2020

TOWARDS THE RATIONAL DESIGN OF ORGANIC SEMICONDUCTORS THROUGH COMPUTATIONAL APPROACHES

Qianxiang Ai
University of Kentucky, alex.qai@outlook.com
Author ORCID Identifier:
https://orcid.org/0000-0002-5487-2539
Digital Object Identifier: https://doi.org/10.13023/etd.2020.311

STUDENT AGREEMENT:

I represent that my thesis or dissertation and abstract are my original work. Proper attribution has been given to all outside sources. I understand that I am solely responsible for obtaining any needed copyright permissions. I have obtained needed written permission statement(s) from the owner(s) of each third-party copyrighted matter to be included in my work, allowing electronic distribution (if such use is not permitted by the fair use doctrine) which will be submitted to UKnowledge as Additional File.

I hereby grant to The University of Kentucky and its agents the irrevocable, non-exclusive, and royalty-free license to archive and make accessible my work in whole or in part in all forms of media, now or hereafter known. I agree that the document mentioned above may be made available immediately for worldwide access unless an embargo applies.

I retain all other ownership rights to the copyright of my work. I also retain the right to use in future works (such as articles or books) all or part of my work. I understand that I am free to register the copyright to my work.

REVIEW, APPROVAL AND ACCEPTANCE

The document mentioned above has been reviewed and accepted by the student's advisor, on behalf of the advisory committee, and by the Director of Graduate Studies (DGS), on behalf of the program; we verify that this is the final, approved version of the student's thesis including all changes required by the advisory committee. The undersigned agree to abide by the statements above.

Qianxiang Ai, Student

Dr. Chad Risko, Major Professor

Dr. Yinan Wei, Director of Graduate Studies

TOWARDS THE RATIONAL DESIGN OF ORGANIC SEMICONDUCTORS THROUGH COMPUTATIONAL APPROACHES

DISSERTATION

A dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the College of Arts and Sciences at the University of Kentucky

By
Qianxiang Ai

Lexington, Kentucky

Director: Dr. Chad Risko, Professor of Chemistry

Lexington, Kentucky

2020

https://orcid.org/0000-0002-5487-2539

ABSTRACT OF DISSERTATION

TOWARDS THE RATIONAL DESIGN OF
ORGANIC SEMICONDUCTORS THROUGH
COMPUTATIONAL APPROACHES

Though organic semiconductors have illustrated potential as industry-relevant materials for electronics applications, there are few guidelines that can take one from molecular design to functional materials. This limitation is, in part, due to incomplete understanding as to how the atomic-scale construction of the π-conjugated molecules that comprise the organic semiconductors determines the nature and strength of both the noncovalent intramolecular interactions that govern molecular conformation and noncovalent intermolecular interactions that regulate the energetic preference for solid-state packing. Hence, there remain several fundamental questions that need to be resolved in order to design organic semiconductors from *a priori* knowledge, including: What is the relevance of the relatively weak noncovalent intramolecular interactions on determining molecular structure, are current hypotheses put forward as to important interactions valid, and how does chemical substitution as various positions along the π-conjugated backbone impact these interactions? How do the intermolecular noncovalent interactions regulate solid-state packing, are there features of the molecular structure – *e.g.* the π-conjugated backbone, heteroatoms, or pendent alkyl chains – that play a more important role? What connections can be made between the structures/properties of the π-conjugated molecules and the resulting organic semiconductors?

In this dissertation, Chapter 1 provides an introductory discussion of these questions and a brief review of previous studies. Chapter 2 details the computational approaches that were implemented throughout the course of the thesis work. Chapter 3 describes the investigation of a series of pyrene-acene molecules to illustrate the importance of choosing the right molecular structure in π-conjugated

chromophores. In Chapter 4, S…F noncovalent intramolecular interactions are systematically investigated in two separate cases to highlight the varied impact that these interactions can have on molecular and solid-state packing structures. Chapter 5 describes the investigation of an oscillatory crystal packing structure observed for a series of oligothiophenes that follow the odd-even carbon-atom counts of the pendant alkyl chains. In Chapter 6, the polymorphism of functionalized pentacene molecules is studied to reveal how seemingly simple atomic substitutions can drastically alter solid-state packing. To systematically address the aforementioned fundamental questions, Chapter 7 describes the construction and application of a database of crystalline molecular organic semiconductors. Finally, perspectives regarding future research are provided in Chapter 8.

KEYWORDS: Organic Semiconductor, Polymorphism, Data-driven Approaches, Organic Field Effect Transistor, Density Functional Theory

Qianxiang Ai

2020/07/15

TOWARDS THE RATIONAL DESIGN OF
ORGANIC SEMICONDUCTORS THROUGH
COMPUTATIONAL APPROACHES

By

Qianxiang Ai

Chad Risko

Director of Dissertation

Yinan Wei

Director of Graduate Studies

2020/07/15

Date

ACKNOWLEDGEMENTS

It has been almost five years since I started my graduate work, and I have had the fortunate opportunity to learn from, work with, and be inspired by a great number of people throughout this time. I would first like to thank my research adviser, (newly tenured) Professor Chad Risko, who guided me into the fascinating realm of computational materials chemistry, and has taught me so much, not just about research but, more importantly, about how to be a humble, perpetual learner. Thus I would like to express my sincere gratitude and deep appreciation to Professor Risko for being a fantastic mentor, as well as a stimulating group leader.

I am indebted to Professor John Anthony and his group, the "go-to" guys when we need expertise and, more often, "chemical intuition". A significant portion of this dissertation could not have been completed without access to the hundreds of molecules Professor Anthony and his group made over the last decades. I also owe my profound gratitude to Professor Sean Parkin, for the crystal structure dataset we relied on, and his efforts in (re)teaching me about crystallography even I kept forgetting what a glide plane was. I want to thank my collaborators, Professor Lynn Loo and Jeni Sorli, who indulged me by letting me, who is barely allowed to cook, into the clean room to learn about device fabrication. I am also thankful for having collaborators like Professor Oana Jurchescu and her group, as they keep making high-performance devices that exceed our prediction.

I would like to thank the past and present members of Risko group, along with other LAB 2 members, who made me feel welcome and helped me during my studies. Specifically, I would like to thank: Dr. Karol Jarolimek for helping me go through my "noob" phase in research, Dr. Sean Ryno for "bootstrapping" my coding career, Dr. Karl Thorley for all these inspiring and wandering conversations, Dr. Ling-yi Huang for our struggles with *BerkeleyGW*, and Dr. Ruben Sarabia-Riquelme for the torture we enjoyed in solid-state physics class. I would also like

to acknowledge the *OCELOT* developer team, including Vinayak Bhat, Andrew Smith, Stephen Goodlett, Parker Sornberger, for the fun time we spent on reading documentations and writing bugs…

I am also grateful for the support from faculty members of University of Kentucky. In particular, I am appreciative for the one-semester stay in Professor Dong-sheng Yang's lab, where the members helped me to get accustomed to a new country. While I can vividly remember the "ordeal" of taking multiple solid-state related classes in one semester, courses given by Professor Beth Guiton, Professor Lance De Long and Professor Ribhu Kaul indeed prepared me for studying solid-state systems. I would also like to thank the technical support from Vikram Gazula and other members of Center for Computational Sciences, who equipped me to deal with computational demanding problems.

Portions of the work presented in this dissertation were supported by the National Science Foundation (NSF) Designing Materials to Revolutionize and Engineer our Future (DMREF) program in the Division of Materials Research (award number DMR-1627428). Supercomputing resources on the Lipscomb High Performance Computing Cluster were provided by the University of Kentucky Information Technology Department and Center for Computational Sciences.

Finally, to my wife, Pan, and my mother, Jiangyu, who deserve more praise and adoration than I can effectively convey in words. I am eternally grateful for your unwavering support and unfailing patience.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

While the semiconductor industry has been dominated by inorganic materials since the invention of the transistor in the late 1940s, organic semiconductors have been making a strong push into commercial applications, having first been prototyped in organic photovoltaic cells, organic light-emitting diodes, and organic thin-film transistors in the 1980s.[1,2] Thanks to progress in material synthesis and processing, organic semiconducting materials are now not only a research focus but have also become a commercial reality, and have taken a significant market share in display technologies since the early 2000s.[3] Such rapid development was driven by the fact that organic semiconductors can be fabricated into thin, light-weight and mechanically flexible devices via low-temperature, fully solution processes. However, even after decades of effort, relatively limited performance and low stability are generally observed for devices containing organic semiconductors. For example, one of the record charge-carrier mobilities achieved by an organic thin film transistor is 43 $cm^2/(V \cdot s)$, which made use of a meta-stable crystalline phase of C8-BTBT,[4] whereas the charge-carrier mobility of thin film, flexible silicon-based transistors can reach 200 $cm^2/(V \cdot s)$.[5] Thus, explorations of novel organic materials and understanding their intrinsic chemistry are vital to development of next-generation organic semiconductors with performance comparable with inorganic materials.

Organic semiconductors have also attracted interest for the intriguing physics behind their electronic and optical properties. The discovery of the first conductive organic polymer, halogen doped polyacetylene in 1977,[6] was honored with the Nobel Prize in 2000 due in part to the interest that it spurred from a fundamental perspective. Such a "synthetic metal" inspired wide interest in the

natures of excitons, charge-carrier transport (*e.g.* polarons, bipolarons, solitons), and the general electronic structure of organic materials. At first glance, the electronic structure of an organic solid should not be complicated: The active part of the electronic structure consists of coupled π orbitals that lead to delocalized electrons. However, several significant features make organic semiconductors different from their inorganic counterparts. In organic semiconductors, charge carriers almost entirely come from doping, as their large fundamental gap (usually around 2$\sim$3 eV) prevents any considerable thermal excited carriers. Organics also tend to have relatively small dielectric constants and limited intermolecular electronic couplings such that electronic correlations cannot be neglected.[7] As organic solids are held together by weak van der Waals forces, their soft nature makes electronic polarization and molecular and solid-state vibrations crucial to electron dynamics, thus pronounced electron-phonon couplings are expected. In addition to these complexities that preclude in-depth descriptions, let alone accurate predictions, of optoelectronic properties, organic chemists, while are able to realize various molecular structures, cannot predict the resulting macroscopic solid-state structure, which is formally known as, for crystalline materials, the crystal structure prediction (CSP) problem.[8,9] Such a lack of predictive capabilities, along with time-consuming device fabrication that is sensitive to a wide range of processing conditions, are major obstacles towards rational design of high-performance organic semiconducting materials. To provide context for later chapters, the rest of this chapter is organized as follows: First, a summary of noncovalent interactions in organic semiconductors is provided. Molecular structure and packing of organic semiconductors are then discussed. Finally, we introduce data infrastructures for materials science studies, followed by a synopsis of this thesis.

1.1 Noncovalent Interaction

Noncovalent interactions are interactions among atoms or molecular fragments that do not involve intimate sharing of electrons, and are usually an order

of magnitude weaker than covalent bonds.[10] For organic molecules, noncovalent interactions play crucial roles in determining their ground state conformations (intramolecular interactions), as well as how molecules self-assemble into extended structures (intermolecular interactions). This can be exemplified by the complex, versatile three-dimensional structure of a protein, which is developed by folding covalently connected amino acids through noncovalent interactions, including directed hydrogen bonds, hydrophilic/hydrophobic interactions, and van der Waals forces.[11,12] Just like a protein, the properties of an organic solid not only depends on the chemical composition, but are also subject to conformational change of molecules and their assemblies, which highlights the significance of noncovalent interactions.

While the history of noncovalent interactions can be traced back to research on describing real gases by van der Waals in 1870s,[13] its classification still remains non-trivial and often come with ambiguities. Excluding metallic bonds, from a perturbation theory perspective, noncovalent interactions can be classified into four categories: electrostatic, induction, dispersion, and exchange-repulsion.[14] Electrostatic interactions come from permanent multipoles and can be calculated via multipole expansion among well-separated systems.[15] Induction interactions result from the polarization field generated by the electronic relaxation of molecules, whereas instantaneous fluctuations of charge densities give rise to dispersion interactions. Exchange-repulsion originates from the effective statistical interaction (Pauli's principle),[16] which appears as repulsive forces between electrons. Considering that most organic semiconducting materials consist of neutral molecules/polymers with low dielectric constant (less polar), in these systems, exchange-repulsion and dispersion tend to be more significant than other noncovalent interactions.

One of the difficulties of investigating noncovalent interactions is that there are only few descriptors can be confidently determined experimentally. While for

molecular complexes rotational spectroscopy can be used to detect very weak interactions,[17,18] such technics are not readily available for the solid-state. Instead, (changes in) vibrational and solid-state nuclear magnetic resonance (NMR) spectroscopies are used.[19–21] In addition, the atomic positions of crystalline materials, solved by diffraction experiments, allow the existence of noncovalent interactions to be inferred.[22–24] For theoreticians, describing noncovalent interactions often implies a trade-off between accuracy and applicability. While chemical accuracy can be achieved with coupled-cluster approximations (usually at the CCSD(T) level), these calculations are challenged by scaling (N^7 for floating-point operations in standard CCSD(T)) and memory requirements. Hence, researchers turn to lower-level wavefunction methods or dispersion-corrected density functional theory (DFT) methods. A brief description of the latter, which is also used in later chapters, can be found in Chapter 2.

1.2 Molecular Structure and Solid-state Packing

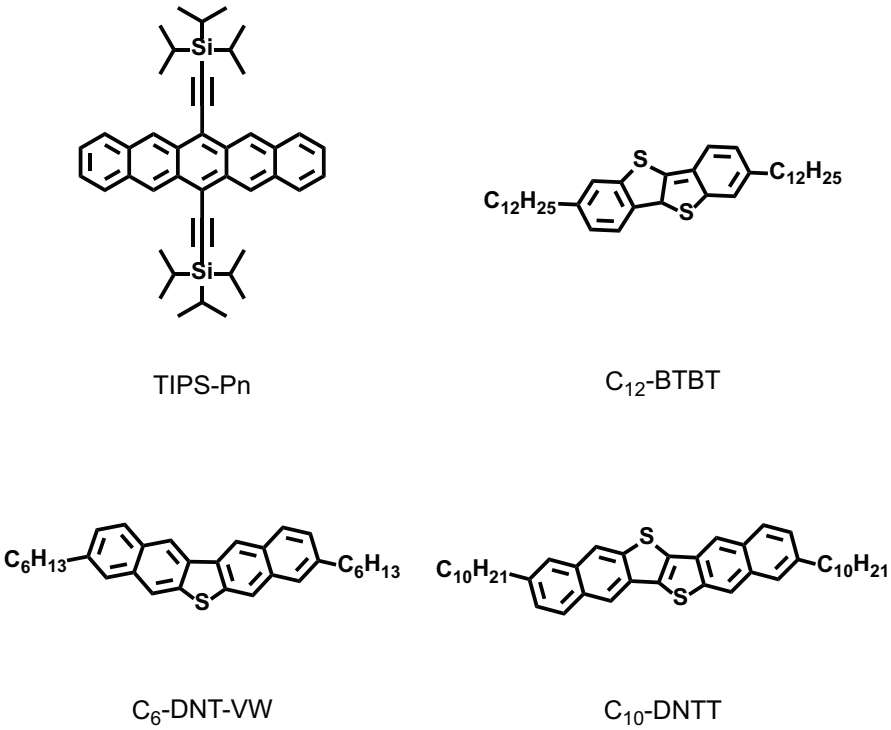


Figure 1.1: Molecular structures of high-mobility organic semiconductors.

In addition to amorphous solids (derived for both molecules and polymers), organic molecular crystals constitute an important class of organic semiconductors. In fact, research on optoelectronic properties of molecular crystals can be traced back to the beginning of 20th century when the photoconductivity of anthracene was studied.[25] Developments over the last few decades have led to a boom of such materials, examples of which are showcased in Fig. 1.1, comprised of molecular structures that lead to high-mobility crystalline organic semiconductors.[26–29] These molecules can be described under the same molecular design paradigm as 6,13-Bis(triisopropylsilylethynyl)pentacene (TIPS-Pn): A π-conjugated core with alkyl-based side chains appended along the periphery. The former, often referred to as the chromophore, contributes π orbitals to the extended solid-state system that are used to form the valence and conduction bands. The side chains are considered to be electronically inert, as their electrons tend to reside in deep energy levels. For single molecules, we are mainly interested in the chromophore, which portends the nature of the electronic, redox, and optical properties. However, the electronic structure of an organic crystal also depends on electronic interactions between molecules, which is subject to molecular packing where side groups also play a role. Chromophore-only molecules, due to high structural rigidity, tend to suffer from poor processability, *e.g.* pentacene is only sparingly soluble in organic solvents.[30] Thus, the purpose of appending side groups is twofold: 1) to improve chromophore processability, such as solubility and air stability, and 2) to achieve molecular packing whose underlying electronic structure is preferred for semiconducting applications.

While advances in organic chemistry allow the synthesis a wide range of chromophores, as well as the ability to append a wide range of side groups, control over molecular packing has never been precise, as it is related to the challenging CSP problem. With enormous progress from CSP community in the last decade, solutions to the CSP challenge appear reachable in near future, as reflected by the promising results of the latest CSP blind test.[31] However, organic semiconductors

like TIPS-pentacene represent a particular class of CSP where intermolecular interactions are dominated by weak dispersion forces and the number of rotatable bonds is large. In addition, crystal configuration can be very sensitive to slight changes in molecular structure, and crystallization method in the case of polymorphism, as shown in the cases of Chapter 5 and Chapter 6. Such lack of predictive capabilities provides unique challenges in the design of organic semiconducting materials.

1.3 Data Infrastructure for Materials Science

The history of data science in organic materials took off in 1960s as the construction of the Cambridge Structural Database started, which has evolved into a popular database for crystallography and pharmaceutical chemistry.[32] Powered by ever-improving computational and data storage resources, along with the wide acceptance of the concept of open science,[33,34] data driven approaches have become increasing popular in a variety of scientific communities, with materials science being no exception. Big-data-driven approaches are proposed as the fourth paradigm of scientific discovery from which previously unknown correlations and dependencies that are invisible in small data set can be revealed.[35,36]

In materials science, in addition to data generation and collection, implementing data-driven approaches requires a set of resources, from physical data storage facilities to software platforms. These are known as data infrastructures, which, ideally, should be able to provide:

- Schemaless data structure: As materials data are collected from various sources, including both experimental and computational results, schemaless data structure allows registering data with arbitrary fields. This is typically implemented with non-relational databases.[37]

- Data curation and standardization: To fully exploit the potential of a large dataset, it is imperative to integrate and consolidate separate information to acquire holistic knowledge. By data curation, *a priori* knowledge of relations between individual data entries can be captured. Data standardization and data integrity assessment are also necessary to ensure high data quality that is crucial for obtaining correct results from data-driven approaches.

- High accessibility and searchability: While data accessibility has been greatly advanced in the last few decades thanks to the open access movement,[38] searchability, that is, the ability to find desired information with a short query, still has room for improvement. This is primarily due to the inconsistent nomenclature gifted by the interdisciplinary nature of materials science.

Recent efforts to develop data infrastructures that meet these demands include the Materials Project,[39] AFLOW,[40] and NOMAD,[41] where results from high-throughput computational workflows are curated and shared with the community. While these studies have been successful in providing access to big data (mainly for inorganic materials) and facilitating data analytics, significant challenges remain, such as integration of experimental data, standardization and searchability.[42]

1.4 Synopsis

This thesis is dedicated to expanding the understanding of organic semiconductors, across the scales from molecules to materials, via computational methods, which are reviewed in Chapter 2. The overarching hypothesis of this thesis is that knowledge derived from computational materials chemistry approaches, generated from a combination of molecular and solid-state electronic structure methods, when integrated with experimental data, can be used both to corroborate and refine existing materials design rules, and obtain new predictive capacities. Within

this context, a series of molecular and crystal systems are examined in the following chapters to provide insights to varied aspects of organic semiconductor design, including the construction of the molecular chromophore, the roles of intramolecular and intermolecular noncovalent interactions on molecular and solid-state packing structures, and how data-centered approaches can advance materials discoveries.

In Chapter 3, we explore a series of pyrene-fused acene molecules to understand how the construction of the π-conjugated chromophore can impact the electronic and optical properties of the molecular system. Here, we reveal how a seemingly large π-conjugated system can consist of almost isolated parts, a feature due in part to the symmetry of the constituents that comprise the system. This feature, while leading to distinct optical response at the molecular level, is also shown to have non-trivial implications for polymer/graphene nano ribbon design.

Intramolecular S…F interactions are believed to promote rigid, planar conformations of extended conjugated systems, and have been introduced into thiophene-based polymers showing improved performance.[43–45] In Chapter 4, such interactions in two thiophene-containing oligomer systems are examined to show the impact of weak, noncovalent interactions on chromophore structure. It is revealed that the stabilizing effect brought by S…F short contact can come from different origins, and the preference towards a planar conformation is subject to the structure of conjugated backbone.

In addition to molecular chromophore, the alkyl side groups, while being electronically inert, can be crucial in determining molecular packing in crystal structures In Chapter 5, an intriguing odd-even oscillation in the crystal structures of functionalized oligothiophenes is studied, which is related to oscillatory noncovalent interactions from incremental changes among the side groups. Discussions on the generality and origin of such oscillation indicate that it could also be observed in other homologous compounds.

The relation between side groups and crystal structure is further discussed in Chapter 6. Here, we focus on an extreme case in which polymorphism is induced by single-atom substitution on side groups. Solid-state structural studies on the polymorphism of 6,13-bis(triisopropylgermanylethynyl) pentacene (TIPGe-Pn) crystals illustrate the impact of lattice vibration on electronic structures through thermal expansion, and how phonons can improve the relative stability of a configuration that is otherwise unfavored at finite temperature.

As an attempt to realize the idea of data-driven material discovery, in Chapter 7, we describe the construction of *OCELOT* (Organic Crystals in Electronic and Light-Oriented Technologies), a digital archive for organic semiconducting materials research. Software tools, as well as curation methods, are developed for research data generation and collection with a focus on crystalline solids. Application of this platform is exemplified by two case studies regarding molecular packing and material screening, respectively.

In summary, these studies reflect our efforts to address various aspects of organic semiconductor design. As will be shown in the following chapters, computational materials chemistry, when properly integrated with experimental insights, is able to not only refine previous understandings, but also provide irreplaceable and interpretable information necessary for prediction. This perception, along with future considerations, will be presented in Chapter 8 to conclude this thesis.

CHAPTER 2

METHODS

In this chapter, the theoretical background and technical models employed in this study are briefly reviewed. This chapter is organized as follows: In Section 2.1 basic concepts of crystal electronic structure are reviewed. The computational workhorse of this study, density functional theory (DFT), is discussed in Section 2.2 along with important approximations and corrections. Relevant charge transport models are introduced in Section 2.3. Atomic units are used throughout this chapter unless otherwise specified. The formalisms presented herein are primarily taken from the following texts: Ashcroft[46], Born[47], Christman[48], Shuai[49], Brandenburg[50] and van Leeuwen[51].

2.1 Electronic Structure of Crystals

2.1.1 Many Body Hamiltonian

The electronic properties of a nucleus-electron system can be depicted by the time independent Schördinger Equation:

$$\hat{H}\Psi(\{\boldsymbol{R}_I\},\{\boldsymbol{r}_i\}) = E_n\Psi(\{\boldsymbol{R}_I\},\{\boldsymbol{r}_i\}) \tag{2.1}$$

where $\boldsymbol{R}_I$ and $\boldsymbol{r}_i$ represent the coordinates of the Ith nucleus and the ith electron, respectively. For now we assume the electrons are spinless. E_n is the energy of the nth eigenstate, and Ψ is the many body wavefunction as an eigenstate of the many body Hamiltonian $\hat{H}$, which is defined as:

$$\hat{H} = \sum_{I=1}^{M}\frac{P_I^2}{2M_I} + \sum_{i=1}^{N}\frac{p_i^2}{2} + \hat{V}(\{\boldsymbol{R}_I\},\{\boldsymbol{r}_i\}). \tag{2.2}$$

Here, $P_I = -i\nabla_R$ and $p_i = -i\nabla_r$ are momentum operators of the Ith nucleus with mass M_I and ith electron, respectively. While the first two terms of Eq. (2.2) are the kinetic energies of the nuclei ($\hat{T}_n$) and electrons ($\hat{T}_e$), the third term comes from the interactions among the $M + N$ particles:

$$\hat{V}(\{\boldsymbol{R}_I\}, \{\boldsymbol{r}_i\}) = \sum_{I \neq J} \frac{Z_I Z_J}{2|\boldsymbol{R}_I - \boldsymbol{R}_J|} + \sum_{i \neq j} \frac{1}{2|\boldsymbol{r}_i - \boldsymbol{r}_j|} - \sum_{i,I} \frac{Z_I}{|\boldsymbol{r}_i - \boldsymbol{R}_I|} \tag{2.3}$$

where Z_I and Z_J represent the nuclear charge of Ith nucleus. The right hand side of Eq. (2.3) consists of, from left to right, the Coulomb interactions of nucleus-nucleus ($\hat{V}_{nn}(\{\boldsymbol{R}_I\})$), electron-electron($\hat{V}_{ee}(\{\boldsymbol{r}_i\})$), and nucleus-electron($\hat{V}_{ne}(\{\boldsymbol{R}_I\}, \{\boldsymbol{r}_i\})$). Consequently, Eq. (2.2) can be written as:

$$\hat{H} = \hat{T}_n + \hat{T}_e + \hat{V}_{nn} + \hat{V}_{ee} + \hat{V}_{ne}. \tag{2.4}$$

While in general there are many E_n of Eq. (2.1), for most situations the most important state is the state with the smallest eigenvalue E_0, *i.e.* the ground state. This is the *quantum many body problem* which is a centerpiece of quantum chemistry.

Considering $\frac{M_H}{m_e} \approx 1840$ where M_H is the mass of the lightest nucleus, proton, and m_e is that of an electron, it is reasonable to assume that the electrons react to a perturbation much faster than nuclei. That is, the motion of electrons may be, approximately, *decoupled* from the motion of nuclei. This motivates us to define the electronic Hamiltonian $\hat{H}_e$ as

$$\hat{H}_e = \hat{T}_e + \hat{V}_{ne} + \hat{V}_{ee} \tag{2.5}$$

where the dynamics of nuclei is absent, but $\hat{H}_e$ still depends on $\{\boldsymbol{R}_I\}$ parametrically. Once the nuclei are clamped to $\boldsymbol{R}_C$ the eigenfunctions of $\hat{H}_e$ can be solved from:

$$\hat{H}_e^{\{\boldsymbol{R}_C\}} \Phi_\nu^{\{\boldsymbol{R}_C\}}(\{\boldsymbol{r}_i\}) = E_{e,\nu} \Phi_\nu^{\{\boldsymbol{R}_C\}}(\{\boldsymbol{r}_i\}) \tag{2.6}$$

where Φ_ν is the νth eigenfunction of $\hat{H}_e$, $E_{e,\nu}$ is the eigenvalue, superscripts represent the parametric dependence on $\{\boldsymbol{R}_C\}$. This is the *electronic structure problem* in which electronic wavefunction Φ are solved at fixed $\{\boldsymbol{R}_C\}$. Let $\Psi(\{\boldsymbol{R}_I\}, (\{\boldsymbol{r}_i\})$ be an exact, ground state (this is not necessary) solution to Eq. (2.1), at an arbitrary $\{\boldsymbol{R}_I\}$ it can be expanded as:

$$\Psi(\{\boldsymbol{R}_I\}, \{\boldsymbol{r}_i\}) = \sum_\nu \Lambda_\nu(\{\boldsymbol{R}_I\})\Phi_\nu^{\{\boldsymbol{R}_C\}}(\{\boldsymbol{r}_i\}) \tag{2.7}$$

where $\Lambda_\nu(\{\boldsymbol{R}_I\})$ is the expansion coefficient for $\Phi_\nu^{\{\boldsymbol{R}_C\}}(\{\boldsymbol{r}_i\})$. Eq. (2.7) comes from the fact that the eigenfunctions $\Phi_\nu^{\{\boldsymbol{R}_C\}}(\{\boldsymbol{r}_i\})$ form a complete basis of the Hilbert space of $\Phi(\{\boldsymbol{r}_i\})$. This property leads to an alternative expansion of $\Psi(\{\boldsymbol{R}_I\}, (\{\boldsymbol{r}_i\})$:

$$\Psi(\{\boldsymbol{R}_I\}, \{\boldsymbol{r}_i\}) = \sum_\nu \lambda_\nu(\{\boldsymbol{R}_I\})\Phi_\nu(\{\boldsymbol{R}_I\}, \{\boldsymbol{r}_i\}) \tag{2.8}$$

which is the "dynamic" version of Eq. (2.7) since Φ_ν explicitly depends on $\{\boldsymbol{R}_I\}$. These expansions allow us to separate the motion of nuclei (λ_ν) from the electronic wavefunctions Φ. Use Eq. (2.8) in Eq. (2.4) along with Eq. (2.5), we have:

$$\hat{H}\Psi = \hat{H}_e\Psi + \hat{T}_n\Psi + \hat{V}_{nn}\Psi \tag{2.9}$$

$$\hat{H}_e\Psi = \sum_\nu \lambda_\nu E_{e,\nu}\Phi_\nu \tag{2.10}$$

$$\hat{V}_{nn}\Psi = \sum_\nu \lambda_\nu \hat{V}_{nn}\Phi_\nu \tag{2.11}$$

but for $\hat{T}_n\Psi$ the expansion is more complicated as $\hat{T}_n$ does not commute with λ_ν:

$$\begin{aligned}\hat{T}_n\Psi &= \sum_{I=1}^{M}\sum_\nu -\frac{1}{2M_I}\nabla_R^2\lambda_\nu\Phi_\nu \\ &= \sum_{I=1}^{M}\sum_\nu -\frac{1}{2M_I}[\lambda_\nu(\nabla_R^2\Phi_\nu) + 2(\nabla_R\lambda_\nu)(\nabla_R\Phi_\nu) + (\nabla_R^2\lambda_\nu)\Phi_\nu]\end{aligned} \tag{2.12}$$

where the chain rule has been invoked. Now we would like to calculate λ_ν so we

just need to solve Φ from Eq. (2.6) to get Ψ. Multiply both sides of Eq. (2.9) with Φ_μ^* (an eigenfunction of $\hat{H}_e$) then integrate over $\boldsymbol{r}_i$, for the left side we have:

$$\langle\Phi_\mu|\,\hat{H}\,|\Psi\rangle = E_0\,\langle\Phi_\mu|\Psi\rangle = E_0\lambda_\mu\,\langle\Phi_\mu|\Phi_\mu\rangle = E_0\lambda_\mu \tag{2.13}$$

and for the right side we obtain:

$$\langle\Phi_\mu|\,\hat{H}_e\,|\Psi\rangle + \langle\Phi_\mu|\,\hat{T}_n\,|\Psi\rangle + \langle\Phi_\mu|\,\hat{V}_{nn}\,|\Psi\rangle = (E_{e,\mu} + \hat{V}_{nn})\lambda_\mu + \langle\Phi_\mu|\,\hat{T}_n\,|\Psi\rangle \tag{2.14}$$

where the third term should be evaluated based on Eq. (2.12):

$$\begin{aligned}\langle\Phi_\mu|\,\hat{T}_n\,|\Psi\rangle = \sum_{I=1}^{M}\sum_{\nu} -\frac{1}{2M_I}[\lambda_\nu\,\langle\Phi_\mu|\,\nabla_R^2\,|\Phi_\nu\rangle \\ + 2(\nabla_R\lambda_\nu)\,\langle\Phi_\mu|\,\nabla_R\,|\Phi_\nu\rangle + \nabla_R^2\lambda_\nu\,\langle\Phi_\mu|\Phi_\nu\rangle].\end{aligned} \tag{2.15}$$

Combining Eq. (2.13), Eq. (2.14) and Eq. (2.15), for each eigenstate Φ_μ of $\hat{H}_e$ we have:

$$\begin{aligned}E_0\lambda_\mu =&(E_{e,\mu} + V_{nn} + \hat{T}_n)\lambda_\mu \\ &+ \sum_{I=1}^{M}\sum_{\nu} -\frac{1}{2M_I}[\lambda_\nu\,\langle\Phi_\mu|\,\nabla_R^2\,|\Phi_\nu\rangle + 2(\nabla_R\lambda_\nu)\,\langle\Phi_\mu|\,\nabla_R\,|\Phi_\nu\rangle]\end{aligned} \tag{2.16}$$

where the terms in rectangle brackets are the electron-phonon coupling terms which correlate the motion of nuclei (∇_R) with electronic states Φ_ν and Φ_μ. We first look at the electron-phonon coupling terms when $\mu = \nu$ (diagonal), the $\langle\Phi_\mu|\,\nabla_R\,|\Phi_\nu\rangle$ terms vanish (in the absence of a magnetic field) since

$$\langle\Phi_\mu|\,\nabla_R\,|\Phi_\nu\rangle = \frac{1}{2}\nabla_R\,\langle\Phi_\mu|\Phi_\nu\rangle \tag{2.17}$$

where the right side is zero. For the $-\frac{1}{2M_I}\,\langle\Phi_\mu|\,\nabla_R^2\,|\Phi_\nu\rangle$ terms if we assume $\Phi_\nu = \Phi_\nu(\{\boldsymbol{R}_I - \boldsymbol{r}_i\})$ we have

$$-\frac{1}{2M_I}\,\langle\Phi_\mu|\,\nabla_R^2\,|\Phi_\nu\rangle = -\frac{1}{M_I}\frac{1}{2}\,\langle\Phi_\mu|\,\nabla_r^2\,|\Phi_\nu\rangle \tag{2.18}$$

which means these terms are at most in the order of 10^{-4} times the kinetic energy of an electron.

Now we introduce the Born-Oppenheimer approximation which states that the electrons will always be in an eigenstate of $\hat{H}_e$ and nuclear motions would not induce transition between these states (adiabatic), which effectively zeros the electron-phonon coupling terms in Eq. (2.16) when $\mu \neq \nu$ (off-diagonal). With this approximation we can rewrite Eq. (2.16) as:

$$(\hat{T}_n + \hat{V}_{nn} + E_{e,\mu})\lambda_\mu = E_0\lambda_\mu \tag{2.19}$$

where the nuclei motions are described by λ_μ which has no $\boldsymbol{r}_i$ dependence, thus Eq. (2.19) is the Schördinger equation for nuclei and $\hat{V}_{nn} + E_{e,\mu}$ is the potential energy surface of eigenstate μ. Now we can calculate Ψ as a simple product

$$\Psi = \lambda_0(\{\boldsymbol{R}_I\})\Phi_0(\{\boldsymbol{R}_I\}, \{\boldsymbol{r}_i\}) \tag{2.20}$$

by solving Eq. (2.19) and Eq. (2.6), where the subscript 0 denotes the ground state.

2.1.2 Electronic Structure in a Periodic Potential

The periodic systems investigated in this study, also known as *crystals*, are infinitely extended atomic systems that can be completely defined by a finite local image (unit cell) and a set of translational vectors. Let R_{ij} be the coordinates of ith atom in jth image, for a crystal we have:

$$R_{ij} + \boldsymbol{\mathcal{T}} = R_{ij'} \tag{2.21}$$

where i, j, j' are the indexes of images and $\boldsymbol{\mathcal{T}}$ is a translational vector of the crystal. In general, $\boldsymbol{\mathcal{T}}$ is not a unique vector but a linear combination of linearly

independent basis vectors. Following Eq. (2.21), we can define the translation operators $\hat{U}_{\boldsymbol{\mathcal{T}}}$, for an arbitrary spatial function $f(\boldsymbol{r})$:

$$\hat{U}_{\boldsymbol{\mathcal{T}}} f(\boldsymbol{r}) = f(\boldsymbol{r} + \boldsymbol{\mathcal{T}}). \tag{2.22}$$

where $f(\boldsymbol{r})$ is called to be *cell periodic* if

$$\hat{U}_{\boldsymbol{\mathcal{T}}} f(\boldsymbol{r}) = f(\boldsymbol{r}) \tag{2.23}$$

which represents the translational symmetry of function $f(\boldsymbol{r})$.

We now look at an electron-nucleus system where nuclei positions are clamped and subject to Eq. (2.21), and *electrons are not interacting with each other*. We can write the electronic Hamiltonian $\hat{H}_{ni}$ of this system as

$$\hat{H}_{ni} = \sum_{i=1}^{i=N} (-\frac{1}{2}\nabla^2_{\boldsymbol{r}_i} + v(\boldsymbol{r}_i)) \tag{2.24}$$

where x_i is the electron coordinates of the ith electron, $v(\boldsymbol{r}_i)$ is the $\hat{V}_{ne}$ term in Eq. (2.5), subscript ni denotes non-interacting electrons. We note the terms in the bracket of Eq. (2.24) can be considered as a single-electron Hamiltonian in the potential induced by fixed nuclei positions, that is $\hat{h} = -\frac{1}{2}\nabla^2_{\boldsymbol{r}} + v(\boldsymbol{r})$ which will be referred to as the Hamilton operator. Inspecting Eq. (2.22), Eq. (2.21) and Eq. (2.24) we noticed that

$$\hat{U}_{\boldsymbol{\mathcal{T}}} v(\boldsymbol{r}) = v(\boldsymbol{r}) \tag{2.25}$$

which means $v(\boldsymbol{r})$ is cell periodic, and

$$\hat{U}_{\boldsymbol{\mathcal{T}}} \nabla^2_{\boldsymbol{r}} f(\boldsymbol{r}) = \nabla^2_{\boldsymbol{r}} f(\boldsymbol{r} + \boldsymbol{\mathcal{T}}) = \nabla^2_{\boldsymbol{r}} \hat{U}_{\boldsymbol{\mathcal{T}}} f(\boldsymbol{r}) \tag{2.26}$$

which leads to

$$[\hat{U}_{\boldsymbol{\mathcal{T}}}, \hat{h}] = 0 \tag{2.27}$$

i.e., the Hamilton operator commutes with the translation operator. Thus we immediately know that $\hat{U}_{\boldsymbol{\mathcal{T}}}$ and $\hat{h}$ share the same eigenfunctions. For $\hat{h}$ we have:

$$\hat{h}\phi(\boldsymbol{r}) = \epsilon\psi(\boldsymbol{r}); \tag{2.28}$$

where ϵ is the eigenvalue of eigenfunction $\psi(\boldsymbol{r})$. Here we also assume there is no degeneracy in solving Eq. (2.28). Similarly, for $\hat{U}_{\boldsymbol{\mathcal{T}}}$ we have:

$$\hat{U}_{\boldsymbol{\mathcal{T}}}\psi(\boldsymbol{r}) = c(\boldsymbol{\mathcal{T}})\psi(\boldsymbol{r}) \tag{2.29}$$

where $c(\boldsymbol{\mathcal{T}})$ is the eigenvalue of $\hat{U}_{\boldsymbol{\mathcal{T}}}$. Considering $\hat{U}_{\boldsymbol{\mathcal{T}}}\hat{U}_{\boldsymbol{R'}}$=$\hat{U}_{\boldsymbol{R+R'}}$, from Eq. (2.29) we obtain:

$$c(\boldsymbol{\mathcal{T}} + \boldsymbol{\mathcal{T}}') = c(\boldsymbol{\mathcal{T}})c(\boldsymbol{\mathcal{T}}'). \tag{2.30}$$

From crystallography we know that $\boldsymbol{\mathcal{T}}$ is a *Bravais lattice vector* and can be decomposed to $\sum_j n_j\boldsymbol{a}_j$ where $\boldsymbol{a}_j$ are primitive vectors and n_j are linear combination coefficients. This implies Eq. (2.30) is true for any $\boldsymbol{a}_j$. Additionally, knowing Eq. (2.30) is exactly in the form of Cauchy functional equation, we can write $c(\boldsymbol{a}_j)$ in an exponential form:

$$c(\boldsymbol{a}_j) = e^{2\pi i x_j} \tag{2.31}$$

where x_j is a complex number in general. Combining Eq. (2.31) and Eq. (2.30) we have:

$$c(\boldsymbol{\mathcal{T}}) = \sum_j c(\boldsymbol{a}_j)^{n_j} = e^{2\pi i \sum_j n_j x_j}, \tag{2.32}$$

which can be written in a more compact form:

$$c(\boldsymbol{\mathcal{T}}) = e^{i\boldsymbol{k}\cdot\boldsymbol{\mathcal{T}}} \tag{2.33}$$

where $\boldsymbol{k}$ is defined as:

$$\boldsymbol{k} = \sum_j x_j\boldsymbol{b}_j \tag{2.34}$$

in which $\boldsymbol{b}_j$ are the basis vectors of reciprocal lattice:

$$\boldsymbol{b}_j \cdot \boldsymbol{a}'_j = 2\pi\delta_{jj'}. \tag{2.35}$$

Based on Eq. (2.33) and Eq. (2.29), we have:

$$\hat{U}_R\psi_{\boldsymbol{k}}(\boldsymbol{r}) = e^{i\boldsymbol{k}\cdot\boldsymbol{\mathcal{T}}}\psi_{\boldsymbol{k}}(\boldsymbol{r}) \tag{2.36}$$

which is the *Bloch theorem.* This allows us to write $\psi_{\boldsymbol{k}}(\boldsymbol{r})$ in a general form. We start by introducing a function $u_{\boldsymbol{k}}(\boldsymbol{r})$ such that

$$\psi_{\boldsymbol{k}}(\boldsymbol{r}) = e^{i\boldsymbol{k}\cdot\boldsymbol{r}}u_{\boldsymbol{k}}(\boldsymbol{r}). \tag{2.37}$$

Substitute Eq. (2.37) into Eq. (2.36) we have:

$$\hat{U}_R\psi_{\boldsymbol{k}}(\boldsymbol{r}) = e^{i\boldsymbol{k}\cdot(\boldsymbol{\mathcal{T}}+\boldsymbol{r})}u_{\boldsymbol{k}}(\boldsymbol{r}+\boldsymbol{\mathcal{T}}) = e^{i\boldsymbol{k}\cdot(\boldsymbol{\mathcal{T}}+\boldsymbol{r})}u_{\boldsymbol{k}}(\boldsymbol{r}) \tag{2.38}$$

which gives

$$u_{\boldsymbol{k}}(\boldsymbol{r}+\boldsymbol{\mathcal{T}}) = u_{\boldsymbol{k}}(\boldsymbol{r}) \tag{2.39}$$

i.e. $u_{\boldsymbol{k}}(\boldsymbol{r})$ is cell periodic. That is, Eq. (2.36) guarantees that the wavefunctions of single-electron Hamiltonian in a periodic potential can be written as a product of a plane wave and a cell periodic function. Such single-electron states/wavefunctions are commonly referred to as Bloch states/waves.

While Bloch's theorem is only rigorously valid for one particle system or non-interacting systems (Bloch electrons), it is surprising that most of realistic correlated systems can be, in one way or another, renormalized to a single-particle picture where the characteristics of the system are nicely explained by the collective behaviors of non-interacting (quasi-)particles.

Bloch's theorem introduces a new quantum number, the wave vector $\boldsymbol{k}$,

which is also formally denoted as crystal momentum of Bloch electrons. By defining the first Brillouin zone as the first Wigner–Seitz cell in the reciprocal space, we can denote any $\boldsymbol{k}'$ based on a $\boldsymbol{k}$ in the first Brillouin zone:

$$\boldsymbol{k}' = \boldsymbol{k} + \boldsymbol{G} \tag{2.40}$$

where $\boldsymbol{G}$ is the reciprocal lattice vector. This relation, along with $e^{i\boldsymbol{G}\cdot\boldsymbol{\mathcal{T}}} = 1$ from the definition of reciprocal lattice, gives us the freedom to approach the electron structure with the eigenstates in the first Brillouin zone only. Furthermore, with Eq. (2.37), we can rewrite the Schördinger equation into:

$$\hat{H}_{\boldsymbol{k}} u_{\boldsymbol{k}}(\boldsymbol{r}) = \epsilon_{\boldsymbol{k}} u_{\boldsymbol{k}}(\boldsymbol{r}) \tag{2.41}$$

by defining the Bloch Hamiltonian $\hat{H}_{\boldsymbol{k}} = \frac{1}{2}(\boldsymbol{k} - i\nabla)^2 + v(\boldsymbol{r})$. In Eq. (2.41) the eigenstates are cell periodic in real space, this means we can solve the eigenvalue problem in finite volume (one primitive cell), which inevitably quantized the eigenspectrum and leads to another quantum number, band index nb. Thus the eigenstates of this one-particle picture can be uniquely denoted by $\boldsymbol{k}$ and nb, and the eigenspectrum $\epsilon_{nb}(\boldsymbol{k})$ is often referred to as the electronic band structure or dispersion relation named after its analogue in wave mechanics.

An important property of the band structure is the effective mass which will be employed in Section 2.3. For a specific band without degeneracy, if we are only interested in the case when $\boldsymbol{k}_0$ is a extreme of the nth band, we can expand $\epsilon(\boldsymbol{k})$ as:

$$\epsilon(\boldsymbol{k}) \approx \epsilon(\boldsymbol{k}_0) + \frac{1}{2}\sum_{i,j} \frac{1}{m^*_{ij}}(k_i - k_{i0})(k_j - k_{j0}) \tag{2.42}$$

where $\frac{1}{m^*_{ij}} = \frac{\partial^2 \epsilon}{\partial k_i \partial k_j}$ is the inverse effective mass tensor and Eq. (2.42) is the effective mass approximation. If the band has spherical symmetry (thus the $\boldsymbol{k}$ components

are orthogonal to each other), the tensor can be reduced to a scalar m^* such that:

$$\epsilon(\boldsymbol{k}) \approx \epsilon(\boldsymbol{k}_0) + \frac{3}{m^*}(\boldsymbol{k} - \boldsymbol{k}_0)^2 \tag{2.43}$$

2.2 Density Functional Theory

2.2.1 Hohenberg-Kohn Theorem (DFT)

Here we introduce the basic formalism of Kohn-Sham density functional theory (DFT), which is a powerful method for solving the electronic structure problem as defined in Eq. (2.6). We assume that the ground state is non-degenerate and the electrons remain spinless. The electronic Hamiltonian is

$$\hat{H}_e = \hat{T}_e + \hat{V}_{ne} + \hat{V}_{ee} = \sum_{i=1}^{N} \frac{-\nabla_i^2}{2} + \sum_{i=1}^{N} V(\boldsymbol{r}_i) + \sum_{i \neq j} \frac{1}{2|x_i - x_j|}) \tag{2.44}$$

where we introduce the potential function V to represent a specific set of nuclei positions, in contrast to Eq. (2.3). We note that the expressions of $\hat{T}_e$ and $\hat{V}_{ee}$ are system-independent.

We start by introducing the density operator as:

$$\hat{n}(\boldsymbol{r}) = \sum_{i=1}^{N} \delta(\boldsymbol{r} - \boldsymbol{r}_i), \tag{2.45}$$

and its expectation value

$$n(\boldsymbol{r}) = \langle \Psi | \, \hat{n} \, | \Psi \rangle \tag{2.46}$$

is the electron density, so we have the constraint $\int n(\boldsymbol{r}) d^3\boldsymbol{r} = N$. Note Eq. (2.45) implies all electrons are identical. This allows us to represent $\hat{V}_{ne}$ as:

$$\hat{V}_{ne} = \sum_{i=1}^{N} V(\boldsymbol{r}_i) = \int V(\boldsymbol{r}) \sum_{i=1}^{N} \delta(\boldsymbol{r} - \boldsymbol{r}_i) d^3\boldsymbol{r} = \int V(\boldsymbol{r}) \hat{n}(\boldsymbol{r}) d^3\boldsymbol{r} \tag{2.47}$$

so the expectation value of $\hat{V}_{ne}$ can be written as:

$$\bar{V}_{ne} = \langle \Psi | \hat{V} | \Psi \rangle = \int V(\boldsymbol{r}) n(\boldsymbol{r}) d^3\boldsymbol{r} \tag{2.48}$$

which indicates $V(\boldsymbol{r})$ and $n(\boldsymbol{r})$ are *conjugate variables.* From Eq. (2.44), Eq. (2.46) and Eq. (2.48) we find the following maps exist

$$V \to \hat{H} \to \Psi \to n \tag{2.49}$$

and it can be proven that the maps are unique.

Based on Eq. (2.48), for non-degenerate ground states, Hohenberg and Kohn proved the bijective nature of Eq. (2.50), which allows unique maps:

$$n \to V \to \hat{H} \to \Psi, \tag{2.50}$$

which is referred to as *Hohenberg-Kohn theorem.* A corollary is we can represent any expectation value as a functional of n:

$$\bar{O}[n] = \langle \Psi[n] | \hat{O} | \Psi[n] \rangle . \tag{2.51}$$

By defining Hohenberg-Kohn functional $F_{\mathrm{HK}} \equiv \langle \Psi[n] | \hat{T} + \hat{W} | \Psi[n] \rangle$, we can rewrite the energy functional as:

$$E[n] = \int V(\boldsymbol{r}) n(\boldsymbol{r}) d^3\boldsymbol{r} + F_{\mathrm{HK}}[n]. \tag{2.52}$$

Since F_{HK} is not system-dependent, considering Rayleigh-Ritz principle, the ground state energy E_0 at a specific $V(\boldsymbol{r})$ can be obtained by minimizing the energy functional of electron/charge densities:

$$E_0 = \mathbf{inf}\{F_{\mathrm{HK}}[n] + \int V(\boldsymbol{r}) n(\boldsymbol{r}) d^3\boldsymbol{r}\}. \tag{2.53}$$

2.2.2 Kohn-Sham Equations

Now the problem is reduced to the minimization of $E[n]$ with constraint $\int n(\boldsymbol{r})d\boldsymbol{r} = N$. The Lagrangian is:

$$L_{V,N}[n] = E[n] - \mu[\int n(\boldsymbol{r})d^3\boldsymbol{r} - N], \tag{2.54}$$

where μ is the Lagrange multiplier. The basic idea in DFT is to find the $n(\boldsymbol{r})$ that zeros the Lagrangian's derivative with respect to $n(\boldsymbol{r})$:

$$\frac{\delta L_{V,N}[n]}{\delta n(\boldsymbol{r})} = \frac{\delta F_{\mathrm{HK}}[n]}{\delta n(\boldsymbol{r})} + V(\boldsymbol{r}) - \mu = 0. \tag{2.55}$$

However, the mathematical form of $F_{\mathrm{HK}}[n]$ is still a mystery. To solve this problem, Kohn and Sham proposed a scheme for a practical form of $F_{\mathrm{HK}}[n]$ with only part of it is approximated by the exchange-correlation functional $E_{xc}[n]$. From (functional) Hellmann-Feynman theorem we have:

$$\frac{\delta E}{\delta v(\boldsymbol{x})} = \langle\Psi| \frac{\delta \hat{H}_e}{\delta V(\boldsymbol{r})} |\Psi\rangle = n(\boldsymbol{r}), \tag{2.56}$$

which allow us to define a Legendre transform based on Eq. (2.52)

$$\mathcal{L}[n] = E[V] - \int V(\boldsymbol{r})n(\boldsymbol{r})d^3\boldsymbol{r}. \tag{2.57}$$

Due to the uniqueness of the map $n \to V \to \mathcal{L}[n]$ provided by Hohenberg-Kohn theorem and Eq. (2.52) we conclude that:

$$\mathcal{L}[n] = F_{\mathrm{HK}}[n]. \tag{2.58}$$

Based on Eq. (2.57), the chain rule of functional derivative gives us:

$$\frac{\delta F_{\mathrm{HK}}}{\delta n(\boldsymbol{r})} = -V(\boldsymbol{r}). \tag{2.59}$$

Now consider an auxiliary system that has N *non-interacting* electrons with a specific external field $V_s(\boldsymbol{r})$ and an electron density $n(\boldsymbol{r})$ *identical* to the interacting system that we are interested in. We can obtain its Hohenberg-Kohn functional $F_s[n]$ with the same procedure based on Eq. (2.57):

$$F_s[n] = E_s[V_s] - \int V_s(\boldsymbol{r})n(\boldsymbol{r})d^3\boldsymbol{r}. \tag{2.60}$$

The beauty of $F_s[n]$ is it is merely the kinetic energy of a non-interacting system, which can be written explicitly as:

$$F_s[n] = -\frac{1}{2}\sum_{i=1}^{N}\int \phi_i^*(\boldsymbol{r})\nabla^2\phi_i(\boldsymbol{r})d^3\boldsymbol{r} \tag{2.61}$$

where ϕ_i stands for the element in Slater determinant of the non-interacting ground state Φ_s (Kohn-Sham wavefunction). Now we try to approximate F_{HK} with F_s by defining $E_{xc}[n]$ within:

$$F_{\mathrm{HK}}[n] = F_s[n] + E_H[n] + E_{xc}[n] \tag{2.62}$$

where $E_H[n] = \frac{1}{2}\int\int \frac{n(\boldsymbol{r})n(\boldsymbol{r}')}{|\boldsymbol{r}-\boldsymbol{r}'|}d^3\boldsymbol{r}d^3\boldsymbol{r}'$ is the Hartree energy functional representing electrostatic energy associated with $n[r]$. Combining Eq. (2.52) and Eq. (2.62), we have:

$$E[n] = F_s[n] + E_H[n] + E_{xc}[n] + \int V(\boldsymbol{r})n(\boldsymbol{r})d^3\boldsymbol{r} \tag{2.63}$$

Combining Eq. (2.59) and Eq. (2.62), the derivative of $F_{\mathrm{HK}}[n]$ with respect to $n(x)$ gives the expression of $V_s(\boldsymbol{r})$:

$$\frac{\delta F_{\mathrm{HK}}[n]}{\delta n(\boldsymbol{x})} = -V(\boldsymbol{r}) = -V_s(\boldsymbol{r}) + \frac{\delta E_H[n]}{\delta n(\boldsymbol{x})} + \frac{\delta E_{xc}[n]}{\delta n(\boldsymbol{x})}. \tag{2.64}$$

This gives an alternative expression for the single-particle Schördinger equations of the non-interacting system, by defining exchange-correlation potential $V_{xc}(\boldsymbol{r})$

as $\frac{\delta E_{xc}[n]}{\delta n(\boldsymbol{x})}$:

$$(-\frac{1}{2}\nabla^2 + V_s(\boldsymbol{r}))\phi_i(\boldsymbol{r}) = (-\frac{1}{2}\nabla^2 + \frac{\delta E_H[n]}{\delta n(\boldsymbol{x})} + V_{xc}(\boldsymbol{r}))\phi_i(\boldsymbol{r}) = \epsilon_i\phi_i(\boldsymbol{r}) \tag{2.65}$$

where ϵ_i stands for the eigenvalue for ϕ_i (also known as Kohn–Sham orbitals). The electron density can be calculated from the non-interacting system:

$$n(\boldsymbol{r}) = \sum_{i=1}^{N} \phi_i^*(\boldsymbol{r})\phi_i(\boldsymbol{r}). \tag{2.66}$$

Eq. (2.63), Eq. (2.65) and Eq. (2.66) constitute the non-degenerate ground state Kohn-Sham equations. Once a suitable $E_{xc}[n]$ is defined, we can calculate $V_{xc}(\boldsymbol{r})$, solve Eq. (2.65) and compute any observable via $n(\boldsymbol{r})$ from Eq. (2.66). Since Kohn-Sham equations are coupled and highly nonlinear, practically, a self-consistent iterative scheme is used to solve them numerically.

2.2.3 Exchange-Correlation Functional

While Eq. (2.63) is in principle exact, $E_{xc}[n]$ remains unknown and so does the exchange-correlation potential V_{xc}. Various approximations have been developed to give a closed form of $E_{xc}[n]$. The "mother" of all approximations is the local density approximation:

$$E_{xc}^{\text{LDA}} = \int \epsilon^{xc\text{-LDA}}(n(\boldsymbol{r}))n(\boldsymbol{r})d^3\boldsymbol{r} \tag{2.67}$$

where $\epsilon^{xc\text{-LDA}}[n]$ is the exchange-correlation energy per particle of the homogeneous electron gas of density $n(\boldsymbol{r})$. Strictly speaking Eq. (2.67) is only valid for slowly varying $n(\boldsymbol{r})$.

An exact form of E_{xc} can be obtained via the *gradient expansion*:

$$E_{xc} = E_{xc}^{\text{LDA}} + \int g_1(n(\boldsymbol{r}))(\nabla n(\boldsymbol{r}))^2 d^3\boldsymbol{r} + \int g_2(n(\boldsymbol{r}))(\nabla^2 n(\boldsymbol{r}))^2 d^3\boldsymbol{r} + \dots \tag{2.68}$$

where g_i are expansion coefficients. This gives rise to the generalized gradient approximation (GGA):

$$E_{xc}^{\mathrm{GGA}} = \int n(\boldsymbol{r})\epsilon_{xc}^{\mathrm{GGA}}(n(\boldsymbol{r}), \nabla n(\boldsymbol{r}))d^3\boldsymbol{r} \tag{2.69}$$

where $\epsilon_{xc}^{\mathrm{GGA}}$ depends on not only electron density but its gradient. A more advanced functional type is the meta-GGA functional:

$$E_{xc}^{m\mathrm{GGA}} = \int n(\boldsymbol{r})\epsilon_{xc}^{m\mathrm{GGA}}(n(\boldsymbol{r}), \nabla n(\boldsymbol{r}), \nabla^2 n(\boldsymbol{r}), \tau(\boldsymbol{r}))d^3\boldsymbol{r} \tag{2.70}$$

where $\tau = \sum_{i=1}^{N_{\mathrm{occu}}} \frac{1}{2}|\nabla\phi_i(\boldsymbol{r})|^2$ is the kinetic energy density. Moreover, E_{xc}^{DFT} can be mixed with the exact Hartree-Fork exchange energy calculated from ϕ_i:

$$E_{xc}^{\mathrm{hybrid}} = (1-a)E_{xc}^{\mathrm{DFT}} + aE_x^{\mathrm{HF}} \tag{2.71}$$

where $0 < a < 1$ is the mixing coefficient and E_{xc}^{hybrid} is the hybrid exchange-correlation functional. Eq. (2.67), Eq. (2.69), Eq. (2.70) and Eq. (2.71) are the first four rungs of Jacob's ladder of the density functional approximation, while E_{xc} of the fifth rung also includes information of unoccupied ϕ_i in addition to Eq. (2.71).

2.2.4 Dispersion Correction

One important caveat of Eq. (2.67), Eq. (2.69), Eq. (2.70) and even Eq. (2.71) is that only local contributions to the electron correlation are included in E_{xc}. One direct consequence is long range dispersion is neglected such that the correct R^{-6} dependence of the interaction energy on interatomic distance in long range cannot be produced. To reduce such errors, especially for molecular/inhomogeneous systems, currently there are two types of correction methods have been frequently employed:

- **vdW-DF** using a nonlocal term E^{NL} to describe the dispersion interaction: $E_{xc}^{\mathrm{vdW-DF}} = E_{xc}^{\mathrm{DFT}} + E^{\mathrm{NL}}$. Generally, E^{NL} is calculated by

$$E^{\mathrm{NL}} = \int\int n(\boldsymbol{r})\chi(\boldsymbol{r},\boldsymbol{r}')n(\boldsymbol{r}')d^3\boldsymbol{r}d^3\boldsymbol{r}' \tag{2.72}$$

  where χ is the nonlocal correlation kernel which is flavor-dependent.

- **DFT-D** is a type of semiclassical corrections that add a correction term E^{disp} to the Kohn–Sham energy (Eq. (2.63))

$$E^{\mathrm{DFT-D}} = -\sum_{AB}\sum_{n=6,8,10,\ldots}\frac{s_n C_n^{AB}}{R_{AB}^n} f_{\mathrm{damp}}(R_{AB}) \tag{2.73}$$

  where AB denotes an atomic pair, R the interatomic distance, s_n the global scaling factor. f_{damp} is the damping function used to control the range of dispersion correction and C_n is the averaged nth order dispersion coefficient. The negative sign comes from the attractive nature of dispersion interactions.

In vdW-DF the dispersion effects are taken into account via E_{xc} which allows charge density corrections if it is applied self-consistently. However, considering its high computational complexities, along with few tests for organic molecular systems, DFT-D scheme is used throughout this study.

2.3 Charge Transport Models

One of the most fundamental processes in an organic semiconductor is the charge-carrier transport process where (excess) charges percolate from site to site. This process is particularly interesting in organic materials due to their soft nature, which leads to potentially large electron-phonon coupling, and the intermediate couplings between molecules/π-fragments. The latter makes the charge transport behaves fall between the band model (strong coupling limit) and hopping model (weak coupling limit). Here we briefly introduce these two limits.

2.3.1 Band Model

Due to the uncertainty principle, the dynamics of microscopic particles can no longer be described with point-mass model. While it is meaningless to ask what the position of an electron is for a specific eigenstate, alternatively, wave packets, constructed from Bloch states, could be used to describe the particle dynamics. For the nth band without degeneracy, a wave packet $\psi_{\boldsymbol{k}_0}(\boldsymbol{r}, t)$ localized around $\boldsymbol{k}_0$ can be constructed as a superposition of Bloch states:

$$\psi_{\boldsymbol{k}_0}(\boldsymbol{r}, t) = \sum_{\boldsymbol{k}} \mathcal{A}_{\boldsymbol{k}_0}(\boldsymbol{k}) u_{\boldsymbol{k}}(\boldsymbol{r}) e^{i\boldsymbol{k}\cdot\boldsymbol{r} - i\epsilon(\boldsymbol{k})t} \tag{2.74}$$

where $\mathcal{A}_{\boldsymbol{k}_0}(\boldsymbol{k})$ is the amplitude function that strongly peaked at $\boldsymbol{k}_0$, $\epsilon(\boldsymbol{k})$ is the dispersion relation. Such wave packet has a group velocity $\mathfrak{v}(\boldsymbol{k}_0)$ as:

$$\mathfrak{v}(\boldsymbol{k}_0) = \nabla_{\boldsymbol{k}} \epsilon(\boldsymbol{k})|_{\boldsymbol{k}=\boldsymbol{k}_0}. \tag{2.75}$$

Considering the one-dimensional cases, the physical quantity of interest here is the mobility μ, which is formally defined as:

$$\mu = \langle \mathfrak{v}(k) \rangle / \mathcal{E} \tag{2.76}$$

where $\langle \mathfrak{v}(k) \rangle$ is the mean velocity of the Bloch wave packets and $\mathcal{E}$ is the applied constant, uniform electric field which is parallel to $\mathfrak{v}(k)$. The mean velocity, by definition, can be calculated by:

$$\langle \mathfrak{v}(k) \rangle = \int \mathfrak{v}(k) f(k) dk \tag{2.77}$$

where $f(k)$ is the distribution function such that $f(k)dk$ is the number of wave packets in the unit volume around k and the integral runs over the whole reciprocal space.

While at thermal equilibrium $f(k)$ is the Fermi-Dirac distribution $f_{\mathrm{F}}(k)$, in general the form of $f(k)$ is unknown. To address this issue, we need to examine the Boltzmann equation, which has the general form:

$$\frac{\partial f}{\partial t} + \mathfrak{v}(k) \cdot \nabla_r f + \dot{k} \cdot \nabla_k f = \frac{\partial f}{\partial t}_{\text{collision}}. \tag{2.78}$$

The right side of the Boltzmann equation is the collision contribution to the time variation of f, noting now f is defined in the phase space. Assuming the system is at a stationary state ($\frac{\partial f}{\partial t} = 0$), spatially uniform ($\nabla_r f = 0$) and weak electric field, with relaxation time approximation $\frac{\partial f}{\partial t}_{\text{collision}} = -\frac{f - f_{\mathrm{F}}}{\tau(\epsilon)}$ ($\tau(\epsilon)$ is the relaxation time), the Boltzmann equation now is:

$$\dot{k} \cdot \nabla_k f(k) = -\frac{f(k) - f_{\mathrm{F}}(k)}{\tau(\epsilon)}. \tag{2.79}$$

Within the semiclassical scheme of electron dynamics, for constant, uniform electric field we have $\dot{k} = -\mathcal{E}$, and the left side of Eq. (2.79) can be written as:

$$-\frac{\partial f_{\mathrm{F}}(\epsilon(k))}{\partial \epsilon(k)} \nabla_k \epsilon(k) \cdot \mathcal{E} = -\frac{\partial f_{\mathrm{F}}(\epsilon)}{\partial \epsilon} \mathfrak{v}(k) \cdot \mathcal{E} \tag{2.80}$$

and now Eq. (2.79) is:

$$f(k) = f_{\mathrm{F}}(k) + \tau(\epsilon) \mathfrak{v}(k) \mathcal{E} \frac{\partial f_{\mathrm{F}}(\epsilon)}{\partial \epsilon} \tag{2.81}$$

which can be considered as a first order perturbation of $f(k)$ by $\mathcal{E}$.

From now till the end of this section we would abandon atomic units for a precise description of mobility. Combining Eq. (2.81) and Eq. (2.77), applying $\epsilon(k) = \frac{1}{2} m^* \mathfrak{v}^2(k)$ (semiclassical scheme) and $\frac{\partial f_{\mathrm{F}}(\epsilon)}{\partial \epsilon} = \frac{f_{\mathrm{F}}(\epsilon)(1 - f_{\mathrm{F}}(\epsilon))}{k_{\mathrm{B}} T}$, we have:

$$\langle \mathfrak{v}(k) \rangle = \frac{2 e_0 \mathcal{E}}{m^* k_{\mathrm{B}} T} \int \epsilon(k) \tau(k) f_{\mathrm{F}}(\epsilon(k)) (1 - f_{\mathrm{F}}(\epsilon(k))) dk \tag{2.82}$$

where the term $\int \mathfrak{v}(k) f_{\mathrm{F}}(k) dk$ is absent since $f_{\mathrm{F}}(k)$ is an even function. When

$f_{\mathrm{F}} \ll 1$ (Boltzmann limit), $f_{\mathrm{F}}(k)$ reduces to a Maxwellian $f_{\mathrm{F}}(k) = e^{-\frac{\epsilon-\epsilon_{\mathrm{F}}}{k_{\mathrm{B}}T}}$ and its derivative now is $\frac{\partial f_{\mathrm{F}}(\epsilon)}{\partial \epsilon} = -\frac{1}{k_{\mathrm{B}}T} e^{-\frac{\epsilon-\epsilon_{\mathrm{F}}}{k_{\mathrm{B}}T}}$, where ϵ_{F} is the Fermi level and k_{B} is the Boltzmann constant. Thus Eq. (2.82) can be rewritten to:

$$\langle \mathfrak{v}(k) \rangle = -\frac{2e_0\mathcal{E}}{m^* k_B^2 T^2} \int \epsilon(k)\tau(k) e^{-\frac{\hbar^2 k^2}{2m^* k_{\mathrm{B}} T}} dk \tag{2.83}$$

where we used the one-dimension effective mass approximation $\epsilon - \epsilon_{\mathrm{F}} = \frac{\hbar^2 k^2}{2m^*}$.

Now the only variable left is the relaxation time $\tau(k)$, which indicates the mean time needed for the system to return to f_{F} from f. In a scattering picture, this is equivalent to the mean time between two scattering events. Here we introduce the one-dimension deformation potential theory to estimate $\tau(k)$ as this approach will be used in the later chapter. For one-dimension system with elastic electron-phonon scattering, $\tau(k)$ is expressed as:

$$\frac{1}{\tau(k)} = \sum_q \Theta(k, k')(1 - \frac{k'}{k}) \tag{2.84}$$

where $\Theta(k, k')$ is the transition probability for the wave packet scattered from k to $k' = k + q$. Fermi's golden rule enables us to calculate $\Theta(k, k')$ by:

$$\Theta(k, k') = \frac{2\pi}{\hbar} |M(k, k')|^2 \delta(\epsilon(k) - \epsilon(k') - \hbar\omega_q) \tag{2.85}$$

where $M(k, k') = \langle \psi_k | \Delta\hat{V}(r) | \psi_{k'} \rangle$ is the electron-phonon coupling element between state ψ_k and $\psi_{k'=k+q}$, $\Delta\hat{V}(r)$ is the perturbation potential used to describe the electron-phonon coupling, q is the phonon momentum and ω_q is the phonon frequency. The deformation potential theory gives an estimation of $M(k, k')$ by assuming $\Delta\hat{V}(r) = E_1 \Delta(r)$ where $\Delta(r)$ is the relative volume change of the unit cell and constant E_1 is the deformation potential constant associated with a band extreme (intravalley scattering).[52] That is, the electron-phonon scattering is simulated via a perturbation manner induced by the change in unit cell. Assuming

the phonon band is fully occupied, we have:

$$|M(k,k')|^2 = \frac{k_{\mathrm{B}}TE_1^2}{2CL} \tag{2.86}$$

where C is the elastic constant and L is the length of the one-dimension unit cell. Combining Eq. (2.85), Eq. (2.84) and Eq. (2.86), we have:

$$\frac{1}{\tau(k)} = \frac{\pi k_{\mathrm{B}}TE_1^2}{\hbar CL}\sum_q (1-\frac{k'}{k})\delta(\epsilon(k)-\epsilon(k')-\hbar\omega_q) \tag{2.87}$$

which can be rewritten into its integral form for intravalley scattering ($\hbar\omega_q \to 0$):

$$\frac{1}{\tau(k)} = \frac{k_{\mathrm{B}}TE_1^2}{2\hbar C}\int_q \frac{q}{k}\delta(\frac{\hbar^2q^2}{2m^*})dq \tag{2.88}$$

where we used $\sum_q q = \frac{L}{2\pi}\int qdq$ from Born-von Karman boundary condition and effective mass approximation. Performing the integration in Eq. (2.88) we obtain:

$$\frac{1}{\tau(k)} = \frac{m^*k_{\mathrm{B}}TE_1^2}{\hbar^3Ck}. \tag{2.89}$$

Finally, with Eq. (2.76), Eq. (2.89) and Eq. (2.83), we obtain the expression for mobility in one-dimension deformation potential theory:

$$\mu = \frac{e_0\hbar^2C}{(2\pi k_{\mathrm{B}}T)^{\frac{1}{2}}m^{*\frac{3}{2}}E_1^2}. \tag{2.90}$$

2.3.2 Hopping Model

In the weak coupling limit the charge becomes localized on a molecular (fragment) site and the charge transport process becomes temperature-driven. The rate of such process can be described via the Marcus charge-transfer theory:

$$k = \frac{2\pi}{\hbar}\frac{J^2}{\sqrt{4\pi\lambda k_{\mathrm{B}}T}}e^{-\frac{\Delta G^0\lambda}{4k_{\mathrm{B}}T}} \tag{2.91}$$

where J is the electronic coupling, λ the reorganization energy, ΔG^0 the free energy change of charge transport process, k_B the Boltzmann constant and T the temperature. Eq. (2.91) can be extended to a more general form:

$$k = \frac{2\pi}{\hbar}\frac{J^2}{\sqrt{4\pi\lambda_s k_\mathrm{B}T}}\sum_{\nu=0}^{\inf} e^{-S}\frac{S^\nu}{\nu!}e^{-\frac{(\lambda_s+\nu\hbar\omega)^2}{4\lambda_s k_\mathrm{B}T}} \tag{2.92}$$

where λ_s is the reorganization energy from environment (outer sphere) and S represents the Huang-Rhys factor calculated at an effective mode ω contributing to molecular (inner sphere) reorganizations. Eq. (2.92) is the Marcus-Levich-Jortner equation which treats molecular vibrations quantum mechanically, unlike the pure classical treatment in Eq. (2.91). Thus, whether using Eq. (2.91) or Eq. (2.92) depends on temperature and the vibrational modes: if $\hbar\omega \gg k_\mathrm{B}T$ then Eq. (2.92) should be used, if $\hbar\omega \ll k_\mathrm{B}T$ then Eq. (2.91) may be considered as a good approximation.

Now we look at mobility calculation. Instead of wave packets we have charge carriers hopping from a site $\mathfrak{S}$ to its neighbors. Assuming Brownian motion of charge carriers, the diffusion coefficient D can be approximated as:

$$D = \frac{1}{2}\sum_i a_i^2 k_i P_i \tag{2.93}$$

where a_i, k_i, $P_i = \frac{k_i}{\sum_i k_i}$ are the distance, rate and probability of hopping to the ith neighbor of $\mathfrak{S}$, respectively. Eq. (2.93) is used in the Einstein–Smoluchowshi equation:

$$\mu = \frac{q}{k_\mathrm{B}T}D \tag{2.94}$$

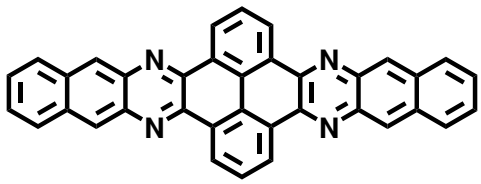
to give mobility, where q is the charge of charge carriers.

CHAPTER 3

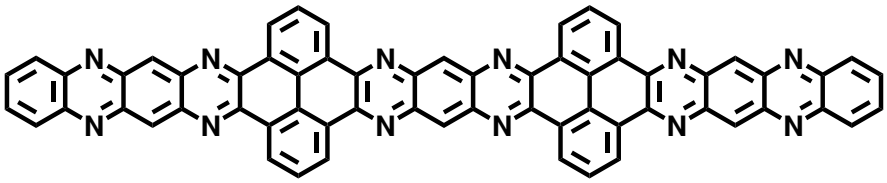
FUSING ACENES TO PYRENE: A LONGER ACENE?

3.1 Introduction

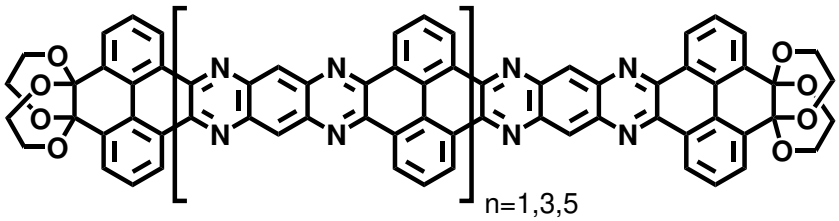
It is remarkable how the simple structures built from laterally fusing benzene rings known as *acenes* have attracted considerable attention from different scientific communities since the term was coined in 1930s by Clar and colleagues.[53] While applications in electronic devices, especially in field-effect transistors,[54–56] contributed enormously to the field of organic electronics, theoreticians are fascinated by phenomena the stem from its unique optoelectronic properties (*e.g.* singlet fission).[57–59] On the other hand, from the perspectives of organic chemistry, long acenes present a synthetic challenge due to their notorious air and light instabilities, which pose strict requirements on reaction environment .[60] As the quest of making longer acenes has led to the successful synthesis of undecacene,[61,62] an interesting byproduct is the development of various chemical modifications used to increase the stability of acene backbone, including attaching functional groups (*e.g.* 6,13-bis(triisopropylsilylethynyl)pentacene),[63] heteroatom substitution (*e.g.* azaacene[64]) and polycyclic aromatic hydrocarbon (PAH) insertion.[65] While the last approach, corroborated by Clar's sextet rule, has been quite successful in stabilizing the molecular system such that molecules with 15 rings in a row can exist as single crystals at room temperature,[66] it is questionable that the resultant can be considered as an acene since the molecular backbone has been significantly altered.[67]



2.42 eV (dichlorobenzene)



1.85 eV (dichloromethane)



1.89 eV (chloroform)

Figure 3.1: Prene-azaacene systems in previous studies with absorption spectra onset of the first peak.[66,68,69]

Among various PAH-inserted acenes and azaacenes,[66,68,70,71] we are intrigued by the pyrene-inserted molecules. Such systems can be extended to almost macromolecule size without suffering from significant temperature/air instability.[68,71,72] Unlike acenes, they are also able to support a stable dianion, whose ground state is believed to be an open-shell singlet.[73] Aromaticity of their molecular backbone is said to be "broken" at the pyrene units where weak magnetic shielding is identified.[66] Anomalies are also present in their UV-vis absorption spectra. Let n_{xrings} be the number of fused benzene rings in an acene or a pyrene-(aza)acene molecule along the long axis, and $\epsilon^{S_1}_{\text{acene}}$, $\epsilon^{S_1}_{\text{PA}}$ be the excitation energy of the first singlet excited state S_1 for acene and pyrene-(aza) acene, respectively: $\epsilon^{S_1}_{\text{PA}}$ is always much higher than $\epsilon^{S_1}_{\text{acene}}$ with the same n_{xrings}, and $\epsilon^{S_1}_{\text{PA}}$ barely decrease within a homologous series, as shown in Fig. 3.1.

In this chapter, we present a systematic computational analysis of coupled

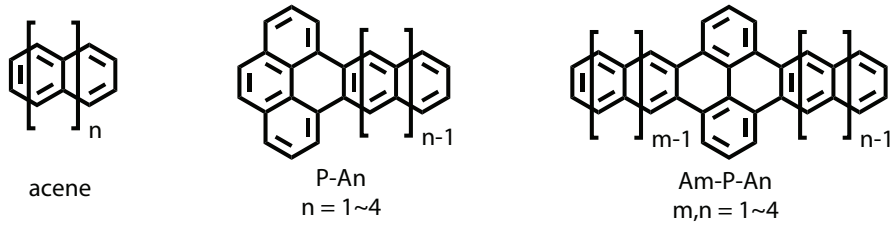


Figure 3.2: Molecular structures and their abbreviations of pyrene-acene systems.

pyrene-acene systems (Fig. 3.2, where m and n denote the lengths of acene segments). We relate this property to local structural relaxations, which in turn leads to anomalous optical responses and triplet localization. Finally, the implication of pyrene insertion in polymer is discussed and another PAH is identified to be able to interrupt the large acene system in a manner similar to pyrene. An earlier version of some parts of this chapter was published in *Chemistry of Materials* **2018** *30*(3), 947-957.

3.2 Methods

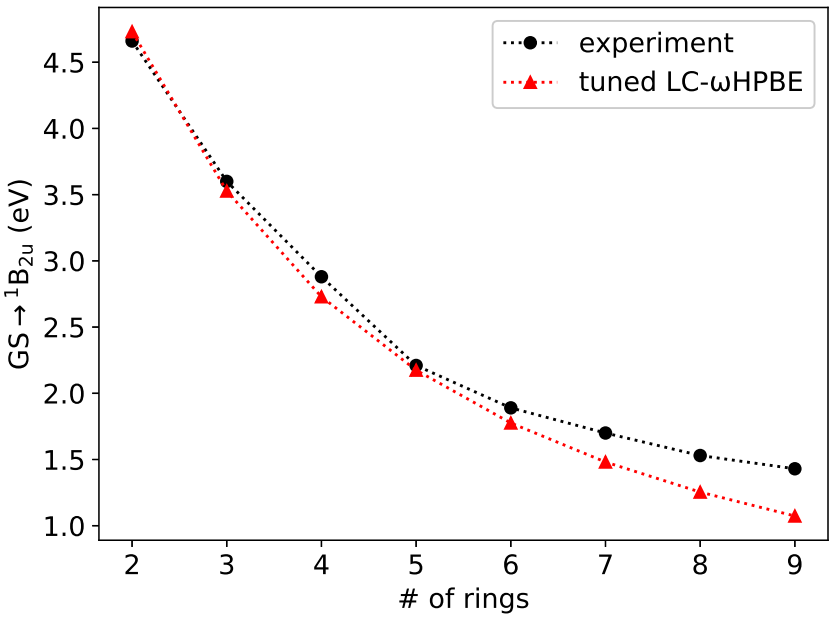


Figure 3.3: Vertical excitation energies of the first bright excited state ($^1B_{2u}$) in acenes.

Unrestricted Kohn-Sham density functional theory (DFT) calculations were performed to investigate the electronic states involved in this study. For molecular systems, the long-range corrected functional LC-ωHPBE was used along with def2tzvp basis set as implemented in *Gaussian 16*.[74–77] Molecular excited states were studied via linear response time-dependent density functional theory (TDDFT). To accurately estimate the excitation energies for conjugated systems with various sizes, the range-separation parameter ω was tuned via minimizing the sum of the ionization energy and the highest occupied molecular orbital (HOMO) energy.[78,79] Such tuning process, combined with geometric optimization, was performed iteratively such that $\Delta\omega < 0.005$ Bohr^{-1}. As a benchmark for this method, the lowest bright state of acene, from naphthalene to nonacene, is calculated and compared with experimental results, as shown in Fig. 3.3. While reasonably good agreement was obtained for acenes up to hexacene, starting from heptacene, the difference between calculated and experimental excitation energy increases with increasing system size, reaching ∼360 meV at nonacene. Such difference may come from the strong static correlations in higher acenes that cannot be accurately described by single-determinant methods. To minimize inconsistency, all molecular calculations are performed at the optimally tuned LC-ωHPBE/def2tzvp level of theory, and spin contamination, if any, is negligible ($< S^2 > -s(s+1) < 0.02$, where $< S^2 >$ is the total spin and s equals 0.5 times the number of unpaired electrons) in all cases.

Nuclear magnetic shielding tensors were calculated with the gauge-including atomic orbital (GIAO) method,[80] and the magnetic shielding tensors were calculated for a scan path 1.7 Å above the molecular plane.[81] For all molecules, the scan path was defined along the long axis of the molecular backbone with a step size of 0.4 Å.

For polymer systems, the Perdew, Burke, and Ernzerhof (PBE) exchange-correlation functional[82] with electron-ion interactions described by the projector

augmented-wave method[83] and full periodic boundary conditions using the Vienna *Ab-initio* Simulation Package.[84] In order to obtain accurate energy-volume relations, the fast Fourier transformation mesh was manually set to $150 \times 150 \times 150$ for all unit cells. The kinetic energy cutoff for the plane-wave basis set was 520 eV, and Gaussian smearing with a width of 0.05 eV was employed. The total energy convergence criterion was set to 0.01 meV in all self-consistent field iterations, and the maximum force allowed on each atom was 0.01 eV/Å. The Brillouin zone, unless specified otherwise, was sampled with a $(300/a) \times 1 \times 1$ Γ-centered grid, where a is the lattice constant of the polymer in Å. Note that vacuum is inserted into the respective unit cells in order to treat each structure as a non-interacting, one-dimensional polymer; enough vacuum is inserted such that the shortest distance between two 1D polymers is larger than 10 Å.

3.3 Results and Discussion

Molecular structures of pyrene-acene systems and their abbreviations are shown in Fig. 3.2. These structures are classified as acenes, pyrene-acenes (P-A), and acene-pyrene-acenes (A-P-A), where A stands for 1∼5 linearly fused rings (from benzene to pentacene). A specific molecule is denoted based on the number of rings in its acene segments, *e.g.* P-A3 represents a pyrene-anthracene molecule, and A2-P-A3 a naphthalene-pyrene-anthracene molecule.

3.3.1 Ground State

DFT calculations were performed to obtain the ground-state structures of molecules in Fig. 3.2 (see Section 3.2 for details). Bond length alternation (BLA) patterns in the π-conjugated backbones are shown in Fig. 3.4, where C-C bonds are colored such that the color linearly changes from blue (long, 1.46 Å) to red (short, 1.35 Å). For the trans-polyacetylene-like edges of acene molecules, the C-C bond length oscillates around the commonly seen aromatic C-C bond length (1.40 Å)

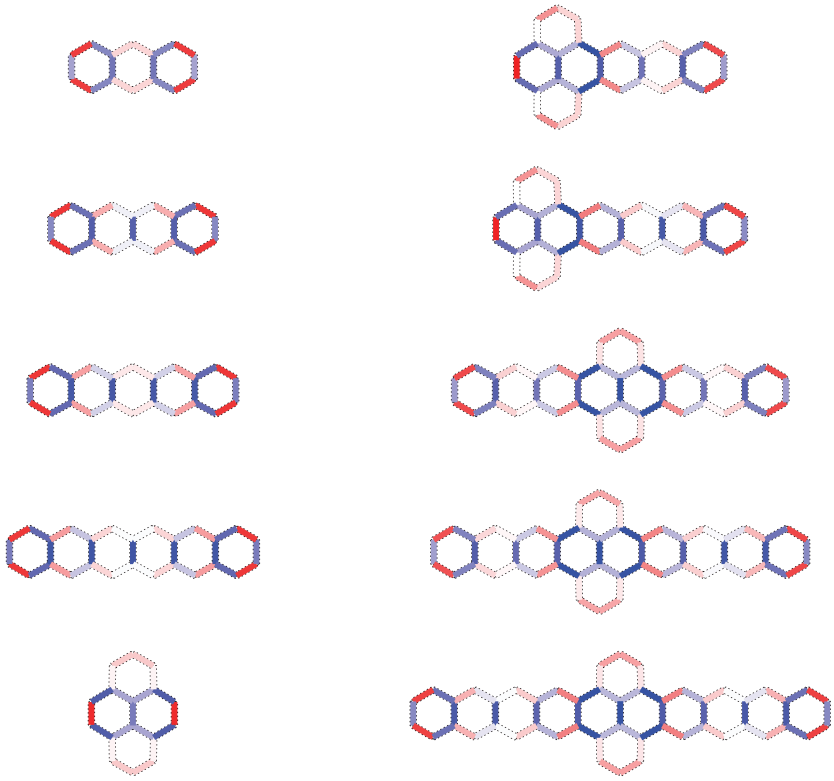


Figure 3.4: Bond length alternation patterns of (left) from top to bottom: anthracene, tetracene, pentacene, hexacene, pyrene, and (right) P-A3, P-A4, A3-P-A3, A3-P-A4, A4-P-A5. Only C-C bonds in molecular backbones are shown.

from the ends to the middles of molecules with decreased amplitude. Interestingly, for both P-A and A-P-A, their BLA patterns can be roughly considered as a direct concatenation of those of pyrene and the corresponding acenes, with the exception of the C-C bond that directly connects these two fragments. In fact, some of the bonds in the pyrene units are so long (1.47∼1.48 Å) that a P-A (A-P-A) molecule can almost be regarded as a combination of biphenyl unit and acene unit(s) by σ bonds. We note that in the A-P-A case such statement remains valid even two acenes with different length were added, *e.g.* A3-P-A4.

A chemically intuitive approach to understand such BLA patterns is the Clar formulae, originated from the Clar sextet rule, where it is proposed that the chemical structure with a maximum number of sextets best represents the (localization) behavior of π electrons.[53] Within this framework, as shown in Fig. 3.5, for acenes of arbitrary length, while degenerate Clar formulae can be drawn based

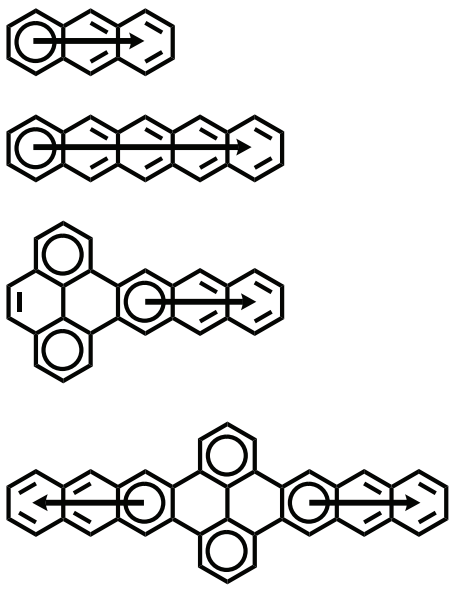


Figure 3.5: Clar formulae of, from top to bottom, anthracene, pentacene, P-A3, A3-P-A3.

on the number of fused rings, only one sextet (*i.e.* the migrating sextet[85,86]) from their Kekulé formula is allowed without bond breaking (*e.g.* forming radicals). In passing we note that while the instability of long acenes can be related to the lack of sextets in their Clar formulae, the migrating sextet cannot accurately represent the nature of long acenes where the central ring exhibit higher reactivity and, based on several local aromaticity descriptors, higher aromaticity, than the terminal rings. For P-A (A-P-A) molecules, just as their BLA patterns, their Clar formulae resulted from merging Clar formulae of pyrene and acenes, leading to three (four) isolated sextet. This is also in agreement with the stability of large pyrene-acene systems, *e.g.* the stability of A3-P-A3 should be comparable to that of anthracene, not that of octacene.

Nucleus-independent chemical shift (NICS) based methods, where the negative of the absolute magnetic shielding tensor ($\mathrm{NICS}_{\alpha\beta}$, where α, β stand for the indices of components for eternal magnetic field and magnetic moment, respectively) are calculated under an external magnetic field, are important computational tools to investigate aromaticity in PAHs. Here, the NICS scan was used as an approach to assess the aromatic characteristics of P-A, A-P-A systems

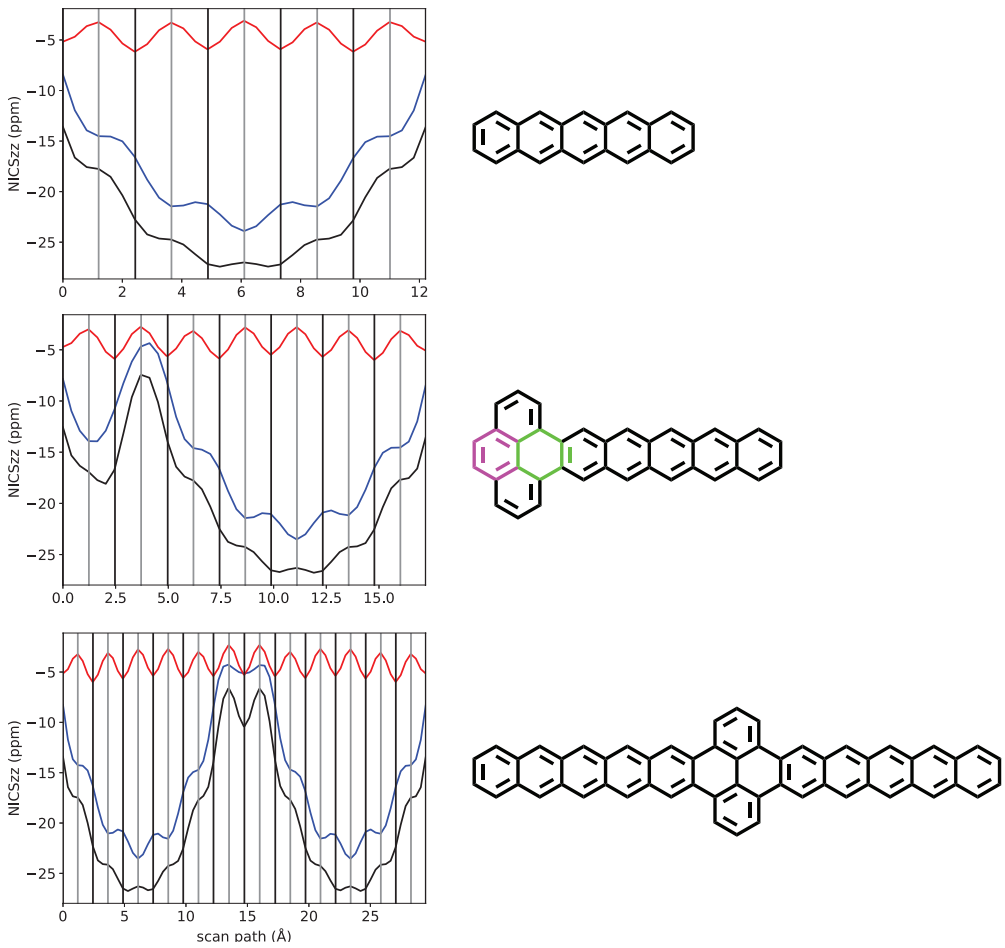


Figure 3.6: NICS-xy scan of pentacene, P-A5, and A5-P-A5. NICS_{zz}, its π and σ components are colored black, blue and red, respectively. For each molecule, the scan path consists of points on a straight line, separated by 0.4 Å, running from the left end to the right end.

and acenes.[81] In particular, the NICS_{zz} component (z represents the normal direction of molecular plane) is further decomposed into the contributions from π orbitals and σ orbitals by calculating NICS_{zz} for an artificial conformation where aromatic carbons are terminated by hydrogen atoms. As shown in Fig. 3.6, using pentacene, P-A5 and A5-P-A5 as examples, we observe a clear resemblance between their pentacene segments in terms of their NICS-xy profiles. For A5-P-A5, the $\mathrm{NICS}_{\pi zz}$ profiles are quite similar ($<$ 5 ppm at ring centers) to the $\mathrm{NICS}_{\sigma zz}$ profile at their pyrene segments, indicating weakly induced ring currents, thus more localized π electrons in these regions. Interestingly, in P-A5, while the small differences between $\mathrm{NICS}_{\pi zz}$ and $\mathrm{NICS}_{\sigma zz}$ present in the pyrene ring which shares an edge with the acene segment (green ring in Fig. 3.6, the second ring

along the scan path), for the other pyrene ring in scan path (magenta ring), values of $NICS_{\pi zz}$ are much lower than $NICS_{\sigma zz}$. This observation, along with the BLA patterns and Clar formulae of P-A and A-P-A, suggest that fusing acene to pyrene induces local structural relaxations at the junction, and concurrent limited electronic communication among pyrene and acene segments.

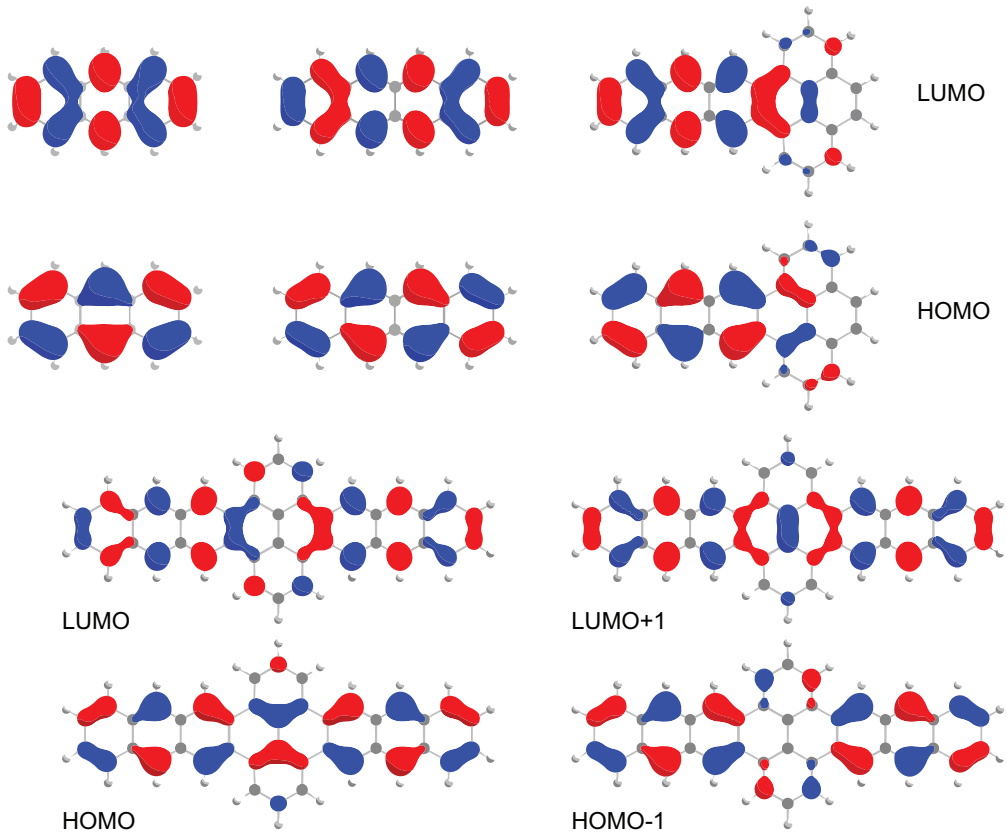


Figure 3.7: Frontier molecular orbitals (isovalue of 0.03 e/Å^3) of anthracene, tetracene, P-A3 and A3-P-A3.

This argument is further confirmed by inspecting the frontier molecular orbitals (FMOs) of, without loss of generality, P-A3 and A3-P-A3. As shown in Fig. 3.7, both the highest occupied molecular orbital (HOMO) and the lowest unoccupied molecular orbital (LUMO) of P-A3 localize heavily at the acene segment and the "junction" ring of pyrene. This, along with the resemblance between HOMO/LUMO of P-A3 and that of tetracene, even that of anthracene to a lesser degree, suggests that the fused pyrene segment, instead of extending the π-conjugation to "five rings in a row", only contributes approximately one additional ring to the anthracene segment. In the case of A3-P-A3, while the HOMO

seems delocalized along the entire molecular backbone, it can also be described as two tail-to-tail tetracene HOMOs, just as the LUMO of A3-P-A3 is similar to two isolated tetracene LUMOs. Considering that the HOMO-1 of A3-P-A3 consists of two isolated anthracene HOMOs, the small energy splitting (35 meV) between the HOMO and HOMO-1 implies limited electronic coupling among the π orbitals of pyrene and acene segments. This scenario is synonymous with a Class II mixed-valence system, as named by Robin and Day,[87,88] where the pyrene acts as a very weak electronic bridge between the two acene (redox center) units. Indeed, such almost degenerate HOMOs also appear in pyrene-azaacene systems, implying the universality of the disruptive impact from pyrene insertion.[68]

3.3.2 Excited States

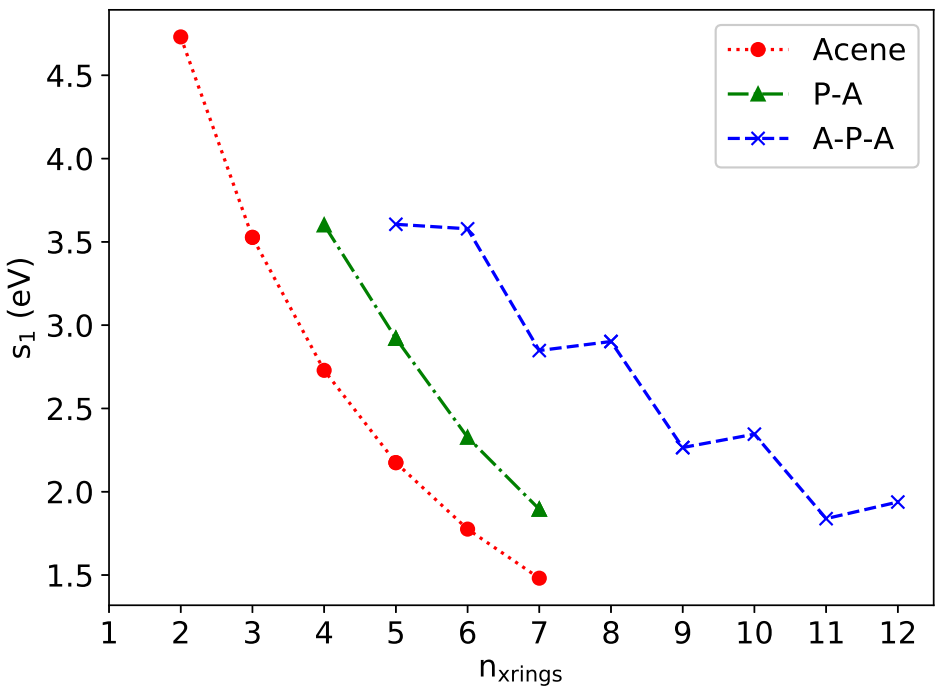


Figure 3.8: ϵ^{S_1} as a function of n_{xrings} for acenes, P-A and A-P-A.

We now turn to the vertical electronic excitations in P-A and A-P-A, which led to our initial suspicions about being a large degree of π-conjugation in these systems. Excitation energies for low-lying singlet states were calculated via linear

response TDDFT, as detailed in the Section 3.2. Fig. 3.8 shows ϵ^{S_1} (excitation energy of the first singlet excited state) plotted against n_{xrings} (the number of rings along the long axis). For acenes longer than naphthalene, the first bright excited state ($^1B_{2u}$) is a singlet with dominant HOMO $\rightarrow$ LUMO transition, whose excitation energy decreases with increasing number of fused rings, in accord with previous studies.[89,90] Approximately, $\epsilon^{S_1}_{\text{PA}}(n_{\text{xrings}})$ can be considered as $\epsilon^{S_1}_{\text{acene}}(n_{\text{xrings}})$ shifted by one along the x-axis, *e.g.* $\epsilon^{S_1}_{\text{P-A3}}$ is close to $\epsilon^{S_1}_{\text{tetracene}}$. This can be understood by previous MO analyses, as the first excited state of P-A is also dominated by HOMO $\rightarrow$ LUMO transition.

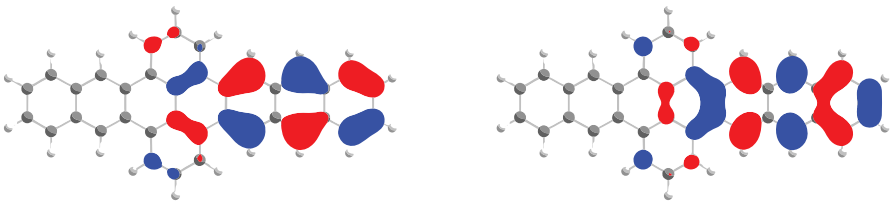


Figure 3.9: The pair of natural transition orbitals (isovalue of 0.03 e/Å^3) for the first singlet excitation in A2-P-A3.

$\epsilon^{S_1}_{\text{A-P-A}}(n_{\text{xrings}})$ is further shifted along the x-axis comparing with $\epsilon^{S_1}_{\text{P-A}}(n_{\text{xrings}})$ as pyrene insertion breaks the molecular backbone into two shorter acene segments. An odd-even oscillation appears in $\epsilon^{S_1}_{\text{A-P-A}}(n_{\text{xrings}})$, as $\epsilon^{S_1}_{\text{An-P-An}}$ is close to or sometimes even slightly larger than $\epsilon^{S_1}_{\text{A(n-1)-P-An}}$. Such oscillation can be explained using A2-P-A3 and A3-P-A3 as examples. For A2-P-A3, the first bright excitation can still be described as a HOMO $\rightarrow$ LUMO transition, with corresponding natural transition orbitals (NTOs) shown in Fig. 3.9 where the electron-hole pair is localized at the side of anthracene segment due to the lower molecular symmetry (C_{2v}). For A3-P-A3 (D_{2h}), there are two dominant excitation configurations of the first excitation, HOMO(g) $\rightarrow$ LUMO+1(u) and HOMO$-$1(u) $\rightarrow$ LUMO(g) with similar coefficients (0.48 and 0.50, see Fig. 3.7 for the MOs), and both of them resemble the HOMO $\rightarrow$ LUMO excitation of anthracene. Thus, the excited state of both A2-P-A3 and A3-P-A3 come from anthracene-like excitations, leading to similar ϵ^{S_1} values. In passing, we note this excitation in An-P-An is always

accompanied by an almost degenerate excitation with equally-weighted configurations HOMO(g) → LUMO(g) and HOMO−1(u) → LUMO+1(u), which is dark due to one-photon absorption selection rules (g → u or u → g).

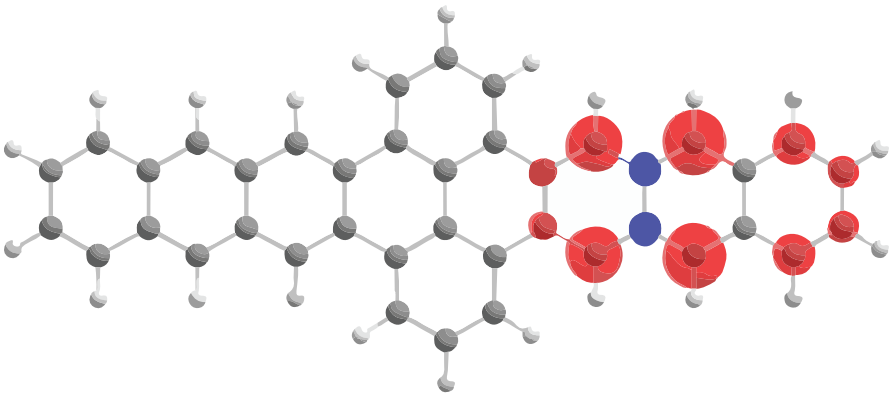


Figure 3.10: Total spin density (isovalue of 0.002 e/Å^3) of A3-P-A3 at the lowest triplet state.

In contrast to a delocalized singlet state, the exciton of An-P-An in its first triplet state is localized at one of the acene segments. Using A3-P-A3 as an example, from BLA patterns and total spin density (Fig. 3.10), the unpaired electrons are localized at the edges of one anthracene segment, while the other anthracene segment remains as in the ground state. The molecular symmetry is further lowered to C_s as two acene segments deviate from the pyrene plane to form a bent butterfly shape. The adiabatic and vertical triplet energies are estimated to be 1.49 eV and 1.51 eV, respectively. Such triplet exciton localization in A3-P-A3 is very similar to the case in platinum-acetylide oligomer,[91,92], where optimized triplet structure features a distortion around platinum that lowered the ground state symmetry.

Spin density localization also applies to the monocation state of A3-P-A3, where structural relaxations and charge are localized at one of the anthracene segments, giving a lower molecular symmetry (C_s). Such localization behaviors may be related to the unusual stability of fused pyrene-azaacene dianions.[73] The electronic ground state of a metastable monocation structure with D_{2h} symmetry (with delocalized spin density) is estimated to be 90 meV higher than that of C_s

symmetry, which can be regarded as the lower bound of charge transfer barrier (from one acene segment to the other). It is thus informative to reconsider the idea of using A-P-A molecular crystals with brickwork packing structure in transistor applications:[93] If the charge tends to localize on the one end of a "brick", the percolation pathways of charge would still be one-dimensional even with brickwork packing.

3.3.3 The Blocker Effect: Origin and Its Implications in Polymers

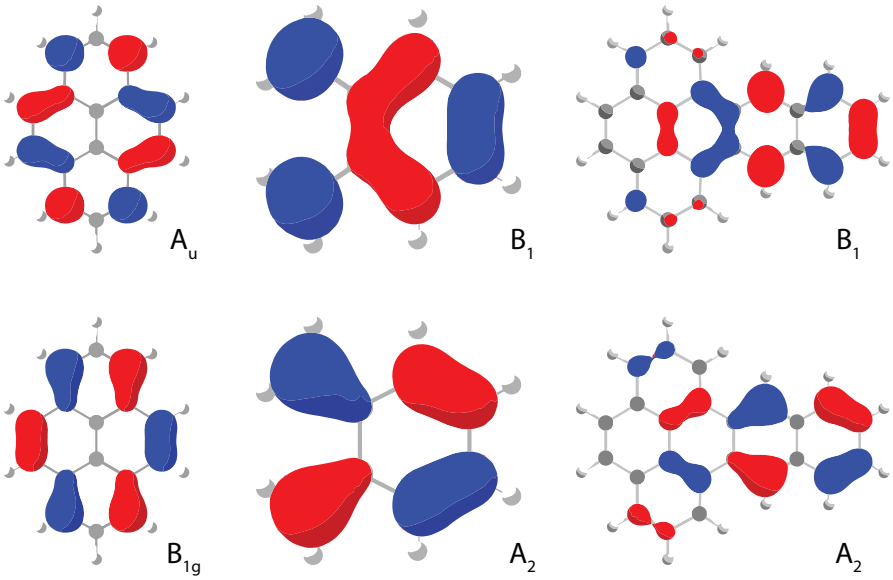


Figure 3.11: (Top) LUMO and (bottom) HOMO of pyrene, 1,2-benzodiene, and P-A2, isovalue was set to 0.03 e/Å^3.

We note pyrene may not be the only PAH that could act as an electronic "blocker" between acenes. In fact, the origin of such blocking effect can be understood by inspecting the molecular orbital (MO) symmetries of each fragment. Fig. 3.11 shows that the FMOs of pyrene and 1,2-benzodiene moieties possess mismatched symmetries, which forbids effective overlap of MOs when they are combined to form P-A2. Thus, fusing a second 1,2-benzodiene to the other side of the pyrene would lead to limited orbital overlap between the naphthalene segments.

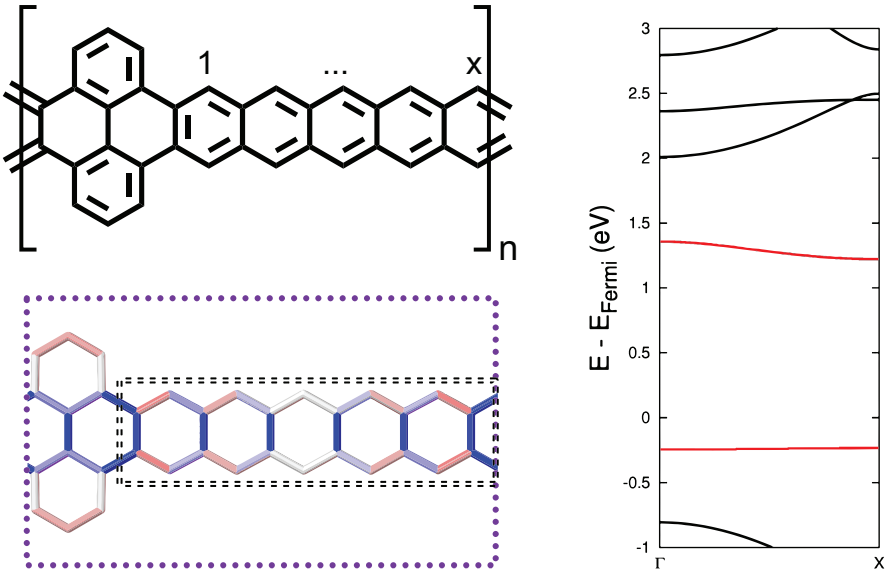


Figure 3.12: Structure of poly-P-Ax (x=5), bond length alternation pattern (purple rectangle represents the repeating unit, while the black box incorporates the pentacene segment), and electronic band structure, where the valence and conduction bands are colored red.

Knowing the nature of pyrene insertion, it is informative to look at the polymer version of pyrene-acene system (poly-P-Ax, *e.g.* poly-P-A5 in Fig. 3.12), which is a subset of delimited polyacenes.[94] While polyacene can be free of Peierls instabilities,[95] thus free of BLA at its edges, BLA pattern of, for example, poly-P-A5 is almost identical to that of P-A5, with conspicuous acene characteristics. The electronic band structure of poly-P-Ax features a flat valence band, of which the charge density localized at the acene segments. Thus, while poly-P-Ax can be a chemically realizable target (stability brought by three Clar sextets per repeating unit) for synthetic chemist, its electronic structure would be similar to a set of oligoacenes, rather than polyacene, thus not suitable for reaching high hole mobility. Interestingly, while the hold mobility of poly-P-Ax vanishes due to infinitely heavy hole, based on the one-dimension deformation potential theory (Section 2.3.1), electron mobility of poly-P-Ax is estimated to be $5 \sim 31$ cm^2/(V $\cdot$ s)for x $= 2 \sim 5$ which is much lower than that of other graphene nanoribbons, but higher than most of molecule-based devices . Following the MO symmetry argument, we found another PAH, dibenzo(bc,kl)coronene (DBC), with whom pyrene shares the

same FMO symmetry. Delimited polyacene formed by DBC insertion indeed provides electronic structure and BLA similar to that of poly-P-Ax, as illustrated in Fig. 3.13.

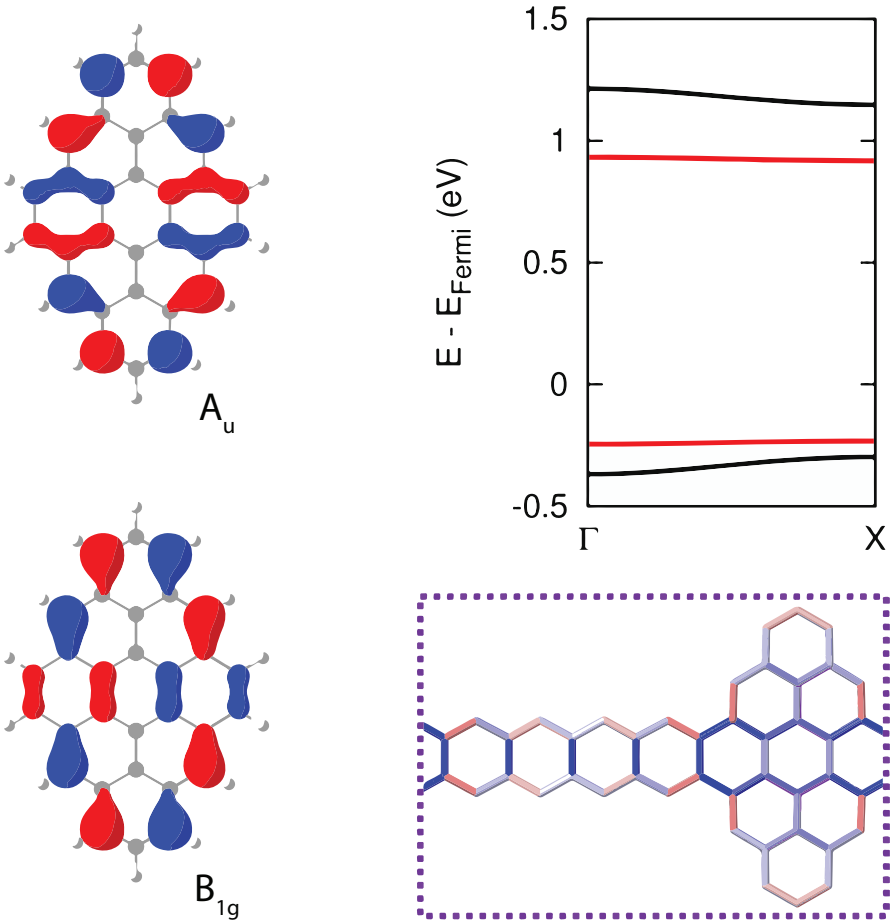


Figure 3.13: Left: HOMO and LUMO (from top to bottom, isovalue of 0.03 e/Å^3) of dibenzo(bc,kl)coronene (DBC). Right: Electronic band structure and BLA pattern of DBC delimited polyacene.

3.4 Conclusion

In this chapter, electronic structures of pyrene-acene systems are investigated with detailed analyses on their structural relaxation, aromaticity, and excited states. It has been shown that the presence of pyrene, as indicated by the local structural relaxations, effectively blocks the electronic communications between acene segments. The implications of this effect for polymers are discussed, and the origin of such effect can be traced back to molecular orbital symmetry

of pyrene FMOs, which is evidenced by DBC delimited polyacene system. These results provide insights on large acene-based conjugated systems, as well as narrow graphene nanoribbons from bottom-up synthesis.

CHAPTER 4

INTRAMOLECULAR S…F INTERACTIONS: THE IMPACT ON MOLECULAR CONFORMATION

4.1 Introduction

Fluorine substitution has emerged as a popular method to alter molecular and electronic structures and the resulting solid-state molecular packings of organic semiconductors.[96,97] The rationale reported from previous studies rely on, besides fluorine being a strong electron-withdrawing agent, the existence of weak but persistent, stabilizing interactions involving fluorine.[98,99] Such noncovalent interactions, usually hydrogen bonding or halogen bonding,[100,101] are able to support a rigid conjugated backbone, which in turn can facilitate consistent, ordered solid-state molecular packings and impact charge-transport process. To explore the chemical space for high-performance organic semiconductors, it is thus necessary to understand the nature and strengths of fluorine-related interactions that are important or, in some cases, determinant of molecular conformation.[102]

Despite the variety of fluorine-related noncovalent interactions in organic molecules and crystals, introducing intramolecular sulfur…fluorine (S…F) interactions has been frequently adopted in designing semiconducting polymers,[43,44] where such interactions are believed to reduce rotational disorder along polymer chain. The success of this strategy, however, is not transparent as the stabilizing nature of S…F interaction remains unclear.[103] The S…F pair consists of two electronegative atoms, where, from an electrostatic perspective, attractive interactions are unlikely to exist. On the other hand, the two atoms are often introduced as substituents on π-conjugated units, thus the stabilizing effect is also subject to the nature of these conjugation systems.

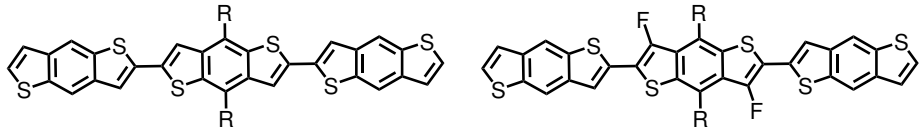
In this chapter, intramolecular S…F interactions in two fluorinated chromophores, namely benzodithiophene (BDT) trimers[104] and EDOT-ph-EDOT[105] (EDOT = 3,4-ethylenedioxythiophene, ph = phenylene), are studied based on computational results. It is revealed that while in BDT systems the S…F interaction features a stabilizing effect to promote a more rigid, planar conjugate backbone, in the case of EDOT-ph-EDOT, in addition to S…F interaction, fluorination on the phenylene unit can also significantly alter the strength of nearby intramolecular O…H interaction (hydrogen bond), which has a greater impact on altering the molecular geometries and packing structures. These cases demonstrate two distinct mechanisms through which conformation is influenced by S…F interactions, thus providing insight as to how to make use of these noncovalent interactions to design organic semiconductors. An earlier version of some parts of this chapter was published in *Chemistry of Materials* **2019** *31*(17), 7070-7079, in collaboration with Dr. Peter Skabara's group at University of Glasgow.

4.2 Methods

Density functional theory (DFT) calculations were carried out on isolated molecules and bulk crystal structures. For the isolated molecules, structures were computed at the ωb97xd/jun-cc-pvdz level of theory as implemented in *Gaussian 16*.[74,106,107] All optimized, unconstrained geometries were confirmed as minima on the potential energy surface through normal mode analyses. Potential energy surface (PES) scans, where the interunit dihedral angles were held constant and the remainder of the molecules allowed to relax, were also performed. Atoms-in-molecules (AIM) analyses were carried out with topological analysis module in *Multiwfn* package.[108,109]

4.3 Results and Discussion

4.3.1 The case of BDT trimer: Stabilizing S…F interactions



(a) BDT-3 *(b)* F-BDT-3

Figure 4.1: Chemical structures of BDT-3 and F-BDT-3.

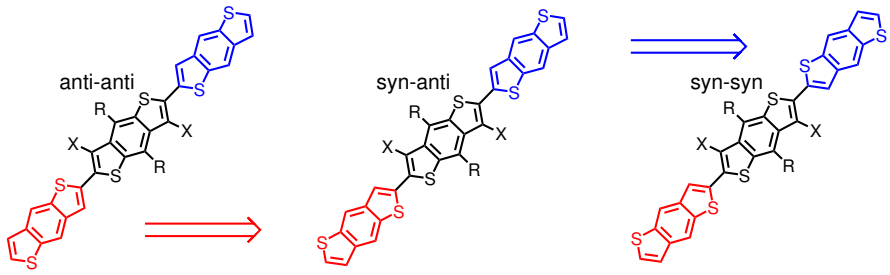
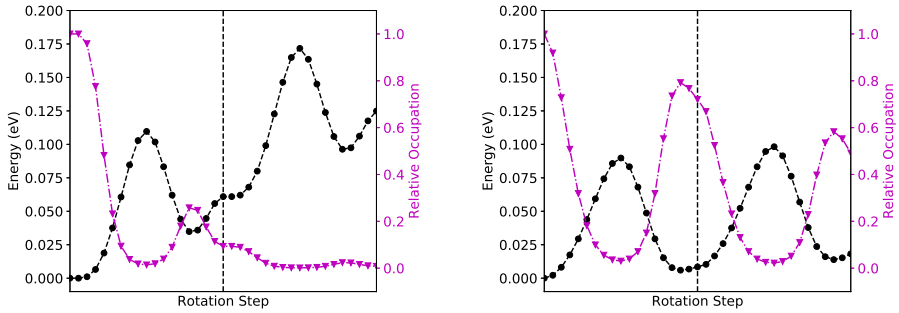


Figure 4.2: Multiple stable conformations of BDT trimer, X=H, F.

The benzodithiophene (BDT) trimer (BDT-3, Fig. 4.1a) has been proposed as a "universal crystal engineering core" due to its simple and scalable preparation method and amenability to addition of functionalized alkynes.[104] While the hole mobility of BDT-3 based field-effect transistors (FETs) can reach as high as 1.6 $cm^2/(V \cdot s)$, the presence of a large amount of (static) conformational disorder (the *syn/anti* disorder from the rotation between BDT units, as shown in Fig. 4.2) in the crystal structures can be a challenge for FET applications as such disorder may lead to an increase in charge trapping, therefore lower device performance.[110] In an attempt to mitigate the disorder in BDT-3 and further improve material performance, fluorine substitutions were introduced in the central BDT unit (F-BDT-3, Fig. 4.1b). Interestingly, while the device performance of fluorinated materials tends to be better than nonfluorinated ones (if preferred molecular packing preserves), the S…F interaction does not always promote a specific backbone conformation, *e.g.* both *syn-syn* and *anti-anti* conformations can both be found in the major configuration of solid-state F-BDT-3. In this section, we present the

computational analyses regarding S…F intramolecular interaction in F-BDT-3 and the implications of fluorination was discussed.



(a) X=H *(b)* X=F

Figure 4.3: PES of (F-)BDT-3 (R=H) defined by dihedral scan. The relative occupation is computed by assuming room temperature (298.15 K) assuming Boltzmann distribution of the conformations. The rotation was carried out as shown in Fig. 4.2: The first data point, the vertical dashed line and the last data point represent perfectly planar *anti-anti*, *syn-anti*, *syn-syn* conformations, respectively.

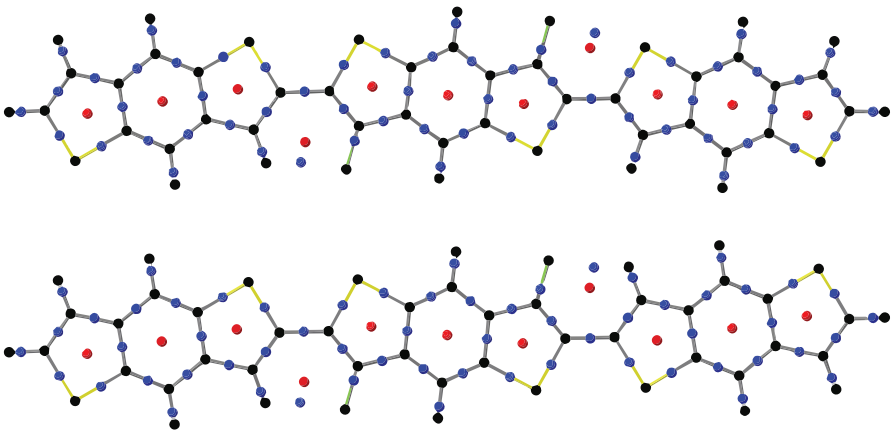
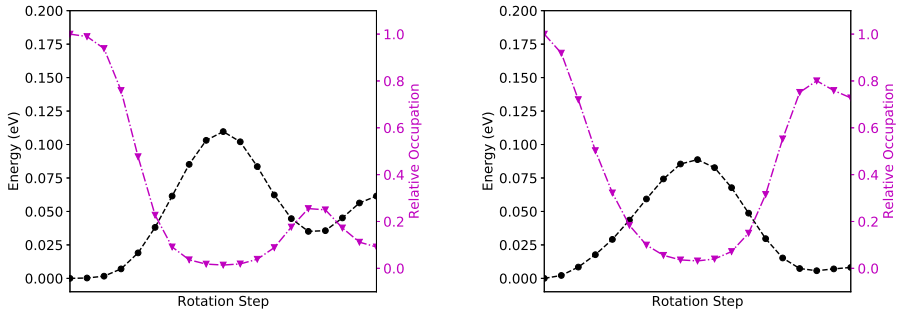


Figure 4.4: Critical point analysis for (top) *syn-anti* conformation and (bottom) *syn-syn* conformation. The NCPs, RCPs and BCPs are colored black, red and blue, respectively. Note BCPs and RCPs between F and H in both conformations.

PES line scans for BDT-3 and F-BDT-3, defined by the dihedral rotation of the C-C bonds connecting BDT units, were performed along the transformation in Fig. 4.2, and the results are shown in Fig. 4.3. For each scan profile, three obvious local minima, corresponding to three quasi-planar conformations, are separated by energies comparable to $k_{\mathrm{B}}T$ at room temperature (k_{B} is the Boltzmann constant, T the temperature). Such small differences between (meta)stable conformations

are related to potential static disorder in solid-state, as observed in the crystal structures of several BDT-3.[104] Notably, *syn-anti* and *syn-syn* conformations are stabilized (using the planar *anti-anti* conformation as the reference), and become more planar after fluorine substitution at the central BDT unit. This can be understood by the formation of F…H hydrogen bond(s) in *syn-anti* and *syn-syn* conformations. Fig. 4.4 shows the electron density critical points (EDCPs) in *syn-anti* and *syn-syn* conformations of F-BDT-3 based on DFT results. EDCPs are points in three-dimensional space where the gradient norm of electron density vanishes. They can be classified, by the Hessian matrix of electron density, into four categories, among which $\mathbf{(3,-3)}$ (nuclear critical point, NCP), $\mathbf{(3,-1)}$ (bond critical point, BCP) and $\mathbf{(3,+1)}$ (ring critical point, RCP) EDCPs are present in Fig. 4.4. In *syn-anti* conformation, two pairs of BCP-RCP appear between BDT units, indicating F…S interaction and F…H hydrogen bond, respectively, while in *syn-syn* conformation such BCP-RCP pairs represent two F…H hydrogen bonds.



(a) X=H *(b)* X=F

Figure 4.5: PES of (F-)BDT-2, which are 2-unit systems derived from BDT-3, defined by dihedral scan from perfectly planar *anti* to perfectly planar *syn* conformation. See the caption of Fig. 4.3 for instructions.

We now turn to the rigidity of BDT-3 backbone, that is, the stability of PES local minima against dihedral rotation of bonds connecting BDT units, which is related to dynamic disorder in crystals. While the transition barrier between local minima decreases slightly ($<$ 25 meV) after fluorination, we note that the rigidity should be measured by the curvature at PES local minima. In order to estimate the impact of S…F interaction on backbone rigidity, a dihedral scan similar to

Table 4.1: Curvature (10^{-5} eV/degree^2), relative energy (meV), short contact between BDT units, and dihedral angle (in degrees) between BDT units for all conformations (local minima in Fig. 4.5) of BDT-2 and F-BDT-2. The relative energy of a conformation was calculated against the energy of *anti* conformation

Molecule	Conformation	Curvature	Energy	Contact	Dihedral
BDT-2	*anti*	1.09	0	None	0.07
F-BDT-2	*anti*	4.08	0	S…F	0.02
BDT-2	*syn*	10.06	35	None	38.51
F-BDT-2	*syn*	2.99	6	H…F	19.32

that shown in Fig. 4.3 is performed for BDT dimer (BDT-2) and fluorinated BDT dimer (F-BDT-2) such that the interference from the other dihedral can be disregarded. The results of the dimer dihedral scans (Fig. 4.5), where the dihedral angle between BDT units was changed from 0°(perfectly planar *anti*) to 180°(perfectly planar *syn*), are summarized in Table 4.1. The order of curvature (rigidity) at local minima follows BDT-2 *syn* > F-BDT-2 *anti* > F-BDT-2 *syn* > BDT-2 *syn*, which suggests (1) comparing with BDT-2, the S…F interaction is able to lock the dihedral between BDT units in the *anti* conformation, and so is H…F hydrogen bond for *syn* conformation, to a less degree, (2) H…F hydrogen bond is able to, to some extent, planarize two BDT units. While BDT-2 *syn* may be less sensitive to rotational disorder, it should be noted that the relative energy is approximately $1.4k_{\mathrm{B}}T$ and the dihedral angle at equilibrium is rather large (38.51°), which is not desired for most semiconducting applications.

In summary, computational results from PES scans and AIM analyses revealed that noncovalent F…S interactions were introduced in fluorinated BDT oligomers (with *syn* conformation for dimer, and *syn-syn* or *syn-anti* for trimer), which leads to more planar, rigid molecular conformations at PES wells. On the other hand, for BDT-3, *syn-anti* and *syn-syn* conformations are stabilized (with respect to *anti-anti* conformation) after fluorination through the formation of F…H bonds, which could promote static disorder in solid-state. We stress that such results should not be directly translated to solid-state where intermolecular

interactions could also affect molecular conformations.

4.3.2 The case of EDOT-ph-EDOT: The illusion of stabilizing S…F interaction

3, 4-ethylenedioxythiophene (EDOT) is a popular moiety for high-performance functional polymers and a frequent subject for research on noncovalent chalcogen-chalcogen interactions.[98,111] For example, in the polymeric form of EDOT, PEDOT features a highly planar structure originates from intramolecular S…O interactions,[99] leading to low optical gap (1.63 eV) and excellent charge transport performance. In this section, we focus on a series of fluorinated EDOT-ph-EDOT and ph-EDOT (ph = phenylene),[105] as shown in Fig. 4.6, with intention to probe stabilizing S…F intramolecular interactions, which could be used as a design strategy in developing functional organic/polymer materials containing EDOT. To our surprise, instead of forming strong, stabilizing F…S interaction, fluorine substitutions impact the molecular conformation via strengthening the dominating O…H hydrogen bond. This case study highlighted the potential unexpected consequences of introducing F…S interactions and emphasized the importance of a holistic view in molecular design.

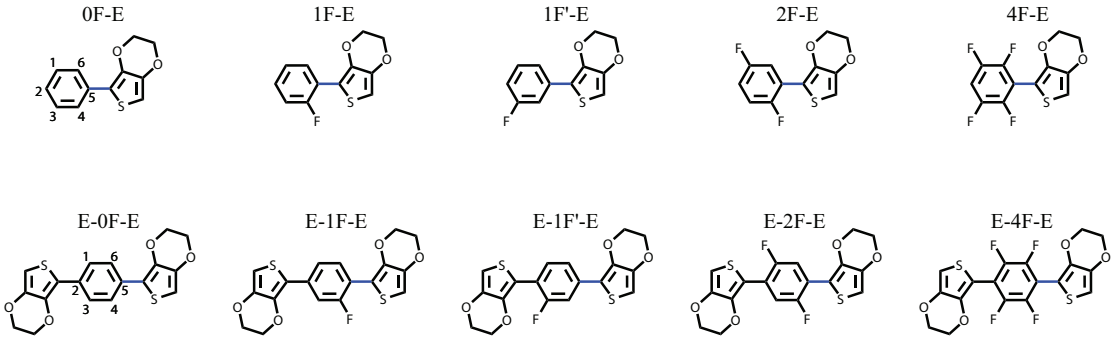


Figure 4.6: Chemical structures of (fluorinated) EDOT-ph-EDOT molecules studied in this section. Five two-unit molecules and five three-unit molecules are shown in top and bottom panel, respectively. The bond related to PES scans (Fig. 4.7 and Fig. 4.8) is denoted by blue color. The sites on phenyl ring are labeled with Arabic numbers for discussion.

To evaluate the impact of fluorine substitution on the molecular conformation, it is illustrative to examine the PES as a function of the dihedral angles among the aromatic units; here, the PES scans were performed by twisting EDOT around the bond between EDOT and the phenyl ring (the blue bonds in Fig. 4.6). The results are shown in Fig. 4.7 and Fig. 4.8 for two-unit and three-unit systems, respectively. Comparing two sets of scan profiles, we noticed that adding an additional EDOT unit only introduce trivial changes. Thus we will focus on the two-moiety structures with their PES scans shown in Fig. 4.7, where two half-scans are performed for each structure based on twisting direction; the yellow dots in the figure represent the minimum position for each half-PES scan. Notably, there is a rapid rise in the PES for 1F-E, 2F-E in the second half-scan, and 4F-E in both half-scans that can be directly attributed to repulsive O…F interactions. In general, fluorine substitution at positions labeled **4** or **6** on the phenyl ring reduces the rigidity of the dihedral torsion. Interestingly, the local minimum in the first half-scan of 1F-E is less stable than that in the second half-scan, where no direct S…F interaction can be established. Hence, instead of acting as an agent to "lock" the structure into a quasi-planar configuration, such substitution tends to reduce the transition barrier among the local minima, as well as the curvature at the local minima on the PES. The only exception is 1F′-E, where fluorination, without establishing S…F interaction, indeed stabilized the local minima with conformations more planar than that of 0F-E. This, as will be shown later, is primarily due to strengthened O…H bond between EDOT and the phenyl ring.

AIM analyses of the noncovalent interactions confirms the existence of S…F BCP at the scan local minima of 1F-E and 2F-E (selected parameters of local minima are presented in Table 4.2, no such BCP is found for 1F-E and 4F-E due to large S…F distance). The S…F distances in these situations are smaller than the sum of their van der Waals radii (3.27 Å). The nature of these interactions is revealed by reduced density gradient (RDG) analyses (Fig. 4.9), where both weak attractive and repulsive interactions are identified by the sign of λ_2, the second

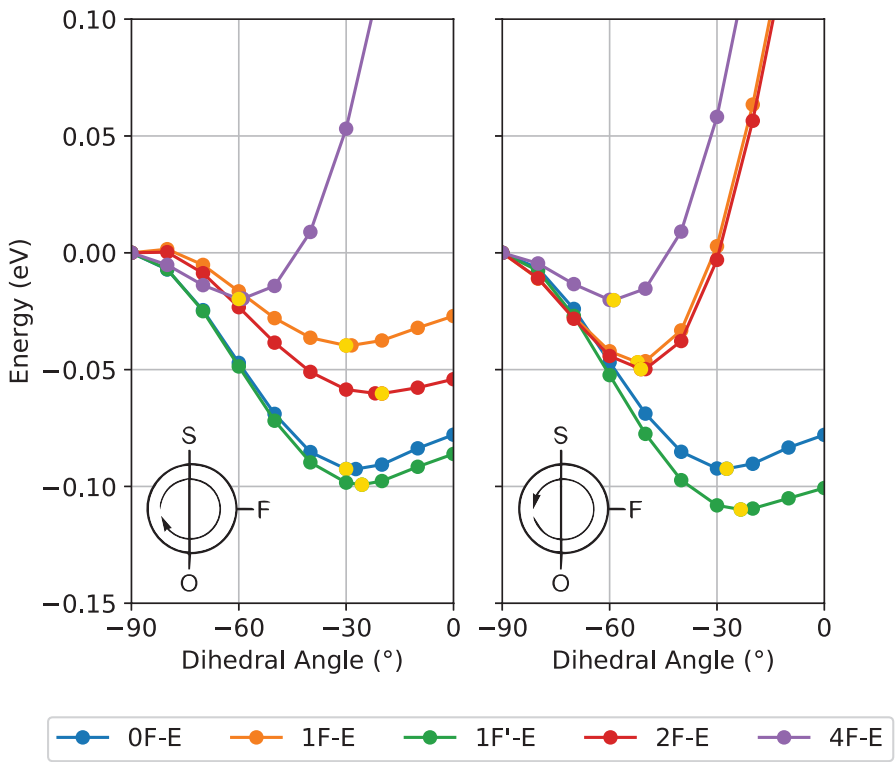


Figure 4.7: PES scans of five two-unit structures where the dihedral angle between EDOT and phenyl ring is being twisted from −90°to 0°. Two subplots represent two different twisting directions as shown in the inset schemes where F represents the fluorine (if any) which is the closest to EDOT unit. Local minima are denoted by yellow dots.

largest eigenvalue of electron density Hessian matrix.[112] The existence of such repulsive interactions, indicative of often-referred steric effects, is in agreement with the RCP found at the center of C-C-C-F…S ring. While previous studies have discovered S…F interactions as halogen-type bonds,[113–115] we note that in these EDOT-ph structures the S…F interaction is limited and the charge depleted region around sulfur does not point towards fluorine, as illustrated by the electron density Laplacian (Fig. 4.10).

On the other hand, fluorine atoms substituted at positions labeled **1** or **3** on the phenyl ring result in more rigid structures. Here, it is apparent, however, that the O…H interaction gains prevalence. The strength of this interaction is significantly affected by the bond acidity of the C-H bond, as the electronic density at the BCP increases significantly after fluorine atoms are introduced onto the phenyl ring. By defining $|V_H|$ as the electrostatic potential at H without the

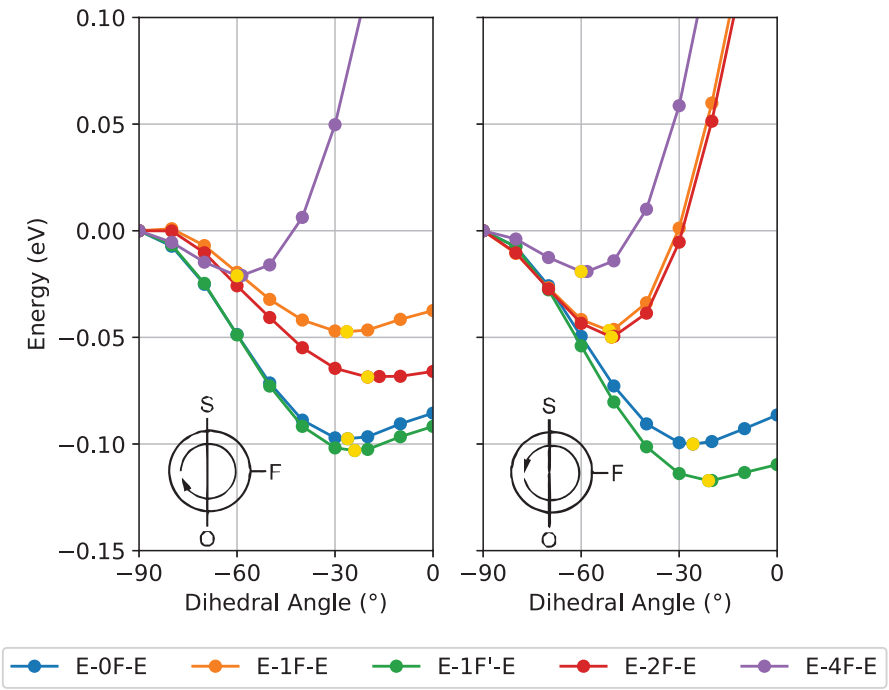


Figure 4.8: PES scans of five three-unit structures where the dihedral angle between EDOT and phenyl ring is being twisted from -90°to 0°. Two subplots represent two different twisting directions as shown in the inset schemes where F represents the fluorine (if any) which is the closest to EDOT unit. Local minima are denoted by yellow dots.

Table 4.2: Parameters of the minimum configuration for two-unit structures, including $\boldsymbol{D}$ (the dihedral angle between two aromatic units in degrees), $\boldsymbol{\rho}$ (the electron density at BCP in 10^{-2} e/Bohr^3), $\boldsymbol{d}$ (the distance between a pair of atoms in Å), $\boldsymbol{\Delta}E$ (the energetic difference between this configuration and scan origin in meV) and $|V_H|$ (the electrostatic potential (in Ha/e) at H without the contribution from the H nucleus of the hydrogen atom involved in the O…H interaction).

Label	**Half-scan**	$\boldsymbol{D}$	$\boldsymbol{\rho}$(O…H)	$\boldsymbol{d}$(O…H)	$\lvert V_H \rvert$	$\boldsymbol{\rho}$(F…S)	$\boldsymbol{d}$(F…S)	$\boldsymbol{\Delta}E$
0F-E	1st	-27.3	1.15	2.40	1.0988	-	-	92
1F-E	1st	-28.5	1.34	2.32	1.0910	1.32	2.87	40
1F′-E	1st	-25.6	1.21	2.37	1.0924	-	-	99
1F′-E	2nd	-23.1	1.29	2.34	1.0815	-	-	110
2F-E	1st	-21.9	1.56	2.24	1.0743	1.44	2.81	60
4F-E	1st	-58.9	-	-	-	-	-	14

contribution from the H nucleus itself, we can evaluate the pK_a of C-H bond, which is inversely correlated with $|V_H|$.[116] Interestingly, while generally a larger C-H bond acidity yields a stronger O…H interaction (reflected by ρ(O…H), charge density at BCP), E-1F seems to be an outlier with a larger $|V_H|$ but also a larger ρ(O…H) comparing to E-1F′ (in the second half-scan). This could be addressed by steric

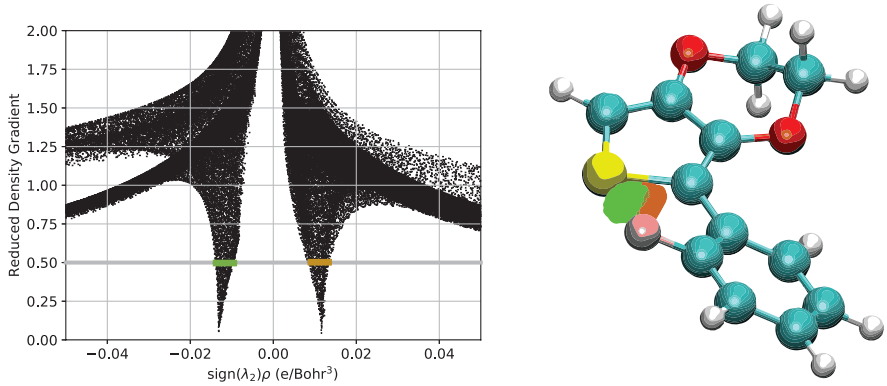


Figure 4.9: Left: RDG isosurface with RDG=0.5 that is colored based on the sign of λ_2 as shown in the scatter plot. Right: Scatter plot of the RDG values for the S…F interaction versus $sign(\lambda_2)\rho$, where λ_2 is the second largest eigenvalue of electron density (ρ) Hessian matrix. The data is collected in a 3 Bohr × 3 Bohr × 3 Bohr box centered at the midpoint between F and S, evenly sampled with 216000 grid points. The horizontal line represents the isovalue that is shown in the RDG isosurface.

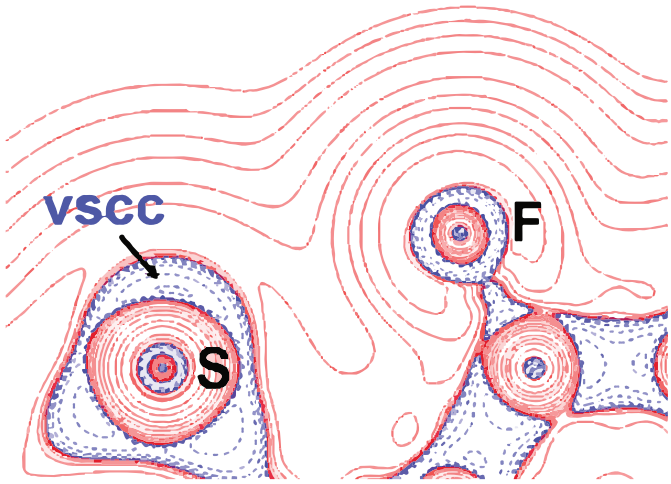


Figure 4.10: Contour map of electron density Laplacian in the S…F region. Contour lines are colored red/blue for negative/positive values. The charge depletion region on sulfur is labeled (valence shell of charge concentration, VSCC).

effects due to S…F interactions in E-1F that partially bend the structure, leading to a closer O…H pair (Fig. 4.11). Here, one can surmise that the relative structural rigidity of the two-moiety structures comes from strong O…H interactions, and not necessarily from the S…F interactions.

Finally, knowing the impact of fluorination on O…H pair, it is informative to

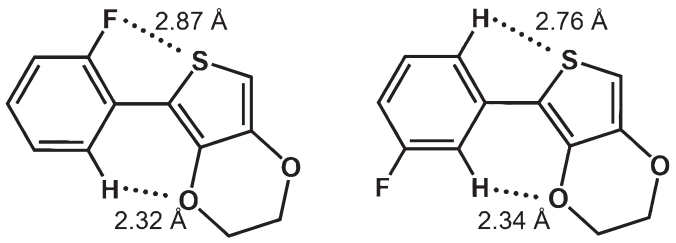


Figure 4.11: Structural features in the optimized structures of (left) E-1F and (right) E-1F′ in the second half-scan.

revisit the PES scans in Fig. 4.7 and Fig. 4.8. The scan profiles of 1F′-E and E-1F′-E (green profiles in Fig. 4.7 and Fig. 4.8) can be understood via the strengthened O…H hydrogen bond, which stablize/planarize the conformation at local minima with respect to that of non-fluorinated molecules. Such understandings can then be extended to the scan profiles of 2F-E and E-2F-E, where a combination of mostly repulsive F…S interaction and strong O…H hydrogen bond (due to difluorination) resulted in a planar conformation that is less rigid than 0F-E. These findings suggest that introducing fluorine to conjugated systems can have multifaceted impact that cannot be simply described as "formation of attractive F…S interactions".

4.4 Conclusion

In this chapter, two case studies were conducted to examine the impact of introducing fluorine substitution and S…F intramolecular interactions in conjugate units commonly adopted in functional material design. In the case of BDT oligomer, after fluorination at the central BDT unit, stabilizing S…F interactions, which lead to a more planar, rigid molecular backbone, were identified. However, fluorination also stabilized *syn-anti* and *syn-syn* conformations through the formation of F…H hydrogen bond. In the case of EDOT-ph-EDOT, however, instead of forming strong, directional S…F interactions, fluorination at the phenyl ring increased the strength of O…H hydrogen bonding on the other side of phenyl ring, which supported a rigid, planar conformation. These results illustrated the

diverse conformational consequences of constructing S…F interactions in organic molecules, and the possibilities of forming/strengthening other noncovalent interactions upon fluorination, thus provided insights for rational molecular/polymer design towards target properties.

CHAPTER 5

ALKYL CHAIN INDUCED STRUCTURE ALTERNATION: CRYSTAL STRUCTURES OF 2,5-BIS-(3-ALKYLTHIOPHEN-2-YL) THIENO[3,2-B]THIOPHENES

5.1 Introduction

Though advances in molecular and materials design has contributed to continued improvements in organic semiconductor performance, the ability to control molecular packing, in molecular or polymeric materials, remains a grand challenge. Such capacity, known as crystal engineering for molecular crystals, is crucial as the electronic and optical properties of organic semiconductors are determined by the packing pattern of π-conjugated chromophores. Among various methods in crystal engineering, it is believed that the alkyl side-chains, while being electronically inert, serve as key contributors to govern the solid-state molecular packing, making them important "handles" in design of functional organic semiconductors.[117–120]

Alternating pattern in material properties can frequently be found for molecular crystals with normal alkyl groups, such as (end-substituted) *n*-alkanes and alkanedicarboxylic acids.[121,122]. Such alternation in material properties is often accompanied by alternation in crystal structures. For example, the melting point alternation of *n*-alkanedithiols originates from varied packing densities of their crystal structures.[123] While research on this phenomenon is reinvigorated by crystallographers thanks to development in instrumentation and characterization techniques,[124–126], little attention has been paid to this effect in the community of organic semiconductor research: Researchers rarely synthesized materials as a homologous series, with those that do often overlooking the possibility of a property alternation. The seminal work of Ebata *et al.* on Cx-BTBT ([1]benzothieno[3,

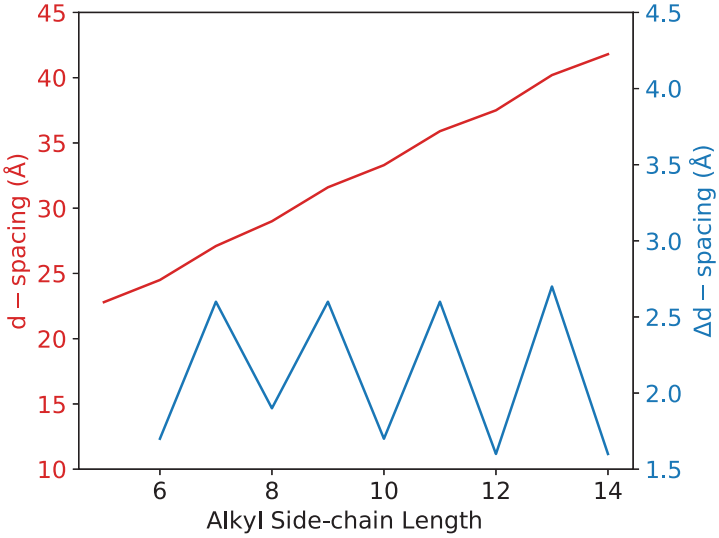


Figure 5.1: Out-of-plane d-spacing of a series of BTBT molecules with systematically varied side-chain length. The nonlinear increase in spacing is highlighted in the d-spacing difference, demonstrating a molecular packing alternation. Data adopted from the study by Ebata *et al.*.[28].

2-b]benzothiophene) molecules with $x = 5 - 14$ (x is the number of methyl units in the alkyl side chains) is one such example.[28] The authors noted that:

> "...d-spacing of the other (BTBT) derivatives depended on the length of the alkyl groups: With longer alkyl groups, larger d-spacings were obtained, indicating that all the derivatives take similar molecular packing structures."

and extrapolated the crystal structure of C12-BTBT to all other molecular crystals. From their d-spacing data, however, a distinct even-odd oscillation pattern can be observed by plotting d-spacing against x in Cx-BTBT, as shown in Fig. 5.1.

In this chapter, we examined the crystal structures of a series of 2,5-bis-(3-alkylthiophen-2-yl) thieno[3,2-b]thiophenes (BTTT-Cx, with $x = 7 \sim 12$ denoting the number of carbon atoms in the alkyl side chains, see Fig. 5.2) to investigate the effect of side-chain length, atom count (even versus odd number of methyl groups), and terminal methyl arrangement on molecular packing. While all molecules form layered structures with BTTT chromophores exhibiting intralayer herringbone

packing, interlayer packing pattern of BTTT-Cx is a function of the parity of x. This observation, along with accompanied melting point alternation, was associated with distinctive interlayer interactions, as highlighted by density functional theory (DFT) investigations and geometric analyses. Finally, we comment on the observed intralayer disorder, which can be related to irregularities in interlayer stacking. The BTTT-Cx series present an extreme case of this crystal packing alternation, highlighting side-chain alternation as an underexploited handle to modify molecular packing.

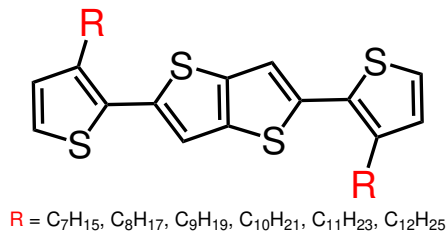


Figure 5.2: Chemical structures of BTTT-Cx.

This work was done in collaboration with the group of Dr. Alejandro Briseno (University of Massachusetts, Amherst), who carried out the molecular and materials synthesis and characterization, and Dr. Sean Parkin (University of Kentucky), who carried out x-ray diffraction experiments for crystal structure determination. An earlier version of some parts of this chapter was published in *Chemistry of Materials* **2019** *31*(17), 6900-6907.

5.2 Methods

5.2.1 DFT calculations for lattice energies

Density functional theory (DFT) calculations for lattice energies DFT calculations were carried out with Vienna *Ab-initio* Simulation Package,[84,127,128] making use of the Perdew, Burke, and Ernzerhof (PBE) exchange-correlation functional.[82] The electron-ion interactions were described with the projector augment wave (PAW) method.[83] The kinetic energy cutoff for the plane-wave basis set was

set to 520 eV, and a Gaussian smearing with a width of 50 meV was employed. The DFT-D3 correction was employed with BJ-damping to describe the dispersion forces.[129,130] The convergence criterion of the total energy was set to 10^{-5} eV in the self-consistent field loop, and that of forces during relaxation was set to 0.001 eV/Å. The Brillouin-zone was sampled with a $4 \times 4 \times 1$ Γ-centered grid. For 2D structures (molecular monolayers), 15 Å vacuum is inserted into the unit cell to block interlayer interactions. The tetrahedron method with Blöchl corrections is used for total energy calculations based on pre-relaxed structure. The lattice energy of a periodic structure S, E_S, is defined by $E_S = -(E_S^{\mathrm{PBE}} - N \cdot E_{\mathrm{mol}}^{\mathrm{PBE}})/N$, where E_S^{PBE} is the total energy per unit cell from DFT calculations, $E_{\mathrm{mol}}^{\mathrm{PBE}}$ is the total energy of a single molecule, N is the number of molecules per unit cell.

5.2.2 Molecular surface

For the DFT relaxed structures, we used the solvent excluded surface to describe the morphology of the interlayer interfaces and to carry out the geometric analyses. Surfaces were built with a 2 Å solvent probe, which is the smallest probe for constructing a pinhole-free monolayer surface, as implemented in *Jmol*.[131,132] One should note that, to correctly calculate the height of a point on these surfaces (defined as H_P in main text), it is necessary to use fractional coordinates as the stacking direction may not be collinear with the z-axis in the Cartesian basis.

5.3 Results and Discussion

5.3.1 Even-odd oscillation regulated by chain-length

We start our analysis by classifying the crystal structures of BTTT-Cx. As shown in Fig. 5.3, oscillation in (interlayer) molecular packing was observed in the crystal structure of BTTT-Cx, as x, the number of carbon atoms in the alkyl side chains, oscillates between odd and even numbers. With $x = 7, 9, 11$,

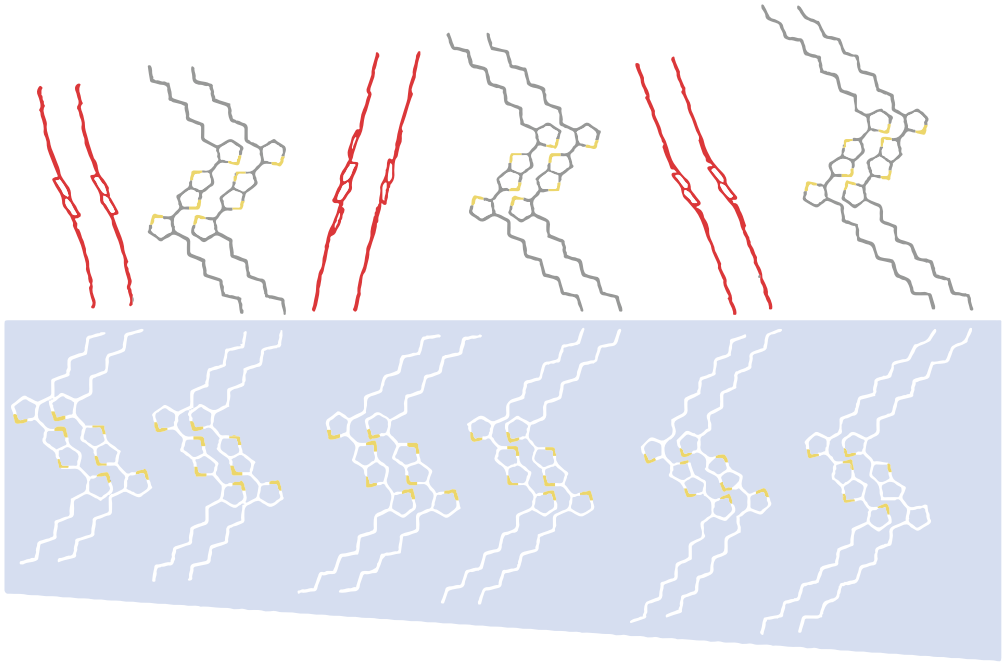


Figure 5.3: Crystal packing motifs of, from left to right, BTTT-C7, BTTT-C8 (major), BTTT-C9, BTTT-C10, BTTT-C11, and BTTT-C12. The red molecules demonstrate the type **I** stacking as defined in the main text.

from one layer to its adjacent layer, BTTT-Cx molecules undergo an approximately 90°rotation along the layer stacking direction. This packing configuration is referred to as type **I** packing. For $x = 8, 10, 12$, however, magnitude of such rotation is approximately 180°, a situation referred to as type **II** packing. One should note that such classification is not based entirely on the symmetry operations connecting the adjacent layers, as among crystals with type **I** packing, two different interlayer arrangements (**I**a and **I**b) were observed. In BTTT-C7 and BTTT-C11, the monoclinic crystals have adjacent layers related by a 2_1 screw axis along the stacking direction (**I**a, space group $P2_1/c$), while for BTTT-C9 the operation is a c-glide (**I**b, space group $C2/c$). Regardless of this differentiation, the interlayer interfaces of type **I** crystals are almost identical to each other. We also note that we later found a few exceptions for crystals of BTTT-C8, which showed a similar c-glide operation as in type **I** packing; this observation was not found for any other even chain length BTTT variant. Nevertheless, the even alkyl side-chain BTTT-Cx tend to show a preference for type **II** packing over type **I**.

Oscillations in the melting point were also identified, with the BTTT-Cx

containing odd-numbered alkyl chains tending to have higher melting points than those even-numbered alkyl side chains.[133] We note that even-odd melting point alternations are well documented for n-alkanes,[121] where, however, the oscillating pattern was opposite (even-numbered alkanes generally melting at higher temperatures than odd-numbered alkanes).

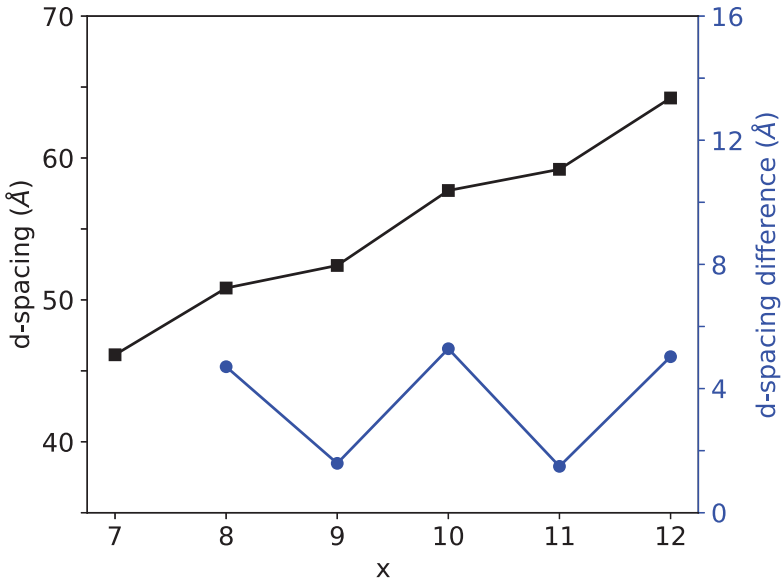


Figure 5.4: Along the layer stacking direction, (black) the d-spacing of the BTTT monomer crystals as a function of chain length, and (blue) the d-spacing difference with the addition of one carbon.

Differences in the terminal ethyl group orientations of the alkyl chains correlate with the crystal packing and material property alternations. Here, we define V_{ethyl} as the vector pointing from the penultimate carbon to the last carbon in the side chain. V_{ethyl} of BTTT-Cx tend to be perpendicular to the stacking direction when x is odd, while the opposite is true when x is even. The crystal d-spacing of the plane defined by the layer stacking direction, derived as a function of increasing side-chain length (Fig. 5.4), shows an alternation pattern similar to the Cx-BTBT case (Fig. 5.1).[28] The relatively tighter interlayer packings for the type **I** configurations, in turn, are reflected by higher melting temperatures when compared to their type **II** counterparts.

5.3.2 Origin of the oscillation: Interlayer interactions

To dig deeper into the underlying interlayer interactions of the effects due to the even or odd alkyl side chains, lattice energies were calculated for both three-dimensional (3D) crystals and two-dimensional (2D) periodic monolayers (extracted from the crystals) via periodic DFT. Here, we use the lattice energy differences between respective 3D and 2D structures as a descriptor to estimate the strengths of interlayer interactions. The intralayer interactions are much stronger than the interlayer interactions (Fig. 5.5), a result expected due to the considerable intralayer overlap of the π-conjugated backbones and alkyl chains. Interestingly, while the 3D and 2D intralayer interactions increase almost linearly with respect to the side chain lengths, there is a definite even-odd oscillation in the interlayer interaction.

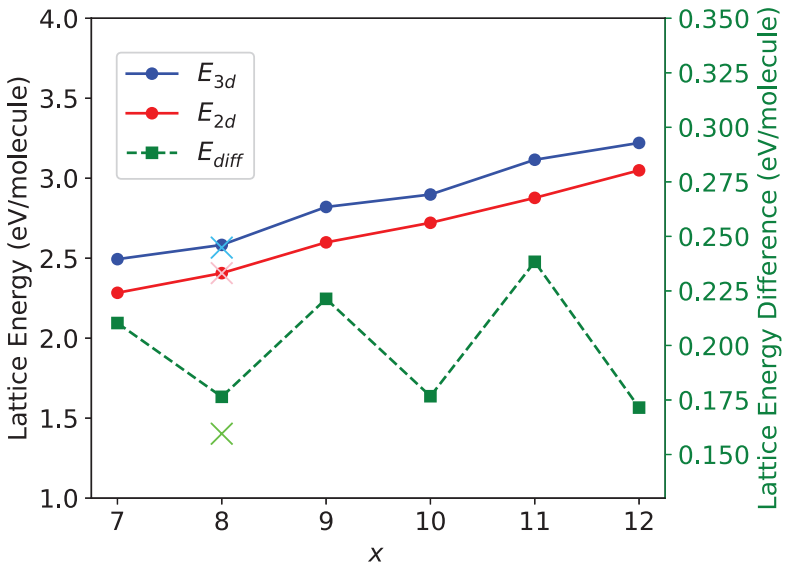


Figure 5.5: Lattice energies of BTTT-Cx crystals. Blue circles denote the lattice energies of BTTT-Cx crystals, red circles denote that of monolayer BTTT-Cx, and green squares denote the difference between them. For BTTT-C8 with the type **I** packing, such quantities are represented by light blue, pink and light green crosses, respectively.

The difference in interlayer interactions can be, to some extent, explained by a purely geometric analysis of the interface morphologies. Here, we implement the concept of the solvent excluded surface, which has been used to describe the structures of macromolecules.[131,134] By constructing the solvent excluded surface

for each monolayer of BTTT-Cx, distinct undulating patterns determined by the orientation of terminal ethyl groups (V_{ethyl}), thus by the parity of x, are revealed (Fig. 5.6).

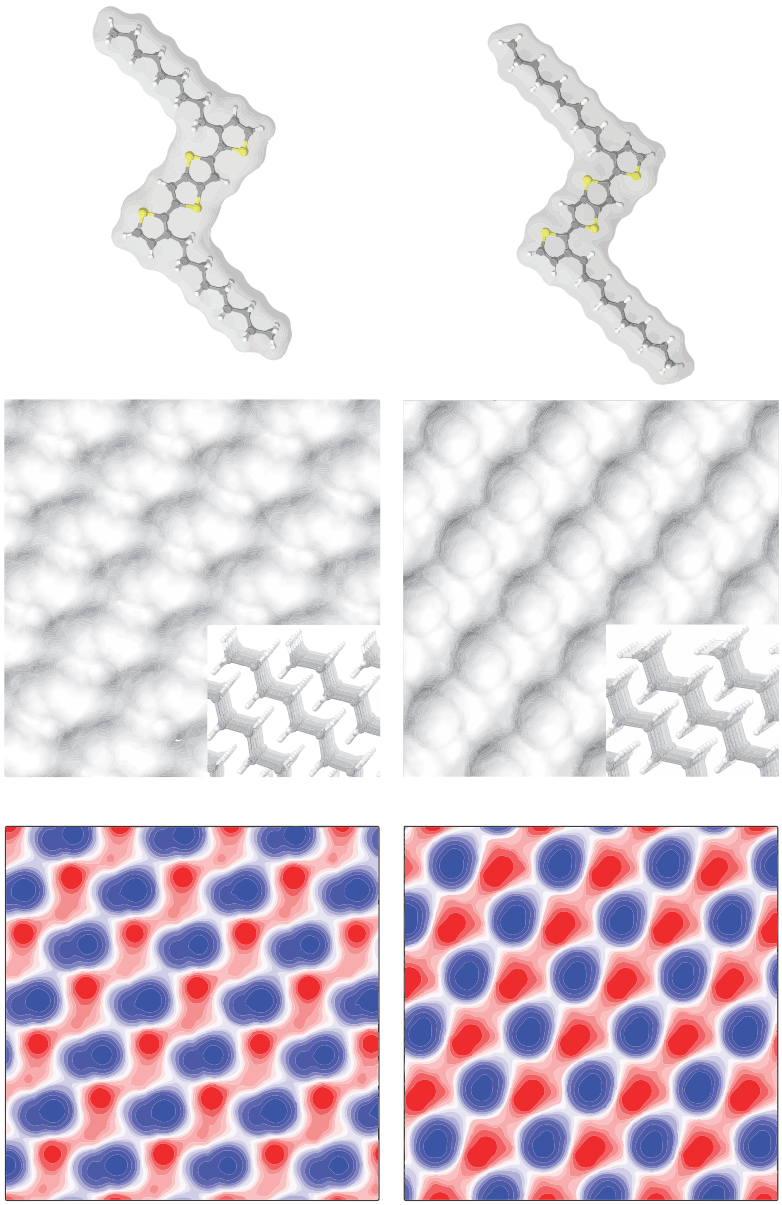


Figure 5.6: Representative pictorial demonstrations and color maps of the distinct, undulating alkyl-chain surface morphologies for odd (left) and even (right) BTTT-Cx.

To quantitatively describe the surface undulation pertaining to the inter-layer interactions, the notion of normalized undulation magnitude is introduced as $Z_P = H_P/H_{\text{max}}$ where Z_P is the normalized undulation magnitude of point P on the surface, H_P is the undulation magnitude at point P, and H_{max} is the

maximum value of H_P on a surface. Following the scheme in Fig. 5.7, the surface morphology can be visualized as a colormap of Z_P.[135] Using the Z_P colormaps of both packing types, we can illustrate the interface interactions by the value of $(Z_P^1 + Z_P^2)$ within a unit cell, where the superscripts denote individual surfaces. Such quantities describe the overlap or contact of two monolayer surfaces. If the sum of Z_P^1 and Z_P^2 is 0 (2), the spatial gap (overlap) between two molecular surfaces at point P is maximized, if the sum is 1, two surfaces merely touch at P: In this case, such a description is valid due to the existence of an inversion center in the molecule, and the absolute value of $(Z_P^1 + Z_P^2)$ is not related to any physical quantity.

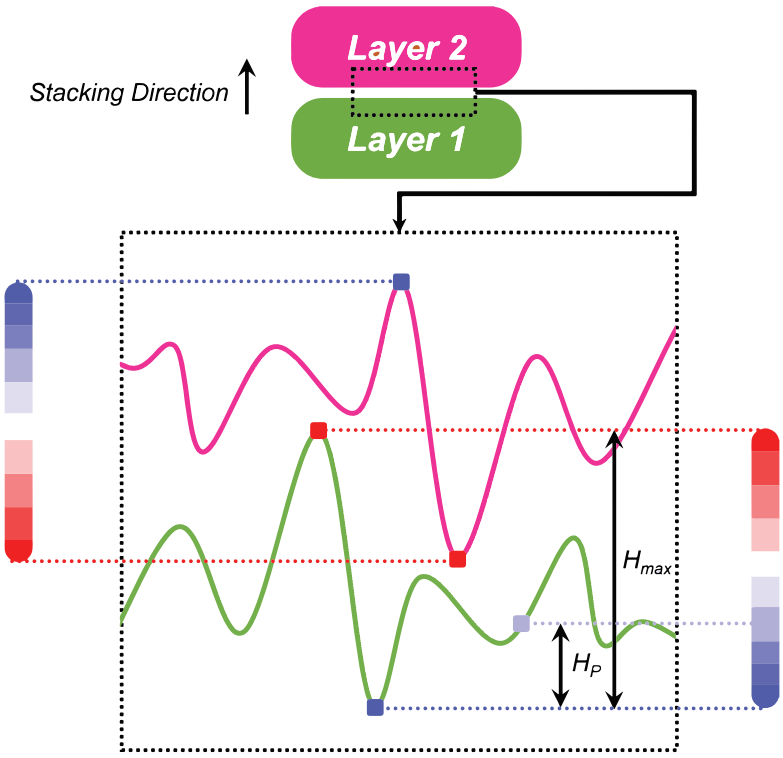


Figure 5.7: Color scheme for the colormap of Z_P, point with $Z_P = 1$ is colored red while that with $Z_P = 0$ is colored blue.

Fig. 5.8 shows the interfaces for four different crystals after periodic DFT relaxations, BTTT-C8-**I**, BTTT-C8-**II**, BTTT-C9-**I**, and BTTT-C9-**II**, where **I** and **II** denote the packing type. The experimentally determined crystal structures of BTTT-C8-**I**, BTTT-C8-**II** and BTTT-C9-**I** are used as the starting point for the relaxations: For BTTT-C9-**II**, an *in silico* polymorph is derived by first trimming

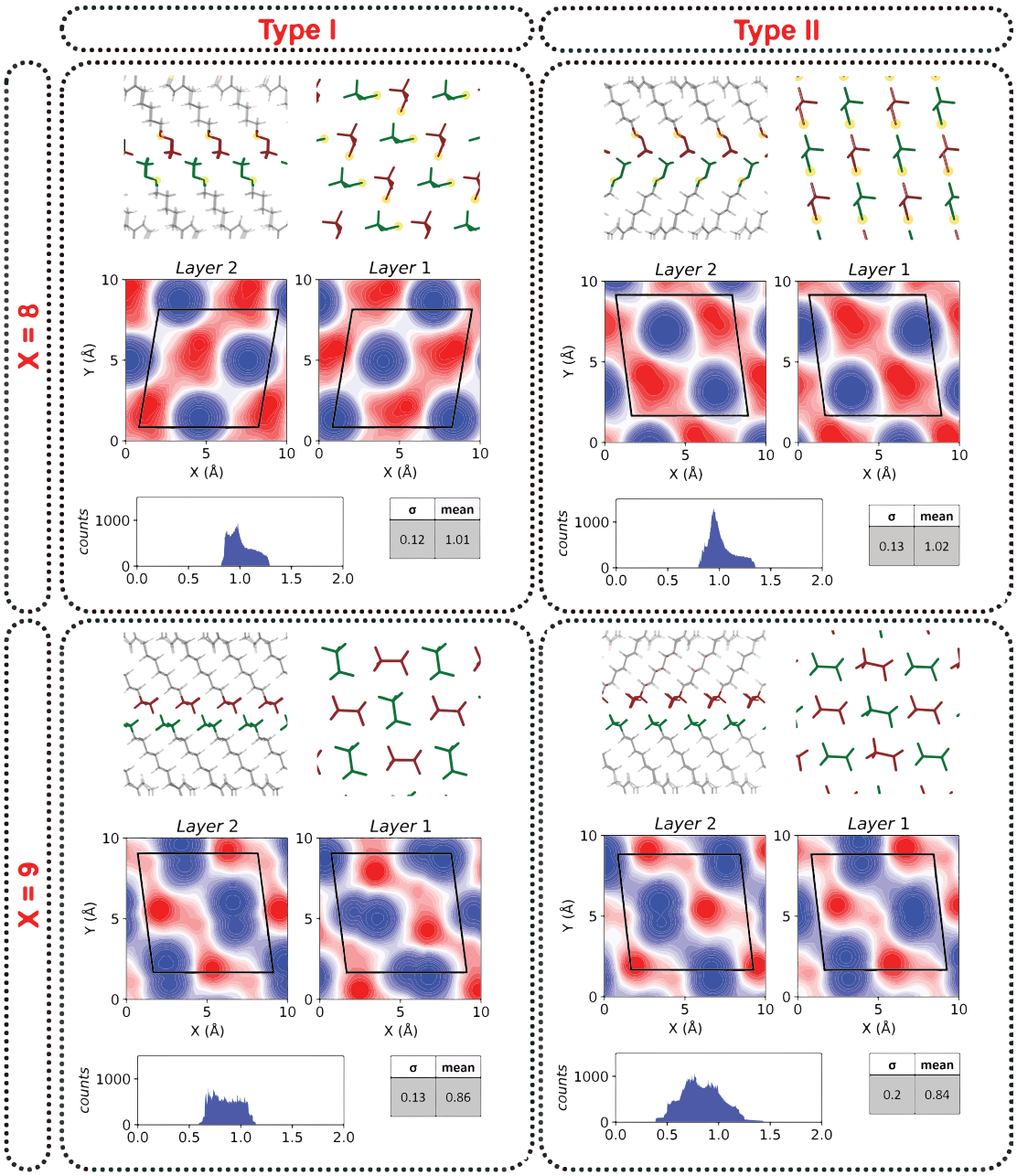


Figure 5.8: Comparison between BTTT-Cx type **I** and type **II** packing for $x = 8$ and $x = 9$. For each subplot, the top panel shows the side-view and the top-view for interlayer interface, where groups near the interface are colored brown or green as shown in Fig. 5.7. Especially for $x = 8$, hydrogen atoms far from the interface are removed and the third carbon (count from the end) is haloed to illustrate the orientation of alkyl chain. The middle panel shows the color map of Z_P^1 and Z_P^2, and the black parallelogram represents the corresponding unit cell. The bottom panel shows the distribution of $(Z_P^1 + Z_P^2)$ within one unit cell.

the side chains in BTTT-C10-**II** to C9 and then relaxing the unit-cell structure. The results of the periodic DFT relaxations are used here, in part, as it is well

known that hydrogen positions are not accurately determined by ordinary X-ray diffraction. Examination of the interlayer interface when $x = 8$ shows that while the distributions of $(Z_P^1 + Z_P^2)$ within a unit cell do differ for the two packing types, the differences in the standard deviations and mean values are not significant, a result consistent with the small difference in interlayer interaction (∼17 meV). When $x = 9$, however, the opposite situation is present: A wide distribution of $(Z_P^1 + Z_P^2)$ is found for the type **II** structure, and the interlayer interactions are much weaker, by ∼149 meV per molecule when compared to type **I**. This is consistent with the results from BTTT-C7, which shares the same interlayer interface but, as mentioned, a different interlayer arrangement with respect to BTTT-C9; For $x = 7$, the type **II** packing configuration is determined to be ∼127 meV per molecule less stable than that for type **I** packing.

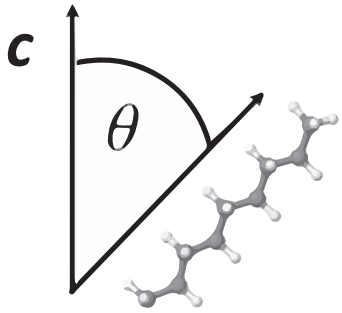


Figure 5.9: Definition of the alkyl side chain tilting angle, θ, where $\boldsymbol{c}$ is the stacking direction of molecular layers.

One plausible argument as to the variation in the surface topologies as a function of the alkyl side-chain length is that the differences are strongly affected by the tilting angle of the side chains, θ, defined as the angle between the (linear) direction of alkyl chain and the layer stacking direction, as shown in Fig. 5.9. It is interesting that while the even-odd effect is still present in crystals where θ is small, such as n-alkanes ($\theta = 15°$∼$18°$),[121] no oscillation in layer stacking is observed. For BTTT-Cx, θ ranges from 42° to 45°. Hence, it seems plausible that there exists a critical value θ_c beyond which the even-odd effect is strong enough to guide layer stacking during crystallization. Notably, in C12-BTBT, $\theta = 30°$; it

would therefore be interesting to revisit the structures of Cx-BTBT to determine if similar alternating orientations are observed.[28]

5.3.3 A note on the intralayer disorder

The diffraction patterns of all crystals contain discrete diffraction spots as well as extended diffuse streaks. Depending on individual crystal quality, some streaks are smooth and featureless whereas others are less uniform. By ignoring the diffuse streaks, all films can be indexed as primitive monoclinic. Depending on crystal quality, however, the occasional crystal (*e.g.* BTTT-C9 or BTTT-C12), can be transformed to a larger C-centered monoclinic unit cell that also accounted for non-uniformity within the streaks, *i.e.* the additional reciprocal lattice points coincided with maxima within the streaks. For the simpler primitive setting, the refined crystal structure models each have the two BTTT conformers superimposed (*i.e.* disordered, as shown in Fig. 5.10). On the other hand, the centered model has the two BTTT orientations alternating within each layer, which results in unit cells that were large enough to incorporate four molecules (two per asymmetric unit), without the need to model overlapped BTTT disorder. The two types of unit cells are related by a transformation matrix, as shown in Fig. 5.11. It seems likely that on a sub-microscopic scale there exist domains in which the two conformers alternate, but that for diffraction from macroscopic crystals this alternation is scrambled, leading to disorder. We note similar phenomenon has also been observed in 2D oligothiophene-based lead iodide perovskites,[136] which suggests the universality of such stacking-induced disorder.

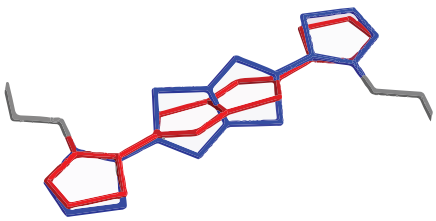


Figure 5.10: Pictorial representations of the disordered groups present in BTTT-C7, BTTT-C8, BTTT-C10, and BTTT-C11.

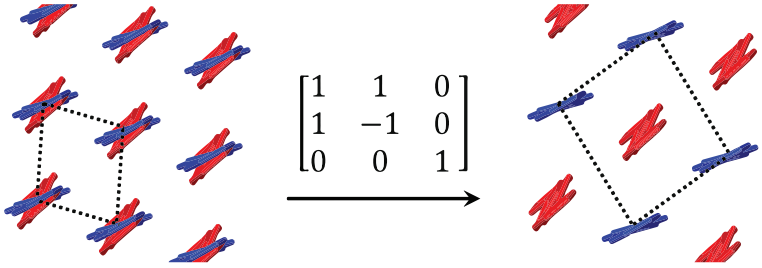


Figure 5.11: the transformation, when viewed along the long axis of bttt, relating the "primitive" two-molecule unit cell on the left and four-molecule "supercell" on the right. note that only one layer is shown here while there are two layers per unit cell.

To verify the intralayer disorder, we evaluated the relative stabilities of two well-defined structures in 2D: The "primitive" structure built from the larger unit cell, and the "supercell" structure generated by applying the aforementioned transformation to the smaller unit cell with only one disorder group. As shown in Fig. 5.12, the primitive structure is 28∼43 meV/molecule more stable than the supercell structure, which suggests that the disorder is likely a result of interlayer stacking. Interestingly, the relative stability may have non-trivial implications on the electronic structure of these 2D crystals. The electronic band structure of "primitive" structure differs from "supercell" structure by the degeneracy at high-symmetry points, as shown in Fig. 5.13, which is a direct result of the higher symmetry in the "supercell" structure.

5.4 Conclusion

We studied the oscillation of molecular packing in a series of BTTT-Cx molecular crystals as a function of x, the number of carbon atoms in the alkyl chains symmetrically appended to the molecules. Their crystal structures are classified into two categories, among which the type **I** packing, predominately adopted by BTTT-Cx with x being odd, exhibits an interlayer twist that enables a smaller interlayer spacing, and in turn a higher melting point than that of type **II** packing found in BTTT-Cx with even x. This phenomenon, accompanied by a melting point alternation, resulted from variations in interlayer interactions, as

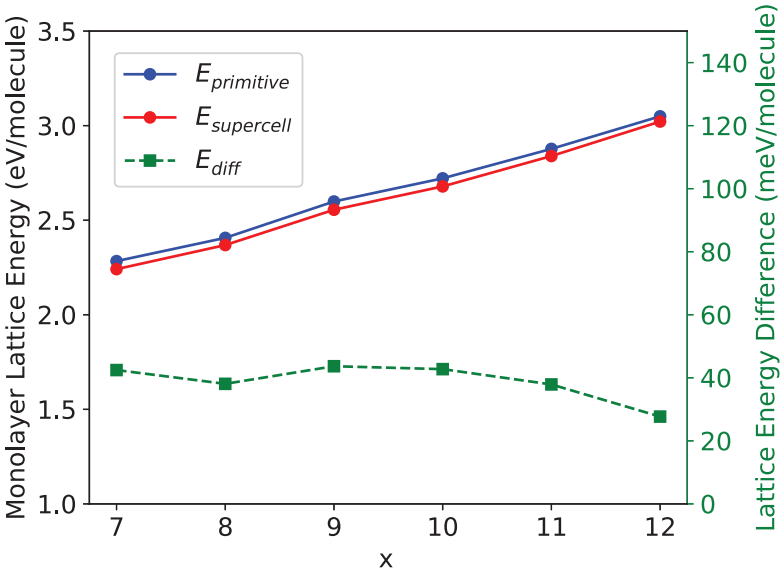


Figure 5.12: Lattice energies and lattice energy differences for two 2d lattice of intralayer structures.

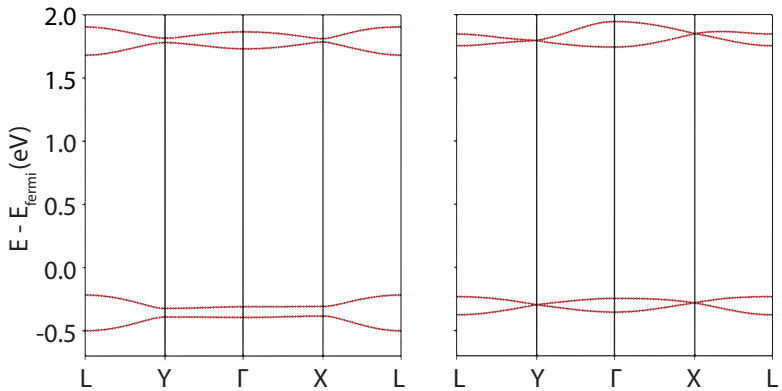


Figure 5.13: Electronic band structure for the (left) "primitive" unit cell and (right) "supercell".

revealed by DFT studies and geometric analyses. The origin of such oscillation is related to the different orientations of terminal ethyl groups caused by the side chain tilting angle θ. Focusing on an extreme case of crystal structure alternation, this study highlighted the effect of side chain parity, as an under-reported handle for controlling molecule packing in organic semiconducting materials.

CHAPTER 6

POLYMORPHISM BY ATOMISTIC SUBSTITUTION: THE CASE OF 6,13-BIS(TRIISOPROPYLGERMANYLETHYNYL) PENTACENE

6.1 Introduction

While the term "polymorphism" was first introduced for phosphate salts,[137] this concept has become a crucial part of solid-state organic chemistry,[138] as the synergy between weak intermolecular interactions and large molecular conformational space promotes the possibility of crystallizing into distinct crystal structures from the same molecule. In practice, at a given temperature and pressure, a molecular crystal has only one thermodynamically stable polymorphic form, which represents the global minimum on free energy landscape. However, other metastable polymorphic forms could be energetically close to the stable form, and transformations between crystalline phases could be kinetically hindered.[139] In the field of organic electronics, the existence of polymorphism can result in great complexity of device fabrication, where there tends to be only one polymorph possessing desired physical/chemical properties.[140] This is related to the molecular design principles behind organic electronic material: While the "functional parts" of target molecule are rigid conjugate backbones to provide charge carriers (π orbitals), bulky, flexible side groups are often attached to improve solubility by minimizing solvation free energy, and, more importantly, to realize a certain type of solid state order. In the crystal structure, such conformationally diverse side groups can lead to dynamic disorder,[141] conformation polymorphism,[142] and, as the main topic of this study, distinct molecular packing motifs.

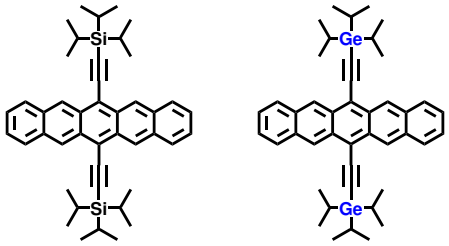


Figure 6.1: Molecular structure of (left) **TIPS-Pn** and (right) **TIPG-Pn**.

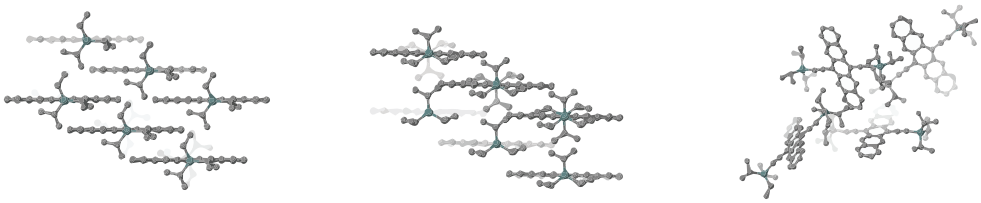


Figure 6.2: Possible molecular packing motifs for **TIPG-Pn**: (left) brickwork, (middle) slip-stack, (right) herribone.

6,13-Bis(triisopropylsilylethynyl) pentacene (**TIPS-Pn**, Fig. 6.1 left), since its discovery in 2001 by Anthony *et al.*[63] has become one of the most popular molecular materials for high mobility organic field-effect transistors.[143] It has been argued that the high performance of **TIPS-Pn** mainly originated from the brickwork packing pattern in the solid-state (Fig. 6.2 left) where the π-conjugated pentacene backbones adopt a face-to-face brickwork packing motif, establishing large intermolecular π orbital overlap. We should note that the brickwork structure was solved via x-ray crystallography for a single crystal bulk **TIPS-Pn** crystal, which is stable from 90 K to 398 K at ambient pressure.[144] While multiple **TIPS-Pn** polymorphs on thin films have been identified in previous studies,[145,146] it is believed that all of them share the brickwork packing motif similar to bulk **TIPS-Pn** and structure differences solely come from different side group conformations. The universality of the brickwork structure of **TIPS-Pn** seems to corroborate the empirical crystal engineering principle proposed in Anthony's original paper: Solid-state face-to-face π stacking can be approached when the diameter of the alkyl substituent is approximately half that of the acene length. Despite the simplicity of this purely geometric-based argument and ambiguities in quantifying parameters, a recent database study revealed that for linear acenes this principle

has > 70% successful rate in reaching face-to-face π stacking.[147]

This work was done in collaboration with the group of Dr. John Anthony (University of Kentucky), who carried out the molecular synthesis, and Dr. Sean Parkin (University of Kentucky), who carried out x-ray diffraction experiments for crystal structure determination.

Table 6.1: Packing parameters of **TIPG-Pn** and **TIPS-Pn** crystal structures. Two unique sets of slips can be defined for **BW** structure

Structure	Long Axis Slip (Å)	Short Axis Slip (Å)	Vertical Slip (Å)	Temperature (K)
TIPS-BW	9.51/6.69	1.88/1.11	3.22/3.35	173
TIPG-BW	9.43/6.89	1.83/1.24	3.26/3.36	240
TIPG-SS	7.54	2.20	3.44	180
TIPG-HB	N/A	N/A	N/A	90

Knowing the tendency of **TIPS-Pn** forming brickwork structure, it is puzzling that, being almost chemically identical to **TIPS-Pn**, the polymorphism of 6,13-bis(triisopropylgermanylethynyl) pentacene (**TIPG-Pn**, Fig. 6.1 right) involves three distinct molecular packing motifs (Fig. 6.2), namely a slipped-stack (**SS**) structure, a herringbone (**HB**) structure, and a brickwork (**BW**) structure that differs only trivially from the crystal structure of **TIPS-Pn**, as illustrated by the packing parameters shown in Table 6.1. Replacing the silicon atom in one TIPS side group with germanium results in only 0.07 Å (2%) increase in the radius of side group (average distance between silicon/germanium and the outmost hydrogens), which should not alter the conclusion from a geometry-based solid-state packing principle. It is worth noting that the slipped-stack and herringbone structures of **TIPG-Pn** are measured at different temperatures due to destructive phase transitions at low temperature ($\geq$ 90 K). This, along with the fact that herringbone structure of **TIPG-Pn** can be obtained by recrystallization while the other two structures can only be accessed via slow evaporation of solution in Petri dishes,[140] seems to suggest the herringbone structure is the most stable crystal form from 90 K to room temperature.

In this study, the polymorphism of **TIPG-Pn** is examined computationally based on bulk single-crystal structures solved by X-ray diffraction. Virtual polymorphs of **TIPS-Pn** constructed by replacing germanium with silicon in **TIPG-Pn** structures were also studied to further understand the role of heavy atom in the side groups. To investigate the stability of **TIPG-Pn** polymorphs at finite temperature, a quasi-harmonic thermal expansion model was used to derive unit-cells at different temperatures, and various descriptors were evaluated. While it is still unclear why **TIPS-Pn** prefers the **BW** motif, these results highlight the significance of the effect from lattice vibrations and the importance of configurational entropy.

6.2 Methods

All computational models with periodic boundary conditions involved in this study were derived from experimentally solved structure by single-crystal x-ray diffraction performed in Dr. Sean Parkin's lab (University of Kentucky). For each structure, relaxation was performed for both cell parameters and atomic positions under a certain external pressure determined by quasi-harmonic approximation (see Section 6.2.2). Throughout the paper, structures are labeled based on side group type and packing motif, *e.g.* **TIPG-BW** denotes the crystal of **TIPG-Pn** molecules with brickwork (**BW**) packing motif. As **TIPS-Pn** crystals have only been identified with **BW** motif, we constructed virtual crystal structures of **TIPS-HB** and **TIPS-SS** from **TIPG-HB** and **TIPG-SS** via atomistic substitution, respectively.

6.2.1 Density functional calculations

DFT calculations were carried out with Vienna *Ab-initio* Simulation Package (VASP),[84,127,148] making use of the Perdew, Burke, and Ernzerhof (PBE) exchange-correlation functional.[82] The electron-ion interactions were described

with the projector augment wave (PAW) method.[83] The kinetic energy cutoff for the plane-wave basis set was set to 1000 eV, and a Gaussian smearing with a width of 50 meV was employed. The D3 correction was employed with BJ-damping to describe the dispersion forces.[129,130] The convergence criterion of the total energy was set to 10^{-6} eV in the self-consistent field loop, and that of forces during relaxation was set to 0.005 eV/Å. For unit cells, the k-meshes used are Gamma centered 4x4x2, 3x3x3, 1x1x1 for **BW**, **SS**, **HB** structures, respectively.

6.2.2 Quasi-harmonic thermal expansion

Thermal expansion due to lattice vibrations was simulated via a quasi-harmonic approximation (QHA) method by relaxing the periodic structures under an external hydrostatic pressure p_{qha}.[149] This pressure at a certain temperature was determined by

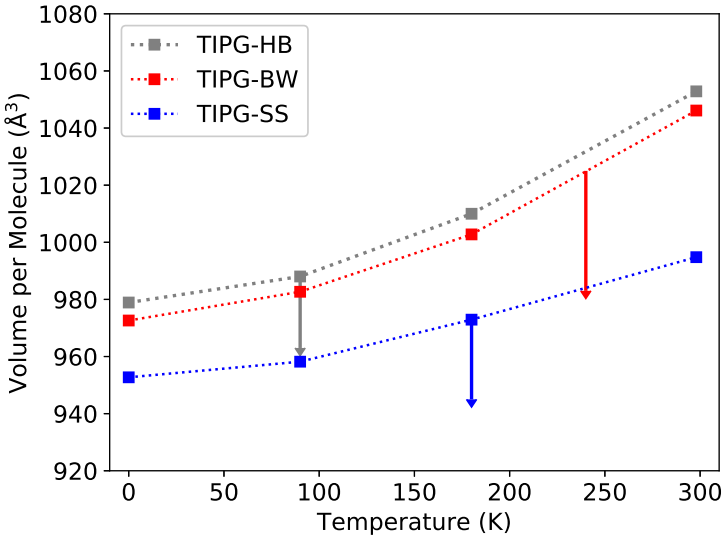
$$p_{\mathrm{qha}} = -\frac{dF_{\mathrm{vib}}}{dV} \tag{6.1}$$

where V is the volume of unit cell and F_{vib} is the vibrational free energy. In this study Eq. (6.1) was evaluated by performing (central) finite difference with a step size of 5%. F_{vib} was approximated by Eq. (6.4) where the harmonic vibrational frequencies were calculated by finite displacement as implemented in *Phonopy*,[150] where supercells ($2 \times 2 \times 1$ for **SS** and **BW**, $1 \times 1 \times 1$ for **HB**) are used with gamma-centered k-mesh, and the finite displacement is set to 0.01 Å.

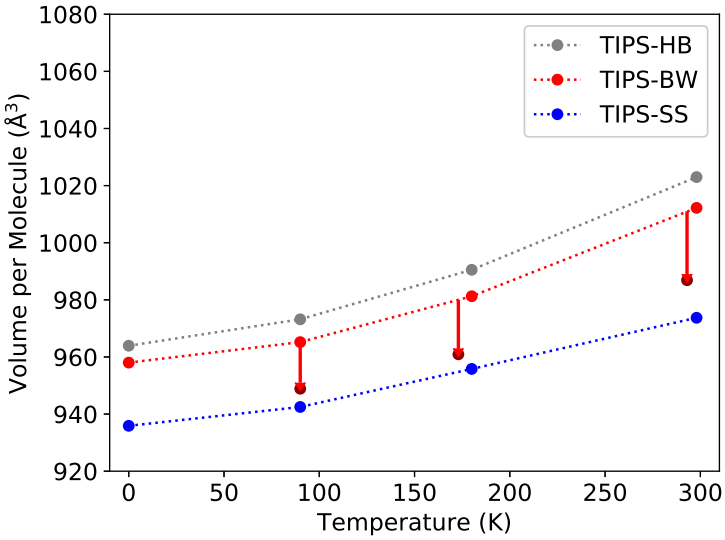
6.3 Results and Discussion

6.3.1 Thermal expansion

For organic molecular crystals, it is crucial to take volumetric expansion into account when calculating the lattice energy, as the cell volume/contraction with temperature is typically about 3% per 100 K.[151] The volumetric expansion was simulated via the QHA as described in Section 6.2.2 and the results are shown



(a) **TIPG-Pn**



(b) **TIPS-Pn**

Figure 6.3: Thermal expansion of **TIPG-Pn** and **TIPS-Pn** crystals where vertical arrows point to experimental values.

in Fig. 6.3. Comparing with experimental data, thermal expansion simulated with QHA and PBE-D3 overestimated the cell volume with an average of $\Delta V = 2.53\%$, where the worst case is **TIPG-BW** (240 K) at 4.2%. This overestimation becomes more severe at higher temperature for **TIPS-Pn**, as ΔV increases from 1.5% at 90 K to 2.2% at 293 K, which could be due to the anharmonic effects at high temperature that are ignored in QHA. In general, at the same temperature cell volume of both **TIPS-Pn** and **TIPG-Pn** follows $V(\mathbf{SS}) < V(\mathbf{BW}) < V(\mathbf{HB})$,

with V(**TIPS-Pn**) $<$ V(**TIPG-Pn**) for each packing motif, which is expected considering a TIPG side group is larger than a TIPS side group if their conformations are similar.

We now focus on **BW** and **SS** structures due to the absence of π orbital overlap in **HB** structures, which is not favorable for transistor applications. By extracting molecular dimers from the crystal structures (Table 6.2), we note that, in general, thermal expansion increases the relative displacements between the molecules in the dimer, and the increment tends to be larger in dimers of **TIPG-Pn** than that in dimers of **TIPS-Pn**. The anisotropic nature of thermal expansion is also reflected in the dimer structures, with long-axis displacement being the largest among three types of displacement, especially for **SS** structures.

Table 6.2: Relative displacements between molecules of dimers extracted from crystal structures at different temperatures. The values in parentheses denote the absolute change of slip with respect to the value at 0 K. Two unique sets of slips can be defined for **BW** structure. Some rows are shaded for readability purposes.

Structure	Temperature (K)	Long Axis Slip (Å)	Short Axis Slip (Å)	Vertical Slip (Å)
TIPS-BW	0	9.53/6.64	1.89/0.88	3.33/3.47
TIPS-BW	298	9.70 (0.17)/ 6.71 (0.07)	1.98 (0.09)/ 0.97(0.09)	3.41 (0.08)/ 3.55 (0.08)
TIPS-SS	0	7.47	2.06	3.51
TIPS-SS	298	7.58(0.11)	2.09 (0.03)	3.58 (0.07)
TIPG-BW	0	9.54/6.69	1.89/0.93	3.33/3.46
TIPG-BW	298	9.76(0.22)/ 6.82(0.13)	2.03 (0.14)/ 1.06 (0.13)	3.43 (0.10)/ 3.57 (0.11)
TIPG-SS	0	7.55	2.04	3.49
TIPG-SS	298	7.68 (0.13)	2.07 (0.03)	3.55 (0.06)

Table 6.3: Hole effective mass (in m_0, electron rest mass) at Γ point for **BW** structures at different temperature calculated by finite difference.

Structure	m_e 0K	m_e 90K	m_e 180K	m_e 298K
TIPS-BW	0.76	0.80	0.90	1.06
TIPG-BW	0.80	0.85	0.95	1.22

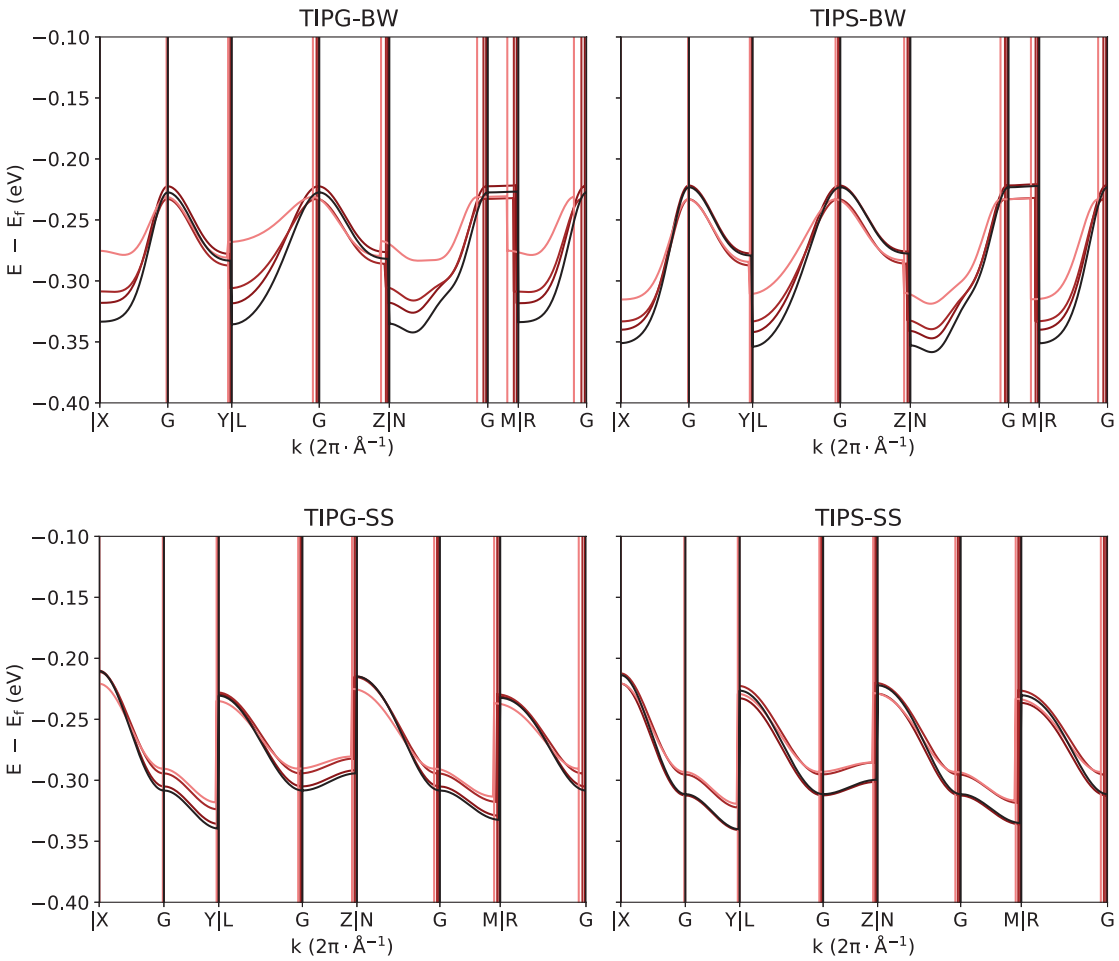


Figure 6.4: Electronic band structure of **TIPS-Pn** and **TIPG-Pn** crystals at difference temperatures. Black, dark red, brown, light coral represent the structure at 0 K, 90 K, 180 K, 298 K, respectively.

The difference in intermolecular displacement can have significant implications on the electronic structure. As shown in Fig. 6.4, for all **SS** and **BW** structures, the valence band becomes progressively flattened as the cell expands. For **BW** structures, the band flattening effect is corroborated by the increase of hole effective mass after expansion, as shown in Table 6.3. Interestingly, this observation could pose additional complexities on the negative temperature coefficient of hold mobility identified in **TIPS-Pn** based field-effect transistors.[152]

6.3.2 Relative Stability

The sublimation free energy ΔG_{sub} (298K) can be used as a descriptor to evaluate the relative stabilities of various polymorphs. ΔG_{sub} can be decomposed

as:

$$\Delta G_{\mathrm{sub}} = G^{\mathrm{mol}} - (E_{\mathrm{ie}}^{\mathrm{cr}} + F_{\mathrm{vib}}^{\mathrm{cr}} + F_{\mathrm{config}}^{\mathrm{cr}}) \tag{6.2}$$

where G^{mol} is the free energy of gas phase molecules. For a crystal structure, $E_{\mathrm{ie}}^{\mathrm{cr}}$, $F_{\mathrm{vib}}^{\mathrm{cr}}$ and $F_{\mathrm{config}}^{\mathrm{cr}}$ are the energy of ion-electron system, the vibrational free energy, and the configurational free energy, respectively. While error from G^{mol} could affect the absolute value of ΔG_{sub}, this term will be canceled when comparing stability among polymorphs of the same molecule. We note that the following assumptions are made:

- The pV term is too small to be included in the free energy of the crystal.
- Electronic excitation is negligible.

Eq. (6.2) can also be written as:

$$\Delta G_{\mathrm{sub}} = E_{\mathrm{latt}} + G_{\mathrm{xie}}^{\mathrm{mol}} - F_{\mathrm{vib}}^{\mathrm{cr}} - F_{\mathrm{config}}^{\mathrm{cr}} \tag{6.3}$$

where $E_{\mathrm{latt}} = E_{\mathrm{ie}}^{\mathrm{mol}} - E_{\mathrm{ie}}^{\mathrm{cr}}$ is the lattice energy, and $G_{\mathrm{xie}}^{\mathrm{mol}}$ is the free energy of gas phase molecule, excluding the electronic energy (of the ion-electron system), which will be cancelled if the comparison is done for polymorphs of the same molecule. E_{latt} is calculated individually as it can be used to identify the difference in static interactions among polymorphs. The vibrational (Helmholtz) free energy $F_{\mathrm{vib}}^{\mathrm{cr}}$ is calculated by sum over the contributions from each phonon mode:

$$F_{\mathrm{vib}}^{\mathrm{cr}} = \sum_{i,\boldsymbol{q}} \left(\frac{\hbar\omega_{i,\boldsymbol{q}}}{2} + k_{\mathrm{B}}T \ln\left(1 - \exp\left\{-\frac{\hbar\omega_{i,\boldsymbol{q}}}{k_{\mathrm{B}}T}\right\}\right)\right) \tag{6.4}$$

where $\omega_{i,\boldsymbol{q}}$ is the phonon frequency of mode i with wave-vector $\boldsymbol{q}$, k_{B} the Boltzmann constant, $\hbar$ the reduced Planck's constant.

Assuming vanishing configurational free energy (here ΔG_{sub} becomes $\Delta G_{\mathrm{sub}}^{0}$, where the superscript "0" denotes $F_{\mathrm{config}}^{\mathrm{cr}} = 0$), all terms in Eq. (6.3) are calculated and shown in Table 6.4 for all six crystal structures at 298 K. From

Table 6.4: Values for terms in equation 2 in eV or eV per molecule. Values in parentheses are differences relative to the reference values in **HB** structures from shaded rows. ΔG^0_{sub} is calculated by assuming $F^{\text{cr}}_{\text{config}}$=0.

Structure	E_{latt}	$G^{\text{mol}}_{\text{xie}}$	$F^{\text{cr}}_{\text{vib}}$	ΔG^0_{sub}
TIPS-BW	2.351 (-0.047)	20.031	21.017 (0.029)	1.366 (-0.076)
TIPS-SS	2.452 (0.053)	20.031	21.078 (0.090)	1.405 (-0.036)
TIPS-HB	2.398	20.031	20.988	1.442
TIPG-BW	2.308 (-0.110)	19.917	20.815 (0.014)	1.419 (-0.116)
TIPG-SS	2.502 (-0.084)	19.917	20.908 (0.107)	1.511 (-0.024)
TIPG-HB	2.418	19.917	20.801	1.535

the lattice energies, it can be argued that the strength of static interaction follows **SS** > **HB** > **BW**, which agrees with the order of cell size (Fig. 6.3). On the other hand, **HB** structure suffered the least destabilizing effect from lattice vibrations ($F^{\text{cr}}_{\text{vib}}$ being the smallest), which makes the **HB** structure the most stable in terms of ΔG^0_{sub}. While including vibrational contributions in free energy corroborated the argument that **TIPG-Pn HB** is the most stable polymorph, considering the similarity between **BW** and **SS** motifs, **BW** structures seem to be too unstable to be observed experimentally.

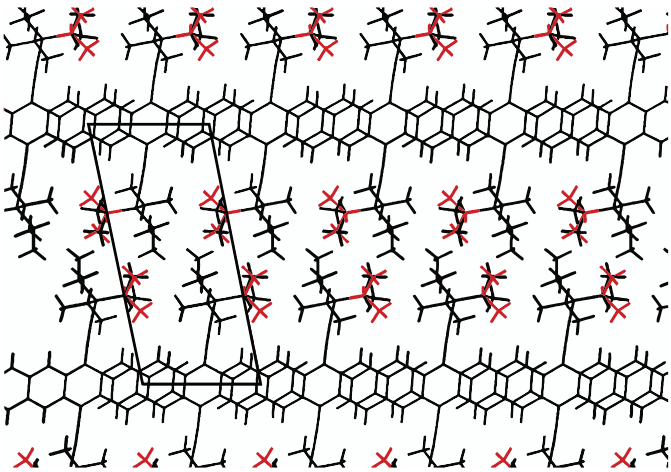


Figure 6.5: Disorder in TIPS/TIPG side groups for **BW** structures. The alternative configuration is colored red.

To address the stability of **BW** structures, it is beneficial to calculate $F^{\text{cr}}_{\text{config}}$ explicitly due to the conformationally disordered TIPS/TIPG side groups in **TIPS-BW** and **TIPG-BW**. In solving the crystal structure at 173 K, such disorder was simulated by two configurations with fractional occupancies for each

asymmetric unit, as shown in Fig. 6.5. Configurational free energy calculation depend on the nature of the disorder, as the number of configurations depends on whether the configurations are "frozen-in" (kinetically stable) or thermally populated. It is also possible that the disordered configurations are correlated due to close spatial vicinity.[153]

For "frozen-in" disorder, the configurational free energy can be calculated by:

$$F^{\mathrm{cr}}_{\text{config-fi}} = \sum_j o_j(E_j - k_{\mathrm{B}}T \ln o_j) \tag{6.5}$$

where o_j and E_j are the experimental occupancy and relative energy of configuration j, respectively. For thermally populated configurations, we have:

$$F^{\mathrm{cr}}_{\text{config-th}} = -k_{\mathrm{B}}T \ln \sum_j \exp(-E_j/(k_{\mathrm{B}}T)) \tag{6.6}$$

according to Boltzmann distribution. We note that in both Eq. (6.5) and Eq. (6.6), the differences between vibrational free energies of different configurations are ignored (that is, E_j only accounts for lattice energy), and E_j should be calculated for a supercell as disorder in the side groups is related to interlayer interactions (Fig. 6.5) that are not captured by unit cell. Table 6.5 shows ΔG_{sub} including configurational contributions where $\Delta G^{\mathrm{fi}}_{\mathrm{sub}} = \Delta G^{0}_{\mathrm{sub}} - F^{\mathrm{cr}}_{\mathrm{config-fi}}$ and $\Delta G^{\mathrm{th}}_{\mathrm{sub}} = \Delta G^{0}_{\mathrm{sub}} - F^{\mathrm{cr}}_{\text{config-th}}$. Values for **SS** and **HB** structures are also presented for comparison only, as disorder is absent in these structures. For both **TIPS-Pn** and **TIPG-Pn**, the **BW** structure is stabilized by the existence of disorder, and **TIPS-BW** is almost degenerate with **TIPS-SS** in terms of ΔG_{sub}. However, at 298 K, **TIPG-SS** is still $\sim$1.8 $k_{\mathrm{B}}T$ more stable than **TIPG-BW**, and **TIPS-HB** is $\sim$1.6 $k_{\mathrm{B}}T$ more stable than **TIPS-BW**, which contradicts the fact that no **HB** structure has been identified for **TIPS-Pn** crystals. To address such contradiction, results from more accurate computational methods could be informative, as well as experimental studies on the thermochemistry and crystallization process of **BW** structures.

Table 6.5: Values of ΔG^0_{sub} after including $F^{\text{cr}}_{\text{config}}$ in eV. Values in parentheses are differences relative to the reference values in **HB** structures from shaded rows.

Structure	ΔG^0_{sub}	$\Delta G^{\text{fi}}_{\text{sub}}$	$\Delta G^{\text{th}}_{\text{sub}}$
TIPS-BW	1.366 (-0.076)	1.400 (-0.042)	1.400 (-0.042)
TIPS-SS	1.405 (-0.036)	1.405 (-0.036)	1.405 (-0.036)
TIPS-HB	1.442	1.442	1.442
TIPG-BW	1.419 (-0.116)	1.463 (-0.072)	1.459 (-0.076)
TIPG-SS	1.511 (-0.024)	1.511 (-0.024)	1.511 (-0.024)
TIPG-HB	1.535	1.535	1.535

6.4 Conclusion

Three **TIPG-Pn** crystals at finite temperature was studied via DFT calculations, and QHA was employed to take thermal expansion into account. It is revealed that, from 0 K to room temperature, hole effective mass of **TIPG-BW** can have a 52.5% increase due to changes in intermolecular spacing promoted by thermal expansion. Vibrational analyses indicate that vibrational free energy is able to change the relative stability order of polymorphs, and so is configurational entropy, to a less degree. While some of the results contradict experimental observations, especially in the case of **TIPS-BW** which is stable at room temperature, by including all contributions to solid-state free energy, new directions are proposed to explain the puzzling polymorphism of **TIPG-Pn** that is absent in its structural analog **TIPS-Pn**.

CHAPTER 7

DEVELOPING *OCELOT*: A DATA-DRIVEN APPROACH FOR ORGANIC SEMICONDUCTOR DESIGN

7.1 Introduction

Thanks to advances in semiconductor fabrication and integrated circuits, the costs of generating and storing data has never been cheaper, giving rise to data-driven approaches in both industrial and scientific fields. While virtual screening and database studies have become routine in pharmaceutical chemistry,[154] similar practices are becoming more common in materials science, though they are scarce for organic semiconductors.[155,156] The scarcity is in part related to the unique complexities in the design of molecular semiconducting materials. Unlike in drug discovery where direct maps between molecular structures/conformations and molecular behavior (*e.g.* docking) can be established,[157] the performance of an organic semiconducting device is a macroscopic response function that directly depends on the solid-state structure formed by small molecules at, presumably, thermodynamic equilibrium.

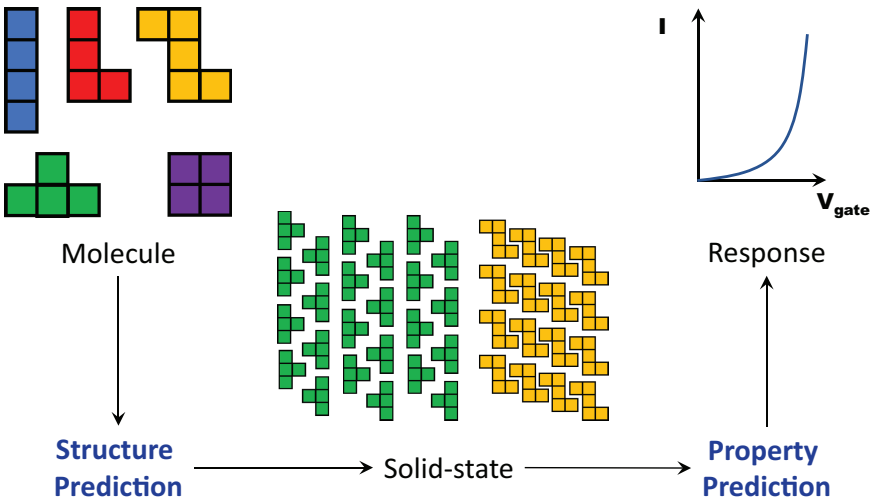


Figure 7.1: Model of designing organic semiconductors.

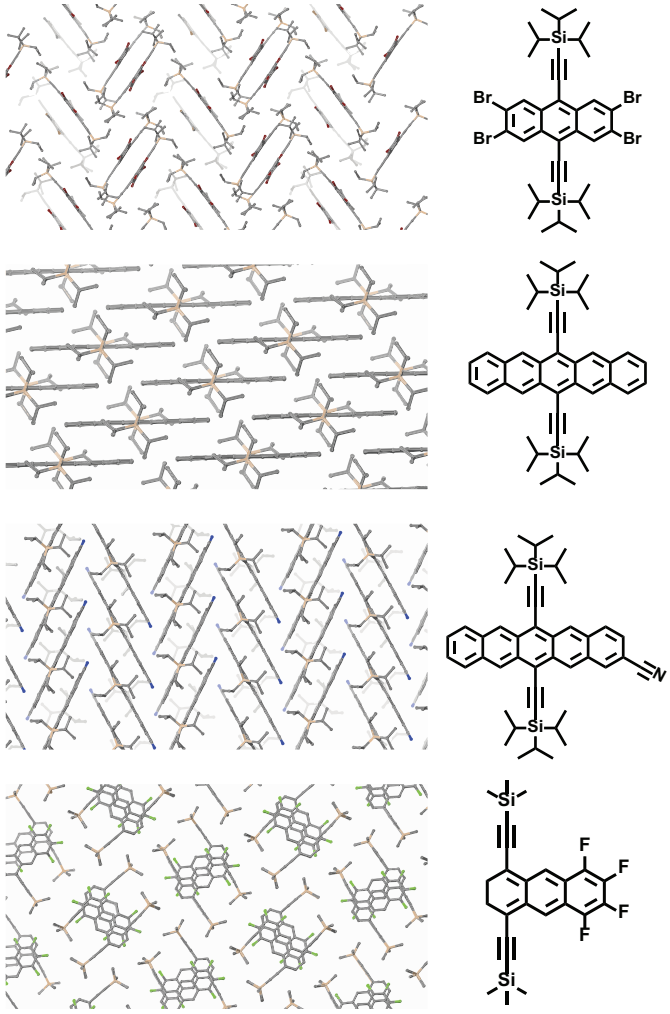


Figure 7.2: Various molecular packing motifs (left) and corresponding molecular structure (right). Note the brickwork packing motif of **TIPS-Pn** in the second row. All crystals were synthesized by John Anthony's group and solved by Dr. Sean Parkin at University of Kentucky.

While atomic connections in molecules can be accurately realized by well-trained organic chemists, control over solid-state structure formation is far from ideal, as crystal structure prediction remains a grand challenge.[31] That is, the bridge between molecular structure and material response is often blocked by the inability to predict the solid-state structure. Thus, the extended structures formed by molecular building blocks should be regarded as a fundamental schema in the model, and two types of predictions should be made to correlate molecular structure with device performance, as shown in Fig. 7.1. The significance of having solid-state structure as an additional layer can be easily understood by inspecting

Fig. 7.2 where four similar molecular structures yield drastically different molecular packing motifs. It is also implied that the "step size" of exploring chemical space should be small, since even a miniature modification of molecular structure could have remarkable impact on how the molecules assemble.[140]

6,13-Bis(triisopropylsilylethynyl)pentacene (**TIPS-Pn**, the second row in Fig. 7.2) can be considered as a prototype small-molecule organic semiconductor: The π-conjugated backbone the pentacene "core" provides polarizable π electrons that can serve as charge carriers, while the bulky side groups with multiple rotatable bonds are added symmetrically to improve solubility and, more importantly, to enable a molecular packing motif (brickwork) that is preferred for charge-carrier transport. Considering the additional rotational degrees of freedom introduced by the side groups, the difficulty of solid-state structure prediction is very high. However, in the seminal work of Anthony and co-workers,[158] based on a small set of crystal structures, it was proposed that face-to-face stacking of the π-conjugated backbones can be reached when the ratio between the "diameter" of substituents and the "length" of conjugated backbone meets particular geometric parameters. While this model may seem overly simple, it has proven powerful even for backbones with nine rings in a row,[104] in spite of both expected and unexpected outliers. The success of this purely geometry-based empirical rule revealed the intrinsic coherence in the structure prediction problem, and its limitations encouraged us to refine our understanding of molecular design with a larger, more complete data set.

As a data-driven approach to meet the challenges implicated by two prediction problems (Fig. 7.1), we introduce *OCELOT* (Organic Crystals in Electronic and Light-Oriented Technologies), a digital archive based on experimental structures and high-throughput computational workflows, offering descriptor-based schema of molecular, crystal, and electronic structures, allowing efficient data processing, data curation, and chemistry interpretability for a large dataset. In the rest of this study we present a brief description of the *OCELOT* application

programming interface (API) followed by details of the computational workflow implementations. Finally, we illustrate the potential applications of *OCELOT* via two case studies with datasets collected from either established database or community contributions.

7.2 Methods

7.2.1 Methods for crystal calculations

Density functional theory (DFT) was used

1. to relax the experimentally solved single-crystal structure so the hydrogen positions can be reliably determined
2. to obtain electronic band structure and carrier effective mass.

These calculations were performed with Vienna *Ab-initio* Simulation Package,[84,127,128] making use of the Perdew, Burke, and Ernzerhof (PBE) exchange-correlation functional.[82] The electron-ion interactions were described with the projector augment wave (PAW) method.[83] The kinetic energy cutoff for the plane-wave basis set was set to 1.3 × ENMAX as suggested to avoid Pulay stress error in relaxations,[159] and a Gaussian smearing with a width of 50 meV was applied to the partial occupancies of Kohn-Sham orbitals. For unit cells, as implemented in the Materials Project,[39] Monkhorst-Pack grids with 64 k-points per $Å^{-3}$ in the reciprocal space were applied to sample the first Brillouin zone, except for hexagonal cells where Gamma centered meshes with a similar k-point density is used. The convergence criterion of the total energy was set to 10^{-6} eV in the self-consistent field loop. For structural relaxations, DFT-D3 correction was employed with BJ-damping to describe the dispersion forces,[129,130] and the convergence criterion of forces was set to 0.02 eV/Å. For band structure calculations, line sampling was

employed to connecting high-symmetry points in the Brillouin zone as defined by *AFLOW*.[160]

7.2.2 Methods for molecular calculations

Molecular DFT calculations were performed with *Gaussian 16*[74] to determine:

(**1**) Gas phase geometries of the chromophores;

(**2**) HOMO-HOMO/LUMO-LUMO electronic couplings of dimers extracted from crystal structures;

(**3**) Vertical ionization energy (VIE);

(**4**) Vertical electron attachment energy (VEA);

(**5**) The first vertical singlet transition ($V_{S_0 \rightarrow S_1}$) via time-dependent DFT (TDDFT);

(**6**) Inner-sphere hole reorganization energies (E_{ihr}).

Geometry optimizations were performed at the LC-ωHPBE/Def2SVP level of theory.[75–77,163] For Item (**2**), while experimental data is lacking, benchmarks against coupled-cluster calculations suggest constrained DFT with 50% exact exchange yield best estimates.[164,165] However, considering the high-throughput nature of *OCELOT* and the computational cost to obtain exact exchange, electronic couplings were thus obtained via fragment molecular orbital method with polarization taken into account11 at the PBE/6-31g(d,p) level.[166,167] To find the proper functional for Item (**3**), Item (**4**), Item (**5**), Item (**6**), such calculations were performed for a set of acenes, and the results were compared with the experimentally determined ones as shown in Table 7.1, where pure functional (PBE), hybrid functionals (B3LYP, M06-2X),[168–172] and long-range-corrected functional

Table 7.1: Difference between DFT and experimentally determined VIE, VEA, $V_{S_0 \to S_1}$, E_{ihr} for acenes from naphthalene to pentacene (2∼5 rings). All values are shown in eV. Functionals with an asterisk are the IP-tuned functionals.

# of rings	B3LYP	M062X	ωB97X	LC-ωHPBE	PBE	ωB97X*	LC-ωHPBE*	expt.[161,162]
				ΔVIE				
2	-0.854	0.105	-0.023	0.140	-0.282	-0.168	-0.013	8.14
3	-0.316	0.014	-0.096	0.066	-0.382	-0.307	-0.129	7.44
4	-0.388	-0.055	-0.150	0.007	-0.447	-0.407	-0.205	6.97
5	-0.436	-0.101	-0.185	-0.032	-0.489	-0.479	-0.248	6.63
				ΔVEA				
2	-0.110	-0.159	-0.219	-0.115	0.001	-0.253	-0.085	-0.18
3	-0.008	-0.040	-0.128	-0.021	0.111	-0.187	0.011	0.53
4	0.028	0.010	-0.095	0.017	0.144	-0.245	-0.014	1.058
5	0.101	0.095	-0.024	0.094	0.214	-0.185	0.059	1.392
				$\Delta V_{S_0 \to S_1}$				
2	-0.025	0.256	0.293	0.347	-0.282	0.198	0.146	3.593
3	-0.121	0.231	0.370	0.486	-0.386	0.167	0.144	2.734
4	-0.195	0.144	0.320	0.428	-0.454	0.081	0.051	2.174
5	-0.226	0.106	0.306	0.407	-0.478	0.042	0.016	1.759
				ΔE_{ihr}				
2	-0.002	0.062	0.097	0.131	-0.045	0.091	0.046	0.182
3	-0.040	0.021	0.065	0.099	-0.080	0.009	-0.005	0.174
4	-0.029	0.031	0.085	0.120	-0.067	0.017	0.000	0.138
5	-0.010	0.051	0.113	0.152	-0.046	0.028	0.014	0.102

(ωB97X, LC-ωHPBE, and their IP-tuned[106,173] versions) were used to generate the dataset with the Def2TZVP basis set.[75,163] From these results we note that the IP-tuned LC-ωHPBE outperforms other functionals, except for VIE where the "out-of-the-box" LC-ωHPBE and ωB97X seem to be better choices. Thus IP-tuned LC-ωHPBE/Def2TZVP was used for Item (**4**), Item (**5**), Item (**6**) while the original LC-ωHPBE is used for Item (**3**).

7.2.3 Data structure and descriptors in *OCELOT* database

Computational results in the *OCELOT* database are organized based on a "Structure-to-Results" scheme. Two basic classes were constructed to realize a relatively flat data structure:

- `DbStructure` represents a molecular/crystal structure. `DbStructure.results` contains all formatted results from *OCELOT* workflow with `DbStructure` as the sole input.

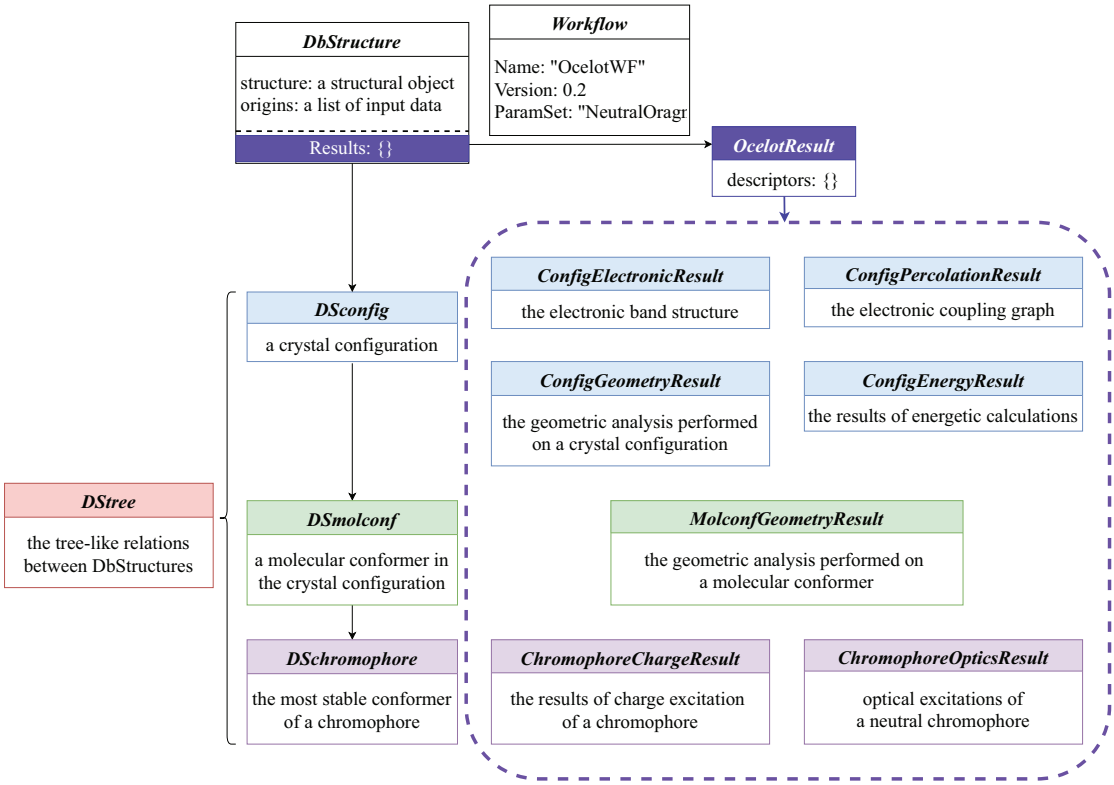


Figure 7.3: Collections in *OCELOT* database.

- `OcelotResult` is the class of the elements in `DbStructure.results`. It has a `descriptors` filed such that all descriptors related to a certain calculation can be exported.

These classes allow us to access a specific set of descriptors via a uniform method as the complicated relations between `DbStructure` have been stored in `DStree`, as shown in Fig. 7.3. All descriptors that can be calculated in *OCELOT* are listed in Table 7.2. We note that *OCELOT* is still under active development thus this list is subject to change.

Table 7.2: Descriptors in OcelotResult classes.

Descriptor	Description
	ConfigGeometryResult
`PackingParameters`	information about molecular packing motif
`SpaceFill`	volume of the space enclosed by the molecular vdW surface divided by unit cell volume
`LattParam`	lattice parameters of the unit cell
`SpaceGroup`	space group number as in the IUC table
`CellVolume`	unit cell volume
	ConfigPercolationResult
`ElectronicCouplingGraph`	a graph with molecular sites as nodes, edge properties are set to include electronic couplings between nodes
`ElectronPercolationDimension`	the dimension of electron percolation pathways
`HolePercolationDimension`	the dimension of hole percolation pathways
`PercolationCutoff`	the electronic coupling cutoff used to determine whether an edge can present in percolation pathways
	ConfigElectronicResult
`HoleLineEffectiveMass`	line effective mass calculated at the maximum of valence band
`ElectronLineEffectiveMass`	line effective mass calculated at the minimum of conduction band
`HoleEffectiveMassTensor`	tensor effective mass calculated at the maximum of valence band
`ElectronEffectiveMassTensor`	tensor effective mass calculated at the minimum of conduction band
`CBDispersion`	max(CB)−min(CB)
`VBDispersion`	max(VB)−min(VB)
`DispersionRelation`	electronic band structure as a 2D array
`Fermi`	Fermi level of the system
	ConfigEnergeticsResult
`EwaldEnergy`	ion-ion electrostatic interaction energy
`DispersionCorrection`	dispersion correction applied to total energy
`TotalEnergy`	total energy of the ion-electron system
`LatticePotentialEnergy`	internal energy change of the sublimation process without vibrational contributions
	MolconfGeometryResult
`Fragments`	a list of FragConformer objects after the fragmentation
`FragmentationMethod`	how the fragments are generated
`Descriptors3D`	3D geometric descriptors
`Descriptors2D`	2D geometric descriptors
	ChromophoreChargeResult
`VIE`	vertical ionization energy
`VEA`	vertical electron attachment energy
`ihreorg`	inner hole reorganization energy
`iereorg`	inner electron reorganization energy
`AIE`	adiabatic ionization energy
`AEA`	adiabatic electron attachment energy
	ChromophoreOpticsResult
`omega`	IP tuned omega value for range separated functionals
`vs0s1`	vertical excitation energy from s0 to s1
`dscfs0t1`	energy difference between s0 and t1 from delta scf

7.2.4 Deployment details

OCELOT is a MongoDB (v 4.2.3) database hosted on a virtual server (Ubuntu 18.04.4) provided by the Center for Computational Sciences (CCS) of University of Kentucky (UK) via the OpenStack platform. This database is also used to support the Fireworks platform for performing high-throughput calculations as specified by *OCELOT* workflow. All heavy-lifting calculations are performed on Lispcomb Computer Cluster in CCS via the *SLURM* workload manager.[174] *OCELOT* Web UI is a Django 2.1 application[175] hosted on UK College of Arts and Sciences server. Connection between *OCELOT* Web UI and *OCELOT* database is established via the *djongo* package. The *Jmol*[131] package was used to present crystal structures while *plotly*[176] was employed for the data visualization module in *OCELOT* Web UI.

7.3 Result and Discussion

7.3.1 *OCELOT* API

To systematically study the design model (Fig. 7.1) and to massively populate data entries in *OCELOT*, a python-3.x based API, which heavily relies on previous published cheminformatics packages,[177–179] *OCELOT* API, was created to provide necessary data schemas such that both chemistry concepts and material properties can be efficiently represented by abstract data types. Fig. 7.4 shows a summary of basic python classes defined for molecular and crystal structures. The `Config` class is introduced to accommodate the fact that most of *ab-initio* calculations start with well-defined atomic positions, rather than the averaged electron densities solved from *e.g.* X-ray diffractions. Instead of using one class to encapsulate both the molecular graph and molecular conformation, we make the distinction between a `MolGraph` and a `MolConformer` to respect the potential changes in molecular conformation from gas-phase to solid-state. Finally,

the `FragmentGraph` and `FragConformer` classes are defined for concepts such as "conjugated backbones" and "side groups" to pave the way for molecular fragmentation analysis. The flexibility brought by defining a general class for molecular fragment allows users to customize fragmentation scheme. As shown in Fig. 7.5, the "fused-ring-based" fragmentation on the left is useful for the structural prediction problem as mentioned in the case of **TIPS-Pn**, while the "conjugation-based" fragmentation is preferred for solving the property prediction problem.

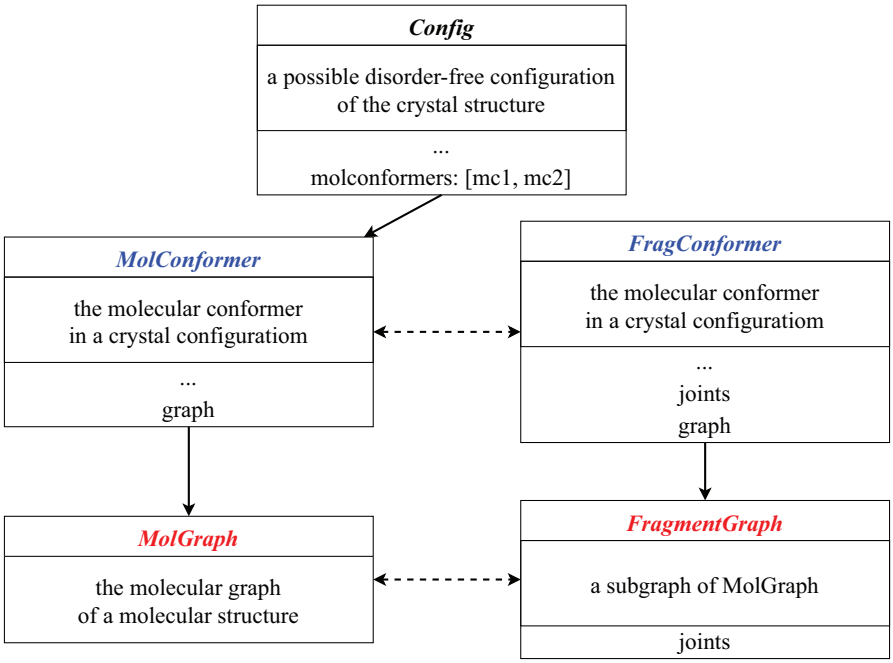


Figure 7.4: Basic structure classes in *OCELOT* API.

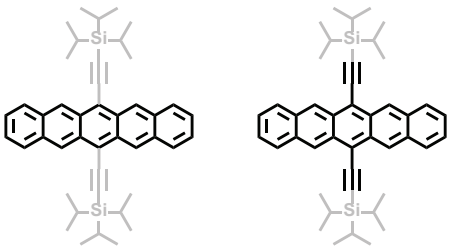


Figure 7.5: Molecular fragmentations: fused-ring-based (left) and conjugation-based (right).

OCELOT API is also capable of performing simple, computationally inexpensive tasks to obtain structural descriptors. For heavy lifting, input/output modules to various external calculators were included. Some of the capacities are listed in Table 7.3 along with a brief description for each task. As an example,

in Fig. 7.6 we showcase the NICS-XY-Scan[81] results of a substituted indolonaphthyridine thiophene[180] by using *OCELOT* API and *Gaussian 16* as the external calculator. This example indicates that while *OCELOT* API was developed as a data generator for *OCELOT*, it can also serve as a toolkit for other, more generic applications in the field of computational chemistry.

Table 7.3: Some of the heavy lifting tasks available in ocelot.task.

Task Name	Description	Calculator	Reference
`bzcal`	Parse electronic band structure calculations.	VASP, AFLOW	40
`calec`	Calculate electronic coupling for dimers.	Gaussian	181
`confgen`	Generate then cluster conformers.	RDKit	182, 183
`emtensor`	Calculate effective mass tensors.	VASP	184
`hop`	Obtain percolation pathways for charge carriers.	ZINDO, Gaussian	181, 185
`nics`	Obtain nucleus independent chemical shift for molecules.	Gaussian	81
`reorg`	Calculate inner-sphere reorganization energies.	Gaussian	186

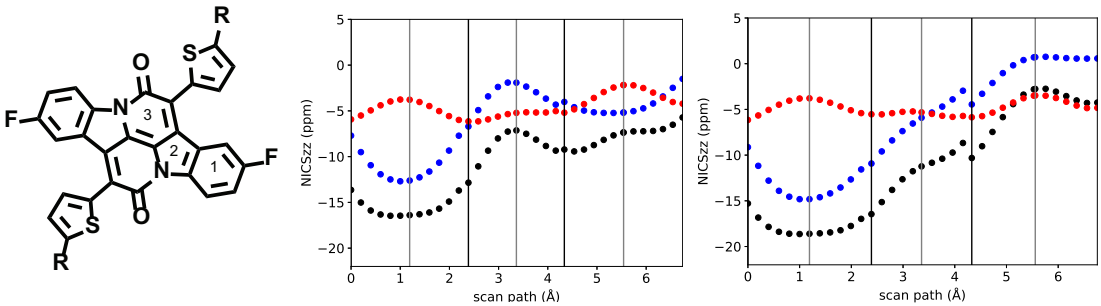


Figure 7.6: NICS-XY-scan of a substituted indolonaphthyridine thiophene. NICS-xy scan with (left) the scan path, from ring #1 to ring #3, for (middle) S_0 and (right) T_1 state. Only the data from three rings are shown due to molecular symmetry. The black, red and blue markers represent the response from the whole molecule, the hydrogenated σ-only model and the π electrons, respectively.

A data curation package, `ocelot.curator`, including a set of parsers and checkers, was introduced into the *OCELOT* API to take different conventions into account and to rule out corrupted, unphysical data. Additionally, the `DataEntry` class provides a generic framework to store both data content and data access information (*e.g.* data provider, license), as well as to push/pull data to/from previously deployed database. Data importation is then standardized via the collection and curation methods of the `Contribution` class. These implementations guaranteed the data integrity of *OCELOT* and can be applied for various data sources.

7.3.2 Computational Workflow

The *OCELOT* workflow is created to calculate various descriptors of both molecular/crystal structure and electronic structure in a high-throughput manner. Fig. 7.7 shows a sketch of the workflow that was deployed on the Fireworks platform[187] with the crystallographic information files[188] (CIFs) as the sole inputs. Data collection relies on previously published structural databases *e.g.* the Cambridge structural database[32] as well as individual contributions from the scientific community. It is important to note that while data from established structural databases is preferred due to the uniform notations prescribed by the publisher, a large amount of confidently solved crystal structures have never published and thus remain as "dark" materials to the community. Just as reaction predictions can be remarkably improved via including unreported, "dark" reactions,[189] these unreported/underreported materials may provide irreplaceable insights to the structural prediction problem. We also note data curation is essential for CIF files as disorder information and other irregularities often lead to faulty, unphysical structural information.

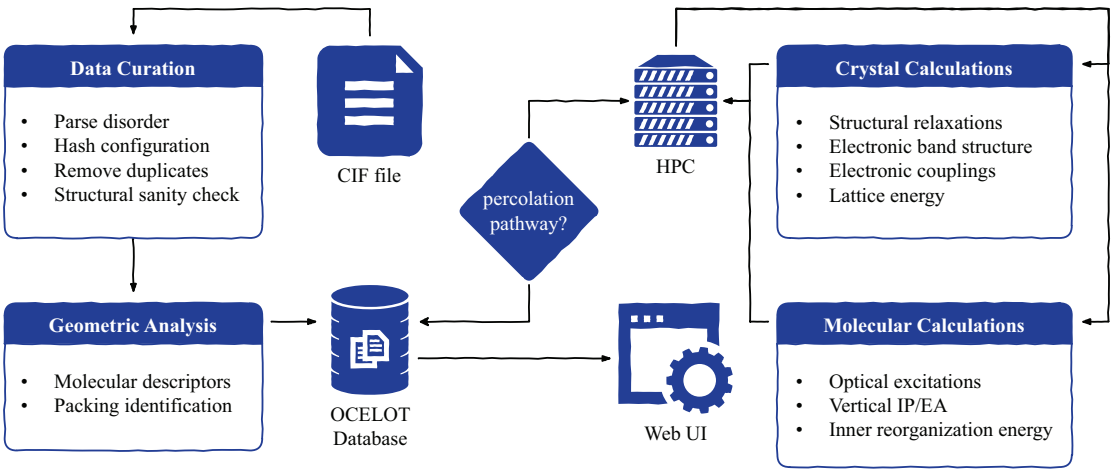


Figure 7.7: Workflow of *OCELOT*.

Fast geometric analyses are performed for all crystal structures and underlying cluster/molecular structures to investigate the structure prediction problem. Descriptors of electronic structure and optical response are only calculated if

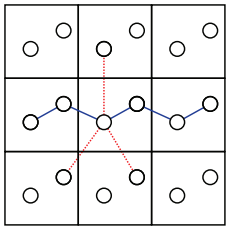


Figure 7.8: Percolation pathway based on transfer integral cutoff. Segments connecting molecular sites are colored based on if the transfer integral between two molecules is larger than a cutoff. The blue segments form a percolation pathway while the transfer integrals of red segments are too low to support effective charge carrier transport.

percolation pathways exist for charge carriers, as crystals that do not have these pathways are unlikely to be high-performing semiconducting materials due to large spatial separations among neighboring chromophores. The existence of such pathways is identified based on a user-defined cutoff for transfer integrals of dimers within a supercell, as illustrated in Fig. 7.8. Considering the solubilizing groups in a molecular structure are often chemically inert, in the step of "molecular calculations", quantum chemistry calculations are carried out only for the largest conjugated subgroup, *i.e.* the "conjugation-based" fragmentation in Fig. 7.5. We note this practice is justified as significant conformational change is not expected for relatively rigid, π-conjugated chromophore from gas phase to solid state. At the end of the *OCELOT* workflow, computational results are collected and converted to a set of descriptor-based schemas defined by *OCELOT* API (see Table 7.2 for a detailed list of descriptors). Finally, after being stored in a non-relational database (*OCELOT* database in Fig. 7.7), these formatted data are presented via a web-based user interface (*OCELOT* Web UI, https://oscar.as.uky.edu/ocelot/, where crystal structure data are generously provided by Dr. John Anthony at University of Kentucky and Dr. Michael Haley at University of Oregon), which was developed to provide easy access for the community, along with a data visualization module to exploit chemical intuition.

7.3.3 Applications

In this section, in addition to previously mentioned functions, we showcase potential applications of *OCELOT* in organic semiconductors research. Among various molecular packing motifs, the brickwork packing (Fig. 7.2, the second row) is of particular interest due to the large spatial overlap among π-conjugated chromophores and multiple charge-carrier percolation pathways. As an attempt to decipher the correlations between molecular structure and brickwork packing motifs, we attempted to answer the question of "how does one obtain an organic crystal with brickwork packing motif through disubstitution?" With the help of *OCELOT* API, a dataset of 423 organic crystals was collected from the Cambridge structural database[32] with the following three criteria:

1. The crystal structure contains only one type of neutral organic molecule.
2. The molecule can be fragmented into a linear conjugated backbone and two side groups that each contain more than three atoms.
3. The molecule has inversion symmetry.

The dataset was further split into two subsets, the "three-ring" set (3R set) where the range of molecular backbone length is set to be 6 Å$\sim$9 Å and the "five-ring" set (5R set) where the range is 11 Å$\sim$14 Å, as backbone lengths of the original set has a bimodal distribution. After performing geometric analyses for each crystal structure, a set of two-dimension colored scatter plots (the "pair plot matrix"), where axes representing different molecular descriptors and marker colors denoting molecular packing, can be obtained as shown in Fig. 7.9 (only subplots with relatively clear clustering are included, see caption for more details). We first note that it is much more likely to find a crystal with brickwork packing motif in 5R set than in 3R set, and the size (volume) of sidechains in 5R set is centered around 200 Å^3, in contrast to the wider distribution in 3R set. From the

pair plots of 3R set, it appears that the sidechains in brickwork motif tends to have small van der Waals (vdW) volumes (around 140 Å^3) and high eccentricity (> 0.75), which is also supported by the data from 5R set. Interestingly, in the 5R set, the position of substitution becomes important, as the molecules almost always crystallize into brickwork structure when the substitution angle (see Fig. 7.9 caption for its definition) is around 26°. While it is clear that such assertions need to be further validated/refined via more data and sophisticated analyses, *OCELOT* provides a flexible way to investigate the correlations between molecular fragment descriptors and crystal structure properties.

While typical materials discovery generally aims to explore unknown structures, previously reported materials with excellent properties may have been missed due to incomplete characterization or issues with device fabrication. Another application of *OCELOT* is to "rediscover" these materials via *OCELOT* workflow. From a collection of molecules synthesized by Dr. John Anthony's research group at the University of Kentucky, molecule A (codename `x12556`) and molecule E (codename `k14065`), which had been left on the shelf previously due to seemingly insufficient spatial overlap between the π-conjugated chromophores, were predicted by *OCELOT* workflow to be excellent hole transporting materials due to their small hole effective mass (0.73 m_0 and 0.78 m_0, respectively, where m_0 is the rest mass of an electron), as shown in Fig. 7.10. These two molecules were thus chosen to be the active layer in field-effect transistors fabricated by Hamna Haneef in the group of Oana Jurchescu (Wake Forest University). The best hole mobility measured from these devices reached 3.5 cm^2/(V $\cdot$ s). Intriguingly, the mobility of molecule E device (1.8 $\pm$ 0.54 cm^2/(V $\cdot$ s)) tends to be higher than that from molecule A device (0.76 $\pm$ 0.37 cm^2/(V $\cdot$ s)). While the exact causes for such difference in performance is still under investigation, this study illustrates the potential of *OCELOT* to discover unexpected properties of materials that remain "dark" to the community.

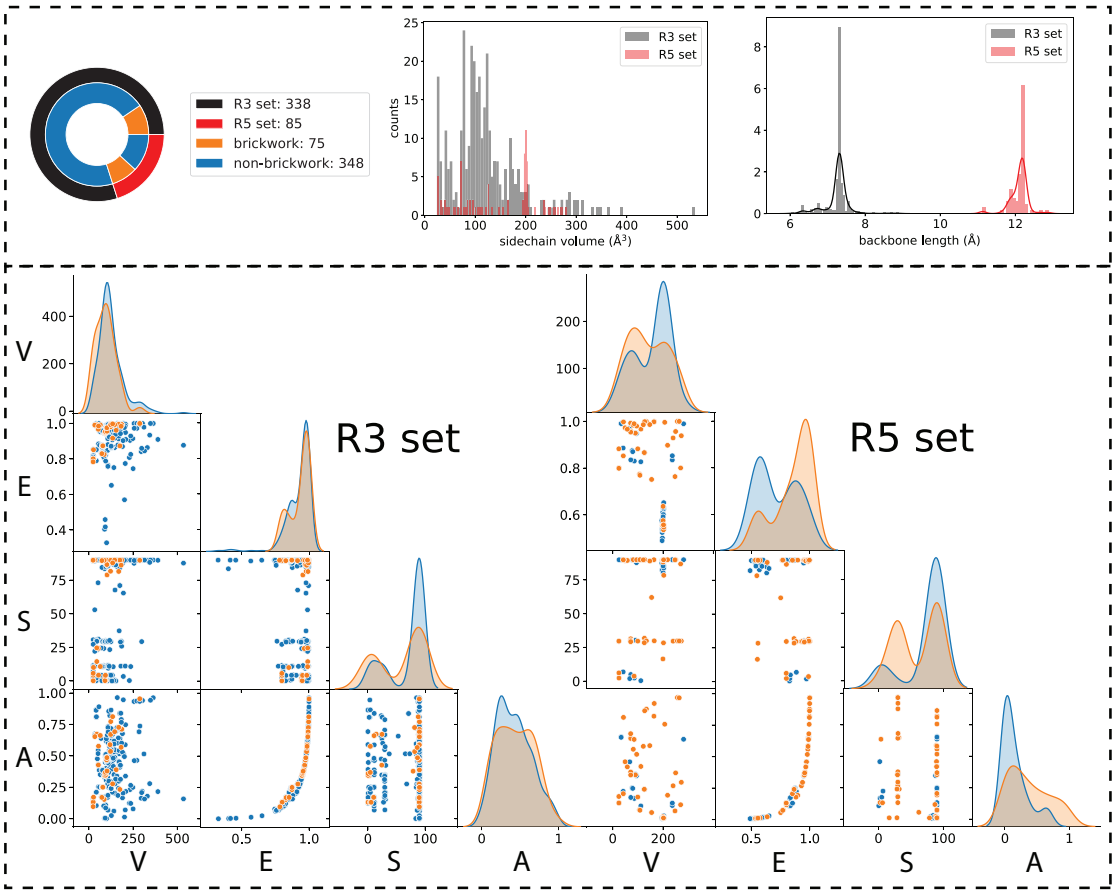


Figure 7.9: Top panel: molecular packing, sidechain volume and backbone length distributions in R3 and R5 datasets. Bottom panel: pair plot matrix generated by *seaborn*,[190] where off-diagonal subplots are the pair plots generated from 4 sidechain geometric descriptors of one sidechain (there is no need to consider both sidechains due to the inversion symmetry). V, E, S, A stand for van der Waals (vdW) volume (in Å^3), eccentricity, substitution angle, asphericity, respectively. Asphericity, eccentricity, and vdW volume were calculated by RDKit,[177,191] while the "substitution angle" is the acute/right angle between the long axis of backbone and $\boldsymbol{V}_{\text{sidechain}}$ (the vector from the geometric center of backbone to the substitution site). The diagonal plots are probability density estimate (Gaussian) of one descriptor. The curve shape of eccentricity-asphericity plot comes from the correlation between the two descriptors by definition.

7.4 Conclusion

The *OCELOT* database is introduced as a data-driven approach to solve prediction problems in the design of organic semiconductors. *OCELOT* API, a python 3.x package, was developed to parse experimental data for descriptor calculations, featuring flexible molecular fragmentation, input generations for complex

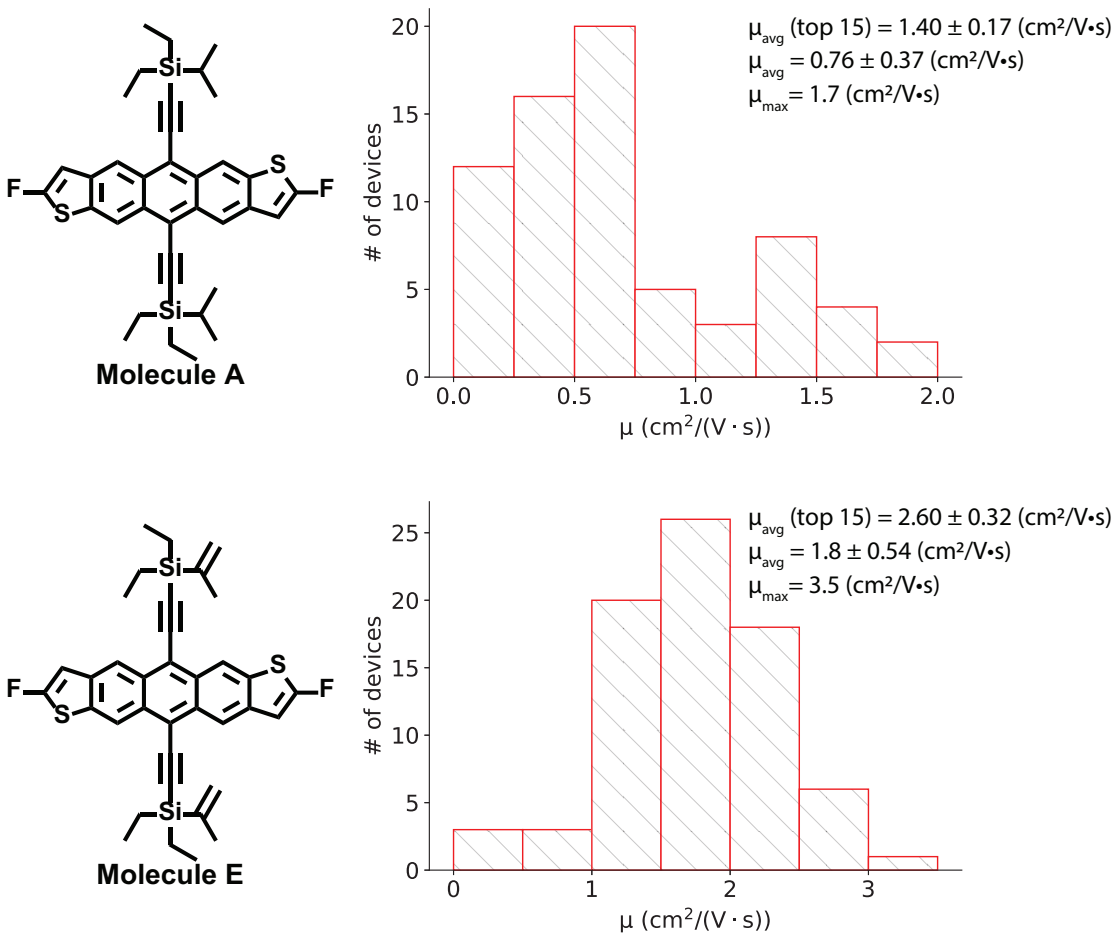


Figure 7.10: Molecular structure and device performance of (top) molecule A and (bottom) molecule E. Mobility data are provided by Hamna Haneef in Oana Jurchescu's research group at Wake Forest University.

computation tasks, and customized data curation. Combined with the *OCELOT* workflow, various descriptors of molecular, crystal, and electronic structures can be calculated in a high-throughput manner, which are then presented in the *OCELOT* Web UI with the experimental data. We believe such capacities of data curation and generation can provide insights to the organic semiconductor materials community, as exemplified by the case-studies on disubstitution and hole transport.

CHAPTER 8

PERSPECTIVES

In this dissertation, several molecular and solid-state organic semiconducting materials were studied through simulation and modeling. We first focused on molecular systems of pyrene-fused acenes (Chapter 3), from which the disruptive effect of pyrene insertion was identified through evaluation of both the molecular geometries and electronic structures. These effects, originating from the molecular orbital (MO) symmetry of pyrene, leads to the paradoxical observation that seemingly large, fused ring systems could appear electronically as shorter acenes. While this work identified another polycyclic aromatic hydrocarbon (PAH) that should behave in a similar way to pyrene in disrupting the electronic conjugation in extended acenes, an intuitive, computationally inexpensive method to find other PAHs could be helpful, especially for synthetic chemists to select synthetic targets. This might be achieved by correlating MO symmetry with molecular symmetry (based on simple quantum chemical models, e.g. the Hückel or tight-binding approximation), or by exploiting the rich chemistry encoded in Clar formulae. Other features worth exploring are the excited states of A-P-A systems: Considering the similarity of these structures to an acene dimer complex, the A-P-A could sever as a testbed suitable for singlet fission materials.[192,193]

In Chapter 4 we investigated intramolecular S…F interactions in two thiophene-containing systems. Potential energy surface (PES) and charge density analyses were employed to address the impact of such interactions on preferred molecular conformations. While these results highlight the stabilizing effect brought by fluorine substitution, either by S…F interaction or by inducing stronger hydrogen bonds, it would be naïve to assume these will be carried over into solid-state, where intermolecular interactions could be equally important in

determining molecular conformation. Furthermore, ambiguities still exist for the nature of such intramolecular interactions. For example, from energy decomposition analysis, qualitatively, it is believed that the stabilizing effect mainly comes from attractive electrostatic interactions, along with weak dispersion and inductive interactions, overcoming exchange repulsion, but accurate quantitative analysis remains elusive due to limitations in methods to segment the electron density in a single molecule.[103,194] Considering there is a wide range of potential intramolecular interactions that can be exploited to establish conformational preference, research on this topic can benefit from more accurate quantitative methods and physical models that can interpret these phenomena at molecular level in the solid state.

The odd-even layer oscillation of crystals derived from BTTT-Cx, where $x = 7 \sim 12$, was studied in Chapter 5. The oscillatory intermolecular/interlayer interactions found here echoed the first odd-even oscillation in organic crystals found by Adolf Baeyer in 1877.[195] By using a combination of periodic density functional theory methods and geometric analysis, this phenomenon is explained based on intermolecular (interlayer) interactions. We also suggest that this oscillation could be a general phenomenon for alkyl-functionalized thiophene-based organic semiconductors with lamellar crystals structures, such as Cx-BTBT.[28] Observed intralayer disorder was also discussed and attributed to irregularity in interlayer stacking. Such conclusions could be more general, as it is analogous to a later study regarding oligothiophene-based hybrid perovskites, where intralayer disorder is observed for the oligothiophene layer.[136]

The polymorphism of **TIPGe-Pn** crystals was addressed in Chapter 6, and results showed that including lattice vibrations could change the relative stability order of polymorphs. While relative stabilities of **TIPGe-Pn** crystals from simulations agree with experimental observations, it is puzzling that the brickwork structure of **TIPS-Pn**, being the most stable structure experimentally, is the least stable from simulations. From a computational perspective, this could be related to insufficient accuracy in describing weak dispersion forces under the

framework of DFT. Experimental data, especially thermochemical data, could be also helpful to understand the contradiction between experiments and simulations. Another interesting question comes from the significant dependence of the electronic structure on thermal expansion. It has been shown that at low temperatures that charge-carrier mobility measured in **TIPS-Pn**-based transistors is inversely correlated with temperature, which is attributed to the suppression of dynamic disorder.[152] How much of this inverse correlation actually come from thermal contraction is an intriguing path to follow.

Chapter 7 represents the initial construction of a data infrastructure for organic semiconductors. The major product, *OCELOT*, was introduced along with detailed implementations in data generation, curation, and presentation. While promising results have been obtained in extracting correlations and materials rediscovery, it is still far from a platform with complete, integrated data that can lead to predictive capabilities. As a long-term project, efforts need to be made in not only expanding the size of *OCELOT*, both the numbers of entries and numbers of descriptors, but also consolidating data from different disciplines (synthetic chemistry, materials processing and characterization, device engineering, etc.) such that the platform covers the full spectrum from molecules to crystals to thin-film materials.

Finally, it is informative to reflect these results in the perspective of data-driven materials research. While theoretical results can provide insights on various aspects of materials, integrating computational data with experimental observations is vital to obtain correct, novel understanding, since most of these studies originate from seemingly irregular experimental results. Thus, in addition to developing more accurate, computationally efficient simulations/approximations that allow facile, high quality data generation for a variety of physical models, large scale data collection/curation of experimental data is required to realize the fourth paradigm in materials science.

REFERENCES

[1] Tarek Zaki. *Short-Channel Organic Thin-Film Transistors : Fabrication , Characterization , Modeling and Circuit Demonstration*. Springer, 2014.

[2] Anna Köhler and Heinz Bässler. *Electronic processes in organic semiconductors: An introduction*. John Wiley & Sons, 2015.

[3] Takatoshi Tsujimura. *OLED Display Fundamentals and Applications*. John Wiley & Sons, 2017.

[4] Yongbo Yuan, Gaurav Giri, Alexander L. Ayzner, Arjan P. Zoombelt, Stefan C.B. Mannsfeld, Jihua Chen, Dennis Nordlund, Michael F. Toney, Jinsong Huang, and Zhenan Bao. Ultra-high mobility transparent organic thin film transistors grown by an off-centre spin-coating method. *Nature Communications*, 5:3005, 2014.

[5] Ying Gao, Mojtaba Asadirad, Yao Yao, Pavel Dutta, Eduard Galstyan, Shahab Shervin, Keon Hwa Lee, Sara Pouladi, Sicong Sun, Yongkuan Li, Monika Rathi, Jae Hyun Ryou, and Venkat Selvamanickam. High-Performance Flexible Thin-Film Transistors Based on Single-Crystal-like Silicon Epitaxially Grown on Metal Tape by Roll-to-Roll Continuous Deposition Process. *ACS Applied Materials and Interfaces*, 8(43):29565–29572, 2016.

[6] Hideki Shirakawa, Edwin J. Louis, Alan G. MacDiarmid, Chwan K. Chiang, and Alan J. Heeger. Synthesis of electrically conducting organic polymers: Halogen derivatives of polyacetylene, (CH)x. *Journal of the Chemical Society, Chemical Communications*, (16):578–580, 1977.

[7] Martin Dressel. Quantum criticality in organic conductors? Fermi liquid versus non-Fermi-liquid behaviour. *Journal of Physics Condensed Matter*, 23(29):293201, 2011.

[8] Artem R. Oganov and Colin W. Glass. Crystal structure prediction using ab initio evolutionary techniques: Principles and applications. *Journal of Chemical Physics*, 124(24):244704, 2006.

[9] Artem R. Oganov. *Modern Methods of Crystal Structure Prediction*. John Wiley & Sons, 2010.

[10] Natalie A Whalley, Sarah Walters, and Kate Hammond. *Molecular Cell Biology*. Macmillan, 2018.

[11] Christopher M. Dobson. Protein folding and misfolding. *Nature*, 426(6968):884–890, 2003.

[12] José Nelson Onuchic and Peter G. Wolynes. Theory of protein folding. *Current Opinion in Structural Biology*, 14(1):70–75, 2004.

[13] van der Waals. *De continuiteit van den gasen Vloeistoftoestand*, volume 1. Sijthoff, 1873.

[14] Christopher Sutton, Chad Risko, and Jean Luc Brédas. Noncovalent Intermolecular Interactions in Organic Electronic Materials: Implications for the Molecular Packing vs Electronic Properties of Acenes. *Chemistry of Materials*, 28(1):3–16, 2016.

[15] A. J. Stone and M. Alderton. Distributed multipole analysis Methods and applications. *Molecular Physics*, 100(1):221–233, 2002.

[16] W. J. Mullin and G. Blaylock. Quantum statistics: Is there an effective fermion repulsion or boson attraction? *American Journal of Physics*, 71(12):1223–1231, 2003.

[17] Cristina Puzzarini, Lorenzo Spada, Silvia Alessandrini, and Vincenzo Barone. The challenge of non-covalent interactions: theory meets experiment for reconciling accuracy and interpretation. *Journal of Physics: Condensed Matter*, 2020.

[18] Pavel Hobza and Klaus Müller-Dethlefs. *Non-covalent interactions: theory and experiment*. Number 2. Royal Society of Chemistry, 2010.

[19] Yijue Xu, Patrick M.J. Szell, Vijith Kumar, and David L. Bryce. Solid-state NMR spectroscopy for the analysis of element-based non-covalent interactions. *Coordination Chemistry Reviews*, 411:213237, 2020.

[20] Ashley M. Wright, Lynn V. Joe, Austin A. Howard, Gregory S. Tschumper, and Nathan I. Hammer. Spectroscopic and computational insight into weak

noncovalent interactions in crystalline pyrimidine. *Chemical Physics Letters*, 501(4-6):319–323, 2011.

[21] Matthew D King, Wayne Ouellette, and Timothy M Korter. Noncovalent interactions in paired DNA nucleobases investigated by terahertz spectroscopy and solid-state density functional theory. *The Journal of Physical Chemistry A*, 115(34):9467–9478, 2011.

[22] Ibon Alkorta and Antonio Elguero, Joséand Frontera. Not only hydrogen bonds: Other noncovalent interactions. *Crystals*, 10(3):180, 2020.

[23] Jane S. Murray, Giuseppe Resnati, and Peter Politzer. Close contacts and noncovalent interactions in crystals. *Faraday Discussions*, 203:113–130, 2017.

[24] Minakshi Bhardwaj, Qianxiang Ai, Sean R. Parkin, and Robert B. Grossman. An unusually short intermolecular N - H...N hydrogen bond in crystals of the hemi-hydrochloride salt of 1-exo-acetamidopyrrolizidine. *Acta Crystallographica Section E: Crystallographic Communications*, 76:77–81, 2020.

[25] A. Pochettino and A. Sella. Photoelectric behavior of anthracene. *Acad. Lincei Rendus*, 15:355–363, 1906.

[26] Toshihiro Okamoto, Chikahiko Mitsui, Masakazu Yamagishi, Katsumasa Nakahara, Junshi Soeda, Yuri Hirose, Kazumoto Miwa, Hiroyasu Sato, Akihito Yamano, Takeshi Matsushita, Takafumi Uemura, and Jun Takeya. V-shaped organic semiconductors with solution processability, high mobility, and high thermal durability. *Advanced Materials*, 25(44):6392–6397, 2013.

[27] Robert Hofmockel, Ute Zschieschang, Ulrike Kraft, Reinhold Rödel, Nis Hauke Hansen, Matthias Stolte, Frank Würthner, Kazuo Takimiya, Klaus Kern, Jens Pflaum, and Hagen Klauk. High-mobility organic thin-film transistors based on a small-molecule semiconductor deposited in vacuum and by solution shearing. *Organic Electronics*, 14(12):3213–3221, 2013.

[28] Hideaki Ebata, Takafumi Izawa, Eigo Miyazaki, Kazuo Takimiya,

Masaaki Ikeda, Hirokazu Kuwabara, and Tatsuto Yui. Highly soluble [1]benzothieno[3,2-b]benzothiophene (BTBT) derivatives for high-performance, solution-processed organic field-effect transistors. *Journal of the American Chemical Society*, 129(51):15732–15733, 2007.

[29] Gaurav Giri, Steve Park, Michael Vosgueritchian, Max Marcel Shulaker, and Zhenan Bao. High-mobility, aligned crystalline domains of tips-pentacene with metastable polymorphs through lateral confinement of crystal growth. *Advanced Materials*, 26(3):487–493, 2014.

[30] Susan Budavari, M J O'Neil, A Smith, P E Heckelman, and J F Kinneary. the Merck Index. *Chemical&Engineering News*, 79(49):6–7, 2001.

[31] Anthony M. Reilly, Richard I. Cooper, Claire S. Adjiman, Saswata Bhattacharya, A. Daniel Boese, Jan Gerit Brandenburg, Peter J. Bygrave, Rita Bylsma, Josh E. Campbell, Roberto Car, David H. Case, Renu Chadha, Jason C. Cole, Katherine Cosburn, Herma M. Cuppen, Farren Curtis, Graeme M. Day, Robert A. DiStasio, Alexander Dzyabchenko, Bouke P. Van Eijck, Dennis M. Elking, Joost A. Van Den Ende, Julio C. Facelli, Marta B. Ferraro, Laszlo Fusti-Molnar, Christina Anna Gatsiou, Thomas S. Gee, Luca M. De Gelder, Renéand Ghiringhelli, Hitoshi Goto, Stefan Grimme, Rui Guo, Detlef W.M. Hofmann, Johannes Hoja, Rebecca K. Hylton, Luca Iuzzolino, Wojciech Jankiewicz, Daniël T. De Jong, John Kendrick, Niek J.J. De Klerk, Hsin Yu Ko, Liudmila N. Kuleshova, Xiayue Li, Sanjaya Lohani, Frank J.J. Leusen, Albert M. Lund, Jian Lv, Yanming Ma, Noa Marom, Artëm E. Masunov, Patrick McCabe, David P. McMahon, Hugo Meekes, Michael P. Metz, Alston J. Misquitta, Sharmarke Mohamed, Bartomeu Monserrat, Richard J. Needs, Marcus A. Neumann, Jonas Nyman, Shigeaki Obata, Harald Oberhofer, Artem R. Oganov, Anita M. Orendt, Gabriel I. Pagola, Constantinos C. Pantelides, Chris J. Pickard, Rafal Podeszwa, Louise S. Price, Sarah L. Price, Angeles Pulido, Murray G. Read, Karsten Reuter, Elia Schneider, Christoph Schober, Gregory P. Shields, Pawanpreet Singh, Isaac J. Sugden, Krzysztof Szalewicz, Christopher R. Taylor,

Alexandre Tkatchenko, Mark E. Tuckerman, Francesca Vacarro, Manolis Vasileiadis, Alvaro Vazquez-Mayagoitia, Leslie Vogt, Yanchao Wang, Rona E. Watson, Gilles A. De Wijs, Jack Yang, Qiang Zhu, and Colin R. Groom. Report on the sixth blind test of organic crystal structure prediction methods. *Acta Crystallographica Section B: Structural Science, Crystal Engineering and Materials*, 72(4):439–459, 2016.

[32] Colin R. Groom, Ian J. Bruno, Matthew P. Lightfoot, and Suzanna C. Ward. The Cambridge structural database. *Acta Crystallographica Section B: Structural Science, Crystal Engineering and Materials*, 72(2):171–179, 2016.

[33] Michael Woelfle, Piero Olliaro, and Matthew H. Todd. Open science is a research accelerator. *Nature Chemistry*, 3(10):745–748, 2011.

[34] Erin C. McKiernan, Philip E. Bourne, C. Titus Brown, Stuart Buck, Amye Kenall, Jennifer Lin, Damon McDougall, Brian A. Nosek, Karthik Ram, Courtney K. Soderberg, Jeffrey R. Spies, Kaitlin Thaney, Andrew Updegrove, Kara H. Woo, and Tal Yarkoni. How open science helps researchers succeed. *eLife*, 5(JULY):e16800, 2016.

[35] Claudia Draxl and Matthias Scheffler. Big Data-Driven Materials Science and Its FAIR Data Infrastructure. *Handbook of Materials Modeling*, pages 49–73, 2020.

[36] Kristin M. Tolle, D. Stewart W. Tansley, and Anthony J.G. Hey. *The fourth Paradigm: Data-intensive scientific discovery*, volume 99. Microsoft research Redmond, WA, 2011.

[37] Rabi Prasad Padhy, Manas Ranjan Patra, and Suresh Chandra Satapathy. RDBMS to NoSQL: Reviewing Some Next-Generation Non-Relational Database's. *International Journal of Advanced Engineering Sciences and Technologies*, 11(11):15–30, 2011.

[38] Peter Suber. Timeline of the open access movement. *FOS Timeline*, (April 1992):1–19, 2009.

[39] Anubhav Jain, Shyue Ping Ong, Geoffroy Hautier, Wei Chen,

William Davidson Richards, Stephen Dacek, Shreyas Cholia, Dan Gunter, David Skinner, Gerbrand Ceder, and Kristin A. Persson. Commentary: The materials project: A materials genome approach to accelerating materials innovation. *APL Materials*, 1(1):11002, 2013.

[40] Stefano Curtarolo, Wahyu Setyawan, Gus L.W. Hart, Michal Jahnatek, Roman V. Chepulskii, Richard H. Taylor, Shidong Wang, Junkai Xue, Kesong Yang, Ohad Levy, Michael J. Mehl, Harold T. Stokes, Denis O. Demchenko, and Dane Morgan. AFLOW: An automatic framework for high-throughput materials discovery. *Computational Materials Science*, 58:218–226, 2012.

[41] Claudia Draxl and Matthias Scheffler. NOMAD: The FAIR concept for big data-driven materials science. *MRS Bulletin*, 43(9):676–682, 2018.

[42] Lauri Himanen, Amber Geurts, Adam Stuart Foster, and Patrick Rinke. Data-Driven Materials Science: Status, Challenges, and Perspectives. *Advanced Science*, 6(21):1900808, 2019.

[43] Shaoqing Zhang, Long Ye, Wenchao Zhao, Delong Liu, Huifeng Yao, and Jianhui Hou. Side chain selection for designing highly efficient photovoltaic polymers with 2D-conjugated structure. *Macromolecules*, 47(14):4653–4659, 2014.

[44] Jingbo Zhao, Yunke Li, Haoran Lin, Yuhang Liu, Kui Jiang, Cheng Mu, Tingxuan Ma, Joshua Yuk Lin Lai, Huawei Hu, Demei Yu, and He Yan. High-efficiency non-fullerene organic solar cells enabled by a difluorobenzothiadiazole-based donor polymer combined with a properly matched small molecule acceptor. *Energy and Environmental Science*, 8(2):520–525, 2015.

[45] Christian B. Nielsen, Andrew J.P. White, and Iain McCulloch. Effect of Fluorination of 2,1,3-Benzothiadiazole. *Journal of Organic Chemistry*, 80(10):5045–5048, may 2015.

[46] U. Dehlinger. *Solid State Physics*, volume 9. Saunders College, Philadelphia, 1956.

[47] Max Born, Kun Huang, and M. Lax. *Dynamical Theory of Crystal Lattices*, volume 23. Clarendon press, 1955.

[48] G. Willems. *Fundamentals of solid state physics for semiconductors*. Wiley New York, 2008.

[49] Linjun; Song Chenchen Zhigang, Shuai; Wang. *Theory of Charge Transport in Carbon Electronic Materials*. Springer Science&Business Media, 2012.

[50] Jan Gerit Brandenburg and Stefan Grimme. Dispersion corrected Hartree–Fock and density functional theory for organic crystal structure prediction. In *Topics in Current Chemistry*, volume 345, pages 1–24. Springer, 2014.

[51] Robert van Leeuwen. Density Functional Approach to the Many-Body Problem: Key Concepts and Exact Functionals. *Advances in Quantum Chemistry*, 43:25–94, 2003.

[52] F. B. Beleznay, F. Bogár, and J. Ladik. Charge carrier mobility in quasi-one-dimensional systems: Application to a guanine stack. *Journal of Chemical Physics*, 119(11):5690–5695, 2003.

[53] E. Clar. Vorschläge zur Nomenklatur kondensierter Ringsysteme (Aromatische Kohlenwasserstoffe, XXVI. Mitteil.). *Berichte der deutschen chemischen Gesellschaft (A and B Series)*, 72(12):2137–2139, 1939.

[54] Yen-Yi Lin, D I Gundlach, Shelby F Nelson, and Thomas N Jackson. Pentacene-based organic thin-film transistors. *IEEE Transactions on Electron Devices*, 44(8):1325–1331, 1997.

[55] Ming L Tang, Anna D Reichardt, Nobuyuki Miyaki, Randall M Stoltenberg, and Zhenan Bao. Ambipolar, high performance, acene-based organic thin film transistors. *Journal of the American Chemical Society*, 130(19):6064–6065, 2008.

[56] Marcia M. Payne, Sean R. Parkin, John E. Anthony, Chung Chen Kuo, and Thomas N. Jackson. Organic field-effect transistors from solution-deposited functionalized acenes with mobilities as high as 1 $cm^2/(V \cdot s)$. *Journal of the American Chemical Society*, 127(14):4986–4987, 2005.

[57] Bettina Sabine Basel, Constantin Hetzer, Johannes Zirzlmeier, Dominik

Thiel, Rebecca Guldi, Frank Hampel, Axel Kahnt, Timothy Clark, Dirk Michael Guldi, and Rik R. Tykwinski. Davydov splitting and singlet fission in excitonically coupled pentacene dimers. *Chemical Science*, 10(13):3854–3863, 2019.

[58] Paul M. Zimmerman, Zhiyong Zhang, and Charles B. Musgrave. Singlet fission in pentacene through multi-exciton quantum states. *Nature Chemistry*, 2(8):648–652, 2010.

[59] Timothy C. Berkelbach, Mark S. Hybertsen, and David R. Reichman. Microscopic theory of singlet exciton fission. II. Application to pentacene dimers and the role of superexchange. *Journal of Chemical Physics*, 138(11):114103, 2013.

[60] Rajib Mondal, Christina Tönshoff, Dmitriy Khon, Douglas C. Neckers, and Holger F. Bettinger. Synthesis, stability, and photochemistry of pentacene, hexacene, and heptacene: A matrix isolation study. *Journal of the American Chemical Society*, 131(40):14281–14289, 2009.

[61] Rafal Zuzak, Ruth Dorel, Marek Kolmer, Marek Szymonski, Szymon Godlewski, and Antonio M. Echavarren. Higher Acenes by On-Surface Dehydrogenation: From Heptacene to Undecacene. *Angewandte Chemie - International Edition*, 57(33):10500–10505, 2018.

[62] Bin Shen, Jörg Tatchen, Elsa Sanchez-Garcia, and Holger F. Bettinger. Evolution of the Optical Gap in the Acene Series: Undecacene. *Angewandte Chemie - International Edition*, 57(33):10506–10509, 2018.

[63] J. E. Anthony, J. S. Brooks, D. L. Eaton, and S. R. Parkin. Functionalized pentacene: Improved electronic properties from control of solid-state order [20]. *Journal of the American Chemical Society*, 123(38):9482–9483, 2001.

[64] Benjamin D. Lindner, Jens U. Engelhart, Olena Tverskoy, Anthony Lucas Appleton, Frank Rominger, Anastasia Peters, Hans Jörg Himmel, and Uwe H.F. Bunz. Stable hexacenes through nitrogen substitution. *Angewandte Chemie - International Edition*, 50(37):8588–8591, 2011.

[65] Alexander H. Endres, Manuel Schaffroth, Fabian Paulus, Hilmar Reiss, Hubert Wadepohl, Frank Rominger, Roland Krämer, and Uwe H.F. Bunz. Coronene-Containing N-Heteroarenes: 13 Rings in a Row. *Journal of the American Chemical Society*, 138(6):1792–1795, 2016.

[66] Zilong Wang, Peiyang Gu, Guangfeng Liu, Huiying Yao, Yishi Wu, Yongxin Li, Ganguly Rakesh, Jia Zhu, Hongbing Fu, and Qichun Zhang. A large pyrene-fused N-heteroacene: Fifteen aromatic six-membered rings annulated in one row. *Chemical Communications*, 53(55):7772–7775, 2017.

[67] Matthias Müller, Lukas Ahrens, Victor Brosius, Jan Freudenberg, and Uwe H.F. Bunz. Unusual stabilization of larger acenes and heteroacenes. *Journal of Materials Chemistry C*, 7(45):14011–14034, 2019.

[68] Diego Cortizo-Lacalle, Juan P. Mora-Fuentes, Karol Strutyński, Akinori Saeki, Manuel Melle-Franco, and Aurelio Mateo-Alonso. Monodisperse N-Doped Graphene Nanoribbons Reaching 7.7 Nanometers in Length. *Angewandte Chemie - International Edition*, 57(3):703–708, 2018.

[69] Sandeep More, Rajesh Bhosale, Sunil Choudhary, and Aurelio Mateo-Alonso. Versatile 2,7-substituted pyrene synthons for the synthesis of pyrene-fused azaacenes. *Organic Letters*, 14(16):4170–4173, 2012.

[70] Wangqiao Chen, Xinxiong Li, Guankui Long, Yongxin Li, Rakesh Ganguly, Mingtao Zhang, Naoki Aratani, Hiroko Yamada, Ming Liu, and Qichun Zhang. Pyrene-Containing Twistarene: Twelve Benzene Rings Fused in a Row. *Angewandte Chemie - International Edition*, 57(41):13555–13559, 2018.

[71] Mikhail Feofanov, Vladimir Akhmetov, Dmitry I. Sharapa, and Konstantin Amsharov. Modular Approach to the Synthesis of Two-Dimensional Angular Fused Acenes. *Organic Letters*, 22(5):1698–1702, 2020.

[72] Ben Lin Hu, Cunbin An, Manfred Wagner, Georgia Ivanova, Anela Ivanova, and Martin Baumgarten. Three-Dimensional Pyrene-Fused N-Heteroacenes. *Journal of the American Chemical Society*, 141(13):5130–5134, 2019.

[73] Xinping Wang, Chao Chen, Huapeng Ruan, Zhongtao Feng, Yong Fang, Yue

Zhao, Yuanting Su, Tan Gengwen, and Shuxuan Tang. Crystalline Diradical Dianions of Pyrene-Fused Azaacenes. *Angewandte Chemie International Edition*, 2020.

[74] M J Frisch, G W Trucks, H B Schlegel, G E Scuseria, M A Robb, J R Cheeseman, G Scalmani, V Barone, G A Petersson, and H Nakatsuji. Gaussian 16, Revision A. 03, Gaussian. *Inc., Wallingford CT*, 2016.

[75] Florian Weigend and Reinhart Ahlrichs. Balanced basis sets of split valence, triple zeta valence and quadruple zeta valence quality for H to Rn: Design and assessment of accuracy. *Physical Chemistry Chemical Physics*, 7(18):3297–3305, 2005.

[76] Thomas M. Henderson, Artur F. Izmaylov, Giovanni Scalmani, and Gustavo E. Scuseria. Can short-range hybrids describe long-range-dependent properties? *Journal of Chemical Physics*, 131(4):44108, 2009.

[77] Oleg A. Vydrov and Gustavo E. Scuseria. Assessment of a long-range corrected hybrid functional. *Journal of Chemical Physics*, 125(23):234109, 2006.

[78] Tamar Stein, Leeor Kronik, and Roi Baer. Prediction of charge-transfer excitations in coumarin-based dyes using a range-separated functional tuned from first principles. *Journal of Chemical Physics*, 131(24):244119, 2009.

[79] Haitao Sun, Sean Ryno, Cheng Zhong, Mahesh Kumar Ravva, Zhenrong Sun, Thomas Körzdörfer, and Jean Luc Brédas. Ionization Energies, Electron Affinities, and Polarization Energies of Organic Molecular Crystals: Quantitative Estimations from a Polarizable Continuum Model (PCM)-Tuned Range-Separated Density Functional Approach. *Journal of Chemical Theory and Computation*, 12(6):2906–2916, 2016.

[80] James R. Cheeseman. A comparison of models for calculating nuclear magnetic resonance shielding tensors. *Journal of Chemical Physics*, 104(14):5497–5509, 1996.

[81] Renana Gershoni-Poranne and Amnon Stanger. The NICS-XY-scan: Identification of local and global ring currents in multi-ring systems. *Chemistry - A European Journal*, 20(19):5673–5688, 2014.

[82] John P. Perdew, J. A. Chevary, S. H. Vosko, Koblar A. Jackson, Mark R. Pederson, D. J. Singh, and Carlos Fiolhais. Atoms, molecules, solids, and surfaces: Applications of the generalized gradient approximation for exchange and correlation. *Physical Review B*, 46(11):6671–6687, 1992.

[83] Peter E Blöchl. Projector augmented-wave method. *Physical Review B*, 50(24):17953, 1994.

[84] G. Kresse and J. Furthmüller. Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set. *Physical Review B - Condensed Matter and Materials Physics*, 54(16):11169–11186, 1996.

[85] Dariusz W. Szczepanik, Miquel Solà, Tadeusz M. Krygowski, Halina Szatylowicz, Marcin Andrzejak, Barbara Pawełek, Justyna Dominikowska, Mercedes Kukułka, and Karol Dyduch. Aromaticity of acenes: The model of migrating π-circuits. *Physical Chemistry Chemical Physics*, 20(19):13430–13436, 2018.

[86] Debojit Bhattacharya, Anirban Panda, Anirban Misra, and Douglas J. Klein. Clar theory extended for polyacenes and beyond. *Journal of Physical Chemistry A*, 118(24):4325–4338, 2014.

[87] Melvin B Robin and Peter Day. Mixed valence chemistry-a survey and classification. In *Advances in inorganic chemistry and radiochemistry*, volume 10, pages 247–422. Elsevier, 1968.

[88] M. Parthey and M. Kaupp. Quantum-chemical insights into mixed-valence systems: Within and beyond the Robin-Day scheme. *Chemical Society Reviews*, 43(14):5067–5088, 2014.

[89] Yang Yang, Ernest R. Davidson, and Weitao Yang. Nature of ground and electronic excited states of higher acenes. *Proceedings of the National Academy of Sciences of the United States of America*, 113(35):E5098–E5107, 2016.

[90] Holger F. Bettinger, Christina Tönshoff, Markus Doerr, and Elsa Sanchez-Garcia. Electronically Excited States of Higher Acenes up to Nonacene: A Density Functional Theory/Multireference Configuration Interaction Study.

Journal of Chemical Theory and Computation, 12(1):305–312, 2016.

[91] Enrique R. Batista and Richard L. Martin. Exciton localization in a Pt-acetylide complex. *Journal of Physical Chemistry A*, 109(43):9856–9859, 2005.

[92] Thomas M. Cooper, Douglas M. Krein, Aaron R. Burke, Daniel G. McLean, Joy E. Rogers, Jonathan E. Slagle, and Paul A. Fleitz. Spectroscopic characterization of a series of platinum acetylide complexes having a localized triplet exciton. *Journal of Physical Chemistry A*, 110(13):4369–4375, 2006.

[93] Zhongbo Zhang and Qichun Zhang. Recent progress in well-defined higher azaacenes (n 6): Synthesis, molecular packings, and applications. *Materials Chemistry Frontiers*, 2020.

[94] Qianxiang Ai, Karol Jarolimek, Samuel Mazza, John E. Anthony, and Chad Risko. Delimited Polyacenes: Edge Topology as a Tool to Modulate Carbon Nanoribbon Structure, Conjugation, and Mobility. *Chemistry of Materials*, 30(3):947–957, 2018.

[95] Holger F. Bettinger. Electronic structure of higher acenes and polyacene: The perspective developed by theoretical analyses. *Pure and Applied Chemistry*, 82(4):905–915, 2010.

[96] Francesco Babudri, Gianluca M. Farinola, Francesco Naso, and Roberta Ragni. Fluorinated organic materials for electronic and optoelectronic applications: The role of the fluorine atom. *Chemical Communications*, (10):1003–1022, 2007.

[97] John R. Tumbleston, Andrew C. Stuart, Eliot Gann, Wei You, and Harald Ade. Fluorinated polymer yields high organic solar cell performance for a wide range of morphologies. *Advanced Functional Materials*, 23(27):3463–3470, 2013.

[98] Gary Conboy, Howard J. Spencer, Enrico Angioni, Alexander L. Kanibolotsky, Neil J. Findlay, Simon J. Coles, Claire Wilson, Mateusz B. Pitak,

Chad Risko, Veaceslav Coropceanu, Jean Luc Brédas, and Peter J. Skabara. To bend or not to bend-are heteroatom interactions within conjugated molecules effective in dictating conformation and planarity? *Materials Horizons*, 3(4):333–339, 2016.

[99] Howard J. Spencer, Peter J. Skabara, Mark Giles, Iain McCulloch, Simon J. Coles, and Michael B. Hursthouse. The first direct experimental comparison between the hugely contrasting properties of PEDOT and the all-sulfur analogue PEDTT by analogy with well-defined EDTT-EDOT copolymers. *Journal of Materials Chemistry*, 15(45):4783–4792, 2005.

[100] Kiamars Eskandari and Mina Lesani. Does fluorine participate in halogen bonding? *Chemistry - A European Journal*, 21(12):4739–4746, 2015.

[101] Craig C. Robertson, James S. Wright, Elliot J. Carrington, Robin N. Perutz, Christopher A. Hunter, and Lee Brammer. Hydrogen bonding: Vs. halogen bonding: The solvent decides. *Chemical Science*, 8(8):5392–5398, 2017.

[102] Brett R. Beno, Kap Sun Yeung, Michael D. Bartberger, Lewis D. Pennington, and Nicholas A. Meanwell. A Survey of the Role of Noncovalent Sulfur Interactions in Drug Design. *Journal of Medicinal Chemistry*, 58(11):4383–4438, 2015.

[103] Karl J Thorley and Iain McCulloch. Why are S–F and S–O non-covalent interactions stabilising? *Journal of Materials Chemistry C*, 6(45):12413–12421, 2018.

[104] Anthony J. Petty, Qianxiang Ai, Jeni C. Sorli, Hamna F. Haneef, Geoffrey E. Purdum, Alex Boehm, Devin B. Granger, Kaichen Gu, Carla Patricia Lacerda Rubinger, Sean R. Parkin, Kenneth R. Graham, Oana D. Jurchescu, Yueh Lin Loo, Chad Risko, and John E. Anthony. Computationally aided design of a high-performance organic semiconductor: The development of a universal crystal engineering core. *Chemical Science*, 10(45):10543–10549, 2019.

[105] Tetiana Kharandiuk, Eman J. Hussien, Joseph Cameron, Romana Petrina,

Neil J. Findlay, Roman Naumov, Wim T. Klooster, Simon J. Coles, Qianxiang Ai, Stephen Goodlett, Chad Risko, and Peter J. Skabara. Noncovalent Close Contacts in Fluorinated Thiophene-Phenylene-Thiophene Conjugated Units: Understanding the Nature and Dominance of O…H versus S…F and O…F Interactions with Respect to the Control of Polymer Conformation. *Chemistry of Materials*, 31(17):7070–7079, 2019.

[106] Jeng Da Chai and Martin Head-Gordon. Systematic optimization of long-range corrected hybrid density functionals. *Journal of Chemical Physics*, 128(8):84106, 2008.

[107] Thom H. Dunning. Gaussian basis sets for use in correlated molecular calculations. I. The atoms boron through neon and hydrogen. *The Journal of Chemical Physics*, 90(2):1007–1023, 1989.

[108] Tian Lu and Feiwu Chen. Multiwfn: A multifunctional wavefunction analyzer. *Journal of Computational Chemistry*, 33(5):580–592, 2012.

[109] Tian Lu and Feiwu Chen. Quantitative analysis of molecular surface based on improved Marching Tetrahedra algorithm. *Journal of Molecular Graphics and Modelling*, 38:314–323, 2012.

[110] Anna K. Hailey, Anthony J. Petty, Jennifer Washbourne, Karl J. Thorley, Sean R. Parkin, John E. Anthony, and Yueh Lin Loo. Understanding the Crystal Packing and Organic Thin-Film Transistor Performance in Isomeric Guest–Host Systems. *Advanced Materials*, 29(23):1700048, 2017.

[111] Cristina Pozo-Gonzalo, Tahir Khan, Joseph J.W. McDouall, Peter J. Skabara, Donna M. Roberts, Mark E. Light, Simon J. Coles, Michael B. Hursthouse, Helmut Neugebauer, Antonio Cravino, and N. Serdar Sariciftci. Synthesis and electropolymerisation of 3 ,4 -bis(alkylsulfanyl)terthiophenes and the significance of the fused dithiin ring in 2,5-dithienyl-3,4-ethylenedithiothiophene (DT-EDTT). *Journal of Materials Chemistry*, 12(3):500–510, 2002.

[112] Erin R. Johnson, Shahar Keinan, Paula Mori-Sánchez, Julia Contreras-García, Aron J. Cohen, and Weitao Yang. Revealing noncovalent interactions. *Journal of the American Chemical Society*, 132(18):6498–6506, 2010.

[113] Keiamars Eskandari and Homam Zariny. Halogen bonding: A lump-hole interaction. *Chemical Physics Letters*, 492(1-3):9–13, 2010.

[114] Mysore S. Pavan, Karothu Durga Prasad, and T. N.Guru Row. Halogen bonding in fluorine: Experimental charge density study on intermolecular F F and F S donor–acceptor contacts. *Chemical Communications*, 49(68):7558–7560, 2013.

[115] Yi Xia, Stéphane Viel, Yang Wang, Fabio Ziarelli, Erik Laurini, Paola Posocco, Maurizio Fermeglia, Fanqi Qu, Sabrina Pricl, and Ling Peng. Rationalizing the F S interaction discovered within a tetrafluorophenylazido-containing bola-phospholipid. *Chemical Communications*, 48(36):4284–4286, 2012.

[116] Neetha Mohan and Cherumuttathu H. Suresh. A molecular electrostatic potential analysis of hydrogen, halogen, and dihydrogen bonds. *Journal of Physical Chemistry A*, 118(9):1697–1705, 2014.

[117] Jianguo Mei and Zhenan Bao. Side chain engineering in solution-processable conjugated polymers. *Chemistry of Materials*, 26(1):604–615, 2014.

[118] Yan Zhou, Tadanori Kurosawa, Wei Ma, Yikun Guo, Lei Fang, Koen Vandewal, Ying Diao, Chenggong Wang, Qifan Yan, Julia Reinspach, Jianguo Mei, Anthony Lucas Appleton, Ghada I. Koleilat, Yongli Gao, Stefan C.B. Mannsfeld, Alberto Salleo, Harald Ade, Dahui Zhao, and Zhenan Bao. High performance all-polymer solar cell via polymer side-chain engineering. *Advanced Materials*, 26(22):3767–3772, 2014.

[119] Lei Fang, Yan Zhou, Yu Xing Yao, Ying Diao, Wen Ya Lee, Anthony L. Appleton, Ranulfo Allen, Julia Reinspach, Stefan C.B. Mannsfeld, and Zhenan Bao. Side-chain engineering of isoindigo-containing conjugated polymers using polystyrene for high-performance bulk heterojunction solar cells. *Chemistry of Materials*, 25(24):4874–4880, 2013.

[120] Jianguo Mei, Do Hwan Kim, Alexander L. Ayzner, Michael F. Toney, and Zhenan Bao. Siloxane-terminated solubilizing side chains: Bringing conjugated polymer backbones closer and boosting hole mobilities in thin-film transistors. *Journal of the American Chemical Society*, 133(50):20130–20133, 2011.

[121] Roland Boese, Hans Christoph Weiss, and Dieter Bläser. The melting point alternation in the short-chain n-alkanes: Single-crystal X-ray analyses of propane at 30 K and of n-butane to n-nonane at 90 K. *Angewandte Chemie - International Edition*, 38(7):988–992, 1999.

[122] Venkat R. Thalladi, Markus Nüsse, and Roland Boese. The melting point alternation in α,ω-alkanedicarboxylic acids. *Journal of the American Chemical Society*, 122(38):9227–9236, 2000.

[123] Venkat R. Thalladi, Roland Boese, and Hans Christoph Weiss. The melting point alternation in α,ω-alkanediols and α,ω- alkanediamines: Interplay between hydrogen bonding and hydrophobic interactions. *Angewandte Chemie - International Edition*, 39(5):918–922, 2000.

[124] Andrew D. Bond. On the crystal structures and melting point alternation of the n-alkyl carboxylic acids. *New Journal of Chemistry*, 28(1):104–114, 2004.

[125] Feng Tao and Steven L Bernasek. Understanding odd− even effects in organic self-assembled monolayers. *Chemical Reviews*, 107(5):1408–1453, 2007.

[126] Hylke B. Akkerman, Stefan C.B. Mannsfeld, Ananth P. Kaushik, Eric Verploegen, Luc Burnier, Arjan P. Zoombelt, Jonathan D. Saathoff, Sanghyun Hong, Sule Atahan-Evrenk, Xueliang Liu, Alán Aspuru-Guzik, Michael F. Toney, Paulette Clancy, and Zhenan Bao. Effects of odd-even side chain length of alkyl-substituted diphenylbithiophenes on first monolayer thin film packing structure. *Journal of the American Chemical Society*, 135(30):11006–11014, 2013.

[127] G. Kresse and J. Furthmüller. Efficiency of ab-initio total energy calculations for metals and semiconductors using a plane-wave basis set. *Computational*

Materials Science, 6(1):15–50, 1996.

[128] D. Joubert. From ultrasoft pseudopotentials to the projector augmented-wave method. *Physical Review B - Condensed Matter and Materials Physics*, 59(3):1758–1775, 1999.

[129] Stefan Grimme, Stephan Ehrlich, and Lars Goerigk. Effect of the damping function in dispersion corrected density functional theory. *Journal of Computational Chemistry*, 32(7):1456–1465, 2011.

[130] Stefan Grimme, Jens Antony, Stephan Ehrlich, and Helge Krieg. A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu. *Journal of Chemical Physics*, 132(15):154104, 2010.

[131] Robert M. Hanson. Jmol-a paradigm shift in crystallographic visualization. *Journal of Applied Crystallography*, 43(5 PART 2):1250–1260, 2010.

[132] Michael L Connolly. Analytical molecular surface calculation. *Journal of applied crystallography*, 16(5):548–558, 1983.

[133] Edmund K. Burnett, Qianxiang Ai, Benjamin P. Cherniawski, Sean R. Parkin, Chad Risko, and Alejandro L. Briseno. Even-Odd Alkyl Chain-Length Alternation Regulates Oligothiophene Crystal Structure. *Chemistry of Materials*, 31(17):6900–6907, 2019.

[134] B. Lee and F. M. Richards. The interpretation of protein structures: Estimation of static accessibility. *Journal of Molecular Biology*, 55(3):379–IN4, 1971.

[135] Qianxiang Ai. wavemap. *Zenodo.*, 10.5281/zenodo.2658795, 2019.

[136] Manoj K. Jana, Chi Liu, Sven Lidin, David J. Dirkes, Wei You, Volker Blum, and David B. Mitzi. Resolving rotational stacking disorder and electronic level alignment in a 2d oligothiophene-based lead iodide perovskite. *Chemistry of Materials*, 31(20):8523–8532, 2019.

[137] E Mitscherlich. *Sur la relation qui existe entre la forme cristalline et les proportions chimiques*, volume 14. 1820.

[138] Kortney Kersten, Ramanpreet Kaur, and Adam Matzger. Survey and analysis of crystal polymorphism in organic structures. *IUCrJ*, 5(2):124–129, 2018.

[139] Graeme M. Day, James Chisholm, Ning Shan, W. D.Sam Motherwell, and William Jones. An assessment of lattice energy minimization for the prediction of molecular organic crystal structures. *Crystal Growth and Design*, 4(6):1327–1340, 2004.

[140] Jeni C. Sorli, Qianxiang Ai, Devin B. Granger, Kaichen Gu, Sean Parkin, Karol Jarolimek, Nicholas Telesz, John E. Anthony, Chad Risko, and Yueh Lin Loo. Impact of Atomistic Substitution on Thin-Film Structure and Charge Transport in a Germanyl-ethynyl Functionalized Pentacene. *Chemistry of Materials*, 31(17):6615–6623, 2019.

[141] Steffen Illig, Alexander S. Eggeman, Alessandro Troisi, Lang Jiang, Chris Warwick, Mark Nikolka, Guillaume Schweicher, Stephen G. Yeates, Yves Henri Geerts, John E. Anthony, and Henning Sirringhaus. Reducing dynamic disorder in small-molecule organic semiconductors by suppressing large-Amplitude thermal motions. *Nature Communications*, 7:10736, 2016.

[142] Hyunjoong Chung and Ying Diao. Polymorphism as an emerging design strategy for high performance organic electronics. *Journal of Materials Chemistry C*, 4(18):3915–3933, 2016.

[143] Chris D. Sheraw, Thomas N. Jackson, Dave L. Eaton, and John E. Anthony. Functionalized pentacene active layer organic thin-film transistors. *Advanced Materials*, 15(23):2009–2011, 2003.

[144] Jihua Chen, John Anthony, and David C. Martin. Thermally induced solid-state phase transition of bis(triisopropylsilylethynyl) pentacene crystals. *Journal of Physical Chemistry B*, 110(33):16397–16403, 2006.

[145] Ying Diao, Kristina M. Lenn, Wen Ya Lee, Martin A. Blood-Forsythe, Jie Xu, Yisha Mao, Yeongin Kim, Julia A. Reinspach, Steve Park, Alán Aspuru-Guzik, Gi Xue, Paulette Clancy, Zhenan Bao, and Stefan C.B. Mannsfeld. Understanding polymorphism in organic semiconductor thin films through

nanoconfinement. *Journal of the American Chemical Society*, 136(49):17046–17057, 2014.

[146] Jie Xu, Ying Diao, Dongshan Zhou, Yisha Mao, Gaurav Giri, Wei Chen, Nan Liu, Stefan C.B. Mannsfeld, Gi Xue, and Zhenan Bao. Probing the interfacial molecular packing in TIPS-pentacene organic semiconductors by surface enhanced Raman scattering. *Journal of Materials Chemistry C*, 2(16):2985–2991, 2014.

[147] Qianxiang Ai, Sean Ryno, and Chad Risko. Organic Crystals in Electronic and Light-Oriented Technologies (OCELOT) Database v0.01. 2019.

[148] Georg Kresse and Daniel Joubert. From ultrasoft pseudopotentials to the projector augmented-wave method. *Physical review b*, 59(3):1758, 1999.

[149] Grygoriy A. Dolgonos, Johannes Hoja, and A. Daniel Boese. Revised values for the X23 benchmark set of molecular crystals. *Physical Chemistry Chemical Physics*, 21(44):24333–24344, 2019.

[150] Atsushi Togo and Isao Tanaka. First principles phonon calculations in materials science. *Scripta Materialia*, 108:1–5, 2015.

[151] Binoy K. Saha. Thermal Expansion in Organic Crystals. *Journal of the Indian Institute of Science*, 97(2):177–191, 2017.

[152] Tomo Sakanoue and Henning Sirringhaus. Band-like temperature dependence of mobility in a solution-processed organic semiconductor. *Nature Materials*, 9(9):736–740, 2010.

[153] Grahame R. Woollam, Marcus A. Neumann, Trixie Wagner, and Roger J. Davey. The importance of configurational disorder in crystal structure prediction: the case of loratadine. *Faraday Discussions*, 211:209–234, 2018.

[154] Douglas B. Kitchen, Hélène Decornez, John R. Furr, and Jürgen Bajorath. Docking and scoring in virtual screening for drug discovery: Methods and applications. *Nature Reviews Drug Discovery*, 3(11):935–949, 2004.

[155] Christian Kunkel, Christoph Schober, Johannes T Margraf, Karsten Reuter, and Harald Oberhofer. Finding the right bricks for molecular legos: A data mining approach to organic semiconductor design. *Chemistry of Materials*,

31(3):969–978, 2019.

[156] Mahmoud Hammouri, Taylor M Garcia, Cameron Cook, Stephen Monaco, Sebastian Jezowski, Noa Marom, and Bohdan Schatschneider. High-throughput pressure-dependent density functional theory investigation of herringbone polycyclic aromatic hydrocarbons: part 1. pressure-dependent structure trends. *The Journal of Physical Chemistry C*, 122(42):23815–23827, 2018.

[157] Andrew L. Hopkins. Network pharmacology: The next paradigm in drug discovery. *Nature Chemical Biology*, 4(11):682–690, 2008.

[158] John E. Anthony, David L. Eaton, and Sean R. Parkin. A road map to stable, soluble, easily crystallized pentacene derivatives. *Organic Letters*, 4(1):15–18, 2002.

[159] G Kresse and O Lebacq. VASP manual, 2013.

[160] Wahyu Setyawan and Stefano Curtarolo. High-throughput electronic band structure calculations: Challenges and tools. *Computational Materials Science*, 49(2):299–312, 2010.

[161] Russell D Johnson. NIST Computation Chemistry Comparison and Benchmark Database, NIST Standard Reference Database Number 101 Release 14. *NIST, Gaithersburg, MD*, 2006.

[162] E A V Edsworth, P J Padley, and K Schofield. *Aromatic character and aromaticity*, volume 6. Cambridge University, Cambridge, 1970.

[163] Florian Weigend. Accurate Coulomb-fitting basis sets for H to Rn. *Physical Chemistry Chemical Physics*, 8(9):1057–1065, 2006.

[164] Adam Kubas, Felix Hoffmann, Alexander Heck, Harald Oberhofer, Marcus Elstner, and Jochen Blumberger. Electronic couplings for molecular charge transfer: Benchmarking cdft, fodft, and fodftb against high-level ab initio calculations. *The Journal of chemical physics*, 140(10):104105, 2014.

[165] Adam Kubas, Fruzsina Gajdos, Alexander Heck, Harald Oberhofer, Marcus Elstner, and Jochen Blumberger. Electronic couplings for molecular charge transfer: benchmarking cdft, fodft and fodftb against high-level ab initio

calculations. ii. *Physical Chemistry Chemical Physics*, 17(22):14342–14354, 2015.

[166] Larry A. Curtiss, Mark P. McGrath, Jean Philippe Blaudeau, Nancy E. Davis, Robert C. Binning, and Leo Radom. Extension of Gaussian-2 theory to molecules containing third-row atoms Ga-Kr. *The Journal of Chemical Physics*, 103(14):6104–6113, 1995.

[167] Michelle M. Francl, William J. Pietro, Warren J. Hehre, J. Stephen Binkley, Mark S. Gordon, Douglas J. DeFrees, and John A. Pople. Self-consistent molecular orbital methods. XXIII. A polarization-type basis set for second-row elements. *The Journal of Chemical Physics*, 77(7):3654–3665, 1982.

[168] Yan Zhao and Donald G. Truhlar. The M06 suite of density functionals for main group thermochemistry, thermochemical kinetics, noncovalent interactions, excited states, and transition elements: Two new functionals and systematic testing of four M06 functionals and 12 other functionals (Theoretical Chemistry Accounts DOI 10.1007/s00214-007-0310-x). *Theoretical Chemistry Accounts*, 119(5-6):525, 2008.

[169] Axel D. Becke. Density-functional thermochemistry. III. The role of exact exchange. *The Journal of Chemical Physics*, 98(7):5648–5652, 1993.

[170] Chengteh Lee, Weitao Yang, and Robert G. Parr. Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. *Physical Review B*, 37(2):785–789, 1988.

[171] S. H. Vosko, L. Wilk, and M. Nusair. Accurate spin-dependent electron liquid correlation energies for local spin density calculations: a critical analysis. *Canadian Journal of Physics*, 58(8):1200–1211, 1980.

[172] P. J. Stephens, F. J. Devlin, C. F. Chabalowski, and M. J. Frisch. Ab Initio calculation of vibrational absorption and circular dichroism spectra using density functional force fields. *Journal of Physical Chemistry*, 98(45):11623–11627, 1994.

[173] Roi Baer, Ester Livshits, and Ulrike Salzner. Tuned Range-Separated Hybrids in Density Functional Theory. *Annual Review of Physical Chemistry*,

61(1):85–109, 2010.

[174] Andy B. Yoo, Morris A. Jette, and Mark Grondona. SLURM: Simple Linux Utility for Resource Management. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, volume 2862, pages 44–60. Springer, 2003.

[175] Django Software Foundation. Django , 2018.

[176] Plotly Technologies Inc. Collaborative data science, 2015.

[177] Greg Landrum. RDKit : A software suite for cheminformatics , computational chemistry , and predictive modeling, 2011.

[178] Shyue Ping Ong, William Davidson Richards, Anubhav Jain, Geoffroy Hautier, Michael Kocher, Shreyas Cholia, Dan Gunter, Vincent L. Chevrier, Kristin A. Persson, and Gerbrand Ceder. Python Materials Genomics (pymatgen): A robust, open-source python library for materials analysis. *Computational Materials Science*, 68:314–319, 2013.

[179] A A Hagberg, D A Schult, and P J Swart. Exploring network structure, dynamics, and function using NetworkX. Technical report, 2008.

[180] Kealan J. Fallon, Peter Budden, Enrico Salvadori, Alex M. Ganose, Christopher N. Savory, Lissa Eyre, Simon Dowland, Qianxiang Ai, Stephen Goodlett, Chad Risko, David O. Scanlon, Christopher W.M. Kay, Akshay Rao, Richard H. Friend, Andrew J. Musser, and Hugo Bronstein. Exploiting Excited-State Aromaticity to Design Highly Stable Singlet Fission Materials. *Journal of the American Chemical Society*, 141(35):13867–13876, 2019.

[181] Edward F. Valeev, Veaceslav Coropceanu, Demetrio A. Da Silva Filho, Seyhan Salman, and Jean Luc Brédas. Effect of electronic polarization on charge-transport parameters in molecular organic semiconductors. *Journal of the American Chemical Society*, 128(30):9882–9886, 2006.

[182] Sereina Riniker and Gregory A. Landrum. Better Informed Distance Geometry: Using What We Know to Improve Conformation Generation. *Journal of Chemical Information and Modeling*, 55(12):2562–2574, 2015.

[183] Darko Butina. Unsupervised data base clustering based on daylight's fingerprint and Tanimoto similarity: A fast and automated way to cluster small and large data sets. *Journal of Chemical Information and Computer Sciences*, 39(4):747–750, 1999.

[184] Alexandr Fonari and Christopher Sutton. Effective Mass Calculator (EMC). 2012.

[185] Michael C Zerner. Semiempirical molecular orbital methods. *Reviews in computational chemistry*, 2:313–365, 1991.

[186] Nadine E. Gruhn, Demetrio A. Da Silva Filho, Tonja G. Bill, Massimo Malagoli, Veaceslav Coropceanu, Antoine Kahn, and Jean Luc Brédas. The vibrational reorganization energy in pentacene: Molecular influences on charge transport. *Journal of the American Chemical Society*, 124(27):7918–7919, 2002.

[187] Anubhav Jain, Shyue Ping Ong, Wei Chen, Bharat Medasani, Xiaohui Qu, Michael Kocher, Miriam Brafman, Guido Petretto, Gian Marco Rignanese, Geoffroy Hautier, Daniel Gunter, and Kristin A. Persson. Fireworks: A dynamic workflow system designed for highthroughput applications. *Concurrency Computation*, 27(17):5037–5059, 2015.

[188] S. R. Hall, F. H. Allen, and I. D. Brown. The crystallographic information file (CIF): a new standard archive file for crystallography. *Acta Crystallographica Section A*, 47(6):655–685, 1991.

[189] Paul Raccuglia, Katherine C. Elbert, Philip D.F. Adler, Casey Falk, Malia B. Wenny, Aurelio Mollo, Matthias Zeller, Sorelle A. Friedler, Joshua Schrier, and Alexander J. Norquist. Machine-learning-assisted materials discovery using failed experiments. *Nature*, 533(7601):73–76, 2016.

[190] M Waskom. seaborn: statistical data visualization. Python 2.7 and 3.5.

[191] R Todeschini and V Consonni. Handbook of chemoinformatics: from data to knowledge in 4 volumes, 2004.

[192] Bettina S. Basel, Johannes Zirzlmeier, Constantin Hetzer, Brian T. Phelan, Matthew D. Krzyaniak, S. Rajagopala Reddy, Pedro B. Coto, Noah E.

Horwitz, Ryan M. Young, Fraser J. White, Frank Hampel, Timothy Clark, Michael Thoss, Rik R. Tykwinski, Michael R. Wasielewski, and Dirk M. Guldi. Unified model for singlet fission within a non-conjugated covalent pentacene dimer. *Nature Communications*, 8(1):1–8, 2017.

[193] Johannes Zirzlmeier, Dan Lehnherr, Pedro B. Coto, Erin T. Chernick, Rubén Casillas, Bettina S. Basel, Michael Thoss, Rik R. Tykwinski, and Dirk M. Guldi. Singlet fission in pentacene dimers. *Proceedings of the National Academy of Sciences of the United States of America*, 112(17):5325–5330, 2015.

[194] Ewa Pastorczak, Antonio Prlj, Jérôme F. Gonthier, and Clémence Corminboeuf. Intramolecular symmetry-adapted perturbation theory with a single-determinant wavefunction. *Journal of Chemical Physics*, 143(22):224107, 2015.

[195] Adolf Baeyer. Ueber Regelmässigkeiten im Schmelzpunkt homologer Verbindungen. *Berichte der deutschen chemischen Gesellschaft*, 10(2):1286–1288, 1877.

VITA

Qianxiang Ai

Education

B.Sc. UESTC Zhongshan Institute, Zhongshan, China, 2012

M.Sc. Nanchang University, Nanchang, China, 2015

Awards

2018 College of Arts & Sciences Fellowship, University of Kentucky

2018 Outstanding Oral Qualifier Award, University of Kentucky

2016 Fast Start Award, University of Kentucky

2016 Steckler Fellowship, University of Kentucky

2015 Steckler Fellowship, University of Kentucky

2012 Outstanding Graduate, UESTC Zhongshan Institute

Publications

12. Evans Onyango, Philip Mannes, Alexandre Pletnev, Devin Granger, **Qianxiang Ai**, Chad Risko, John Anthony, and Gordon Gribble. Synthesis and Electronic Properties of a Linearly Fused Anthracene Dimer *Tetrahedron Letters*, 152182, 2020
11. Namal Wanninayake, **Qianxiang Ai**, Ruixin Zhou, Md Ariful Hoque, Sidney Herrell, Marcelo I Guzman, Chad Risko, and Doo Young Kim. Understanding the effect of host structure of nitrogen doped ultrananocrystalline diamond electrode on electrochemical carbon dioxide reduction. *Carbon*, 157:408–419, 2020.
10. Minakshi Bhardwaj, **Qianxiang Ai**, Sean R Parkin, and Robert B Grossman. An unusually short intermolecular N-H…N hydrogen bond in crystals of the hemi-hydrochloride salt of 1-exo-acetamidopyrrolizidine. *Acta*

Crystallographica Section E: Crystallographic Communications, 76(1):77–81, 2020.

9. Andrew M. Zeidell, Laura Jennings, Conerd K. Frederickson, **Qianxiang Ai**, Justin J. Dressler, Lev N. Zakharov, Chad Risko, Michael M. Haley, and Oana D. Jurchescu. Organic semiconductors derived from dinaphtho-fused s-indacenes: How molecular structure and film morphology influence thin-film transistor performance. *Chemistry of Materials*, 31(17):6962–6970, 2019.
8. Jeni C. Sorli, **Qianxiang Ai**, Devin B. Granger, Kaichen Gu, Sean Parkin, Karol Jarolimek, Nicholas Telesz, John E. Anthony, Chad Risko, and Yueh-Lin Loo. Impact of atomistic substitution on thin-film structure and charge transport in a germanyl-ethynyl functionalized pentacene. *Chemistry of Materials*, 31(17):6615–6623, 2019.
7. Anthony J. Petty, **Qianxiang Ai**, Jeni C. Sorli, Hamna F. Haneef, Geoffrey E. Purdum, Alex Boehm, Devin B. Granger, Kaichen Gu, Carla Patricia Lacerda Rubinger, Sean R. Parkin, Kenneth R. Graham, Oana D. Jurchescu, Yueh-Lin Loo, Chad Risko, and John E. Anthony. Computationally aided design of a high-performance organic semiconductor: the development of a universal crystal engineering core. *Chemical Science*, 10:10543–10549, 2019.
6. Tetiana Kharandiuk, Eman J. Hussien, Joseph Cameron, Romana Petrina, Neil J. Findlay, Roman Naumov, Wim T. Klooster, Simon J. Coles, **Qianxiang Ai**, Stephen Goodlett, Chad Risko, and Peter J. Skabara. Noncovalent close contacts in fluorinated thiophene–phenylene–thiophene conjugated units: Understanding the nature and dominance of O…H versus S…F and O…F interactions with respect to the control of polymer conformation. *Chemistry of Materials*, 31(17):7070–7079, 2019.
5. Kealan J Fallon, Peter Budden, Enrico Salvadori, Alex M Ganose, Christopher N Savory, Lissa Eyre, Simon Dowland, **Qianxiang Ai**, Stephen Goodlett, Chad Risko, and others. Exploiting excited-state aromaticity

to design highly stable singlet fission materials. *Journal of the American Chemical Society*, 141(35):13867–13876, 2019.

4. Edmund K Burnett, **Qianxiang Ai**, Benjamin P Cherniawski, Sean R Parkin, Chad Risko, and Alejandro L Briseno. Even–odd alkyl chain-length alternation regulates oligothiophene crystal structure. *Chemistry of Materials*, 31(17):6900–6907, 2019.
3. **Qianxiang Ai**, Karol Jarolimek, Samuel Mazza, John E Anthony, and Chad Risko. Delimited polyacenes: edge topology as a tool to modulate carbon nanoribbon structure, conjugation, and mobility. *Chemistry of Materials*, 30(3):947–957, 2018.
2. **Qianxiang Ai**, Yulia A Getmanenko, Karol Jarolimek, Raúl Castañeda, Tatiana V Timofeeva, and Chad Risko. Unusual electronic structure of the donor–acceptor cocrystal formed by dithieno [3, 2-a: 2', 3'-c] phenazine and 7, 7, 8, 8-tetracyanoquinodimethane. *The Journal of Physical Chemistry Letters*, 8(18):4510–4515, 2017.
1. Yi Wang, **Qianxiang Ai**, and Jiang Lu. Raft radical copolymerization of beta-pinene with maleic anhydride and aggregation behaviors of their copolymer in aqueous solution. *Journal of Polymer Science Part A: Polymer Chemistry*, 53(12):1422–1429, 2015.