between unfilled states,[107] which allows for mobilities higher than those seen in organic compounds (up to $\sim 10^3$ cm^2/Vs)[26]. Conversely, charge movement between chromophores is accomplished through quantum tunneling from the electron orbital of chromophore i to chromophore j, resulting in mobilities over the range of 10^{-6} to 1 cm^2/Vs[8]. In addition to the differences in charge movement, the electrical permittivity for organic compounds is lower ($\epsilon_{r,\text{organic}} \sim 3$)[67] than inorganic silicon ($\epsilon_{r,\text{silicon}} \sim 11.7$). This results in charges not being screened as efficiently from each other in organic devices compared to inorganic devices, and has effects on generating electricity from light[38].

For OPVs to convert light into electricity, four (simplified) steps must be completed (see also Figure 2.2)[80]:

1. A photon is absorbed in the photoactive layer to create a coulombically bound electron-hole pair called an exciton - which are more strongly coulombically bound because of the low ϵ_r.
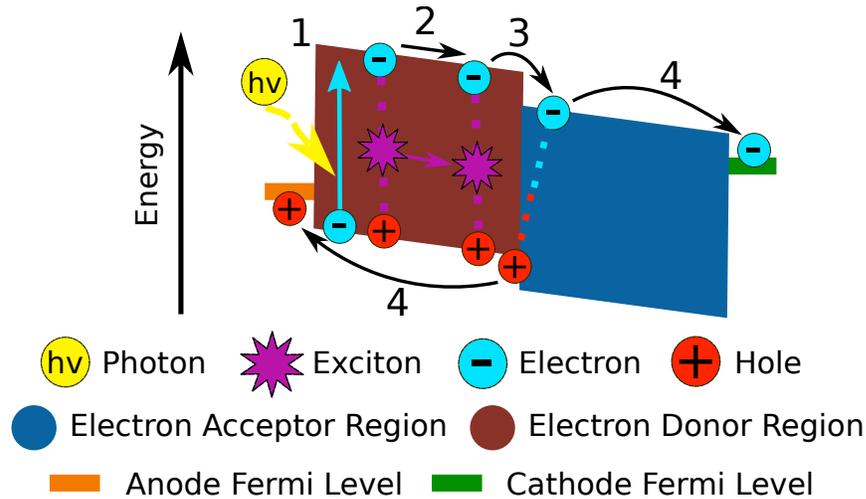
Figure 2.2: The four steps for converting light into electricity in an OPV are (1) coulombically bound electron-hole pair "exciton" creation, (2) exciton diffusion, (3) exciton dissociation into free charges and (4) charge transport to the electrodes.

2. The exciton diffuses to an electron acceptor-donor interface.

3. The exciton dissociates into a free electron and hole at the interface.

4. The free charges travel through the photoactive layer to the electrodes.

Issues can occur at all four of these steps and result in a loss of the power, and there is guaranteed to be significant losses during these steps. This guaranteed loss is described by the Shockley-Queisser limit: the thermodynamic limit of any single-junction terrestrial solar cell at $\sim 34\%$ power conversion efficiency η[98]. One significant loss is in the photon energy that will be used by the solar cell, e.g. light will be reflected rather than absorbed[106]. Additionally, the solar cell can only make use of a portion of the photon energies that are shown on it; energies below that of the HOMO-LUMO gap cannot be used to create power and energies beyond that of the gap will be lost through relaxation processes[28]. Even if the device successfully absorbs a photon, to create a (Frenkel) exciton[38], it is not guaranteed that the

energy will be collected. In order for energy to be gathered from the exciton, the exciton must diffuse to an interface between the donor and acceptor chromophores in its lifetime where it is able to dissociate into a free electron and hole[17,31]. However, there are limits on how far (~ 10 nm[29]) and how long (~ 3 ns[73]) the exciton will travel before the exciton decays back into the ground state. Consequently, obtaining structures in OPVs that allow for excitons to quickly reach interfaces to maximize the number of dissociations will be important for developing efficient OPVs.

Even if an exciton successfully dissociates into an electron and hole, loss can occur in the form of recombination events: geminate and non-geminate (or bimolecular)[21]. Geminate recombinations are recombinations in which charges generated from the same photon recombine. The relaxation of the coulombically bound electron and hole pair exciton, is one form of geminate recombination. After the exciton has successfully diffused and separated at an acceptor-donor interface, it is possible for these newly freed charges to still recombine and the energy to be lost in another form of geminate recombination. If the charges are successfully separated and begin to move towards opposite electrodes, it is still possible for them to encounter a charge of the opposite type, that was either photo-generated in its own event or injected at one of the electrodes, and recombine (bimolecular).

The amount of electrical energy an OPV is able to produce based on the amount of light shining on it is quantified power conversion efficiency η. To calculate η, the electrical current J flowing through the device is compared to voltage (V) across the device in the form of a $J - V$ curve (Figure 2.3). Two properties worth noting in Figure 2.3 are the short-circuit current J_{SC}, the open-circuit voltage V_{OC}. J_{SC} is the electrical current that flows within the device when there is no net voltage across the device i.e. the internal and external electric fields are balanced. V_{OC} is the voltage
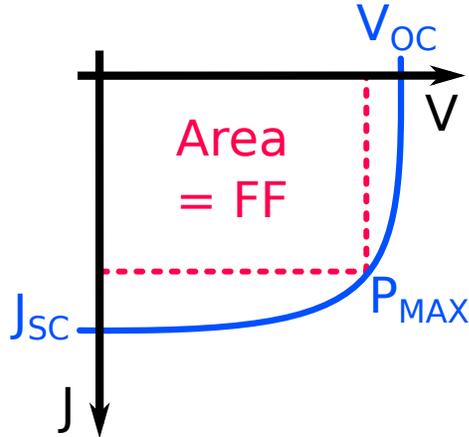
Figure 2.3: A simplified $J - V$ curve. The values on the $J - V$ curve help quantify the η of a solar cell.

across the device that reduces the current J to 0 and represents the internal electrical field (in large part arising from the HOMO-LUMO gap[6]). Preparing devices with high V_{OC} and J_{SC} is important as their trade-off determines the maximum power obtainable from the device (P_{MAX}). The value of P_{MAX} in turn informs the actual operating J and V of the device for calculating the fill-factor FF, which is the area of the largest rectangle that can fit within the $J - V$ curve[14,93]. These properties together provide the power conversion efficiency η:

$$\eta = \frac{FF \times V_{OC} \times J_{SC}}{P_{IN}}, \tag{2.1}$$

in which P_{IN} is the input power spectrum. For consistency devices are often compared against a standard P_{IN}: AM1.5, which corresponds to ~ 1000 W m^{-2} - the approximate power at ground level given by un-obscured sun light at its zenith.

The first OPV had a low η of 1%, produced by C.W. Tang in 1986[105]. Tang's OPV contained a bilayer of electron donor copper pthalocyanine and acceptor perylene tetracarboxyllic derivative. Improvements on this pioneering device show advances
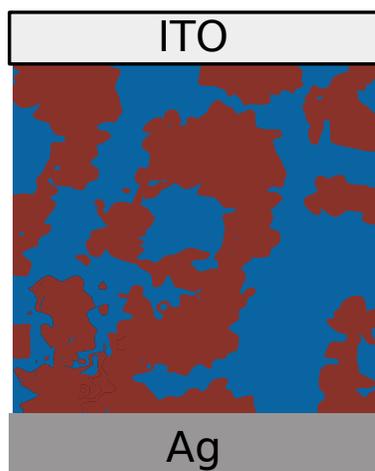
Figure 2.4: Example "inverted" OPV device structure in which the red electron donor and blue electron acceptor types are sandwiched between the transparent Indium Tin Oxide cathode and the Ag anode.

in optimizing the four necessary steps listed in Figure 2.2. One improvement came in the form of better optimizing step 1, by combining components that together absorb photons over the complete solar spectrum. In this device, the perylene derivative was replaced with a buckminsterfullerene (C_{60}), which resulted in an η increase to $\sim 3.6\%$ (note: a hole blocking layer was also implemented which would also affect performance).[94] Rather than optimizing just the choice of compounds, other researches also utilized a bulk heterojunction (BHJ, see Figure 2.4) rather than a bilayer, where the acceptor and donor layers intercalate to increase the surface area of interfaces and allowing more excitons to dissociate into free charges[43,119]. The development of the BHJ increased the maximum η for an OPV to $\sim 5.5\%$ and 6%, and now the BHJ is the standard structure within OPVs. The calculated η necessary for OPVs to be commercially competitive and have short energy payback times is 10-15%[18,30], which is now surpassed in these laboratory devices: over 15% in ternary blend, single junction devices[120] and over 17% in multi-junction devices[81].

However, mass-produced device η must be improved to match their laboratory counterparts if OPVs are to see commercial success[104]. These sample devices highlight two major ways OPVs can be improved: (1) finding better compounds with more complementary photoelectronic properties and (2) optimizing structure. There has been significant work that has gone into identifying molecules with complementary electronic properties[42], however, exploring the processing combinations to identify if and how high efficient structures are obtainable, must be done.

The "brute-force" method to identify which structures can be self-assembled is to create many different devices under many different conditions. Omitting the time and cost that would make such an endeavor unrealistic, this process is unfavorable due to characterization of physical devices providing incomplete pictures of structure[15,24]. Device characterization of OPV thin-films normally fall into two classes: surface and reciprocal-space, scattering techniques. Surface techniques such as atomic force microscopy detect the crystalline and amorphous regions at the surface of the film, but do not reveal structures within the bulk that will be important for charge generation and transport[19]. Conversely, scattering techniques such as grazing incident X-ray scattering (GIXS) probe the bulk material, revealing the averaged periodic features but lack the precisions to resolve a unique position of the atoms[24]. Additionally, the presence of amorphous regions and residual solvents complicate the analysis. In contrast to "wet" methods, some structural computer simulations offer explicit resolution into the structure within the device and are not subject to the same time and cost restraints. Simulated structures can be verified against experimental findings through experimental and simulated GIXS patterns[59]. As such, computer simulations offer a way to much more quickly explore the factors governing OPV self-assembly, and can act as a way to guide experimentalists in creating devices.

However, only predicting self-assembly is insufficient as we must also connect the self-assembled structures to their ability to convert light into electricity. Although some debate exists[41] on whether geminate[82] or bimolecular[21,102,117] recombinations are the largest factor affecting OPV performance, preparing devices that allow for efficient charge collection will be important for commercial OPVs. As such, one metric that represents how "good" a structure is the zero-field mobility μ_0, because structures with high μ_0 are less likely to have charges become stuck until a recombination event occurs. Generally speaking, OPVs with faster carrier mobilities exhibit faster response times and better overall performance[101], and some researchers have proposed that devices with higher charge transport rates are necessary for higher FF[95]. Additionally, a key advantage of selecting the mobility as the performance metric is that it is widely applicable to all organic electronic devices - such as transistors and diodes, and can be directly compared to time-of-flight experiments in which charge movement is measured in the absence of an external field[113].

2.2 Simulations

If we are to use computers to investigate self-assembly and charge transport, we must understand what computational techniques are available. In this section, we discuss the possible simulation methods that can be used, and provide a stronger theoretical background into those applied in this dissertation.

2.2.1 Structural Simulations

One challenge we will face predicting self-assembly in OPVs stems from the length-scales in real devices (~ 100 nm) and the detail that is compromised to achieve these
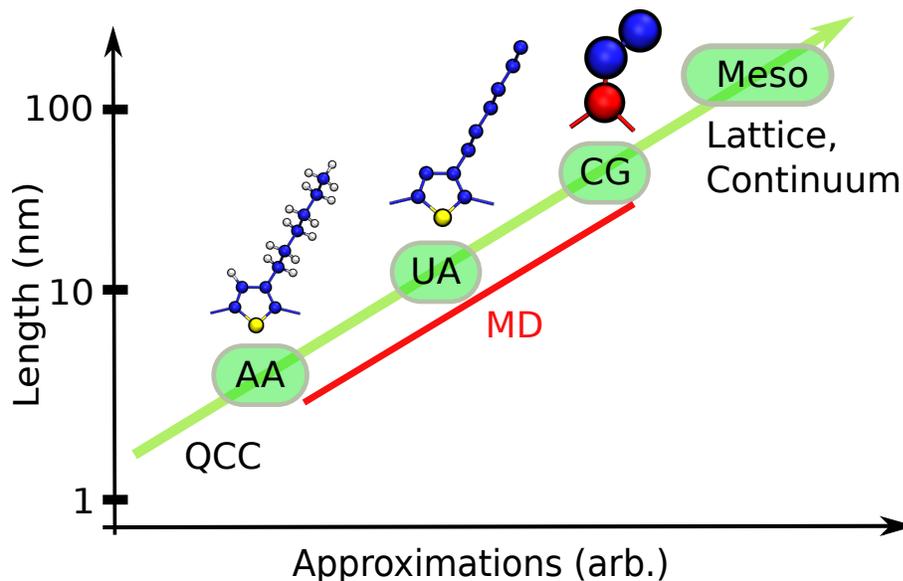
Figure 2.5: Structural simulations can be conducted at a variety of length-scales, but achieving larger simulation sizes requires using techniques with fewer details. QCCs require the fewest approximations, but the computational cost limits this technique to a few hundred atoms. MD iteratively integrates Newton's equation of motion for simulation elements, which allows it to simulate systems over a variety of length-scales depending on the detail of the simulation elements. Lattice models such as metropolis MC can reach hundreds of nm with Cahn-Hilliard or Ising models, but must represent interspecies interactions as an analytical function.

scales (see Figure 2.5)[83]. For instance, quantum chemical calculations (QCC) try to solve the Schrödinger equation for the atom nuclei and electrons[35]. These QCC have the highest resolution and require only atom types and positions, but their computational demands preclude simulating thousands of atoms—far too small for studying bulk morphological features that emerge during self-assembly. On the other end of the spectrum, we have lattice based[45] and finite-element techniques[114] that can easily extend past device scales. However, in these simulation types molecular resolution (such as rotations) is lost, and calculating structures relies on an analytically function that must be informed by more detailed simulations or experiments[4].

Molecular dynamics (MD) exists between these two extremes, and can encompass a range of detail: from atomistic to coarse models which represent groups of atoms as a single element [34]. MD still requires that the interactions are either derived from experimental measurements or more detailed simulations such as QCC, however, there have been multiple libraries created from which molecules can be assigned "off the shelf" interactions based on atom types [64].

Molecular Dynamics
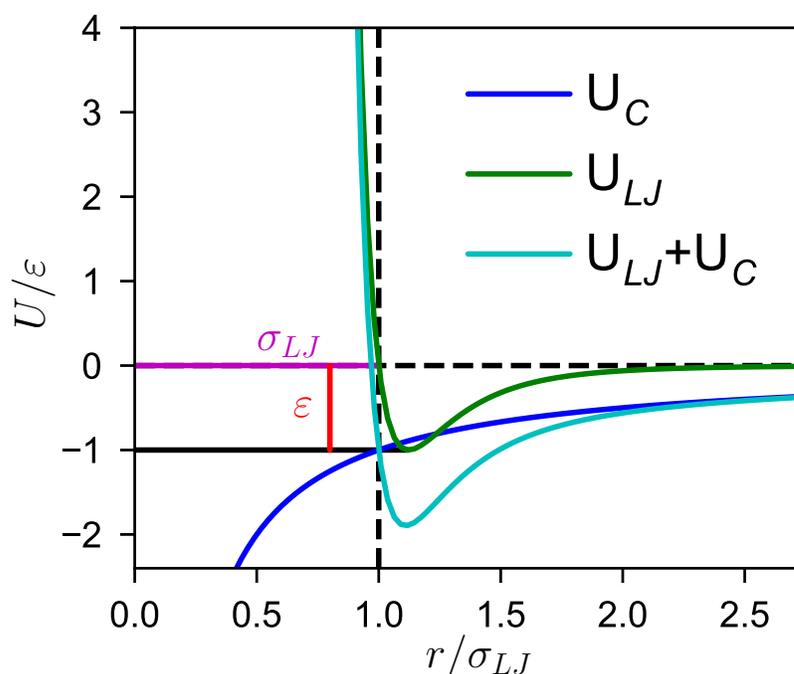
Figure 2.6: Non-bonded interactions considered in this work are the Lennard-Jones and Coulombic. Lennard-Jones interactions dominate at values close to $\sigma_{LJ} = 1$, whereas at longer distances the Coulombic interaction dominates.

MD samples equilibrium ensembles of configurations over time by numerically

integrating Newton's equations of motion for simulation elements - in this case, the simulation elements are atoms[34]. The first MD simulations were conducted by Alder and Wainwright looking at hard-spheres during the 1950s[1]. Shortly after this, Rahman simulated Argon with Lennard-Jones interactions (2.2), and correctly predicted diffusion coefficients[96]. Although there are multiple functional forms that can be used to describe the interactions between elements, the Lennard-Jones interaction is widely used[64] and will be discussed here. The Lennard-Jones interaction considers energetic contributions as a function of distance r arising from the Pauli-exclusion principle (12th power term) and van der Waals attractions (6th power term)[54]:

$$U_{\text{LJ}}(r) = 4\varepsilon \left(\left(\frac{\sigma_{LJ}}{r} \right)^{12} - \left(\frac{\sigma_{LJ}}{r} \right)^{6} \right), \tag{2.2}$$

in which ε is the depth of the Lennard-Jones well and σ_{LJ} is the van der Waals radius. This functional form results in very strong repulsive forces at short distances, an energy minimum at $2^{1/6}\sigma_{LJ}$ and decreasing attractions at longer distances.

In addition to considering the LJ interactions, MD can consider non-bonded Coulombic interactions interactions with the form:

$$U_{\text{C}}(r) = \frac{q_i q_j}{4\pi\epsilon_0\epsilon_r r}, \tag{2.3}$$

in which q_i and q_j are the charges on particles i and j, ϵ_0 is the permittivity of free space, and ϵ_r is the relative permittivity of the material[69]. Many different force-fields, which are parameterizations of ε, σ_{LJ} and q, are based on these two non-bonded interactions: Optimized Performance of Liquid Simulations (OPLS)[60], Transferable Potentials for Phase Equilibria (TraPPE)[78], Generalized Amber Force Field (GAFF)[111] and Chemistry at Harvard Macromolecular Mechanics (CHARMM)[76].

In comparing the functional forms of Equations 2.2 and 2.3 in Figure 2.6, we see that total energy functional form is dominated by the LJ term close to $\sigma_{LJ} = 1$, and as r increases it is dominated more by the longer distance Coulombic term. However, we note that the curves shown in Figure 2.6 are simplified and the exact parameterizations of σ_{LJ}, ε and q will affect the magnitude of U_{LJ} and U_C contributions to the total energy.

In addition to non-bonded constraints, these force-fields can capture bonded constants in the form of two-body bonds, three-body angles and four-body dihedrals. Bonds are often represented with a harmonic oscillator with the form:

$$U_{\text{bond}}(r) = \frac{1}{2} k_b (r - r_0)^2, \tag{2.4}$$

in which k_b is the spring constant and r_0 is the equilibrium bond length[112]. Similarly to the bonds, angles are also often harmonic in the form:

$$U_{\text{angle}}(\theta) = \frac{1}{2} k_\theta (\theta - \theta_0)^2, \tag{2.5}$$

with k_θ and θ_0 being the spring angle constant and the equilibrium angle, respectively[112].

Dihedrals can be described by multiple forms depending on how the dihedral equation is fit to the *ab initio* calculations, but typically involve a summation over a series of dihedral constants k_{d_n} and cosines of the dihedral angle φ. In this work, dihedrals are in the form of a multi-harmonic series[5]:

$$U_{\text{dihedral}}(\varphi) = \sum_{n=0}^{4} k_{d_n} cos^n \varphi, \tag{2.6}$$

the GAFF harmonic form[111]:

$$U_{\text{dihedral}}(\varphi) = \sum \frac{k_{d_n}}{2} \left[1 + cos(n\varphi - \gamma)\right], \tag{2.7}$$

in which γ is the phase, or the OPLS form of a Fourier series[60]:

$$U_{\text{dihedral}}(\varphi) = \sum_{n=1}^{4} \frac{k_{d_n}}{2}(1 - (-1)^n cos(n \cdot \varphi)), \tag{2.8}$$

and the form used is based on the parameterization provided in the literature. As such the total energy of a particle is given by the summation of all non-bonded and bonded interactions:

$$U_i = \sum_{j=i+1}^{N} U_{\text{LJ}}(r_{i,j}) + U_{\text{C}}(r_{i,j})$$
$$+ \sum U_{i_{\text{bonds}}}(r) + \sum U_{i_{\text{angles}}}(\theta) + \sum U_{i_{\text{dihedrals}}}(\varphi). \tag{2.9}$$

With the energy stemming from all other particles known, the force that is exerted on this particle is calculated with:

$$F = -\frac{dU}{dr}, \tag{2.10}$$

which is implemented numerically using the velocity-Verlet algorithm[79]:

$$v_i(t + \Delta t) = v_i(t) + \frac{F(t + \Delta t) + F(t)}{2m_i}\Delta t, \tag{2.11}$$

where Δt signifies the next step forward in time and m_i is the mass of the particle. Each particle's position can then be changed by some $v \cdot dt$, in which dt is some step

forward in time at which point the energies and forces are again calculated. This is the fundamental mechanism of MD: The iterative calculation of energy and forces time t, followed by movement of the particles based on a step forward in time dt, then recalculation of energy/forces based on new positions[34]. Repeating these steps many times allows the system to eventually relax into a free-energetically minimizing structure.

The computational cost of relaxing this system to the free-energy minimum comes in the form of calculating the forces at every timestep then updating positions based on that force. In the simple implementation of the force calculation, calculating the force between every pair of particles scales as $O(N^2)$, where N is the number of particles. Because the force and energy are decaying to 0 at longer distances, it is also possible to implement cell-lists in which the total simulation volume is partitioned into smaller sections or cells. The force calculations are then only calculated between particles in the same or neighboring cells, which reduces the computational cost to $O(N * ln(N))$[46]. The computational cost can be reduced further by reducing the number of particles in the simulation by aggregating multiple particles into a single, larger particle[36,49,66]. This process is called "coarse-graining", and depending on the degree of agglomeration, allows simulation sizes to match those of experimental thin-film thicknesses[11]. As noted earlier, increasing the degree of coarse-graining requires that the interactions are derived from more detailed simulations (such as atomistic MD) and that some important details may be abstracted away during this process. One common degree of coarse-graining is the united-atom (UA) model in which hydrogen atoms are absorbed into their parent atoms, and sizes and interactions are scaled to account for their presence[61,78]. Conducting coarse-graining in this way is desirable due to hydrogen atoms being plentiful in organic compounds, but not

contributing much to the final self-assembled morphology, allowing for studies to obtain experimentally accurate results[13,92].

In addition to coarse-graining strategies, further assumptions can be applied to the molecular models to decrease the computational complexity of the model. This can be accomplished by replacing bonded constraints with "rigid bodies" which constrains bonded interactions to be constant[89]. Another method is to exclude electrostatic interactions, which decay to zero much more slowly than the Lennard-Jones potential, requiring much larger distances to be considered for the force calculation and has been observed to be irrelevant for morphology control[50]. However, quantifying the trade-off in performance and accuracy is important, and these effects will be the focus of Chapters 3 and 4.

2.2.2   Charge Transport Simulations

There are three common methods used for calculating charge transport through OPV materials: Drift-diffusion, Master equation, and kinetic Monte Carlo (KMC). Conducting an in-depth analysis and explanation of each method is beyond the scope of this dissertation, but the reader is referred to a recent review by Groves on the subject[39]. Of the methods mentioned, drift diffusion is the most approximate but most computationally efficient. Drift diffusion functions by numerically solving the motion of charges based on an applied field and the random walk of charges, however, this comes at the cost of capturing energetic or morphological disorder within the film[40]. Master equation is able to capture this energetic and morphological disorder affect on charge transport, however, this is in the form of a morphology-averaged value. Molecular KMC (as opposed to mesoscale) is the most computationally intense

method, but provides explicit detail into the charge hopping and transport rates based on the morphology and is the focus in this dissertation.

Kinetic Monte Carlo

KMC is an event-based simulation method where each event has a time τ associated with it and this work utilizes the implementation in MorphCT[55]. The time τ is determined based on the negative natural log of a random number over the interval $[0, 1)$ x and a rate coefficient $k_{i,j}$:

$$\tau = -\frac{ln(x_r)}{k_{i,j}}.\tag{2.12}$$

In this dissertation, the only event we consider is the hop of a charge between neighboring chromophores i and j. Neighboring chromophores are identified by a Voronoi tessellation of the morphology (based off the Delaunay triangulation implementation in Scipy[53]); any chromophores with adjacent Voronoi faces are considered neighbors. The hopping rate coefficient $k_{i,j}$ is determined from an expression based on the semi-classical Marcus theory[77]:

$$k_{i,j} = \frac{|J_{i,j}|^2}{\hbar}\sqrt{\frac{\pi}{\lambda k_B T_{KMC}}}\exp\left(-\frac{r}{\alpha}\right)\exp\left(-\frac{(\Delta E_{i,j} - \lambda)^2}{4\lambda k_B T_{\text{KMC}}}\right),\tag{2.13}$$

in which $\hbar$ is Planck's reduced constant, k_B is Boltzmann's constant, and T_{KMC} is the temperatures in the KMC simulation (293 K). λ is the re-organization energy, the energy required to polarize and depolarize a chromophore in response to a carrier hopping to and from the chromophore[52]. λ is compound specific, and is based on the electronic structure calculations for a single chromophore. The electronic coupling, transfer integral $J_{i,j}$ is calculated using the energy-splitting-in-dimer method[7,25]:

$$J_{i,j} = \frac{1}{2}\sqrt{(E_{\text{HOMO}} - E_{\text{HOMO-1}})^2 - (\Delta E_{i,j})^2}, \tag{2.14}$$

in which E_{HOMO} and $E_{\text{HOMO-1}}$ are the highest and second highest HOMO energy levels, and $\Delta E_{i,j}$ is the site-to-site energy difference, calculated by $E_{\text{HOMO},i} - E_{\text{HOMO},j}$ for donor compounds. Note: for acceptor compounds it is instead $E_{\text{LUMO+1}} - E_{\text{LUMO}}$. We also include an additional exponential term in the hopping rate equation based on the center-of-mass separation between chromophores r and a tuning parameter $\alpha = 2$ Å[84,110]. This term originated from Mott's variable range hopping theory (VRH)[86], which is often used in polymer hopping theory[33,109]. VRH accounts for deficiencies in the prediction of $J_{i,j}$ within the amorphous systems, which do not sufficiently suppress $J_{i,j}$ between chromophores with large separations, leading to unphysical carrier motion. The energy levels for $\Delta E_{i,j}$ and $J_{i,j}$ are chromophore pair specific and must be calculated using QCC. Because this calculation must be done for every pair within the simulation volume, it becomes very computationally expensive. This computational expense can be reduced by relying on semi-empirical methods, which utilize experimentally derived coefficients to predict electronic orbitals, rather than computing them explicitly. In this study, we use Zerner's neglect of differential overlap (ZINDO/S) used to predict the energy levels[63,97]. And in Chapter 7 we increase the performance further by using machine learning to predict $J_{i,j}$ values.

With $k_{i,j}$ known for every possible hop within the simulation, a charge can be "placed" on a random chromophore within the simulation. The charge will then hop to an adjacent chromophore that has the shortest τ associated, with both the displacement and time recorded, with the KMC algorithm that the shortest τ corresponds to the most probable event. This process of charge hopping, recording the time and

displacement at each hop, can be repeated millions of times for the charge. The simulations are periodically bound, allowing charges to move hundreds of nanometers through the periodically repeating system comprised of ~ 10 nm unit cells. The mean squared displacement (MSD), averaged over the thousands of charges initialized on randomly selected chromophores, for different simulation lifetimes t. The gradient of the MSD as a function of t gives the carrier diffusivity, D:

$$D = \frac{1}{2n} \cdot \frac{d\text{MSD}}{dt}, \tag{2.15}$$

where $n = 3$ is the number of dimensions. D can be related to the mobility, μ, through the three-dimensional Einstein-Smoluchowski relation:

$$\mu = \frac{qD}{k_B T_{\text{KMC}}}. \tag{2.16}$$

The relation in Equation 2.15 is frequently employed in charge transport investigations[10,20,25], and is expected to provide a reasonable upper-bound for carrier diffusivity for systems with no external driving force[91]. In this dissertation, charge carrier trajectories are obtained in isolation, i.e. there are no Coulombic interactions with other charge carriers and no external electric field is applied. Therefore, this methodology is expected to produce a "best-case" zero-field carrier mobility, μ_0, that describes the diffusion of the carriers at low charge density, similar to time-of-flight experiments. As such, morphologies can be directly compared on their "ease" of charge-transport, with morphologies with the easiest charge transport likely corresponding to better devices.

Study	Method	Model	Repeat Units	Time	Size	CT
Northrup[90]	DFT	AA	4	N/A	1.8 nm*	Y
Xie[116]	DFT	AA	4	N/A	1.8 nm	N
Dag[22]	MD/DFT	AA	144*	100 ps	5 nm	N
Huang[47]	MD	AA	720	5 - 35 ns	6 nm	N
Moreno[85]	MD	AA	300	2 ns	5 nm	N
Alexiadis[3]	MD	AA	2700*	20-45 ns	8 nm	N
Bhatta[5]	MD	AA	320-1280	5 ns	9* nm	N
D'Avino[23]	MD	UA	1600	60 ns	6 nm	Y
Miller[83,84]	MD	UA	1500-15000	0.3-3 μs	15 nm	Y
Alessandri[2]	MD	5CG	11000*	6 μs	30 nm	N
Munshi[87]	MD	5CG	3600-18000	3 μs	15 nm	N
Huang[47]	MD	3CG	36864	10 ns	25 nm	N
Jankowski[51,59]	MD	3CG	2250-3750	1.7 μs	8* nm	Y
Du[27]	MD	3CG	13000*	2.5 ns	24 nm	N
Jones[56]	MD	3CG	4600-17000*	8 ns	10 nm	Y
Scherer[100]	MD	3CG	8000	80 ns	17 nm	N
Lee[70]	MD	1CG	X	10 ns	33 nm	N
Carrillo[11]	MD	1CG	3.2×10^6	400 ns	127 nm	N
Greco[37]	MC	1CG	8000-16000	N/A	25 nm	Y
Kawashima[62]	MC	Meso	N/A	N/A	150 nm	Y
Neupane[88]	MC	Meso	N/A	N/A	100 nm	Y
Lyons[57,58,74,75]	CH	Meso	N/A	N/A	130 nm	Y
Wodo[114,115]	CH	Meso	N/A	N/A	800 nm	Y
Finck[32]	CH	Meso	N/A	N/A	100 nm	Y

Table 2.1: Overview of recent computational studies of P3HT, including Method: Density Functional Theory - DFT, Molecular dynamics - MD, Monte Carlo - MC or Cahn-Hilliard - CH. Model: AA - all-atom, UA - united-atom, 5CG - coarse-grained with 5 simulation elements per repeat unit, 3CG - coarse-grained with 3 simulation elements per repeat unit, 1CG - coarse-grained with one element per repeat unit, or Meso which does not consider molecules only domain type, Repeat Units: approximate number of repeat units simulated, Time: simulation time, Size: longest simulation axis, and if the structures were used for charge-transport calculations, CT: Y-yes or N-no. *Explicit numbers were not provided in the report, but are estimated here. X-No numbers provided and insufficient information to estimate. The cyan line indicates works that are presented later in this dissertation.

2.2.3 Simulation Field

To briefly summarize how simulations have been applied to understand structure-performance relationships in OPVs we present simulation studies for the electron donor polymer Poly-(3-hexylthiophene) (P3HT). P3HT is the "fruit fly" of OPV research and is often used as a bench-mark due to its inexpensive nature, well-known morphology, and easily produced variants (although some argue expanding the set of OPV "fruit flies" to be more representative of other chemistries)[65]. Table 2.1 shows a variety of simulation studies for P3HT using a variety of different models with the works presented in Chapters 4 and 6 highlighted in cyan. In Table 2.1 we can see a series of trends that support our earlier assertion in Section 2.2.1: all-atom simulations are too small and short to simulate self-assembly and large ("Meso" Model) simulations lose molecular resolution.

In all-atom simulations, two methods are used: DFT and MD. These studies typically fall into two categories: investigating energy minimizing lattice parameters or developing force-field interactions and primarily utilize "pre-built" rather than self-assembled structures. On the other end of the size scale we have Monte Carlo and Cahn-Hilliard simulations which achieve device-scales by simulating pure domains of acceptor or donor type and then calculating the charge transfer through the systems based on parameterized rates,[62,68] thereby omitting molecular features. The only MD simulation that was able to achieve device scales was conducted by Carrillo[11]. In this study, Carrillo obtained 400 ns of simulation time, which was insufficient to reach an equilibrated structure and to our knowledge has not been utilized for charge transport calculations or simulated until equilibrium.

Excluding the mesoscale simulations, there are relatively few simulations that have

also been used to predict charge transport. For those studies that have calculated charge transport properties, the works of Northrup, D'Avino and Greco all utilized pre-built morphologies[23,37,90]. Jones was the first to self-assemble a P3HT morphology using a 3 coarse-grained site model then conducted charge transport properties for that simulation[56]. However, the morphologies produced by Jones ran for only 8 ns, which was insufficient to obtain an equilibrated structure. Jankowski was able to create equilibrated structures of P3HT using a 3 site model[51]. These equilibrated morphologies produced by Jankowski were then utilized by Jones for charge transport calculations[59]. However, in this study, the morphologies investigated consisted of four disordered structures and one perfect crystal, which is not representative of the amorphous and crystalline domains seen in experimental P3HT structures[12,103]. As such, we build on the works of Jankowski and Jones (who are large contributers to this dissertation), and present experimentally relevant self-assembled structures and directly link these structures to their charge-transport properties.

BIBLIOGRAPHY

[1] B. J. Alder and T. E. Wainright. Phase Transition for a Hard Sphere System. *Journal of Chemical Physics*, 27(1957):1208, 1957.

[2] Riccardo Alessandri, Jaakko J. Uusitalo, Alex H. De Vries, Remco W. A. Havenith, and Siewert J. Marrink. Bulk Heterojunction Morphologies with Atomistic Resolution from Coarse-Grain Solvent Evaporation Simulations. *Journal of the American Chemical Society*, 139(10):3697–3705, 2017.

[3] Orestis Alexiadis and Vlasis G. Mavrantzas. All-Atom Molecular Dynamics Simulation of Temperature Effects on the Structural, Thermodynamic, and Packing Properties of the Pure Amorphous and Pure Crystalline Phases of Regioregular P3HT. *Macromolecules*, 46(6):2450–2467, 2013.

[4] Stavros Athanasopoulos, James Kirkpatrick, Diego Martínez, Jarvist M. Frost, Clare M. Foden, Alison B. Walker, and Jenny Nelson. Predictive Study of Charge Transport in Disordered Semiconducting Polymers. *Nano Letters*, 7(6):1785–1788, 2007.

[5] Ram S. Bhatta, Yeneneh Y. Yimer, David S. Perry, and Mesfin Tsige. Improved Force Field for Molecular Modeling of Poly(3-hexylthiophene). *The Journal of Physical Chemistry B*, 117(34):10035–10045, 2013.

[6] Christoph J. Brabec, Antonio Cravino, Dieter Meissner, N. Serdar Sariciftci, Thomas Fromherz, Minze T. Rispens, Luis Sanchez, and Jan C. Hummelen. Origin of the Open Circuit Voltage of Plastic Solar Cells. *Advanced Functional Materials*, 11(5):374–380, 2001.

[7] Jean-Luc Brédas, David Beljonne, Veaceslav Coropceanu, and Jérôme Cornil. Charge-Transfer and Energy-Transfer Processes in π-Conjugated Oligomers and Polymers: A Molecular Picture. *Chemical Reviews*, 104(11):4971–5004, 2004.

[8] Jean-Luc Brédas, Joseph E Norton, Jérôme Cornil, and Veaceslav Coropceanu. Molecular Understanding of Organic Solar Cells: The Challenges. *Accounts of Chemical Research*, 42(11):1691–1699, 2009.

[9] William D. Callister and David G. Rethwisch. *Materials Science and Engineering: An Introduction*. John Wiley and Sons, 9 edition, 2013.

[10] Paola Carbone and Alessandro Troisi. Charge Diffusion in Semiconducting Polymers: Analytical Relation between Polymer Rigidity and Time Scales for Intrachain and Interchain Hopping. *The Journal of Physical Chemistry Letters*, 5(15):2637–2641, 2014.

[11] Jan-Michael Y. Carrillo, Rajeev Kumar, Monojoy Goswami, Bobby G. Sumpter, and W. Michael Brown. New Insights Into the Dynamics and Morphology of P3HT:PCBM Active Layers in Bulk Heterojunctions. *Physical Chemistry Chemical Physics*, 15(41):17873, 2013.

[12] Jui Fen Chang, Baoquan Sun, Dag W. Breiby, Martin M. Nielsen, Theis I. Sölling, Mark Giles, Iain McCulloch, and Henning Sirringhaus. Enhanced Mobility of Poly(3-hexylthiophene) Transistors by Spin-Coating from High-Boiling-Point Solvents. *Chemistry of Materials*, 16(23):4772–4776, 2004.

[13] Chunxia Chen, Praveen Depa, Victoria García Sakai, Janna K. Maranas, Jeffrey W. Lynn, Inmaculada Peral, and John R. D. Copley. A Comparison of United Atom, Explicit Atom, and Coarse-grained Simulation Models for Poly(ethylene oxide). *Journal of Chemical Physics*, 124(23):234901:1–234901:11, 2006.

[14] Hsiang-Yu Chen, Jianhui Hou, Shaoqing Zhang, Yongye Liang, Guanwen Yang, Yang Yang, Luping Yu, Yue Wu, and Gang Li. Polymer Solar Cells with Enhanced Open-Circuit Voltage and Efficiency. *Nature Photonics*, 3(11):649–653, 2009.

[15] Wei Chen, Maxim P. Nikiforov, and Seth B. Darling. Morphology Characterization in Organic and Hybrid Solar Cells. *Energy & Environmental Science*, 5(8):8045–8074, 2012.

[16] Jim Clark. Bonding in Benzene, 2000.

[17] Tracey M. Clarke and James R. Durrant. Charge Photogeneration in Organic Solar Cells. *Chemical Reviews*, 110(11):6736–6767, 2010.

[18] Kevin M. Coakley and Michael D. McGehee. Conjugated Polymer Photovoltaic Cells. *Chemistry of Materials*, 16(23):4533–4542, 2004.

[19] David C. Coffey, Obadiah G. Reid, Deanna B. Rodovsky, Glenn P. Bartholomew, and David S. Ginger. Mapping Local Photocurrents in Polymer/-Fullerene Solar Cells with Photoconductive Atomic Force Microscopy. *Nano Letters*, 7(3):738–744, 2007.

[20] Veaceslav Coropceanu, Jérôme Cornil, Demetrio A. da Silva Filho, Yoann Olivier, Robert Silbey, and Jean-Luc Brédas. Charge Transport in Organic Semiconductors. *Chemical Reviews*, 107(4):926–952, 2007.

[21] Dan Credgington, Fiona C. Jamieson, Bright Walker, Thuc-Quyen Nguyen, and James R. Durrant. Quantification of Geminate and Non-Geminate Recombination Losses Within a Solution-Processed Small-Molecule Bulk Heterojunction Solar Cell. *Advanced materials*, 24(16):2135–2141, 2012.

[22] Sefa Dag and Lin-Wang Wang. Packing Structure of Poly(3-hexylthiophene) Crystal: Ab Initio and Molecular Dynamics Studies. *The Journal of Physical Chemistry B*, 114(18):5997–6000, 2010.

[23] Gabriele D'Avino, Sébastien Mothy, Luca Muccioli, Claudio Zannoni, Linjun Wang, Jérôme Cornil, David Beljonne, and Frédéric Castet. Energetics of Electron-Hole Separation at P3HT/PCBM Heterojunctions. *Journal of Physical Chemistry C*, 117(25):12981–12990, 2013.

[24] Dean M. DeLongchamp, R. Joseph Kline, and Andrew Herzing. Nanoscale Structure Measurements for Polymer-Fullerene Photovoltaics. *Energy & Environmental Science*, 5(3):5980–5993, 2012.

[25] Wei-Qiao Deng and William A. Goddard. Predictions of Hole Mobilities in Oligoacene Organic Semiconductors from Quantum Mechanical Calculations. *The Journal of Physical Chemistry B*, 108(25):8614–8621, 2004.

[26] Letian Dou, Jingbi You, Ziruo Hong, Zheng Xu, Gang Li, Robert A. Street, and Yang Yang. 25th Anniversary Article: A Decade of Organic/Polymeric Photovoltaic Research. *Advanced Materials*, 25(46):6642–6671, 2013.

[27] Chunmiao Du, Yujin Ji, Junwei Xue, Tingjun Hou, Jianxin Tang, Shuit-Tong Lee, and Youyong Li. Morphology and Performance of Polymer Solar Cell

Characterized by DPD Simulation and Graph Theory. *Scientific Reports*, 5(October):16854, 2015.

[28] Ala'a F. Eftaiha, Jon-Paul Sun, Ian G. Hill, and Gregory C. Welch. Recent Advances of Non-fullerene, Small Molecular Acceptors for Solution Processed Bulk Heterojunction Solar Cells. *Journal of Materials Chemistry A*, 2(5):1201–1213, 2014.

[29] Energy Environ, Oleksandr V. Mikhnenko, Hamed Azimi, Markus Scharber, Mauro Morana, W. M. Blom, and Maria Antonietta. Environmental Science Exciton Diffusion Length in Narrow Bandgap Polymers. *Energy and Environmental Science*, 5:6960–6965, 2012.

[30] Nieves Espinosa, Markus Hösel, Dechan Angmo, and Frederik C. Krebs. Solar Cells with One-Day Energy Payback for the Factories of the Future. *Energy & Environmental Science*, 5(1):5117, 2012.

[31] Krishna Feron, Warwick Belcher, Christopher Fell, and Paul Dastoor. Organic Solar Cells: Understanding the Role of Förster Resonance Energy Transfer. *International Journal of Molecular Sciences*, 13(12):17019–17047, 2012.

[32] Benjamin Y. Finck and Benjamin J. Schwartz. Drift-Diffusion Studies of Compositional Morphology in Bulk Heterojunctions: The Role of the Mixed Phase in Photovoltaic Performance. *Physical Review Applied*, 6(5):1–16, 2016.

[33] Rocco P. Fornari and Alessandro Troisi. Theory of Charge Hopping Along a Disordered Polymer Chain. *Physical Chemistry Chemical Physics*, 16(21):9997, 2014.

[34] Daan Frenkel and Berend Smit. *Understanding Molecular Simulations: From Algorithms to Applications*. Elsevier, 2 edition, 2008.

[35] Richard A. Friesner. Ab Initio Quantum Chemistry : Methodology and Applications. *Chemical Theory and Compution*, 102(19):6648–6653, 2005.

[36] Thomas E. Gartner and Arthi Jayaraman. Modeling and Simulations of Polymers: A Roadmap. *Macromolecules*, 52:755–786, 2019.

[37] Cristina Greco, Anton Melnyk, Kurt Kremer, Denis Andrienko, and Kostas Ch. Daoulas. Generic Model for Lamellar Self-Assembly in Conjugated Polymers: Linking Mesoscopic Morphology and Charge Transport in P3HT. *Macromolecules*, 52:968–981, 2019.

[38] Brian A. Gregg and Mark C. Hanna. Comparing Organic to Inorganic Photovoltaic Cells: Theory, Experiment, and Simulation. *Journal of Applied Physics*, 93(3605), 2003.

[39] C. Groves. Simulating Charge Transport in Organic Semiconductors and Devices: A Review. *Reports on Progress in Physics*, 80(2):026502, 2017.

[40] C. Groves, R. A. Marsh, and N. C. Greenham. Monte Carlo Modeling of Geminate Recombination in Polymer-Polymer Photovoltaic Devices. *Journal of Chemical Physics*, 129(11):114903, 2008.

[41] Chris Groves, James C. Blakesley, and Neil C. Greenham. Effect of Charge Trapping on Geminate Recombination and Polymer Solar Cell Performance. *Nano Letters*, 10(3):1063–1069, 2010.

[42] Johannes Hachmann, Roberto Olivares-Amaya, Sule Atahan-Evrenk, Carlos Amador-Bedolla, Roel S. Sánchez-Carrera, Aryeh Gold-Parker, Leslie Vogt, Anna M. Brockway, and Alán Aspuru-Guzik. The Harvard Clean Energy Project: Large-scale Computational Screening and Design of Organic Photovoltaics on the World Community Grid. *Journal of Physical Chemistry Letters*, 2(17):2241–2251, 2011.

[43] J. J. M. Halls, C. A. Walsh, N. C. Greenham, E. A. Marseglia, R. H. Friend, S. C. Moratti, and A. B. Holmes. Efficient Photodiodes from Interpenetrating Polymer Networks. *Letters to Nature*, 376(6540):498–500, 1995.

[44] Anne Marie Helmenstine. The Octet Rule Explanation in Chemistry, 2018.

[45] Ian C. Henderson and Nigel Clarke. On Modelling Surface Directed Spinodal Decomposition. *Macromolecular Theory and Simulations*, 14(7):435–443, 2005.

[46] Michael P. Howard, Joshua A. Anderson, Arash Nikoubashman, Sharon C. Glotzer, and Athanassios Z. Panagiotopoulos. Efficient Neighbor List Calcula-

tion for Molecular Simulation of Colloidal Systems Using Graphics Processing Units. *Computer Physics Communications*, 203:45–52, 2016.

[47] David M. Huang, Roland Faller, Khanh Do, and Adam J. Moulé. Coarse-Grained Computer Simulations of Polymer/Fullerene Bulk Heterojunctions for Organic Photovoltaic Applications. *Journal of Chemical Theory and Computation*, 6(2):526–537, 2010.

[48] Hui Huang and Jinsong Huang. Organic and Hybrid Solar Cells. *Organic and Hybrid Solar Cells*, 9783319108:1–337, 2014.

[49] Helgi I. Ingólfsson, Cesar A. Lopez, Jaakko J. Uusitalo, Djurre H. de Jong, Srinivasa M. Gopal, Xavier Periole, and Siewert J. Marrink. The Power of Coarse Graining in Biomolecular Simulations. *Wiley Interdisciplinary Reviews: Computational Molecular Science*, 4(3):225–248, 2014.

[50] Nicholas E. Jackson, Kevin L. Kohlstedt, Brett M. Savoie, Monica Olvera de la Cruz, George C. Schatz, Lin X. Chen, and Mark A. Ratner. Conformational Order in Aggregates of Conjugated Polymers. *Journal of the American Chemical Society*, pages 6254–6262, 2015.

[51] Eric Jankowski, Hilary S. Marsh, and Arthi Jayaraman. Computationally Linking Molecular Features of Conjugated Polymers and Fullerene Derivatives to Bulk Heterojunction Morphology. *Macromolecules*, 46(14):5775–5785, 2013.

[52] Erik Johansson and Sven Larsson. Electronic Structure and Mechanism for Conductivity in Thiophene Oligomers and Regioregular Polymer. *Synthetic Metals*, 144(2):183–191, 2004.

[53] Eric Jones, Travis Oliphant, Pearu Peterson, and Others. SciPy: Open Source Scientific Tools for Python.

[54] J. E. Jones. On the Determination of Molecular Fields. II. From the Equation of State of a Gas. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 106(738):463–477, 1924.

[55] M. L. Jones. MorphCT, 2018.

[56] M. L. Jones, D. M. Huang, B. Chakrabarti, and Chris Groves. Relating Molecular Morphology to Charge Mobility in Semicrystalline Conjugated Polymers. *Journal of Physical Chemistry C*, 120(8):4240–4250, 2016.

[57] Matthew L. Jones, Buddhapriya Chakrabarti, and Chris Groves. Monte Carlo Simulation of Geminate Pair Recombination Dynamics in Organic Photovoltaic Devices: Multi-Exponential, Field-Dependent Kinetics and Its Interpretation. *The Journal of Physical Chemistry C*, 118(1):85–91, 2014.

[58] Matthew L. Jones, Reesha Dyer, Nigel Clarke, and Chris Groves. Are Hot Charge Transfer States the Primary Cause of Efficient Free-Charge Generation in Polymer:Fullerene Organic Photovoltaic Devices? A Kinetic Monte Carlo Study. *Physical Chemistry Chemical Physics*, 16(38):20310–20320, 2014.

[59] Matthew Lewis Jones and Eric Jankowski. Computationally Connecting Organic Photovoltaic Performance to Atomistic Arrangements and Bulk Morphology. *Molecular Simulation*, 43(10-11):1–18, 2017.

[60] William L. Jorgensen, David S. Maxwell, and Julian Tirado-Rives. Development and Testing of the OPLS All-Atom Force Field on Conformational Energetics and Properties of Organic Liquids. *Journal of the American Chemical Society*, 118(45):11225–11236, 1996.

[61] William L. Jorgensen and Julian Tirado-Rives. The OPLS [Optimized Potentials for Liquid Simulations] Potential Functions for Proteins, Energy Minimizations for Crystals of Cyclic Peptides and Crambin. *Journal of the American Chemical Society*, 110(6):1657–1666, 1988.

[62] Eisuke Kawashima, Mikiya Fujii, and Koichi Yamashita. Thermal Effect on the Morphology and Performance of Organic Photovoltaics. *Physical Chemistry Chemical Physics*, 18(38):26456–26465, 2016.

[63] James Kirkpatrick. An Approximate Method for Calculating Transfer Integrals Based on the ZINDO Hamiltonian. *International Journal of Quantum Chemistry*, 108(1):51–56, 2008.

[64] Christoph Klein, Andrew Z. Summers, Matthew W. Thompson, Justin Gilmer, Clare McCabe, Peter T. Cummings, Janos Sallai, and Christopher R. Iacovella.

Formalizing Atom-typing and the Dissemination of Force Fields with Foyer. *arXiv*, pages 1–38, 2018.

[65] Andrew T. Kleinschmidt, Samuel E. Root, and Darren J. Lipomi. Poly(3-hexylthiophene) (P3HT): Fruit fly or Outlier in Organic Solar Cell Research? *Journal of Materials Chemistry A*, 5(23):11396–11400, 2017.

[66] Sebastian Kmiecik, Dominik Gront, Michal Kolinski, Lukasz Wieteska, Aleksandra Elzbieta Dawid, and Andrzej Kolinski. Coarse-Grained Protein Models and Their Applications. *Chemical Reviews*, 116(14):7898–7936, 2016.

[67] Martin Knipper, Jürgen Parisi, Kevin Coakley, Christoph Waldauf, Christoph J. Brabec, and Vladimir Dyakonov. Impedance Spectroscopy on Polymer-Fullerene Solar Cells. *Zeitschrift für Naturforschung*, 62a:490–494, 2007.

[68] Hari K. Kodali and Baskar Ganapathysubramanian. Solar Energy Materials & Solar Cells Sensitivity Analysis of Current Generation in Organic Solar Cells Comparing Bilayer, Sawtooth, and Bulk Heterojunction Morphologies. *Solar Energy Materials and Solar Cells*, 111:66–73, 2013.

[69] David N. LeBard, Benjamin G. Levine, Philipp Mertmann, Stephen A. Barr, Arben Jusufi, Samantha Sanders, Michael L. Klein, Athanassios Z. Panagiotopoulos, A. Barr, Arben Jusufi, Samantha Sanders, and L. Klein. Self-Assembly of Coarse-Grained Ionic Surfactants Accelerated by Graphics Processing Units. *Soft Matter*, 8(8):2385–2397, 2012.

[70] Cheng-Kuang Lee, Chun-wei Pao, and Chih-Wei Chu. Multiscale Molecular Simulations of the Nanoscale Morphologies of P3HT:PCBM Blends for Bulk Heterojunction Organic Photovoltaic Cells. *Energy & Environmental Science*, 4(10):4124–4132, 2011.

[71] Yongfang Li. Molecular Design of Photovoltaic Materials for Polymer Solar Cells: Toward Suitable Electronic Energy Levels and Broad Absorption. *Accounts of Chemical Research*, 45(5):723–733, 2012.

[72] LibreTexts. Sigma and Pi Bonds - Chemistry, 2016.

[73] Richard R. Lunt, Noel C. Giebink, Anna A. Belak, Jay B. Benziger, and Stephen R. Forrest. Exciton Diffusion Lengths of Organic Semiconductor Thin Films Measured by Spectrally Resolved Photoluminescence Quenching. *Journal of Applied Physics*, 105(5), 2009.

[74] Benjamin P. Lyons, Nigel Clarke, and Chris Groves. The Quantitative Effect of Surface Wetting Layers on the Performance of Organic Bulk Heterojunction Photovoltaic Devices. *The Journal of Physical Chemistry C*, 115(45):22572–22577, 2011.

[75] Benjamin P. Lyons, Nigel Clarke, and Chris Groves. The Relative Importance of Domain Size, Domain Purity and Domain Interfaces to the Performance of Bulk-Heterojunction Organic Photovoltaics. *Energy & Environmental Science*, 5(6):7657, 2012.

[76] Alexander D. MacKerell and Nilesh K. Banavali. All-Atom Empirical Force Field for Nucleic Acids: II. Application to Molecular Dynamics Simulations of DNA and RNA in Solution. *Journal of Computational Chemistry*, 21(2):105–120, 2000.

[77] R. A. Marcus. On the Theory of Oxidation-Reduction Reactions Involving Electron Transfer. I. *The Journal of Chemical Physics*, 24(5):966, 1956.

[78] Marcus G. Martin and J. Ilja Siepmann. Transferable Potentials for Phase Equilibria. 1. United-Atom Description of n-Alkanes. *Journal of Physical Chemistry B*, 102(14):2569–2577, 1998.

[79] Glenn J. Martyna, Douglas J. Tobias, and Michael L. Klein. Constant pressure molecular dynamics algorithms. *The Journal of Chemical Physics*, 101(5):4177–4189, 1994.

[80] Katherine A. Mazzio and Christine K Luscombe. The Future of Organic Photovoltaics. *Chemical Society Reviews*, 44(1):78–90, 2015.

[81] Lingxian Meng, Yamin Zhang, Xiangjian Wan, Chenxi Li, Xin Zhang, Yanbo Wang, Xin Ke, Zuo Xiao, Liming Ding, Ruoxi Xia, Hin-Lap Yip, Yong Cao, and Yongsheng Chen. Organic and Solution-Processed Tandem Solar Cells With 17.3% Efficiency. *Science*, 2612:1094–1098, 2018.

[82] V. D. Mihailetchi, L. J. A. Koster, J. C. Hummelen, and P. W. M. Blom. Photocurrent Generation in Polymer-Fullerene Bulk Heterojunctions. *Physical Review Letters*, 93(21):216601, 2004.

[83] Evan Miller, Matthew Jones, Michael Henry, Paul Chery, Kyle Miller, and Eric Jankowski. Optimization and Validation of Efficient Models for Predicting Polythiophene Self-Assembly. *Polymers*, 10(12):1305, 2018.

[84] Evan Miller, Matthew Jones, and Eric Jankowski. Tying Together Multiscale Calculations for Charge Transport in P3HT: Structural Descriptors, Morphology, and Tie-Chains. *Polymers*, 10(12):1358, 2018.

[85] Margherita Moreno, Mosè Casalegno, Guido Raos, Stefano V. Meille, and Riccardo Po. Molecular Modeling of Crystalline Alkylthiophene Oligomers and Polymers. *The Journal of Physical Chemistry. B*, 114(4):1591–1602, 2010.

[86] N. F. Mott. Conduction in Glasses Containing Transition Metal Ions. *Journal of Non-Crystalline Solids*, 1(1):1–17, 1968.

[87] Joydeep Munshi, Rabindra Dulal, Teyu Chien, Wei Chen, and Ganesh Balasubramanian. Solution Processing Dependent Bulk Heterojunction Nanomorphology of P3HT/PCBM Thin Films. *ACS Applied Materials and Interfaces*, 11(18):17056–17067, 2019.

[88] Upendra Neupane, Behzad Bahrami, Matt Biesecker, Mahdi Farrokh Baroughi, and Qiquan Qiao. Kinetic Monte Carlo modeling on organic solar cells: Domain size, donor-acceptor ratio and thickness. *Nano Energy*, 35(November 2016):128–137, 2017.

[89] Trung Dac Nguyen, Carolyn L. Phillips, Joshua A. Anderson, and Sharon C. Glotzer. Rigid Body Constraints Realized in Massively-parallel Molecular Dynamics on Graphics Processing Units. *Computer Physics Communications*, 182(11):2307–2313, 2011.

[90] John E. Northrup. Atomic and Electronic Structure of Polymer Organic Semiconductors: P3HT, PQT, and PBTTT. *Physical Review B*, 76(24):245202, 2007.

[91] S. V. Novikov. Diffusion of a Particle in the Gaussian Random-Energy Landscape: Einstein Relation and Analytical Properties of Average Velocity and Diffusivity as Functions of Driving Force. *Physical Review E*, 98(1):012128, 2018.

[92] Wolfgang Paul, Do Y. Yoon, and Grant D. Smith. An Optimized United Atom Model for Simulations of Polymethylene Melts. *Journal of Chemical Physics*, 103(4):1702–1709, 1995.

[93] J. Peet, J. Y. Kim, N. E. Coates, W. L. Ma, D. Moses, A. J. Heeger, and G. C. Bazan. Efficiency Enhancement in Low-Bandgap Polymer Solar Cells by Processing with Alkane Dithiols. *Nature Materials*, 6(7):497–500, 2007.

[94] P. Peumans and S. R. Forrest. Very-High-Efficiency Double-Heterostructure Copper Phthaloyanine/C60 Photovoltaic Cells. *Applied Physics Letters*, 79(1):126–128, 2001.

[95] Christopher M. Proctor, John A. Love, and Thuc-Quyen Nguyen. Mobility Guidelines for High Fill Factor Solution-Processed Small Molecule Solar Cells. *Advanced Materials*, 26(34):5957–5961, 2014.

[96] A. Rahman. Correlations in the Motion of Atoms in Liquid Argon. *Physical Review A*, 136:A405–A411, 1964.

[97] J. Ridley and Michael Zerner. An Intermediate Neglect of Differential Overlap Technique for Spectroscopy: Pyrrole and the Azines. *Theoretica Chimica Acta*, 32(2):111–134, 1973.

[98] Sven Rühle. Tabulated Values of the Shockley-Queisser Limit for Single Junction Solar Cells. *Solar Energy*, 130:139–147, 2016.

[99] R.T. Sanderson. Carbon-Carbon Bond Lengths. *Science*, 116(3002):41–42, 1952.

[100] C. Scherer and D. Andrienko. Comparison of Systematic Coarse-Graining Strategies for Soluble Conjugated Polymers. *European Physical Journal: Special Topics*, 225(8-9):1441–1461, 2016.

[101] Sean E. Shaheen, David S Ginley, and Ghassan E Jabbour. Organic-Based Photovoltaics: Toward Low-Cost Power Generation. *MRS Bulletin*, 30(01):10–19, 2005.

[102] C. G. Shuttle, B. O'Regan, A. M. Ballantyne, J. Nelson, D. D. C. Bradley, and J. R. Durrant. Bimolecular Recombination Losses in Polythiophene: Fullerene Solar Cells. *Physical Review B*, 78(11):113201, 2008.

[103] H. Sirringhaus, P. J. Brown, R. H. Friend, M. M. Nielsen, K. Bechgaard, B. M. W. Langeveld-Voss, a. J. H. Spiering, R. A. J. Janssen, E. W. Meijer, P. Herwig, and D. M. de Leeuw. Two-Dimensional Charge Transport in Self-Organized, High-Mobility Conjugated Polymers. *Nature*, 401(6754):685–688, 1999.

[104] Roar R. Søndergaard, Markus Hösel, and Frederik C. Krebs. Roll-to-Roll Fabrication of Large Area Functional Organic Materials. *Journal of Polymer Science, Part B: Polymer Physics*, 51(1):16–34, 2013.

[105] C. W. Tang. Two-Layer Organic Photovoltaic Cell. *Applied Physics Letters*, 48(2):183–185, 1986.

[106] Zheng Tang, Wolfgang Tress, and Olle Ingana. Light Trapping in Thin Film Organic Solar Cells. *Materials Today*, 17(8):389–396, 2014.

[107] Brian K. Tanner. Electrical Conduction in Semiconductors and Insulators. In *Introduction to the Physics of Electrons in Solids*, chapter 6, pages 111–138. Cambridge University Press, 1 edition, 1995.

[108] Brian K. Tanner. Semiconductor Devices. In *Introduction to the Physics of Electrons in Solids*, chapter 7, pages 138–166. Cambridge University Press, 1 edition, 1995.

[109] D. N. Tsigankov and A. L. Efros. Variable Range Hopping in Two-Dimensional Systems of Interacting Electrons. *Physical Review Letters*, 88(17):176602, 2002.

[110] Ellen Van, Matthew Jones, Eric Jankowski, and Olga Wodo. Using Graphs to Quantify Energetic and Structural Order in Semicrystalline Oligothiophene Thin Films. *Molecular Systems Design & Engineering*, 1:273–277, 2018.

[111] Junmei Wang, Romain M. Wolf, James W. Caldwell, Peter A. Kollman, and David A. Case. Development and Testing of a General Amber Force Field. *Journal of Computational Chemistry*, 25(9):1157–1174, 2004.

[112] Scott J. Weiner, Peter A. Kollman, David A. Case, U. Chandra Singh, Caterina Ghio, Giuliano Alagona, Salvatore Profeta, and Paul Weinerl. A New Force Field for Molecular Mechanical Simulation of Nucleic Acids and Proteins. *Journal of the American Chemical Society*, 106(17):765–784, 1984.

[113] Martin Weis, Jack Lin, Dai Taguchi, Takaaki Manaka, and Mitsumasa Iwamoto. Analysis of Transient Currents in Organic Field Effect Transistor: The Time-of-Flight Method. *Journal of Physical Chemistry C*, 113(43):18459–18461, 2009.

[114] Olga Wodo and Baskar Ganapathysubramanian. Modeling Morphology Evolution During Solvent-Based Fabrication of Organic Solar Cells. *Computational Materials Science*, 55:113–126, 2012.

[115] Olga Wodo, Srikanta Tirthapura, Sumit Chaudhary, and Baskar Ganapathysubramanian. Computational Characterization of Bulk Heterojunction Nanomorphology. *Journal of Applied Physics*, 112(6):064316, 2012.

[116] Weiyu Xie, Y. Y. Sun, S. B. Zhang, and John E. Northrup. Structure and Sources of Disorder in Poly(3-hexylthiophene) Crystals Investigated by Density Functional Calculations with van der Waals Interactions. *Physical Review B - Condensed Matter and Materials Physics*, 83(18):1–6, 2011.

[117] Liang Xu, Jian Wang, and Julia W. P. Hsu. Structural Order: The Dominant Factor for Nongeminate Recombination in Organic Photovoltaic Devices. *Journal of Physical Chemistry C*, 121(17):9242–9248, 2017.

[118] Shujiang Yang, Pavel Olishevski, and Miklos Kertesz. Bandgap Calculations For Conjugated Polymers. *Synthetic Metals*, 141(1-2):171–177, 2004.

[119] G. Yu, J. Gao, J. C. Hummelen, F. Wudl, and Alan J. Heeger. Polymer Photovoltaic Cells: Enhanced Efficiencies via a Network of Internal Donor-Acceptor Heterojunctions. *Science*, 270(5243):1789–1791, 1995.

[120] Jun Yuan, Yunqiang Zhang, Liuyang Zhou, Guichuan Zhang, Hin-Lap Yip, Tsz-Ki Lau, Xinhui Lu, Can Zhu, Hongjian Peng, Paul A. Johnson, Mario Leclerc, Yong Cao, Jacek Ulanski, Yongfang Li, and Yingping Zou. Single-Junction Organic Solar Cell with over 15% Efficiency Using Fused-Ring Acceptor with Electron-Deficient Core. *Joule*, pages 1–12, 2019.

CHAPTER 3

ENHANCED COMPUTATIONAL SAMPLING OF PERYLENE AND PERYLOTHIOPHENE PACKING WITH RIGID-BODY MODELS[1]

3.1 Introduction

Polycyclic aromatic hydrocarbons (PAHs) are an important class of materials, not least because they can self-assemble into structures that exhibit long-range order[7]. This spontaneous ordering can be beneficial, as in the case of organic electronic devices (particularly plastic solar cells), where ordering maximizes the inter-molecular charge transport between molecules, resulting in high carrier mobilities and device efficiencies[3,6,16,37,44,62]. Spontaneous aggregation of PAHs is detrimental in oil extraction operations, where asphaltene self-assembly leads to the rapid precipitation of tar-like aggregates that foul heat exchangers and limit crude oil processing[18,24,40]. In order to control PAH aggregation, whether for new organic electronic device design or to prevent fouling in oil extraction, the we require a fundamental understanding of PAH self-assembly physics.

As mentioned in Section 2.2, Elucidating links between PAH chemistry and self-

---

assembled morphologies has been limited by the difficulties associated with measuring PAH structure[13,25]. Computational models of molecular self-assembly are not subject to the same space, time, and material cost constraints of wet labs and therefore permit broader explorations of large parameter spaces. The practical limitations to performing MD simulations of every system of interest are the computational costs of relaxing a system to equilibrium and the subsequent sampling of the equilibrium ensemble of states. Consequently, it is essential that the molecular models, computational algorithms, and computational hardware used to perform MD simulations are optimized to minimize computational cost. Minimizing computational cost is especially relevant for high-throughput, "Big Data" screening studies wherein thousands of simulations are performed and a 10% improvement in efficiency can save weeks of computing[8,20].

One way longer simulation times and larger systems can be accesed is by applying coarse-graining strategies, but this comes at the cost of resolution that has been abstracted away. One commonly used coarse-graining scheme is the UA model, where an explicit consideration of hydrogens is omitted[30,36]. UA models have been used in simulations to accurately predict the behavior of various systems, including polymers[5], proteins[30] and other hydrocarbons[38,43,54,60], demonstrating no significant structural differences between the UA and all-atom simulations, or experimental data. In particular, UA investigations have shown good agreement with the structural and thermodynamic properties predicted by experiment for a variety of polycyclic aromatic hydrocarbons[11].

Another approximation technique arises from assuming that the *intra-molecular* positional fluctuations of atoms *within* certain small molecules or functional groups do not significantly affect the long-range self-assembly characteristics. As such, these groups of atoms can be assumed to belong to a rigid body, where atoms

are fixed in space with respect to their bonded neighbors. This can dramatically reduce the computational overhead of determining all of the bond, angle and dihedral constraints within the system, permitting significantly longer length and timescales to be accessed[41]. Rigid bodies are frequently implemented in systems that investigate the morphology of nano-particle systems, using amphiphilic block co-polymers[23,33] or DNA[31,32] to guide their self-assembly, and are shown to accurately describe the experimentally observed structures. Studies of more general "metaparticles" have shown that various structural features that are relevant to organic materials can be generated using rigid body simulations[19]. Rigid models have also been used to simulate aromatic hydrocarbons, showing excellent agreement with experiment for a wide range of structural properties and particle dynamics[4]. Despite the potential for rigid models to lower the computational cost of simulating planar, polycyclic aromatic hydrocarbons, systematic comparisons between flexible and rigid models used across broad sets of state points and molecular chemistries are lacking.

To put the advances of the present work in clearer context, we briefly summarize recent computational work focusing on polyaromatic molecules at the center of this work. Asphaltenes have polyaromatic cores similar to the molecules studied here, and require lengthy simulations of aggregate structure at many thermodynamic state points in order to understand their assembly physics. It is common to see simulations using the LINCS algorithm to rigidly constrain bond lengths to help access longer simulation times[49,51,58], but there are not yet examples of *rigid bodies* described above to simplify simulations of polyaromatic cores. The longest simulations, with trajectories of ~ 2000 ns, observe the assembly of a dozen asphaltenes in water and methane under four different conditions[63]. Atomistic simulations of 400 asphaltene molecules in vacuum, access only 6 ns of simulation time[34]. Indeed, simulations that

include atomistic representations or explicit solvents generally access fewer than 150 ns for fewer than 30 molecules[10,49,51,59,61]. Wang and Ferguson use coarse-grained models and GPU acceleration to access 500 ns trajectories for 200 asphaltenes at 27 state points, demonstrating the sampling advantages of simplified models and hardware accelerators[58]. The work presented here extends the ideas of Wang and Ferguson, demonstrating that rigid body models, combined with GPU accelerators can be used to efficiently sample *thousands* of thermodynamic state points, by equilibrating systems of 200 molecules for over 100 ns each. Prior molecular simulation work on perylene has investigated pre-assembled crystal structures[57], whereas here we observe for the first time the spontaneous self-assembly of structures large enough and with sufficient order to be directly compared against experimental GIXS data.

In order to apply high-throughput simulation techniques to study the self-assembly of rigid organic molecules, we require an understanding of which models and methods are necessary and sufficient for predicting experimentally observed structures. Therefore, in this work, UA MD simulations are performed to characterize rigid and flexible models of two types of small molecules.

In this work, the polycyclic aromatic hydrocarbons perylene and perylo[1,12-bcd]thiophene (perylothiophene) are considered as test cases, due to both their applicability to organic electronics[3] and similarity to the cores of complex asphaltene molecules[50]. A "flexible" model is considered, in which the topological connectivities and geometries are enforced by bond, angle and dihedral constraints, and compare to the computational performance and resultant packing morphologies when the molecules are described as rigid bodies. The rigid model considerably improves computational efficiency by permitting a larger quantity of timesteps to be accessed per second, without significantly increasing the relaxation or autocorrelation

timescales, or affecting the obtained morphologies. In the Methods section we describe the simulation methodology, along with details of the molecular topologies and the analysis metrics we employ. In the Results section we discuss the thermodynamically stable morphologies predicted with the two molecules and two models, demonstrating that experimentally observed phase transitions are reproduced. Finally, we conclude by considering the quantitative implications of modeling organic molecules as rigid bodies, comparing our predictions with experimental measurements, and highlighting opportunities for future work.

3.2   Methods

MD simulations are performed to determine the thermodynamically stable morphologies of perylene and perylothiophene as a function of temperature and packing fraction for both rigid and flexible models. Equilibration and sampling metrics are computed from time series of potential energy and structural metrics. Structural metrics are computed from the resulting morphologies to distinguish "vapor", "droplet", "ordered", and "eclipsed" morphologies, which are summarized in phase diagrams.

3.2.1   Molecular Dynamics Simulations

MD simulations are performed using the HOOMD-blue simulation suite in order to predict the thermodynamically stable morphologies for perylene and perylothiophene (Figure 3.1a and b respectively)[1,17]. These are performed in the canonical ensemble (constant number of molecules N, volume V, and temperature T), using a Nosé-Hoover thermostat to control the temperature[26]. Particle positions and velocities are updated with two-step velocity-Verlet integration of Newton's equations

of motion with a time step of dt = 0.001 dimensionless time units[52]. We run simulations of systems containing 200 to 1500 molecules with periodic boundary conditions until equilibrium is reached based on stabilization and autocorrelation of observables including potential energy and structural crystallinity order parameter (Appendix B-Section 1). Each MD simulation is initialized with a random, unique initial configureation except those at the lowest temperatures, which are initialized from the final configureation of the next-highest temperature run with the same density. After simulations have reached equilibrium they are continued until at least 20 statistically independent snapshots are obtained, requiring on average 1.05×10^7 timesteps in total.

The molecular topologies (atomic positions and bond connectivities) for perylene are constructed using Avogadro-1.1.1 (Figure 3.1c)[22]. The base units of length $\sigma_{LJ} = 3.8$ Å, mass $M = 12.011$ amu, and energy $\varepsilon = 121.7$ cal/mol (0.51 kJ/mol) are taken from the OPLS-UA force field, where σ_{LJ} is the van der Waals radius and ε is the depth of the Lennard-Jones potential, modeling the pair-wise interactions between aromatic carbons[30]. The complete bond, angle, and dihedral force-field parameters are tabulated in Appendix B-Section 2, alongside a description of the computational infrastructure used in this investigation. The rigid body constraints are implemented as described in Nguyen et al[41]. Perylene simulations have time scale

$$\tau_s = \sqrt{\frac{M\sigma_{LJ}^2}{\varepsilon}} = 1.84 \times 10^{-12} \text{ s,} \tag{3.1}$$

and consequently, dt = 1.84 fs. The extensive work of Tsuzuki et al has previously emphasized the importance of electrostatic interactions in the packing of aromatic molecules containing a thiophene moiety, however, electrostatic interations between

the partial atomic charges of simulation elements are omitted in this investigation [55]. This is in the interest of computational efficiency, as preliminary results have shown that simulations run 3-4× more slowly when electrostatic interactions are included, as described in Appendix B-Section 3. In this work, we are primarily concerned with finding the optimal compromise between accurately generating morphologies and maximizing the number of computed simulation timesteps per second, and so we do not consider partially charged atoms on the molecules. To ascertain the effect of heteroatoms on molecular packing, a sulfur atom is included in the coarse-grained perylene model to simulate perylothiophene (Figure 3.1d), which is selected due to its viability in organic photovoltaic devices [3] and also due the importance of heteroatoms in asphaltene molecules [50].
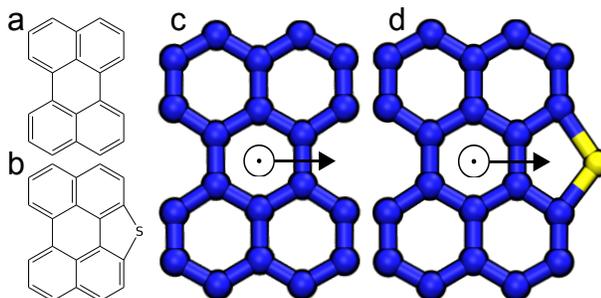
Figure 3.1: Molecular structures of (a) perylene and (b) perylothiophene. United Atom topologies for (c) perylene and (d) perylothiophene. Blue spheres represent C-H beads, yellow spheres represent S, and blue/yellow cylinders represent bonds of the respective bead type. Arrows are drawn to indicate the vectors that was used to describe the orientation of the molecules, both in-plane and normal to the plane of the molecules.

The Lennard-Jones parameters $\varepsilon_S = 358$ cal/mol (1.5 kJ/mol) and $\sigma_{LJ,S} = 3.5$ Å for sulfur-sulfur interactions are informed by OPLS-UA and TraPPE-UA force fields that show $\varepsilon_S \approx 3\varepsilon_C$ [30,35]. The mass of the sulfur atom is $M_S = 32.06$ amu. Simulations of both perylene and perylothiophene are conducted for a range

of: $61.25 \leq T \leq 1837$ K. A range of desities are considered between 0.01 g/cm^3 and 1.7 g/cm^3, which includes experimentally observed solid densities of 1.286 and 1.483 g/cm^3 for perylene and perylothiophene respectively[9,47]. Appendix B-Section 4 summarizes the range of temperatures and densities studied in this work along with temperature and density steps.

3.2.2 Structural Analysis

Molecular packing is quantified by four metrics: the center-of-mass radial distribution function $g(r)$, orientational correlations between neighboring molecules ξ, a degree-of-crystallinity metric ψ, and with simulated grazing-incidence X-Ray scattering (GIXS) characterization. The metrics $g(r)$, ξ, and ψ are used in concert to distinguish the "vapor", "droplet", "ordered", and "eclipsed" phases. Simulated GIXS patterns are used to identify periodic length scales, structural symmetries, and to compare simulated morphologies against experimentally synthesized materials. These GIXS patterns are computed using numerical Fourier techniques described in Refs [29] and [48].

Calculating $g(r)$ (using the MDTraj implementation[39]) between molecular centers of mass provides data to justify molecular clustering criteria. Two key length scales are determined from $g(r)$ of ordered morphologies, $R_{in,cut}$ and $R_{out,cut}$, which are used to distinguish molecules that are nearest neighbors in a "stack", and molecules that are in neighboring stacks respectively. The in-stack cutoff distance $R_{in,cut} = 5.0$ Å is chosen as a clustering criterion based on the minimum following the first-neighbor peak in $g(r)$ located at $r = 3.5$ Å (Figure 3.2a). The neighboring-stack cutoff distance $R_{out,cut} = 11.6$ Å is determined by the geometric average

$$R_{out,cut} = \sqrt{r_c^2 + r_\pi^2}, \tag{3.2}$$

of the first-neighbor peak ($r_\pi = 5.0$ Å) and the stack-center to stack-center distance ($r_c = 10.5$ Å) determined from the third-neighbor peak of $g(r)$ (Figure 3.2a).
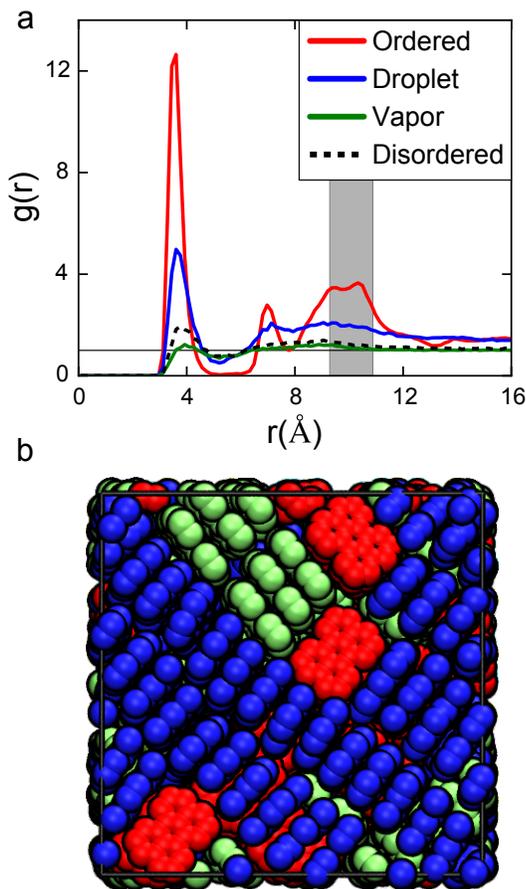
Figure 3.2: (a) Radial distribution functions for different temperature runs with a solid line at $g(r) = 1$ as reference for a completely random system. $T_{\text{Vapor}} > T_{\text{Disordered}} > T_{\text{Droplet}} > T_{\text{Ordered}}$. Temperature decreases give rise to more distinct $g(r)$s from that of a random system. The dashed black line describes a disordered system that lies between the droplet and vapor phases. Additional analysis is needed to distinguish the phase of this system, see text for details. (b) Clusters colored blue, green and red based on molecule distances and orientations, for a representative ordered morphology of rigid perylene, performed at $T = 1290$ K and $\rho = 1.77$ g/cm^3.

The orientational order parameter ξ is calculated for a simulation snapshot by averaging the local orientational order

$$O_{i,j} = \vec{o_i} \cdot \vec{o_j} \qquad (3.3)$$

over all pairs of neighboring molecules, where o_i and o_j are the unit orientation vectors for molecules i and j whose COM separation is less than $R_{in,cut}$[57]. The in-plane orientation vector o for perylene is defined to be orthogonal to the bonds connecting naphthalene rings and is depicted by a black arrow in 3.1c. In perylothiophene, the direction of orientation vector o is defined from the center-of-mass of the perylene core to the sulfur atom, as shown in 3.1d. A system is considered "eclipsed" when $\xi \geq 0.9$. This cutoff is chosen based on analysis of ξ dependence with T at the densities studied here, indicating that 0.9 is a sensitive cutoff distinguishing molecular stacks that are in and out of register (Appendix B-Section 5).

The crystallinity order parameter ψ is calculated for a simulation snapshot by the fraction of molecules in a large "clusters". Here, two molecules i and j are considered part of the same cluster when $r_{i,j} \leq R_{out,cut}$ and when $\theta_{i,j} < \theta_{cut}$. The alignment of two molecules $\theta_{i,j}$ is calculated from the dot product of their out-of-plane unit vectors (Figure 3.1c,d). The θ_{cut} was determined from the distribution of the dot-products between the normal vectors of molecules (Appendix B-Section 6). Clusters comprising 30% or more of the system (60+ molecules) are considered "large". This metric is chosen to distinguish "ordered" morphologies containing a few large clusters, whereas "droplet" and "vapor" configureations contain many small clusters. 3.2b shows a configureation with three clusters (colored blue, green, and red) identified using this clustering protocol. The "large cluster" cutoff of 30% is chosen to balance

overcounting of small clusters against undercounting large clusters, based on visual inspection using Visual Molecular Dynamics[27].

The gaseous "vapor" phase distinguished here is characterized by low molecular association (no defined peaks in the $g(r)$), low oriential correlations $\xi \leq 0.9$, and low crystallinity $\psi \leq 0.5$ as expected at low densies and high temperatures. The liquid "droplet" phase is distinguished by significant aggregation (first peak is only defined peak in $g(r)$), low orientational correlations $\xi \leq 0.9$, and low crystallinity $\psi \leq 0.5$. At high temperatures and pressures, these liquid and gaseous phases become indistinguishable and so are characterized instead as a "fluid" phase. The solid "ordered" phase is distinguished by clearly defined first-, second- and third-nearest neighbor peaks in $g(r)$ (Figure 3.2a), low orientational correlations $\xi < 0.9$, and high crystallinity $\psi > 0.5$. In this work, ordered phases are characterized by hexagonally-packed columns of molecules (Figure 3.3) demonstrating significant π-stacking (Figure 3.2b). The "eclipsed" phase shares the $g(r)$ peak structure of the ordered phase, but has high orientational correlations $\xi \geq 0.9$, and high crystallinity $\psi \geq 0.5$. As in the ordered phase, π-stacking is prevalent in the eclipsed phase, but differs in that molecules in a stack share the same orientation.

At low densities (≤ 0.65 g/cm^3) we consider an additional criterion to distinguish vapor and droplet phases. Here, average $g(r)$ values in the range of $r = 9.3$ to 10.9 Å (gray highlight in 3.1a) are compared to those present in the highest temperature, most dispersed system. This range is chosen to allow long-range order comparisons between simulations, without taking into account the short-range π-stacking present in all samples. If the average $g(9.3 < r < 10.9)$ for the lower temperature sample is within 10% of that of highest temperature, most dispersed trial, then the system was considered a vapor. An example of this characterization methodology is shown
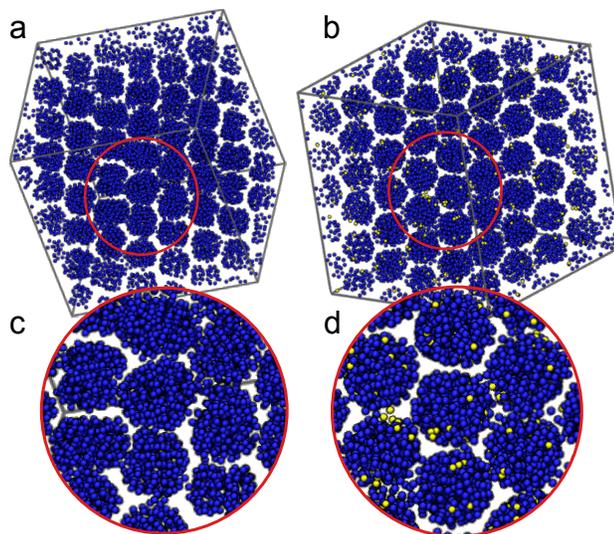
Figure 3.3: Hexagonally packed columns extending into the page, characteristic of ordered morphologies of (a) perylene and (b) perylothiophene with zoomed in regions (c) and (d). Carbon atoms are represented with blue spheres and sulfur with yellow.

in 3.2a. Here, for $9.3 \leq r < 10.9$ Å, the 1163 K system (black dashed line) is within 10% of the 1837 K system (green line) representing the vapor system, and so the 1163 K system is also considered a vapor. At lower temperatures, the 918 K system (blue line) is shown to be more than 10% of the $g(r)$ of the vapor and so this morphology more closely represents a droplet. As the temperature decreases further (red line, 612 K), the morphology is significantly more ordered than the vapor case, and the $g(r)$ peaks become even more defined. Thus, $g(r)$ is used to determine the phase transition temperatures from strongly ordered to droplet to vapor.

Using these criteria to distinguish the structures resulting from MD simulations, phase diagrams are constructed for flexible and rigid perylene, and for flexible and rigid perylothiophene. Note that the transition temperatures indicated on phase diagrams presented here are approximations whose utility is derived from their low cost to generate. For more accurate phase coexistence curves, Gibbs ensemble simulation

techniques are more appropriate[42].

3.3 Results

Initial simulation results demonstrate no differences in structural metrics ψ, ξ, and $g(r)$ for systems of 200 to 1500 molecules (Appendix B-Section 7). Due to the significantly faster relaxation times and simulation times, we therefore focus on systems of 200 molecules, and report the results in this section. Systems of 200 molecules require on average 20 ns of simulation time to reach equilibrium, followed by on average 115 ns of sampling time. Potential energy and ψ autocorrelation times are calculated to ensure at least 20 uncorrelated snapshots are generated during the sampling period. If fewer than 20 independent snapshots are generated, simulations are run longer until this criterion is satisfied.

3.3.1 Phase Diagrams

Using $g(r)$, ξ, and ψ to differentiate morphologies resulting from our MD simulations we summarize how assembled morphologies of perylene and perylothiophene depend on temperature, density, and model flexibility in four phase diagrams (Figure 3.4). Each phase diagram is divided into five regions: "vapor" (green), "fluid" (cyan), "droplet" (blue), "ordered" (red), and "eclipsed" (magenta). At high temperatures ($T > 1700$ K), vapor phases are observed at low densities and fluid phases at high densities. All four phase diagrams show a phase boundary between fluid phases and ordered phases at higher densities ($\rho > 0.65$) g/cm^3, though the disorder-order transition temperatures T_{DO} vary. Another common feature is the phase boundary between the ordered and eclipsed phases at lower temperatures across the full range

of densities studied. Finally, all four phase diagrams indicate regions of phase space with droplets located between the ordered and vapor phases at low densities ($\rho \leq$ 0.65) g/cm^3. At densities greater than 0.65 g/cm^3 the droplet and vapor phases are indistinguishable.
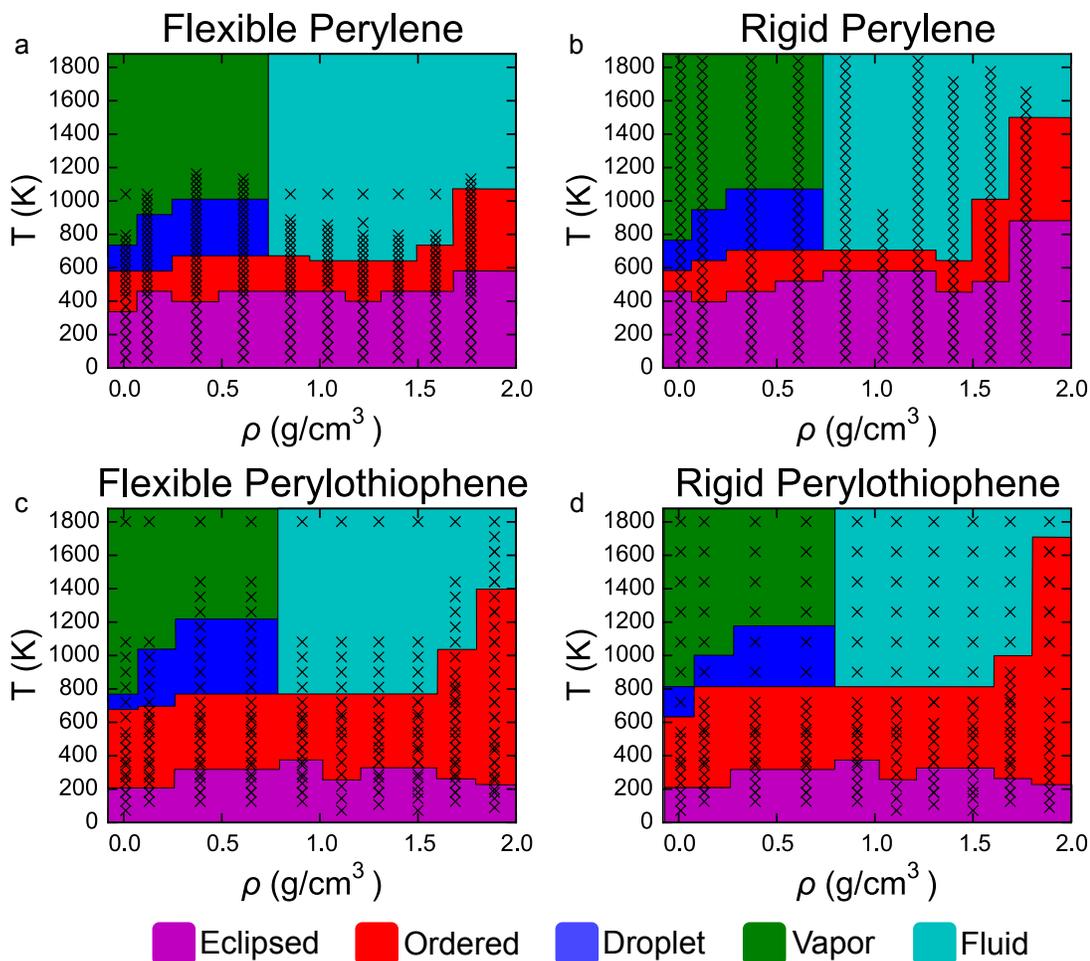
Figure 3.4: Phase diagrams of (a) flexible perylene, (b) rigid perylene, (c) flexible perylothiophene, and (d) rigid perylothiophene. Regions are colored to describe phases of decreasing order as explained in the text, where magenta describes eclipsed columns, red describes ordered π-stacked columns, blue describes droplets, green describes vapors, and cyan describes a disordered fluid phase. Each black "x" represents a state point calculated in this investigation.

We find that the phase diagrams for rigid models of perylene and perylothiophene

are similar to their flexible counterparts. For perylene, the phase boundaries between the droplet phase and the vapor, fluid, and ordered phases are nearly identical with both models. The phase boundary between fluid and ordered phases are also similar between rigid and flexible perylene, though we observe significantly higher T_{DO} for the rigid model at high densities ($\rho > 1.5$ g/cm^3). The transition temperatures between ordered perylene and eclipsed perylene also shift higher slightly for the rigid model relative to the flexible model (and more so at $\rho = 1.77$ g/cm^3), but generally are within 90 K of each other. For perylothiophene the rigid and flexible phase diagrams are nearly identical. As with perylene, there is a shift up in T_{DO} observed at high densities. At low densities we observe discrepancies between the rigid and flexible phase boundaries between the droplet, vapor, and ordered phases.

The phase diagrams for perylene and perylothiophene are qualitatively similar in that they each contain the same five phases arranged similarly. One difference is that the phase boundaries between ordered and fluid/vapor phases of perylothiophene occur at higher temperatures than for perylene. Additionally, the phase boundary between the ordered and eclipsed phases for perylothiophene is at lower temperatures than observed for perylene. This demonstrates that the addition of a single heteroatom (the difference between perylothiophene and perylene) is enough to significantly broaden the thermodynamic stability of an ordered morphology (here, hexagonally-packed cylinders) over a wide region of state space.

Overall we find the rigid models of perylene and perylothiophene give comparable structural results to their flexible counterparts. The discrepancy between rigid and flexible models at high densities is not surprising. At high densities, flexible molecules can compress to relieve stresses imposed by the crowded molecules nearby, while rigid molecules cannot. This observation is consistent with entropic considerations

(which depend sensitively on degree-of-freedom constraints) determining phase behavior of densely-packed shapes[21,56]. The densities at which the rigid model differs substantially from the flexible model are significantly higher than experimentally observed densities for these molecules. We therefore conclude that rigid models of perylene and perylothiophene give comparable structural results to their flexible model counterparts and are appropriate for mapping out phase space.

This is further justified by the correspondence between our simulated perylene transition temperatures and those from experiments. In work by Botoshansky, Herbstein, and Kapon, the phase boundary between melted perylene and crystalline α-perylene was found at ≈ 600 K and the phase boundary between α-perylene and β-perylene was found at ≈ 450 K[2]. These temperatures correspond well to the ≈ 600 K and ≈ 400 K phase boundaries between fluid/ordered and ordered/eclipsed phases observed in this work at experimental densities (1.28 g/cm^3).

3.3.2 Molecular packings

In the ordered phase, rigid and flexible models of perylene and perylothiophene self-assemble hexagonally packed columns (Figure 3.3a and b), at temperatures near 600 K. The center-of-mass radial distribution function $g(r)$ shows nearly identical structural correlations among these four models, with flexible perylene demonstrating slightly different ordering length scales (Figure 3.5. For these ordered structures, shown in 3.3a and b, the first $g(r)$ peak corresponds to π-stacking and is located at 3.50 Å in flexible perylene. We measure π-stacking distance for both models of perylothiophene and the rigid model of perylene to be 3.44 Å. We therefore conclude that the rigid model faithfully reproduces the major packing features of its flexible

counterpart for perylene and perylothiophene, with no $g(r)$ peak varying more than 0.1 Å between the two models.
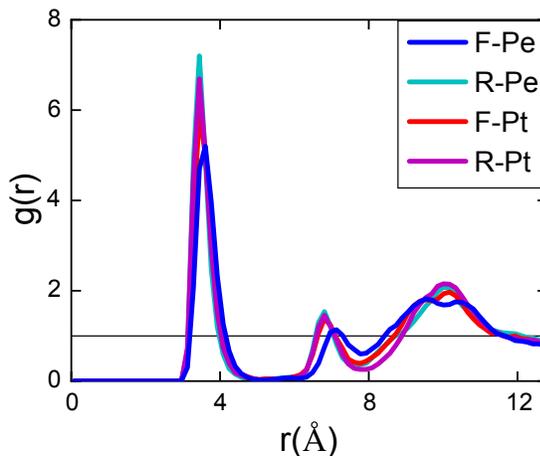
Figure 3.5: Radial distribution function for the ordered simulations of Perylene and Perylothiophene. Flexible perylene (F-Pe) is represented with a blue line, rigid perylene (R-Pe) with a cyan line, flexible (F-Pt) perylothiophene with a red line, and rigid perylothiophene (R-Pt) with a magenta line.

3.3.3 Perylene Morphologies

At temperatures below the disorder-order transition temperature T_{DO} we observe perylene systems with $\rho > 0.85$ g/cm^3 self-assembling into ordered morphologies. We observe a second phase transition between ordered and eclipsed morphologies at even lower temperatures. The ordered morphologies are characterized by stacks of perylene molecules arranged hexagonally, as shown in 3.3. The stacks of molecules in ordered morphologies have no orientational correlations along the stacking axis. In eclipsed morphologies all of the perylenes in a stack have identical orientations, as shown in 3.6. Furthermore, we observe the orientationally ordered stacks in eclipsed morphologies can aggregate in both checkerboard (Figure 3.6a) and aligned (Figure

3.6b) motifs. In the checkerboard case, adjacent stacks have orientations offset by 90°. In the aligned case, adjacent stacks have identical orientations. The potential energy of the checkerboard motif is slightly lower $(-276.13 \pm 0.17\varepsilon)$ but very close to that of the unidirectional grains $(-275.95 \pm 0.19\varepsilon)$ (Appendix B-Section 8).
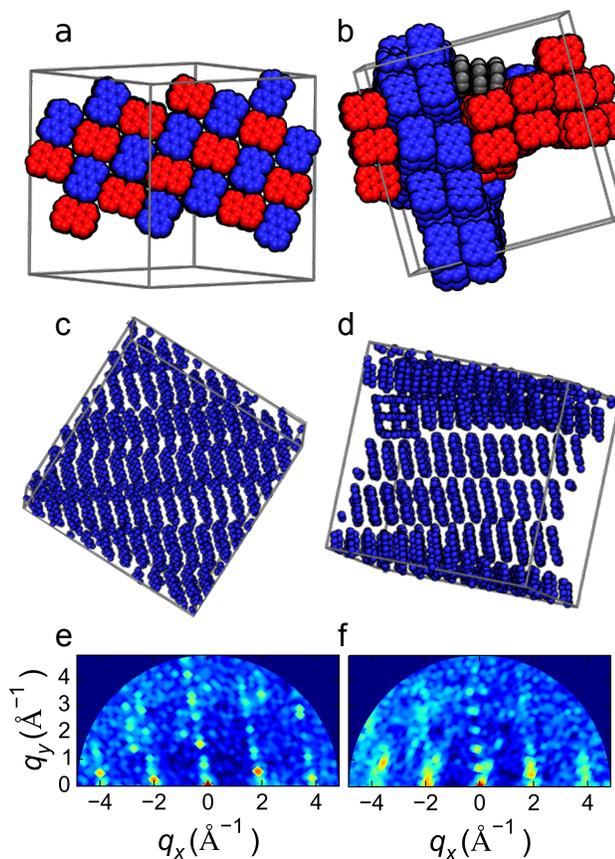
Figure 3.6: (a) Alternating orientations and (b) separated orientations found in low temperature runs for both rigid and flexible perylene. These are colored red and blue to highlight the different orientations. In the separated simulation, there exists a small, differently oriented crystal that is colored gray. Structures for low temperature (c) alternating and (d) separated systems of perylene, with the corresponding simulated GIXS patterns for (e) alternating and (f) separated systems.

Simulated GIXS patterns of ordered and eclipsed morphologies (Figure 3.7a) are qualitatively similar to experimental GIXS patterns of β-perylene crystals (Figure

3.7b, adapted with permission from Ishii and Miyasaka)[28]. We observe in-plane reflections (labeled 110 in 3.7b) which correspond to π-stacking and out-of-plane reflections (labeled 001 in 3.7b) corresponding to columnar packings. We do not observe ($11\bar{1}$) and (120) peaks, corresponding to the glide planes that describe the symmetry between columns in our simulated GIXS patterns. We find, through both GIXS and $g(r)$ analysis, a π-stacking distance of 3.4 Å which is in good agreement with the 3.47 Å found in α-perylene[12,15].
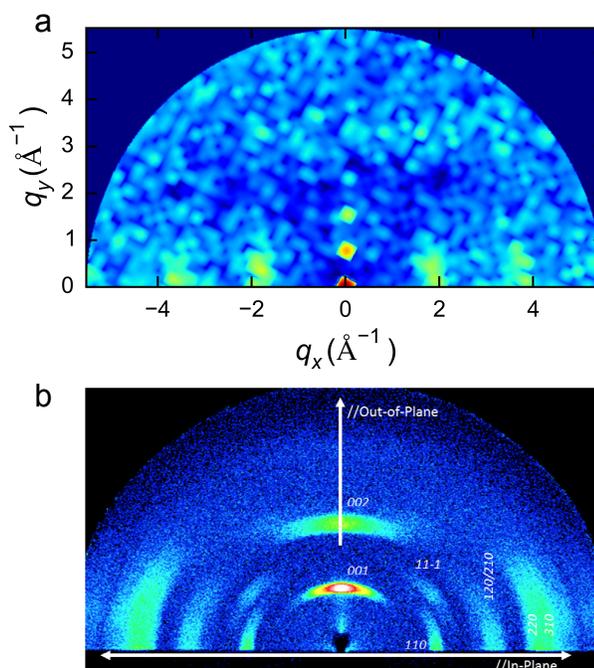
Figure 3.7: (a) Simulated scattering results for hexagonally packed columns, (b) Experimental XRD pattern for β-perylene, reproduced with permission from Ayumi Ishii and Tsutomu Miyasaka [28]. Copyright 2012 Royal Society of Chemistry.

The local structures of the ordered and eclipsed perylene morphologies observed here differ from the experimentally observed α and β polymorphs of perylene[12,15,53]. The herringbone arrangements observed in α and β perylene are not observed in our simulations and are the primary structural difference. Lattice parameters for

β-perylene have been measured experimentally as: $a = 9.76$ Å, $b = 5.84$ Å, and $c = 10.61$ Å[45,46]. We calculate the lattice parameters for aligned eclipsed morphologies to be: $a = 7.36$ Å, $b = 3.33$ Å, and $c = 9.12$ Å. The parameter b is significantly smaller in the simulated morphologies, as the perylene stacks are orthogonal to the b-axis in this investigation, instead of the flattened-out herringbone/γ-structure often observed experimentally in the β phase[14]. For the checkerboard eclipsed morphologies, we calculate lattice parameters: $a = 11.46$ Å, $b = 3.33$ Å, and $c = 16.22$ Å, which are further from β-perylene parameters than the aligned eclipsed morphology. Scattering patterns for the two structures are also shown in 3.6e and f with the corresponding morphologies in c and d. The scattering patterns shown in 3.6e and f are similar to that of the β-perylene structure shown in 3.7b wherein we observe reflections due to π-stacking and columnar formations.

It is not surprising that the rigid and flexible models of perylene studied here do not perfectly capture the crystalline structures of experimentally characterized perylene polymorphs. Our modeling choice to neglect long-range electrostatics is expected to impact the thermodynamic stability of molecular packings[55]. Furthermore, the Lennard-Jones interaction parameters and bond lengths chosen in this work are "off-the-shelf" OPLS parameters that have not been optimized for perylene. With these modeling assumptions in mind, we conclude that our simple models of perylene do a satisfactory job predicting the broad phase behavior and reproduce many of the structural features of perylene polymorphs observed in experiments.

3.3.4   Perylothiophene Morphologies

As with perylene, below T_{DO}, systems with $\rho \geq 0.91$ g/cm^3 self-assemble into ordered morphologies that are characterized by hexagonally arranged columns of π-stacked

perylothiophene molecules (Figure 3.3b). We observe another similarity between perylene and perylothiophene in the π-stacked columns shown in 3.8a, which is quantified in the simulated GIXS pattern (Figure 3.8c). Hexagonally packed features (Figure 3.8b) are also present, as can be seen in the corresponding simulated GIXS pattern (Figure 3.8d). The stacks of perylothiophene molecules have no orientational bias within the ordered stacks at higher temperatures, $T > 400$ K. However, as temperature decreases (below ~ 400 K) sulfur atoms aggregate along the perylothiophene stacks to form an eclipsed phase in which perylothiophene molecules within the stack have the same in-plane orientation (Figure 3.8e and f). Beyond increased aggregation and order within columns, we additionally find that the perylothiophene columns within the eclipsed phase arrange in a staggered formation, with increased anisotropy between them as shown in 3.8g.

Unlike perylene, perylothiophene tends not to assemble into an α-herringbone structure and so only exists as a single, flattened-herringbone β polymorph. This difference is likely due to the more strongly ordered π-stacking arising from the balance between the dispersion force and electrostatic (dipole-dipole) interactions between the sulfur atoms in the thiophene ring[55]. The local structure of the ordered and eclipsed phases are similar to that of the crystal structure determined by Santos and coworkers containing regular π-stacking of perylothiophene molecules within the system[47]. Within the ordered phase, we measure an inter-planar distance of 3.44 Å and an inter-columnar distance of 10.50 Å. However, the eclipsed phase resembles the experimentally determined structure more closely due to the presence of aligned perylothiophene molecules that are in a staggered conformation. We measure lattice parameters $a = 9.34$, $b = 3.49$, $c = 8.42$ Å in the eclipsed state. These results broadly agree with those reported from experiment by Santos and coworkers: $a = 8.46$ and

$c = 8.95$ Å, whereas the b parameter is in excellent agreement $b = 3.52$ Å [47]. Inter-columnar distances, along with a and c lattice parameters in our system are different due to our model predicting that perylothiophene molecules stack perpendicularly to the b-axis whereas they are seen to form a 39.84° angle with the b-axis in experiment. Despite these structural differences, the similarity of our results and experimental structures indicates that our model successfully describes much of the key physics governing self-assembly in perylothiophene.

3.3.5 Performance

	Perylene				Perylothiophene			
	U_{LJ}		ψ		U_{LJ}		ψ	
	Flexible	Rigid	Flexible	Rigid	Flexible	Rigid	Flexible	Rigid
	Relaxation, t_r				Relaxation, t_r			
Disordered	1.00	0.49	0.28	0.17	1.00	0.38	0.60	0.18
Ordered	1.00	0.66	0.31	0.19	1.00	1.03	0.43	0.33
Eclipsed	1.00	0.70	0.20	0.07	1.00	1.20	0.20	0.29
	Autocorrelation, t_a				Autocorrelation, t_a			
Disordered	1.00	1.47	3.67	1.14	1.00	0.96	5.50	0.48
Ordered	1.00	4.14	5.17	3.51	1.00	16.23	16.50	12.10
Eclipsed	1.00	1.60	20.00	12.18	1.00	2.70	10.25	4.65

Table 3.1: Computational efficiency comparisons for rigid and flexible models of perylene and perylothiophene using the non-bonded component of the potential energy (U_{LJ}) and crystallinity order parameter (ψ) as structure metrics. The time taken to relax to equilibrium t_r and autocorrelation time t_a are determined by normalizing τ_r and τ_a by the TPS for each case, respectively. Values of t_r and t_a are normalized to the performance of the flexible model for each molecule calculated using U_{LJ}. Smaller numbers indicate better performance.

The absolute performance of a computational model depends on three factors: the number of timesteps per second (TPS) that can be calculated, the number of timesteps required for the random initial configureation to relax to equilibrium (τ_r) and the
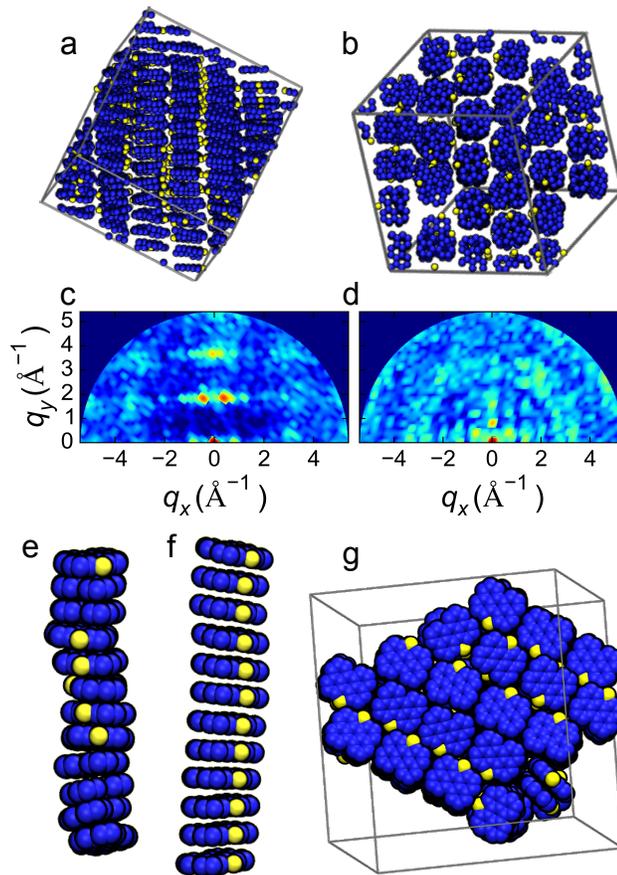
Figure 3.8: Perylothiophene structural features. (a) Side-view simulation snapshot showing π-stacked columns. (b) Top-view snapshot showing hexagonally-packed columns. (c) Simulated GIXS pattern taken from the side and (d) from the top. (e) A single, orientationally uncorrelated stack of perylothiophene at 550 K and (f) 360 K. (g) Top-view simulation snapshot at 360 K in which columns of perylothiophene form staggered conformations.

number of timesteps between statistically independent frames (τ_a) determined from autocorrelations in potential energy and structural metric time series. It is useful to consider these parameters in terms of the "wall-clock time", normalizing τ_r and τ_a by TPS to determine t_r and t_a. To evaluate the computational performance of rigid and flexible models, we compare t_r and t_a between the models in disordered, ordered, and eclipsed regimes. Performance is summarized in Table 3.1, where the non-bonded

contributions to potential energy U_{LJ} and the crystallinity order parameter ψ are each used to calculate relaxation and autocorrelation times. Values of t_r and t_a are normalized by the flexible model U_{LJ}-determined values to give absolute relative performance values. That is, a rigid ψ-calculated value of $t_r = 0.17$ for perylene indicates the rigid model relaxes to equilibrium 5.88 times faster than the flexible model, and a rigid ψ-calculated value of $t_a = 3.51$ for perylene indicates the rigid model requires we wait 3.51 times as long between independant samples than if we used the U_{LJ} autocorrelation times in the ordered phase. For each flexible-rigid pair of data in the system, the lower number of the two indicates the more efficient simulation methodology as relaxation or a statistically independent frame has been obtained more quickly.

Sampling times t_a are generally larger when determined by autocorrelations in the degree of crystallinity ψ than when determined by U_{LJ} autocorrelations. This observation is important because potential energy is frequently used as a proxy for molecular structure and we show here that a direct measure of structure decorrelates more slowly than potential energy. We therefore use autocorrelation times calculated from ψ as the more conservative and accurate metric for both structural sampling and performance comparisons.

For the state points studied here, initial relaxation to equilibrium is a small fraction of the total simulation time compared to sampling the equilibrium distribution. We find that the time required to relax to equilibrium as measured by ψ is faster than that measured by U_{LJ}. Furthermore, we find that the rigid model is generally more efficient than the flexible model at relaxing to equilibrium, independent of whether the simulations are performed at disordered, ordered, or eclipsed regions of phase space (Table 3.1). The only outlier to this is the case of eclipsed perylothiophene

morphologies, where we find the rigid model to be take 1.2 (U_{LJ}) to 1.45 (ψ) times as long as the flexible model to reach equilibrium.

Comparing the autocorrelation times t_a as measured by ψ in Table 3.1 we find the rigid model offers significant performance benefits to sampling independent configureations. The rigid model is 3.2 times faster sampling disordered perylene than the flexible model, 1.47 times faster sampling ordered perylene, and 1.64 times faster sampling eclipsed perylene. For perylothiophene, the rigid model is 11.5, 1.36, and 2.2 times faster sampling disordered, ordered, and eclipsed phases, respectively. The factor of 11.5 times better performance sampling disordered configureations of perylothiophene could be particularly important in the simulation of asphaltenes that include heteroatoms and have interesting disordered aggregation behavior. Because the sampling portion of these simulations is the majority of simulation time and because the rigid model offers a factor of 1.36 to 11.5 times better performance sampling statistically independent structures, we recommend using rigid bodies to model polyaromatic hydrocarbons when appropriate.

3.4 Conclusions

Molecular dynamics simulations are used to investigate the self-assembly of perylene and perylothiophene systems using a simplified molecular model. Perylothiophene is shown to self-assemble hexagonally-packed columns over a broader region of phase space than perylene, a consequence of the added sulfur atom. Our models reproduce the structural features of these molecules observed in experiments (where available), reproducing the π-stacking distances of α-perylene and simulated GIXS analysis show morphologies with features observed in β-perylene. The predicted structures of pery-

lothiophene match well the lattice parameters determined experimentally, particularly at lower temperatures where we see the emergence of staggered perylothiophene rings wherein sulfurs aggregate along the columns. While the simplified UA model used here captures much of the key physics governing perylene and perylothiophene self-assembly, we note differences in lattice parameters and local molecular structure between the morphologies observed here and experiments. In particular, the herringbone arrangement of perylene molecules observed in α and β polymorphs are not observed in this work. Preliminary simulations indicate that herringbone arrangements emerge when long-range electrostatics are included.

We find close agreement between the phase transitions in experimental work and the phase diagrams generated with both flexible and rigid models. The only caveat found is that at high densities where confinement effects are most pronounced, the rigid model and flexible models differ in the thermodynamically-stable packings. We show that potential energy autocorrelations under-predict how fast structure is evolving in these systems and demonstrate the rigid model offers 1.36 to 11.5 times enhanced sampling per second. We therefore propose the use of rigid models to accelerate the simulation of polyaromatic hydrocarbons. In particular, these models offer promise to investigate assemblies of asphaltenes in solution and alternative electron acceptors in photovoltaics, two fields that require access to long aggregation timescales.

BIBLIOGRAPHY

[1] Joshua A. Anderson, Chris D. Lorenz, and A. Travesset. General Purpose Molecular Dynamics Simulations Fully Implemented on Graphics Processing Units. *Journal of Computational Physics*, 227(10):5342–5359, 2008.

[2] Mark Botoshansky, Frank H. Herbstein, and Moshe Kapon. Towards a Complete Description of a Polymorphic Crystal: The Example of Perylene Redetermination of the Structures of the (z = 2 and 4) Polymorphs. *Helvetica Chimica Acta*, 86(4):1113–1128, 2003.

[3] Alejandro L. Briseno, Stefan C. B. Mannsfeld, Samson A. Jenekhe, Zhenan Bao, and Younan Xia. Introducing Organic Nanowire Transistors Semiconductors. *Materials Today*, 11(4):38–47, 2008.

[4] Gianni Cardini. A Comparison Between the Rigid and Flexible Model of Cyclohexane in the Plastic Phase by Molecular Dynamic Simulations. *Chemical Physics*, 193(1-2):101–108, 1995.

[5] Chunxia Chen, Praveen Depa, Victoria García Sakai, Janna K. Maranas, Jeffrey W. Lynn, Inmaculada Peral, and John R. D. Copley. A Comparison of United Atom, Explicit Atom, and Coarse-grained Simulation Models for Poly(ethylene oxide). *Journal of Chemical Physics*, 124(23):234901:1–234901:11, 2006.

[6] Wei Chen, Maxim P. Nikiforov, and Seth B. Darling. Morphology Characterization in Organic and Hybrid Solar Cells. *Energy & Environmental Science*, 5(8):8045–8074, 2012.

[7] Xiaoyu Chen, Paola Carbone, Giuseppe Santangelo, Andrea Di Matteo, Giuseppe Milano, and Florian Müller-Plathe. Backmapping Coarse-Grained Polymer Models Under Sheared Nonequilibrium Conditions. *Physical Chemistry Chemical Physics*, 11(12):1977, 2009.

[8] L. Cheng, R. S. Assary, X. Qu, A. Jain, S. P. Ong, N. N. Rajput, K. Persson, and L. A. Curtiss. Accelerating Electrolyte Discovery for Energy Storage with High Throughput Screening. *The Journal of Physical Chemistry Letters*, 6(2):283–291, 2014.

[9] Jürgen Christen and André Strittmatter. Perylene, C20H12, Alpha-Form Crystal Structure, Lattice Parameters, Density, Melting Point. In *Ternary Compounds, Organic Semiconductors*, volume III/17H, pages 1–4. Springer-Verlag, 2014.

[10] J. L. L. F. S. Costa, D. Simionesie, Z. J. Zhang, and P. A. Mulheran. Aggregation of Model Asphaltenes: a Molecular Dynamics Study. *Journal of Physics: Condensed Matter*, 28(39):394002, 2016.

[11] Benoit Creton, Theodorus De Bruin, Véronique Lachet, and Carlos Nieto-Draghi. Extension of a Charged Anisotropic United Atoms Model to Polycyclic Aromatic Compounds. *Journal of Physical Chemistry B*, 114(19):6522–6530, 2010.

[12] Ayan Datta, Sasmita Mohakud, and Swapan K. Pati. Comparing the Electron and Hole Mobilities in the α and β Phases of Perylene: Role of π-stacking. *Journal of Materials Chemistry*, 17(19):1933–1938, 2007.

[13] Dean M. DeLongchamp, R. Joseph Kline, and Andrew Herzing. Nanoscale Structure Measurements for Polymer-Fullerene Photovoltaics. *Energy & Environmental Science*, 5(3):5980–5993, 2012.

[14] G. R. Desiraju and A. Gavezzotti. Crystal Structures of Polynuclear Aromatic Hydrocarbons. Classification, Rationalization and Prediction from Molecular Structure. *Acta Crystallographica Section B*, 45(5):473–482, 1989.

[15] D. M. Donaldson, J. M. Robertson, and J. G. White. The Crystal and Molecular Structure of Perylene. *Proceedings of the Royal Society A*, 220(1142):311–321, 1953.

[16] Ala'a F. Eftaiha, Jon-Paul Sun, Ian G. Hill, and Gregory C. Welch. Recent Advances of Non-fullerene, Small Molecular Acceptors for Solution Processed Bulk Heterojunction Solar Cells. *Journal of Materials Chemistry A*, 2(5):1201–1213, 2014.

[17] Jens Glaser, Trung Dac Nguyen, Joshua A. Anderson, Pak Lui, Filippo Spiga, Jaime A. Millan, David C. Morse, and Sharon C. Glotzer. Strong Scaling of General-Purpose Molecular Dynamics Simulations on GPUs. *Computer Physics Communications*, 192:97–107, 2014.

[18] Lamia Goual, Mohammad Sedghi, Xiaoxiao Wang, and Ziming Zhu. Asphaltene Aggregation and Impact of Alkylphenols. *Langmuir*, 30(19):5394–5403, 2014.

[19] Michael Grünwald and Phillip L. Geissler. Patterns without Patches: Hierarchical Self-Assembly of Complex Structures from Simple Building Blocks. *ACS Nano*, 8(6):5891–5897, 2014.

[20] Johannes Hachmann, Roberto Olivares-Amaya, Sule Atahan-Evrenk, Carlos Amador-Bedolla, Roel S. Sánchez-Carrera, Aryeh Gold-Parker, Leslie Vogt, Anna M. Brockway, and Alán Aspuru-Guzik. The Harvard Clean Energy Project: Large-scale Computational Screening and Design of Organic Photovoltaics on the World Community Grid. *Journal of Physical Chemistry Letters*, 2(17):2241–2251, 2011.

[21] Amir Haji-Akbari, Michael Engel, and Sharon C. Glotzer. Phase Diagram of Hard Tetrahedra. *The Journal of Chemical Physics*, 135(19):194101:1–194101:10, 2011.

[22] Marcus D. Hanwell, Donald E. Curtis, David C. Lonie, Tim Vandermeerschd, Eva Zurek, and Geoffrey R. Hutchison. Avogadro: An Advanced Semantic Chemical Editor, Visualization, and Analysis Platform. *Journal of Cheminformatics*, 4(17):1–17, 2012.

[23] Jie He, Xinglu Huang, Yan Chun Li, Yijing Liu, Taarika Babu, Maria A. Aronova, Shouju Wang, Zhongyuan Lu, Xiaoyuan Chen, and Zhihong Nie. Self-assembly of Amphiphilic Plasmonic Micelle-like Nanoparticles in Selective Solvents. *Journal of the American Chemical Society*, 135(21):7974–7984, 2013.

[24] Michael P. Hoepfner, Cláudio Vilas Bôas Fávero, Nasim Haji-Akbari, and H. Scott Fogler. The Fractal Aggregation of Asphaltenes. *Langmuir*, 29(28):8799–8808, 2013.

[25] Michael P. Hoepfner and H. Scott Fogler. Multiscale Scattering Investigations of Asphaltene Cluster Breakup, Nanoaggregate Dissociation, and Molecular Ordering. *Langmuir*, 29(49):15423–15432, 2013.

[26] William G. Hoover. Canonical Dynamics: Equilibrium Phase-Space Distributions. *Physical Review A*, 31(3):1695–1697, 1985.

[27] William Humphrey, Andrew Dalke, and Klaus Schulten. VMD: Visual Molecular Dynamics. *Journal of Molecular Graphics*, 14(1):33–38, 1996.

[28] Ayumi Ishii and Tsutomu Miyasaka. A High Voltage Organic-Inorganic Hybrid Photovoltaic Cell Sensitized with Metal-Ligand Interfacial Complexes. *Chemical Communications*, 48(79):9900–2, 2012.

[29] Eric Jankowski, Hilary S. Marsh, and Arthi Jayaraman. Computationally Linking Molecular Features of Conjugated Polymers and Fullerene Derivatives to Bulk Heterojunction Morphology. *Macromolecules*, 46(14):5775–5785, 2013.

[30] William L. Jorgensen and Julian Tirado-Rives. The OPLS [Optimized Potentials for Liquid Simulations] Potential Functions for Proteins, Energy Minimizations for Crystals of Cyclic Peptides and Crambin. *Journal of the American Chemical Society*, 110(6):1657–1666, 1988.

[31] C. Knorowski and A. Travesset. Materials Design by DNA Programmed Self-assembly. *Current Opinion in Solid State and Materials Science*, 15(6):262–270, 2011.

[32] Ting I. N. G. Li, Rastko Sknepnek, Robert J. MacFarlane, Chad A. Mirkin, and Monica Olvera De La Cruz. Modeling the Crystallization of Spherical Nucleic Acid Nanoparticle Conjugates with Molecular Dynamics Simulations. *Nano Letters*, 12(5):2509–2514, 2012.

[33] Jian Liu, Xiaoxia Li, Li Guo, Mo Zheng, Junyi Han, Xiaolong Yuan, Fengguang Nie, and Xiaolong Liu. Reaction Analysis and Visualization of ReaxFF Molecular Dynamics Simulations. *Journal of Molecular Graphics & Modelling*, 53C:13–22, 2014.

[34] Juan Liu, Yapu Zhao, and Sili Ren. Molecular Dynamics Simulation of Self-Aggregation of Asphaltenes at an Oil/Water Interface: Formation and Destruction of the Asphaltene Protective Film. *Energy and Fuels*, 29(2):1233–1242, 2015.

[35] Nusrat Lubna, Ganesh Kamath, Jeffrey J. Potoff, Neeraj Rai, and J. Ilja Siepmann. Transferable Potentials for Phase Equilibria. 8. United-Atom De-

scription for Thiols, Sulfides, Disulfides, and Thiophene. *The Journal of Physical Chemistry B*, 109(50):24100–24107, 2005.

[36] Marcus G. Martin and J. Ilja Siepmann. Transferable Potentials for Phase Equilibria. 1. United-Atom Description of n-Alkanes. *Journal of Physical Chemistry B*, 102(14):2569–2577, 1998.

[37] Katherine A. Mazzio and Christine K Luscombe. The Future of Organic Photovoltaics. *Chemical Society Reviews*, 44(1):78–90, 2015.

[38] John D. McCoy and John G. Curro. Mapping of Explicit Atom onto United Atom Potentials. *Macromolecules*, 31(26):9362–9368, 1998.

[39] Robert T. McGibbon, Kyle A. Beauchamp, Matthew P. Harrigan, Christoph Klein, Jason M. Swails, Carlos X. Hernández, Christian R. Schwantes, Lee Ping Wang, Thomas J. Lane, and Vijay S. Pande. MDTraj: A Modern Open Library for the Analysis of Molecular Dynamics Trajectories. *Biophysical Journal*, 109(8):1528–1532, 2015.

[40] Oliver C Mullins, Hassan Sabbah, Joëlle Eyssautier, Andrew E. Pomerantz, Loïc Barré, A. Ballard Andrews, Yosadara Ruiz-Morales, FarshiPomerand Mostowfi, Richard McFarlane, Lamia Goual, Richard Lepkowicz, Thomas Cooper, Jhony Orbulescu, Roger M. Leblanc, John Edwards, and Richard N. Zare. Advances in Asphaltene Science and the YenMullins Model. *Energy & Fuels*, 26(7):3986–4003, 2012.

[41] Trung Dac Nguyen, Carolyn L. Phillips, Joshua A. Anderson, and Sharon C. Glotzer. Rigid Body Constraints Realized in Massively-parallel Molecular Dynamics on Graphics Processing Units. *Computer Physics Communications*, 182(11):2307–2313, 2011.

[42] Athanassios Z. Panagiotopoulos. Gibbs Ensemble Techniques. In *Observation, Prediction and Simulation of Phase Transitions in Complex Fluids*, pages 463–501. Kluwer Academic Publishers, 1995.

[43] Wolfgang Paul, Do Y. Yoon, and Grant D. Smith. An Optimized United Atom Model for Simulations of Polymethylene Melts. *Journal of Chemical Physics*, 103(4):1702–1709, 1995.

[44] Peter Peumans, Soichi Uchida, and Stephen R. Forrest. Efficient Bulk Hetero-junction Photovoltaic Cells Using Small-Molecular-Weight Organic Thin Films. *Nature*, 425(6954):158–162, 2003.

[45] André Pick, Michael Klues, Andre Rinn, Klaus Harms, Sangam Chatterjee, and Gregor Witte. Polymorph-Selective Preparation and Structural Characterization of Perylene Single Crystals. *Crystal Growth and Design*, 15(11):5495–5504, 2015.

[46] Anupama Ranganathan and G. U. Kulkarni. Aromaticity in Benzene-Like Rings An Experimental Electron Density Investigation. *Journal of Chemical Sciences*, 115(5-6):637–647, 2003.

[47] I. C. Santos, M. Almeida, J. Morgado, M. T. Duarte, and L. Alcácer. Perylo[1,12-b,c,d]thiophene. *Acta Crystallographica Section C*, 53(11):1640–1642, 1997.

[48] Klaus Schmidt-Rohr. Simulation of Small-Angle Scattering Curves by Numerical Fourier Transformation. *Journal of Applied Crystallography*, 40:16–25, 2007.

[49] Hugo Santos Silva, Ana Carolina Rennó Sodero, Brice Bouyssiere, Hervé Carrier, Jean-Pierre Pierre Korb, Ahmad Alfarra, Germain Vallverdu, Didier Bégué, and Isabelle Baraille. Molecular Dynamics Study of Nanoaggregation in Asphaltene Mixtures: Effects of the N, O, and S Heteroatoms. *Energy & Fuels*, 30(7):5656–5664, 2016.

[50] Johan Sjöblom, Sébastien Simon, and Zhenghe Xu. Model Molecules Mimicking Asphaltenes. *Advances in Colloid and Interface Science*, 218:1–16, 2015.

[51] Ana C. R. Sodero, Hugo Santos Silva, Patricia Guevara Level, Brice Bouyssiere, Jean Pierre Korb, Hervé Carrier, Ahmad Alfarra, Didier Bégué, and Isabelle Baraille. Investigation of the Effect of Sulfur Heteroatom on Asphaltene Aggregation. *Energy and Fuels*, 30(6):4758–4766, 2016.

[52] W. C. Swope, H. C. Andersen, P. H. Berens, and K. R. Wilson. A Computer Simulation Method for the Calculation of Equilibrium Constants for the Formation of Physical Clusters of Molecules: Application to Small Water Clusters. *Journal of Chemical Physics*, 76(1):637–649, 1982.

[53] Jiro Tanaka. The Electronic Spectra of Aromatic Molecular Crystals. II. The Crystal Structure and Spectra of Perylene. *Bulletin of the Chemical Society of Japan*, 36(10):1237–1249, 1963.

[54] Mesfin Tsige, John G. Curro, Gary S. Grest, and John D. McCoy. Molecular Dynamics Simulations and Integral Equation Theory of Alkane Chains: Comparison of Explicit and United Atom Models. *Macromolecules*, 36(6):2158–2164, 2003.

[55] Seiji Tsuzuki, Kazumasa Honda, and Reiko Azumi. Model Chemistry Calculations of Thiophene Dimer Interactions: Origin of π-stacking. *Journal of the American Chemical Society*, 124(41):12200–12209, 2002.

[56] Greg van Anders, Daphne Klotsa, N. Khalid Ahmed, Michael Engel, and Sharon C. Glotzer. Understanding Shape Entropy Through Local Dense Packing. *Proceedings of the National Academy of Sciences*, 111(45):E4812–E4821, 2014.

[57] Gustavo Fabián Velardez, Henrik T. Lemke, Dag W. Breiby, Martin M. Nielsen, Klaus Braagaard Møller, and Niels E. Henriksen. Theoretical Investigation of Perylene Dimers and Excimers and their Signatures in X-ray Diffraction. *Journal of Physical Chemistry A*, 112(35):8179–8187, 2008.

[58] Jiang Wang and Andrew L. Ferguson. Mesoscale Simulation of Asphaltene Aggregation. *The Journal of Physical Chemistry B*, pages 8016–8035, 2016.

[59] Shuang-Mei Xin, Qing-Kun Liu, Kai Wang, Yi Chen, Pei-Qing Yuan, Zhen-Min Cheng, and Wei-Kang Yuan. Solvation of Asphaltenes in Supercritical Water: A Molecular Dynamics Study. *Chemical Engineering Science*, 146:115–125, 2016.

[60] Do Y. Yoon, Grant D. Smith, and Tsunetoshi Matsuda. A Comparison in Simulations of a United Atom and an Explicit of polymethylene Atom Model. *The Journal of Chemical Physics*, 98(12):10037–10043, 1993.

[61] Xinzhe Zhu, Daoyi Chen, and Guozhong Wu. Molecular Dynamic Simulation of Asphaltene Co-aggregation with Humic Acid During Oil Spill. *Chemosphere*, 138:412–421, 2015.

[62] Taojun Zhuang, Takeshi Sano, and Junji Kido. Efficient Small Molecule-Based Bulk Heterojunction Photovoltaic Cells with Reduced Exciton Quenching in Fullerene. *Organic Electronics*, 26:415–419, 2015.

[63] Mucong Zi, Daoyi Chen, Haoqing Ji, and Guozhong Wu. Effects of Asphaltenes on the Formation and Decomposition of Methane Hydrate: A Molecular Dynamics Study. *Energy & Fuels*, 30(7):5643–5650, 2016.

CHAPTER 4

OPTIMIZATION AND VALIDATION OF EFFICIENT MODELS FOR PREDICTING POLYTHIOPHENE SELF-ASSEMBLY[1]

4.1   Introduction

Poly-3(hexylthiophene) (P3HT) is one of the polymers that kick-started research into solution-phase self-assembly of OPVs[4,58], and is the focus of this work. Often referred to as the "bench-mark" OPV compound[36], the relative ease of working with P3HT has led to hundreds of studies linking P3HT's structure to its performance in OPV devices[7]. This work on P3HT over nearly two decades highlights the difficulty and opportunity in optimizing self-assembly, and underscores the potential utility of informing experimentation with theory. How many of these experiments could have provided additional insight if equilibrium phase diagrams of P3HT were known in 2002? Would more promising ingredients have been identified earlier if a theoretical maximum η for P3HT blends was known? Answering these questions requires calculating phase diagrams and predicting η, which requires models of P3HT capable of predicting self-assembly.

---

[1]This chapter is published in the journal *Polymers* and is referenced as "Miller, E. D.; Jones, M.L.; Henry, M.M.; Chery, P.; Miller, K.; Jankowski, E. Optimization and Validation of Efficient Models for Predicting Polythiophene Self-Assembly. *Polymers* **2018**, *10*, 1305."

Predicting P3HT self-assembly in particular, and OPV assemblies in general, is difficult because of the multiple length-scales that matter: atomic orbitals, molecular packing, alignment of crystallites, and thermodynamic phase separation all impact OPV device performance. First principles calculations have the highest resolution and can provide insight into charge transport relationships in P3HT[28,33,56], but their computational demands preclude simulating thousands of atoms–far too small to gain insight into the bulk morphological features that arise from thermodynamic self-assembly. Macroscopic models are successful in predicting device-scale morphologies with thickness ~ 100 nm both on-lattice[9,12,13,32,45] and off-lattice[52,54,55], but cannot represent important structural features such as crystallite grain orientations and energetic differences between molecules. Molecular models implemented in molecular dynamics (MD) or Monte Carlo simulations fill the gap between first-principle and macroscopic models, though the system size versus relaxation time trade off significantly hinders investigations of self-assembly[5,11,16,18,20,21,30,48,53]. At the largest scales, the structural evolution of 5 million coarsely-modeled P3HT monomers can be accessed on > 100 nm length-scales[5], but the computational cost of evaluating each step meant that equilibration was inaccessible over the 400 ns simulation trajectory. At 11-nm scales, equilibration of coarse-grained P3HT models is achievable over ~ 2 μs simulation trajectories[20], but such coarse models miss the π-stacking details of P3HT rings, which can have implications for charge transport calculations[22,31]. Long relaxation times can be avoided in MD simulations through carefully selected initial conditions[2,3,19,43,48], but these simulations can only check if a structure is locally stable, not whether it will robustly self-assemble at a particular state-point. Determining the optimal "sweet spot" between system size, model resolution, and computational cost of predicting equilibrium is therefore essential if MD simulations

of thousands of candidate materials and conditions are to be used to inform OPV experimentation[57].

The aims of this work are to (1) describe an optimized model of P3HT that is efficient and meets structural prediction needs, (2) resolve ambiguity around what "big enough" and "fast enough" means for P3HT, and (3) discuss transferable recommendations for simulating other OPV materials. This work is organized as follows: We present our P3HT model in Section 4.2 and characterization techniques in Section 4.3. We explain the important performance metrics and discuss current and future requirements for predicting OPV self-assembly in Section 4.4.1. We employ small-scale simulations to evaluate P3HT self-assembly over ~ 350 state-points in Section 4.4.2 and evaluate the impact of simulated solvent evaporation in Section 4.3.1. We evaluate structural predictions with large-scale simulations in Section 4.4.4, and finally validate against experimental measurements in Section 4.4.5.

4.2   Model

We represent P3HT molecules with an UA model[35] with three types of simulation elements ("beads"): aromatic carbon (CA), aliphatic carbon (CT), and sulfur (S) (Figure 4.1). This level of coarse-graining is convenient for modeling OPV materials, as the reduction of simulation elements from 25 atoms to 11 united-atom sites per monomer reduces computational cost[24,25,35], while simplifying back-mapping of atomic coordinates for charge transport calculations[22]. The base units of mass $M = 32$ amu, energy $\varepsilon = 0.32$ kcal/mol, and length $\sigma_{LJ} = 3.905$ Å, used to describe interactions within the simulation, are adapted from the Optimized Potentials for Liquid Simulations (OPLS)-United Atom (UA) force-field[23]. The pairwise non-bonded

interaction potentials derived from these base units are presented in Table 4.1 in which ε is the depth of the Lennard-Jones potential, σ_{LJ} is the van der Waals radius, and M is the mass of the bead.
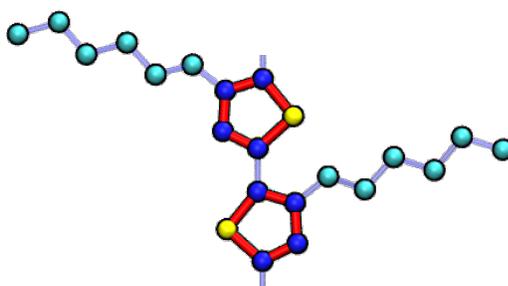
Figure 4.1: P3HT is modeled in this work with a united-atom representation. Sulfur beads (S) are yellow, aromatic carbon beads (CA) are dark blue, and aliphatic carbons beads (CT) are in cyan. Red bonds indicate thiophene rings modeled as rigid bodies, whereas the light blue indicate harmonic bonds.

Table 4.1: Optimized OPLS-UA interaction parameters for CA, CT, and S simulation elements used in this work. ε is the depth of the Lennard-Jones well, σ_{LJ} is the van der Waals radius, and M is the mass.

Bead Type	σ_{LJ} (Å)	ε (kcal/mol)	M (amu)
CA	3.436	0.11	13.0
CT	3.905	0.17	15.0
S	3.436	0.32	32.0

The pairwise bonded constraints (bond lengths, triplet angles, and quadruplet dihedrals) are taken from a modified atomistic force-field based on OPLS-All Atom, as parameterized by Bhatta et al.[2] Since this force-field is atomistic, we adapt it to account for implicit hydrogens and the reduced number of element types in our model (see Appendix C-Section 1 for full details). We also model aromatic thiophene rings as rigid: the bonds, angles, and dihedrals are fixed, maintaining the relative positions of the elements of the rigid bodies throughout the simulation[42]. We further optimize

this model by adjusting interaction parameters to better predict P3HT structure, and lower computational cost with implicit solvent and electrostatics.

The unit of time, τ_s, can be calculated from the base units:

$$\tau_s = \sqrt{\frac{M\sigma_{LJ}^2}{\varepsilon}} = 1.8 \times 10^{-12} \text{ s.} \tag{4.1}$$

We use a timestep of $0.001\tau_s$, so each simulation timestep corresponds to 1.8 fs. The "base case" model considered here handles solvent and long-range electrostatics implicitly, and each oligomer comprises 15 monomers (15mers). Comprehensive evaluation of the optimized model assumptions, including explicit consideration of electrostatic interactions and short oligomer chains are included in Appendix C-Sections 2 and 3.

4.3 Methods

In this work we conduct molecular dynamic (MD) simulations using the Graphical Processing Unit (GPU)-accelerated HOOMD-blue simulation package[1,10], performed on NVIDIA K80 and P100 GPUs. The code used to produce this data is open-source and freely available at Ref [39]. The complete dataset from this investigation is available at Ref [38]. Simulations are conducted in the canonical ensemble (NVT), in which the total number of particles, volume, and temperature are kept constant. The Nosé-Hoover thermostat, in which the system is coupled to a heat bath, is applied to maintain the temperature[15]. Particle positions and velocities are updated with the two-step velocity-Verlet integration of Newton's equations of motion with a timestep of 1.8 fs[51].

Each simulation is initialized from a unique random configuration within a cubic volume with periodic boundary conditions. We accomplish this by first placing molecules created with the mBuild software package[26] at random positions in a large simulation volume, where molecules are sufficiently separated so that they can be placed without overlapping their neighbors. A short MD simulation (1.8 ns) is performed at high temperature (T$\sim$ 1300 K) to randomize the molecule positions and orientations. The system volume is then reduced during another short simulation (1.8 ns and 1300 K) until the target density is reached. This process of "initializing", "mixing", and "shrinking" has been previously used to initialize independent snapshots at arbitrary densities[14,20,34,40]. Unless otherwise specified, every simulation presented herein is instantaneously quenched from high temperature to the target temperature for the duration of its NVT simulation. We consider target temperatures from $T = 80$ to 1300 in steps of 80 K. These temperatures span the glass transition (300 K) and melting (490 K) temperatures expected for P3HT[2]. Of course, real P3HT degrades at the higher temperatures in this range. Our simulations at these high-T conditions are performed to provide unique independent snapshots from which to initialize independent simulations, and to check if high-temperature structural transitions might exist if P3HT did not degrade.

We consider a range of relevant film densities, $\rho = 0.56, 0.72, 0.89, 1.05, 1.11$ g/cm^3, with the largest $\rho = 1.11$ g/cm^3 corresponding to the experimental thin-film density for P3HT[41]. For the lowest densities, as much as 40% of the simulation volume is occupied by the implicit solvent, whereas the volume occupied by the implicit solvent is negligible for the highest densities. We employ an extremely simplified model of solvent quality: The we define the parameter ε_s to represent how poor the solvent is for P3HT, and scale all of the pairwise interaction potentials

by this amount ($\varepsilon_{ij} \rightarrow \varepsilon_s \times \varepsilon_{ij}$). We explore ε_s values from 0.2 to 1.2 in this work. Low values of $\varepsilon_s \leq 0.7$ correspond to solvents in which P3HT is highly soluble (e.g. chloroform, chlorobenzene or 1,2-dichlorobenzene), whereas larger values $\varepsilon_s > 0.7$ describe solvents where P3HT is less soluble (e.g. acetone)[44]. While this simplified model cannot capture complex or entropic solvent phenomena, it provides a significant computational advantage[8,47,49]. Furthermore, adjusting ε_s while holding T constant enables exploration of how equilibrium structure depends on molecular attractions at fixed kinetic energy. In this work, we perform simulations at the combinations of T, ρ, and ε_s described above to understand how these parameters in concert influence thermodynamic phase behavior.

4.3.1 Solvent Evaporation

Each simulation performed herein utilizes one of two simulation protocols to sample microstates at the target state-point. Protocol (1) ignores solvent evaporation: disordered initial configurations at the target density, ρ, and solvent quality, ε_s, are instantaneously quenched from $T = 1300$ K to the target temperature, after which equilibration progress is monitored[40]. Protocol (2) is a very basic, qualitative model of solvent evaporation that helps to sample configurations at experimental densities ($\rho = 1.11$ g/cm^3): First, a system is equilibrated at $\rho = 0.72$ g/cm^3 and the target temperature and ε_s using Protocol (1), followed by a linear compression to $\rho = 1.11$ g/cm^3 over 280 ns. After the shrinking step of Protocol (2), equilibration progress is monitored as in Protocol (1).

Modeling implicit solvent removal in this way creates two fundamental tensions with our claimed thermodynamic approach. Firstly, invoking Protocol (2) suggests that microstate sampling with Protocol (1) is non-ergodic over practical time scales.

It is well established in both experiments and simulations that polymer dynamics are kinetically arrested at higher densities, so Protocol (2) can be viewed as a sampling acceleration scheme that assumes structures arising from enthalpy minimization at low densities are representative of equilibrium at high densities. Of course, steric entropic effects are known to contribute most interestingly to the free energy at high densities, with the striking diversity of entropically stabilized hard-polyhedra phases as just one example[6]. This leads to the second tension: Invoking an implicit solvent model assumes entropic contributions of the solvent either (a) can be effectively represented in the coarse Hamiltonian, or (b) are negligible compared to enthalpic contributions. It is therefore implied in the present work—and every multiscale study invoking coarse-graining—that potential energy minimization dominates the free energy minimization of the coarse Hamiltonian, whose emergent coarse structures represent the underlying atomistic description with all of its encoded entropic contributions. In the cases where such coarse-graining is not predictive of the more detailed representation's structure, there are interesting open questions about how to include entropic contributions within the coarse Hamiltonian explicitly, or whether back-mapping fine-grained structure is sufficient, both beyond the scope of this work.

Protocol (2) does not capture evaporation-driven dynamics that occur in real systems, including the alignment of linear molecules induced by hydrodynamic flows and steric effects at interfaces. In principle, such dynamics could be used to facilitate self-assembly, but are beyond the equilibrium approach of this work.

4.3.2 Morphology Characterization

To characterize the molecular packings obtained in our simulations we use two structural metrics: an order parameter and simulated GIXS using the Diffractometer

simulation software[22,37]. GIXS patterns are used to identify and quantify periodic morphological features and are used to validate predicted structures directly against experiments. We obtain a set of patterns by simulating diffraction on each cubic morphology from 60 unique orientations uniformly distributed on a sphere. We identify orientations with clearly resolved peaks and align crystallographic directions along the same axes before averaging these orientations into a final diffraction pattern. Treating the diffraction patterns in this way improves signal-to-noise ratio of periodic features, allowing detection of periodic length-scales more precisely.

The order parameter, ψ, is used to describe the proportion of thiophene rings in "large" clusters. The clustering algorithm is described in full in our previous work and presented, with examples, in the Appendix C-Section 4[40]. Briefly, two thiophene rings are considered "clustered" if their centers-of-mass are within 6.6 Å of each other and if the planes of the thiophene rings are oriented within 20° degrees of each other. The value of 6.6 Å is informed by the radial distribution function of the thiophene centroids in ordered P3HT, and the 20° cut-off is taken due to rotations under this having small effect on the transfer integral (a measure of the electron orbital overlap) between two rings[28]. A cluster must contain at least six thiophene rings to be considered "large" and contribute towards ψ, a cut-off that is selected to distinguish morphologies with fewer large clusters from those with many small clusters.

4.4   Results and Discussion

Here we benchmark P3HT simulations using our optimized model to provide context for the system sizes that are practically accessible, perform experiments with simulated solvent evaporation as potential way to avoid long relaxation times, and

evaluate the system sizes needed to validate predictions against experiments.

4.4.1 Computational Performance and Scaling

The time it takes to predict self-assembly of a material with MD primarily depends upon the size of the simulated volume, which affects two key metrics:

1. Relaxation time: The number of timesteps that must be evaluated before the system reaches equilibrium. Larger volumes generally mean larger relaxation times because more molecules must rearrange before the system has converged to the equilibrium distribution of microstates.

2. Computational performance: The number of timesteps that can be evaluated per each second that elapses on a clock on the wall, here measured as Timesteps Per Second (TPS). TPS scales between $O(N^{-1})$ and $O(N^{-2})$.

We measure relaxation time and TPS in order to quantify the practicality of performing equilibrium simulations as a function of system size. We perform instantaneous quenches to $T = 600$ K at $\rho = 0.72$ g/cm^3 and $\varepsilon_s = 0.8$ for our base case model with N ranging from $N = 16,500$ to $N = 600,000$.

Figure 4.2a shows TPS decreases monotonically with N, closely matching the $O(N^{-1})$ reference slope (orange). For the smallest systems ($N = 16,500$), this corresponds to being able to perform 400 ns per day, and, for the largest systems ($N = 600,000$), 17 ns per day. In Figure 4.2b we show a characteristic time evolution of the Lennard-Jones pair potential energy, which we use as one proxy for structure. At equilibrium, measurements of potential energy are observed to fluctuate about a stable, time-invariant average (Region 3 in Figure 4.2b). Before equilibrium is reached, we observe a fast initial change in structure (Region 1), followed by a slower
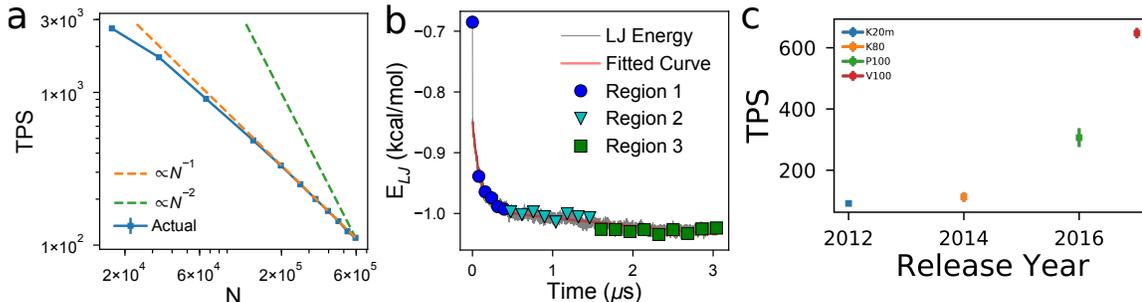
Figure 4.2: GPU-accelerated simulations of P3HT presented here achieve ideal performance scaling, but practical equilibration of a volume is limited by relaxation times. (a) Computational performance measured by Timesteps Per Second (TPS, blue data) scales $O(N^{-1})$ (dashed orange line for simulations performed at $T = 600$ K at $\rho = 0.72$ g/cm^3 and $\varepsilon_s = 0.8$). (b) The time evolution of non-bonded potential energy shows a fast initial structural rearrangement (blue), a slower relaxation to equilibrium (cyan), followed by sampling of equilibrium microstates (green). The divergence of relaxation time (cyan) with system size, N, and density, ρ, puts practical limitations on the systems that can be equilibrated, despite high TPS values. (c) Computational performance, measured by TPS, has roughly doubled with each new hardware release for the last four generations of Nvidia hardware (K20m, K80, P100, and V100 cards). Error bars indicate one standard deviation over 5 independent simulations per cluster across multiple clusters.

relaxation time (Region 2). We detail the automatic detection of these regions and present the curves for multiple systems in Appendix C-Section 5. Here we observe Region 1 is insensitive to N, occurring within the first 0.5 μs of simulation time. The relaxation time (Region 2), however, strongly depends on N. We measure relaxation times of ~ 0.2 μs for the $\sim 16,000$ beads, ~ 0.4 μs for $\sim 29,000$ beads, and ~ 1.0 μs for 40,000 beads. For system sizes larger than $N = 40,000$ we do not observe equilibration of the base case model 15mers at $T = 600$ K, $\rho = 0.72$ g/cm^3, $\varepsilon_s = 0.8$. Empirically these observations suggest relaxation time scales $O(N^2)$, though the longer relaxation times for larger N tested here precludes detailed evaluation.

Once a system has come to equilibrium, we measure decorrelation times, explained in detail in previous works[14,40]. The 100 15mer simulation requires ~ 80 ns for

each independent measurement to be generated, the 175 15mer simulation requires ~ 50 ns per measurement, and the 250 15mer simulation requires ~ 300 ns per measurement. Therefore, simulations of around one μs are needed to sample the equilibrium distribution of microstates after the relaxations of Regions 1 and 2.

To put these performance numbers in context with advances in computational hardware, we benchmark P3HT systems with $N = 165,000$ on 4 different Nvidia GPUs and 6 different high performance computing systems. Figure 4.2c shows a factor-of-two improvement in TPS roughly every two years. The TPS scaling and relaxation scaling as a function of N, and evolution of TPS over GPU release year data presented so far allows us to answer "How many years must we wait before we can equilibrate a system with twice the spatial dimensions of the largest practical dimensions today?" Doubling the size of the simulation volume along each axis results in 8 times the volume and therefore 8 times higher N (given the same density), so we would expect the TPS to drop by a factor of ~ 8, given that TPS $\propto O(N^{-1})$. However, as relaxation time scales as roughly $O(N^2)$, we would require 64 times as many simulation time steps to equilibrate before sampling. This means that doubling the linear dimensions of a system requires 512 times the TPS to equilibrate it in the same amount of time. Extrapolating current hardware trends, a new GPU 18 years from now would meet this performance need. Of course, performance scaling will vary significantly, depending on model details (e.g. chain length), T, ρ, and ε_s, so the precise numbers reported here will have limited transferability to other chemistries and conditions.

Even so, we draw two takeaways from these data: The first dispels the idea that significantly larger volumes can be equilibrated with incremental advances in hardware. Rather, doubling the dimensions of a system requires decades of hardware

improvement, all other factors being equal. Consequently, the second takeaway is that techniques that mitigate relaxation times will be essential in predicting OPV morphologies relevant to device scales. Such techniques include modeling at multiple scales, modeling the minimal necessary physics at each scale, efficiently sampling parameter space, and advanced sampling techniques[17,46,50].

4.4.2 Identifying optimal assembly conditions

Despite the divergent behavior of simulation time as a function of N observed in the previous section, it is computationally tenable to efficiently sample the state space of P3HT self-assembly using a base case system of 100 15mers, using our OPLS-UA model. We therefore perform an ensemble of MD simulations over a range of 350 unique state-points (depicted by a black "x" in Figure 4.3) each defined by T, ρ, and ε_s, to determine which combinations are correlated with self-assembly. Doing so generates the rough phase diagram of P3HT structure as a function of T, ρ, and ε_s. Each simulation employs cubic volumes with edge length ~ 7 nm. These volumes relax to equilibrium within ~ 180 ns, after which the ordering measured by ψ is constant. Decorrelated equilibrium microstates are drawn from trajectories after this initial relaxation, with an additional 180 ns of simulation time generating ~ 8 microstates per state-point. The colorbars in Figure 4.3 quantify the degree of ordering measured by the order parameter ψ. In each case, more ordered systems appear red, whereas systems with less ordering and fewer ordered clusters appear in blue. The order parameter values depicted between simulated state-points are linearly interpolated.

In Figure 4.3 we observe two major trends in P3HT ordering: (1) increasing the density limits the ordering and (2) there exists a narrow band of T-ε_s combinations that produce a high degree of order, independent of density. The first trend arises
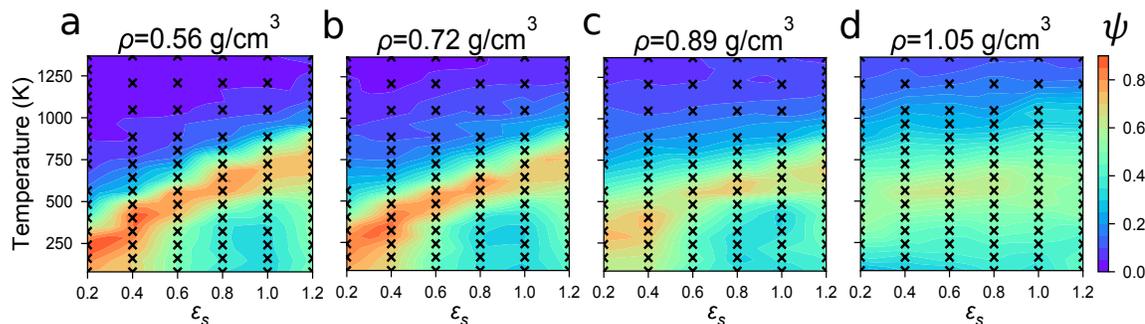
Figure 4.3: The degree of ordering, ψ, for Protocol (1) shows the most robust assembly occurs at lower densities, with more temperature-solvent combinations resulting in high ψ. (a) $\rho = 0.56$ g/cm^3, (b) $\rho = 0.72$ g/cm^3, (c) $\rho = 0.89$ g/cm^3, (d) $\rho = 1.05$ g/cm^3. Red regions denote order, whereas blue denotes disorder. Each black "x" indicates a measurement from an MD trajectory, and ψ values between measurements are linearly interpolated.

from systems becoming kinetically arrested: Chains have little room to rearrange at high densities after being instantaneously quenched below the melting temperature. The second trend arises from the relationships between T, kinetic energy, and the scaling of the Lennard-Jones well-depths through ε_s. When the ratio $\frac{T}{\varepsilon_s}$ is sufficiently high, simulation elements have sufficient kinetic energy to routinely break out of the short-range pairwise potential energy wells of their neighbors. Conversely, in systems with deep potential wells, beads are more likely to get stuck in local potential energy minima, resulting in longer relaxation times. As expected, we observe that P3HT orders most robustly when it has both sufficient free volume and kinetic energy to rearrange, providing the temperature is below the melting temperature. These requirements are consistent with experimental annealing practices used to increase order, where energy is added (thermal annealing) or interaction strengths are decreased while increasing free space for polymers to rearrange (solvent annealing).

In Figure 4.3, we also observe that P3HT is able to robustly self-assemble over a range of a couple of hundred Kelvin. This self-assembly occurs just below the melting

temperature, given a particular solvent quality. In systems with sufficient free space to order, the model predicts melting temperatures in the range of ~ 400 to 600 K (depending on the solvent strength), which corresponds well to the experimentally observed melting temperature at 490 K. The experimental melting temperature of P3HT in the absence of a solvent is reproduced when $\varepsilon_s = 0.5$, indicating that the optimized OPLS force-field used here overpredicts P3HT's melting temperature and that varying ε_s can be thought of as either varying solvent quality, or correcting for systematic attraction offsets in the force-field.

4.4.3 Modeling Solvent Evaporation Facilitates Equilibration

We observe that P3HT simulations at $\rho \geq 1.05$ g/cm^3 show a low degree of order, ψ (Figure 4.3), when instantaneously quenched from $T = 1300$ K to the target simulation temperature. However, highly ordered P3HT has been observed in experiments at and near this density. We explain this discrepancy by kinetic arrest over simulation timescales: Closely-packed P3HT volumes with negligible solvent have long rearrangement times. To avoid such trapping and to more faithfully model solvent evaporation, we perform "shrinking" simulations using a simple model of solvent evaporation (Protocol (2)) from $\rho = 0.72$ g/cm^3 to $\rho = 1.11$ g/cm^3 over 36 ns, and compare the resultant systems to the base case in the previous section. The initial density $\rho = 0.72$ g/cm^3 is chosen because it is the highest density at which highly-ordered morphologies are robustly assembled. When solvent evaporation is modeled in this way we generally observe negligible change in ψ (Figure 4.4) as the system transitions from $\rho = 0.72$ g/cm^3 to $\rho = 1.11$ g/cm^3. At high temperatures, $600 \leq T < 900$K, we observe increased ordering as a result of solvent evaporation (Figure 4.4), which is consistent with previous work showing that increased density

at constant temperature can lead to a higher degree of order[40]. In aggregate, these results indicate that our OPLS-UA model is efficient enough to identify the temperature-solvent-density combinations that result in molecular self-assembly.
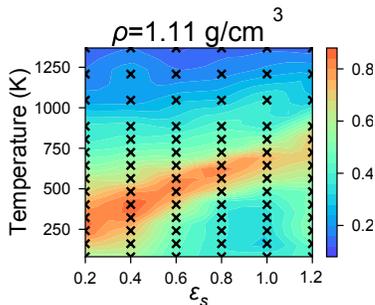
Figure 4.4: Morphologies sampled with Protocol (2) are observed to have higher ordering, ψ than Protocol (1) at the same state-points (compare Figure 4.3d), which suggests that simulating solvent evaporation helps to avoid long relaxation times.

The results presented in Figure 4.3 are generated with Protocol (1): Low-solvent (high P3HT density) systems display less order because of longer rearrangement times. The results presented in Figure 4.4 are generated with Protocol (2): Equilibrating and then shrinking the simulation volume while holding temperature constant results in structures that are as ordered as those at $\rho = 0.72$ g/cm^3, but at the experimental density of $\rho = 1.11$ g/cm^3, and with GIXS in quantitative agreement with experiments (Figure 4.6). We therefore recommend using Protocol (2) for simulating solvent evaporation where appropriate, because otherwise long rearrangement times at high densities can be avoided.

4.4.4 Large volumes are needed for experimental validation

Here, we combine the results of the previous two sections and perform solvent evaporation simulations of large volumes at specific state-points to evaluate which advantages in structural insight, if any, are afforded with larger volumes. We compare the base

case "small" systems of 100 15mers ($N = 16,500$, $L = 7$ nm) against "large" 1000 15mer ($N = 165,000$, $L = 15$ nm) systems. The large simulations are initialized at $T = 600$ K, $\rho = 0.72$ g/cm^3, and $\varepsilon_s = 0.8$ using Protocol (1). During the evolution of the large systems, we record atom positions at three different degrees of order: when the system is disordered, when some crystallites have formed but disordered regions still exist (semi-ordered), and when it has ordered. These times are chosen based on the degree of structural evolution discussed in Section 4.4.1. Each of these snapshots is used to initialize independent simulations using Protocol (2) to reach $\rho = 1.11$ g/cm^3 over a 180 ns simulation trajectory. We compare these three large morphologies at experimental densities, "disordered": $\psi \sim 0.4$, "semi-ordered": $\psi \sim 0.6$, and "ordered": $\psi \sim 0.8$), to the smaller base cases. Note, we present our analysis for only the ordered system here in the main text and in Appendix C-Section 6 we present the analysis for the semi-ordered and disordered systems.
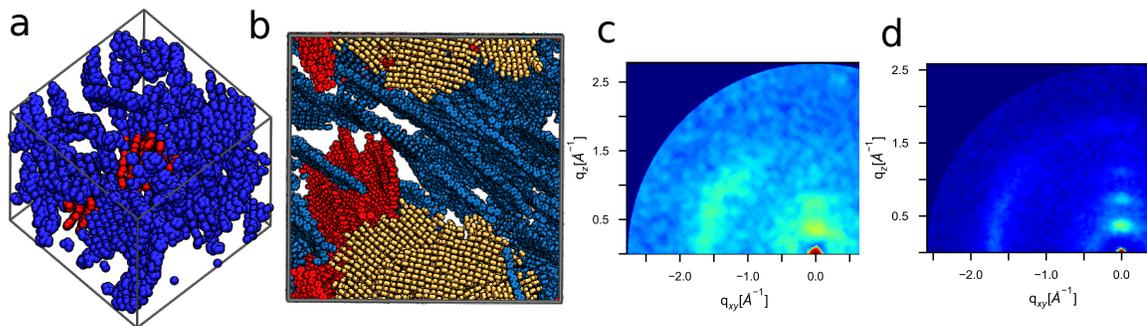
Figure 4.5: While small systems (a) are sufficient for identifying key structural features of ordered P3HT, large volumes (b) are needed to resolve structural periodicities (c-small, d-large) and therefore enable experimental validation.

The large and small ordered systems shown in Figure 4.5a and b, in which only "large" clusters, identified using the cluster analysis discussed in section Section 4.3.2, are shown (large clusters ≥ 6 monomers, and side chains are omitted). The

large ordered system contains a few large crystallites, colored blue, red, and yellow. This contrasts with the small morphology, which primarily consists of a single large crystallite (shown in blue in Figure 4.5a), with the next largest having significantly fewer members (shown in red). These results indicate that smaller systems will tend to have fewer ordered crystallites, which limits the opportunity to observe periodic organization of these structures. Despite this difference, GIXS patterns show that the same periodic distances are present in both system sizes (Figure 4.5 c vs. d), albeit with significantly increased noise in the case of the smaller system. As such, small morphologies can be used to identify state-points of structural interest, however, large simulations are better at characterizing crystal structure and quantifying morphological order.

4.4.5 Experimental Validation of Optimized P3HT Model

To validate our model, we perform simulations of 1,000 15mers with Protocol (2) and compare simulated GIXS patterns against experimental P3HT patterns ($T = 600$ K, $\varepsilon_s = 0.8$). Predicted and experimental GIXS patterns are presented in Figure 4.6a (averaged over 18 simulation orientations) and Figure 4.6b (Reprinted with permission from [27]. Copyright 2012 American Chemical Society). Both experimental and predicted structures are characterized by bright reflections extending vertically along the out-of-plane axis with reciprocal spacing of 0.38 ± 0.02 Å^{-1} (corresponding to real-space separation of 16.5 Å) and the narrow peak perpendicular to the [100] direction at 1.68 ± 0.02 Å^{-1} (corresponding to a real-space separation of 3.74 Å). To connect these scattering features to morphological features, we present the ordered morphology in Figure 4.7a, which shows lamellae of π-stacked thiophene rings (shown with dark blue CA and yellow S), and aliphatic tails (cyan CT). It is the periodic

π-stacking at ~ 3.7 Å and perpendicular alkyl-stacking at 16.5 Å responsible for the ~ 1.7 and 0.4 Å^{-1} features that are observed in the GIXS patterns. The agreement between experimental and predicted structures demonstrates the present OPLS-UA model is capable of efficiently and quantitatively predicting ordered P3HT structures within three weeks of simulation on a single GPU. Also similar to the structures seen experimentally, the lamellae in the ordered system do not represent a single, perfect crystal, but rather multiple crystallites with various grain orientations. The thiophene rings in these grains are depicted by red, blue, and yellow in Figure 4.7b.
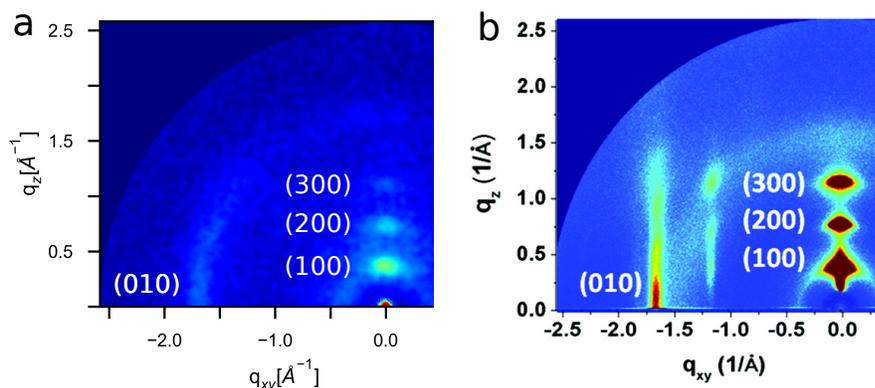
Figure 4.6: Our model produces (a) simulated GIXS patterns that closely match (b) experiment with π-stacking features along the (010) plane at 1.68 ± 0.02 Å^{-1} and alkyl-stacking features along the (100) plane with a spacing of 0.38 ± 0.02 Å^{-1}. (Experimental GIXS pattern (b) Reprinted with permission from Ref [27]. Copyright 2012 American Chemical Society).

Within each layer, the thiophene rings primarily stack co-facially in either an "aligned" (Figure 4.7c) or "anti-aligned" (Figure 4.7d) conformation, in which the sulfur atoms of adjacent rings are on the same side or opposite sides of the stack respectively. We calculate the radial distribution function ($g(r)$, Figure 4.7e) between monomer centers to characterize short-range packing. A monomer center is defined by the geometric average position of the sulfur and two furthest carbons in the thiophene
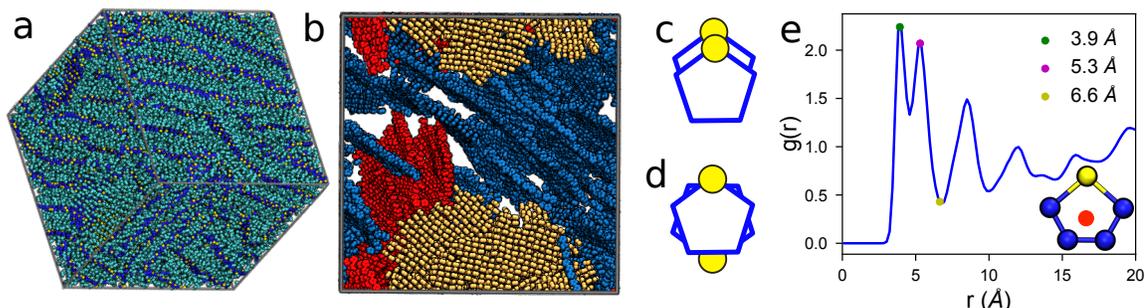
Figure 4.7: (a) A representative ordered molecular morphology of P3HT 15mers (CA-dark blue, S-yellow, CT-cyan) depicting π- and alkyl-stacked lamellae (state-point: implicit charges, $T \sim 600$ K, $\varepsilon_s = 0.8$, $\rho = 1.11$ g/cm^3). (b) The locations of the three largest crystallites in the system (colored blue, red, and yellow in order of descending size). Small crystallites and side chains are omitted for clarity. Within each crystallite, thiophene rings stack in (c) an aligned or (d) anti-aligned conformation, which are observed in (e) the $g(r)$ of the thiophene centroid (e-inset) as the green (3.9 Å) and magenta (5.3 Å) dots respectively. The $g(r)$ minimum at 6.6 Å (yellow dot) is used as a clustering criterion describing the maximum separation of two rings in the same cluster.

rings (see Figure 4.7e inset), and the spacing between two centers is used to distinguish aligned and anti-aligned π-stacking. The first peak in the $g(r)$ describes π-stacking of the thiophene heads and is split into two features at 3.9 and 5.3 Å corresponding to the aligned and anti-aligned cases, respectively. As evidenced by the $g(r)$ peak magnitudes, we observe a slight preference for aligned thiophene stacking vs. anti-aligned stacking. Generally, more ordered morphologies of P3HT are expected to provide faster charge transport characteristics. As such, these results show that sufficient amounts of good solvent, which is then evaporated off just below the P3HT melting temperature, are expected to produce ordered morphologies with beneficial electronic device properties.

4.5 Conclusions

In this work we presented insight into semiconducting polymer assembly aimed at both molecular simulators and experimentalists. Specifically for P3HT, we demonstrated excellent quantitative agreement with P3HT nanostructure investigated by GIXS and we found temperature and solvent combinations where robust self-assembly into ordered structures is expected. In doing so, we validated the predictive accuracy of our optimized OPLS-UA model, which implicitly includes solvent, charges, and abstracts away fast degrees of freedom in the thiophene rings.

Our analysis of computational efficiency scaling with simulation size showed that projected improvements to computational hardware over the next decades will not enable the equilibration of significantly larger organic semiconductor volumes than those presented, using the techniques demonstrated in this work. Since relaxation times were identified as the limiting factor to polymer equilibration, multi-scale techniques and model approximations must be used in order to predict OPV morphologies at experimentally relevant length scales. For instance, we demonstrated that modeling solvent molecules implicitly by modulating the inter-molecular interactions in our forcefield, and implementing a very basic technique to simulate solvent evaporation leads to good experimental predictions at relatively low computational cost. Based on our observations, we therefore propose the following simulation guidelines for predicting the morphologies of OPV candidate molecules:

1. Benchmark performance to identify the system size N that is practical for equilibrating hundreds of systems.

2. Generate coarse phase diagrams with these inexpensive simulations to identify rough phase boundaries and interesting structures.

3. Use simulated solvent evaporation to generate morphologies at experimental densities, with sufficiently large volumes.

4. Validate predictions against experimental GIXS patterns, when available.

These guidelines can be applied to any OPV active layer material, and will help to ensure that the most information about model validity and OPV morphology are gained per unit of simulation time. Combining these guidelines with automatic identification methodologies and more detailed to more efficiently search parameter space[29,46] will further improve information gained per Central Processing Unit cycle. Extending the current investigation and applying these methods to a broader range of OPV candidate materials with potential for mitigating climate change will be the focus of future work.

BIBLIOGRAPHY

[1] Joshua A. Anderson, Chris D. Lorenz, and A. Travesset. General Purpose Molecular Dynamics Simulations Fully Implemented on Graphics Processing Units. *Journal of Computational Physics*, 227(10):5342–5359, 2008.

[2] Ram S. Bhatta, Yeneneh Y. Yimer, David S. Perry, and Mesfin Tsige. Improved Force Field for Molecular Modeling of Poly(3-hexylthiophene). *The Journal of Physical Chemistry B*, 117(34):10035–10045, 2013.

[3] Natalia I. Borzdun, Sergey V. Larin, Stanislav G. Falkovich, Victor M. Nazarychev, Igor V. Volgin, Alexander V. Yakimansky, Alexey V. Lyulin, Vikas Negi, Peter A. Bobbert, and Sergey V. Lyulin. Molecular Dynamics Simulation of Poly(3-hexylthiophene) Helical Structure In Vacuo and in Amorphous Polymer Surrounding. *Journal of Polymer Science, Part B: Polymer Physics*, 54(23):2448–2456, 2016.

[4] N. Camaioni, G. Ridolfi, G. Casalbore-Miceli, G. Possamai, and M. Maggini. The Effect of a Mild Thermal Treatment on the Performance of Poly(3-alkylthiophene)/Fullerene Solar Cells. *Advanced Materials*, 14(23):1735–1738, 2002.

[5] Jan-Michael Y. Carrillo, Rajeev Kumar, Monojoy Goswami, Bobby G. Sumpter, and W. Michael Brown. New Insights Into the Dynamics and Morphology of P3HT:PCBM Active Layers in Bulk Heterojunctions. *Physical Chemistry Chemical Physics*, 15(41):17873, 2013.

[6] P. F. Damasceno, M. Engel, and S. C. Glotzer. Predictive Self-Assembly of Polyhedra into Complex Structures. *Science*, 337(6093):453–457, 2012.

[7] Minh Trung Dang, Lionel Hirsch, and Guillaume Wantz. P3HT:PCBM, Best Seller in Polymer Photovoltaic Research. *Advanced Materials*, 23(31):3597–3602, 2011.

[8] Michael Feig and Charles L. Brooks. Recent Advances in the Development and Application of Implicit Solvent Models in Biomolecule Simulations. *Current Opinion in Structural Biology*, 14(2):217–224, 2004.

[9] Jun-ichi Fukuda, Makoto Yoneya, and Hiroshi Yokoyama. Numerical Treatment of the Dynamics of a Conserved Order Parameter in the Presence of Walls. *Physical Review E*, 73(6):066706, 2006.

[10] Jens Glaser, Trung Dac Nguyen, Joshua A. Anderson, Pak Lui, Filippo Spiga, Jaime A. Millan, David C. Morse, and Sharon C. Glotzer. Strong Scaling of General-Purpose Molecular Dynamics Simulations on GPUs. *Computer Physics Communications*, 192:97–107, 2014.

[11] Jonathan Gross, Momchil Ivanov, and Wolfhard Janke. Comparing Atomistic and Coarse-Grained Simulations of P3HT. *Journal of Physics: Conference Series*, 750(1):012009, 2016.

[12] Chris Groves, Robin G. E. Kimber, and Alison B. Walker. Simulation of Loss Mechanisms in Organic Solar Cells: A Description of the Mesoscopic Monte Carlo Technique and an Evaluation of the First Reaction Method. *The Journal of Chemical Physics*, 133(14):144110, 2010.

[13] Ian C. Henderson and Nigel Clarke. On Modelling Surface Directed Spinodal Decomposition. *Macromolecular Theory and Simulations*, 14(7):435–443, 2005.

[14] Michael M. Henry, Matthew Lewis Jones, Stefan D. Oosterhout, Wade A. Braunecker, Travis W. Kemper, Ross E. Larsen, Nikos Kopidakis, Michael F. Toney, Dana C. Olson, and Eric Jankowski. Simplified Models for Accelerated Structural Prediction of Conjugated Semiconducting Polymers. *The Journal of Physical Chemistry C*, 121(47):26528–26538, 2017.

[15] William G. Hoover. Canonical Dynamics: Equilibrium Phase-Space Distributions. *Physical Review A*, 31(3):1695–1697, 1985.

[16] David M. Huang, Roland Faller, Khanh Do, and Adam J. Moulé. Coarse-Grained Computer Simulations of Polymer/Fullerene Bulk Heterojunctions for Organic Photovoltaic Applications. *Journal of Chemical Theory and Computation*, 6(2):526–537, 2010.

[17] Helgi I. Ingólfsson, Cesar A. Lopez, Jaakko J. Uusitalo, Djurre H. de Jong, Srinivasa M. Gopal, Xavier Periole, and Siewert J. Marrink. The Power of

Coarse Graining in Biomolecular Simulations. *Wiley Interdisciplinary Reviews: Computational Molecular Science*, 4(3):225–248, 2014.

[18] Ahmed E. Ismail, George Stephanopoulos, and Gregory C. Rutledge. Topological Coarse Graining of Polymer Chains Using Wavelet-Accelerated Monte Carlo. II. Self-Avoiding Chains. *The Journal of Chemical Physics*, 122(23):234902, 2005.

[19] Nicholas E. Jackson, Kevin L. Kohlstedt, Brett M. Savoie, Monica Olvera de la Cruz, George C. Schatz, Lin X. Chen, and Mark A. Ratner. Conformational Order in Aggregates of Conjugated Polymers. *Journal of the American Chemical Society*, pages 6254–6262, 2015.

[20] Eric Jankowski, Hilary S. Marsh, and Arthi Jayaraman. Computationally Linking Molecular Features of Conjugated Polymers and Fullerene Derivatives to Bulk Heterojunction Morphology. *Macromolecules*, 46(14):5775–5785, 2013.

[21] M. L. Jones, D. M. Huang, B. Chakrabarti, and Chris Groves. Relating Molecular Morphology to Charge Mobility in Semicrystalline Conjugated Polymers. *Journal of Physical Chemistry C*, 120(8):4240–4250, 2016.

[22] Matthew Lewis Jones and Eric Jankowski. Computationally Connecting Organic Photovoltaic Performance to Atomistic Arrangements and Bulk Morphology. *Molecular Simulation*, 43(10-11):1–18, 2017.

[23] William L. Jorgensen, Jeffry D. Madura, and Carol J. Swenson. Optimized Intermolecular Potential Functions for Liquid Hydrocarbons. *Journal of the American Chemical Society*, 106(22):6638–6646, 1984.

[24] William L. Jorgensen and Julian Tirado-Rives. The OPLS [Optimized Potentials for Liquid Simulations] Potential Functions for Proteins, Energy Minimizations for Crystals of Cyclic Peptides and Crambin. *Journal of the American Chemical Society*, 110(6):1657–1666, 1988.

[25] Hojin Kim, Dmitry Bedrov, and Grant D. Smith. Molecular Dynamics Simulation Study of the Influence of Cluster Geometry on Formation of C60 Fullerene Clusters in Aqueous Solution. *Journal of Chemical Theory and Computation*, 4(2):335–340, 2008.

[26] Christoph Klein, János Sallai, Trevor J. Jones, Christopher R. Iacovella, Clare McCabe, and Peter T. Cummings. A Hierarchical, Component Based Approach to Screening Properties of Soft Matter. *Foundations of Molecular Modeling and Simulation*, 2016.

[27] Sangwon Ko, Eric T. Hoke, Laxman Pandey, Sanghyun Hong, Rajib Mondal, Chad Risko, Yuanping Yi, Rodrigo Noriega, Michael D. McGehee, Jean-Luc Brédas, Alberto Salleo, and Zhenan Bao. Controlled Conjugated Backbone Twisting for an Increased Open-Circuit Voltage while Having a High Short-Circuit Current in Poly(hexylthiophene) Derivatives. *Journal of the American Chemical Society*, 134(11):5222–5232, 2012.

[28] Yi-kang Lan and Ching-i Huang. A Theoretical Study of the Charge Transfer Behavior of the Highly Regioregular Poly-3-hexylthiophene in the Ordered State. *The Journal of Physical Chemistry B*, 112(47):14857–14862, 2008.

[29] Thomas J. Lane, Diwakar Shukla, Kyle A. Beauchamp, and Vijay S. Pande. To Milliseconds and Beyond: Challenges in the Simulation of Protein Folding. *Current Opinion in Structural Biology*, 23(1):58–65, 2013.

[30] Cheng-Kuang Lee, Chun-wei Pao, and Chih-Wei Chu. Multiscale Molecular Simulations of the Nanoscale Morphologies of P3HT:PCBM Blends for Bulk Heterojunction Organic Photovoltaic Cells. *Energy & Environmental Science*, 4(10):4124–4132, 2011.

[31] Stephen Loser, Sylvia J. Lou, Brett M. Savoie, Carson J. Bruns, Amod Timalsina, Matthew J. Leonardi, Jeremy Smith, Tobias Harschneck, Riccardo Turrisi, Nanjia Zhou, Charlotte L. Stern, Amy A. Sarjeant, Antonio Facchetti, Robert P. H. Chang, Samuel I. Stupp, Mark A. Ratner, Lin X. Chen, and Tobin J. Marks. Systematic Evaluation of StructureProperty Relationships in Heteroacene Diketopyrrolopyrrole Molecular Donors for Organic Solar Cells. *Journal of Materials Chemistry A*, 5(19):9217–9232, 2017.

[32] Benjamin P. Lyons, Nigel Clarke, and Chris Groves. The Quantitative Effect of Surface Wetting Layers on the Performance of Organic Bulk Heterojunction Photovoltaic Devices. *The Journal of Physical Chemistry C*, 115(45):22572–22577, 2011.

[33] C. F. N. Marchiori and M. Koehler. Density Functional Theory Study of the Dipole Across the P3HT:PCBM Complex: the Role of Polarization and Charge Transfer. *Journal of Physics D: Applied Physics*, 47(21):215104, 2014.

[34] Hilary S. Marsh, Eric Jankowski, and Arthi Jayaraman. Controlling the Morphology of Model Conjugated Thiophene Oligomers through Alkyl Side Chain Length, Placement, and Interactions. *Macromolecules*, 47(8):2736–2747, 2014.

[35] Marcus G. Martin and J. Ilja Siepmann. Transferable Potentials for Phase Equilibria. 1. United-Atom Description of n-Alkanes. *Journal of Physical Chemistry B*, 102(14):2569–2577, 1998.

[36] Katherine A. Mazzio and Christine K Luscombe. The Future of Organic Photovoltaics. *Chemical Society Reviews*, 44(1):78–90, 2015.

[37] Evan D. Miller, Michael M. Henry, and Eric Jankowski. Diffractometer, 2018.

[38] Evan D. Miller, Matthew L. Jones, Michael M. Henry, and Eric Jankowski. Molecular Dynamics Data for Optimization and Validation of Modeling Techniques for Predicting Structures and Charge Mobilities of P3HT, 2018.

[39] Evan D. Miller, Matthew L. Jones, Michael M. Henry, and Eric Jankowski. Poly-(3-hexylthiophene) Model and Code for Molecular Dynamic Simulations, 2018.

[40] Evan D. Miller, Matthew L. Jones, and Eric Jankowski. Enhanced Computational Sampling of Perylene and Perylothiophene Packing with Rigid-Body Models. *ACS Omega*, 2(1):353–362, 2017.

[41] Gregory M. Newbloom, Katie M. Weigandt, and Danilo C. Pozzo. Structure and Property Development of Poly(3-hexylthiophene) Organogels Probed with Combined Rheology, Conductivity and Small Angle Neutron Scattering. *Soft Matter*, 8(34):8854, 2012.

[42] Trung Dac Nguyen, Carolyn L. Phillips, Joshua A. Anderson, and Sharon C. Glotzer. Rigid Body Constraints Realized in Massively-parallel Molecular Dynamics on Graphics Processing Units. *Computer Physics Communications*, 182(11):2307–2313, 2011.

[43] Shigeaki Obata and Yukihiro Shimoi. Control of Molecular Orientations of Poly(3-hexylthiophene) on Self-Assembled Monolayers: Molecular Dynamics Simulations. *Physical Chemistry Chemical Physics*, 15(23):9265, 2013.

[44] Jong Hwan Park, Jong Soo Kim, Ji Hwang Lee, Wi Hyoung Lee, and Kilwon Cho. Effect of Annealing Solvent Solubility on the Performance of Poly(3-hexylthiophene)/Methanofullerene Solar Cells. *The Journal of Physical Chemistry C*, 113:17579–17584, 2009.

[45] Peter Peumans, Soichi Uchida, and Stephen R. Forrest. Efficient Bulk Heterojunction Photovoltaic Cells Using Small-Molecular-Weight Organic Thin Films. *Nature*, 425(6954):158–162, 2003.

[46] Carolyn L. Phillips and Gregory A. Voth. Discovering Crystals Using Shape Matching and Machine Learning. *Soft Matter*, 9(35):8552, 2013.

[47] Benoît Roux and Thomas Simonson. Implicit Solvent Models. *Biophysical Chemistry*, 78(1-2):1–20, 1999.

[48] C. Scherer and D. Andrienko. Comparison of Systematic Coarse-Graining Strategies for Soluble Conjugated Polymers. *European Physical Journal: Special Topics*, 225(8-9):1441–1461, 2016.

[49] Hyeyoung Shin, Tod A. Pascal, William A. Goddard, and Hyungjun Kim. Scaled Effective Solvent Method for Predicting the Equilibrium Ensemble of Structures with Analysis of Thermodynamic Properties of Amorphous Polyethylene Glycol-Water Mixtures. *Journal of Physical Chemistry B*, 117(3):916–927, 2013.

[50] Hythem Sidky, Yamil J. Colón, Julian Helfferich, Benjamin J. Sikora, Cody Bezik, Weiwei Chu, Federico Giberti, Ashley Z. Guo, Xikai Jiang, Joshua Lequieu, Jiyuan Li, Joshua Moller, Michael J. Quevillon, Mohammad Rahimi, Hadi Ramezani-Dakhel, Vikramjit S. Rathee, Daniel R. Reid, Emre Sevgen, Vikram Thapar, Michael A. Webb, Jonathan K. Whitmer, and Juan J. De Pablo. SSAGES: Software Suite for Advanced General Ensemble Simulations. *Journal of Chemical Physics*, 148(4), 2018.

[51] W. C. Swope, H. C. Andersen, P. H. Berens, and K. R. Wilson. A Computer Simulation Method for the Calculation of Equilibrium Constants for the For-

mation of Physical Clusters of Molecules: Application to Small Water Clusters. *Journal of Chemical Physics*, 76(1):637–649, 1982.

[52] Ellen Van, Matthew Jones, Eric Jankowski, and Olga Wodo. Using Graphs to Quantify Energetic and Structural Order in Semicrystalline Oligothiophene Thin Films. *Molecular Systems Design & Engineering*, 1:273–277, 2018.

[53] Reid C. Van Lehn, Prabhani U. Atukorale, Randy P. Carney, Yu-Sang Yang, Francesco Stellacci, Darrell J. Irvine, and Alfredo Alexander-Katz. Effect of Particle Diameter and Surface Composition on the Spontaneous Fusion of Monolayer-Protected Gold Nanoparticles with Lipid Bilayers. *Nano Letters*, 13(9):4060–4067, 2013.

[54] Olga Wodo and Baskar Ganapathysubramanian. Modeling Morphology Evolution During Solvent-Based Fabrication of Organic Solar Cells. *Computational Materials Science*, 55:113–126, 2012.

[55] Olga Wodo, Srikanta Tirthapura, Sumit Chaudhary, and Baskar Ganapathysubramanian. A Graph-Based Formulation for Computational Characterization of Bulk Heterojunction Morphology. *Organic Electronics*, 13(6):1105–1113, 2012.

[56] Weiyu Xie, Y. Y. Sun, S. B. Zhang, and John E. Northrup. Structure and Sources of Disorder in Poly(3-hexylthiophene) Crystals Investigated by Density Functional Calculations with van der Waals Interactions. *Physical Review B - Condensed Matter and Materials Physics*, 83(18):1–6, 2011.

[57] Pengfei Xuan, Yueli Zheng, Sapna Sarupria, and Amy Apon. SciFlow: A Dataflow-Driven Model Architecture for Scientific Computing using Hadoop. In *2013 IEEE International Conference on Big Data*, pages 36–44. IEEE, 2013.

[58] G. Yu, J. Gao, J. C. Hummelen, F. Wudl, and Alan J. Heeger. Polymer Photovoltaic Cells: Enhanced Efficiencies via a Network of Internal Donor-Acceptor Heterojunctions. *Science*, 270(5243):1789–1791, 1995.

CHAPTER 5

STRUCTURAL AND CHARGE TRANSPORT PREDICTIONS FOR DIBENZO-TETRAPHENYL PERIFLANTHENE AND FULLERENE MIXTURES

5.1 Introduction

In this chapter we begin our transition from answering the question "how can we self-assemble OPV structures?" to "what self-assembled structures are good for efficient devices?" To link these questions, we self-assemble morphologies containing dibenzo-tetraphenyl-periflanthene (DBP)–a perylene derivative which has been of significant interest in recent years–then calculate charge mobilities through these structures. DBP is frequently identified as amorphous and horizontally oriented in devices[27], however the crystalline structure has been observed to be triclinic[6]. DBP has been used to create devices such as transistors[4], organic light emitting diodes[22] and also high-efficiency OPVs. When used to create an OPV, DBP can been utilized as an electron acceptor such was the case in the work of Bartynski[3], in which the highest power conversion efficiency η obtained was 2.8%. More often, DBP acts as an electron donor in devices, where it is combined with a fullerene derivative electron acceptor[7,13,28,30,31]. These devices have obtained remarkable $\eta = 8.1\%$ for a single junction device (with compound buffers)[29] and $\eta = 11.1\%$ for tandem devices[5].

In terms of photo-electronic properties, DBP is desirable due to complementary absorbency with fullerene derivative C_{70}, a widely used electron acceptor–completely covering the spectral wavelength from near UV to 650 nm.[28] This complementary absorption allows for the ability to create thin but highly absorbing devices[30]. DBP has a peak absorption at 585 nm,[22] stemming from a Highest Occupied Molecular Orbital (HOMO) and a Lowest Occupied Molecular Orbital (LUMO) of -5.5 eV and -3.5 eV, respectively[28]. Additionally, electronic measurements have also determined that DBP has a reorganization energy of 0.17 eV[3], and a measured hole mobility of 8.3×10^{-4} cm^2/Vs[31]. However, these electronic properties have been shown to change under heating or solvent annealing of the material, due to increases in crystallinity. In some cases, this can enhance the electronic properties[3], but in the case of OPVs can also decrease the overall photoconversion efficiencies stemming from too-large crystalline domains[10]. Currently, there has been no study to investigate how crystallinity affects charge transport properties, or the processing methods that would achieve optimal crystallinity.

To investigate crystallinity effects, we use MD to determine the thermodynamically stable structures for mixtures of DBP and four common fullerene derivatives at multiple solvent strengths, temperatures and solvent amounts to create a library of morphologies with differing crystallinity. We then conduct KMC simulations on these morphologies to link structural features to charge transport behavior. The report will be as follows: in Section 5.2 we will detail the force fields and simulation protocols used in the MD simulations. In Sections 5.3.1 and 5.3.2 we discuss the structural predictions of these simulations, by comparing simulations to experiment and quantifying the self-assembly behavior of our simulations. In Section 5.3.3 we present the charge transport behaviors of these structures and discuss the factors

affecting charge transport before concluding.

5.2   Methods

In this work we simulate binary mixtures of the electron donor species DBP, with the four fullerene electron acceptors: C_{60}-Fullerene, C_{70}-Fullerene, Phenyl-C_{60}-butyric acid methyl ester ($PC_{60}BM$), and Phenyl-C_{70}-butyric acid methyl ester ($PC_{70}BM$). The thermodynamically favored structures of these mixtures is determined at a variety of temperatures (T), densities (ρ) and solubilities (ε_s) in which a value of 0 corresponds to complete miscibility and higher values to non-miscibility. Molecular dynamic simulations with periodic boundary conditions are performed using the HOOMD-blue 1.3 simulation package on Tesla P100 and K20 graphical processing units in the canonical ensemble (constant number of particles, volume, and temperature)[1,8]. The Nosè-Hoover thermostat is used to regulate temperature[14]. Particle positions and velocities are updated with two-step velocity-Verlet integration of Newton's equations of motion[24].

Simulations are either initialized from a configuration with random molecular positions and orientations or a previous run. Those started from random configurations are run to equilibration and then at least 5 independent snapshots are obtained, as determined by the time series of potential energy, and $\sim 1 \times 10^9$ timesteps in total. In the runs initialized from a previous run, we simulate evaporation by shrinking the simulation volume, thereby increasing the density. Density is increased in steps of $\Delta\rho = 0.15$ g/cm^3 over 2×10^7 timesteps, then allowing 3×10^7 additional timesteps to re-equilibrate.

Two molecular quantity simulation sizes are done in this work: a phase sweep of

T, ρ and ε_s containing 100 DBP and 100 C_{70} molecules and evaporating simulations of 500 DBP and 500 C_{60}, C_{70}, $PC_{60}BM$ or $PC_{70}BM$ molecules at select statepoints. Previous work has shown that 200 molecules is sufficient size to observe stacking behaviors, however, we take a more conservative approach when studying the charge carrier movement[20]. In all simulations, an implicit solvent is used to reduce the computational burden. Additionally, all simulation elements are considered charge-neutral.

The units in the simulation are dimensionless, having been normalized by the fundamental units of length: $\sigma_{LJ} = 3.905$ Å, energy: $\varepsilon = 0.210$ kcal/mol and mass: $M = 16.0$ amu. From these fundamental units, the time unit $\tau_s = 1.67 \times 10^{-12}$ s (see Equation 3.1), which produces a timestep $dt = 0.001\tau_s = 1.67$ fs and ~ 1.67 μs total of simulation time.

The molecular topologies, i.e. atom positions and bonds are constructed using the Avogadro-1.1.1 program[12]. Non-bonded interactions are adapted from the OPLS-UA forcefields, in which explicit consideration of the hydrogen atoms is neglected and the other interactions are tuned to implicitly consider their presence[17]. The interaction parameters are given in Table 5.1. The quantities given in Table 5.1 were adapted

Element	σ_{LJ} (Å)	M (amu)	ε (kcal/mol)
Aliphatic Carbon (CT)	3.905	15.0	0.066
Aromatic Carbon (CA)	3.80	13.0	0.70
Fullerene Carbon (FCA)	3.0	12.0	0.066
Oxygen (O)	2.96	16.0	0.210

Table 5.1: OPLS-UA non-bonded parameters for the simulation elements.

directly from the OPLS-UA forcefield except for the FCA which size is tuned to better match the scattering patterns seen experimentally. Bonded interactions (bonds,

angles and dihedrals) are either considered to be part of rigid bodies or are modelled by harmonic potentials. These bonded potentials are presented in full in Appendix D. Within rigid bodies, the bonds, angles and dihedrals between simulation elements are all constrained to be constant, and the implementation is described in full by Nguyen et al.[21]. By constraining these interactions, the calculations between these elements can be neglected, thereby increasing computational performance. These approximations have previously been shown to have a negligible effect on the morphology[20]. Specifically, the following bodies are considered rigid: phenyl rings, fullerene cages and periflanthene cores (see Figure 5.1).
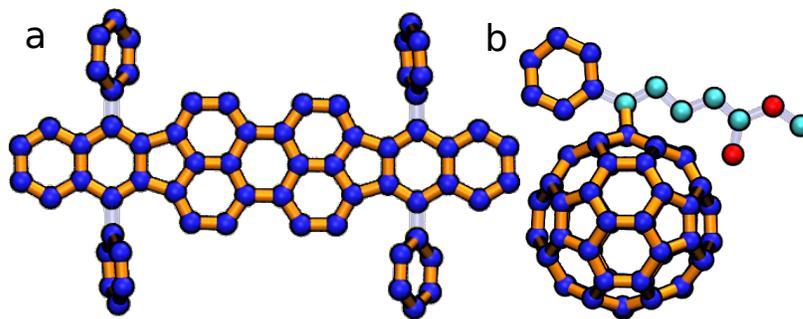
Figure 5.1: The MD models used for (a) DBP and (b) $PC_{60}BM$. Aromatic carbons atom types are shown with blue, aliphatic carbons with cyan, and oxygen with red. Rigid bonds are identified with thick, orange bonds and flexible bonds with light blue.

Two structural metrics are used to quantify molecular packing: degree of stacking and simulated GIXS characterization. The degree of stacking is used to quantify the ratio of DBP molecules that are comprised into stacks, whereas the GIXS patterns are used to identify periodic length scales and symmetries. The methodology for quantifying structure is presented in full in Section 3.2.2, and as such, the reader is invited to see this work for a more detailed explanation[20]. Briefly, in order to consider two DBP molecules stacked, their centers of mass must be within $2 \sigma_{LJ}$ of each other. This value is calibrated from the radial distribution function of ordered, neat DBP.

Additionally, no orientational requirement is applied to consider two molecules as stacked - simply being within this distance cut-off is only possible if the two molecules are stacked co-facially.

The theory behind the charge transport simulations are presented in Section 2.2.2. In these KMC simulations, we utilize simulations containing 500 DBP molecules and 500 C_{60}, $PC_{60}BM$, C_{70} or $PC_{70}BM$ molecules in which each molecule is considered a chromophore. Each of these simulations is at a density of ~ 1.5 g/cm^3, which is an interpolation between neat DBP and C_{60} densities. Further, the structures used for charge transport were created over the temperature range of ~ 150 to 650 K, in steps of ~ 100 K (although the KMC temperature is constant at 293 K). As such, six MD temperatures and four acceptor type combinations are explored, resulting in 24 state points. For each temperature-acceptor combinations, five independent samples are generated (through applying the autocorrelation function to the potential energy), totaling 120 charge transport simulations.

For each simulation, the morphology is returned to the atomistic from the united-atom representation. Here, fine-graining is somewhat trivial, as the hydrogens can be placed around the appropriate beads based on sp hybridization rules and typical element-hydrogen bond-lengths. Using the atomistic representation, we then use Zerner's neglect of differential overlap (ZINDO/S) used to predict the energy levels[18,23]. In applying the Marcus hopping in Section 2.2.2, we utilize constants found in literature for the reorganization energy λ. As such, λ is taken to be 0.17 eV for DBP from the work of Bartynski[3], and 0.15 eV for the fullerene derivatives[15]. In the simulations, 100,000 charges are placed within the simulation and are allowed to run for five different times 1×10^{-12} to 1×10^{-8} (in steps of 1×10^{-1}) s.

5.3 Results And Discussions

5.3.1 Neat Systems
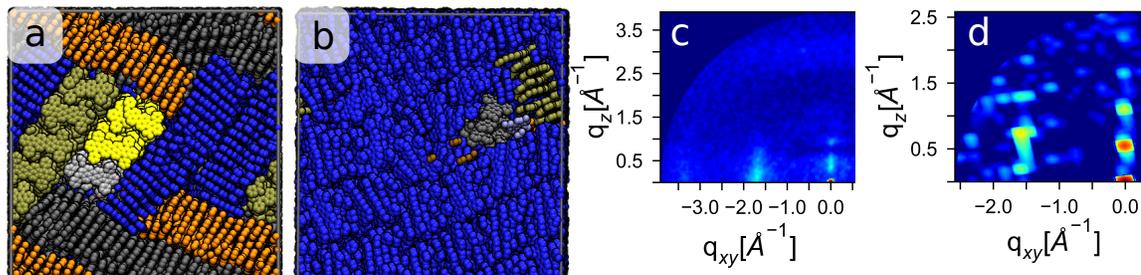
Figure 5.2: The thermodynamically favored structure for neat DBP is predicted to be π-stacked columns, which exist both as (a) multiple orientations and (b) primarily one orientation with defects. These stacks are identified through the clustering analysis and colored based on the cluster. (c) The GIXS pattern of (b), averaged over 9 reoriented GIXS patterns chosen from the presence of features. (d) The GIXS pattern of experimental DBP. This pattern was generated by replicating the crystal structure reported by Debad[6] then using the simulated diffraction program on the structure.

Simulated structures at $\sim$ 150 and 350 K and experimental density 1.25 g/cm^3 for neat DBP are shown in Figure 5.2a and b. These structure are dominated by columns of π-stacked DBP molecules, which exist with a variety of column orientations (a) or a single crystal with small, defect crystallites (b). As a first step, the neat DBP structure in (b) is compared to the experimentally observed crystal structure by comparing GIXS patterns. To compare the simulated and experimental systems, crystal data must be generated as no neat DBP GIXS pattern was available. To overcome the absence of a pattern, a perfect crystal structure is taken from the Cambridge Crystallographic Database. This crystal structure, reported by Debad[6], is a triclinic crystal system with lattice parameters of 10.6, 11.5 and 13.0 Å and angles of 95.0, 111.2, 100.5°. Using these lattice parameters, a super-cell is created by duplicating the cell in the three different dimensions using Avogadro's super-cell

builder[12] The in-house diffraction software is then used on the super-cell to create the reference pattern[16].

The simulated and experimental patterns are shown in Figure 5.2c and d, respectively. The simulated and experimental patterns show quantitatively similar features with a peak along the q_{xy} axis around ~ 1.6 Å^{-1} and peaks along the q_z axis around ~ 0.5 Å^{-1}. The peak along q_{xy}, corresponds to π-stacking with real distances of ~ 3.7 Å. The peaks along q_z, correspond to intercolumnar distances. Here, the simulation overpredicts the intercolumnar stacking slightly with the simulated stacking at ~ 0.5 Å$^{-1} \approx 12.6$ Å and the experimental at ~ 0.6 Å$^{-1} \approx 10.5$ Å. This differences arises from the difference in stacking angle between the experimental and simulated morphologies, similar to what was observed in perylene, DBP stacks in a β-herringbone conformation which is not obtained in our simulation[20].

It is also useful to compare the structure of DBP to that of five-membered ring perylene, and sulfated perylene: perylothiophene conducted in previous work (Section 3). DBP similar to perylene and perylothiophene forms discotic columns of molecules in an aligned conformation. However, DBP unlike perylene and perylothiophene, which have two columnar phases: a rotationally aligned and non-aligned phase, DBP only presents an aligned phase. This is due to the presence of the phenyl rings extending from the periflanthene core, which only allow for molecules to form columns when the molecules are rotated by 0 or 90° relative to each other. However, when the molecules are aligned, the π-overlap is maximized, and thereby this is the most favored conformation.

Turning now from DBP to PC$_{60}$BM and PC$_{70}$BM, the patterns of the simulated and experimental as-cast GIXS are shown in Figures 5.3a-d. All of these patterns exhibit a strong, radially symmetric band at ~ 1.4 Å^{-1}, corresponding to a distance of
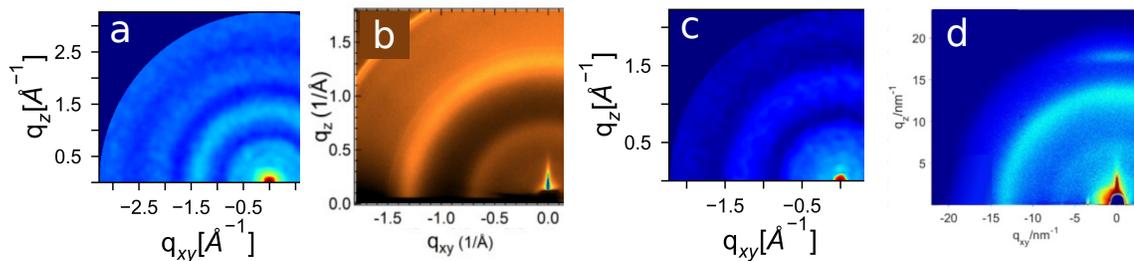
Figure 5.3: (a) The GIXS pattern of simulated $PC_{60}BM$. (b) The experimental pattern of $PC_{60}BM$, adapted from Ref [9] with permission from The Royal Society of Chemistry. (c) The GIXS pattern of simulated $PC_{70}BM$. (d) The experimental pattern of $PC_{70}BM$, adapted with permission from Ref [19]. Copyright 2018 American Chemical Society.

~ 4.5 Å, with $PC_{70}BM$ extending to shorter wavenumbers. These patterns correspond to amorphous structures, with no long range order. However, it should be noted, that crystallinity has been observed in thermally annealed $PC_{60}BM$ neat films[26].

5.3.2 Mixtures

With the neat DBP and fullerene systems showing experimental accuracy, the effect of mixing these systems together is now investigated. As such, DBP is mixed with C_{60}, C_{70}, $PC_{60}BM$ or $PC_{70}BM$ in 1:1 molar ratios and simulated at different temperatures ($160 \rightarrow 670$ K, in steps of ~ 100 K) and with each system shrunk from $\rho = 1.22$ to 1.53 g/cm^3 in steps of 0.15 g/cm^3. As is seen in Figure 5.4a-d, each of the systems has extensive π-stacked columns of DBP molecules. The functionalized, $PC_{60/70}BM$ systems phase separate more strongly at moderate temperatures (370 K) than that of the non-functionalized systems. This can be seen with visual inspection of Figures 5.4a-d. The separation is also seen in the GIXS patterns of the DBP molecules (fullerene atoms omitted from GIXS) in Figures 5.4e and f in which the non-functionalized C_{60} containing system in Figure 5.4e does not show periodic features along q_z, whereas
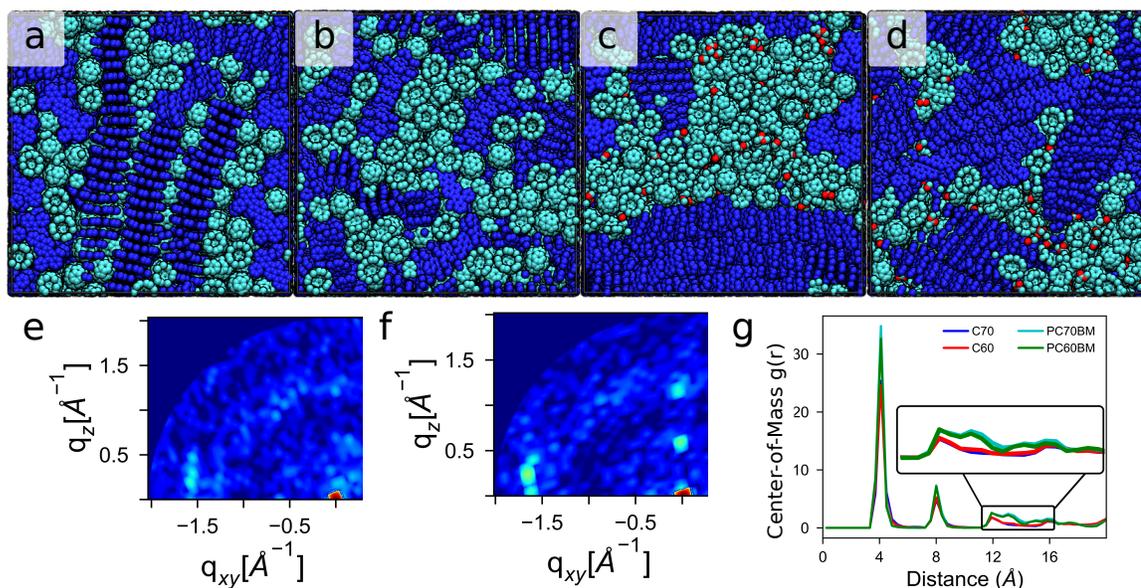
Figure 5.4: Moderate temperature (~ 370 K) morphologies for (a) C_{60}, (b) C_{70}, (c) $PC_{60}BM$ and (d) $PC_{70}BM$ showing that the functionalized fullerenes phase separate. This difference in phase-separation can be observed in the GIXS patterns of (e) C_{60} and (f) $PC_{60}BM$: only (f) shows periodic features along the q_z axis. Additionally, this difference in phase-separation can be seen in the $g(r)$ with a difference over the range 12 - 16 Å, corresponding to intercolumnar distances.

the $PC_{60}BM$ containing system in Figure 5.4f does. Lastly, in the $g(r)$ in Figure 5.4g in which $PC_{60/70}BM$ diverge from $C_{60/70}$ over the range of 12-16 Å, also shows this separation. The higher degree of phase-separation for the functionalized systems is perhaps unexpected as fullerenes are often functionalized to increase the solubility for polymeric systems[11].

In higher temperature systems (590 K), π-stacked columns are still present in the morphology (see Figure 5.5a) rather than being completely disordered. The presence of π-stacked columns, but not phase-separated DBP in the functionalized systems, can be seen in the GIXS pattern of $PC_{60}BM$ in Figure 5.5b: there are faint peaks at ~ 1.6 Å^{-1} along the q_{xy} axis, but no repeated features along the q_z axis. The absence
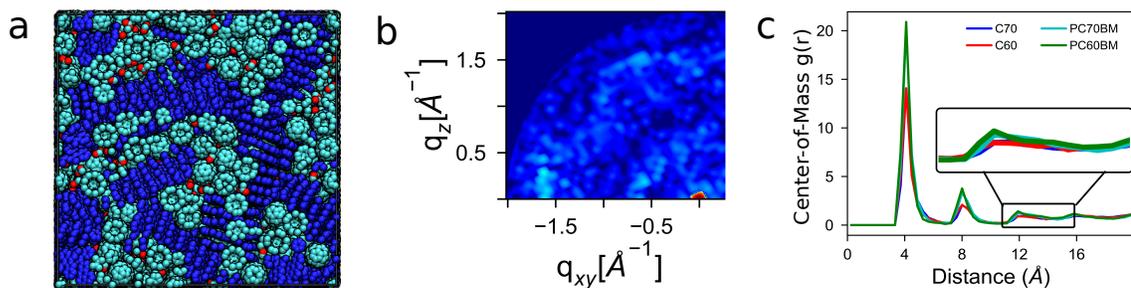
Figure 5.5: (a) The higher temperature (590 K) $PC_{60}BM$ system has lost the aggregated columns observed at lower temperatures, which absence is also seen in the GIXS pattern (b). (c) The $g(r)$ for the different systems at a high temperature no longer show a distinction between any of the functionalized and non-functionalized fullerene systems.

of the phase-separation is also seen in the $g(r)$ (Figure 5.5c) in which there are still peaks at ~ 4 Å increments indicating π-stacking, but the distinction seen in Figure 5.4g over the range 12-16 Å is no longer observed. The continuation of π-stacking is due to the density induced self-assembly that occurs as the system shrinks to higher densities, and in order to obtain this density, short-range ordering must occur.
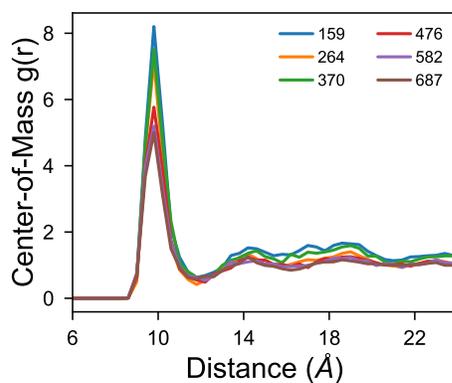
Figure 5.6: The $g(r)$ of $PC_{60}BM$ shows a small decrease in the first peak (10 Å) as a function of temperature, indicating there is no large change.

In addition to investigating the structure of DBP as a function of temperature, the

structure of the fullerene derivatives is also considered. Figure 5.6 shows the $g(r)$ of $PC_{60}BM$ as a function of temperature. All of these systems are primarily dominated by a nearest-neighbor peak at ~ 10 Å. Additionally, Figure 5.6 shows that there is a small decrease in the amount of order as a function of temperature. More specifically, at temperatures ≤ 370 K the nearest-neighbor peak is slightly greater than at higher temperatures. This change in intensity is likely due to the change in phase separation that occurs in the $PC_{60}BM$ system. However, the small changes as a function of temperature is not surprising as the structure was observed to be amorphous in the GIXS patterns in Figures 5.3a-d. And as such, not much discussion will be spent on the structure of the fullerene systems.

We now turn to investigating the structure behavior of DBP with C_{70} at a variety of T, ρ, and ε_s. For these tests we use a ratio of 1:1 for DBP and C_{70}, and choose C_{70} because it is the most commonly used acceptor combination in literature. Additionally, to quantify the structure we use the clustering metric ψ, which measured the number of large π-stacked DBP columns within the simulation as was done in Section 4.3.2. Note, contrary to how ψ was used in combination with $g(r)$ and ξ in Section 3 we do not identify distinct phases, rather, we purely utilize ψ as our structural descriptor as was done with P3HT.

Figure 5.7, shows the order ψ as a function of the three thermodynamic variables T, ρ, and ε_s in which red colors indicate more order and blue little order. We see that the highest degree of ordering occurs at lower ρ, stronger solvents (low ε_s), and lower temperatures. Order occurring at low ρ is unsurprising as in Section 3 we observed that high densities result in much more kinetic trapping. Additionally, this kinetic trapping is exacerbated since the rigid bodies within DBP are larger than those in perylene and perylothiophene. The optimal conditions for assembly are moderately
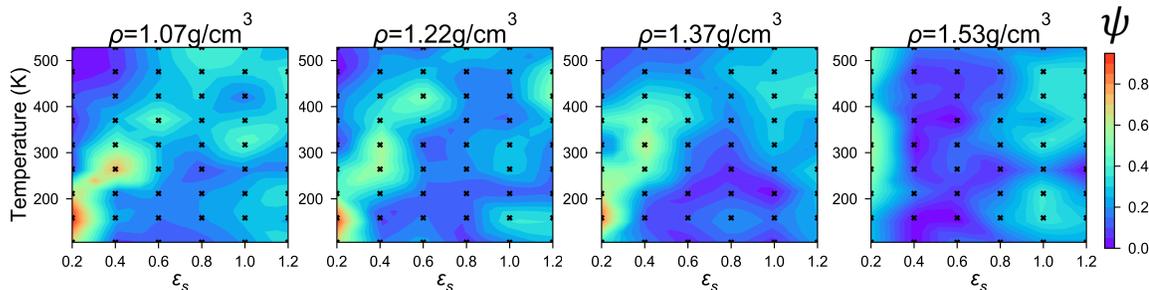
Figure 5.7: Heatmaps showing ψ for the various ρ, ε_s and T. Red corresponds to more ordering whereas blue corresponds to disorder. Black 'x's indicate where simulations were run and data is linearly interpolating between these points.

low temperatures (150-200 K) and "good" solvents ($\varepsilon_s = 0.2$). If temperatures are too low (100 K), the systems become kinetically trapped. Further, if solvents are not sufficiently strong, $\varepsilon_s \geq 0.6$, then the periflanthene and fullerene do not have sufficient energy to break from local minima to form extended stacks. As such, the envelope in which ordering occurs is restricted to strong solvents to lubricate periflanthene movement and low temperatures to drive extended ordering, and generating devices with these structures will need to utilize these conditions. However, as seen in the Figure 5.4 systems, we note that an additional way to induce ordering is by shrinking the volume of the system, which is not explored in Figure 5.7. Therefore, additional ordering in devices would be obtained if the solvents were removed slowly, and these systems are likely to have the fastest charge transport.

5.3.3 Charge Transport

We now turn to comparing the structures shown in Figures 5.4 and 5.5 to their corresponding charge transport properties. Figure 5.8 shows the (a) electron and (b) hole mobilities: $\mu_{E,0}$ and $\mu_{H,0}$, respectively. As expected, many of the systems show decreasing charge movement movement with increasing temperature. However,
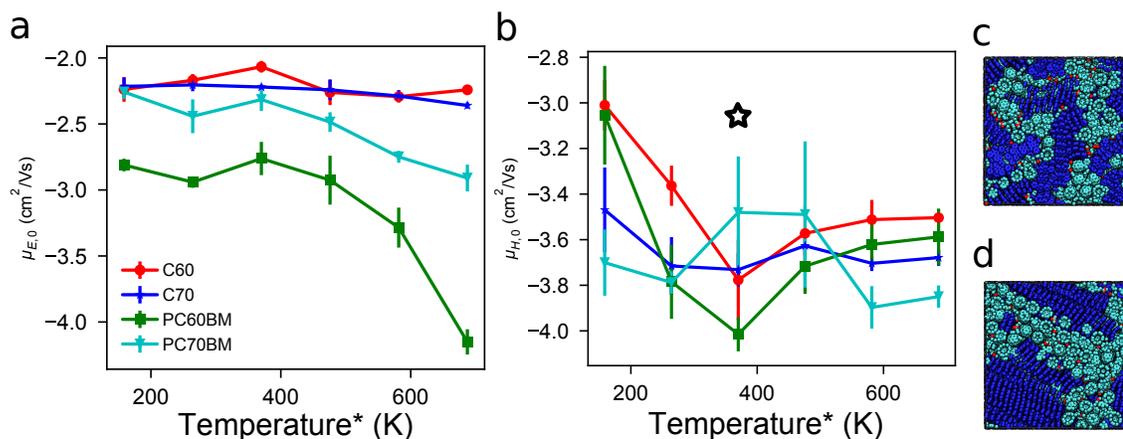
Figure 5.8: (a) The electron mobilities through the fullerene derivatives show that the non-functionalized derivatives (red and blue) do not change much over the temperature range, whereas the functionalized (cyan and greed) decrease with increasing temperature. (b) The hole mobilities through DBP primarily show a decrease in μ_0 as a function of temperature, but agree with experimentally measured mobilities (shown with a black star). This decrease is largest in the C_{60} systems, which decrease by a factor of ~ 5. The increase-then decrease behavior seen in the $PC_{70}BM$ system can be explained by kinetic trapping at low temperatures (c, ~ 250 K) then ordering at moderate temperatures (d, ~ 350 K).

$\mu_{H,0}$ does not always follow this same trend, as $\mu_{H,0}$ in the $PC_{70}BM$ system is lower $(1 \times 10^{-3.7}$ cm^2/Vs$)$ at low temperatures (~ 150 K) then increases to $1 \times 10^{-3.4}$ cm^2/Vs at ~ 370 K, before decreasing back to $1 \times 10^{-3.8}$ cm^2/Vs at ~ 600 K. This trend can be understood by identifying the kinetic trapping of the morphology at low temperatures (Figure 5.8c) which then orders at higher temperatures (Figure 5.8d), which matches the expected trend of increasing order increasing μ_0.

We now analyse $\mu_{E,0}$ and $\mu_{H,0}$ in more detail to link the morphological characteristics to charge transport. The $\mu_{E,0}$ values are close to the μ values measured experimentally in fullerene systems of 1×10^{-2} cm^2/Vs[25] (however we would expect the zero-field mobilities to be somewhat lower than the transistor mobilities). $\mu_{E,0}$ shows two different trends: the non-functionalized fullerene derivatives are fairly insensitive

to the increase in temperature, whereas the functionalized fullerenes are sensitive to temperature increases. The relative insensitivity for the non-functionalized systems is perhaps unsurprising as the fullerene cages are represented with rigid bodies and these are dense systems i.e. there is no extra space in which the fullerenes can "spread-out". As such, $J_{i,j}$ (as described in Equation 2.14) between fullerenes is unlikely to be decreased with increasing temperature because $\Delta E_{i,j}$ is constant at zero due to the rigid bodies, and the $E_{\text{LUMO}+1} - E_{\text{LUMO}}$ energy splitting will not change because the molecules are the same distance apart.
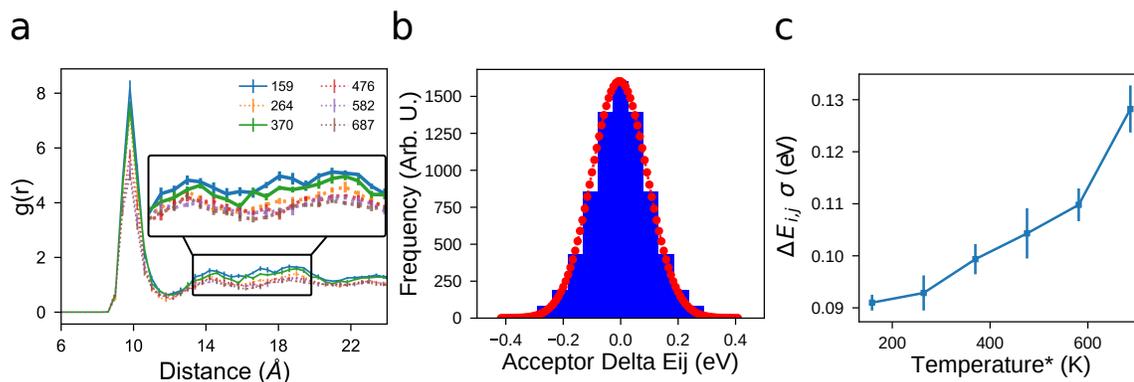
Figure 5.9: (a) The $g(r)$ for the fullerene cage centers show that there is slightly more order in the ~ 160 and 370 K systems than the other temperatures. (b) The distribution of the $\Delta E_{i,j}$ values calculated within the simulation with a fitted Gaussian curve. (c) The σ value describing the standard deviation of the $\Delta E_{i,j}$ as a function of preparation temperature for the five independent snapshots. Both (a) and (c) affect to $\mu_{E,0}$.

In the functionalized systems, we do observe a decrease in $\mu_{E,0}$ with increasing temperature. This decrease arises from structural and energetic disorder that arises at different temperatures (see Figure 5.9). For example, in Figure 5.8a, we see that the PC$_{60}$BM systems at ~ 160 and 370 K have slightly higher $\mu_{E,0} \approx 1 \times 10^{-2.3}$ cm^2/Vs. In Figure 5.9a, the $g(r)$ shows that these two systems have slightly higher ordering over the range (14,18) Å. This range is indicative of the second nearest

neighbor, and for these systems represents that these systems are more likely to have contiguous paths along which the electrons can travel. However, the other systems do not show structural changes that would explain differences in $\mu_{E,0}$. The other lens through which we can understand this difference is energetic differences between the systems. In these systems, the fullerene cages are represented with rigid bodies, however, the bonds in the aliphatic chain are flexible and can have different energy levels. The distribution of $\Delta E_{i,j}$ values for the 370 K system, is shown in Figure 5.9b and shows that at this temperature, the $\Delta E_{i,j}$ distribution is roughly Gaussian with an average at zero and a standard deviation σ of ~ 0.1 eV. We can quantify the increase in energetic disorder by comparing σ for the different temperatures, shown in Figure 5.9c. In Figure 5.9c, we see that σ increases with temperature. This increase in σ will have two effects: first as expressed by the energy-splitting-in-dimer method (Equation 2.14), this will decrease $J_{i,j}$ with increasing σ, and second, as seen in the Marcus hopping expression (Equation 2.13) this energy will affect the movement of electrons.
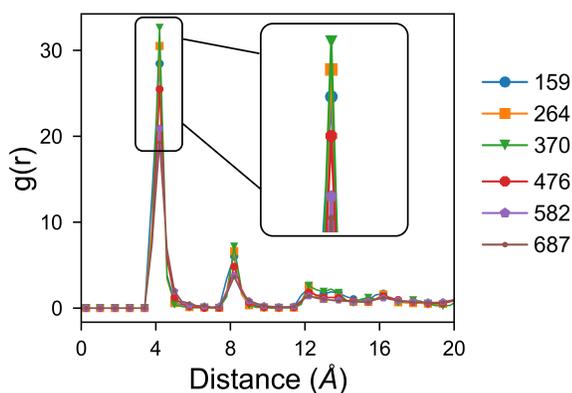
Figure 5.10: The radial distribution of DBP centers-of-mass for different temperatures within the $PC_{60}BM$ mixtures shows that the order increases from 159 K to 370 K before decreasing again.

We now shift focus from $\mu_{E,0}$ to $\mu_{H,0}$ through DBP molecules. $\mu_{H,0}$ is in fair agreement with the experimental measurements for DBP (represented with a black star in Figure 5.8b). As noted earlier, $\mu_{H,0}$ for the different systems primarily initially decreases as a function of temperature ($\sim 160 \rightarrow 370$ K) before leveling out (or even increasing) above this temperature. However, this trend in $\mu_{H,0}$ does not follow the same trend in order as a function of temperature. For example, Figure 5.10 shows that the 370 K system has the highest order. The change $\mu_{H,0}$ is unlikely due to energetic changes as was seen in the functionalized fullerene systems as there are only four flexible bonds in DBP as very few $\Delta E_{i,j}$ values are greater than 0.1 eV. However, the change in $\mu_{H,0}$ can be understood by understanding the continuity of the DBP stacks.

In Figure 5.11 we examine the DBP stacks and the movement of holes through them. Figure 5.11a shows the stacks within the ~ 160 K system. These stacks typically span the simulation volume and correspond to the observation in Figure 5.10 that these systems are ordered. In Figure 5.11b we show the net hole hops through the system; which reveals that the holes move along these π-stacked columns and that there is little movement between stacks. In Figure 5.11c we plot the displacement of the 1,000 farthest moving holes, which agrees to our observation about Figure 5.11b that charges are primarily moving along the columns and not between columns, as the displacements are primarily along the direction of the stacks. This bias in trajectory can be quantified with the anisotropy value, which when close to one indicates that all charges are moving along the same axis and when close to zero the displacement is spherically symmetric. With this meaning in mind, the 160 K system having an anisotropy value of 0.77 again supports the observation that charges are likely to move along the same axis for a long period of time.
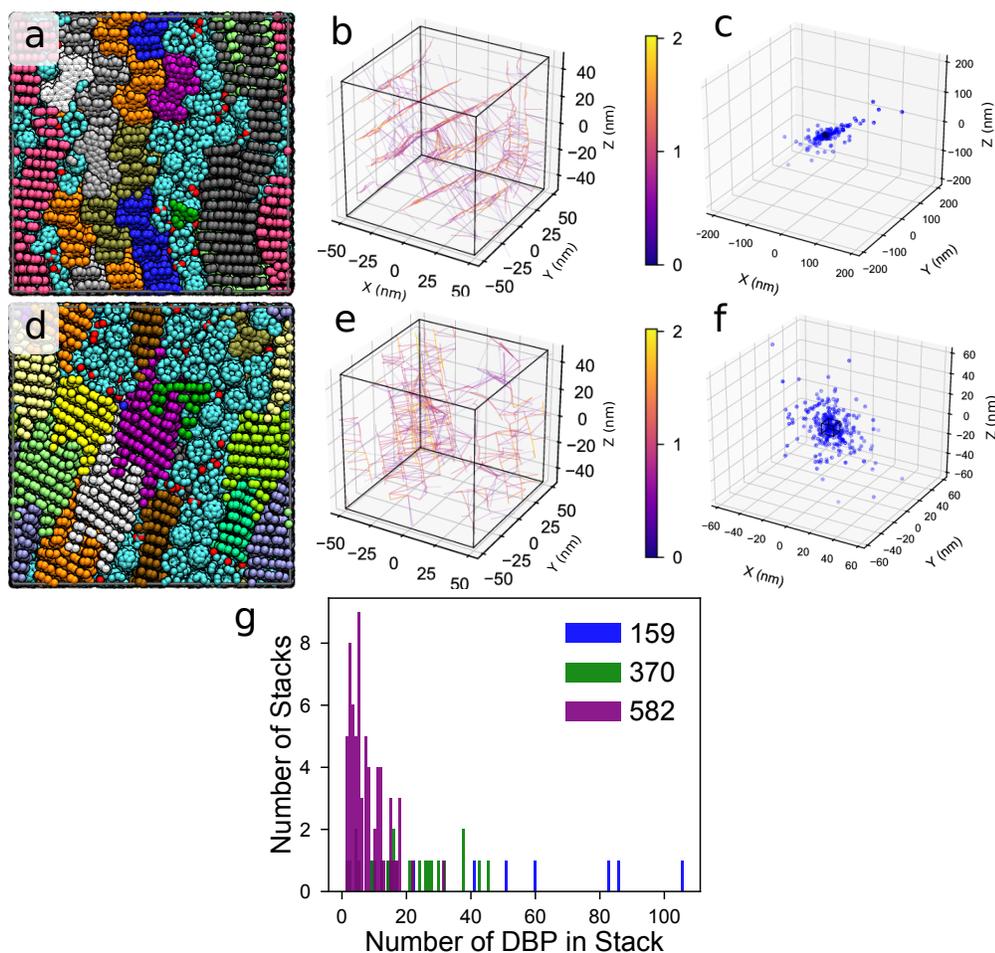
Figure 5.11: (a) The unique stacks of DBP within the simulation are colored based on stack (cyan and red are carbons and oxygen in the $PC_{60}BM$, respectively) in the ~ 160 K system. (b) The network-graph of charge movement between chromophore centers shows that holes primarily move along these stacks, but rarely between stacks in the ~ 160 K system. The color bar corresponds to the $\log_{10}$ number of net hops that occur along a path. (c) The displacement of the 1,000 farthest moving holes within the simulation within the ~ 160 K system. (d, e, f) The same analysis for the 370 K system. (g) The number of unique stacks identified within the 160, 370 and 580 K systems.

Figure 5.10 showed that the 370 K system is ordered, which is confirmed in Figure 5.11d. However, contrary to as was seen in Figure 5.11a, this system has many highly-ordered small crystallites rather than long continuous stacks. The network

graph in Figure 5.11e shows how this affects the charge transport properties: holes primarily move along the π-stacked columns but also frequently hop between columns. The reason for more frequent intercolumnar hops is that the holes can only move so far without making a slow hop to a new column. This intercolumnar hopping can also be seen in the trajectory of the holes in Figure 5.11f which shows a much more visually symmetric trajectory (anisotropy = 0.23). The anisotropy of 0.23 is much closer to the high temperature systems with have anisotropies of ~ 0.06. Further, although the system is highly ordered, the number of unique stacks is more similar to disordered systems that have many small unique stacks compared to the 160 K system which have few, large stacks (see Figure 5.11g). As such, it is the continuity of the stacks that is important for hole transport, more so than the "order" of the stacks.

We note that the importance of having contiguous stacks may cause some pitfalls in periodically bound simulations. These finite size effects come in the form of an infinitely long stack because a stack is able to link to itself across a periodic boundary. This can cause issues as a charge placed on a perfect, infinitely long stack is likely to have an unphysically high mobility. Further, because all structures within the simulation are periodic, this also means any defects will be encountered each time a charge travels through the morphology. The orientation of the stack within the simulation volume will dictate if a stack can be infinitely long or will have a periodic defect. As such, a researcher simulating poly-aromatic hydrocarbons will need to be aware of these potential periodic boundary effects (and other finite size effects[2]) and account for them. Likely the most straight forward way to do this is to simulate systems large enough that periodic features are smaller than the periodic boundaries of the simulation volume.

5.4 Conclusion

In this work, we investigated the effects of solvent strength, temperature and density on the self assembly of DBP-fullerene mixtures. If high crystallinities are desired, we found that devices should be created using good solvents ($\varepsilon_s = 0.2$) and low temperatures (100 K). Additionally, the highest crystallinities were obtained when the functionalized fullerenes were used to create morphologies. From the radial distribution function we found that the size of stacks of DBP decreases with increasing temperature, which correlates to smaller stacks. Fullerenes, in contrast, didn't show the same temperature dependence as they are fairly constant over the range of temperatures investigated.

DBP, unlike perylene and perylene derivative perylothiophene, formed only in aligned stacks rather than the discotic columns observed in previous work. Lastly, we note that our simplified (United-Atom, charge-neutral) model produced small changes in comparison to the experimentally observed structure. We found that these differences primarily emerged due to the difference in the stacking angle of the DBP backbones which resulted in slightly farther π-stackings and larger inter-columnar spacings, but still maintained similar periodic features.

We also explore the factors that govern charge transport through DBP-Fullerene mixtures as a function of temperature. Because of the rigid body representation of the fullerene cages and the constrained density of the simulations, the C_{60} and C_{70} systems are unaffected by the different temperatures. The functionalized $PC_{60}BM$ and $PC_{70}BM$ systems however are affected by temperatures. This can be in the form of having more continuous structures at lower temperatures and the decrease in hopping rates because of larger distributions of $\Delta E_{i,j}$. In DBP we are able to

predict $\mu_{H,0}$ close to experimental measurements. In these DBP systems, the most important factor for charge transport is having continuous π-stacked columns along which charges can travel. Highly ordered systems that contain many highly ordered, small(er) crystallites will have poor charge movements because of slow intercolumnar hops. Lastly, for these systems with periodic features (i.e. π-stacked columns) that have feature sizes the same as the periodic boundary lengths, periodic effects must be carefully considered as to not produce spurious results.

BIBLIOGRAPHY

[1] Joshua A. Anderson, Chris D. Lorenz, and A. Travesset. General Purpose Molecular Dynamics Simulations Fully Implemented on Graphics Processing Units. *Journal of Computational Physics*, 227(10):5342–5359, 2008.

[2] Denis Andrienko. Multiscale Concepts in Simulations of Organic Semiconductors. *Handbook of Materials Modeling*, pages 1–12, 2018.

[3] Andrew N. Bartynski, Stefan Grob, Theresa Linderl, Mark Gruber, Wolfgang Brütting, and Mark E. Thompson. Organic Solar Cells with Open Circuit Voltage over 1.25 V Employing Tetraphenyldibenzoperiflanthene as the Acceptor. *The Journal of Physical Chemistry C*, 120(34):19027–19034, 2016.

[4] W. Boukhili, M. Mahdouani, R. Bourguiga, and J. Puigdollers. Temperature Dependence of the Electrical Properties of Organic Thin-Film Transistors Based on Tetraphenyldibenzoperiflanthene Deposited at Different Substrate Temperatures: Experiment and Modeling. *Microelectronic Engineering*, 150:47–56, 2016.

[5] Xiaozhou Che, Xin Xiao, Jeramy D. Zimmerman, Dejiu Fan, and Stephen R. Forrest. High-Efficiency, Vacuum-Deposited, Small-Molecule Organic Tandem and Triple-Junction Photovoltaic Cells. *Advanced Energy Materials*, 4(18), 2014.

[6] Jeff D. Debad, Jonathan C. Morris, Vince Lynch, Philip Magnus, and Allen J. Bard. Dibenzotetraphenylperiflanthene: Synthesis, Photophysical Properties, and Electrogenerated Chemiluminescence. *Journal of the American Chemical Society*, 118(3):2374–2379, 1996.

[7] Daisuke Fujishima, Hiroshi Kanno, Toshihiro Kinoshita, Eiji Maruyama, Makoto Tanaka, Makoto Shirakawa, and Kenichi Shibata. Organic Thin-Film Solar Cell Employing a Novel Electron-Donor Material. *Solar Energy Materials and Solar Cells*, 93(6-7):1029–1032, 2009.

[8] Jens Glaser, Trung Dac Nguyen, Joshua A. Anderson, Pak Lui, Filippo Spiga, Jaime A. Millan, David C. Morse, and Sharon C. Glotzer. Strong Scaling of General-Purpose Molecular Dynamics Simulations on GPUs. *Computer Physics Communications*, 192:97–107, 2014.

[9] Enrique D. Gomez, Katherine P. Barteau, He Wang, Michael F. Toney, and Yueh Lin Loo. Correlating the Scattered Intensities of P3HT and PCBM to the Current Densities of Polymer Solar Cells. *Chemical Communications*, 47(1):436–438, 2011.

[10] Stefan Grob, Andrew N. Bartynski, Andreas Opitz, Mark Gruber, Florian Grassl, Eduard Meister, Theresa Linderl, Ulrich Hörmann, Christopher Lorch, Ellen Moons, Frank Schreiber, Mark E. Thompson, and Wolfgang Brütting. Solvent Vapor Annealing on Perylene-Based Organic Solar Cells. *Journal Material Chemistry A*, 3(30):15700–15709, 2015.

[11] Serap Günes, Helmut Neugebauer, and Niyazi Serdar Sariciftci. Conjugated Polymer-Based Organic Solar Cells. *Chemical Reviews*, 107(4):1324–1338, 2007.

[12] Marcus D. Hanwell, Donald E. Curtis, David C. Lonie, Tim Vandermeerschd, Eva Zurek, and Geoffrey R. Hutchison. Avogadro: An Advanced Semantic Chemical Editor, Visualization, and Analysis Platform. *Journal of Cheminformatics*, 4(17):1–17, 2012.

[13] Masaya Hirade and Chihaya Adachi. Small Molecular Organic Photovoltaic Cells with Exciton Blocking Layer at Anode Interface for Improved Device Performance. *Applied Physics Letters*, 99(15):2009–2012, 2011.

[14] William G. Hoover. Canonical Dynamics: Equilibrium Phase-Space Distributions. *Physical Review A*, 31(3):1695–1697, 1985.

[15] Julien Idé, Daniele Fazzi, Mosè Casalegno, Stefano Valdo Meille, and Guido Raos. Electron Transport in Crystalline PCBM-Like Fullerene Derivatives: a Comparative Computational Study. *Journal of Materials Chemistry C*, pages 7313–7325, 2014.

[16] Matthew Lewis Jones and Eric Jankowski. Computationally Connecting Organic Photovoltaic Performance to Atomistic Arrangements and Bulk Morphology. *Molecular Simulation*, 43(10-11):1–18, 2017.

[17] William L. Jorgensen and Julian Tirado-Rives. The OPLS [Optimized Potentials for Liquid Simulations] Potential Functions for Proteins, Energy Minimizations

for Crystals of Cyclic Peptides and Crambin. *Journal of the American Chemical Society*, 110(6):1657–1666, 1988.

[18] James Kirkpatrick. An Approximate Method for Calculating Transfer Integrals Based on the ZINDO Hamiltonian. *International Journal of Quantum Chemistry*, 108(1):51–56, 2008.

[19] Cheng Kun Lyu, Fei Zheng, B. Hari Babu, Meng Si Niu, Lin Feng, Jun Liang Yang, Wei Qin, and Xiao Tao Hao. Functionalized Graphene Oxide Enables a High-Performance Bulk Heterojunction Organic Solar Cell with a Thick Active Layer. *Journal of Physical Chemistry Letters*, 9(21):6238–6248, 2018.

[20] Evan D. Miller, Matthew L. Jones, and Eric Jankowski. Enhanced Computational Sampling of Perylene and Perylothiophene Packing with Rigid-Body Models. *ACS Omega*, 2(1):353–362, 2017.

[21] Trung Dac Nguyen, Carolyn L. Phillips, Joshua A. Anderson, and Sharon C. Glotzer. Rigid Body Constraints Realized in Massively-parallel Molecular Dynamics on Graphics Processing Units. *Computer Physics Communications*, 182(11):2307–2313, 2011.

[22] Kenji Okumoto, Hiroshi Kanno, Yuji Hamada, Hisakazu Takahashi, and Kenichi Shibata. High Efficiency Red Organic Light-Emitting Devices Using Tetraphenyldibenzoperiflanthene-Doped Rubrene as an Emitting Layer. *Applied Physics Letters*, 89(1):1–4, 2006.

[23] J. Ridley and Michael Zerner. An Intermediate Neglect of Differential Overlap Technique for Spectroscopy: Pyrrole and the Azines. *Theoretica Chimica Acta*, 32(2):111–134, 1973.

[24] W. C. Swope, H. C. Andersen, P. H. Berens, and K. R. Wilson. A Computer Simulation Method for the Calculation of Equilibrium Constants for the Formation of Physical Clusters of Molecules: Application to Small Water Clusters. *Journal of Chemical Physics*, 76(1):637–649, 1982.

[25] S. M. Tuladhar, D. Poplavskyy, S. A. Choulis, J. R. Durrant, D. D. C. Bradley, and J. Nelson. Ambipolar Charge Transport in Films of Methanofullerene

and Poly(phenylenevinylene)/Methanofullerene Blends. *Advanced Functional Materials*, 15(7):1171–1182, 2005.

[26] Eric Verploegen, Rajib Mondal, Christopher J Bettinger, Seihout Sok, Michael F Toney, and Zhenan Bao. Effects of Thermal Annealing Upon the Morphology of Polymer-Fullerene Blends. *Advanced Functional Materials*, 20(20):3519–3529, 2010.

[27] Julia Wagner, Mark Gruber, Alexander Hinderhofer, Andreas Wilke, Benjamin Bröker, Johannes Frisch, Patrick Amsalem, Antje Vollmer, Andreas Opitz, Norbert Koch, Frank Schreiber, and Wolfgang Brütting. High Fill Factor and Open Circuit Voltage in Organic Photovoltaic Cells with Diindenoperylene as Donor Material. *Advanced Functional Materials*, 20(24):4295–4303, 2010.

[28] Zhongqiang Wang, Daisuke Yokoyama, Xiao-Feng Wang, Ziruo Hong, Yang Yang, and Junji Kido. Highly Efficient Organic p-i-n Photovoltaic Cells Based on Tetraphenyldibenzoperiflanthene and Fullerene C70. *Energy Environ. Sci.*, 6(1):249–255, 2013.

[29] Xin Xiao, Kevin J. Bergemann, Jeramy D. Zimmerman, Kyusang Lee, and Stephen R. Forrest. Small-Molecule Planar-Mixed Heterojunction Photovoltaic Cells with Fullerene-Based Electron Filtering Buffers. *Advanced Energy Materials*, 4(7):1301557, 2014.

[30] Daisuke Yokoyama, Zhong Qiang Wang, Yong Jin Pu, Kenta Kobayashi, Junji Kido, and Ziruo Hong. High-Efficiency Simple Planar Heterojunction Organic Thin-Film Photovoltaics with Horizontally Oriented Amorphous Donors. *Solar Energy Materials and Solar Cells*, 98:472–475, 2012.

[31] Yan-qiong Zheng, William J. Potscavage, Takeshi Komino, Masaya Hirade, Junji Adachi, and Chihaya Adachi. Highly Efficient Bulk Heterojunction Photovoltaic Cells Based on C70 and Tetraphenyldibenzoperiflanthene. *Applied Physics Letters*, 102(14):143304, 2013.

CHAPTER 6

TYING TOGETHER MULTISCALE CALCULATIONS FOR CHARGE TRANSPORT IN P3HT: STRUCTURAL DESCRIPTORS, MORPHOLOGY, AND TIE-CHAINS[1]

6.1 Introduction

In this chapter, we continue investigating structure-charge transport relations, but shift focus from small molecule DBP to the P3HT structures presented in Chapter 4. P3HT is the benchmark donor material in organic electronics, largely due to reasonable hole mobilities, ideal energy level placement and bandgap, excellent optical absorption, and good solution processability[20]. It is also widely studied experimentally–especially in the context of P3HT:fullerene OPVs where device power conversion efficiencies as high as 6.5% have been obtained[7]. In neat P3HT field effect transistors, high mobilities of 0.1-0.4 cm^2/Vs have been measured for devices which contained microcrystalline domains surrounded by an amorphous matrix[4,27]. Time-of-flight mobilities for P3HT tend to be lower, $\mu = 1 \times 10^{-3}$ to 1×10^{-5} cm^2/Vs, due to the lower charge density and absence of an electric field to drive the movement of charges[1,16,19,25]. These experiments have made excellent progress in linking the

---

[1]This chapter is published in the journal *Polymers* and is referenced as "Miller, E. D.; Jones, M.L.; Jankowski, E. Tying Together Multiscale Calculations for Charge Transport in P3HT: Structural Descriptors, Morphology, and Tie-Chains. *Polymers* **2018**, *10*, 1358."

nanoscale polymer structure to device performance. For instance, high regioregularity[16,19,25] (i.e. a large proportion of the monomers in each chain have a consistent placement of the aliphatic side chain attached to the thiophene ring (see Figure 6.1a)) and shorter polymer chains are both expected to result in advantageous molecular packing resulting in a high degree of crystalline order[1]. However, comprehensive experimental investigation of the relationship between morphology and charge motion is prohibited by time, expense, and difficulty.
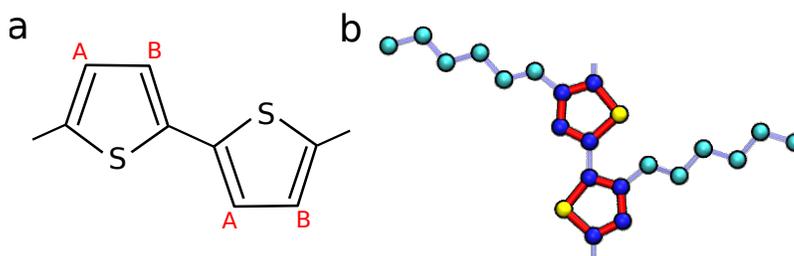
Figure 6.1: (a) The chemical structure of a poly-alkylthiophene chain. If the aliphatic side chains are consistently located only at the A sites or the B sites along the backbone, then the polymer is regioregular. In regiorandom chains, the placement of the aliphatic side chain is randomly placed at an A or B site on each thiophene. (b) P3HT united-atom model used in previous work. Sulfur, aromatic carbons, and aliphatic carbons are represented by yellow, dark blue, and cyan beads respectively. The rigid bonds are shown with thick red lines and flexible bonds are shown in light blue. (Figure replicated from [22] with authors' permission.)

Computational modeling provides insight into morphology and mobility that is inaccessible in experiments, and can more efficiently be used to evaluate how changes to processing parameters (temperature, solvent quality) tune performance. Techniques such as drift-diffusion[2,17], master equation[8,26], and kinetic Monte Carlo (KMC)[5,10,13,15] have all been successfully employed to investigate charge transport of model OPV morphologies. Drift-diffusion and lattice-based "mesoscale" KMC[14] can investigate device performance properties, but lose important details at molec-

ular length-scales. Conversely, master equation and "molecular" KMC maintain the molecular resolution but require approximations such as periodic boundary conditions to investigate charge motion over distances relevant for devices[13,15,26]. Such methods have been used to investigate time-of-flight mobilities, some reporting values a few orders of magnitude higher than expected (1×10^0 to 1×10^3 cm^2/Vs)[13,28,29], and others focusing on transfer integrals and inferred mobility without predicting mobility values[5,18]. For this investigation, we implement molecular KMC simulations, which are more computationally expensive than master equation techniques, but offer explicit spatial resolution of charges within the morphology[13,15]. Molecular KMC uses the positions, orientations, and energetics of electronically active portions of the molecules (chromophores), to determine the rates at which carriers can perform quantized "hops" between chromophore pairs. The hopping rates between every pair of chromophores in the system can be calculated in order to predict the expected motion of carriers through the system and the overall μ_0 (subscript "0" signifies that there is no electric field, similar to time-of-flight experiments).

In this work we utilize morphologies generated in Section 4 using a model that provides state-of-the-art prediction accuracy validated by experiments, while still providing sufficient computational efficiency to facilitate the investigation of a large number of processing parameter combinations[22]. With these morphologies, we perform semi-empirical QCCs to obtain the chromophore energy levels and molecular KMC simulations to obtain hole mobilities for pristine P3HT thin films at ~ 100 different state-points. We find that the structural order parameter developed previously, ψ, does not satisfactorily predict the observed charge carrier mobility in the thin films. Modifying the descriptor by including the variation of aliphatic bond lengths as a proxy for system-wide disorder, ψ', provides much better quantitative agreement be-

tween order parameter and charge mobility for the small "parameter sweep" systems. That said, we show that ψ' is less predictive of mobility for larger monodisperse systems ($10\times$ the number of chains) at experimentally interesting state-points. We propose that this is due to ψ' not taking into account the difficulty in charges hopping between crystallite grains of different orientations, effectively trapping carriers in the ordered crystallites. We therefore investigate the effect of a polydisperse distribution of chain lengths on mobility. We show that the longest chains in these distributions can span multiple clusters as "tie-chains" and mitigate the carrier trapping within crystals, bringing mobilities back in line with the predictions offered by ψ'. This allows us to predict the processing conditions that result in the highest performing devices. Our finding that highly ordered structures may have low mobilities if connecting paths between ordered domains are absent paves the way for new analytical techniques to help link structure to device performance[28].

This article is structured as follows: in Section 6.2 we cover the important details of previous molecular dynamics (MD) work and discuss the methodology in using the MorphCT simulation package[12] to conduct KMC simulations and obtain zero-field mobilities. In Section 6.3.1 we then explore charge transport properties for a large set of morphologies generated using an optimized molecular model, and in Section 6.3.2 identify structural features important for charge transport in larger systems.

6.2   Methods

6.2.1   Molecular Dynamics Simulations

The P3HT morphologies studied here were previously predicted using MD simulations, and so only salient information will be covered here[22]. We investigate ~ 100

morphologies generated from simulations using an adapted Optimized Performance for Liquid Simulations - United Atom forcefield to govern the non-bonded pair interactions (see Figure 6.1b).

Here, all systems are at experimentally measured thin-film density, $\rho = 1.11$ g/cm^3. However, we still consider a wide range of temperatures ($80 \leq T \leq 1300$ K in steps of 80 K) and solvents ($0.2 \leq \varepsilon_s \leq 1.2$ in steps of 0.2). We find that the highest degrees of order are observed in a band located from low temperatures and "good" solvents ($T \sim 250$ K, $\varepsilon_s = 0.2$) to high temperatures and "poor" solvents ($T \sim 750$ K, $\varepsilon_s = 1.2$). We note that the T defined here corresponds only to the temperatures of the MD simulation, which affects the energetic disorder of the polymer chains but not subsequent charge transport calculations. We also curate larger systems (cubic simulation volumes of side 15 nm containing 1,000 chains with degree of polymerization 15 - 15mers) with differing degrees of order (as quantified by the order parameter, ψ): "amorphous", "semi-crystalline", and "crystalline" by terminating the evolution of an experimentally relevant state-point after different degrees of equilibration. This ensures that different degrees of structural order can be obtained without changing the energetic disorder arising from thermal vibrations, allowing us to divorce the effects of structure and energetic disorder on the charge transport properties.

To quantify the structural order of these systems, we develop an order parameter, ψ, which is defined as the fraction of thiophene rings comprising "large" clusters containing more than six thiophene rings out of the total number of thiophene rings in the system - a measure of the proportion of crystallinity. A key aim of this work is to ascertain the efficacy of using ψ - a purely structural property of the morphology - to predict the charge transport of an arbitrary morphology.

6.2.2   Kinetic Monte Carlo Simulations

The charge transport calculations are performed using the MorphCT software package, running on Intel Xeon central processing units[12,15]. Mapping the structure to the mobility requires several processes operating over multiple length-scales. These are combined into an automated simulation pipeline, permitting us to analyse the molecular structure on Ångström length-scales and femtosecond time-scales, to carrier motion over hundreds of nanometers and microseconds. Firstly, the morphology is returned to the atomistic from the united-atom representation. For more strongly coarse-grained systems than those explored here, the interested reader is referred to the fine-graining methodology described in Ref [15]. In the case of the united-atom morphologies, fine-graining is somewhat trivial, as the hydrogens can be placed around the appropriate beads based on sp hybridization rules and typical element-hydrogen bond-lengths. The molecules are then split into electronically active chromophores, which are defined as individual monomers for simplicity. Although the carrier delocalization length for P3HT is around 7 monomers[21,23], we have found that using individual monomers broadly captures the delocalization behaviour by calculating fast electronic couplings between adjacent monomers compared to adjacent chains. Using single-monomer chromophores is advantageous as it removes the requirement of knowing the delocalization length of the simulated polymer beforehand, increasing the transferability and applicability of the model to other polymeric systems[15]. The neighbors of every chromophore in the system are calculated by performing a Voronoi analysis that treats adjacent Voronoi cells as direct neighbors. The molecular orbital energy levels of each pair of neighboring chromophores as a dimer, as well as each chromophore in isolation, are calculated using fast, semi-empirical QCCs. MorphCT

uses the Zerner's intermediate neglect of differential overlap method (ZINDO/S), which has been shown to provide good agreement of relative orbital energies when compared to more rigorous DFT techniques (See Appendix E.1 for details).

As was discussed in Chapter 5, to apply the Marcus expression and determine hopping rates, we must still the reorganization energy value from literature. The reorganization energy, λ, is the energy required to polarize and depolarize a chromophore, in response to a carrier hopping from one to another. This is material specific, and for P3HT we set $\lambda = 306$ meV based on electronic structure calculations for a single monomer[11]. Given the absence of other charges in the system restricting carrier movement, and the small, periodically repeating simulation volume enhancing the effect of order in the system, we expect our mobility values to be somewhat larger than those determined experimentally. However, we argue that the simulated mobilities calculated in this work are still an important proxy for semi-conducting electronic device performance, and can be compared to each other to make predictions about expected charge transport trends in physical devices.

6.3   Results and Discussion

Here we calculate the zero-field hole mobilities in P3HT morphologies predicted with molecular simulations in[22]. With both morphologies and predicted mobilities in-hand, we first evaluate structural correlations with mobility by comparing two order parameters. Second, we perform simulations of polydisperse P3HT chains to investigate a mobility anomaly observed for semi-ordered monodisperse chains.

6.3.1 Structure and mobility in "small" morphologies

In previous work, we predicted equilibrium morphologies of P3HT at ~ 100 combinations of temperature, T, and solvent strength, ε_s[22]. Each of these model systems is monodisperse, with 100 15-mers in a cubic periodic volume with 7 nm edges. At each (T, ε_s) state-point, we calculate the degree of order, ψ, in the system by identifying clusters of π-stacked thiophene rings with close positions and orientations. The resultant phase diagram is shown in Figure 6.2a. A band of highly ordered morphologies is visible spanning ε_s from low $T \sim 300$ K and good solvent $\varepsilon_s = 0.2$ to high $T \sim 700$ K and poor solvent $\varepsilon_s = 1.2$. This band is surrounded by state-points at $T < 300$ K and $T > 1000$ K that show a poor degree of order regardless of the solvent quality.

The zero-field hole mobilities, as calculated by MorphCT for each morphology state-point (~ 100 systems), are shown in Figure 6.2b. These mobilities span an order of magnitude from $0.01 \leq \mu_0 < 0.15$ cm^2/Vs as T and ε_s. We note these mobilities are roughly two orders of magnitude higher than observed in experiment ($\mu = 1 \times 10^{-5}$ to 1×10^{-3} cm^2/Vs for time-of-flight measurements[1,16,19,25]), but are still an improvement over similar P3HT calculations ($\mu = 1 \times 10^0$ to 1×10^3 cm^2/Vs[13,28,29]). First and foremost, we attribute the overprediction to the periodic volume with only 7 nm edges; there is little opportunity for boundaries between transport domains to emerge. That is, periodic volumes may overpredict mobility because grain boundaries (or their analogues) are rare. Second, contaminants such as residual solvent are not represented in our molecular model, and would otherwise lead to restricted mobility in experiments. The lowest μ_0 are seen at the highest processing temperatures (> 1100 K) in poor solvents ($\varepsilon_s = 1.2$). Conversely, the highest μ_0

values are seen in the morphologies prepared at low temperatures (< 200 K) in good solvents ($\varepsilon_s \simeq 0.2$). A band of high mobility is seen in a qualitatively similar region to Figure 6.2a, spanning from $T \sim 300$ K and $\varepsilon_s = 0.2$ to $T \sim 700$ K and $\varepsilon_s = 1.2$. This suggests that the strong ordering of thiophene backbones into large crystalline clusters is an important prerequisite for efficient carrier transport. However, it is clear that this is not the only factor affecting the transport, as state-points with low $T < 300$ K and $\varepsilon_s < 0.5$ are also shown to have high mobilities, despite a reduced ψ value. We therefore deduce that our simple definition for ψ, which only considers the conjugation and crystallization of the thiophene backbones, does not sufficiently encode all of the morphological features required to describe charge transport in the system.
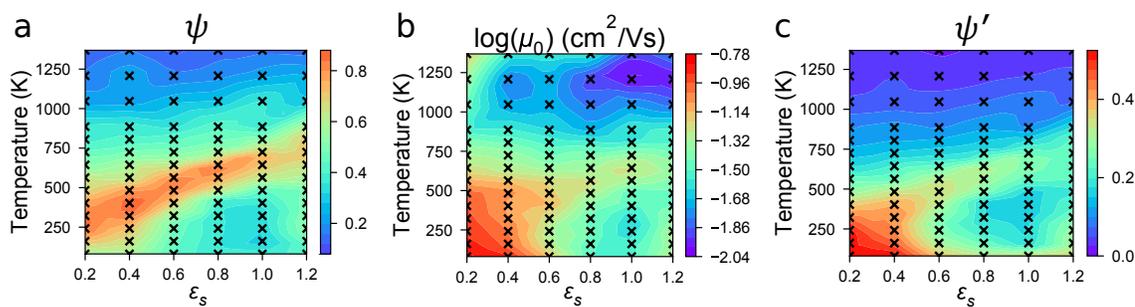
Figure 6.2: Heatmaps of the various properties explored for each simulation. In all cases, black "x"s correspond to state-points where simulations were conducted, with values in between determined by linear interpolation. Red regions correspond to a large value of the property, whereas blue regions correspond to a smaller value. Color bars are normalized to the maximum value of each parameter. (a) The structural order of each system given by the order parameter, ψ, as in the previous work (Figure replicated from [22] with authors' permission). (b) The hole mobility, μ_0, varying between red (~ 0.15 cm^2/Vs) and blue (~ 0.01 cm^2/Vs) There is not a perfect mapping between ψ and μ_0 - lower and higher temperature systems have higher and lower μ_0 respectively, which is not captured effectively by ψ. (c) The modified order parameter, ψ', created by normalizing ψ by the standard deviation of aliphatic bond lengths. ψ' is a significant improvement over ψ, as it captures the presence of additional disorder in the system.

In an attempt to better correlate nanostructure to mobility, we propose a new order parameter that supplements ψ with additional structural information not otherwise encoded by the calculation of structural order. Here we utilize disorder in aliphatic side-chains: it is straightforward to calculate bond-stretching statistics, these are structural metrics not included in the clustering criteria of ψ, and it is plausible that high variance in bond-stretching is a proxy for disorder between the thiophene rings that leads to lower mobility. Note that only the bonds in the aliphatic sidechains are considered in this way, as the backbone thiophene rings are rigid and so have fixed bond, angle, and dihedral constraints. We define $\tilde{\sigma}_i$ for the morphology at each unique state-point (subscript "i"), which is the standard deviation of the bond length distribution for the state-point, σ_i, normalized by the minimum value of σ_i across all state-points:

$$\tilde{\sigma}_i = \frac{\sigma_i}{\min\{\sigma_i\}}. \tag{6.1}$$

Our new order parameter ψ' is defined by:

$$\psi'_i = \frac{\psi_i}{\tilde{\sigma}_i}. \tag{6.2}$$

The aliphatic side-chain bond-length distribution is therefore being used as a proxy for disorder within each cluster, with ψ' weighting clusters with a narrow bond-length distribution as more highly ordered than those with broader distributions. This normalization of ψ provides a new lens for structure, as shown in Figure 6.2c. Qualitatively, ψ' better matches the mobility trends in Figure 6.2b than Figure 6.2a. This agreement suggests that ψ' alleviates shortcomings in ψ (a binary classifier that considers two molecules as only clustered or not), by instead allowing us to

quantify *how ordered* a cluster is. We note that ψ' still tends to underpredict μ_0 at high temperatures > 750 K, especially for systems dissolved in poor solvents with $\varepsilon_s \leq 0.6$. However, for experimentally relevant temperatures and good solvents, there is broad agreement between ψ' and μ_0, which is encouraging for the development of a structural metric that can predict charge transport properties of a morphology without performing KMC. The correlation between ψ' and μ_0 is quantified in figure Figure 6.3. The two properties are shown to track better than random, with a correlation coefficient $R^2 = 0.62$ across all state-points, although we note increased variability in μ_0 for systems with low ψ'. Many of the low ψ' values correspond to systems at high temperatures (> 750 K), suggesting that ψ' tends to overpredict the backbone disorder arising from thermal contributions. Generally however, correlation between ψ' and μ_0 suggests that ψ' can be used to quickly identify processing protocols that are expected to have good charge transport properties to submit for further investigation.

6.3.2 Structure and mobility in "large" and polydisperse cases

To investigate our hypothesis that the high charge mobilities predicted above in "small" volumes is a consequence of these volume we perform charge mobility calculations on systems with 10 times as many simulation elements. Although they contain 1,000 oligomers, each edge of these volumes is just over twice as long (15 nm) as the 100 oligomer simulations (7 nm) because the cubic box length scales as the cube root of the number of elements. These volumes are expected to be larger than the average size of the P3HT crystallites, allowing for multiple crystalline and amorphous domains in the same sample in accordance with experiment, but are still smaller than the thickness of the thin-films developed experimentally. While these
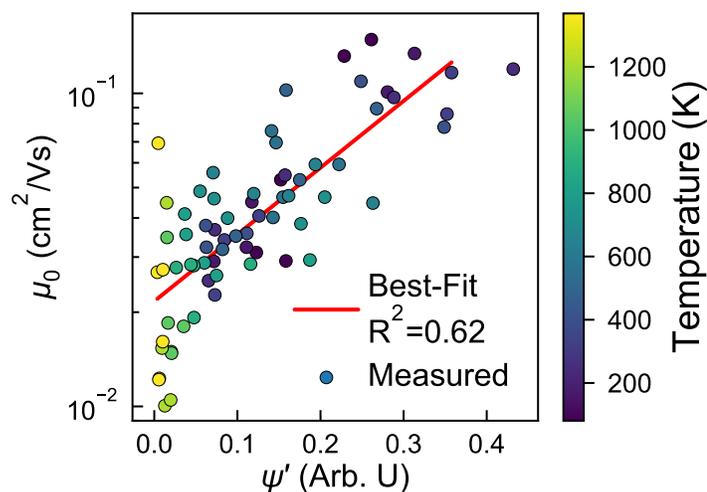
Figure 6.3: The zero-field mobility, μ_0 shows reasonable correlation to ψ' (R^2=0.62), indicating it can be useful as a purely structural metric to broadly predict interesting processing state-points to investigate further.

larger systems hold promise for giving better insight into charge transfer pathways, we recognize the periodic volumes could still lead to inflated absolute mobilities compared to time-of-flight measurements.

Mobility and Carrier Behavior

The mobilities μ_0 for the three classes of monodisperse 1,000-molecule P3HT morphologies: amorphous ($\psi' \sim 0.17$), semi-crystalline ($\psi' \sim 0.25$), and crystalline ($\psi' \sim 0.33$) are shown in Figure 6.4a, along with error bars representing the standard deviations from 10 independent microstates. Additionally, all calculated charge transport parameters, along with their associated standard errors, are listed in Table 6.1. Unlike in the "small" systems, we observe no correlation between mobility and ψ'. The amorphous and crystalline cases have mobilities commensurate with the "small" systems (Figure 6.2), whereas the semi-crystalline system exhibits a
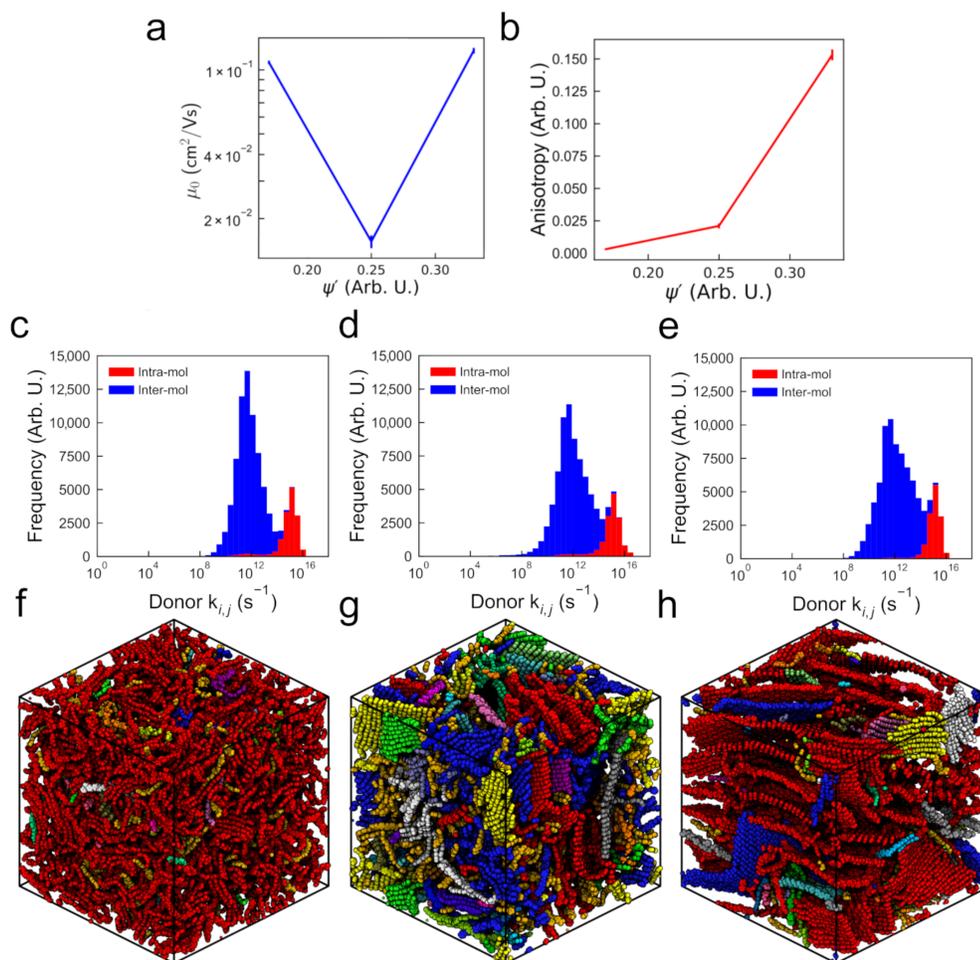
Figure 6.4: (a) Zero-field hole mobility (inset: representative morphology visualization with sidechains omitted for clarity) and (b) carrier trajectory anisotropy for the 1,000 oligomer simulations. In both cases, error bars are calculated based on the standard error arising from 10 repeat simulations of structurally decorrelated morphologies simulated under the same conditions. The hopping-rate distributions for the (c) amorphous, (d) semi-crystalline, and (e) crystalline systems have stacked bars (no obfuscation) where red describes hops along the chain and blue describes hops between chains. Regions of high connectivity in the (f) amorphous, (g) semi-crystalline, and (h) crystalline morphologies are denoted by colored clusters. The clusters are defined based on the frequency of hops performed between chromophore pairs in the simulations (further details in Appendix E.3.

significantly lower mobility. We can interpret this observation in two opposing ways: On one hand, zero-field charge mobility of $\mu_0 = 1.56 \times 10^{-2}$ cm^2/Vs is nearing

the 1×10^{-3} cm^2/Vs observed in experiments, and seems to support our working hypothesis that boundaries between crystallites should inhibit charge transport. On the other hand, this observation is surprising because P3HT is widely regarded to form a semi-crystalline structure in experimental devices, which we expect to have higher charge mobility than the amorphous case[24]. We note that further modifications to ψ' that explicitly accounts for the variation in transfer integrals across chromophores does not address the fact that the lowest mobility comes from medium order (See Appendix E.2 for details). Throughout this work (e.g., Figure 6.4) we color backbones of P3HT based upon the cluster to which they belong, which depends on charge hops, and is discussed in detail in Appendix E.3.

Table 6.1: Charge transport metrics calculated for the three degrees of ordering in systems of 1,000 monodisperse P3HT chains. Average values for 10 statistically independent samples are listed, along with the corresponding standard error over the 10 measurements. Clusters are defined based on a hopping frequency cut-off as described in the text.

Property	Amorphous	Semi-Crystalline	Crystalline
Mobility (cm^2/Vs$\times 10^1$)	1.085 ± 0.006	0.156 ± 0.003	1.23 ± 0.01
Anisotropy (Arb. U.)	0.0031 ± 0.0001	0.0210 ± 0.0006	0.153 ± 0.001
Intra-molecular rate (s$^{-1} \times 10^{-15}$)	1.813 ± 0.001	2.493 ± 0.001	1.8703 ± 0.0006
Inter-molecular rate (s$^{-1} \times 10^{-13}$)	0.834 ± 0.001	2.208 ± 0.005	2.642 ± 0.005
ΔE_{ij} std (eV)	0.06252 ± 0.00006	0.1114 ± 0.0001	0.0571 ± 0.0001
Total clusters (Arb. U.)	500 ± 10	1540 ± 10	467 ± 6
Large (> 6) clusters (Arb. U.)	134 ± 1	209 ± 1	73 ± 1
Largest cluster size (Arb. U.)	9600 ± 200	2100 ± 100	8300 ± 300

To further investigate the anomalous semi-crystalline case, we consider the directions charges move during the KMC simulations. It might be expected that the carrier trajectory anisotropy controls the overall mobility–a high anisotropy suggests that carriers are restricted to a particular direction, making it more likely to increase its mean squared displacement over the same amount of time than in a system where transport in three dimensions is equally likely. In Figure 6.4b, carrier transport is

shown to be anisotropic in the crystalline morphology, indicating a consistent grain orientation between the crystalline regions. The anisotropy is significantly lower in the semi-crystalline case, where a variety of grain orientations are present. Perhaps unsurprisingly, the amorphous systems exhibit near-spherical carrier transport, which is consistent with the lack of ordered crystallines in the morphologies. The anisotropy is somewhat higher in the semi-crystalline case, where a variety of grain orientations are present. In Figure 6.4b, carrier transport is shown to be anisotropic in the crystalline morphology, indicating a consistent grain orientation between the crystalline regions. Given that the anisotropy of the semi-crystalline morphology is intermediate between the more and less ordered systems (unlike the calculated mobilities), we deduce that anisotropy is not the sole factor governing carrier mobility.

Our calculated hopping rate distributions presented in Figure 6.4c-e in isolation would also suggest intermediate mobilities for the semi-crystalline system, as the availability of fast (high $k_{i,j} \sim 10^{14}$ s^{-1}) inter-molecular hops appears to decrease with decreasing ψ'. Therefore, the distribution of hopping rates alone is insufficient to predict performance - the rate, location, and neighborhood of those hops in the morphology are all required in order to make predictions.

The amorphous morphology (Figure 6.4f) explicitly shows no crystallization, instead forming a disordered matrix of entangled polymer chains. However, with the chains colored based on charge hopping, it is clear that there is one large (red) percolating cluster that connects most chains to most other. The structure of the crystalline morphology (Figure 6.4h) is expectedly lamellar, with one large, ordered percolating (red) cluster. The semi-crystalline system (Figure 6.4g) shows small regions of crystallized lamellae, interspersed within an amorphous matrix. The prevalence of multiple clusters indicates that charges have trouble hopping between crystallites. This is the

first evidence we observe of non-intermediate properties of the semi-crystalline system compared to the crystalline and amorphous morphologies. Further analysis of the hops occurring within the ordered crystallites of the semi-crystalline and crystalline morphologies reveals that charges can travel in fast "loops" within ordered regions wherein hops are fast, but total displacement is low (Appendix E.4). That is, carriers in ordered regions have a high probability of spending long periods of time hopping between the same set of chromophores within the same plane, without increasing their displacement from their start position. Because these carriers are effectively "trapped", if the ordered regions are not connected, overall mobility suffers.

Considering these factors in aggregate, we conclude that the crystalline morphology mobility of 1.16×10^{-1} cm^2/Vs is due primarily to fast carrier transport along the ordered crystallites, and note that it would be even higher if carriers did not frequently "loop around" within the large ordered regions rather than travel ballistically. The mobility through the amorphous morphology is restricted by slower overall isotropic carrier motion, but the lack of traps explains higher mobility than the semi-crystalline case. The proximity of the amorphous case mobility (1.02×10^{-1} cm^2/Vs) to the crystalline case highlights the importance of trapping to overall carrier mobility. This also bears out some recent investigations that have shown beneficial carrier behavior in less conventionally-ordered systems[9]. The semi-crystalline morphology ranks highly in isotropic transport, *and* low in ordered domain alignment, resulting in an order-of-magnitude lower mobility of 1.64×10^{-2} cm^2/Vs.

In summary, we find that a convolution of different structural and transport metrics is required in order to correctly predict carrier mobilities - no one factor is sufficient to explain the observed trend. Carrier transport is strongly dependent on the local neighborhood around each chromophore - if a carrier has easy access to

the surrounding chromophores but not beyond, then it will become trapped, even if the average cluster and chromophore characteristics of the whole morphology are favourable. KMC simulations are the current best way to convolve the structural metrics and obtain the device performance behavior - it is presently not possible to map directly between structure and performance otherwise. For the three degrees of order considered here, we have shown that the amorphous morphology has stronger connections (characterized by a smaller number of larger clusters) than the more-ordered semi-crystalline one, leading to a higher carrier mobility and improved charge transport. That this disordered charge mobility is higher than expected in experiments suggests there are improvements to the absolute value charge hopping rates, or assumptions about chromophore size and electron delocalization that could improve mobilities calculated with KMC.

Tie chains in polydisperse systems

We hypothesize that the difference in mobility between the semi-crystalline system and the crystalline and amorphous morphologies is due to the monodispersity and short length of the chains studied here. Generally in experimental devices, P3HT is obtained with a molecular weight in excess of 50 kDa, corresponding to chain polymerisation lengths of many hundreds or thousands of monomers[13]. In such systems, the chains are long enough to fold back on themselves several times, forming sheet-style crystallites in the system, where a single chain can form multiple layers of the same crystallite[3]. Previous work has shown that, while 15mer chains were able to reproduce experimental scattering patterns, they were too short to undergo self-folding. Conversely, chains with 50 repeat units were able to undergo self-folding, but required untenably long simulation times to order into the experimentally ob-

served structures[22]. In the case of longer chains, multiple folds and multiple chains can stack together to increase the size of the crystallite regions, with portions of the outermost chains remaining outside of the crystallite, forming an amorphous matrix between the grains. An example of this is seen in Figure 6.5a and b in which a series of longer chains or a single long chains spans multiple clusters. In some cases, these "tails" may connect to a different crystallite, effectively forming a "tie-chain" between two crystallites[24]. Since carrier motion is faster along a chain than between neighboring chains, this provides a fast and efficient route for carriers to transport between crystallites, which is sometimes known as a "carrier highway". Such routes are not unique to P3HT; tie-chains are found to be critical in other polymeric systems, for instance in complimentary semiconducting polymer blends[30]. In our work, chains with 15 repeat units do not self-assemble tie-chains as evidenced by the semi-crystalline system explored here.
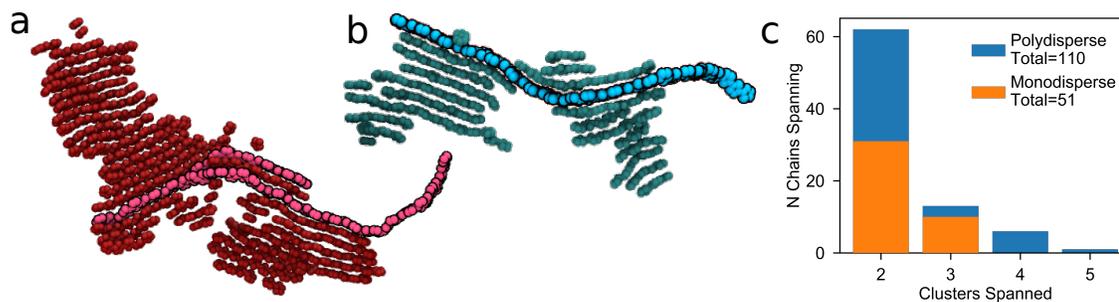
Figure 6.5: Long polymer chains are able to link clusters together to enhance charge transport between them. The links can either consist of (a) multiple chains or (b) a single chain extending through the surrounding amorphous matrix. (c) The semi-crystalline polydisperse systems, with chains up to 50 monomers and polydispersities of 1.8, have double the amount of tie-chains spanning two clusters as the semi-crystalline monodisperse system consisting only 15mers. Additionally, some chains in the polydisperse system span four or five clusters. The bars in the histogram overlap so that the frequency of chains spanning multiple clusters is given by the top of the orange and blue bars for the mono- and polydisperse systems respectively.

To test whether longer chains will serve as tie-chains between otherwise charge-trapping crystallites, we investigate morphologies made with polydisperse chains up to 50 repeat units (~ 8 kDa). The maximum of 50 repeat units is chosen to prevent any individual chain from interacting with itself in more than one image of the periodic volume. Due to these size constraints, we are unable to achieve experimentally relevant chain lengths (for example, 20-100 kDa from Sigma-Aldrich or Reike Metals). We can, however, achieve experimental polydispersities of 1.8 (See Appendix E.5 for details). Furthermore, simulating a polydisperse distribution of chain lengths allows us to introduce chains that may be long enough to span several crystallites, while still maintaining appropriate length-scales to obtain good agreement with experimental scattering patterns (also demonstrated in Appendix E.5).

Polydisperse morphologies are generated using the same process as the monodisperse cases as explained previously, and result in three similar degrees of ordering: amorphous ($\psi' \sim 0.18$), semi-crystalline ($\psi' \sim 0.27$), and crystalline ($\psi' \sim 0.31$). We calculate mobilities of these polydisperse morphologies with KMC and present them in Figure 6.6a. By including a distribution of chain lengths, the expected order-mobility trend has been reclaimed - mobility increases with additional order. Generally, μ_0 is slightly higher in the polydisperse systems than in the monodisperse 15mer systems, as the increased average molecular weight (2.9 ± 0.1 kDa for the polydisperse and 2.5 kDa for the monodisperse systems) leads to a higher proportion of fast intra-chain hops. Figure 6.6b-d show that, unlike the monodisperse systems in Figure 6.4f-h, all three of the systems are highly connected and form a single, large cluster spanning the entire system (colored red). This higher connectivity is due to the presence of more chains spanning between crystallites in the polydisperse case than the short monodisperse case (Figure 6.5c). The improved connectivity

is quantified in Table 6.2, where the number of large clusters and the size of the largest cluster are both intermediate between the amorphous and crystalline systems. Additionally, Table 6.2 shows a significantly lower carrier trajectory anisotropy in the case of the semi-crystalline and crystalline polydisperse systems than in the monodisperse case (Table 6.1). This suggests that charges are no longer restricted by grain boundaries and are able to change direction more easily - a process that was prohibitively slow in the monodisperse case. These results are in good agreement with previous investigations that show tie-chains are a dominating factor in carrier transport through polymer devices[6,13].
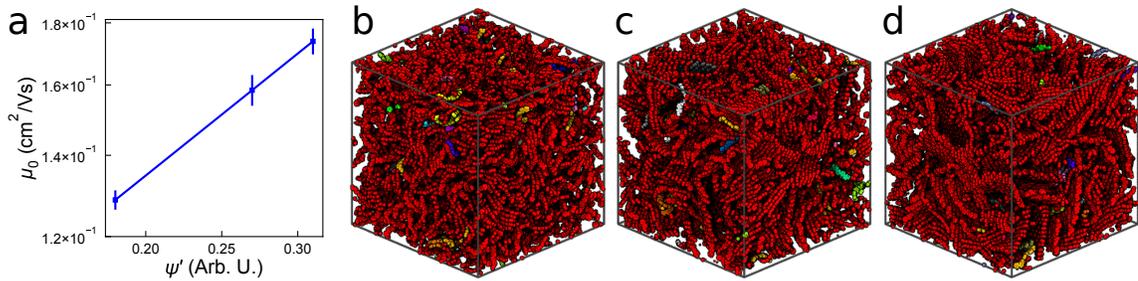
Figure 6.6: (a) Zero-field mobilities for the polydisperse P3HT simulations based on the modified order parameter ψ'. Morphologies showing regions of high connectivity for the (b) amorphous, (c) semi-crystalline, and (d) crystalline systems.

Table 6.2: Charge transport metrics calculated for three degrees of order in polydisperse P3HT systems. Average values over 10 statistically independent samples are listed, along with the corresponding standard error over the 10 measurements. Clusters are defined based on a hopping frequency cut-off as described in the text.

Property	Amorphous	Semi-Crystalline	Crystalline
Mobility ($cm^2/Vs \times 10^1$)	1.29 ± 0.02	1.58 ± 0.04	1.74 ± 0.04
Anisotropy (Arb. U.)	0.0040 ± 0.0005	0.016 ± 0.001	0.020 ± 0.004
Intra-molecular rate ($s^{-1} \times 10^{-15}$)	1.3413 ± 0.0003	1.4670 ± 0.0007	1.5137 ± 0.0002
Inter-molecular rate ($s^{-1} \times 10^{-13}$)	0.700 ± 0.004	1.231 ± 0.0007	1.590 ± 0.007
ΔE_{ij} std (eV)	0.0554 ± 0.0002	0.0549 ± 0.0001	0.0538 ± 0.0001
Total clusters (Arb. U.)	400 ± 23	350 ± 17	380 ± 21
Large (> 6) clusters (Arb. U.)	130 ± 10	70 ± 6	60 ± 3
Largest cluster size (Arb. U.)	11200 ± 260	13200 ± 200	13500 ± 100

The observation that ψ' and μ_0 are strongly correlated in large, polydisperse systems (Figure 6.6), somewhat correlated in small, monodisperse systems (Figure 6.3), and poorly correlated in large, monodisperse systems (Figure 6.4a), highlights a shortcoming in using purely structural metrics to predict charge transport. In isolation, structure can provide some insight into the average rate at which hops can occur in the morphology - of the hopping criteria studied in this investigation, only the hopping rate is described by ψ'. This relationship is quantified by the increase of average inter-molecular hopping rates for both the monodisperse (Table 6.1 shows $0.834 \rightarrow 2.208 \rightarrow 2.642 \times 10^{13}$ for the amorphous, semi-crystalline, and crystalline structures respectively) and polydisperse systems (Table 6.2: $0.700 \rightarrow 1.231 \rightarrow 1.590 \times 10^{13}$). Graphically this is also demonstrated by the shift of the inter-molecular hopping rate peak towards the intra-molecular peak in Figure 6.4c-e. However, considering only the hopping rate distributions fails to take into account the local neighborhood of hops available. Therefore, ψ' is unable to distinguish between regions where charges may be trapped within crystallites, or able to flow along a fast extended path. This is confirmed by our clustering analysis in Appendix E.2 - no combination of purely structural cluster criteria was able to produce the same cluster distributions observed in our simulations. We therefore conclude that knowledge of the carrier hopping rates in the chromophore network is insufficient - one must also know how these rates are distributed in order to identify regions of trapping that will reduce carrier mobilities. This is a key advantage of computational methods such as KMC–even though carriers have no knowledge of the surrounding hop neighborhood (all hops are performed on a chromophore-by-chromophore basis to first order), the extensive statistical averaging of the method allows us to probe the local hopping neighborhood and identify crystallites.

6.4   Conclusions

Using QCC to inform KMC simulations of charge transport in P3HT morphologies currently gives the best insight into how nanostructure influences charge mobility. These calculations confirm that charges move most quickly along P3HT backbones and second-most quickly between aligned backbones. However, because charges rarely hop between distinct crystallites, tie-chains connecting ordered crystallites are essential to mitigating the trapping of charges that would otherwise lower mobility. By combining the large volumes from optimized MD simulations of P3HT with QCC-informed charge transport, this is the first work to definitively show the impact tie-chains have on charge mobility. The computational techniques demonstrated in this manuscript are applicable to other organic semiconducting materials (including non-polymeric small molecules) and we expect to detect a similar relationship between charge transport and the presence of tie-chains for other conjugated polymer systems.

Looking to the future, this work highlights two areas for improving mobility predictions. Firstly, the present work shows that purely structural metrics miss important factors for charge transport, but this does not preclude the existence of better metrics that are more predictive than those studied here. That is, discovery of structural metrics that are good enough to predict mobility without having to perform KMC simulations would save a lot of time. Secondly, while the mobilities predicted with KMC are the current state-of-the-art, they are systematically about two orders of magnitude higher than in experiments. Whether this is due to inaccurate assumptions about what comprises a chromophore, or whether improvements to calculating charge hopping rates are needed, or something else, it seems like quantitative predictions of mobility are on the horizon. Exploring these improvements to the KMC calculations

presented here and investigating a broader range of chemistries to further validate these techniques is the subject of future work.

BIBLIOGRAPHY

[1] Amy M. Ballantyne, Lichun Chen, Justin Dane, Thomas Hammant, Felix M. Braun, Martin Heeney, Warren Duffy, Iain McCulloch, Donal D. C. Bradley, and Jenny Nelson. The Effect of Poly(3-hexylthiophene) Molecular Weight on Charge Transport and the Performance of Polymer:Fullerene Solar Cells. *Advanced Functional Materials*, 18(16):2373–2380, 2008.

[2] J. A. Barker, C. M. Ramsdale, and N. C. Greenham. Modeling the Current-Voltage Characteristics of Bilayer Polymer Photovoltaic Devices. *Physical Review B*, 67(7):075205, 2003.

[3] Martin Brinkmann and Jean Claude Wittmann. Orientation of Regioregular Poly(3-hexylthiophene) by Directional Solidification: A Simple Method to Reveal the Semicrystalline Structure of a Conjugated Polymer. *Advanced Materials*, 18(7):860–863, 2006.

[4] Jui Fen Chang, Baoquan Sun, Dag W. Breiby, Martin M. Nielsen, Theis I. Sölling, Mark Giles, Iain McCulloch, and Henning Sirringhaus. Enhanced Mobility of Poly(3-hexylthiophene) Transistors by Spin-Coating from High-Boiling-Point Solvents. *Chemistry of Materials*, 16(23):4772–4776, 2004.

[5] David L. Cheung, David P. McMahon, and Alessandro Troisi. Computational Study of the Structure and Charge-Transfer Parameters in Low-Molecular-Mass P3HT. *The Journal of Physical Chemistry B*, 113(28):9393–9401, 2009.

[6] Edward J. W. Crossland, Kim Tremel, Florian Fischer, Khosrow Rahimi, Günter Reiter, Ullrich Steiner, and Sabine Ludwigs. Anisotropic Charge Transport in Spherulitic Poly(3-hexylthiophene) Films. *Advanced Materials*, 24(6):839–844, 2012.

[7] Minh Trung Dang, Lionel Hirsch, and Guillaume Wantz. P3HT:PCBM, Best Seller in Polymer Photovoltaic Research. *Advanced Materials*, 23(31):3597–3602, 2011.

[8] Mario Einax, Marcel Dierl, and Abraham Nitzan. Heterojunction Organic Photovoltaic Cells as Molecular Heat Engines: A Simple Model for the Performance Analysis. *The Journal of Physical Chemistry C*, 115(43):21396–21401, 2011.

[9] Juan Felipe Franco-Gonzalez, Nicolas Rolland, and Igor V. Zozoulenko. Substrate-Dependent Morphology and Its Effect on Electrical Mobility of Doped Poly(3,4-ethylenedioxythiophene) (PEDOT) Thin Films. *ACS Applied Materials & Interfaces*, 10(34):29115–29126, 2018.

[10] Chris Groves, Robin G. E. Kimber, and Alison B. Walker. Simulation of Loss Mechanisms in Organic Solar Cells: A Description of the Mesoscopic Monte Carlo Technique and an Evaluation of the First Reaction Method. *The Journal of Chemical Physics*, 133(14):144110, 2010.

[11] Erik Johansson and Sven Larsson. Electronic Structure and Mechanism for Conductivity in Thiophene Oligomers and Regioregular Polymer. *Synthetic Metals*, 144(2):183–191, 2004.

[12] M. L. Jones. MorphCT, 2018.

[13] M. L. Jones, D. M. Huang, B. Chakrabarti, and Chris Groves. Relating Molecular Morphology to Charge Mobility in Semicrystalline Conjugated Polymers. *Journal of Physical Chemistry C*, 120(8):4240–4250, 2016.

[14] Matthew L. Jones, Reesha Dyer, Nigel Clarke, and Chris Groves. Are Hot Charge Transfer States the Primary Cause of Efficient Free-Charge Generation in Polymer:Fullerene Organic Photovoltaic Devices? A Kinetic Monte Carlo Study. *Physical Chemistry Chemical Physics*, 16(38):20310–20320, 2014.

[15] Matthew Lewis Jones and Eric Jankowski. Computationally Connecting Organic Photovoltaic Performance to Atomistic Arrangements and Bulk Morphology. *Molecular Simulation*, 43(10-11):1–18, 2017.

[16] Youngkyoo Kim, Steffan Cook, Sachetan M. Tuladhar, Stelios A. Choulis, Jenny Nelson, James R. Durrant, Donal D. C. Bradley, Mark Giles, Iain McCulloch, Chang-Sik Sik Ha, and Moonhor Ree. A Strong Regioregularity Effect in Self-Organizing Conjugated Polymer Films and High-Efficiency Polythiophene:Fullerene Solar Cells. *Nature Materials*, 5(3):197–203, 2006.

[17] L. J. A. Koster, E. C. P. Smits, V. D. Mihailetchi, and P. W. M. Blom. Device Model for the Operation of Polymer/Fullerene Bulk Heterojunction Solar Cells. *Physical Review B*, 72(8):085205, 2005.

[18] Yi-kang Lan and Ching-i Huang. A Theoretical Study of the Charge Transfer Behavior of the Highly Regioregular Poly-3-hexylthiophene in the Ordered State. *The Journal of Physical Chemistry B*, 112(47):14857–14862, 2008.

[19] Ralf Mauer, Marcel Kastler, and Frédéric Laquai. The Impact of Polymer Regioregularity on Charge Transport and Efficiency of P3HT:PCBM Photovoltaic Devices. *Advanced Functional Materials*, 20(13):2085–2092, 2010.

[20] Katherine A. Mazzio and Christine K Luscombe. The Future of Organic Photovoltaics. *Chemical Society Reviews*, 44(1):78–90, 2015.

[21] David P. McMahon, David L. Cheung, Ludwig Goris, Javier Dacuña, Alberto Salleo, and Alessandro Troisi. Relation between Microstructure and Charge Transport in Polymers of Different Regioregularity. *The Journal of Physical Chemistry C*, 115(39):19386–19393, 2011.

[22] Evan Miller, Matthew Jones, Michael Henry, Paul Chery, Kyle Miller, and Eric Jankowski. Optimization and Validation of Efficient Models for Predicting Polythiophene Self-Assembly. *Polymers*, 10(12):1305, 2018.

[23] D. H. K. Murthy, M. Gao, M. J. W. Vermeulen, L. D. A. Siebbeles, and T. J. Savenije. Mechanism of Mobile Charge Carrier Generation in Blends of Conjugated Polymers and Fullerenes: Significance of Charge Delocalization and Excess Free Energy. *The Journal of Physical Chemistry C*, 116(16):9214–9220, 2012.

[24] Rodrigo Noriega, Jonathan Rivnay, Koen Vandewal, Felix P. V. Koch, Natalie Stingelin, Paul Smith, Michael F. Toney, and Alberto Salleo. A General Relationship Between Disorder, Aggregation and Charge Transport in Conjugated Polymers. *Nature Materials*, 12(11):1038–1044, 2013.

[25] Shyam S. Pandey, Wataru Takashima, Shuichi Nagamatsu, Takeshi Endo, Masahiro Rikukawa, and Keiichi Kaneto. Regioregularity vs Regiorandomness: Effect on Photocarrier Transport in Poly(3-hexylthiophene). *Japanese Journal of Applied Physics*, 39(Part 2, No. 2A):L94–L97, 2000.

[26] Nicolas Rolland, Juan Felipe Franco-Gonzalez, Riccardo Volpi, Mathieu Linares, and Igor V. Zozoulenko. Understanding Morphology-Mobility Dependence in PEDOT:Tos. *Physical Review Materials*, 2(4):045605, 2018.

[27] H. Sirringhaus, P. J. Brown, R. H. Friend, M. M. Nielsen, K. Bechgaard, B. M. W. Langeveld-Voss, a. J. H. Spiering, R. A. J. Janssen, E. W. Meijer, P. Herwig, and D. M. de Leeuw. Two-Dimensional Charge Transport in Self-Organized, High-Mobility Conjugated Polymers. *Nature*, 401(6754):685–688, 1999.

[28] Ellen Van, Matthew Jones, Eric Jankowski, and Olga Wodo. Using Graphs to Quantify Energetic and Structural Order in Semicrystalline Oligothiophene Thin Films. *Molecular Systems Design & Engineering*, 1:273–277, 2018.

[29] Linjun Wang, Qikai Li, Zhigang Shuai, Liping Chen, and Qiang Shi. Multiscale Study of Charge Mobility of Organic Semiconductor with Dynamic Disorders. *Physical Chemistry Chemical Physics*, 12(13):3309, 2010.

[30] Yan Zhao, Xikang Zhao, Michael Roders, Ge Qu, Ying Diao, Alexander L. Ayzner, and Jianguo Mei. Complementary Semiconducting Polymer Blends for Efficient Charge Transport. *Chemistry of Materials*, 27(20):7164–7170, 2015.

CHAPTER 7

MACHINE LEARNING PREDICTIONS OF ELECTRONIC COUPLINGS FOR CHARGE TRANSPORT CALCULATIONS OF P3HT[1]

7.1 Introduction

Finding a needle in a haystack is hard because of all the hay: Inspecting each straight, pointy object drawn from a large haystack rarely reveals needles and it is impractical to inspect all the pointy objects. Searching haystacks is analogous to finding optima in large problem spaces—such as the identification of the best ingredients for high-efficiency, low-cost organic photovoltaics (OPVs) for sustainable power generation, in which, progress is hindered by the experimental and computational expense of enumerating the combination of factors that govern a candidate's viability. Replacing experiments with computer simulations increases the rate of candidate inspection, as computer simulations can be performed at a lower cost and in less time, but does not wholly alleviate the time burden. Here we focus on strategies for further increasing the rate at which candidates can be inspected by lowering the computational cost of connecting OPV structure to its performance.

---

[1]This chapter has been submitted the *AIChE* journal and is referenced as "Miller, E. D.; Jones, M.L.; Henry, M.H.; Stanfill, B.; Jankowski, E. Machine Learning Predictions of Electronic Couplings for Charge Transport Calculations of P3HT. *AIChE* **2019**–Submitted"

OPVs are a focus of sustainable energy development because devices with 15% η are theorized as sufficient for one-day energy-pay-back times[10], which would circumvent economic barriers to widespread deployment. A key difficulty in mass-producing high η devices is controlling the self-assembled active-layer morphology (the spontaneously forming microstructure within the electricity generating portion of the device). The majority of active layers are primarily composed of two components: an electron donating and an electron accepting species, and the microstructural order of these two components determines the device's overall efficiency[43]. Recent developments in new OPV ingredients have demonstrated power conversion efficiencies in excess of 15%[29,46], however mass-produced OPVs still fall below the efficiencies required for widespread commercial viability, and the precise origins of the higher efficiencies are not fully understood. To make OPVs with one-day payback times a reality, a fundamental understanding of how ingredient chemistry and processing determines the active layer morphology and how the resulting features influence η is needed.

Here we describe machine learning (ML) efforts towards speeding calculations linking OPV morphology to the mobility of charges through it, which in turn determines the fill factor and η[45] of OPV devices. To validate our approach, we focus on the benchmark donor polymer poly-(3-hexylthiophene) (P3HT), which is the archetype for linking the self-assembled morphology to efficiency[8,28] due to its solution processability and history in breakthrough (in 2006, 5% η) OPVs[25]. In P3HT devices, faster charge movement (which corresponds to better η) can be obtained by creating devices that maximize the degree of crystalline order,[6] which can be accomplished by using high regioregularity[41] and shorter polymer chains[5,21]. Time-of-flight measurements of hole mobility in P3HT experiments range from $\mu = 1 \times 10^{-5}$ to 1×10^{-3}

cm^2/Vs[2,19,27,35]. Computational work has helped to explain the role of thiophene ring orientation on charge transport[22], and kinetic Monte Carlo (KMC) simulations of charge transport have predicted mobilities ranging from $\mu = 1 \times 10^{-4}$ to 0.6 cm^2/Vs[13,16,18,31,42], depending on the degree of ordering in of the P3HT morphologies. These experimental and computational predictions of mobility provide references for validation: Calculated hole mobilities in P3HT should fall between $\mu = 1 \times 10^{-4}$ to 0.6 cm^2/Vs and increase with increasing P3HT crystallinity.

In our own prior work, we predict charge transport through P3HT by first predicting P3HT morphologies at ~ 350 processing state points[30], then calculating charge mobility through ~ 100 of these structures[31] using KMC simulations. Doing so requires hopping rates between P3HT chromophores, which we calculate with Marcus semi-classical hopping theory[26] using quantum chemical ZINDO/S[20,40] calculations to obtain the electronic transfer integrals between chromophores (couplings, $J_{i,j}$), which describe the amount of frontier molecular orbital overlap between pairs chromophores. Completely connecting all the neighbors in a representative system requires $\sim 2 \times 10^5$ ZINDO/S calculations per morphology, corresponding to about 26 CPU hours of computation time. We aim to determine the efficacy of using ML to predict $J_{i,j}$ and bypass the numerous, expensive ZINDO/S calculations required to characterise the charge transport properties of a morphology. We take inspiration from recent studies in which ML based on first-principle calculations has been used to accelerate development of organic light-emitting diodes[12], OPV candidate compounds[38], and electronic predictions based on coarse-grained sites[14]. The use of ML to accelerate materials discovery has grown recently due to advances in enabling hardware, algorithms, and open-source libraries[11,33,44]. The $J_{i,j}$ prediction problem approached here is well-suited to supervised learning algorithms where ample data

can inform classification or regression schemes relating inputs features to output properties, especially if discerning these relations would be difficult or tedious for a human[23,24,37,39].

7.2   Methods

We focus on two ways of generating electronic transfer integrals ($J_{i,j}$); the control case of quantum chemical ZINDO/S calculations using orca[34], and the test cases of machine learning methods trained to predict them. Either method of transfer integral generation is used in a sequence of computational methods:

1. Sample OPV morphologies using molecular simulations.

2. Generate transfer integrals between chromophores in each morphologies (this work)

3. Predict charge mobilities from transfer integrals using KMC simulations

In prior work we describe combining these steps into the MorphCT[15] software pipeline, the details of said implementation[18], and applications to P3HT (Chapter 6)[31].

To determine charge mobilities with kinetic Monte Carlo (KMC) simulations, morphologies are treated as a weighted network in which each P3HT monomer is considered an electronically active chromophore and charges may hop to neighboring chromophores as defined by neighboring cells from a Voronoi tessellation of thiophene

ring centers of mass. We calculate electronic transfer integrals between chromophores using the energy-splitting-in-dimer method (ESD)[4,9]:

$$J_{i,j} = \frac{1}{2}\sqrt{(E_{HOMO} - E_{HOMO-1})^2 - (\Delta E_{i,j})^2}, \tag{7.1}$$

where the magnitude of the splitting of the highest occupied molecular orbital to a new E_{HOMO} and E_{HOMO-1} in the dimer state is compared to the difference in $HOMO$ level of the isolated, individual chromophores:

$$\Delta E_{i,j} = E_{\text{HOMO},j} - E_{\text{HOMO},i}. \tag{7.2}$$

ZINDO/S requires atom positions and types of each chromophore to calculate $(E_{HOMO} - E_{HOMO-1})$ and $\Delta E_{i,j}$.

7.2.1 Machine learning

To predict $J_{i,j}$ using any machine learning approach we select input features that are then related to $J_{i,j}$ calculated by ZINDO/S. Because ZINDO/S requires only atom types and positions, we select nine spatial features that we expect to be predictors of $J_{i,J}$ between P3HT monomers:

1. Whether the monomers are chemically bonded to each other

2. The distance between their thiophene ring centers of mass

3. The relative "pitch" between thiophene rings (Figure 7.1)

4. The relative "roll" between thiophene rings

5. The distance between sulfur atoms on the thiophene rings

6. The x-component of the thiophene ring center separations

7. The y-component of the thiophene ring center separations

8. The z-component of the thiophene ring center separations

9. $\Delta E_{i,j}$
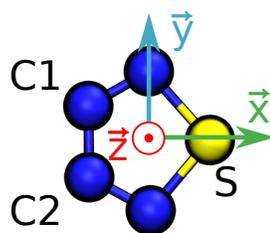
Figure 7.1: Reference thiophene ring and local coordinate axes used to determine relative spatial features between P3HT monomers. The thiophene ring center of mass is used as the origin of the local coordinates. A thiophene ring's rotation about its local y-axis relative to another thiophene ring in the reference frame is used to calculate "pitch". A thiophene's rotation about its local x-axis relative to the reference ring defines "roll".

Note that the "yaw" angle about the thiophene's local z-axis is missing from this list of features as preliminary work has shown that its effect on the transfer integral is negligible. This is expected as the electron density is delocalized above and below the plane of the thiophene ring, so rotations around the local z-axis do not affect the amount of molecular orbital overlap. We test ordinary least squares (OLS), support vector machines (SVM), K-nearest neighbors (KNN), artificial neural networks (ANN), and random forests (RF) as machine learning implementations for predicting $J_{i,j}$ from the above nine features. The review article of Ref [11] provides a comprehensive overview of ML techniques in soft matter, and is a recommended starting place for understanding the taxonomy of ML techniques. Briefly, OLS

determines coefficients for linear combinations of input features by minimizing error on a training data set; SVM classifies possible outcomes based on hyperplanes dividing the feature space of a training set; KNN uses determines "proximity" in feature-space between elements of a training set and predicts $J_{i,j}$ based on members of clusters that emerge from this grouping. ANN are composed of "layer" matrices that transform inputs into outputs through matrix multiplication, with iterative re-weighting matrix elements performed by gradient descent optimization using a training set of known features and $J_{i,j}$. The ANN is implemented in the Python package Tensorflow[1] (version 1.9.0, see SI-Section 1 for ANN details), and all other methods are conducted with the package Scikit-Learn (version 0.19.1) with the default argument values[36]. The code used in this study is available at Ref [32] and the data set at Ref [17].

We explain RFs in more detail, due to their focus in the discussion that follows. RFs are an ensemble technique in which the prediction from many decision trees are combined into an output. A decision tree operates by partitioning the data, based on the features and their values, into progressively smaller subgroups to determine an average outcome ($\bar{y}$) for the group. The decision tree implementation in Scikit-Learn[36] is based on the classification and regression tree (CART) algorithm, which creates a binary split based on a threshold (t_f) for a feature (f) at a "leaf", creating two "branches":

$$d_{f_x} = \begin{cases} \text{Left Branch if } f_x < t_f \\ \text{Right Branch if } f_x \geq t_f \end{cases}, \qquad (7.3)$$

in which d_{f_x} signifies the branch decision for sample x. The threshold t_f is determined by minimizing the cost function:

$$C(d_f) = \frac{n_{\text{left}}}{N} E_{\text{left}}(d_f) + \frac{n_{\text{right}}}{N} E_{\text{right}}(d_f), \qquad (7.4)$$

where n_{left} and n_{right} are the number of samples on each branch (based on the decision d_f), N is the total number samples on the leaf, and $E_{\text{left,right}}(d_f)$ is the error from assigning the samples to the left and right branches. This error is measured as the mean-squared error:

$$E(d_f) = \frac{1}{n_m} \sum_i^{n_m} (y_i - \bar{y})^2, \tag{7.5}$$

where y_i is the true output and n_m is the number of samples in the left or right branch. This processes is repeated further with additional cut-offs, thereby growing the tree and partitioning the data into smaller and smaller partitions, reducing the error on each leaf, until a stopping criteria (such as a maximum depth) is met. RFs avoid over-fitting by providing each tree with a different subset of the total training data, then taking the ensemble average over each tree "voting" on the outcome.

Here we draw training set chromophore pairs from one "disordered", one "semi-crystalline", and one "crystalline" morphology from prior work[31]. Each morphology is composed of 15,000 P3HT repeat units, giving about 230,000 chromophore pairs (as defined by the Voronoi tessellation around thiophene centers). The ML techniques are trained against some or all of these 700,000 chromophore pairs and their associated ZINDO/S calculations of $J_{i,j}$. The ML techniques are tested against 6.48 million chromophore pairs from 9 additional "disordered", 9 "semi-crystalline", and 9 "crystalline" morphologies.

7.3   Results and Discussion

In this section we first summarize the accuracy of five machine learning techniques for correlating our nine chosen structural features with $J_{i,j}$ calculated using ZINDO/S. We show that Random Forests are the optimal choice here for their ease of imple-

mentation and accuracy. We then evaluate the KMC charge mobility calculations from the RF-predicted $J_{i,j}$. We discuss the time saved through using RFs in place of ZINDO/S. Finally, we determine which features matter most for $J_{i,j}$ and investigate the relationship between the training set population and RF's prediction capabilities to understand the minimal information needed for accurate RF training

7.3.1 Comparison of ML techniques
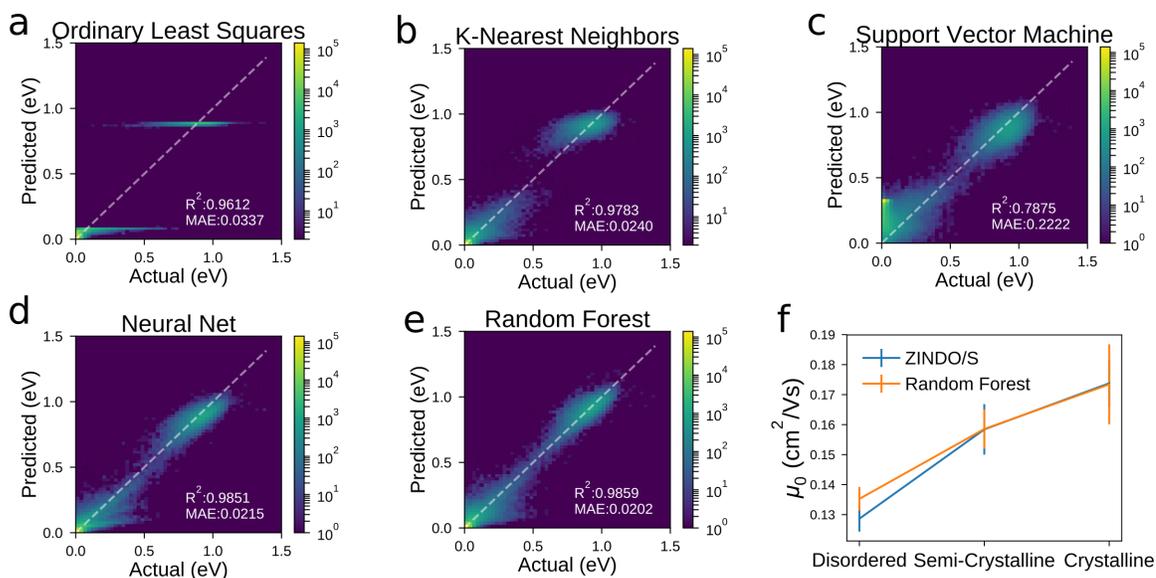
Figure 7.2: Accuracy of predictions of ZINDO/S $J_{i,j}$ from (a) OLS, (b) KNN, (c) SVM, (d) ANN, and (e) RF. The x-axes of each plot describe $J_{i,j}$ calculated with ZINDO/S and the y-axis corresponds to the predicted $J_{i,j}$ for a ML technique, with each chromophore pair from the training set occupying one pixel on these axes. The number of chromophore pairs at a particular location is represented by the purple-to-yellow colorbar. (f) The mobilities from RF $J_{i,j}$ are commensurate with those using ZINDO/S $J_{i,j}$. In the disordered morphology case, the RF-informed mobilities are $\sim$ 5% higher than ZINDO/S-informed mobilities. Error bars show the standard error of the mobility calculations.

Prediction accuracies of OLS, KNN, SVM, ANN, and RF techniques are shown in Figure 7.2. We orient the reader to two regions in each accuracy plot: There

is a cluster of bonded chromophore pairs with $0.6 < J_{i,j} < 1.1$ and a cluster of non-bonded pairs with $J_{i,j} < 0.5$. The more test pairs that are not on the diagonal line indicating perfect agreement between predicted and actual $J_{i,j}$, and the further their distance from the diagonal line of agreement, the worse the method. The poor predictive capabilities of OLS (Figure 7.2a), despite the surprisingly high $R^2 = 0.96$, suggests nonlinear relationships between features determines $J_{i,j}$. SVM accurately predicts bonded $J_{i,j}$ but fails when the chromophores are non-bonded (yellow region near $(Actual = 0, Predicted = 0.4)$). This results in a large number of $J_{i,j} \sim 0.4$ eV predictions for hops that should have zero coupling, leading to a low R^2 value and high mean-absolute-error (MAE). KNN provides predictions that are more accurate than OLS and SVMs and with better predictions of non-bonded pairs, but with over-prediction of bonded interactions, which can be seen as a "tail" extending above the perfect match diagonal around (Actual=0.6,Predicted=0.8). Both the RF and the ANN outperform the aforementioned techniques, with RF slightly outperforming ANN. Because the ANN has a larger number of hyper-parameters to tune (number of hidden layers, neurons per layer, activation function type, optimization method (See SI)) and is less accurate than RF, we focus on RFs henceforth.

7.3.2   Mobility Predictions

The predicted $J_{i,j}$'s from the random forest closely track the actual values, with an R^2 value of 0.986 and a MAE of 0.020 eV, though there exist outliers (Figure 7.2e). For example, the predicted average non-bonded $J_{i,j}$ value is slightly higher (.0015 eV) than the actual mean (<0.001 eV) (see SI-Section 2). With the ultimate goal of determining the efficacy of ML in predicting overall charge carrier mobilities through a morphology, we test the significance of these deviations by using predicted $J_{i,j}$

values in KMC simulations to calculate the final hole mobility for the system (Figure 7.2f). The mobilities calculated from the RF predictions are slightly higher than those determined with ZINDO/S for the disordered system. We hypothesize this over-prediction stems from our features incompletely describing structural perturbations that occur more frequently in disordered systems. For example, it is known that the dihedral angle between two chromophores will affect the charge transport along the chain[22], so trying out explicit dihedral angle features rather than the present combinations of rotations may provide marginal accuracy gains. Despite the small over-prediction of disordered P3HT mobility, the resultant mobilities are close (within 5% of ZINDO/S-informed mobilities), and follow the expected trend of increasing mobility with increasing crystallinity. These agreements are encouraging, as mobilities can vary by several orders-of-magnitude for different chemistries and processing conditions, and suggest that RF-predicted transfer integrals are an effective replacement for the relatively expensive ZINDO/S calculations.

7.3.3 Performance Benefit

To quantify the computational burden alleviated by using random forests we consider representative times for training the RF, generating $J_{i,j}$ with ZINDO/S for one morphology, and the frequency of calculating $J_{i,j}$ for multiple morphologies. Applying a trained RF to a representative system of $\sim$ 200,000 chromophore pairs (with unknown energy levels and transfer integrals) requires 4 minutes on an Intel Haswell CPU, compared to $\sim$ 26 CPU hours using Intel Xeon CPUs with ZINDO/S calculations. This factor of 390$\times$ speedup for a single simulation snapshot is multiplied in ensemble sampling studies: It is gained for each of the independent samples in an equilibrated simulation trajectory. This transferability of RFs trained across

disordered, semi-crystalline and crystalline P3HT demonstrates that a single RF can be used to accurately infer ensemble charge mobilities across hundreds of state points, each with hundreds of morphology snapshots. Using RFs therefore *enables* such screening studies, replacing 1.08×10^4 CPU-days of ZINDO/S calculations with 28 CPU-days of RF lookups.

7.3.4   RF Training Requirements

We consider here the minimal training set (the fewest ZINDO/S calculations) needed for accurate RF prediction of $J_{i,j}$, helping to gauge what "plenty of data to train against" means for the present problem. We evaluate the performance of several RFs, calibrated with different sizes of training data. In each case, the number of samples was selected randomly from the complete database of $\sim 700,000$ samples. Figure 7.3a shows that R^2 and MAE converge exponentially to high and low values, respectively, with as few as 100 training samples. The fast convergence is due to the algorithm quickly learning that bonded chromophores typically result in high $J_{i,j}$ (> 0.7 eV) and non-bonded chromophores resulting in low $J_{i,j}$ (< 0.3 eV).
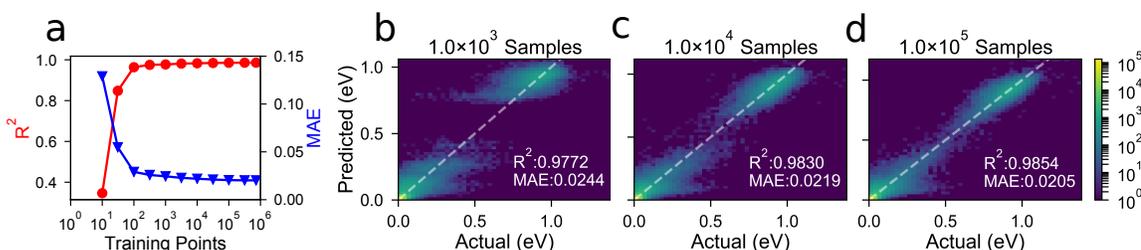
Figure 7.3: (a) Dependence of the R^2 and MAE on number of training examples shows that prediction accuracy converges around tens-of-thousands of pairs. (b - d) Despite relatively "good" R^2 and MAE values, significant deviations from the diagonal of perfect prediction are seen below $\sim 100,000$ training samples.

Although convergence to a fairly accurate prediction ($R^2 \sim 0.977$) is quickly achieved based on bonded/non-bonded chromophores, it can be seen in Figure 7.3 that with 1×10^3 samples, the distribution between bonded/non-bonded transfer integrals is bimodal, with high non-bonded $J_{i,j}$ and low bonded $J_{i,j}$ that occur in the range (0.4, 0.7) eV being missed. When 1×10^4 samples are used, the (0.4, 0.7) eV gap begins to fill in (Figure 7.3c), but it is not until 1×10^5 samples are used that the high/low non-bonded/bonded are correctly captured by the RF (Figure 7.3d). Extracting and training on these features from a simulation takes a negligible amount of time ($\sim$ 2 minutes for extracting, 14 seconds for training on 1×10^5 samples). The most expensive part of the process will be conducting the ZINDO/S calculations to train on these 1×10^5 samples ($\sim$ 13 hours).

7.3.5   Feature Comparison

We compare the relative importances of the nine features we currently use in predicting $J_{i,j}$, relying on the RF's advantage of feature transparency. Specifically, we use permutation importance, which compares the accuracy of the RF (R^2 value) on a validation set with true values and when the features' values have been shuffled. The importance is then the difference in R^2 caused by permuting that feature. The permutation mechanism is more computationally expensive than the mean decrease in impurity (or Gini importance) which is built into Scikit-Learn's RF algorithm but is more reliable. We note that the X, Y, and Z displacements are permuted in aggregate, i.e. in testing the X, Y, and Z importances, all three columns are permuted at the same time so that their importance relative the COM-COM feature can be better distinguished. The calculated feature importances, normalized to sum to one, are shown in Figure 7.4. By far, the most important feature in predicting $J_{i,j}$ is whether

or not two chromophores are directly bonded to each other. This is due to charges being delocalized over neighboring chromophores, which result in very high $J_{i,j}$ values. When the "bonded" feature is missing, many low, bonded $J_{i,j}$ are over-predicted and high non-bonded $J_{i,j}$ are under-predicted.
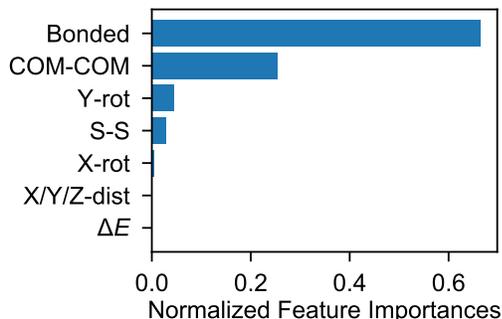
Figure 7.4: The feature importances for the RF algorithm. The X, Y, and Z distances are all combined into one feature importance.

In Figure 7.5 we summarize the prediction accuracies of RFs trained, but with select features omitted from the training sets. The biggest deviation from champion accuracy ($R^2 = 0.9858$) is observed when the bonded feature is omitted, as expected. Removing the COM-COM feature results in an over-prediction of the "bonded" $J_{i,j}$ values–transfer integrals in the 0.8-1.0 eV region are shifted closer to 1.0 eV (Figure 7.5b). The importance of having close chromophores is somewhat unsurprising as the transfer integrals decreases rapidly as the two chromophores move away from each other.[3,7,9,22] We note that the COM-COM feature is directly dependent on the X, Y, and Z displacements as it is the square-root of the squared-sums of the X, Y, and Z offsets. Although it is a composite feature, explicitly training on the COM-COM distance is very important for predicting the $J_{i,j}$. The individual X-, Y-, and Z-dist features have negligible feature importance, even when permuted in aggregate (Figure 7.4). This is likely to be due to the small size and relative symmetry of the thiophene

ring, and the non-linear relationship between the individual features and the aggregate COM-COM feature. If larger or asymmetric chromophores were used, such as a coronene or a perylene derivative, the displacements along the different axes are likely to dominate and increase relative feature importance (see Figure 7.5c) [7].
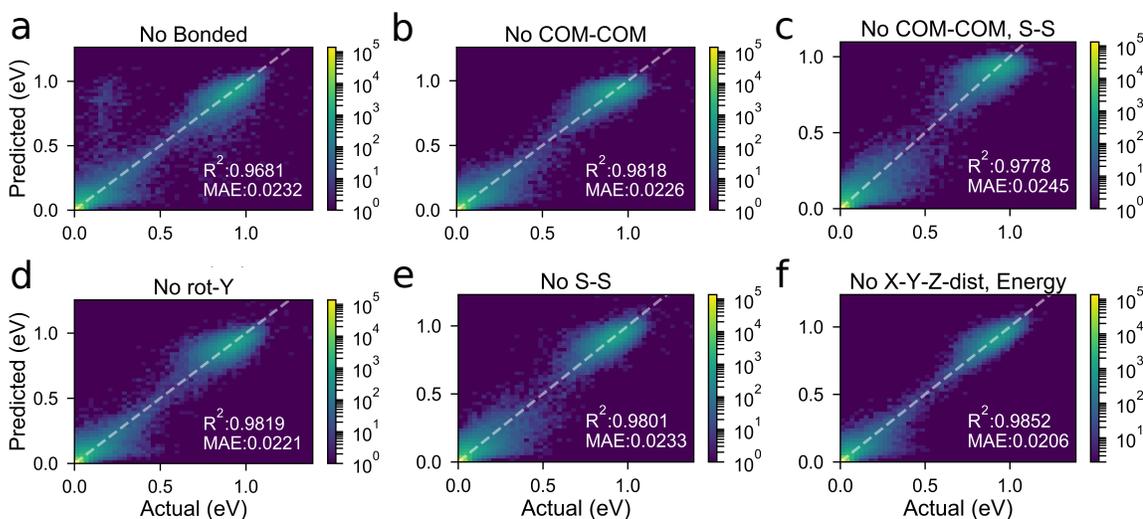
Figure 7.5: (a) Removing the Bonded feature results in a high number of outliers as both bonded and non-bonded $J_{i,j}$ values are under and over predicted. (b) Removing the COM-COM constraint results in a flattening and broadening of the "bonded" $J_{i,j}$ distribution. (c) If both the COM-COM and S-S distances are removed (and therefore only the displacements along the X, Y and Z axes are considered) the distribution of $J_{i,j}$ is much more split between "bonded" and "non-bonded". (d, e) Removing the rotation around Y and the S-S distance create more noise. (f) The X, Y, and Z displacements and the $\Delta E_{i,j}$ can all be omitted in training and result in high accuracies.

Relative rotation around the Y-axis ("pitch") is the third most important feature, and is more important than rotation around the X-axis ("roll") (Figure 7.5d). This is likely because rotations around Y move the sulfur atom in the ring, as opposed to rotations around X in which the sulfur is stationary. The importance of the relative sulfur positions is further highlighted by the S-S distance being the fourth most important feature, and this feature is responsible for obtaining correct predictions for

high non-bonded $J_{i,j}$ and low bonded $J_{i,j}$ (Figure 7.5e). This indicates that in order to have high $J_{i,j}$, electronegative atoms within the chromophores must be proximal in order to act as bridges between the two chromophores.

The $\Delta E_{i,j}$ feature in this experiment is unimportant for predicting $J_{i,j}$. This unimportance is not surprising as the MD simulations represent the thiophene ring with a rigid-body, which means the relative positions of all the atoms in the ring are fixed throughout the simulation. With this model, differences in energies can only arise based on conformational differences of the aliphatic tails. The effect of these tails on energy is likely to be small, and many studies omit the tails as a way to reduce computational burden and still obtain correct results. Consequently, if $\Delta E_{i,j}$ is small compared to the HOMO and HOMO-1 splitting in Equation 7.1, it becomes negligible for $J_{i,j}$. If flexible thiophene rings were used, the importance of the $\Delta E_{i,j}$ feature would increase (although thiophene ring perturbations are still likely to be small because of the aromatic structure of the ring). Despite the insignificance of $\Delta E_{i,j}$ in predicting $J_{i,j}$, we do not argue that $\Delta E_{i,j}$ will be unimportant for predicting mobility values as Equation 2.13 explicitly considers $\Delta E_{i,j}$ within an exponential and it will likely still have non-negligible effects on the hopping rate. Here, we show that omitting the X-Y-Z displacements and $\Delta E_{i,j}$ features entirely has a negligible effect on the accuracy of only our $J_{i,j}$ predictions (Figure 7.5f).

7.3.6   Curating A Training Set

Here we consider the possibility of curating a "universal" training set of chromophore pairs that inform an RF with predictive capabilities for P3HT morphologies with disparate degrees of order. To curate the training data, we duplicate a chromophore (parent) to create a child chromophore, resulting in all $\Delta E_{i,j}$ values being 0. The child
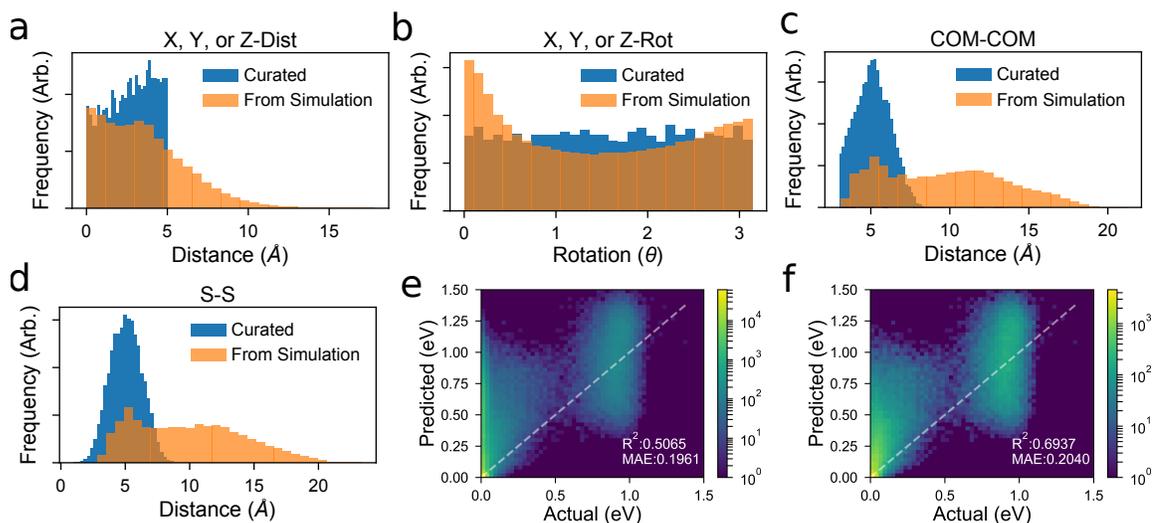
Figure 7.6: The normalized distributions of the training features in the curated set (blue) and in the simulation (orange). (a) Displacements and (b) Rotations are determined based on a uniform distribution along each axis; $0 \leq$ distance ≤ 5 Å for displacements along the axes and $0 \leq$ rotation $\leq \pi$ for rotations and show that a uniform distribution fails to capture more energetically favorable close configurations and similar alignments. (c) COM-COM and (d) S-S distances are then calculated based on the displacements along the various axes. The curated training set does a poor job of predicting $J_{i,j}$ whether (e) the curated set is used to evaluate all chromophore pairs or (f) the chromophore pairs that lie within the range of the curated set.

chromophore is then moved along each axis (≤ 5 Å) and rotated around the x- and y-axes ($\leq \pi$) resulting in 1×10^4 training pairs. The child movement and rotation is done in two ways: at distinct steps, e.g. steps of 0, 1, 2 Å and uniformly distributed over the range (shown in Figure 7.6). For each offset we apply the constraint that the COM-COM distance must be greater than 3 Å, as COM-COM distances shorter than this are unphysical. With this uniform sampling of positions and orientations, close packings and large separations observed in simulations are underrepresented (Figure 7.6a), as are aligned and anti-aligned orientations of thiophene rings (Figure 7.6b). We expect that the undersampling of pi-stacked configurations will most

negatively impact accuracy, as $J_{i,j}$ is negligible for large separations. This data curation generates COM-COM and S-S distributions similar in shape around 5Å, though missing pairs separated at larger distances that are observed in simulations (Figure 7.6c and d). Though these larger spacings are prevalent in the simulated structures, we find they contribute negligibly to charge transport.

We train the RF using this curated training set and validate it against the simulation produced $J_{i,j}$. As is seen in Figure 7.6e, the RF trained on the curated set does a poor job of predicting $J_{i,j}$. The largest error in the predictions arises from the over-prediction of the low (≤ 0.2 eV) $J_{i,j}$ in the system. This error can be reduced somewhat by considering only chromophore pairs that lie within the range of the curated dataset (within 5 Å along each axis). This restriction of the validation data improves the R^2 value ($0.5 \rightarrow 0.7$) while the MAE decreases slightly (both ~ 0.2 eV), however, will come at the cost of missing long-range pairs or inflating/diluting the training set with pairs that are likely to be negligibly small. Despite the small improvement, these curated data provide low predictive utility (Figure 7.6f). This failure of the curated set serves as a reminder that equilibrium simulations efficiently perform importance samplings of configurations, and that a uniform sampling of configurations in a similar range is an insufficient proxy for those configurations that matter most. Related, if training samples are selected from only a single simulation snapshot, it is best here to select them from crystalline morphologies because the relative absence of high $J_{i,j}$ in other morphologies disproportionately lowers the RF prediction accuracy (SI-Section 3).

7.4 Conclusion

The expensive quantum chemical calculation of electronic couplings ($J_{i,j}$) between P3HT chromophores need not be repeated if a representative training set of chromophores is used to train a machine to infer the couplings from chromophore features. We have shown that artificial neural networks and random forests are sufficiently predictive of $J_{i,j}$, resulting in expected charge mobilities. Here, random forests are recommended over artificial neural networks because we begin with a physical intuition for the features salient to $J_{i,j}$, so the RF ability to transparently rank feature importances and the ease of implementing RFs in Scikit-Learn give benefits at no added cost. We show that $J_{i,j}$ are obtained $\sim 390\times$ faster when the RF is used to look up ZINDO/S calculations, and we identify chromophore bonding, distance, "pitch", and sulfur-separation between chromophores to be the strongest predictors. Two conclusions arose from our investigations into minimal training sets: (1) The failure to accurately predict $J_{i,j}$ from a training set curated on chromophore separations and rotations informed by the ranked feature importance highlights the importance of drawing training data from a thermodynamic simulation method in which importance samplings of configurations are performed, and (2) Training sets as small as 1×10^5 chromophore pairs are sufficient to generate $J_{i,j}$ and resultant mobilities in agreement with prior work. In sum, this work demonstrates one example of where significant computational speedups can be gained in exchange for a small amount of machine learning tuning. In future work we look towards identifying other bottlenecks where RFs and ANNs will provide similar speedups, towards the automatic identification of molecular descriptors that allow the prediction of $\Delta E_{i,j}$, and extending this work to additional chemistries.

BIBLIOGRAPHY

[1] Martín Abadi, Ashish Agarwal, Paul Barham, Eugene Brevdo, Zhifeng Chen, Craig Citro, Greg S. Corrado, Andy Davis, Jeffrey Dean, Matthieu Devin, Sanjay Ghemawat, Ian Goodfellow, Andrew Harp, Geoffrey Irving, Michael Isard, Yangqing Jia, Rafal Jozefowicz, Lukasz Kaiser, Manjunath Kudlur, Josh Levenberg, Dan Mane, Rajat Monga, Sherry Moore, Derek Murray, Chris Olah, Mike Schuster, Jonathon Shlens, Benoit Steiner, Ilya Sutskever, Kunal Talwar, Paul Tucker, Vincent Vanhoucke, Vijay Vasudevan, Fernanda Viegas, Oriol Vinyals, Pete Warden, Martin Wattenberg, Martin Wicke, Yuan Yu, and Xiaoqiang Zheng. TensorFlow: Large-Scale Machine Learning on Heterogeneous Distributed Systems, 2016.

[2] Amy M. Ballantyne, Lichun Chen, Justin Dane, Thomas Hammant, Felix M. Braun, Martin Heeney, Warren Duffy, Iain McCulloch, Donal D. C. Bradley, and Jenny Nelson. The Effect of Poly(3-hexylthiophene) Molecular Weight on Charge Transport and the Performance of Polymer:Fullerene Solar Cells. *Advanced Functional Materials*, 18(16):2373–2380, 2008.

[3] J. L. Bredas, J. P. Calbert, D. A. da Silva Filho, and J. Cornil. Organic Semiconductors: A Theoretical Characterization of the Basic Parameters Governing Charge Transport. *Proceedings of the National Academy of Sciences*, 99(9):5804–5809, 2002.

[4] Jean-Luc Brédas, David Beljonne, Veaceslav Coropceanu, and Jérôme Cornil. Charge-Transfer and Energy-Transfer Processes in π-Conjugated Oligomers and Polymers: A Molecular Picture. *Chemical Reviews*, 104(11):4971–5004, 2004.

[5] M. Brinkmann and P. Rannou. Effect of Molecular Weight on the Structure and Morphology of Oriented Thin Films of Regioregular Poly(3-hexylthiophene) Grown by Directional Epitaxial Solidification. *Advanced Functional Materials*, 17(1):101–108, 2007.

[6] Jui Fen Chang, Baoquan Sun, Dag W. Breiby, Martin M. Nielsen, Theis I. Sölling, Mark Giles, Iain McCulloch, and Henning Sirringhaus. Enhanced Mobility of Poly(3-hexylthiophene) Transistors by Spin-Coating from High-Boiling-Point Solvents. *Chemistry of Materials*, 16(23):4772–4776, 2004.

[7] Veaceslav Coropceanu, Jérôme Cornil, Demetrio A. da Silva Filho, Yoann Olivier, Robert Silbey, and Jean-Luc Brédas. Charge Transport in Organic Semiconductors. *Chemical Reviews*, 107(4):926–952, 2007.

[8] Minh Trung Dang, Lionel Hirsch, and Guillaume Wantz. P3HT:PCBM, Best Seller in Polymer Photovoltaic Research. *Advanced Materials*, 23(31):3597–3602, 2011.

[9] Wei-Qiao Deng and William A. Goddard. Predictions of Hole Mobilities in Oligoacene Organic Semiconductors from Quantum Mechanical Calculations. *The Journal of Physical Chemistry B*, 108(25):8614–8621, 2004.

[10] Nieves Espinosa, Markus Hösel, Dechan Angmo, and Frederik C. Krebs. Solar Cells with One-Day Energy Payback for the Factories of the Future. *Energy & Environmental Science*, 5(1):5117, 2012.

[11] Andrew L Ferguson. Machine Learning and Data Science in Soft Materials Engineering. *Journal of Physics: Condensed Matter*, 30(4):043002, 2018.

[12] Rafael Gómez-Bombarelli, Jorge Aguilera-Iparraguirre, Timothy D. Hirzel, David Duvenaud, Dougal Maclaurin, Martin A. Blood-Forsythe, Hyun Sik Chae, Markus Einzinger, Dong-Gwang Ha, Tony Wu, Georgios Markopoulos, Soonok Jeon, Hosuk Kang, Hiroshi Miyazaki, Masaki Numata, Sunghan Kim, Wenliang Huang, Seong Ik Hong, Marc Baldo, Ryan P. Adams, and Alán Aspuru-Guzik. Design of Efficient Molecular Organic Light-Emitting Diodes by a High-Throughput Virtual Screening and Experimental Approach. *Nature Materials*, 1(15):1120–12127, 2016.

[13] Cristina Greco, Anton Melnyk, Kurt Kremer, Denis Andrienko, and Kostas Ch. Daoulas. Generic Model for Lamellar Self-Assembly in Conjugated Polymers: Linking Mesoscopic Morphology and Charge Transport in P3HT. *Macromolecules*, 52:968–981, 2019.

[14] Nicholas E. Jackson, Alec S. Bowen, Lucas W. Antony, Michael A. Webb, Venkatram Vishwanath, and Juan J. de Pablo. Electronic Structure at Coarse-Grained Resolutions from Supervised Machine Learning. *Science Advances*, 5(3):1–38, 2019.

[15] M. L. Jones. MorphCT, 2018.

[16] M. L. Jones, D. M. Huang, B. Chakrabarti, and Chris Groves. Relating Molecular Morphology to Charge Mobility in Semicrystalline Conjugated Polymers. *Journal of Physical Chemistry C*, 120(8):4240–4250, 2016.

[17] Matthew L. Jones, Evan D. Miller, Eric Jankowski, Matthew L. Jones, and Eric Jankowski. Machine Learning for Structure-Performance Relationships in Organic Semiconducting Devices. *Computational Materials Engineering Laboratory*, 3, 2018.

[18] Matthew Lewis Jones and Eric Jankowski. Computationally Connecting Organic Photovoltaic Performance to Atomistic Arrangements and Bulk Morphology. *Molecular Simulation*, 43(10-11):1–18, 2017.

[19] Youngkyoo Kim, Steffan Cook, Sachetan M. Tuladhar, Stelios A. Choulis, Jenny Nelson, James R. Durrant, Donal D. C. Bradley, Mark Giles, Iain McCulloch, Chang-Sik Sik Ha, and Moonhor Ree. A Strong Regioregularity Effect in Self-Organizing Conjugated Polymer Films and High-Efficiency Polythiophene:Fullerene Solar Cells. *Nature Materials*, 5(3):197–203, 2006.

[20] James Kirkpatrick. An Approximate Method for Calculating Transfer Integrals Based on the ZINDO Hamiltonian. *International Journal of Quantum Chemistry*, 108(1):51–56, 2008.

[21] R. J. Kline, M. D. McGehee, E. N. Kadnikova, Jinsong Liu, and J. M. J. Fréchet. Controlling the Field-Effect Mobility of Regioregular Polythiophene by Changing the Molecular Weight. *Advanced Materials*, 15(18):1519–1522, 2003.

[22] Yi-kang Lan and Ching-i Huang. A Theoretical Study of the Charge Transfer Behavior of the Highly Regioregular Poly-3-hexylthiophene in the Ordered State. *The Journal of Physical Chemistry B*, 112(47):14857–14862, 2008.

[23] Mitchell H. Leibowitz, Evan D. Miller, Michael M. Henry, and Eric Jankowski. Application of Artificial Neural Networks to Identify Equilibration in Computer Simulations. *Journal of Physics: Conference Series*, 921(1):012013, 2017.

[24] Yue Liu, Tianlu Zhao, Wangwei Ju, Siqi Shi, Siqi Shi, and Siqi Shi. Materials Discovery and Design Using Machine Learning. *Journal of Materiomics*, 3(3):159–177, 2017.

[25] W Ma, Cuiying Yang, Xiong Gong, Kwanghee Lee, and Alan J. Heeger. Thermally Stable, Efficient Polymer Solar Cells with Nanoscale Control of the Interpenetrating Network Morphology. *Advanced Functional Materials*, 15(10):1617–1622, 2005.

[26] R. A. Marcus. On the Theory of Oxidation-Reduction Reactions Involving Electron Transfer. I. *The Journal of Chemical Physics*, 24(5):966, 1956.

[27] Ralf Mauer, Marcel Kastler, and Frédéric Laquai. The Impact of Polymer Regioregularity on Charge Transport and Efficiency of P3HT:PCBM Photovoltaic Devices. *Advanced Functional Materials*, 20(13):2085–2092, 2010.

[28] Katherine A. Mazzio and Christine K Luscombe. The Future of Organic Photovoltaics. *Chemical Society Reviews*, 44(1):78–90, 2015.

[29] Lingxian Meng, Yamin Zhang, Xiangjian Wan, Chenxi Li, Xin Zhang, Yanbo Wang, Xin Ke, Zuo Xiao, Liming Ding, Ruoxi Xia, Hin-Lap Yip, Yong Cao, and Yongsheng Chen. Organic and Solution-Processed Tandem Solar Cells With 17.3% Efficiency. *Science*, 2612:1094–1098, 2018.

[30] Evan Miller, Matthew Jones, Michael Henry, Paul Chery, Kyle Miller, and Eric Jankowski. Optimization and Validation of Efficient Models for Predicting Polythiophene Self-Assembly. *Polymers*, 10(12):1305, 2018.

[31] Evan Miller, Matthew Jones, and Eric Jankowski. Tying Together Multiscale Calculations for Charge Transport in P3HT: Structural Descriptors, Morphology, and Tie-Chains. *Polymers*, 10(12):1358, 2018.

[32] Evan Miller, Matthew Jones, and Eric Jankowski. Machine Learning Predictions of Electronic Couplings for Charge Transport Calculations - Code, 2019.

[33] T. Mueller, A. Kusne, and R. Ramprasad. Machine Learning in Materials Science: Recent Progress and Emerging Applications. In Abby L. Parrill and Kenny B. Lipkowitz, editors, *Reviews in Computational Chemistry*, number 1, chapter 4, pages 186–273. John Wiley & Sons, 2016.

[34] F. Neese. The ORCA Program System. *Wiley Interdisciplinary Reviews: Computational Molecular Science*, 2(1):73–78, 2012.

[35] Shyam S. Pandey, Wataru Takashima, Shuichi Nagamatsu, Takeshi Endo, Masahiro Rikukawa, and Keiichi Kaneto. Regioregularity vs Regiorandomness: Effect on Photocarrier Transport in Poly(3-hexylthiophene). *Japanese Journal of Applied Physics*, 39(Part 2, No. 2A):L94–L97, 2000.

[36] Fabian Pedregosa, Gaël Varoquaux, Alexandre Gramfort, Vincent Michel, Bertrand Thirion, Olivier Grisel, Mathieu Blondel, Peter Prettenhofer, Ron Weiss, Vincent Dubourg, Jake Vanderplas, Alexandre Passos, David Cournapeau, Matthieu Brucher, Matthieu Perrot, and Édouard Duchesnay. Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research*, 12:2825–2830, 2012.

[37] Ghanshyam Pilania, Chenchen Wang, Xun Jiang, Sanguthevar Rajasekaran, and Ramamurthy Ramprasad. Accelerating Materials Property Predictions using Machine Learning. *Scientific Reports*, 3:1–6, 2013.

[38] Edward O. Pyzer-Knapp, Kewei Li, and Alan Aspuru-Guzik. Learning from the Harvard Clean Energy Project: The Use of Neural Networks to Accelerate Materials Discovery. *Advanced Functional Materials*, 25(41):6495–6502, 2015.

[39] Rampi Ramprasad, Rohit Batra, Ghanshyam Pilania, Arun Mannodi-Kanakkithodi, and Chiho Kim. Machine Learning in Materials Informatics: Recent Applications and Prospects. *npj Computational Materials*, 3(1):1–27, 2017.

[40] J. Ridley and Michael Zerner. An Intermediate Neglect of Differential Overlap Technique for Spectroscopy: Pyrrole and the Azines. *Theoretica Chimica Acta*, 32(2):111–134, 1973.

[41] H. Sirringhaus, P. J. Brown, R. H. Friend, M. M. Nielsen, K. Bechgaard, B. M. W. Langeveld-Voss, a. J. H. Spiering, R. A. J. Janssen, E. W. Meijer, P. Herwig, and D. M. de Leeuw. Two-Dimensional Charge Transport in Self-Organized, High-Mobility Conjugated Polymers. *Nature*, 401(6754):685–688, 1999.

[42] Ellen Van, Matthew Jones, Eric Jankowski, and Olga Wodo. Using Graphs to Quantify Energetic and Structural Order in Semicrystalline Oligothiophene Thin Films. *Molecular Systems Design & Engineering*, 1:273–277, 2018.

[43] Koen Vandewal, Scott Himmelberger, and Alberto Salleo. Structural Factors that Affect the Performance of Organic Bulk Heterojunction Solar Cells. *Macromolecules*, 46:6379–6387, 2013.

[44] Venkat Venkatasubramanian. The Promise of Artificial Intelligence in Chemical Engineering: Is it Here, Finally? *AIChE Journal*, 65(2):466–478, 2019.

[45] Xueping Yi, Bhoj Gautam, Iordania Constantinou, Yuanhang Cheng, Zhengxing Peng, Erik Klump, Xiaochu Ba, Carr Hoi Yi Ho, Chen Dong, Seth R. Marder, John R. Reynolds, Sai Wing Tsang, Harald Ade, and Franky So. Impact of Nonfullerene Molecular Architecture on Charge Generation, Transport, and Morphology in PTB7-Th-Based Organic Solar Cells. *Advanced Functional Materials*, 28(32):1–9, 2018.

[46] Jun Yuan, Yunqiang Zhang, Liuyang Zhou, Guichuan Zhang, Hin-Lap Yip, Tsz-Ki Lau, Xinhui Lu, Can Zhu, Hongjian Peng, Paul A. Johnson, Mario Leclerc, Yong Cao, Jacek Ulanski, Yongfang Li, and Yingping Zou. Single-Junction Organic Solar Cell with over 15% Efficiency Using Fused-Ring Acceptor with Electron-Deficient Core. *Joule*, pages 1–12, 2019.

CHAPTER 8

EFFICIENTLY SIMULATING ARBITRARY CHEMISTRIES AND MIXTURES: PLANCKTON

8.1 Introduction

So far, we have explored a few different types of OPV molecules. However, these chemistries are only the proverbial "scratching the surface" of all OPV candidate compounds. Additionally, predicting self-assembly in these systems had their own-pitfalls and challenges. For example, simulating perylene was approached from a "minimal physics" mindset i.e. sensible default values were selected, but extensive model optimizations were not conducted[20]. This methodology was able to reproduce the bulk structure of experimental perylene, but the exact crystal structure was not predicted. Conversely, systematically tuning the interactions in P3HT resulted in experimentally accurate structures but is not extensible to other chemistries[17]. Both of these studies also relied on in-house code to generate the compounds, initialize systems and manage simulations. In short, these studies have limited "*TRUE*-ness" which hamper efficiently simulating many OPV chemistries.

To obtain more *TRUE* OPV simulations, we have developed a new Python package: `Planckton`. The goal of `Planckton` is to build on peer-reviewed tools to simulate new OPV chemistries[19]. The first challenge in simulating many chemistries is having

an automated way to obtain accurate, verified force-field parameters. As noted in Section 2.2.1, there are many peer-reviewed force-fields that have been developed, however, assigning these force-fields is tedious for a human to do. The challenge in transferring force-fields between systems has been noted by other researchers in the simulation community and has lead to the creation of an automatic force-field parser and assigner: Foyer[11]. In addition to incorporating Foyer into our work flow, we also replace in-house code used to initialize structures and manage simulations with peer-reviewed tools developed in the simulation community[2,3,10].

In this chapter, we present a preliminary self-assembly—charge transport study using `Planckton` to predict the structures. To verify `Planckton` we employ simulations with a Foyer typed and generated P3HT model and compare the results to structures generated in Section 4.4.2, which used our modified OPLS-UA force-field. We then investigate the self-assembly and charge transport of a class of electron acceptors that has generated significant interest in recent years[24].
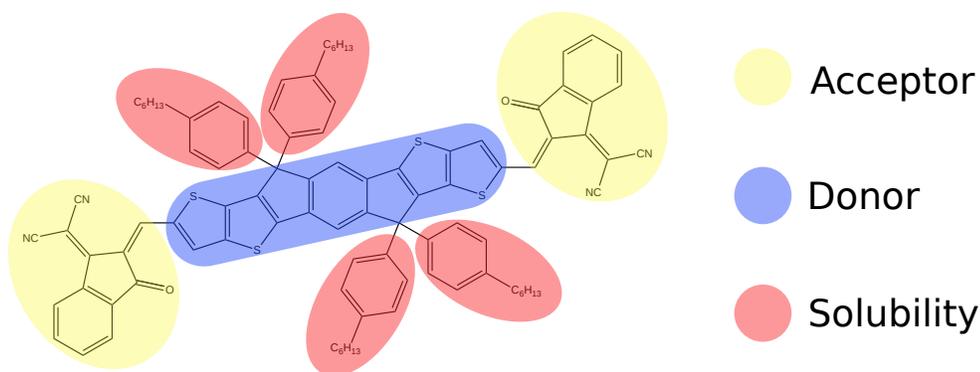
Figure 8.1: The three regions of ITIC that act as an electron acceptor (yellow) and electron donor (blue) and control solubility (red).

These acceptor compounds have increased η in OPVs over 15% in ternary blend, single junction devices[29] and over 17% in multi-junction devices[16]. These com-

pounds are often derivatives of, or inspired by, 3,9-bis(2-methylene-(3-(1,1-dicyanomethylene)-indanone))- 5,5,11,11-tetrakis (4-hexylphenyl)-dithieno[2,3-d:2',3'-d']-s-indaceno [1,2-b:5,6-b'] dithiophene (ITIC). ITIC was first synthesized by Lin et. al in 2015[14], and ITIC derivatives are desirable due to their better absorption of visible wavelengths and better energy level variability, compared to fullerene acceptors[5]. ITIC derivatives are composed of three portions: electron accepting end-capping groups, electron donor core group and core functionalizations to control solubility (see Figure 8.1)[27]. These molecules are "ladder-type" in that they are composed of longer aromatic bodies (see Figure 8.2). Here we focus on two main types: Those with 180° rotational symmetry (C_2) and those with 120° rotational symmetry (C_3, Truxenes). ITIC derivatives are subject to the same processing-performance relationships as other OPV compounds[21] and are prime candidates to act as a test-case for `Planckton`'s ability to type, initialize and run simulations.

8.2 Methods

8.2.1 MD with `Planckton`

To run an MD simulation with `Planckton`, there are four steps:

1. Assign compound force-fields with Foyer

2. Initialize simulations with mBuild

3. Create a simulation state-point with Signac

4. Distribute the simulation on a HPC cluster with Signac-Flow

In the first step, we use Foyer to assign force-field parameters by parsing a molecule and assigning atom interactions from a tabulated force-field file. To assign these
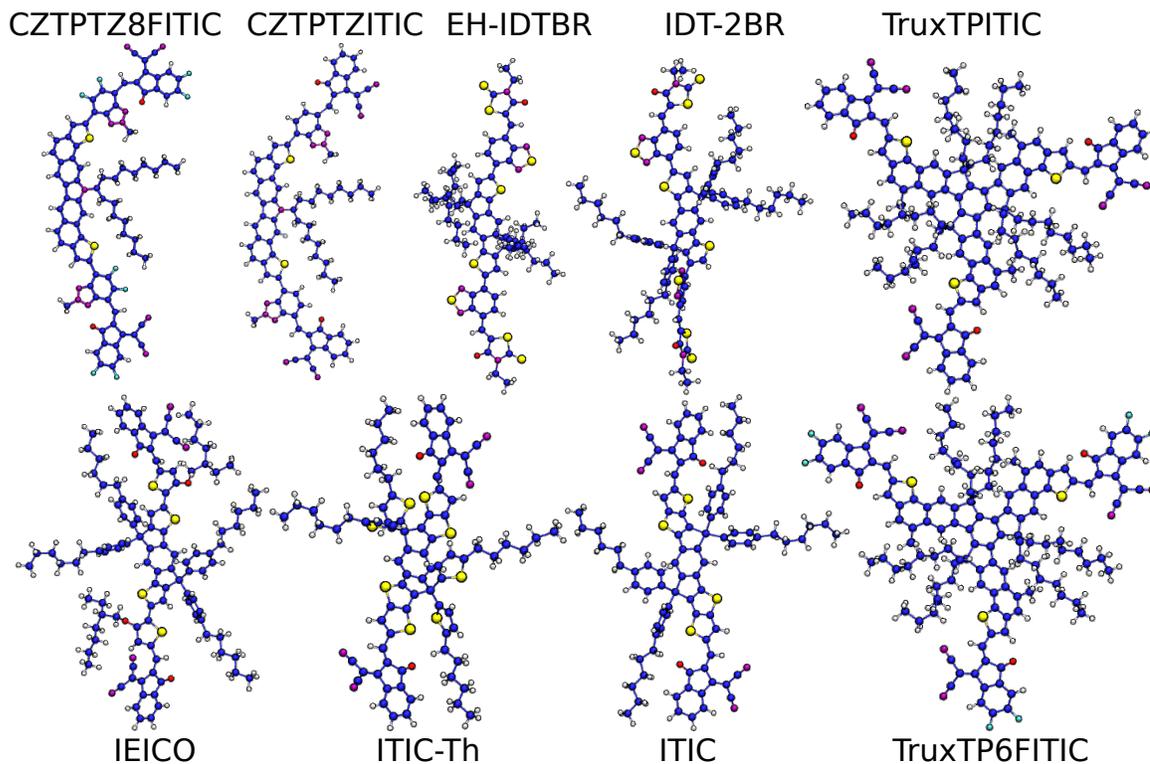
Figure 8.2: The nine electron acceptor types that are automatically typed and simulated with `Planckton`. Hydrogens are white, carbons are dark blue, sulfurs are yellow, nitrogens are magenta, oxygens are red and fluorines are light blue.

interactions, Foyer must know the "atom types" which describes the elements and their chemical contexts. For example, as reported by Klein, the element carbon has ~ 350 "atom types" in the OPLS, GROMACS force-field[11]. To encode atom type information, Foyer can use SMARTS strings, which define chemical patterns[1]. These SMARTS strings can be generated manually (which is unfavored), or alternatively, a program such as Antechamber[25] which considers the element and bonds to assign the atom types which are then interpreted by Foyer[7]. Once atom types are determined for the input molecule, Foyer assigns the atom interactions. In this study, we use GAFF (general Amber force-field)[26], which was the most stable in preliminary tests.

In the second step, mBuild initializes a starting configuration for the simula-

tion[10]. This initialization is accomplished by mBuild placing N copies of the input molecule(s) at random positions within the simulation "box" while avoiding molecule overlap. We use a simulation box larger than the target density (by a factor of 5) so that the molecules can be packed quickly. We also add an additional step during the system initialization: Hydrogen atom removal. This is done to reduce the computational burden arising from many hydrogen atoms, similar to UA models. In ITIC this results in the removal of 82 simulation elements ($186 \rightarrow 104$), a reduction of $\sim 40\%$. However, we note that the naïve removal of hydrogen atoms is not the same as an UA model i.e. UA models implicitly account for hydrogen presence by tuning the other interactions, which tuning is not done here.

During the third step, we create a "statepoint" for the simulation based on the thermodynamic variables. Signac manages statepoints by creating a unique hash directory for each statepoint and then keeping a record of which statepoint and hash belong together[2,3]. As thermodynamic variables defining the statepoint, we use the molecule, density ρ, temperature T and "solvent quality" ε_s. Each statepoint is then distributed over the HPC cluster in the fourth step with Signac-Flow. All simulations are run using the HOOMD-blue[4,6] MD package on NVIDIA K80, P100 and V100 GPUs. Simulations are done in the canonical, NVT ensemble in which the total number of particles, volume and temperature are held constant. This temperature is controlled with the Nosé-Hoover thermostat[9], and particle positions are updated with the two-step velocity-Verlet integration of Newton's equations of motion[23]. For the results presented here, a dimensionless time unit of $dt = 0.001$ is used, and with $\tau_s = 1.94 \times 10^{-12}$s , this corresponds to a timestep of 1.94 fs. Because in Step 2, simulations are initialized into volumes larger than the target density, each simulation undergoes a "shrink step" in which the simulation volume is reduced to the desired

density over 2 ns at ~ 1250 K; producing a randomized initial configuration at the target density.

8.2.2   Characterization

We characterize these systems with the same methodologies presented in previous chapters: GIXS patterns, $g(r)$ and clustering based on COM-COM distances. Because each ITIC derivative have a different indexing of atoms, atom types and positions it is difficult to identify distinct molecular regions *a priori* as was done in P3HT to use in $g(r)$ and clustering. To overcome this, we segment each ITIC derivative into different regions using K-Means clustering of the particle positions, which assigns the atoms to a group based on euclidean distance to the nearest centroid. These centroids are iteratively moved to minimize the average distance to the particles' assigned centroid until the grouping converges. Using this methodology (with four centroids for Truxenes and thee centroids for ITIC compounds), we show the segmentation determined in Figure 8.3b. We note that this methodology is not ideal, and better optimizations are likely to be found, regardless this allows for convenient assignments without human involvement. Using these segmentations, we calculate the COMs of each segment and use the COMs for clustering groups of molecules and $g(r)$.

8.3   Results and Discussion

8.3.1   P3HT Comparison

We begin our discussion of the structures produced with `Planckton` by comparing the automatically typed P3HT structures with those seen in Section 4. The non-bonded interactions are shown in Table 8.1. Foyer (from the Antechamber typing) assigns
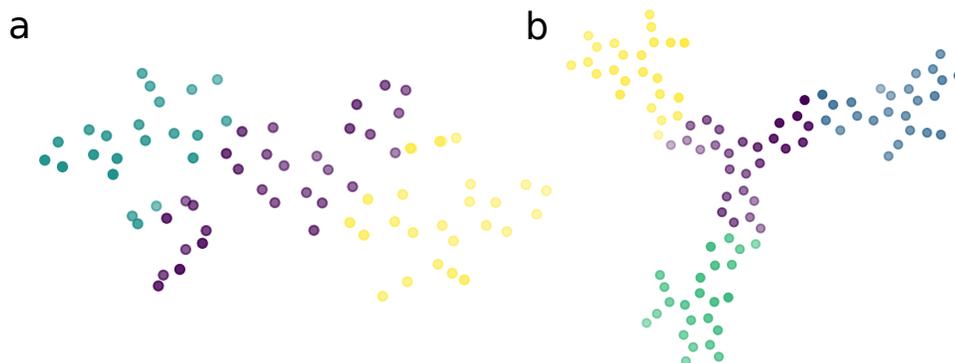
Figure 8.3: The K-means separation into different segments for (a) ITIC and (b) TruxTP6FITIC compound with the donor (purple) and different acceptor portions. Although separated in the figure, all acceptor portions are treated equally in the analysis.

Element	σ_{LJ} (Å)	M (amu)	ε (kJ/mol)
Aliphatic Carbon (c3)	3.340	12.01	0.458
Aromatic Carbon (ca)	3.340	12.01	0.360
Aromatic Carbon (cc)	3.340	12.01	0.360
Sulfur (ss)	3.563	32.06	1.046

Table 8.1: GAFF non-bonded parameters for the various simulation elements predicted with Foyer.

interactions for four atom types in the simulation: sulfurs (ss), aliphatic carbons (c3) and two kinds of aromatic carbons (ca, cc). The ca and cc carbon types have the same non-bonded interactions, despite being typed differently. The distinction is likely needed to distinguish inter and intra thiophene bonded constraints. The aromatic carbons are similar in size to the CA beads used in Section 4.2, but the interactions for these aromatic carbons are weaker than the model used in Section 4.2 (0.360 kJ/mol versus 0.460 kJ/mol). The aliphatic carbons are significantly smaller with the Foyer predicted interactions (3.340 Å for c3 versus 3.905 Å for CT), however, this is not surprising as the CT atoms were "true" UA and interactions were tuned to implicitly

consider hydrogen atoms. Additionally, the c3 carbon interactions are weaker than the CT carbons (0.458 kJ/mol for c3 and 0.711 kJ/mol for CT). Conversely, for sulfur, the ss σ_{LJ} values are larger than that of S (3.563 vs. 3.436 Å), but the interactions are weaker (1.045 kJ/mol vs. 1.34 kJ/mol). In total, the non-bonded interactions are weaker in the Foyer system than they were in the UA system of Section 4.2.
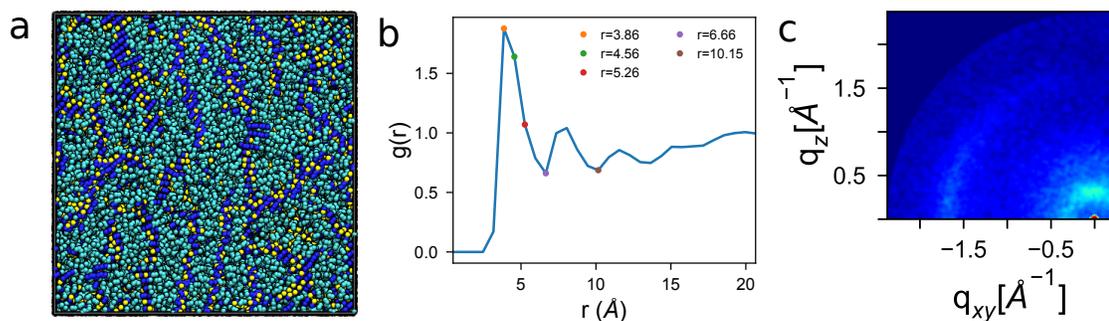
Figure 8.4: (a) The self-assembled P3HT morphology predicted with automatic typing and force-field assignment with `Planckton`. Yellow beads represent sulfur, dark-blue beads represent aromatic carbons and cyan represents aliphatic carbons. (b) The $g(r)$ of the thiophene centers for the morphology in (a). (c) The GIXS pattern for the morphology in (a), averaged over 13 rotated patterns selected for showing π- and aliphatic-stacking.

We initialize and simulate the self-assembly with the `Planckton` predicted interactions. The `Planckton` and the previously used UA models have comparable TPS performance (~ 350 TPS), however, we note that the `Planckton` simulations are done with a more recent version of the HOOMD-Blue software package (1.3.3 vs. 2.3.5, both on Nvidia P100 GPUs). An "ordered" simulation snapshot is shown in Figure 8.4a in which sulfurs are yellow, aliphatic carbons are cyan and aromatic carbons are dark blue. The morphology in Figure 8.4a is visually less ordered than the structure seen previously in Figure 4.7a. We note that some of the additional disorder may be due to the `Planckton` P3HT job having run ~ 200 ns (although LJ-potential energy evolution is steady), whereas the structure in Figure 4.7a was run longer (1.5 μs).

We are currently running the P3HT simulation longer to test is this will increase the order in the `Planckton` simulation. It would not be surprising if the `Planckton` based simulations require longer to reach equilibrium as they do not yet take advantage of accelerating rigid-bodies since we have not yet developed a way to automatically identify which bodies should be rigid.

To quantify structure, we plot $g(r)$ for the thiophene ring geometric centers, Figure 8.4b, which shows a max peak at 3.86 Å and a minimum at 6.66 Å. This is similar to the $g(r)$ in Figure 4.7e in that there is a maximum at ~ 4.0 Å and a minimum at 6.6 Å. However, the peak in Figure 8.4b is not split like it was in Figure 4.7e due to the decreased amount of order in Figure 8.4a. The presence of π-stacking and weak aliphatic stacking can be seen in the GIXS pattern of the morphology (averaged over 13 rotated GIXS patterns): the peak at ~ 1.6 Å^{-1} along q_{xy} corresponds to π-stacking and the peak at ~ 0.4 Å^{-1} along q_z corresponds to alkyl stacking. Both these wavenumbers agree with the distances that were observed in our previous study[17] and experiment[12]. However, the distinct, repeat spots along q_z are not seen like they were in the UA system, which corroborates the visual interpretation that there are not strong, repeating lamellae.
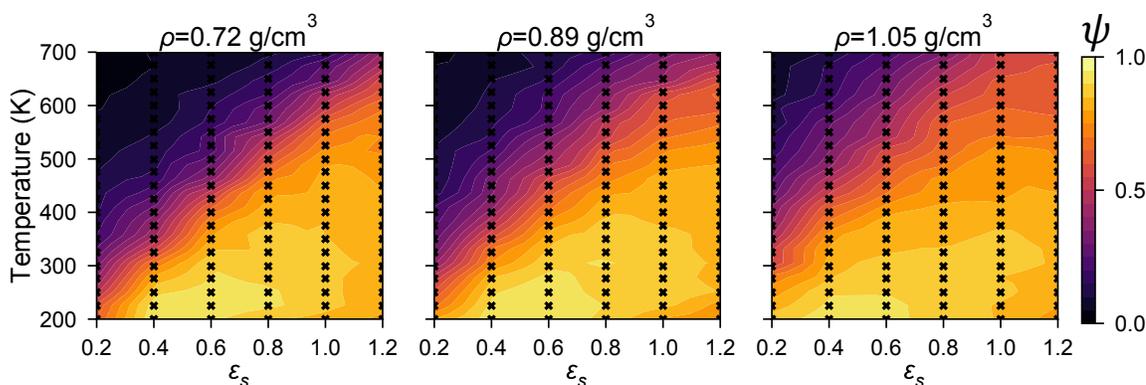
Figure 8.5: The order as a functions of ρ, T and ε_s.

Similarly to Section 4.4.2, we plot the self-assembly (represented by ψ) as a function of T, ε_s and ρ in Figure 8.5. In Figure 8.5 we see that the most ordering occurs at lower temperatures (< 300 K) and moderately low ε_s ($0.4 \leq \varepsilon_s \leq 0.6$). Unsurprisingly, these systems also undergo kinetic arresting with increasing ρ, however, this kinetic arresting is not as severe as it was in the UA case as there is more room with the c3 carbons being 85% of the size of the CT elements. In Figure 8.5 we also do not see the strong band that was observed in Figure 4.3. For example, at low ε_s we do not see very ordered structures. This is likely due to the interactions in the `Planckton` model being weaker than those of the UA model and when the interactions are reduced further by a small ε_s, results in the kinetic energy over-powering the potential energy wells. This weaker interaction can also help explain why the `Planckton` model disorders at a lower temperature and is less likely to become kinetically arrested at low temperatures. Specifically, at low temperatures (T ~ 200 K), kinetic arresting occurs less frequently because the potential wells are shallower and can be escaped. Regardless, the structures predicted with the Foyer determined interactions still represent the P3HT structures that are seen in experiment indicating that `Planckton` can accurately predict structures.

8.3.2   Acceptor Structure

We turn now to investigating the structures and charge transport of ITIC type molecules. These systems are simulated over a range of temperatures 250 to 425 K in steps of 25 K. All systems are at a density of 1.0 g/cm^3 and have $\varepsilon_s = 0.4$ were run for 1×10^8 timesteps (~ 200 ns). Each system contained 400 molecules, and as such, simulations had cubic dimensions of ~ 7 nm and $\sim 10{,}000$ simulation elements. These systems each ran for ~ 200 ns, which was sufficient for the potential energy to

become constant with time.
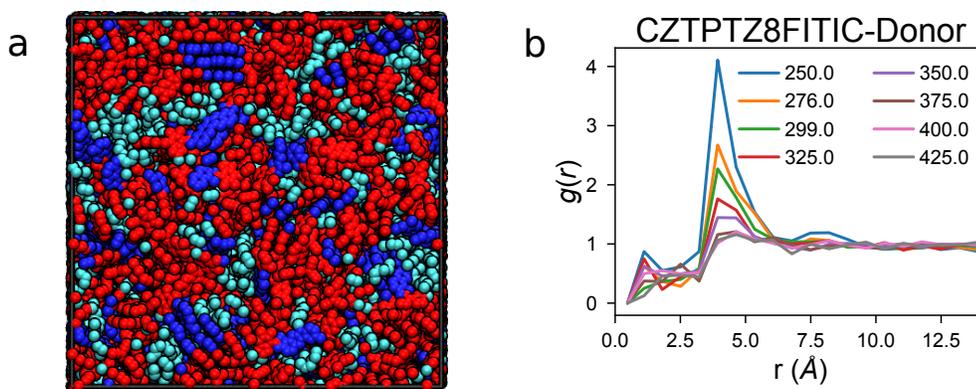
Figure 8.6: (a) The morphology of the CZTPTZ8FITIC system with the donor cores colored blue, the acceptor moieties colored red and the aliphatic chains colored cyan. Within the system small stacks of π-stacked cores can be observed. (b) The presence of this π-stacking can be seen in the donor core average position $g(r)$ as a strong first peak around 4 Å.

Visually, and from the $g(r)$, these systems appear to be primarily disordered at the state points simulated. The exception to this are the CZTPTZITIC and CZTPTZ8FITIC systems, which undergo at least short range π-stacking of the donor cores at lower temperatures (see Figure 8.6). We note that the non-zero values at short distances in Figure 8.6 arise from the average position of the donor core not being exactly within the plane of the donor region. These are able to order due to less bulky core functionalizations. For example, CZTPTZITIC and CZTPTZ8FITIC contain a pyrrole in their core (see Figure 8.2). The nitrogen in this pyrrole is bonded to a branching aliphatic chain which extends from the core. Contrarily, in ITIC two (sp^3) carbon atoms are bonded to benzene rings. These two benzene rings likely create steric hindrance that then prevents other ITIC molecules from π-stacking with this molecule, despite these functional groups primarily present to tune solubility[27].

With this steric hindrance in mind, it is not surprising that the ITIC systems do
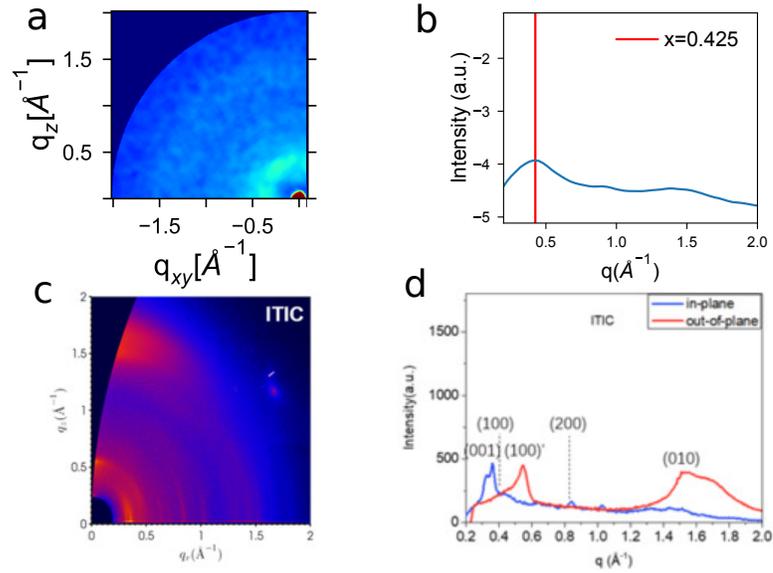
Figure 8.7: (a) Simulated GIXS pattern for ITIC (averaged over 60 views) shows a strong peak at ~ 0.4 Å^{-1} and a faint peak at ~ 1.5 Å^{-1}. These features can also been seen in the structure factor (a radial average of the GIXS pattern) in (b). (c) The experimental GIXS pattern for ITIC shows features at ~ 0.5 Å^{-1} and ~ 1.6 Å^{-1}. (d) The s_q for the pattern in (c) highlights the asymmetry present in (c) with the in-plane s_q having a peak closer to 0.3 Å^{-1} and the out-of-plane s_q having features at 0.5 Å^{-1} 1.5 Å^{-1}. Figures (c) and (d) reproduced with permission from Ref [27].

not order much; however, experimental ITIC devices have been shown to order into periodic structures. We compare GIXS of the ITIC system (250 K) to experiments in Figure 8.7. Our simulated GIXS in Figure 8.7a primarily shows a ring at ~ 0.4 Å^{-1} and a faint ring at ~ 1.5 Å^{-1}. These features can be seen on the structure factor (s_q) plot (Figure 8.7), which is a radial average of the GIXS pattern. The experimental GIXS pattern for ITIC is shown in Figure 8.7c, which shows significantly stronger periodic features than are seen in Figure 8.7a. The experimental features are different between the in- and out-of-plane axes, and the experimental s_q shows that the in-plane orientation as a feature at 0.3 Å^{-1}, where as the out-of-plane axis shows one at $\sim$ 0.5 and 1.55 Å^{-1}. As such, it is likely that we are observing similar features to those

seen in experiment, and our GIXS pattern is acting as a spherical average of these features e.g. the 0.3 and 0.5 Å^{-1} peaks observed in Figure 8.7d are averaged into the single peak at 0.4 Å^{-1} in Figure 8.7b. Generating more comparable morphologies can be accomplished by simulating systems for longer to see if the system relaxes to a more ordered state or simulating larger systems to reduce the noise that occurs with smaller systems. Additionally, as was seen in Chapters 3 and 4, thermal annealing and "evaporation" of the system may result in more ordering. Increasing the order of these systems will be the subject of future work.

8.3.3   Charge Transport

We also conduct KMC simulations for the ITIC-like systems. These KMC simulations are done with the MorphCT package as is explained in Section 2.2.2. Here, we treat each molecule as an individual chromophore so that we utilize values presented in literature for charge transport calculations. Specifically, the reorganization energy λ for all ITIC-like systems is set to 0.155 eV and is calculated for ITIC.[22] Each system has 10,000 charges allowed to run for 1×10^{-10} to 1×10^{-7} seconds.

We present the average $\mu_{E,0}$ values for these systems at different temperatures in Figure 8.8. The results shown in Figure 8.8 are encouraging as most mobility values are within an order-of-magnitude of experimental measurements. The predicted mobilities are all higher than the experimental ones - likely due to the periodic boundaries[18]. Despite this overprediction, the results presented in Figure 8.8 show that using `Planckton` to automatically identify interactions produces structures with similar charge transport properties to experiments. Future work will address how or if these molecules should be segmented based on the different electro-active portions of the molecule, i.e. we must explore our assumption that each molecule can be
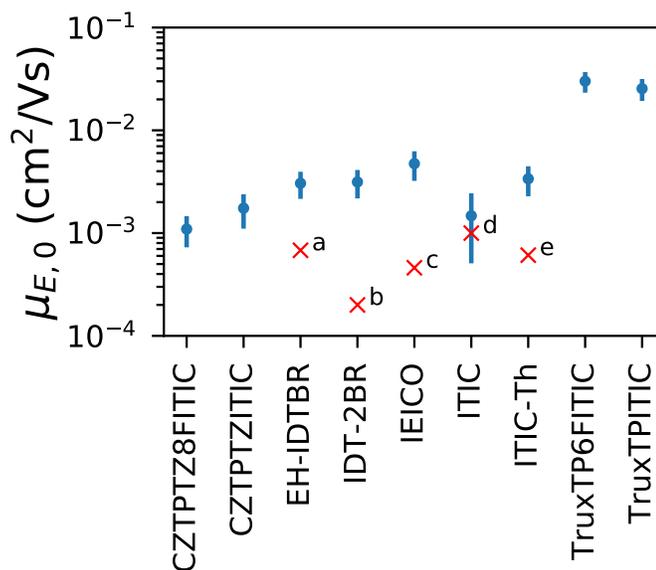
Figure 8.8: The $\mu_{E,0}$ for the new acceptor molecules show roughly order-of-magnitude agreement with experimental measurements. The red "x"s indicate literature values: a-Ref [8], b-Ref [13], c-Ref [28], d-Ref [30] and e-Ref [15].

represented as a single chromophore rather than treating the acceptor and donor portions separately.

8.4 Conclusions

In conclusion, we have presented on a new *TRUE* simulation tool `Planckton` which combines peer-reviewed, maintained tools from the simulation community to allow for many chemistry-processing combinations to be explored. We have verified the structures produced by `Planckton` by simulating structures of P3HT and comparing these structures to previous simulations and experimental GIXS. The P3HT structures produced by `Planckton` show the same periodic features at the same length-scales as experiment and previous simulations. The phase behavior for the `Planckton` simulations is also similar to our previous study, however, the `Planckton` based

simulations are less likely to become kinetically trapped due to softer particle-particle interactions and smaller simulation elements.

We also simulate a new class of electron acceptor-donor compound based on ITIC. Of the candidate compounds only CZTPTZITIC and CZTPTZ8FITIC showed a propensity for forming π-stacking of the electron donating cores. Despite experimental ITIC showing more order than was observed here, we show that the weak periodic features observed here are the same as those seen in experiment. Future work will focus on finding the correct simulation protocol needed to reproduce the structural orders that are seen in experiment. Despite, not having "exact" structural predictions, charge transport calculations show order-of-magnitude agreement with experimental measurements. Additionally, charge transport predictions can likely be improved by better understanding how molecules should be segmented into distinct chromophores based on their electronic properties. Regardless, we have shown how `Planckton` can automatically identify interactions that are able to achieve the correct periodic structures and charge transport properties.

BIBLIOGRAPHY

[1] Daylight Theory: SMARTS - A Language for Describing Molecular Patterns http://www.daylight.com/dayhtml/doc/theory/theory.smarts.html.

[2] Carl S. Adorf, Paul M. Dodd, and Sharon C. Glotzer. Signac - A Simple Data Management Framework. *arXiv*, pages 1–14, 2016.

[3] Carl Simon Adorf, Vyas Ramasubramani, Bradley D. Dice, Michael M. Henry, Paul M. Dodd, and Sharon C. Glotzer. glotzerlab/signac, 2019.

[4] Joshua A. Anderson, Chris D. Lorenz, and A. Travesset. General Purpose Molecular Dynamics Simulations Fully Implemented on Graphics Processing Units. *Journal of Computational Physics*, 227(10):5342–5359, 2008.

[5] Liang Gao, Zhi-Guo Zhang, Haijun Bin, Lingwei Xue, Yankang Yang, Cheng Wang, Feng Liu, Thomas P. Russell, and Yongfang Li. High-Efficiency Non-fullerene Polymer Solar Cells with Medium Bandgap Polymer Donor and Narrow Bandgap Organic Semiconductor Acceptor. *Advanced Materials*, 28(37):8288–8295, 2016.

[6] Jens Glaser, Trung Dac Nguyen, Joshua A. Anderson, Pak Lui, Filippo Spiga, Jaime A. Millan, David C. Morse, and Sharon C. Glotzer. Strong Scaling of General-Purpose Molecular Dynamics Simulations on GPUs. *Computer Physics Communications*, 192:97–107, 2014.

[7] Marcus D. Hanwell, Donald E. Curtis, David C. Lonie, Tim Vandermeerschd, Eva Zurek, and Geoffrey R. Hutchison. Avogadro: An Advanced Semantic Chemical Editor, Visualization, and Analysis Platform. *Journal of Cheminformatics*, 4(17):1–17, 2012.

[8] Sarah Holliday, Raja Shahid Ashraf, Andrew Wadsworth, Derya Baran, Syeda Amber Yousaf, Christian B. Nielsen, Ching Hong Tan, Stoichko D. Dimitrov, Zhengrong Shang, Nicola Gasparini, Maha Alamoudi, Frédéric Laquai, Christoph J. Brabec, Alberto Salleo, James R. Durrant, and Iain McCulloch. High-Efficiency and Air-Stable P3HT-Based Polymer Solar Cells with a New Non-Fullerene Acceptor. *Nature Communications*, 7:1–11, 2016.

[9] William G. Hoover. Canonical Dynamics: Equilibrium Phase-Space Distributions. *Physical Review A*, 31(3):1695–1697, 1985.

[10] Christoph Klein, János Sallai, Trevor J. Jones, Christopher R. Iacovella, Clare McCabe, and Peter T. Cummings. A Hierarchical, Component Based Approach to Screening Properties of Soft Matter. *Foundations of Molecular Modeling and Simulation*, 2016.

[11] Christoph Klein, Andrew Z. Summers, Matthew W. Thompson, Justin Gilmer, Clare McCabe, Peter T. Cummings, Janos Sallai, and Christopher R. Iacovella. Formalizing Atom-typing and the Dissemination of Force Fields with Foyer. *arXiv*, pages 1–38, 2018.

[12] Sangwon Ko, Eric T. Hoke, Laxman Pandey, Sanghyun Hong, Rajib Mondal, Chad Risko, Yuanping Yi, Rodrigo Noriega, Michael D. McGehee, Jean-Luc Brédas, Alberto Salleo, and Zhenan Bao. Controlled Conjugated Backbone Twisting for an Increased Open-Circuit Voltage while Having a High Short-Circuit Current in Poly(hexylthiophene) Derivatives. *Journal of the American Chemical Society*, 134(11):5222–5232, 2012.

[13] Qing Ya Li, Jingyang Xiao, Lu Ming Tang, Hua Chun Wang, Ziming Chen, Zhiyong Yang, Hin Lap Yip, and Yun Xiang Xu. Thermally Stable High Performance Non-Fullerene Polymer Solar Cells with Low Energy Loss by Using Ladder-Type Small Molecule Acceptors. *Organic Electronics: Physics, Materials, Applications*, 44:217–224, 2017.

[14] Yuze Lin, Jiayu Wang, Zhi Guo Zhang, Huitao Bai, Yongfang Li, Daoben Zhu, and Xiaowei Zhan. An Electron Acceptor Challenging Fullerenes for Efficient Polymer Solar Cells. *Advanced Materials*, 27(7):1170–1174, 2015.

[15] Yuze Lin, Fuwen Zhao, Qiao He, Lijun Huo, Yang Wu, Timothy C. Parker, Wei Ma, Yanming Sun, Chunru Wang, Daoben Zhu, Alan J. Heeger, Seth R. Marder, and Xiaowei Zhan. High-Performance Electron Acceptor with Thienyl Side Chains for Organic Photovoltaics. *Journal of the American Chemical Society*, 138(14):4955–4961, 2016.

[16] Lingxian Meng, Yamin Zhang, Xiangjian Wan, Chenxi Li, Xin Zhang, Yanbo Wang, Xin Ke, Zuo Xiao, Liming Ding, Ruoxi Xia, Hin-Lap Yip, Yong Cao,

and Yongsheng Chen. Organic and Solution-Processed Tandem Solar Cells With 17.3% Efficiency. *Science*, 2612:1094–1098, 2018.

[17] Evan Miller, Matthew Jones, Michael Henry, Paul Chery, Kyle Miller, and Eric Jankowski. Optimization and Validation of Efficient Models for Predicting Polythiophene Self-Assembly. *Polymers*, 10(12):1305, 2018.

[18] Evan Miller, Matthew Jones, and Eric Jankowski. Tying Together Multiscale Calculations for Charge Transport in P3HT: Structural Descriptors, Morphology, and Tie-Chains. *Polymers*, 10(12):1358, 2018.

[19] Evan D. Miller, Michael M. Henry, Matthew L. Jones, and Eric Jankowski. Planckton, 2019.

[20] Evan D. Miller, Matthew L. Jones, and Eric Jankowski. Enhanced Computational Sampling of Perylene and Perylothiophene Packing with Rigid-Body Models. *ACS Omega*, 2(1):353–362, 2017.

[21] Robert S. Gurney, Wei Li, Yu Yan, Dan Liu, Andrew J. Pearson, and Tao Wang. Morphology and Efficiency Enhancements of PTB7-Th:ITIC Nonfullerene Organic Solar Cells Processed via Solvent Vapor Annealing. *Journal of Energy Chemistry*, 37:148–156, 2019.

[22] Steven M. Swick, Weigang Zhu, Micaela Matta, Thomas J. Aldrich, Alexandra Harbuzaru, J. Teodomiro Lopez Navarrete, Rocio Ponce Ortiz, Kevin L. Kohlstedt, George C. Schatz, Antonio Facchetti, Ferdinand S. Melkonyan, and Tobin J. Marks. Closely Packed, Low Reorganization Energy π-extended Postfullerene Acceptors for Efficient Polymer Solar Cells. *Proceedings of the National Academy of Sciences*, 115(36):E8341–E8348, 2018.

[23] W. C. Swope, H. C. Andersen, P. H. Berens, and K. R. Wilson. A Computer Simulation Method for the Calculation of Equilibrium Constants for the Formation of Physical Clusters of Molecules: Application to Small Water Clusters. *Journal of Chemical Physics*, 76(1):637–649, 1982.

[24] Andrew Wadsworth, Maximilian Moser, Adam Marks, Mark S. Little, Nicola Gasparini, Christoph J. Brabec, Derya Baran, and Iain McCulloch. Critical Review of the Molecular Design Progress in Non-Fullerene Electron Acceptors

Towards Commercially Viable Organic Solar Cells. *Chemical Society Reviews*, 48(6):1596–1625, 2019.

[25] Junmei Wang, Wei Wang, Peter A Kollman, and David A. Case. Antechamber, An Accessory Software Package For Molecular Mechanical Calculations. *Molecules*, 222(2):U403–U403, 2001.

[26] Junmei Wang, Romain M. Wolf, James W. Caldwell, Peter A. Kollman, and David A. Case. Development and Testing of a General Amber Force Field. *Journal of Computational Chemistry*, 25(9):1157–1174, 2004.

[27] Y. Xiao and X. Lu. Morphology of Organic Photovoltaic Non-Fullerene Acceptors Investigated by Grazing Incidence X-ray Scattering Techniques. *Materials Today Nano*, 5:100030, 2019.

[28] Huifeng Yao, Yu Chen, Yunpeng Qin, Runnan Yu, Yong Cui, Bei Yang, Sunsun Li, Kai Zhang, and Jianhui Hou. Design and Synthesis of a Low Bandgap Small Molecule Acceptor for Efficient Polymer Solar Cells. *Advanced Materials*, 28(37):8283–8287, 2016.

[29] Jun Yuan, Yunqiang Zhang, Liuyang Zhou, Guichuan Zhang, Hin-Lap Yip, Tsz-Ki Lau, Xinhui Lu, Can Zhu, Hongjian Peng, Paul A. Johnson, Mario Leclerc, Yong Cao, Jacek Ulanski, Yongfang Li, and Yingping Zou. Single-Junction Organic Solar Cell with over 15% Efficiency Using Fused-Ring Acceptor with Electron-Deficient Core. *Joule*, pages 1–12, 2019.

[30] Zhong Zheng, Omar M. Awartani, Bhoj Gautam, Delong Liu, Yunpeng Qin, Wanning Li, Alexander Bataller, Kenan Gundogdu, Harald Ade, and Jianhui Hou. Efficient Charge Transfer and Fine-Tuned Energy Level Alignment in a THF-Processed Fullerene-Free Organic Solar Cell with 11.3% Efficiency. *Advanced Materials*, 29(5):3–8, 2017.

CHAPTER 9

CONCLUSIONS AND SUGGESTIONS FOR FUTURE WORK

9.1 Conclusions

Creating low-cost, mass-produced organic solar cells requires understanding how active-layer structure affects OPV performance and how processing dictates the self-assembled active layer morphology. Addressing these question with computer simulations requires the ability to investigate relationships between compounds, processing, structure and properties at length- and time-scales relevant to OPV performance.

This dissertation demonstrates that MD simulations with united-atom models and accelerating assumptions are computationally efficient enough to predict self-assembly under many conditions pertinent to OPV manufacturing. Utilizing an united-atom model with accelerating assumptions, these morphologies are large enough to be directly validated against experimental structures, but still maintain sufficient resolution to directly link the self-assembled structure to charge transport properties. In sum, we show that the efficient computer simulations conducted here are able to inform us of compound-processing-structure-property relationships vital to mass-producing efficient solar cells.

With this ability in mind, we demonstrate the most broad and experimentally-

validated computational study of self-assembly in the active layer structure and corresponding charge transport properties. We find that continuous pathways of overlapping π-orbitals are vital for fast charge transport. Further, we build the foundation for future OPV studies by developing *TRUE* tools and alleviating computational bottlenecks so that many chemistry-processing-property combinations can be investigated and relationships important for OPV mass-production can be identified.

9.1.1 Chapter-Based Conclusions

In Chapter 3 we evaluate the self-assembly and computational performance of small molecule perylene and perylothiophene in an united-atom model with flexible and rigid bonds. We quantify the self-assembly by creating phase diagrams in which five phases exist: gas, liquid, fluid, ordered and eclipsed. We find that rigid models are more computational efficient and have close agreement in phase transitions as the flexible models—excluding very dense systems where confinement effects are more pronounced. We predict that the self-assembled structure for perylene and perylothiophene is ordered discotic columns, which is similar to experimentally observed flattened herringbone structure observed experimentally for these compounds. However, we note that our model does not exactly capture these herringbone motifs - possibly due to the exclusion of electrostatic interactions.

In Chapter 4 we continue investigating the effects of accelerating assumptions with the polymer poly-(3-hexylthiophene) (P3HT). We find that experimental thin-film sizes are untenable with united-atom models and are not going to be accessible for many years. As such, we suggest applying further accelerating assumptions, such as coarser models, to achieve larger simulation volumes. With this in mind, we suggest guidelines for simulating OPV compounds: (1) benchmark performance to determine

the practical simulation size, (2) generate coarse phase diagrams to identify interesting structures, (3) use simulated solvent evaporation to obtain structures at experimental densities and (4) validate the structures against experiment. Following this "recipe" we generate phase maps of 15-mer P3HT self-assembly using our order metric ψ as a function of density, temperature and solvent "goodness". We then simulate "large" 1,000 15-mer systems with solvent evaporation to achieve morphologies that can be more clearly compared to experiment. In doing this, we find that our short chain, implicit charge model achieves structures nearly identical to those observed in experiment.

In Chapter 5 we begin our transition from purely structural predictions and quantifying model assumptions to predicting the self-assembly of perylene derivative dibenzo-tetraphenyl-periflanthene (DBP):fullerene mixtures then using KMC to calculate charge transport through these structures. We verify that both the neat DBP and fullerene systems show structures that are similar to those seen in experiment. We find that the highest degrees of order (as measured by ψ) are observed at low temperatures and stronger solvents. We also find that the movement of electrons is dominated by energetic effects of the fullerenes as the rigid fullerenes are insensitive to morphology preparation temperature whereas functionalized fullerenes decrease with increasing temperature. Conversely, in DBP the movement of holes is not dominated by energetic factors, but rather the continuity of DBP stacks, more so than even the "order" of the simulation. The dependence on stack-continuity highlights a pit-fall in conducting charge transport simulations with periodic simulation volumes: it is possible to obtain unphysical, infinitely long stacks due to columns spanning the periodic volume. As such, researchers will need to be aware of this possible issue, and account for it in simulations.

In Chapter 6 we return to the P3HT system and conduct charge transport calculations for the structures presented in Chapter 4. For smaller systems with 100 15-mers, we find that our order parameter ψ can be improved by considering the deviations in bond-length to capture disorder missed by the binary clustering method. We find that systems with small highly-ordered and amorphous domains with short chains can result in strong charge trapping—greatly reducing the charge mobility. This charge trapping can be alleviated by the presence of longer P3HT chains such as those present in polydisperse simulations. These longer chains act as tie-chains—allowing for a charge-highway between crystallites in an amorphous matrix.

In Chapter 7 we implement machine learning, specifically a random forest algorithm, to replace the redundant ZINDO/S calculations used in determining electronic couplings. The random forest is able to make electronic coupling predictions based on tabulated bonded, spatial, rotational and energetic features $\sim 400\times$ faster than repeating the ZINDO/S calculations for every system. These random forest predictions result in a maximum 5% error for future mobility calculations compared to the explicit ZINDO/S coupling calculations. In developing the machine learning training set, we find that extracting chromophore pairs from MD simulations acts as an importance sampling of physical, close configurations compared to naively moving two chromophores relative to each other. Additionally, as few as 1×10^5 pairs are sufficient to accurately train the machine learning algorithm when these examples are taken from a representative trajectory.

Lastly, in Chapter 8 we work towards more *TRUE* simulations by developing `Planckton` which combines peer-reviewed tools maintained by the simulation community to quickly assign interaction parameters, initialize systems, then manage and distribute those systems on high-performance clusters. One major benefit of

`Planckton` over the tools used in our previous studies is the ability to easily extend simulations to new chemistries. We validate this tool by simulating morphologies of P3HT over a variety of densities, temperatures, and solvent strengths. We find that the structures and phase behavior are similar to our united-atom model and experiment. However, we also note some differences in that the P3HT behavior predicted by `Planckton` is less likely to become kinetically trapped due to softer potentials and smaller simulation elements. We also demonstrate how arbitrary chemistries can be easily simulated by automatically typing and simulated a new class of electron acceptor compound based on ITIC. The periodic features for these ITIC systems are similar to those observed in experiment. However, we find that these simulations do not achieve structures with as much order as those seen in experiment. Despite the decreased amount of order, the electron mobilities through these morphologies are close those measured in experiment.

9.2   Suggestions for Future Work

I believe the most pressing future work is to achieve simulation sizes of ~ 100 nm, while not losing chemical specificity. Until 100 nm length-scales are achieved, we can only obtain a partial picture of structure, and consequently, an incomplete picture of the structure-charge transport relationship. Our picture of the structure-charge transport relationship is incomplete for at least three reasons: (1) There is less charge trapping in smaller systems (e.g. the semi-crystalline structures in Section 6), (2) exciton diffusion lengths are of the order ~ 10 nm[2] and our simulation sizes are too small to know if the simulation is macro- or micro-phase separated and (3) the energy variation is unrealistically low in small, periodically bound simulation

volumes[1]. Increasing system sizes to 100 nm will require applying more coarse-graining strategies that will reduce the number of particles in the simulation while maintaining chemical specificity. This could be achieved by coarse-graining with Multi-State Iterative Boltzmann Inversion[4], possibly implemented into `Planckon` with the VOTCA package[5] and then "fine-graining" as was done by Jones[3].

After we are able to achieve 100 nm simulation sizes for arbitrary chemistries, we would need to expand our charge transport pipeline to include more quantum chemical calculations. For example, we currently utilize reorganization values presented in literature to determine hopping rates in the KMC simulations. Needing literature values limits the speed and compounds we can utilize in our simulations. As such, expanding our KMC pipeline to calculate values such as reorganization energies will enable wide-spread studies.

BIBLIOGRAPHY

[1] Denis Andrienko. Multiscale Concepts in Simulations of Organic Semiconductors. *Handbook of Materials Modeling*, pages 1–12, 2018.

[2] Energy Environ, Oleksandr V. Mikhnenko, Hamed Azimi, Markus Scharber, Mauro Morana, W. M. Blom, and Maria Antonietta. Environmental Science Exciton Diffusion Length in Narrow Bandgap Polymers. *Energy and Environmental Science*, 5:6960–6965, 2012.

[3] Matthew Lewis Jones and Eric Jankowski. Computationally Connecting Organic Photovoltaic Performance to Atomistic Arrangements and Bulk Morphology. *Molecular Simulation*, 43(10-11):1–18, 2017.

[4] Timothy C. Moore, Christopher R. Iacovella, and Clare McCabe. Derivation of Coarse-Grained Potentials via Multistate Iterative Boltzmann Inversion. *The Journal of Chemical Physics*, 140(22):224104, 2014.

[5] Victor Rühle, Christoph Junghans, Alexander Lukyanov, Kurt Kremer, and Denis Andrienko. Versatile Object-Oriented Toolkit for Coarse-Graining Applications. *Journal of Chemical Theory and Computation*, 5(12):3211–3223, 2009.

APPENDIX A

RESEARCH AND TOOL CONTRIBUTIONS IN THIS WORK

A special thanks to those who helped in producing this dissertation. Here, we use this appendix to explicitly state the work contributed by other researchers that make this dissertation possible. Contributions are separated into sections by the Chapter in which the contribution is present. We note that these contributions are not exhaustive as there has been extensive amounts of help from many in debugging code, writing reports and thinking through problems.

A.1   Chapter 3

In Chapter 3 we conduct MD simulations. The MD code to run simulations: HOOMD-Blue including integrators and thermostats was written by the Glotzer group at the University of Michigan[3,4]. The MD wrapper used to run OPV MD simulations: `opv_cg` was written by Dr. Eric Jankowski. In analyzing the MD simulations, the $g(r)$ code was implemented in the MDTraj package[10], the diffraction and autocorrelation code was written by Dr. Eric Jankowski[7,11], and the clustering code was written in-house.

A.2 Chapter 4

In the P3HT model chapter, we primarily use the same tools as Chapter 4, however, a special note for Paul Chery and Kyle Miller who implemented a filter to remove bonded thiophenes from the $g(r)$ analysis and Mike Henry who analyzed the performance on the P3HT simulations on various clusters with various GPUs.

A.3 Chapter 5

In Chapter 5 we begin to also do electronic property calculations with KMC simulations. To do this we use the MorphCT package[6], which is primarily developed by Dr. Matthew Jones, and all personal contributions have been on the data analysis end of the package rather than the charge transport predictions.

A.4 Chapter 6

In Chapter 6 we conduct the KMC simulations using the MorphCT package. In this work, Dr. Matthew Jones was primarily responsible for analyzing the charge transport characteristics in the 15mer, 1000 molecule simulations. As part of this, he also explored various clustering metrics based on $J_{i,j}$ and hopping.

A.5 Chapter 7

In Chapter 7 we conduct machine learning to predict $J_{i,j}$ values. Dr. Matthew Jones wrote the code that enables extracting the information from KMC simulations so that machine learning could be conducted. He was also instrumental in developing

the features used in the training process. Dr. Bryan Stanfill and Mike Henry were also instrumental in setting up the machine learning.

A.6   Chapter 8

In Chapter 8 we present `Planckton`. Mike Henry wrote the bulk of `Planckton`, including tieing together tools from other groups: Antechamber[12], Foyer[9], mBuild[8], and Signac[1,2]. Dr. Matthew Jones implemented dihedral interactions into `Planckton`. My personal efforts in `Planckton` have been in removing hydrogens and implementing shrinking steps into the simulation. In this chapter, we use the Freud analysis package[5] to calculate $g(r)$s rather than MDTraj as it functions better with the HOOMD-Blue output files. Lastly, the charge transport simulations were conducted by Dr. Matthew Jones.

BIBLIOGRAPHY

[1] Carl S. Adorf, Paul M. Dodd, and Sharon C. Glotzer. Signac - A Simple Data Management Framework. *arXiv*, pages 1–14, 2016.

[2] Carl Simon Adorf, Vyas Ramasubramani, Bradley D. Dice, Michael M. Henry, Paul M. Dodd, and Sharon C. Glotzer. glotzerlab/signac, 2019.

[3] Joshua A. Anderson, Chris D. Lorenz, and A. Travesset. General Purpose Molecular Dynamics Simulations Fully Implemented on Graphics Processing Units. *Journal of Computational Physics*, 227(10):5342–5359, 2008.

[4] Jens Glaser, Trung Dac Nguyen, Joshua A. Anderson, Pak Lui, Filippo Spiga, Jaime A. Millan, David C. Morse, and Sharon C. Glotzer. Strong Scaling of General-Purpose Molecular Dynamics Simulations on GPUs. *Computer Physics Communications*, 192:97–107, 2014.

[5] Eric S. Harper, Matthew Spellings, Joshua Anderson, and Sharon C. Glotzer. harperic/freud: Zenodo DOI release, 2016.

[6] M. L. Jones. MorphCT, 2018.

[7] Matthew Lewis Jones and Eric Jankowski. Computationally Connecting Organic Photovoltaic Performance to Atomistic Arrangements and Bulk Morphology. *Molecular Simulation*, 43(10-11):1–18, 2017.

[8] Christoph Klein, János Sallai, Trevor J. Jones, Christopher R. Iacovella, Clare McCabe, and Peter T. Cummings. A Hierarchical, Component Based Approach to Screening Properties of Soft Matter. *Foundations of Molecular Modeling and Simulation*, 2016.

[9] Christoph Klein, Andrew Z. Summers, Matthew W. Thompson, Justin Gilmer, Clare McCabe, Peter T. Cummings, Janos Sallai, and Christopher R. Iacovella. Formalizing Atom-typing and the Dissemination of Force Fields with Foyer. *arXiv*, pages 1–38, 2018.

[10] Robert T. McGibbon, Kyle A. Beauchamp, Matthew P. Harrigan, Christoph Klein, Jason M. Swails, Carlos X. Hernández, Christian R. Schwantes, Lee Ping

Wang, Thomas J. Lane, and Vijay S. Pande. MDTraj: A Modern Open Library for the Analysis of Molecular Dynamics Trajectories. *Biophysical Journal*, 109(8):1528–1532, 2015.

[11] Evan D. Miller, Michael M. Henry, and Eric Jankowski. Diffractometer, 2018.

[12] Junmei Wang, Wei Wang, Peter A Kollman, and David A. Case. Antechamber, An Accessory Software Package For Molecular Mechanical Calculations. *Molecules*, 222(2):U403–U403, 2001.

APPENDIX B

ENHANCED COMPUTATIONAL SAMPLING OF PERYLENE AND PERYLOTHIOPHENE PACKING WITH RIGID-BODY MODELS—SUPPORTING INFORMATION[1]

B.1   Determination of Equilibrium

The potential energy can be used to measure when a morphology has reached equilibrium by determining when its evolution becomes constant as a function of time $(\mathrm{d}E/\mathrm{d}t = 0)$. The calculation is accomplished by separating the total potential energy for each timestep into 10 equally sized bins. The average and standard deviation over each bin is calculated and compared to the average and standard deviation of the final bin. When the average potential energy of a bin is within the standard deviation of the final bin, the energy in no longer is changing and the system is considered to have reached equilibrium. In the example shown in Figure B.1, the red line indicates the critical point τ_r beyond which the average value of the potential energy is shown to be within one standard deviation of the potential energy of the final bin. τ_r therefore represents the timestep at which equilibrium has occured.
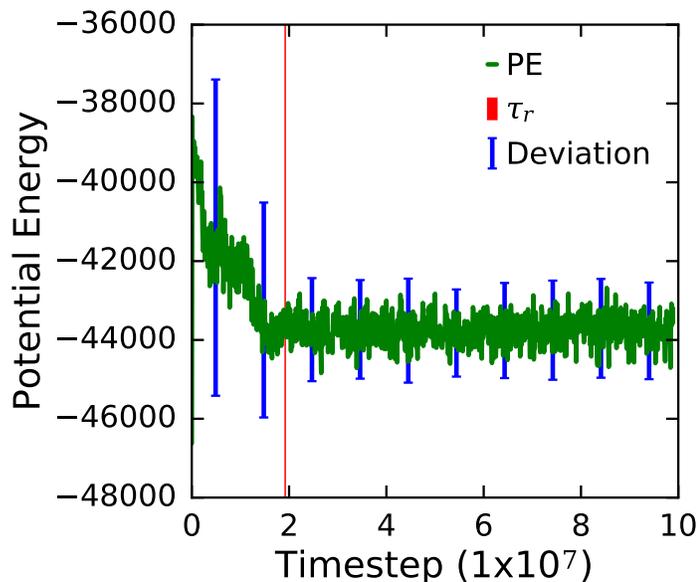
Figure B.1: Potential energy as a function of increasing timestep (green). The standard deviation for each bin is shown in blue. The red vertical line is the timestep at which the system is considered relaxed.

B.2 Molecular Dynamics Force Field and Computing Infrastructure

For this investigation, the Optimized Potential for Liquid Simulations (OPLS) Force Field is used[4]. The atomic masses are defined as $M_C = 12.01$ amu and $M_S = 32.06$ amu. The masses of the hydrogen atoms are not considered during the molecular dynamics simulations themselves, but are taken into account as $M_H = 1.00$ amu in the density calculations. The constants used in the force field for the two molecules are defined in Tables B.1-B.4. Note that, in the interest of computational efficiency, all carbon united atoms in the simulation are treated as C-H groups, and so the atom type 'C' describes the diatomic species.

The simulations leverage the Kestrel and Maverick supercomputers at Boise State

University (BSU) and the Texas Advanced Computing Center (TACC) respectively, using a single NVIDIA K40 graphics processing unit (GPU) per simulation at TACC and a single NVIDIA K20 GPU per simulation at BSU.

Non-bonded interaction parameters used in the MD simulations. Heterogeneous atom pair interactions $\sigma_{LJ_{i,j}} = \sqrt{\sigma_{LJ_i}\sigma_{LJ_j}}$ and $\varepsilon_{i,j} = \sqrt{\varepsilon_i\varepsilon_j}$. Pair interactions are defined by a Lennard-Jones potential (Equation 2.2).

Atom Type	σ_{LJ} (Å)	ε (kcal mol^{-1})
C	3.8	0.122
S	3.5	0.359

Table B.1: Non-bonded interaction parameters used in the perylene and perylothiophene simulations.

Bond-stretching parameters used in the MD simulations. Bonds are defined by a harmonic potential (see Equation 2.4). Note that bond equlibrium distances r_0 are given in terms of the carbon atom diameter σ_C.

Bond Type	r_0 (Å)	k_b (kcal mol^{-1} Å^{-2})
Perylene		
C-C	1.52	253.5
Perylothiophene		
C-C	1.52	745.8
C-S	1.71	745.8

Table B.2: Bond-stretching parameters used in the perylene and perylothiophene simulations.

Angle-bending parameters used in the simulations. Angles are defined by the harmonic potential (see Equation 2.5).

Torsional parameters used in the MD simulations. Torsions are defined by the OPLS form in Equation 2.8.

Angle Type	θ_0 (rad)	k_θ (kcal mol^{-1} rad^{-2})
Perylene		
C-C-C	2.09	46.36
Perylothiophene		
C-C-C	2.09	136.42
C-C-S	2.09	136.42
C-S-C	1.60	136.42

Table B.3: Angle-bending parameters used in the perylene and perylothiophene simulations.

Dihedral Type	k_{d_1} (kcal mol^{-1})	k_{d_2} (kcal mol^{-1})	k_{d_3} (kcal mol^{-1})	k_{d_4} (kcal mol^{-1})
Perylene				
C-C-C-C	0.00	6.10	0.00	0.00
Perylothiophene				
C-C-C-C	0.00	17.95	0.00	0.00
C-C-C-S	0.00	17.95	0.00	0.00
C-C-S-C	0.00	17.95	0.00	0.00
C-S-C-C	0.00	17.95	0.00	0.00

Table B.4: Torsional parameters used in the perylene and perylothiophene simulations.

B.3 Omission of Electrostatic Calculations

The goal of this paper is to highlight how high-throughput MD can be used to quickly generate phase behavior of planar aromatic systems and provide comparisons to experimental work. In it, we omit electrostatic interactions between partially charged atoms in order to reduce the number of required calculations and allow for higher-throughput in this work. As an example, preliminary simulations of perylene conducted with partial atomic charges and their electrostatic interactions showed an average of 273.44 time steps per second whereas simulations excluding the electrostatic interactions averaged 1434.69 time steps per second. Therefore, we are able to increase the speed of the calculations by 3-4× by omitting the electrostatic forces.

We see that the omission of charges likely leads to differences in packing angle

compared to what is typically seen in planar aromatic molecules[2]. For instance, the work of Tsuzuki et al. has shown, using quantum chemical calculations applied to dimers of thiophene-based molecules, that these electrostatic interactions perform a crucial role in favoring the perpendicular orientations observed in herringbone structures[5]. It could therefore be expected that the bulk structural behavior of perylene and perylothiophene would be similarly affected, leading to the herringbone structures observed in the α-polymorphs. We note, however, that the diffraction patterns for these systems obtained in this investigation are in good agreement with those obtained experimentally, and so, for a bulk material, a charge neutral model appears to describe the most important physics required to obtain realistic morphological structures at a wide variety of state points. It is therefore left to the reader to decide whether the computational benefit of omiting the electrostatic charge calculations will outweigh the small-scale differences in packing information for their own systems.

B.4 Unit Conversions

In the HOOMD-blue simulation suite[1,3] unitless temperatures are related to real temperatures by the equation:

$$T_{SI} = \frac{T_{unitless} * \varepsilon}{k_B} \tag{B.1}$$

where T_{SI} is the physical temperature in kelvin, $T_{unitless}$ is the unitless temperature, ε is the energy scale factor (0.122 and 0.359 kcal mol^{-1} for perylene and perylothiophene respectively), and k_B is the Boltzmann constant. The values obtained for T_{real} are presented below in Table B.5.

$T_{unitless}$	T_{SI}, Perylene (K)	T_{SI}, Perylothiophene (K)
1	61.25	180.16
2	122.49	360.33
3	183.74	540.49
4	244.98	720.65
5	306.23	900.81
6	367.47	1080.98
7	428.72	1261.14
8	489.96	1441.30
9	551.21	1621.46
10	612.45	1801.63
11	673.70	1981.79
12	734.94	2161.95
13	796.19	2342.11
14	857.43	2522.28
15	918.68	2702.44
16	979.92	2882.60
17	1041.17	3062.76
18	1102.41	3242.93
19	1163.66	3423.09
20	1224.90	3603.25
21	1286.15	3783.41
22	1347.39	3963.58
23	1408.64	4143.74
24	1469.89	4323.90
25	1531.13	4504.07
26	1592.38	4684.23
27	1653.62	4864.39
28	1714.87	5044.55
29	1776.11	5224.72
30	1837.36	5404.88

Table B.5: The simulation temperature with the corresponding actual temperature in Kelvin for Perylene and Perylothiophene.

Densities are calculated by:

$$\rho = \frac{N * M_w}{V},$$
(B.2)

where N is the number of molecules, M_w is the molecular weight, and V is the simulation volume. Densities values investigated are presented below in Table B.6.

Perylene (g/cm^3)	Perylothiophene (g/cm^3)
0.01	0.01
0.12	0.13
0.37	0.39
0.61	0.65
0.85	0.91
1.04	1.11
1.22	1.30
1.40	1.50
1.59	1.69
1.77	1.89

Table B.6: Densities in g/cm^3 at which the calculations were conducted.

B.5 Determination of ξ Cut Off

As perylene and perylothiophene are annealed, intracolumnar order increases due to intermolecular attractive forces surpassing thermal vibrations. This leads to the formation of an eclipsed phase in which the perylene/perylothiophene molecule covers the one behind it. To measure when this phase emerges, the ξ values are measured over the complete range of temperatures tested. It is observed that ξ converges to $\sim$ 0.95 as temperature is lowered, for all ρ (Figure B.2). The ξ cut-off value is taken as 0.90 to allow for fluctuations in ξ.
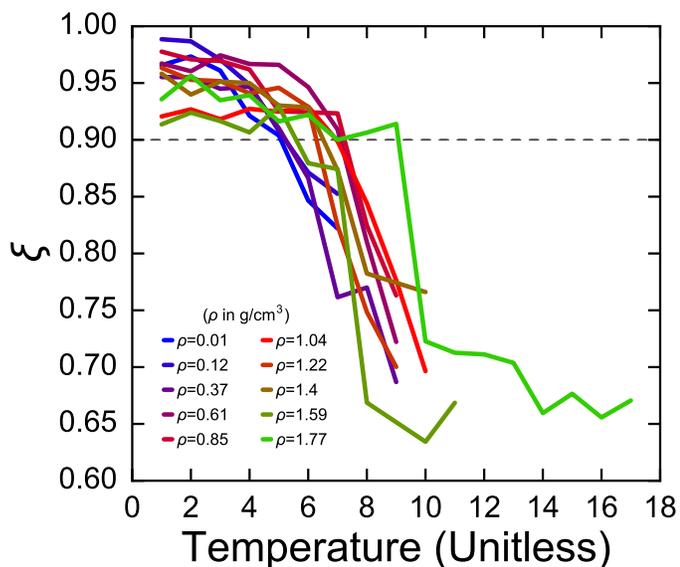
Figure B.2: As temperatures decrease, the degree of eclipsing (ξ) for all ρ measured converge to $\sim$ 0.95. The horizontal dashed black line is considered the cut off for a system to be considered in the eclipsed phase.

B.6   Distribution of θ

The cut off of the dot product of 0.96 is chosen from the distribution of the dot product values. The distribution of dot product values for the system shown in Figure 2 in the text is shown in Figure B.3b. The dot product reaches a maximum at 1 and if an angle is $\pm$ 16° the dot product will still be 0.96. Therefore, we assume the distribution of values is symmetric around one, and split our distribution across one accordingly (as shown in Figure B.3b). We then fit a Gaussian curve to the dot product distribution. We find that at $\sigma = 4.5$ that the dot product value is 0.96.
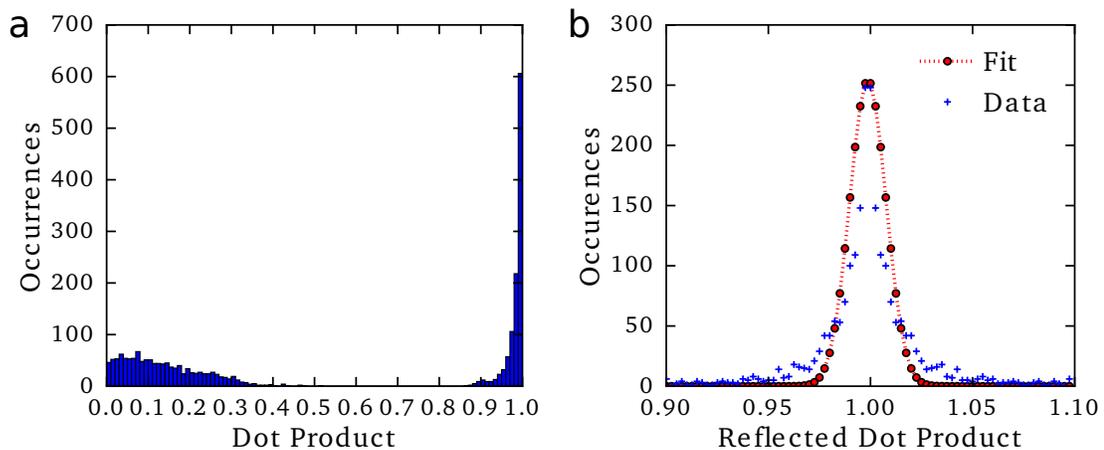
Figure B.3: (a) Histogram of dot product values in the ordered system shown in the text Figure 2. (b) Dot product values approximated to a Gaussian curve. We find that the dot product equals 0.96 at a σ of 4.5.

B.7   System Size Comparison

Simulations are run with 200 to 1500 molecules. Figure B.4a and b show that both systems relax to the energetically favored, hexagonally-packed columns. Both systems are also shown to be very ordered by visual inspection (Figure B.4a,b). Due to the structural similarities (including a near indistinguishable radial distribution function as shown in Figure B.4c), only the more computationally efficient simulations of 200 molecules are considered for further analysis in the main text.
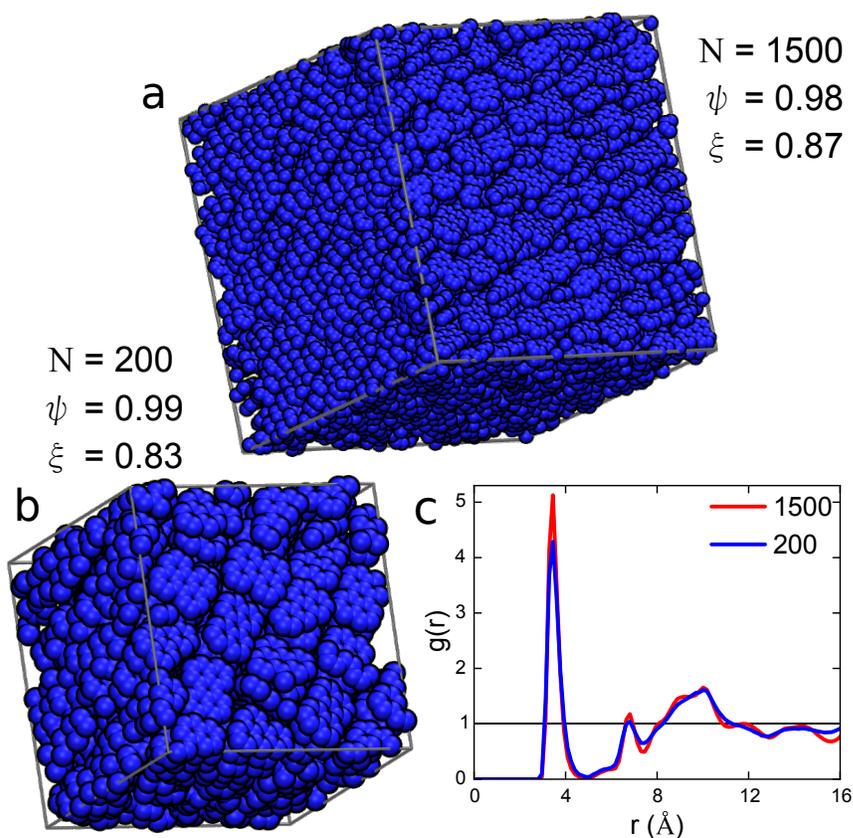
Figure B.4: Example morphologies and calculated ξ and ψ values for (a) N = 1500 and (b) N = 200 molecule systems. (c) The comparison of the radial distribution function.

B.8 Checkerboard-Aligned Energies
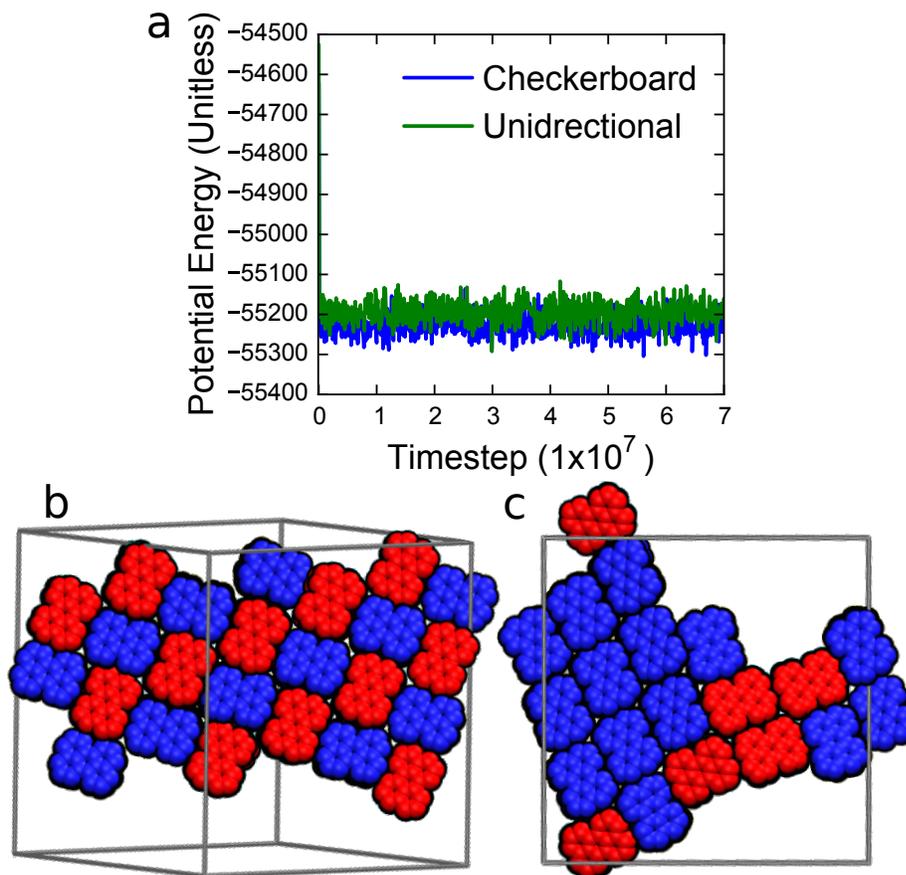
Figure B.5: (a) Potential energies of the two structures showing significant overlap of the potential energies in the (b) checkerboard structure and (c) aligned structure.

The potential energies of the checkerboard and aligned strucural motifs are found to be nearly equal (Figure B.5a). The checkerboard morphology (Figure B.5b) has a potential energy of -55226 ± 34 whereas the aligned system (Figure B.5c) has a potential energy of -55193 ± 37 (the plus/minus is the standard deviation). The morphologies shown in Figure B.5 are simulations of rigid perylene conducted at ρ = 1.22 g/cm^3. The runs were executed for 12 hours at $T = 10$ ($\sim$ 600 K), then the

temperature was lowered in ΔT increments of 1 ($\sim$ 60 K) to the final temperature: $T = 60$ K. The only difference between the two runs is the initial configuration.

BIBLIOGRAPHY

[1] Joshua A. Anderson, Chris D. Lorenz, and A. Travesset. General Purpose Molecular Dynamics Simulations Fully Implemented on Graphics Processing Units. *Journal of Computational Physics*, 227(10):5342–5359, 2008.

[2] G. R. Desiraju and A. Gavezzotti. Crystal Structures of Polynuclear Aromatic Hydrocarbons. Classification, Rationalization and Prediction from Molecular Structure. *Acta Crystallographica Section B*, 45(5):473–482, 1989.

[3] Jens Glaser, Trung Dac Nguyen, Joshua A. Anderson, Pak Lui, Filippo Spiga, Jaime A. Millan, David C. Morse, and Sharon C. Glotzer. Strong Scaling of General-Purpose Molecular Dynamics Simulations on GPUs. *Computer Physics Communications*, 192:97–107, 2014.

[4] William L. Jorgensen and Julian Tirado-Rives. The OPLS [Optimized Potentials for Liquid Simulations] Potential Functions for Proteins, Energy Minimizations for Crystals of Cyclic Peptides and Crambin. *Journal of the American Chemical Society*, 110(6):1657–1666, 1988.

[5] Seiji Tsuzuki, Kazumasa Honda, and Reiko Azumi. Model Chemistry Calculations of Thiophene Dimer Interactions: Origin of π-stacking. *Journal of the American Chemical Society*, 124(41):12200–12209, 2002.

APPENDIX C

OPTIMIZATION AND VALIDATION OF EFFICIENT MODELS FOR PREDICTING POLYTHIOPHENE SELF-ASSEMBLY - SUPPORTING INFORMATION[1]

C.1 Force-Field Constraints for Bonded Atoms

Here we present the intra-molecular interactions, e.g. two-body bond, three-body angle, and four-body dihedral constraints governing our P3HT model. These bonded interactions are adapted from the atomistic models developed by Huang and Bhatta[2,10]. As we employ united-atom models in this work, we compare our force-field values with the Optimized Potentials for Liquid Simulations - United Atom (OPLS-UA) force-field to ensure that we are not introducing spurious interactions when adapting the atomistic model[12].

Bond potentials are harmonic (see Equation 2.4). The k_b and l_0 values are presented in Table C.1. It should be noted that the CA-CA bond parameters presented in Table C.1 only apply to bonds between thiophene rings. The thiophene rings themselves are represented as rigid bodies, and so bonds within the rings do not change during the simulation. If a flexible model were to be used for the thiophene

---

ring, then distinct bead types would be needed to distinguish bonds between and within rings.

Table C.1: The bond constraints used in our force-field to simulate P3HT. [†]Bonds between two monomers, rather than within a single, rigid thiophene. [‡]Bonds completely enclosed by a rigid body, which are therefore fixed at l_0.

Bond	l_0 (Å)	k_b (kcal mol^{-1} Å^{-2})
CA-CA[†]	1.43	392.0
CT-CT	1.53	268.0
CA-CT	1.51	300.0
CA-S[‡]	1.71	291.0

Angles constraints are also harmonic (see Equation 2.5). The angle constraints are presented in Table C.2. As before, the CA-CA-CA and CA-CA-S angles specified describe angle constraints between monomers, rather than those in the rigid thiophene ring.

Table C.2: The angle constraints used in our force-field to simulate P3HT. [†]Angles between two monomers, rather than within a single, rigid thiophene. [‡]Angles completely enclosed by a rigid body, which are therefore fixed at θ_0.

Angle	θ (rad)	k_θ (kcal mol^{-1} rad^{-2})
CA-CA-CA[†]	2.27	54.7
CA-CA-S[†]	2.09	41.7
CA-S-CA[‡]	1.62	86.0
CA-CT-CT	2.16	70.0
CA-CA-CT	2.15	70.0
CT-CT-CT	1.97	37.5

The dihedral parameters are defined by a multi-harmonic function (see Equation 2.6). The dihedral constraints are presented in Table C.3.

Due to the rigid bodies used in this investigation, some dihedrals become ambiguously defined if the force-field from the literature are used. For instance, the CA-CA-CA-CT and the CA-CA-CA-CA dihedral constraints are already considered by the CA-CA-CT-CT and S-CA-CA-S dihedrals respectively. Furthermore, in the

Table C.3: The dihedral constraints used in our force-field to simulate P3HT. [†]: Dihedrals between two monomers, rather than within a single, rigid thiophene.

Dihedral	k_{d_0} (kcal mol^{-1})	k_{d_1}	k_{d_2}	k_{d_3}	k_{d_4}
S-CA-CA-S	2.9533	0.1571	-4.2326	0.39979	1.8855
CA-CA-CA-S[†]	2.9533	-0.1571	-4.2326	-0.39979	1.8855
CA-CA-CT-CT	0.3175	1.127	14.143	-22.297	6.7188
CA-CT-CT-CT	2.4469	-6.3946	10.747	30.695	11.139
CT-CT-CT-CT	1.8922	-3.4904	1.4665	7.1418	0.2859
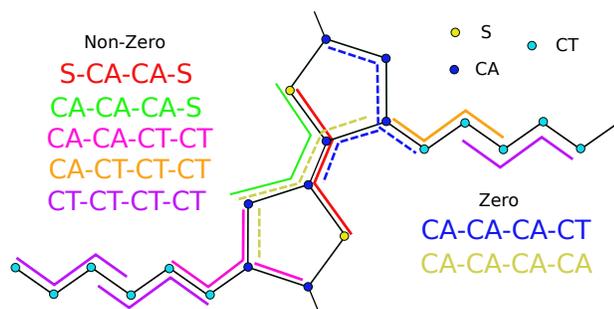
Figure C.1: The flexible dihedrals used in this investigation that are not completely enclosed by a rigid body. Reducing the number of distinct atom types in this model (compared to the all-atom models from which they are derived) leads to some conflicts in the dihedral constraints. To avoid this, the coefficients for some of these dihedrals (CA-CA-CA-CT and CA-CA-CA-CA) are set to zero.

case of CA-CA-CA-CT, two possible definitions are applicable depending on whether the CA is located in the same ring as the alkyl sidechain or not. When the CA is in a neighboring ring to the monomer containing the CT, the dihedral is in the cis conformation, however, when CA is in the same ring as the specified side chain, the dihedral is in the trans conformation. This ambiguity leads to instabilities in the system when the dihedral parameters are not set to 0. As such, we set the CA-CA-CA-CT and CA-CA-CA-CA dihedral coefficients to zero, effectively deactivating them in our force-field. Figure C.1 shows which dihedral constraints are considered (solid lines) and which are omitted (dashed lines), to highlight the redundancies that exist due to reduction to three atom types and the incorporation of the rigid body. We therefore

assume that the other carbon aromatic to chain (CA-CA-CT-CT and CA-CT-CT-CT) dihedrals are sufficient in describing the position of the chains relative to the rings and that the (S-CA-CA-S and CA-CA-CA-S) dihedrals are sufficient in describing the orientation of the thiophene rings along the backbone.

C.2   The Effect of Including Explicit Charges in the Model

Modeling the long-range electrostatic interactions between P3HT chains is both computationally expensive and challenging to do accurately. Including explicit partial charges results in a factor of 3 increase in simulation time for the systems studied here. Additionally, small conformational changes in conjugated polymers such as P3HT give rise to significant changes in electron densities and therefore the partial charges associated with each atom fluctuate over time and space. Furthermore, prior calculations for modeling P3HT electrostatics do not reach consensus on which partial charges are correct, or sufficient[2,7,22] In this section we compare our base case implicit charge model against one with explicit partial charges whose forces are calculated with the Fourier based particle-particle-particle-mesh Ewald summation method[16]. We quantify structural and performance differences between these approaches and show the implicit charge model sufficiently captures the relevant assembly physics.

The charges used in this study are determined through first-principle calculations with the NWChem software[23] with the Becke, three-parameter, Lee-Yang-Parr (B3YLP) hybrid functional[1] with the 6-311++g** basis set[14]. The partial charges for the atoms within the thiophene ring from literature and those calculated in this study are shown in Table C.4.

From the values presented in Table C.4 we utilize the pentamer values for our

Table C.4: The charges for atoms given in literature for all-atom simulations. *Indicates that this was a united-atom simulation. †Indicates that these were calculated as part of this work by the NWChem program [23].

Study	S	C1	C2	C3	C4
Bhatta [2]	-0.22	0.18	-0.04	-0.31	0.03
Obata [21]	-0.07	0.12	-0.07	-0.27	0.05
Moreno [20]	-0.12	-0.01	-0.01	-0.18	-0.03
Huang [10]	-0.15	-0.14	0.075	-0.18	-0.18
Borzdun [3]	0.33	-0.24	-0.28	-0.16	0.08
D'Avino* [8]	-0.06	-0.21	0.22	0.10	-0.29
Trimer†	-0.17	0.28	-0.03	-0.37	-0.11
Pentamer†	-0.16	0.24	0.01	-0.37	-0.14

simulations (Figure C.3a). Further, we note that we also zero-out the charge in the simulation by subtracting the average per-particle charge from every simulation bead to account for rounding errors that may occur on individual atoms.
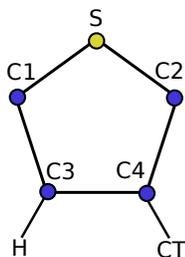
Figure C.2: Labels for identifying the atoms in the thiophene in Table C.4.

Since a united-atom model is used in this work, we sum the partial charges of the hydrogen atoms into the atom to which they are directly bonded. These values are shown in Table C.5. The CTN labels, in which $N=1\rightarrow6$, are used to identify the aliphatic carbons with CT1 being bonded to the thiophene ring and CT6 being the farthest away from the thiophene ring. We note that the C3 presented in Table C.5 is the same as that shown in Table C.4, but with the hydrogen considered. Because the works of Moreno and Huang consider polythiophenes with no alkyl sidechains, we

exclude these works from the list of the side-chain charges.

Table C.5: The charges from literature summed up so the hydrogens are considered with the carbons. C3 is the same as above. CT1 is bonded to the thiophene ring and each CT$N = 1 \rightarrow 6$ is extending from the ring.

Study	C3	CT1	CT2	CT3	CT4	CT5	CT6
Bhatta[2]	-0.08	0.14	-0.04	0.03	0.04	-0.03	-0.01
Obata[21]	-0.11	0.07	0.03	-0.04	0.01	0.09	-0.08
Borzdun[3]	0.06	0.02	-0.01	-0.01	0.0	0.0	0.0
D'Avino[8]	0.10	0.22	0.08	-0.07	0.01	0.04	-0.03
Trimer	-0.12	0.19	0.06	-0.16	0.11	0.07	-0.10
Pentamer	-0.13	0.19	0.07	-0.17	0.10	0.09	-0.10

We assign partial charges to simulation beads (Figure C.3a). We then employ scattering, $g(r)$, and cluster analysis described above to to determine the differences in molecular packing. In our optimized implicit-charge model, we reduce the thiophene ring bead diameters to ~ 3.44 Å from the ~ 3.7 Å used in OPLS-UA and Amber[12,24], which permits the thiophene rings to π-stack at the same length-scale as in experiments. In the explicit charge model benchmarked here, we employ the partial charges, but do not modify the Lennard-Jones ϵ or σ parameters from our optimized implicit-charge model. These parameters *should* be updated to create an optimized explicit-charge model, but we hypothesize that such efforts are not worth their cost and check here to see how the addition of explicit charges affects structure and performance.

To test whether explicit charges destabilize an already-equilibrated structure from an implicit-charge model (Figure 7a), we instantaneously "turn on" electrostatics with the above partial charges and re-equilibrate a 100 15-mer $T = 600$ K, $\rho = 1.11$ g/cm^3 Protocol (2) system over 0.3 μs. The resulting morphology (Figure C.3b) is not significantly destabilized, and shows qualitative agreement with GIXS and $g(r)$ metrics of the implicit-charge case. The primary quantitative difference arises from
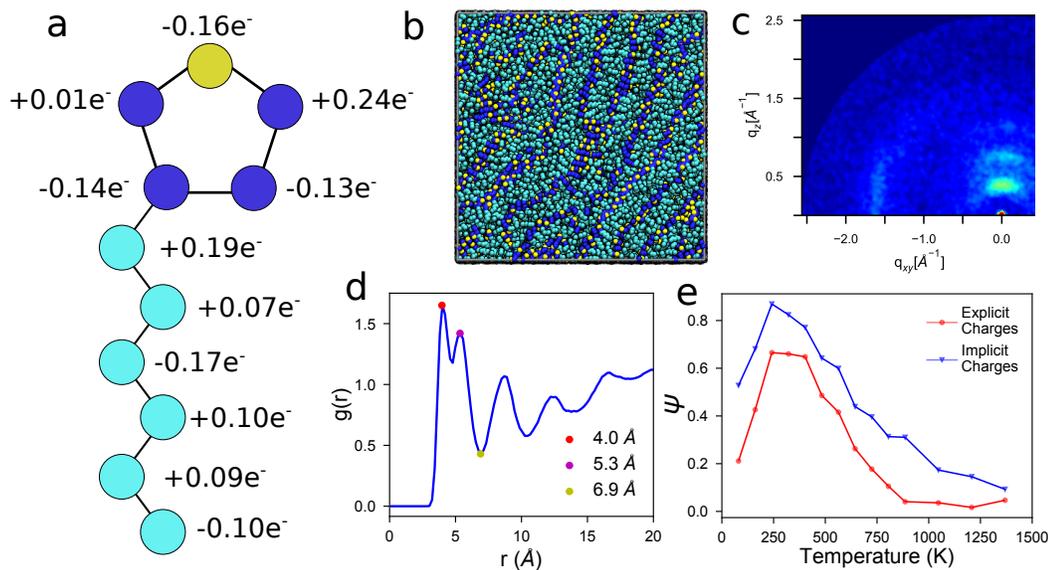
Figure C.3: (a) The partial charges applied to the thiophene ring in the explicit-charge model. These charges were generated from a pentamer chain using first-principle calculations. Incorporating these into the equilibrated implicit-charge structure at the same state-point Figure 7a results in (b), which is visually indistinguishable (CA-dark blue, S-yellow, CT-cyan; state-point: explicit charges, $T \sim 600$ K, $\varepsilon_s = 0.8$, $\rho = 1.11$ g/cm^3). Short- and long-range structural features observed in the (c) GIXS (averaged over 6 orientations) and (d) thiophene centroid $g(r)$ (compared to Figure 6a and Figure 7e respectively) show that the explicit-charge model with $\sigma_{LJ} = 3.7$ Å is consistent with the implicit-charge model with a reduced $\sigma_{LJ} = 3.44$. (e) The explicit- (red circles) and implicit-charge (blue triangles) systems exhibit similar trends in order (ψ) over different temperatures (state-point: $\varepsilon_s = 0.2$ and $\rho = 0.72$ g/cm^3). However, explicit-charge systems systematically obtain a lower degree of order.

the electrostatic repulsion of identical beads when thiophene rings are aligned: With explicit charges, aligned rings are on average 4.0 Å apart, instead of 3.9 Å as in the implicit charge case.

To test differences in self-assembled structure we perform new simulations at $\varepsilon_s = 0.2$ with Protocol (2) and the base-case 100 15-mers, but with explicit electrostatic charges active throughout. Figure C.3e shows that order parameter ψ dependence on temperature is qualitatively the same for both implicit- and explicit-charge models.

Independent of charge consideration, we observe kinetic trapping at low temperatures (< 250 K), increased order at moderate temperatures (250 K to 500 K), and decreasing order as temperatures increase (> 750 K). The explicit-charge model is systematically less ordered ($\psi_{explicit} \sim \psi_{implicit} + 0.1$) than the implicit model across all temperatures studied here, but otherwise assembles the same structures: lamellae of π-stacked backbones below a transition temperature and disordered melts above. On one hand, we interpret these results to justify the factor of 3 performance benefit of the implicit charge model. On the other hand, we acknowledge that the self-assembled structures are not identical between these two cases, and this *could* have significant ramifications, for example in predicting charge mobilities. We therefore recommend using implicit charge models for screening large parameter spaces as performed here, and then subsequently performing more expensive simulations when they are warranted for additional insight.

C.3 The Effect of Considering Longer Chains

In this section we compare performance and packing of "long" 50 monomer (50mer) chains relative to the 15mer base-case used throughout this work. Commercially available P3HT chains (from e.g. Sigma-Aldrich, Reike Metals) are significantly longer, typically ranging from 125 to 625 monomers, and more polydisperse (PDI~ 2) despite the feasibility of synthesizing narrow chain length distributions[6,9,13]. However, 50mers are the longest chains that we can simulate at the same conditions as the 15mers, while ensuring a chain cannot interact with itself through the periodic simulation boundaries. This is a conservative modeling choice, as we do not anticipate that such self-interactions would significantly impact self-assembled morphology. Ad-

ditionally, we expect the 50mer chains to provide useful information about the effect of simulating longer chains compared to the 15mer base case, despite being shorter than commercially available. We simulate 300 50mer chains at $\rho = 1.11$ g/cm^3 using Protocol (2) at $T \sim 600$ K and $\varepsilon_s = 0.8$, and compare against 1000 15-mers at the same conditions as subsection 4.5 of the main text.

The 50mer systems (Figure C.4a) demonstrate less ordering than their 15mer counterparts under the same conditions. Using clustering analysis, we still see short-range order exists in the form of small crystallites that are randomly oriented with respect to each other (Figure C.4b). The lack of long-range order for 50mers is apparent from GIXS (Figure C.4c), which lacks distinct reflections corresponding to lamellae of π-stacked backbones as seen before. The thiophene ring $g(r)$ (Figure C.4d) shows some short-range ordering, with anti-aligned thiophene rings (expected peak at 5.3 Å) less common in 50mer systems. The major structural difference between 50mers and 15mers is that the 50mers are observed to "fold" and form π-stacks with themselves (Figure C.4e) Such self-stacking allows a single chain to form multiple layers in one crystallite, and is observed experimentally [5,11,18].

We estimate the time required for the system containing 300×50 monomer long chains (50mer) to relax to an equilibrated structure compared to the systems containing 1000, 15 monomers long chains (15mer) by comparing the evolution of the Lennard-Jones potential energies and also the structure factor over time. As a first test, we assume that the systems have the same equilibrated structure and that these two structures will have the same non-bonded energies. Therefore, we fit the 50mer and 15mer Lennard-Jones energy over time to equations and determine where the 50mer's equation will equal the final energy of the 15mer system. In choosing what equation should be used to represent the energy evolution, we qualitatively observe
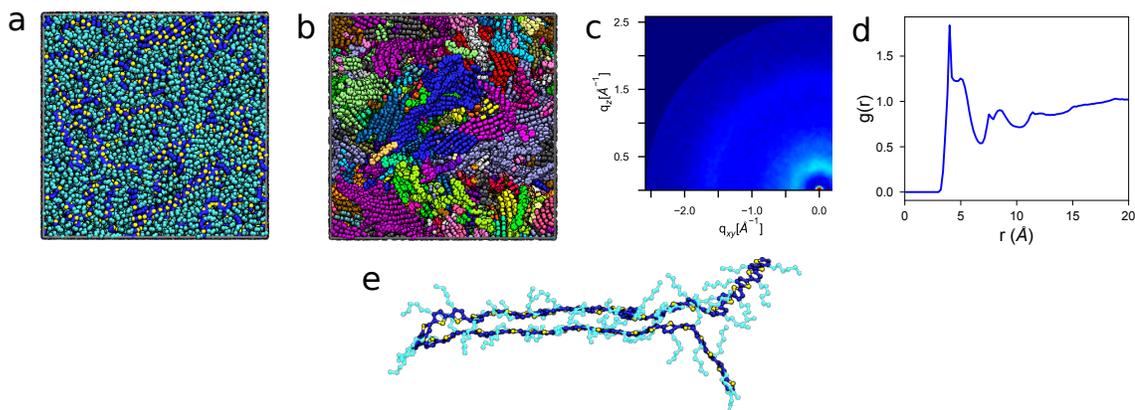
Figure C.4: (a) 300×50mer chains produce morphologies containing no visually discernible periodic order at $T \sim 600$ K and $\varepsilon_s = 0.8$ (CA-dark blue, S-yellow, CT-cyan). (b) However, short-range order is visible when thiophene rings are colored by cluster. (c) Short-range order is confirmed through simulated GIXS (averaged over 60 orientations) and (d) the $g(r)$ of the thiophene centroids, which both show less defined peaks than in the 15mer case (Figure 6a and Figure 7e respectively). (e) The 50mer chains are long enough to undergo π-stacking with themselves, which is observed experimentally[5,11,18] and not observed with the shorter chains.

that the P3HT systems undergoes three stages during relaxation: fast exponential decay, slow approximately linear decay and constant. As such, we fit the per-particle potential decay over simulation time t_s to:

$$f(t_s) = a \times exp(b \times t_s) + c \times t_s + d, \qquad (C.1)$$

to get the coefficients a, b, c, and d for the exponential and linear terms; the fitted curve is showed by the dashed red line in Figure C.5a. With the coefficients known, we solve the equation to determine when it would equal the final energy of the 15mer system, which results in approximately 9,000 times longer for the 50mer system to run to equal the final energy as the 15mer system.

However, it is not necessarily valid to assume that the 50 and 15mer systems will have the same final energy. For instance, it is possible that the 50 system (which

is more prone to entanglements) will reach a metastable state and will require the lifetime of the universe to relax out of the metastable state. Such behavior often occurs in real systems, in that P3HT devices have crystalline and amorphous regions rather than the more energetically favored perfect crystal. As such, if we do not assume that the two systems will have the same equilibrium structure and hence same energy, we can instead utilize the rate of decay of the energy as our descriptor for system evolution. Because the linear decay is the slowest process, we can compare the ratio of the linear coefficients between the 50mer and 15mer system to describe the difference in relaxation times. As such, we find that the 50mer's coefficient is -1.94 $\times$ 10^{-5} and the 15mer's is -3.6 $\times$ 10^{-5}. This suggests a $\sim 2\times$ faster decrease in energy in the 15mer system than the 50mer system.
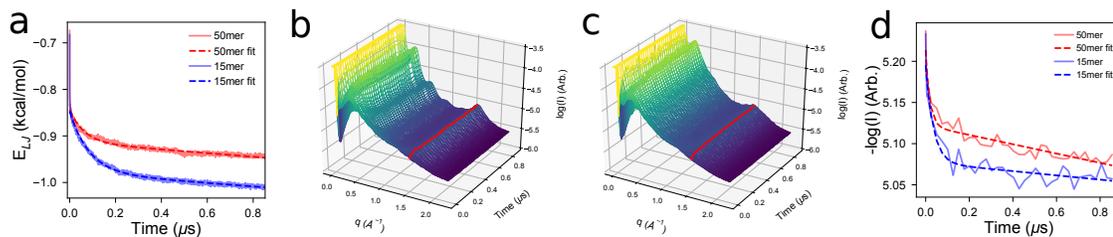
Figure C.5: (a) The potential energy in the (blue) 15mer system decays faster and achieves a final lower per-particle than the (red) 50mer system. The structure factor - a radial average of the GIXS pattern - shows the growth of peaks corresponding to (red) π-stacking over time in the (b) 15mer and the (c) 50mer system. (d) The evolution of the red peaks in b and c show that the 15mer system reaches a stable state faster than the 50mer system.

Rather than using the energy, which is an indirect measure of the structure's evolution, we can instead utilize a direct measurement of the structure such as the structure factor. We therefore conduct the scattering experiment to obtain the structure factor - a radial average of the GIXS pattern - every 25 ns of simulation run time (Figure C.5b and c). We record the height of the peak in the structure factor

located at 1.6 Å^{-1}, corresponding to π-stacking, over time. We again fit the evolution of the structure factor peak over time to an exponential function and compare the relative changes in structure evolution between the 50mer and 15mer systems (Figure C.5d). From this, we again predict that the 50mer system will require twice as long to reach equivalent scattering intensities of that of the 15mer system. Because the structure factor is the most direct calculation of structure and it agrees with the change of energy we conclude that the 50mer system requires twice as long as the 15mer system to order. In summary, the longest chains that can be practically equilibrated should be used, 15mers are sufficient for predicting experimental GIXS patterns, and systems of 50mers can in principle be equilibrated, but the factor of two increase to simulation time precludes routine sweeps of large parameter spaces.

C.4 Order Parameter Explanation and Example
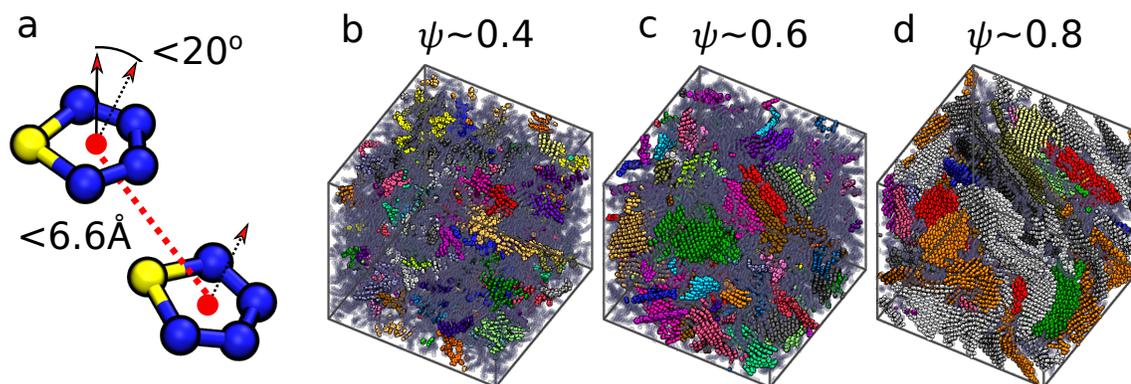
Figure C.6: (a) Two thiophene rings are considered clustered if the distance between their geometric centers is less than 6.6 Å and the angle between the ring normal vectors is less than 20°. Clusters that are considered large (greater than 6 members) are shown with colored, opaque beads, whereas thiophenes in small clusters are colored with gray, diffuse beads. With increasing ψ, $0.4 \rightarrow 0.6 \rightarrow 0.8$ (b-d), the system goes from crystallites within an amorphous matrix to primarily crystallites with intermittent amorphous regions.

The order parameter ψ presented here was developed in previous work as a way to distinguish systems with few, large crystallites from systems with many small crystallites[19]. ψ is a measurement of the ratio of rings comprised into large clusters to all rings in the system e.g. a system with $\psi = 0.8$ would have 80% of the thiophene rings contained in "large" clusters. To identify unique crystallites, we identify thiophene rings that are clustered together based on two criteria (Figure C.6a):

1. The centers of geometry of two rings are separated by less than 6.6 Å.

2. The vector normals to the thiophene ring planes are oriented less than 20° from each other.

These separation and orientational values are based on the minimum of the radial distribution function and the decay of the transfer integral between two rings (a measure of the electron orbital overlap), respectively[15]. Clusters are considered large when they contain more than six members, a cut-off based on the distribution of cluster sizes for different systems at various degrees of order, in which only $\sim 10\%$ of clusters being larger than six members. Representative systems with ψ values of $\sim$ 0.4, 0.6, and 0.8 are shown in Figure C.6b-d. These systems reveal that systems with $\psi \sim 0.4$ are composed of an amorphous matrix with crystallites dispersed throughout and the $\psi \sim 0.8$ system is primarily crystalline with intermittent amorphous regions.

C.5 Linking System Evolution to Energy

In this section we present our analysis of the Lennard-Jones energy to infer structural evolution. We detect three transients in the energy curve corresponding to different
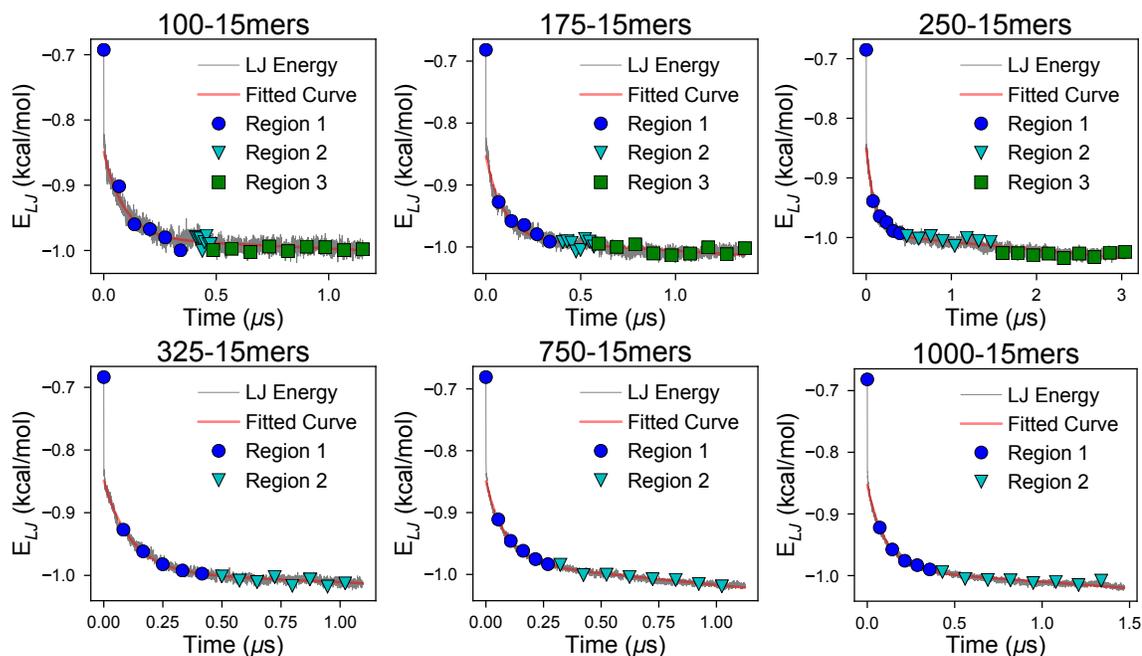
Figure C.7: The per-particle Lennard-Jones potential energies for the various systems shows three regions, highlighted with blue circles, cyan triangles, and green squares for the systems with various sizes.

phases of morphology evolution: exponential decay representing formation of crystals from a disordered configuration, slow decay signifying the process of going from many small crystallites to few, highly ordered crystals and constant energy signifying equilibrium. The regions corresponding to each transient are shown in Figure C.7 for simulations containing various amounts of 15mer chains. We note that not all the final energies are exactly equal. This is due to in part to systems reaching metastable states, which will require indefinite run times to relax out of. Additionally, changes in the simulation volume between the simulations can lead to unfavorable box lengths. For example, one system may be able to form a more perfect crystal when the simulation box length is commensurate with the crystal structure - resulting in a lower energy, whereas another system may never be able to form a perfect crystal because

its periodicity is interrupted by the periodic boundary of the simulation volume. To quantify the duration of the exponential decrease, we conduct a least-squares fit of the per-particle Lennard-Jones energy to C.1, which is also used to compare the 15mer and 50mer systems. We consider a linear term in addition to an exponential term to better account for the slow transformation in Region 2, which is qualitatively linear. We calculate the derivate of C.1 to obtain the exponential and linear components of the slope, and identify the transition time t_s as the point at which the linear component becomes a larger than the exponential component.

To determine when equilibrium is reached and Region 2 is ended, we must first assume that the final energy measurement is in an equilibrium state (which is only a valid assumption for the systems with fewer beads, < 325 15mers). We next assign the measured energies to bins containing 200 measurements each, which equates to 36 ns, and calculate the average $\bar{E}_i$ and standard deviation σ_i of each bin. We then iterate through the bins in reverse order and compare $\bar{E}_i$ and σ_i the final bin's energy $\bar{E}_{final}$ and deviation σ_{final}. When $\bar{E}_i$ is calculated to be more than the sum of the standard deviations higher than $\bar{E}_{final}$, we consider the system as not equilibrated, i.e.:

$$\text{System is} \begin{cases} \text{Not Equilibrated,} & \text{if } \bar{E}_i - \bar{E}_{final} > \sigma_i + \sigma_{final} \\ \text{Equilibrated,} & \text{otherwise.} \end{cases} \tag{C.2}$$
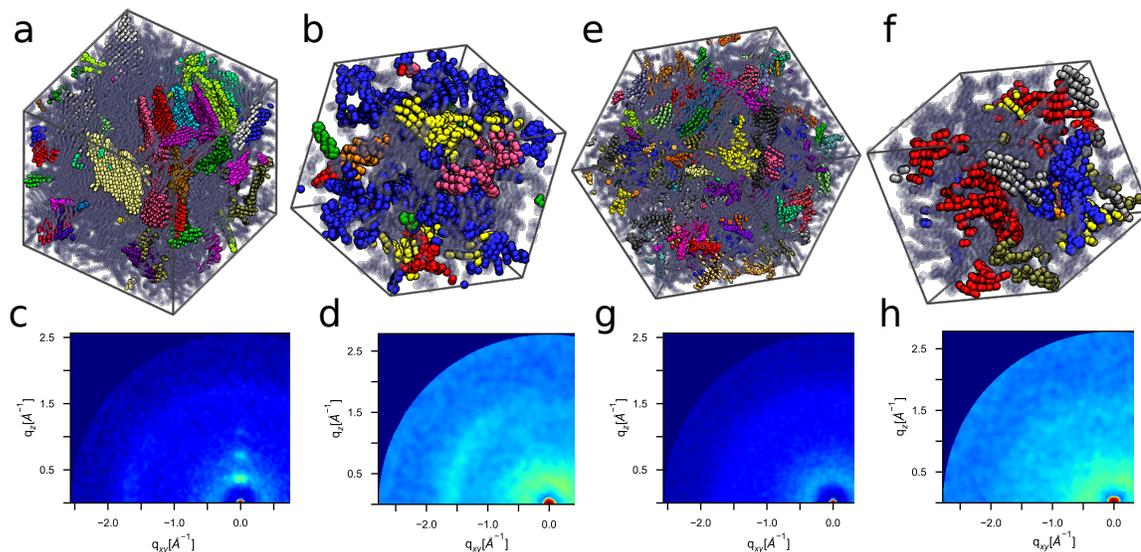
Figure C.8: In cases where amorphous and structured regions exist, large volumes (a) provide more insight than their smaller counterparts (b) in that systems containing 100 15mers do not show periodic lamellae features seen in larger systems (c, d). Conversely in disordered systems, large volumes do not provide significant structural insight as compared to small volumes (e-h).

C.6 Small and Large Comparison for Semi-Ordered and Disordered Systems

In the main text we consider size effect differences between the ordered large and small systems, here we expand this analysis to large and small semi-ordered and disordered systems (Figure C.8). In the large semi-ordered system, some crystallites have aggregated to form alkyl-stacked lamellae (shown with colored, opaque beads), corresponding to the experimentally observed crystalline domains surrounded by an amorphous matrix (shown with gray, diffuse beads)[4]. However, in the small system, one large crystallite interspersed with smaller crystallites dominates the morphology, and the alkyl-stacked lamellae are not observed in the smaller system. The difference between these two structures is confirmed with GIXS patterns: periodic

alkyl stackings are only observed in the large system. That said, the π-stacking feature along the (010) planes forms a diffuse band rather than a distinct peak - likely due to the wide range of crystallite orientations in the large morphology in the large system, whereas this feature is much stronger in the smaller system's pattern. Therefore, there is some dependence on the system size when looking at semi-ordered systems: larger systems are able to have regions of high order and amorphous regions, which are missing in smaller systems that tend to form fewer, larger crystals. This corresponds to crystallite sizes measured experimentally ($\sim$ 10-20 nm), in that the small system with a box length of $\sim$ 7 nm is smaller than the typical crystallite, whereas the larger systems with length 15 nm are the same as the average experimentally observed crystals[17]. We therefore suggest using larger system sizes ($N = 165,000$) to more accurately represent the long-range ordering characteristics of polymers.

In the case of disordered systems, both the 100 and 1000 chain systems are similar: they are comprised of many small, randomly-oriented regions of π-stacking, consisting of just a couple of chains in each case. The GIXS patterns corroborate the visual observations: there is little periodic order within the system. As such there is minimal system size dependence when the structures are more disordered; smaller simulations can be used to accurately investigate high temperatures. In summary, smaller system sizes are beneficial for fast investigations of the relationship between state-point and structure, whereas larger systems including more molecules are be needed to explore large-scale structural features.

BIBLIOGRAPHY

[1] Axel D. Becke. Density-Functional Thermochemistry. III. The Role of Exact Exchange. *The Journal of Chemical Physics*, 98(7):5648–5652, 1993.

[2] Ram S. Bhatta, Yeneneh Y. Yimer, David S. Perry, and Mesfin Tsige. Improved Force Field for Molecular Modeling of Poly(3-hexylthiophene). *The Journal of Physical Chemistry B*, 117(34):10035–10045, 2013.

[3] Natalia I. Borzdun, Sergey V. Larin, Stanislav G. Falkovich, Victor M. Nazarychev, Igor V. Volgin, Alexander V. Yakimansky, Alexey V. Lyulin, Vikas Negi, Peter A. Bobbert, and Sergey V. Lyulin. Molecular Dynamics Simulation of Poly(3-hexylthiophene) Helical Structure In Vacuo and in Amorphous Polymer Surrounding. *Journal of Polymer Science, Part B: Polymer Physics*, 54(23):2448–2456, 2016.

[4] M. Brinkmann and P. Rannou. Effect of Molecular Weight on the Structure and Morphology of Oriented Thin Films of Regioregular Poly(3-hexylthiophene) Grown by Directional Epitaxial Solidification. *Advanced Functional Materials*, 17(1):101–108, 2007.

[5] Martin Brinkmann and Jean Claude Wittmann. Orientation of Regioregular Poly(3-hexylthiophene) by Directional Solidification: A Simple Method to Reveal the Semicrystalline Structure of a Conjugated Polymer. *Advanced Materials*, 18(7):860–863, 2006.

[6] Hugo A. Bronstein and Christine K. Luscombe. Externally Initiated Regioregular P3HT with Controlled Molecular Weight and Narrow Polydispersity. *Journal of the American Chemical Society*, 131(36):12894–12895, 2009.

[7] Sefa Dag and Lin-Wang Wang. Packing Structure of Poly(3-hexylthiophene) Crystal: Ab Initio and Molecular Dynamics Studies. *The Journal of Physical Chemistry B*, 114(18):5997–6000, 2010.

[8] Gabriele D'Avino, Sébastien Mothy, Luca Muccioli, Claudio Zannoni, Linjun Wang, Jérôme Cornil, David Beljonne, and Frédéric Castet. Energetics of Electron-Hole Separation at P3HT/PCBM Heterojunctions. *Journal of Physical Chemistry C*, 117(25):12981–12990, 2013.

[9] Robert B. Grubbs and Robert H. Grubbs. 50th Anniversary Perspective: Living Polymerization-Emphasizing the Molecule in Macromolecules. *Macromolecules*, 50(18):6979–6997, 2017.

[10] David M. Huang, Roland Faller, Khanh Do, and Adam J. Moulé. Coarse-Grained Computer Simulations of Polymer/Fullerene Bulk Heterojunctions for Organic Photovoltaic Applications. *Journal of Chemical Theory and Computation*, 6(2):526–537, 2010.

[11] Kyo Jin Ihn, Jeff Moulton, and Paul Smith. Whiskers of Poly(3-alkylthiophene)s. *Journal of Polymer Science Part B: Polymer Physics*, 31(6):735–742, 1993.

[12] William L. Jorgensen, Jeffry D. Madura, and Carol J. Swenson. Optimized Intermolecular Potential Functions for Liquid Hydrocarbons. *Journal of the American Chemical Society*, 106(22):6638–6646, 1984.

[13] Felix P. V. Koch, Paul Smith, and Martin Heeney. Fibonacci's Route to Regioregular Oligo(3-hexylthiophene)s. *Journal of the American Chemical Society*, 135(37):13695–13698, 2013.

[14] R. Krishnan, J. S. Binkley, R. Seeger, and J. A. Pople. Self-Consistent Molecular Orbital Methods. XX. A Basis Set for Correlated Wave Functions. *The Journal of Chemical Physics*, 72(1):650–654, 1980.

[15] Yi-kang Lan and Ching-i Huang. A Theoretical Study of the Charge Transfer Behavior of the Highly Regioregular Poly-3-hexylthiophene in the Ordered State. *The Journal of Physical Chemistry B*, 112(47):14857–14862, 2008.

[16] David N. LeBard, Benjamin G. Levine, Philipp Mertmann, Stephen A. Barr, Arben Jusufi, Samantha Sanders, Michael L. Klein, Athanassios Z. Panagiotopoulos, A. Barr, Arben Jusufi, Samantha Sanders, and L. Klein. Self-Assembly of Coarse-Grained Ionic Surfactants Accelerated by Graphics Processing Units. *Soft Matter*, 8(8):2385–2397, 2012.

[17] Feng Liu, Dian Chen, Cheng Wang, Kaiyuan Luo, Weiyin Gu, Alejandro L. Briseno, Julia W. P. Hsu, and Thomas P. Russell. Molecular Weight Dependence of the Morphology in P3HT:PCBM Solar Cells. *ACS Applied Materials and Interfaces*, 6(22):19876–19887, 2014.

[18] Elena Mena-Osteritz, Alexander Meyer, Bea M. W. Langeveld-Voss, Rene A. J. Janssen, E. W. Meijer, and Peter Bäuerle. Two-Dimensional Crystals of Poly(3-Alkyl- thiophene)s: Direct Visualization of Polymer Folds in Submolecular Resolution. *Angewandte Chemie International Edition*, 39(15):2679–2684, 2000.

[19] Evan D. Miller, Matthew L. Jones, and Eric Jankowski. Enhanced Computational Sampling of Perylene and Perylothiophene Packing with Rigid-Body Models. *ACS Omega*, 2(1):353–362, 2017.

[20] Margherita Moreno, Mosè Casalegno, Guido Raos, Stefano V. Meille, and Riccardo Po. Molecular Modeling of Crystalline Alkylthiophene Oligomers and Polymers. *The Journal of Physical Chemistry. B*, 114(4):1591–1602, 2010.

[21] Shigeaki Obata and Yukihiro Shimoi. Control of Molecular Orientations of Poly(3-hexylthiophene) on Self-Assembled Monolayers: Molecular Dynamics Simulations. *Physical Chemistry Chemical Physics*, 15(23):9265, 2013.

[22] Tran Thinh To and Stefan Adams. Accurate Poly(3-hexylthiophene) Force-field from First-Principle Modeling. *Nanoscience and Nanotechnology Letters*, 4(7):703–711, 2012.

[23] M. Valiev, E. J. J. Bylaska, N. Govind, K. Kowalski, T. P. P. Straatsma, H. J.J. J. J. Van Dam, D. Wang, J. Nieplocha, E. Apra, T. L. L. Windus, and W. A. A. De Jong. NWChem: A Comprehensive and Scalable Open-Source Solution for Large Scale Molecular Simulations. *Computer Physics Communications*, 181(9):1477–1489, 2010.

[24] Lijiang Yang, Chun Hu Tan, Meng Juei Hsieh, Junmei Wang, Yong Duan, Piotr Cieplak, James Caldwell, Peter A. Kollman, and Ray Luo. New-Generation Amber United-Atom Force Field. *Journal of Physical Chemistry B*, 110(26):13166–13176, 2006.

APPENDIX D

STRUCTURAL AND CHARGE TRANSPORT PREDICTIONS FOR DIBENZO-TETRAPHENYL PERIFLANTHENE AND FULLERENE MIXTURES - SUPPORTING INFORMATION

D.1 Bonded Force Field Interactions

In this section we list the bonded parameters for the DBP-Fullerene simulations. These interactions are taken from the study of biphenyl[1] and the OPLS-UA force field[2]. The non-bonded interactions are presented in Section 5. Likewise, the rigid bodies are explained in Figure 5.1. With these rigid bodies, the bonded constraints listed here only apply to the flexible bodies: the bands between the periflanthene backbone and the benzene rings and within the side chain of the fullerenes.

The bond coefficients are shown in Table D.1 and these bonds follow the harmonic form shown in Equation 2.4.

Table D.1: The bond constraints used in our force-field to simulate DBP and Fullerene derivatives. †CA-CA bond between the periflanthene backbone and benzene rather than within an aromatic body.

Bond	l_0 (Å)	k_b (kcal mol^{-1} Å^{-2})
CA-CA†	1.52	309.6
CA-CT	1.51	430.0
CT-CT	1.53	363.5
CT-O	1.23	705.3

The angle coefficients are shown in Table D.2 and these angles follow the harmonic form shown in Equation 2.5.

Table D.2: The angle constraints used in our force-field to simulate DBP and Fullerene derivatives. †CA-CA-CA angle between the periflanthene backbone and benzene rather than within an aromatic body.

Angle	θ (rad)	k_θ (kcal mol^{-1} rad^{-2})
CA-CA-CA†	2.09	42.6
CA-CA-CT	2.09	46.4
CA-CT-CT	1.91	27.1
CT-CT-CT	1.91	27.1
CT-CT-O	2.06	56.3
CT-O-CT	2.04	56.3
O-CT-O	2.15	56.3

The dihedral coeffecients are shown in Table D.3. These dihedral coefficients are based on the OPLS form shown in Equation 2.8.

Table D.3: The dihedral constraints used in our force-field to simulate DBP and Fullerene derivatives. [†]CA-CA-CA-CA dihedral between the periflanthene backbone and benzene rather than within an aromatic body.

Dihedral	k_{d_1} (kcal mol^{-1})	k_{d_2}	k_{d_3}	k_{d_4}
CA-CA-CA-CA	0.0	3.0	0	0.6
CA-CA-CA-CT	1.4	-0.7	0.2	0.0
CA-CA-CT-CT	1.4	-0.7	0.2	0.0
CA-CT-CT-CT	1.4	-0.7	0.2	0.0
CT-CT-CT-CT	1.3	-0.1	0.4	0.0
CA-CA-CT-CA	0.0	2.2	0.0	0.0
CT-CT-O-CT	4.6	4.7	0.0	0.0
CT-O-CT-O	0.0	4.7	0.0	0.0

BIBLIOGRAPHY

[1] Sven Hauptmann, Thomas Mosell, Stephan Reiling, and Jürgen Brickmann. Molecular Dynamics Simulations of the Bulk Phases of 4-cyano-4'-n-pentyloxybiphenyl. *Chemical Physics*, 208(1):57–71, 1996.

[2] William L. Jorgensen and Julian Tirado-Rives. The OPLS [Optimized Potentials for Liquid Simulations] Potential Functions for Proteins, Energy Minimizations for Crystals of Cyclic Peptides and Crambin. *Journal of the American Chemical Society*, 110(6):1657–1666, 1988.

APPENDIX E

TYING TOGETHER MULTISCALE CALCULATIONS FOR CHARGE TRANSPORT IN P3HT: STRUCTURAL DESCRIPTORS, MORPHOLOGY, AND TIE-CHAINS - SUPPORTING INFORMATION[1]

E.1 ZINDO and DFT Comparison

In this section, we compare the ZINDO calculation of electronic properties used in this investigation, to a more rigorous DFT method to determine the accuracy of semi-empirical frontier molecular orbital energy calculations for P3HT. We use three representative P3HT chromophore pairs selected from an equilibrated, ordered test morphology, visualizations of which are depicted in Figure E.1. The DFT calculations were performed using the B3LYP functional[1] and the 6311++g** basis set[3].

Table E.1: A comparison of the HOMO splitting and calculated transfer integrals for three representative P3HT chromophore pairs

HOMO Splitting	DFT (eV)	ZINDO/S (eV)
0469-3714	0.196	0.095
0841-1237	0.199	0.058
2032-2900	0.086	0.008
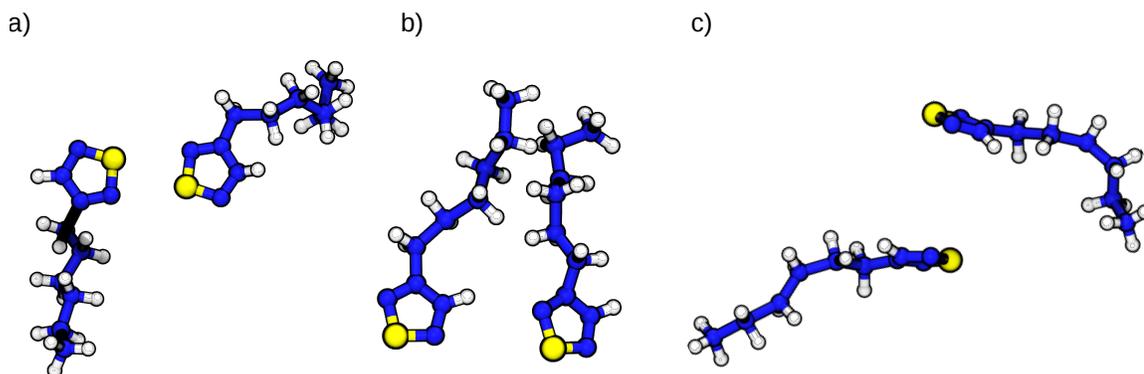
Figure E.1: The three representative chromophore pairs used to investigate the accuracy of the ZINDO/S semi-empirical method. a) 0469-3714, b) 0841-1237, c) 2032-2900. Terminating hydrogens were added for the QCCs based on position in the thiophene ring.

The calculated electronic properties of the chromophore pairs are shown in Table E.1. ZINDO appears to consistently underpredict the HOMO splitting, which would lead to lower transfer integrals and slower transport than expected from more rigorous DFT methods. However, the ZINDO calculations provide good agreement with the DFT results to within ~ 100 meV, which is already the rough cutoff for DFT accuracy. Furthermore, changes in transfer integral of factors of 2-3 are not expected to significantly affect the charge transport properties given that morphological changes can result in orders of magnitude differences. ZINDO calculations can be performed within 5-10 seconds for a chromophore pair, compared to several minutes to half an hour in the case of more rigorous DFT calculations (depending on the DFT level desired). As such, the computational throughput is significantly improved at the cost of the smaller reduction in accuracy - a critically important point given that each morphology can contain upwards of 40,000 chromophore pairs to be considered.

In summary, the vastly improved computational efficiency at the cost of a small reduction in accuracy of ZINDO/S justifies our use of the semi-empirical calculations

for our charge transport properties instead of more rigorous DFT methodologies.

E.2 Developing ψ' to Explicitly Consider Transfer Integrals

One possible short-coming in our modified order parameter ψ' is that it is an indirect predictor of the electronic transfer integral $J_{i,j}$ between two chromophores. The hopping rate between two chromophores is strongly dependent on the electronic transfer integral $J_{i,j}$ (see Equation 2 of the main text) and the energy level difference $\Delta E_{i,j}$ of the two chromophores. Additionally, $\Delta E_{i,j}$ is also partially encoded into $J_{i,j}$, in which chromophores with incompatible energy levels (for instance, a large $\Delta E_{i,j}$) will also reduce $J_{i,j}$ (see Equation 1 of the main text). As such, the transfer integral seems like a good "one-size-fits-all" parameter to describe clustering.
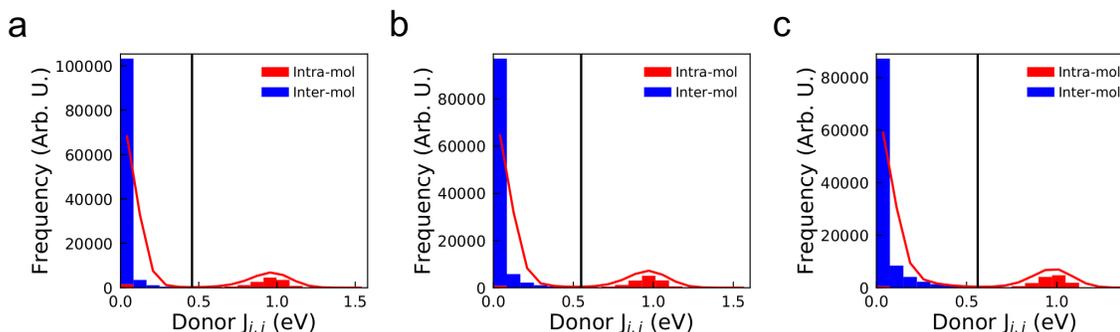
Figure E.2: Distributions of chromophore Voronoi neighbor transfer integrals for the representative 1,000 molecule a) amorphous, b) semi-crystalline, and c) crystalline morphologies. The red line shows the Gaussian filtered distribution shape that was used to determine the cluster cut-off criterion. The black vertical line shows the value of the cut-off criterion, which was automatically determined to be at the minimum for each system - $J_{i,j} > 0.562$, 0.549, and 0.457 eV for the crystalline, semi-crystalline, and amorphous morphologies respectively.

The transfer integral distributions for each representative system are shown in Figure E.2. In all three cases, the distribution has a large spike at very low transfer

integrals and a bump at high TI corresponding to pairs within the same P3HT chain. Initially, we set the transfer integral cut-off to the location of the minimum for each morphology, such that only connections with transfer integrals greater than the cut-off are added to the same cluster. It is convenient to set cut-offs to maxima and minima as these can be determined automatically, rather than being calibrated manually for each separate system. For the crystalline, semi-crystalline, and amorphous morphologies, the cut-offs were set to $J_{i,j} > 0.562$, 0.549, and 0.457 eV respectively.
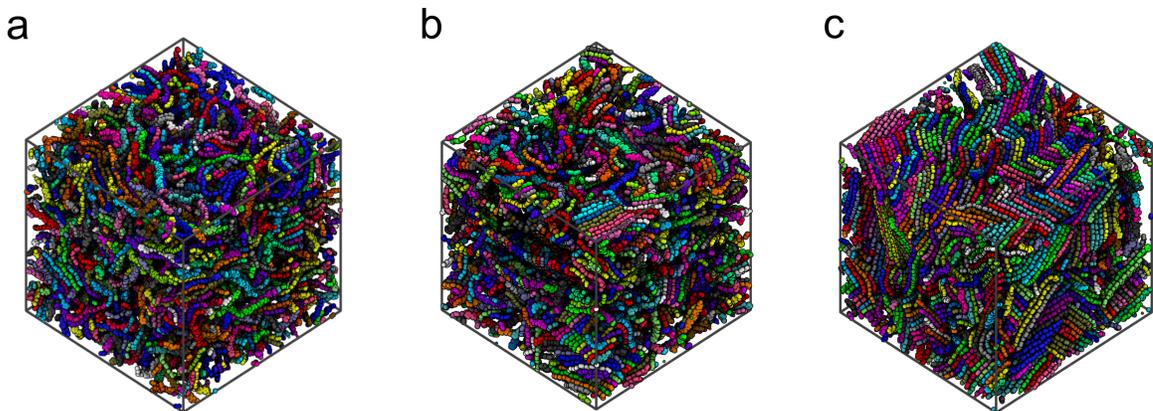
Figure E.3: Visualizations of the clusters in the a) amorphous, b) semi-crystalline, and c) crystalline systems with size > 6 monomer units. Clusters were determined based on an automatically-defined transfer integral cut-off for each system based on the distributions in Figure E.2.

Table E.2: Table of cluster statistics for the three systems, given the automatically-determined transfer integral cut-off criteria.

Property	Amorphous	Semi-Crystalline	Crystalline
Mobility (cm^2/Vs)	1.02×10^{-1}	1.63×10^{-2}	1.16×10^{-1}
$J_{i,j}$ cut-off (eV)	0.457	0.549	0.562
Total clusters (Arb. U.)	1067	1065	972
Large (> 6) clusters (Arb. U.)	964	941	873
Largest cluster size (Arb. U.)	60	60	90

The resultant cluster visualization in Figure E.3 suggests that these cut-off values are too large - in all morphologies, hops with $J_{i,j} > \sim 0.5$ eV are generally only

intra-molecular hops (red region in Figure E.2). This leads to nearly every chain in the system being considered an individual cluster, with few occurrences of clusters forming between multiple chains. There is no significant difference in the cluster distribution throughout the morphology between the three systems, suggesting that a different transfer integral cut-off should be used.
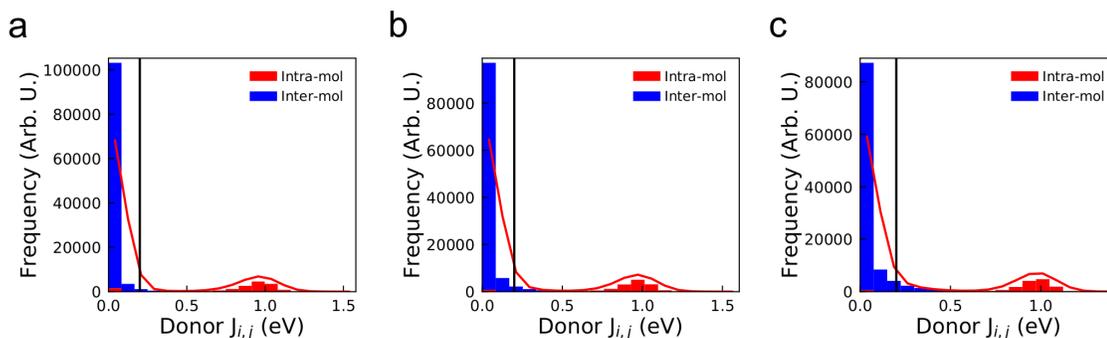
Figure E.4: Distributions of chromophore Voronoi neighbor transfer integrals for the representative 1,000 molecule a) amorphous, b) semi-crystalline, and c) crystalline morphologies. The red line shows the Gaussian filtered distribution shape that was used to determine the cluster cut-off criterion. The black vertical line shows the value of the cut-off criterion, $J_{i,j} > 0.2$ eV.

We can, for instance, reduce the cut-off to something smaller in order to include higher $J_{i,j}$ inter-molecular hops. This however, has the short-coming in that such a selection will likely be arbitrarily chosen, rather than an automatically identified minimum. Regardless, reducing the $J_{i,j}$ cut-off to 0.2 eV (Figure E.4) provides significantly improved results as now a non-negligible proportion of inter-molecular hops have $J_{i,j} >$ cut-off, thereby, allowing clusters to form between molecules.

Now, we compare the clusters identified with the $J_{i,j}$ cut-off between the three systems. The crystalline morphology shows one large cluster (shown in red) and a few smaller clusters with opposing grain orientations, indicating that the crystalline system will have a high connectivity. Conversely, the amorphous morphology is
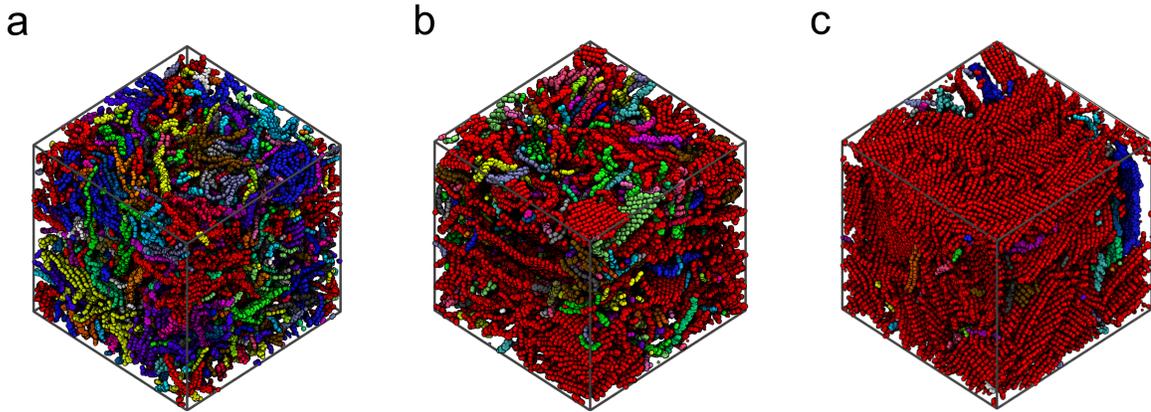
Figure E.5: Visualizations of the clusters in the a) amorphous, b) semi-crystalline, and c) crystalline systems with size > 6 monomer units, given the following clustering criteria: transfer integral > 0.2 eV.

Table E.3: Table of cluster statistics for the three systems, given the following clustering criterion: transfer integral $J_{i,j} > 0.2$ eV.

Property	Amorphous	Semi-Crystalline	Crystalline
Mobility (cm^2/Vs)	1.02×10^{-1}	1.63×10^{-2}	1.16×10^{-1}
$J_{i,j}$ cut-off (eV)	0.200	0.200	0.200
Total clusters (Arb. U.)	289	163	64
Large (> 6) clusters (Arb. U.)	273	151	51
Largest cluster size (Arb. U.)	2564	9914	12837

predicted to have poor connectivity based on this clustering metric stemming from the larger number of small clusters. However, the connectivity in the semi-crystalline morphology again shows a cluster arrangement intermediate between the other two. This cluster arrangement would again predict a high mobility for the crystalline morphology, a low mobility for the amorphous morphology and an intermediate morphology in the semi-crystalline case, which is contrary to our mobility calculations. Varying the cut-off to any consistent value between the three morphologies always results in this conclusion, suggesting that the transfer integral distribution is not an adequate way to identify clusters in the morphology.

E.3 Clustering Based on Hops

One short-coming of the previous clustering algorithms is that it considers charge transport between two chromophores in isolation. However, in the KMC algorithm, hops to all neighboring chromophores are considered and the preferential hop (based on the hopping rate between i and j and the random number x) will be chosen. As such, a "good" hop may not occur because there is a better hop.
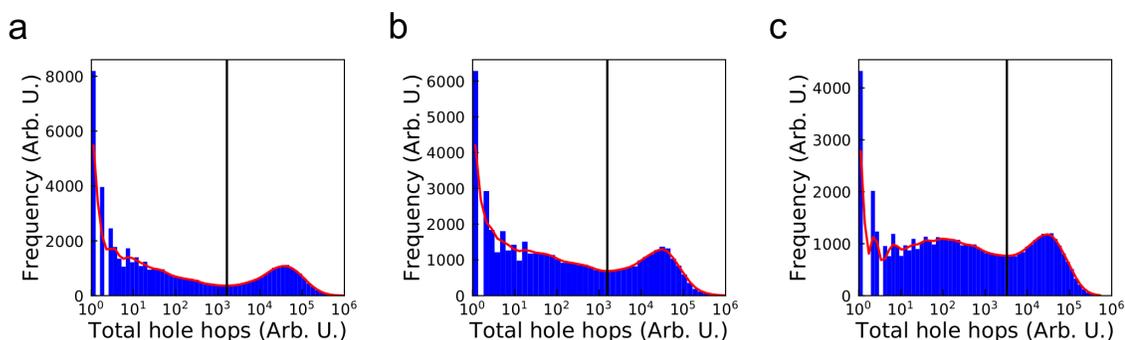
Figure E.6: Distributions of the frequencies with which carriers hop between chromophore Voronoi neighbors for the representative 1,000 molecule a) amorphous, b) semi-crystalline, and c) crystalline morphologies. The red line shows the Gaussian filtered distribution shape that was used to determine the cluster cut-off criterion. The black vertical line shows the value of the cut-off criterion, which was automatically determined to be at the final minimum of the frequency distribution: a total of 3264, 1566, and 1635 hops for the crystalline, semi-crystalline, and amorphous systems respectively.

As such, defining clusters based on regions in which charges will freely move is prudent, however, we must still identify a sensible cut-off in hopping frequency to separate these regions. The distributions of total hole hops between chromophore pairs in the three representative systems are shown in Figure E.6. Note that the x-axis in these plots is logarithmic, leading to quantization of the bins on the left-hand side of the plot. In all three systems, a second peak appears at high hop frequencies. This leads to a local minimum at 3264 hops in the crystalline case, 1566 hops in the

semi-crystalline case, and 1635 hops in the amorphous case. We therefore use these values as the clustering criteria - only chromophores with connections that are used more than this number during the simulation will be added to the same cluster. We note that the exact values of the cut-off criteria are strongly dependent on the duration of the KMC simulation; the value may change significantly if fewer carriers iterations are performed or if simulation times are reduced. In this study, all three systems used the same simulation time-scales for KMC and the same number of carriers were averaged over in order to obtain the charge transport properties.
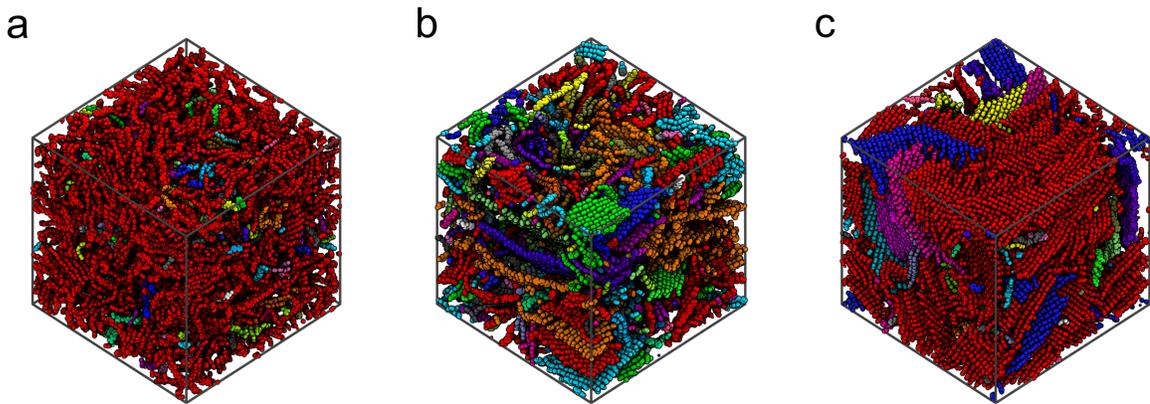
Figure E.7: Visualizations of the clusters in the a) amorphous, b) semi-crystalline, and c) crystalline systems with size > 6 monomer units. Clusters were determined based on an automatically-defined hopping frequency cut-off for each system based on the distributions in Figure E.6.

Table E.4: Table of cluster statistics for the three systems, given the automatically-defined total hop frequency cut-off criteria.

Property	Amorphous	Semi-Crystalline	Crystalline
Mobility (cm^2/Vs)	1.02×10^{-1}	1.63×10^{-2}	1.16×10^{-1}
Hop frequency cut-off (s^{-1})	1635	1566	3264
Total clusters (Arb. U.)	410	1376	418
Large (> 6) clusters (Arb. U.)	134	209	73
Largest cluster size (Arb. U.)	11867	3164	10254

The cluster visualizations using the hop frequency cut-off are shown in Figure E.7 are very promising. The crystalline morphology shows different crystalline grains very clearly - although the majority of the simulation is a single cluster (red), a large cluster with a different grain orientation is clearly visible in the morphology (blue). It is harder to distinguish the cluster distributions of the semi-crystalline and amorphous systems using the visualizations, although the crystals present in the semi-crystalline morphology are clearly resolvable from the amorphous matrix surrounding them. However, Table E.4 shows the first set of cluster properties where the semi-crystalline morphology is not intermediate between the crystalline and amorphous system, in terms of the number of total clusters and the largest cluster size. These clusters describe regions of the morphology that carriers are frequently hopping within. With this definition, hops within the regions are more common than those between clusters and so carriers are effectively trapped in this region - time is still progressing as they hop around, but their mean squared displacement is not significantly increasing. Therefore, a small number of large clusters is advantageous, whereas a large number of small clusters will strongly restrict charge transport properties.

E.4   Intra-cluster trapping

In our investigation, we record the location history of every carrier as it hops through the system. Using the carrier hopping history, we can construct network connectivity diagrams (Figure E.8) to observe the most frequently travelled paths for charges through the morphology. These network connectivity diagrams are constructed by identifying the frequency with which holes in the KMC simulation hop between pairs of chromophores. The centre-of-mass locations of the chromophores then form the
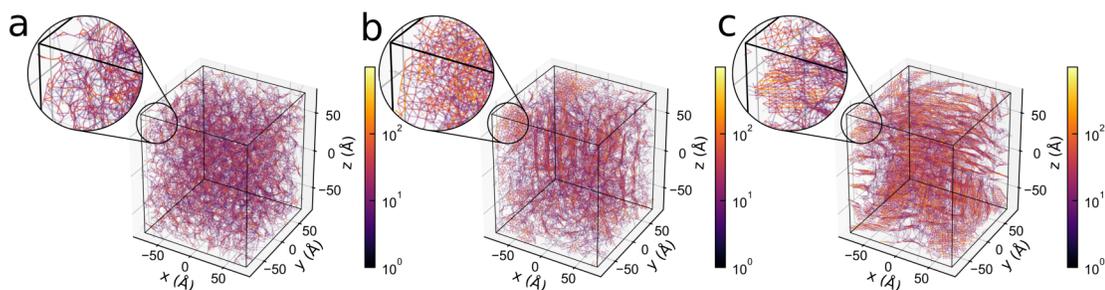
Figure E.8: The network diagrams for the (a) amorphous, (b) semi-crystalline, and (c) crystalline systems show carrier pathways between connected chromophores (insets: zoomed regions). Connections are colored based on a perceptually uniform, logarithmic heatmap, where brighter zones correspond to more frequently used pathways.

nodes of the network, and the shortest paths between each of the chromophore pairs become the edges. The "net hopping frequency" is calculated by subtracting the frequency of forward hops from backward hops and taking the absolute value. These values are normalized to the highest net hopping frequency in the system, and then assigned a color based on the logarithmic color map to highlight preferred carrier transport routes through the morphology.

The differences in structure between the three classes of morphology are clearly evident in Figure E.8. The amorphous network graph (Figure E.8a) shows that no crystallites have formed in the system. There are several high-traffic nodes spread homogeneously throughout the system, explaining the highly isotropic carrier trajectory presented in the main text. The crystalline network graph (Figure E.8c) shows the lamellar structure of the system, with nearly all chains aligned in layers moving left-to-right across the morphology. The most frequently used pathways are along chains, and there are many connections in the π-stacking direction between chains within the crystal. The semi-crystalline network graph (Figure E.8b) exhibits behaviour intermediate between the other two - crystallites with varying grain orientations are

clearly visible, within an amorphous matrix.

The insets in Figure E.8 show a zoomed region in the corner of the morphology, to highlight an area of 'cross-hatching' in the network, where carriers frequently loop around the same subset of chromophores (located at the vertices of the patterns), without increasing mean squared displacement (MSD) from their initial position. In the amorphous morphology (Figure E.8a), no loops are observed, and therefore every hop (no matter how slow) is contributing to the MSD, increasing mobility. In the semi-crystalline (Figure E.8b) and crystalline (Figure E.8c) systems, significant looping can be seen. Carriers in these regions are becoming 'trapped' by the loops - even though transport may be fast between the chromophores, it becomes more difficult for the carriers to leave the crystal along the slow transport routes due to the wealth of fast hops available within. For the crystalline case, the morphology is dominated by one large crystallite that extends across the full simulation volume. Therefore, carriers getting trapped inside this crystallite are still able to move long distances, and the penalty to the mobility from the trapping is lessened (reflected by a high mobility and a larger anisotropy in the main text). However, in the semi-crystalline case, the morphology is composed of multiple crystallites with various orientations, with loops present across all three dimensions. Trapping therefore has a more significant effect - carriers get stuck in the small loops and are unable to increase their mean squared displacement over time in a single direction, restricting the carrier mobility within the system. These conclusions are supported by the cluster maps presented in Figure E.8, as well as the cluster properties presented in Table 6.1; the crystalline and amorphous systems are dominated by a single, well-connected cluster of chromophores permitting a high mobility, whereas the semi-crystalline system is composed of many clusters with differing grain orientations. The visualizations of the network in Figure

E.8 serve to provide additional evidence as to why the clusters described in the main text form within these morphologies.

E.5   Polydisperse Simulations
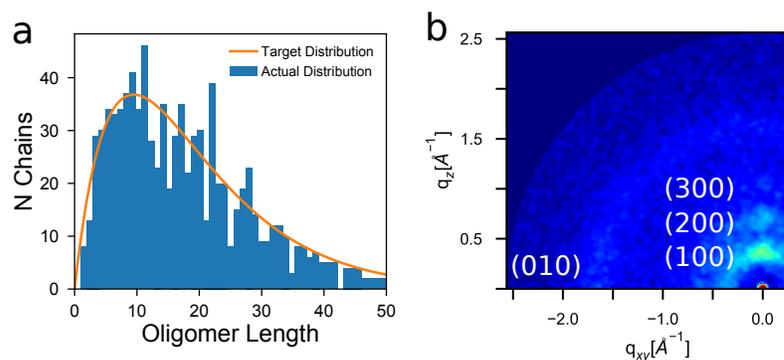
Figure E.9: (a) The scaled target distribution of chain lengths and the histogram of the actual chain lengths used in the polydisperse simulations. (b) The equilibrated polydisperse systems is able to produce periodic features not seen in systems where all the chains are longer.

Here we present how we generate polydisperse P3HT simulations. This can be broken into two steps: first creating a dictionary of P3HT oligomers of varying lengths from 1 to 50 monomers long. Second is using a distribution to determine the amount of each chain length to place into the simulation. To produce chains of arbitrary length, we use the open-source program mBuild in which a polymer can be easily created using monomer building blocks. We limit the chain length used in this study at 50 monomers long as to avoid unphysical interactions of chains feeling themselves across periodic boundaries. To generate the distribution of chain-lengths, we use the Schulz-Flory distribution which is a commonly used mathematical description for polymer lengths in the form[2]:

$$P_L = \alpha^2 D_P((1-\alpha)^{D_P - 1}), \tag{E.1}$$

in which P_L is the probability of seeing a chain of a given length, D_P is the degree of polymerization of a particular chain, and α is a tunable parameter which affects the shape of the distribution. The value for α used in this study was 0.1 and was chosen as this value produces polydispersities of ~ 1.8.

To create the actual distribution of chain lengths we utilize a simple Monte Carlo algorithm. In this algorithm we select a random chain length between 1 and 50 and a random number (x) associated with this chain length between 0 and 1. If x is less than the probability of seeing a chain of that length $P(L)$ we accept the chain otherwise the chain is rejected. In addition to this, to ensure that we have the same number of monomers as the other simulations (15,000 monomers) we keep track of the number of monomers which have been added to the simulation. When the number of monomers added to the simulation via the Monte Carlo algorithm is less than 50 monomers from 15,000 monomers, we terminate the Monte Carlo algorithm and add the remaining monomers to the simulation via one more single chain so that the total number of monomers is 15,000 monomers.

After the Monte Carlo algorithm is finished and a distribution of the chain lengths suggested, we calculate the polydispersity of the simulation with:

$$PDI = \frac{M_w}{M_n}, \tag{E.2}$$

in which M_w is the weight average molecule weight and M_n is the number average molecule weight. M_w and M_n can be calculated with:

$$M_w = \frac{\sum N_i M_i^2}{\sum N_i M_i} \tag{E.3}$$

$$M_n = \frac{\sum N_i M_i}{\sum N_i}, \tag{E.4}$$

in which N_i is the number of chains of that length and M_i is the molecular weight of that chain length. If the PDI of the stochastically generated distribution of chain lengths is below 1.8, we reject the distribution and regenerate the distribution until PDI $\geq$ 1.8. A comparison between the target distribution and the histogram of chain lengths is shown in Figure E.9a.

The distribution of chains presented in Figure E.9a is able to produce ordered morphologies with periodic features (Figure E.9b) along (100) and (010) signifying π- and alkyl-stacking. These features are seen in experimental and 15mer scattering patterns[4]. However, when simulating systems that contain only 50 membered chains, the system requires much longer to relax into these periodic structures.

BIBLIOGRAPHY

[1] Axel D. Becke. Density-Functional Thermochemistry. III. The Role of Exact Exchange. *The Journal of Chemical Physics*, 98(7):5648–5652, 1993.

[2] Paul J. Flory. Random Reorganization of Molecular Weight Distribution in Linear Condensation Polymers. *Journal of the American Chemical Society*, 64(9):2205–2212, 1942.

[3] R. Krishnan, J. S. Binkley, R. Seeger, and J. A. Pople. Self-Consistent Molecular Orbital Methods. XX. A Basis Set for Correlated Wave Functions. *The Journal of Chemical Physics*, 72(1):650–654, 1980.

[4] Evan Miller, Matthew Jones, Michael Henry, Paul Chery, Kyle Miller, and Eric Jankowski. Optimization and Validation of Efficient Models for Predicting Polythiophene Self-Assembly. *Polymers*, 10(12):1305, 2018.

APPENDIX F

MACHINE LEARNING PREDICTIONS OF ELECTRONIC COUPLINGS FOR CHARGE TRANSPORT CALCULATIONS OF P3HT—SUPPORTING INFORMATION[1]

F.1 Artificial Neural Net Construction
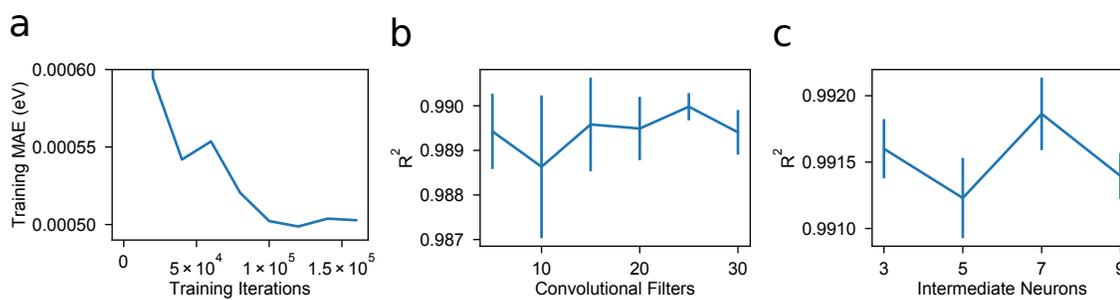
Figure F.1: (a) The ANN requires $\sim 1 \times 10^5$ training iterations (each iteration uses a batch of 10,000 training examples) to minimize the error. (b) 25 filters during the convolution produces slightly better results (averaged over 5 repeat trainings). (c) 7 neurons within the hidden layer produce the best results.

Here, we describe the topology and training requirements for the ANN. The ANN was trained using batches of 10,000 randomly chosen training examples without replacement. Using this batch style, the ANN required 1×10^5 training iterations to

---

[1]This appendix has been submitted the *AIChE* journal and is referenced as "Miller, E. D.; Jones, M.L.; Henry, M.H.; Stanfill, B.; Jankowski, E. Machine Learning Predictions of Electronic Couplings for Charge Transport Calculations of P3HT. *AIChE* **2019**—Submitted"

converge to a minimum error, at which point the training was terminated to prevent overfitting (Figure F.1). When data is supplied to the ANN, it is first passed into a max-pool convolutional layer. The purpose of the convolution layer and pooling is to help account for differences in the units between the features, e.g. distances in Å and rotations in radians. This convolutional layer consists of 25 filters with a window size of 1×1 and "valid" padding, which showed the most accurate answers without overfitting. Additionally, the max-pool layer also had a 1×1 window, but utilized the "same" padding so that the vector size was unchanged. The pooled layer is then subjected to a rectified linear unit activation function before being passed into the next layer.

The hidden layers of the ANN consisted of two hidden layers, containing 9 and 7 neurons respectively. The 9 was chosen due to that being the number of unique features used in training the ANN, whereas the 7 was chosen as it produced the most accurate training. On each neuron, an exponential linear unit activation function was applied, which allowed for greater stability than the rectified linear unit functions. To define the error in the system, Huber loss was used in combination with the Adam optimizer for gradient descent. We note that this topology may still be improved to produce more accurate predictions, however, one of the main strengths of the RF algorithm is that the out-of-box RF implementation produces nearly the same accuracies without the need for human optimization. Additionally, the RF implementation only requires a few minutes to train on a local workstation (Intel Haswell CPU), whereas the ANN requires more than an hour.

F.2 Error in $J_{i,j}$ Predictions

In this section we investigate the error that exists for the bonded and non-bonded chromophores. The RF has a low MAE, however, the trend in $J_{i,j}$ predictions does not perfectly match the actual values. This error in trend can be seen in two asymmetries: low $J_{i,j} < 0.4$ values are often skewed towards underpredictions whereas $J_{i,j} \sim 0.8$ are skewed towards overpredictions. This inaccuracy likely arises from $J_{i,j}$ being dominated by whether two chromophores are directly bonded to each other. As such, we separate the bonded and non-bonded chromophores and investigate the errors in $J_{i,j}$ predictions.

Separating the bonded and non-bonded chromophores results in much poorer predictions: $R^2 = 0.713$, MAE$= 0.018$ eV for the non-bonded and $R^2 = 0.701$, MAE$= 0.042$ eV for the bonded chromophores. The error in the predictions is shown in Figure F.2.
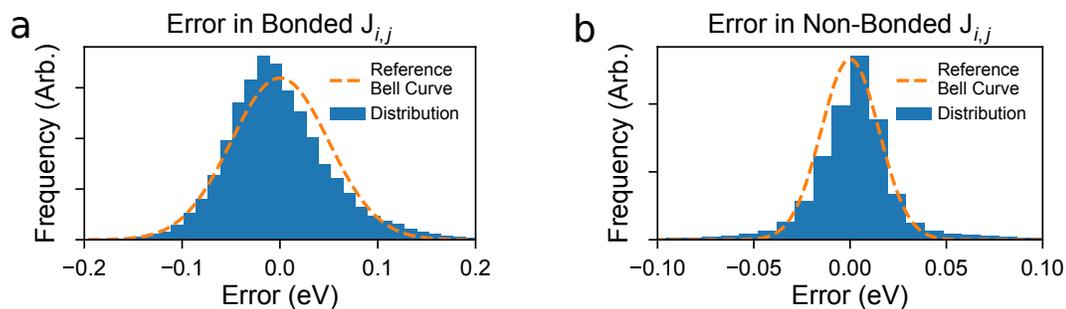
Figure F.2: The distribution of the $J_{i,j}$ error (blue) for the bonded chromophores shows that many values are slightly underpredicted, but that the distribution is slightly skewed from a symmetrical (orange) distribution with larger (< 0.1 eV) errors. Conversely, the distribution of non-bonded $J_{i,j}$ shows that many are slightly over predicted, however, the number of underpredicted outliers is not readily visible.

F.3 Selecting A Representative System

To accurately train the RF, we require 1×10^5 training samples. These training samples can be curated manually by moving and rotating two chromophores relative to each other or by drawing the chromophore pairs from an equilibrated system. In this study, we have drawn the training samples from three systems with varying degrees of order: a disordered system, a semi-crystalline system, and a more ordered system. This results in $\sim 700,000$ training samples. However, if we were to select a single system to use for training, does the degree of order in the system matter and which would be the best system to choose?
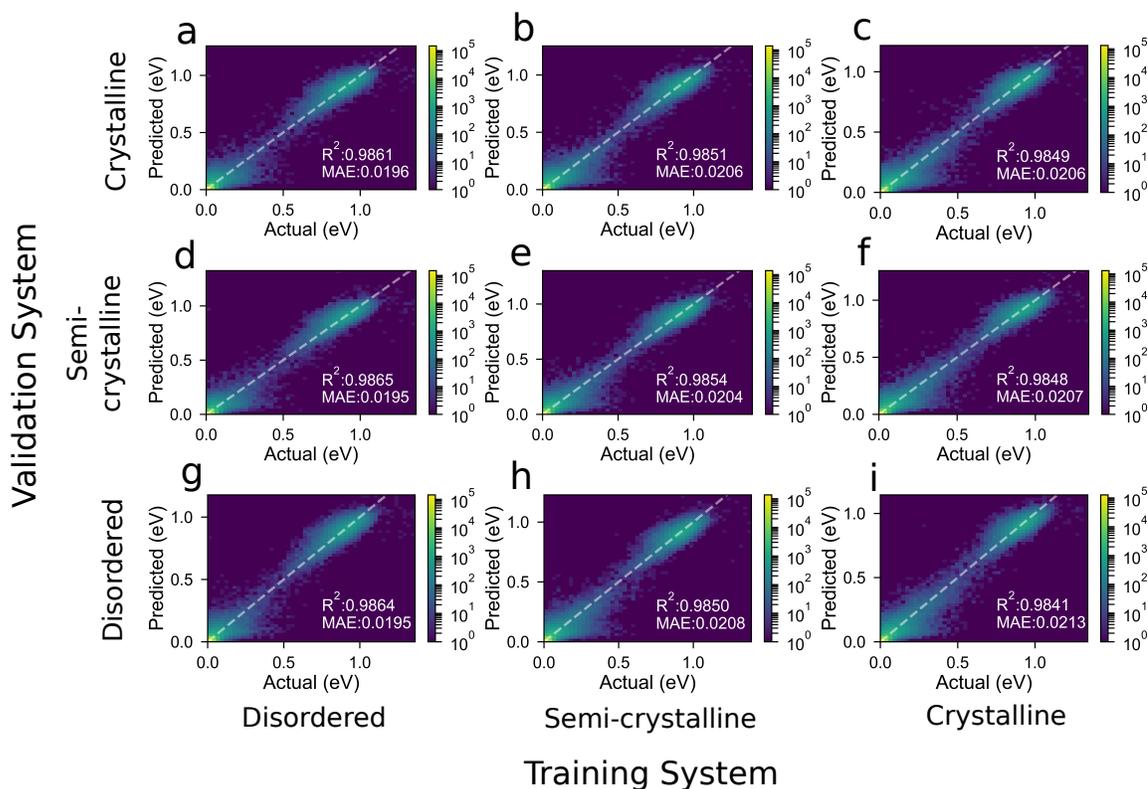
Figure F.3: The changes in accuracy when a system is trained on a single systems and validated on another system shows that training on a more ordered system produces slightly higher accuracies than when trained on a disordered system.

We have trained the RF with a single system of either disordered, semi-crystalline, or ordered and validated the RF on a decorrelated from of the same degrees of order. We present the RF predicted $J_{i,j}$ versus the ZINDO/S determined, actual $J_{i,j}$ in Figure F.3. In comparing these systems, we see that training with a more ordered system results in slightly better results. This is likely due to the ordered system having more examples of high value $J_{i,j}$, which results in the RF being able to more accurately predict these types of $J_{i,j}$. However, we note that a combination of both high and low $J_{i,j}$ is required for accurate training–a perfect crystal will not generalize well and accurately predict transfer integrals in disordered films where COM separations and angle distributions are more broad.