Imperial College London

Predicting Structure-Property Relationships to Accelerate Chiral Organic Electronic Materials Discovery

Julia Alexandra Schmidt

2021

Department of Chemistry

Imperial College London

Submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy of Imperial College London

Declaration of Originality

I hereby declare that the work presented in this thesis is my own and where information has been derived by other sources, I have clearly referenced all external sources accordingly. For all third party copyright works, permission to reproduce the work has been sought and has been attached to this thesis. This work has not been previously submitted for the award of any other degree.

Julia A. Schmidt

London, August 2021.

Copyright Statement

The copyright of this thesis rests with the author. This thesis is licensed and available under a Creative Commons Attribution Non-Commercial No Derivatives licence. Under this license, researchers are free to copy, distribute or transmit the material in this thesis on the condition that they attribute it, that they do not use it for commercial purposes and that they do not alter, transform or build upon it. When reusing or sharing this work, researchers must adhere to and clearly state the licence terms to others by naming the licence and linking to the licence text. Where a work has been adapted, it should be indicated that the work has been changed and the changes should be described. Please seek permission from the copyright holder for uses of this work that are not included in this licence or permitted under UK Copyright Law.

Acknowledgements

I would like to express my profound gratitude to my supervisor Dr Kim E. Jelfs for not only giving me the opportunity to pursue a PhD in an area which I am deeply passionate about, but also for her patience in guiding me throughout this journey. Furthermore, I would like to thank my second supervisor Prof Jenny Nelson for helping me overcome various challenges along the way and for her continuous support of my work. I would like to thank Prof Claire S. Adjiman and Prof Sarah L. Price for providing me with their crystal structure prediction software. In particular, I want to express my gratitude towards Dr Isaac J. Sugden and Dr Louise Price for patiently answering all my questions.

The helicene research would have not been the same without my collaborators, which is why I would like to thank Prof Matt J. Fuchter and Prof Erin R. Johnson for insightful discussions and endless support with my publication (they thankfully never placed an upper bound on the number of draft iterations). I am grateful to Dr Luc LeBlanc and Joseph A. Weatherby for relaxing hundreds of crystal structures without asking too many questions. I would like to say thank you to my helicene buddy, Dr Francesco Salerno, for insightful (non)helicene discussions.

Furthermore, I would like to acknowledge all Jelfs Group members, past and present, for their support. In particular, I want to thank Dr Marcin Miklitz, Dr Enrico Berardo and Dr Lukas Turcani for welcoming me to the group. I would like to acknowledge Dr Andrew Tarzia and Dr Emma Wolpert for kindly taking the time to proofread my thesis. Furthermore, I would like to thank Dr Alejandro Santana-Bonilla, Dr Beth Rice and Dr Jarvist Frost for charge transport discussions.

I would like to acknowledge David Reger from the Jux group at the University of Erlangen-Nürnberg for a very stimulating collaboration. I would like to thank Dr Robert L. Peach for inspiring me with his passion for research and for introducing me to graph theory. My time in the com-

putational offices would have not been the same without Tristan Mackenzie, Francesca Vianello, Léonie Strömich and Astrid Stroobants. I am very grateful to Dr Jordi Aribas-Ariño, Dr Tommaso Francese, Dr Mercè Deumal and Prof Juan J. Novoa from the University of Barcelona for introducing me to computational materials research and for encouraging me to pursue a PhD.

I would like to whole heartedly thank my grandpa, Hugo Schmidt, who let me live in his summer house for almost a year so that I had a quiet place to write my thesis during the pandemic. Furthermore, I would like to acknowledge my grandma, Ulrike Schmidt, who from an early age inspired me with her eagerness to learn and her hard work to continuously push myself. I would like to thank my best friend, Julika Kramm, for encouraging me to work harder, but also showing me ways to enjoy and reward myself for small achievements along the way. Furthermore, I would like to thank the *WolvieG* gang for being the best home I could have had in London. I would like to acknowledge Scotty, Ellie and Ernst Hütchen for the relaxing breaks from thesis writing. There are many friends not mentioned here, but I am sincerely grateful to everyone of them who has supported me throughout this journey. I would like to thank Sam M. Werner for his endless belief in me and for constantly making me laugh. Finally, I would like to thank my parents, Martina and Bernd Schmidt, and my family for all the support and love I have received from them - I could not have completed my PhD without you guys.

Abstract

Molecular materials are challenging to design from scratch since the packing arrangements are based on a balance of the interplay of subtle soft intermolecular interactions. The specific packing arrangement a molecular material adopts determines its properties. Rational material design aims to create new materials with tailored properties (*e.g.*, charge-carrier mobility) but is hampered by the difficulty to predict the most stable solid-state structure for a given molecule theoretically and to obtain it experimentally. Once the structure prediction problem is resolved, structure-property relationships for each polymorph can be obtained.

This thesis predicts structure-property relationships of helicenes to assesses their organic semiconductor (OSC) suitability. First, two 2,2'-disubstituted [6]helicenes were computationally predicted to investigate the structure-property relationship of functionalised [6]helicenes arranging in (homo)chiral columns. The crystal structure prediction (CSP) results were analysed with respect to the accuracy of the methodology, the (dis)similarity of the crystal packing arrangements obtained across the CSP landscapes, and its charge transport properties.

Second, a dimer-based screening approach to assess the suitability of ~1,300 of terminally substituted [6]helicenes for their suitability as *n*-type OSC was developed and the screening results are presented. By applying CSP to the most promising molecules, the presence of the dimer motif used at the screening stage was confirmed. Furthermore, fluorination significantly improved electron transport in the molecular crystal by up to 200%, while side groups containing triple bonds largely lead to improved transfer integrals.

Finally, CSP was performed for naphthalene and helicenes of varying helicene length n (n = 3-12). Thousands of crystal structures were assessed for their degree of π-π, C-H$\cdots\pi$ stacking and the degree of parallel alignment of the long helical axes. It is found that helicenes more frequently

arrange in parallel to each other as the helicene length n increases, supposedly enhancing the formation of homochiral columnar packing arrangements.

List of Publications

Peer-reviewed publications and discussions stemming from this PhD thesis are:

- **Schmidt, J. A.**, Weatherby, J. A., Sugden, I. J., Santana-Bonilla, A., Salerno, F., Fuchter, M. J., Johnson, E. R., Nelson, J., & Jelfs, K. E., Computational Screening of Organic Semiconductors: Exploring Side-Group Functionalization and Assembly to Optimize Charge Transport in Chiral Molecules, Crystal Growth & Design, 2021, Accepted Article, https://doi.org/10.1021/acs.cgd.1c00473.

- Reger, D., Haines, P., Amsharov, K. Y., **Schmidt J. A.**, Ullrich, T., Bönisch, S., Hampel, F., Görling, A., Nelson, J., Jelfs, K. E., Guldi, D. M., & Jux, N., A family of superhelicenes – Easily tunable, chiral nanographenes by merging helicity with planar π-systems, Angew. Chem. Int. Ed., 2021, 60, 18073. https://doi.org/10.1002/anie.202103253

- Peach, R. L., Arnaudon, A., **Schmidt, J. A.**, Palasciano, H. A., Bernier, N. R., Jelfs, K. E., Yaliraki, S. N. & Barahona, M., HCGA: Highly Comparative Graph Analysis for Network Phenotyping, Patterns, 2021, 2, 4, 100227. https://dx.doi.org/10.2139/ssrn.3720300.

- Salerno, F., Rice, B., **Schmidt, J. A.**, Fuchter M. J., Nelson, J., & Jelfs, K. E., The influence of nitrogen position on charge carrier mobility in enantiopure aza[6]helicene crystals, Physical Chemistry Chemical Physics, 2019, 21, 5059-5067, https://doi.org/10.1039/C8CP07603K.

- Addicoat. M., Adjiman, C. S., Arhangelskis, M., Beran, G. J. O., Brandenburg, J. G., Braun, D. E., Burger, V., Burow, A., Collins, C., Cooper, A., Day, G. M., Deringer, V. L., Dyer, M. S., Hare, A., Jelfs, K. E., Keupp, J., Konstantinopoulos, S., Li, Y., Ma, Y., Marom, N., McKay, D., Mellot-Draznieks, C., Mohamed, S., Neumann, M., Nilsson Lill, S., Nyman, J., Oganov, A. R., Price, S. L., Reutzel-Edens, S., Ruggiero, M., Sastre, G., Schmid, R., **Schmidt, J. A.**, Schön, C., Spackman, P., Tsuzuki, S., Woodley, S. M., Yang, S., & Zhu, Q., Structure searching methods: general discussion, Faraday Discussions, 2018, 211, 133-180

Contents

Abstract **7**

List of Figures **11**

List of Tables **15**

Abbreviations **17**

Atom Colouring Legend **19**

1 Introduction **20**

1.1 Introduction to Polymorphism in Molecular Crystals 20

1.1.1 Polymorphism 21

1.1.2 Is This Material Polymorphic? 24

1.1.3 Crystal Structure Prediction 25

1.1.4 Crystal Packing Analysis 28

1.1.5 Challenges 31

1.1.6 Future Outlook 32

1.2 Organic Semiconductors 33

1.2.1 Applications 34

1.2.2 Polymorphism in Organic Semiconductors 34

1.2.3 Modelling Charge Transport 35

1.2.4 Experimental and Computational Work to Maximise Charge Mobility of Molecular Semiconductors 39

1.3 Chiral molecular materials 41

1.3.1 Helicenes 43

1.3.2 Columnar Packing in Chiral Materials 47

1.4 Thesis Outline 48

2 Theoretical Background **51**

2.1 Force Fields 51

2.1.1 The Williams W99 Intermolecular Force Field 54

2.2 Density Functional Theory 55

2.3 Dispersion Corrected Electronic Structure Methods . . . 58
2.4 GFN-xTB . . . 60
2.5 DLPNO-CCSDT . . . 62
2.6 Crystal Structure Prediction . . . 62
2.6.1 CRYSTALPREDICTOR . . . 63
2.6.2 DMACRYS . . . 65
2.7 Machine Learning in the Chemical Sciences . . . 66
2.7.1 Supervised Learning . . . 67
2.7.2 Graph Representations . . . 68
2.7.3 Kernels . . . 68
2.8 Charge-Carrier Mobility Calculations Using Semi-classical Marcus Theory . . . 70
2.8.1 Transfer Integrals . . . 71
2.8.2 Marcus Rate . . . 71
2.8.3 Inner-sphere Reorganisation Energy . . . 73
2.8.4 Mobility Anisotropy . . . 74

3 Structure-property Relationships of 2,2'-Diethynyl- and 2,2'-Dicyano-[6]helicenes: Do Chiral Columns Promote Charge Transport? 75
3.1 Introduction . . . 75
3.1.1 Chapter Outline . . . 79
3.2 Methods . . . 80
3.2.1 Crystal Structure Prediction . . . 80
3.2.2 Crystal Packing Analysis . . . 82
3.2.3 Crystal Charge-carrier Mobilities . . . 87
3.3 Results and Discussion . . . 88
3.3.1 Prediction of Experimentally Observed Polymorphs with CSP . . . 88
3.3.2 Identifying Recurring Structural Motifs Across the 2,2'-Dicyano- and 2,2'-Diethynyl-[6]helicene Crystal-energy Landscapes . . . 97
3.3.3 Graph Kernel-based Crystal Similarity Map . . . 101
3.3.4 Charge-carrier Mobilities Across the Crystal-energy Landscape . . . 106
3.4 Conclusions and Future Work . . . 116

4 Computational Screening of Chiral Organic Semiconductors: Exploring Side-group Functionalisation and Assembly to Optimise Charge Transport 118
4.1 Introduction . . . 119
4.1.1 Chapter Outline . . . 121
4.2 Methods . . . 122
4.2.1 Compound Screening . . . 122
4.2.2 CSP Analysis of Screening Candidates . . . 132
4.3 Results and Discussion . . . 134

4.3.1 Screening of Side-group Functionalised Molecules for Maximal Transfer Integrals . . . 134
4.3.2 Screening of Side-group Functionalised Molecules for Minimal Inner-sphere Reorganisation Energies . . . 139
4.3.3 Identifying OSC Suitability Trends from Screening . . . 141
4.3.4 Analysis of CSP Results . . . 148
4.4 Conclusions . . . 157

5 On the Effect of [n]helicene Length on Crystal Packing Interactions and Polymorphism 160
5.1 Introduction . . . 160
5.2 Methods . . . 164
5.2.1 Crystal Structure Prediction . . . 164
5.2.2 Crystal Packing Analysis . . . 164
5.3 Results and Discussion . . . 165
5.3.1 General CSP Results . . . 165
5.3.2 Crystal Stacking Interactions Behaviour Across the [n]helicene Polymorphic Landscapes . . . 175
5.4 Conclusions . . . 188

6 Conclusions 191

A Appendix to Chapter 3: Structure-property Relationships of 2,2'-Diethynyl- and 2,2'-Dicyano-[6]helicenes: Do Chiral Columns Promote Charge Transport? 224

B Appendix to Chapter 4: Computational Screening of Chiral Organic Semiconductors: Exploring Side-group Functionalisation and Assembly to Optimise Charge Transport 226

C Appendix to Chapter 5: On the Effect of [n]helicene Length on Crystal Packing Interactions and Polymorphism 234
C.1 CRYSTACK Technical Details . . . 234

List of Figures

1.1 Schematic illustration of polymorphism in *p*-aminobenzoic acid. 21
1.2 Illustration of charge transport regimes. 36
1.3 Axially chiral anticlockwise *M*- and clockwise *P*-[6]helicene enantiomers. 44
1.4 Illustrations of helicene applications in chiral optoelectronics and molecular recognition. 45
1.5 [6]helicene recurring structural motifs. 47
1.6 Illustration of chiral columnar packing in [6]helicenes. 49

2.1 Rigid-molecule CSP workflow. 63
2.2 Illustration of the four-point calculation scheme for the inner-sphere reorganisation energy. 74

3.1 Illustration of experimentally observed 2,2'-disubstituted-[6]helicene crystal structures with columnar packing. 78
3.2 Illustration of the three motifs (interlocked, back-to-back and homochiral translational) previously identified for [6]helicene. 83
3.3 Computational workflow of the crystal packing analysis. 85
3.4 The 2,2'-dicyano-[6]helicene crystal-energy landscape at the W99 and DFT-XDM level. 88
3.5 A molecular overlay of the experimentally observed and CSP predicted 2,2'-dicyano-[6]helicene crystal structures. 89
3.6 The relative energy and RMSD change of the 38 lowest-lying 2,2'-dicyano-[6]helicene crystals upon DFT-XDM relaxation. 90
3.7 The 2,2'-dicyano-[6]helicene crystal geometry change with respect to the associated energetic reranking error upon DFT-XDM relaxation. 91
3.8 The 2,2'-dicyano-[6]helicene 2D hirschfeld fingerprint plots of the 12 lowest-lying crystal structures. 92
3.9 The 2,2'-diethynyl-[6]helicene crystal-energy landscape at the W99 and DFT-XDM level. 93
3.10 A molecular overlay of the experimentally observed and CSP predicted 2,2'-diethynyl-[6]helicene crystal structures. 94

3.11 The relative energy and RMSD change of the lowest-lying 2,2'-diethynyl-[6]helicene crystal structures upon DFT-XDM relaxation. 94
3.12 The 2,2'-diethynyl-[6]helicene geometry change with respect to the associated energetic reranking error upon DFT-XDM relaxation. 95
3.13 The 2,2'-diethynyl-[6]helicene 2D hirschfeld fingerprint plots of the 12 lowest-lying crystal structures. 96
3.14 The 2,2'-diethynyl-[6]helicene 2D hirschfeld fingerprint plots of crystal structures containing chiral columnar packing. 97
3.15 The 2,2'-dicyano- and 2,2'-diethynyl-[6]helicene crystal-energy landscapes depicting whether the dimer-based motifs were found. 100
3.16 Graph PM kernel-based similarity map of the 2,2'-dicyano-[6]helicene polymorphic landscape. 102
3.17 Illustration of 2,2'-dicyano-[6]helicene crystal structures arranging in chiral columns. 103
3.18 Graph PM kernel-based similarity map of the 2,2'-diethynyl-[6]helicene polymorphic landscape. 104
3.19 Illustration of 2,2'-diethynyl-[6]helicene crystal structures arranging in chiral columns.105
3.20 Illustrations of [6]helicene crystalline packings. 108
3.21 Comparison of hole and electron mobilities across the lowest-lying [6]helicene, 2,2'-dicyano- and 2,2'-diethynyl-[6]helicene crystal structures. 110
3.22 Polar plots of the highest hole-transporting 2,2'-dicyano-[6]helicene polymorph (structure 6) in $cm^2\ V^{-1}\ s^{-1}$. Charge transport is highest along the back-to-back dimers of the chiral column, as indicated by the arrow in the *bc*-plane. There is no transport along the *ac*-plane. 112
3.23 Angular-dependent mobility of the highest electron-transporting 2,2'-dicyano-[6]helicene polymorph. 112
3.24 Angular-dependent hole and electron mobility of the experimentally observed 2,2'-dicyano-[6]helicene polymorph. 113
3.25 Illustration of the experimentally observed 2,2'-dicyano-[6]helicene polymorph. . . 113
3.26 Angular-dependent mobility of the highest hole-transporting 2,2'-diethynyl-[6]helicene polymorph. 114
3.27 Angular-dependent mobility of the highest electron-transporting 2,2'-diethynyl-[6]helicene polymorph. 115

4.1 Introduction of the homochiral translational motif and the intended dimer-based screening strategy. 120
4.2 The relative abundance of the recurring [6]helicene dimer motifs across the terminally substituted [6]helicenes obtained from the CSD. 123
4.3 The functional groups considered in the computational screening (part I). 124
4.4 The functional groups considered in the computational screening (part II). 125

4.5 Overview of the computational screening procedure. 126
4.6 The [6]helicene translational dimer separation scan and the separation-dependent transfer integral. 128
4.7 The 2,2'-dicyano-[6]helicene translational dimer separation scan and the separation-dependent transfer integral. 128
4.8 The 2,2'-diethynyl-[6]helicene translational dimer separation scan and the separation-dependent transfer integral. 128
4.9 The 7,10-dialdehyde-[6]helicene contour plot of 2D potential energy surface. 133
4.10 Distribution of transfer integrals at different minimum energy separations for both mono- and disubstituted helicenes. 135
4.11 Boxplot of minimum energy separations across the translational dimer model grouped by functional group. 136
4.12 Boxplot of transfer integrals grouped by functional group. 137
4.13 Illustration of the 6-triacetylene-[6]helicene LUMO. 138
4.14 Boxplot of inner electronic reorganisation energies per side group. 140
4.15 Boxplot of inner electronic reorganisation energies with respect to the π–system size. 142
4.16 Boxplot of the OSC suitability of functionalised [6]helicene molecules by comparison of the J/λ_i scores. 144
4.17 Boxplot illustrating the disubstitution position dependence of the $J_{elec}/\lambda_i^{elec}$ scores. 145
4.18 Boxplot illustrating the monosubstitution position dependence of the $J_{elec}/\lambda_i^{elec}$ scores. 146
4.19 Chemical diagrams of the four promising OSC molecular material candidates identified by the screening. 149
4.20 Crystal-energy landscapes and crystal structure illustrations for 13-fluoro-, 4,13-difluoro-, 6-triacetylene- and 7,10-dialdehyde-[6]helicene. 150
4.21 Comparison of electron mobilities across the lowest-lying 13-fluoro-, 4,13-difluoro-, 6-triacetylene- and 7,10-dialdehyde-[6]helicene polymorphs. 151
4.22 Distribution of electron mobilities for 13-fluoro-, 4,13-difluoro-, 6-triacetylene- and 7,10-dialdehyde-[6]helicene. 153

5.1 Molecular overlay and overview of [n]helicenes of varying helicene length n (n=3-12). 161
5.2 Chemical Diagram illustrating π-π and C-H$\cdots\pi$ stacking. 164
5.3 The number of (enantiopure) polymorphs obtained within the set of lowest-lying crystal structures per CSP landscape. 166
5.4 CSP landscapes for naphthalene and [3]-[7]helicenes. 167
5.5 CSP landscapes for [8]-[12]helicenes. 168
5.6 Molecular overlay of the experimentally observed structures with the respective predicted naphthalene and [3]helicene structures. 172

5.7 Molecular overlay of the experimentally observed structures with the respective predicted [6]- and [7]helicene structures. 172
5.8 Molecular overlay of the experimentally observed structures with the respective predicted [10]- and [11]helicene structures. 172
5.9 Illustration of the lowest-lying [12]helicene polymorphs with their long helical axes largely arranged in parallel to one another. 175
5.10 Crystal-energy landscapes for naphthalene and [3]-[7]helicenes coloured by the degree of π-π stacking. 177
5.11 Crystal-energy landscapes for [8]-[12]helicenes coloured by the degree of π-π stacking. 178
5.12 Histogram of the degree of π-π and C-H$\cdots\pi$ stacking interactions for naphthalene and [3]-[12]helicenes. 179
5.13 Crystal-energy landscapes for naphthalene and [3]-[7]helicene coloured by the degree of C-H$\cdots\pi$ stacking. 184
5.14 Crystal-energy landscapes for [8]-[12]helicene coloured by the degree of C-H$\cdots\pi$ stacking. 185
5.15 Crystal-energy landscapes for naphthalene and [3]-[7]helicene coloured by the proportion of molecules within each crystal arranged with their long axes parallel to one another. 186
5.16 Crystal-energy landscapes for [8]-[12]helicene coloured by the proportion of molecules within each crystal arranged with their long axes parallel to one another. 187

A.1 Louvain clustering of the 2,2'-dicyano[6]helicene PM graph similarity matrix. 225

B.1 Boxplot of the electron mobilities in the low-field limit grouped by side group. 233

C.1 Molecular surface area and SASA per [n]helicene molecule. 234

List of Tables

3.1 Comparison of relative energies of experimentally observed crystal structures at the W99 and DFT-XDM level. . . . 97
3.2 The maximum hole mobilities obtained within the polymorphic landscapes of [6]helicene, 2,2'-dicyano- and 2,2'-diethynyl-[6]helicene. . . . 107
3.3 Comparison of electron mobilities across the polymorphic landscapes of [6]helicene, 2,2'-dicyano- and 2,2'-diethynyl-[6]helicene. . . . 111
3.4 Comparison of hole and electron mobilities across the set of experimentally observed crystal structures. . . . 113

4.1 Comparison of XGBoost and kernel ridge regression scores. . . . 131
4.2 Top candidates for the mono- and disubstituted [6]helicenes according to both OSC measures ($J_{\text{elec}}/\lambda_{\text{i}}^{\text{elec}}$, low-field μ). . . . 147

5.1 Comparison of geometric parameters of the experimentally observed and geometry optimised [n]helicene molecules. . . . 163
5.2 Relative energy rankings of the experimentally observed structure within the respective crystal-energy landscape. . . . 166
5.3 Pearson correlation scores between the degree of π-π stacking and C-H$\cdots\pi$ stacking, lattice energy and density. . . . 176

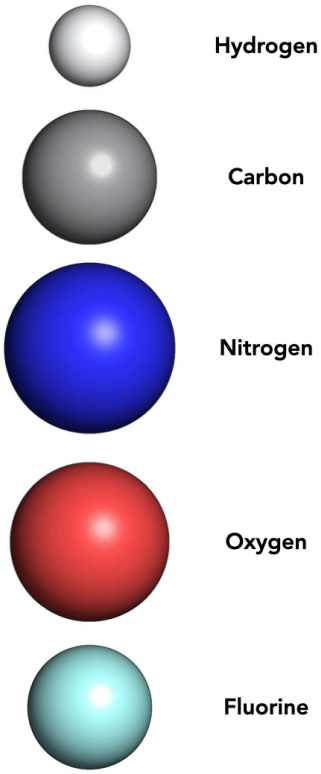
B.1 Results of the dimer-based motif search across the set of terminally functionalised [6]helicene molecular crystals obtained from the CSD. . . . 226
B.2 A list of the separations at which the translational dimer was found across the 6-triacetylene-[6]helicene polymorphs. . . . 230

Abbreviations

CCDC	Cambridge Crystallographic Data Centre
CCSD	Coupled Cluster Approximation with Singles and Doubles
CD	Circular Dichroism
CHELPG	Charges from Electrostatic Potentials using a Grid-based Method
CIF	Crystallographic Information File
COM	Centre Of Mass
CP	Circularly Polarised
CPL	Circularly Polarised Luminescence
CSD	Cambridge Structural Database
CSP	Crystal Structure Prediction
CT	Charge Transport
DFT	Density Functional Theory
DFTB	Density Functional Tight-Binding
DLC	Discotic Liquid Crystal
DLPNO	Domain Based Local Pair-Natural Orbital
DMA	Distributed Multipole Analysis
DMACRYS	Distributed Multipole Analysis (DMA) of Crystals
DSSC	Dye-Sensitised Solar Cell
DZP	Double Zeta Polarisation
EN	Electronegativity
FF	Force Field
GGA	Generalise Gradient Approximation
GPR	Gaussian Process Regression
HF	Hartree Fock
HOMO	Highest Occupied Molecular Orbital
KRR	Kernel ridge regression
LCP	Left-handed Circularly Polarised Light
LDA	Local Density Approximation
LUMO	Lowest Unoccupied Molecular Orbital
MAE	Mean Absolute Error

MEP	Molecular Electrostatic Potential
MD	Molecular Dynamics
ML	Machine Learning
OFET	Organic Field-Effect Transistor
OLED	Organic Light-Emitting Diode
OPLS	Optimised Potentials for Liquid Simulations
OPV	Organic Photovoltaic
OSC	Organic Semiconductor
PC	Packing Coefficient
PM	Pyramid Match
Pn	Pentacene
PES	Potential Energy Surface
RCP	Right-handed Circularly Polarised Light
RMSD	Root Mean Square Deviation
SASA	Solvent Accessible Surface Area
SCF	Kohn-Sham Self-Consistent Field
SMILES	Simplified Molecular Input Line Entry System
TES	5,11-bis(triethylsilylethynyl)
TIPS	6,13-bis(triisopropylsilylethynyl)
TS	Tkatchenko-Scheffler
TZVPP	Valence Triple-Zeta with Two Sets of Polarisation Functions
UFF	Universal Force Field
vdW	Van der Waals
W99	D. E. Williams' 1999 Potential Parameters
XDM	Exchange-Dipole Model
1D	1-Dimensional
2D	2-Dimensional
3D	3-Dimensional

Atom Colouring Legend



Chapter 1

Introduction

In this chapter, I will give an introduction to polymorphism, with the main focus on crystal structure prediction, organic semiconductors and chiral molecular materials. Polymorphic organic molecular materials with potential applications as organic semiconductors, in particular helicenes, will be introduced and the historical advances in the relevant areas of crystal structure prediction, chiral polymorphic materials, organic semiconductors and structure-property mapping will be reviewed. Polymorphism is a common phenomenon in crystalline materials, which adds an extra layer of complexity to the careful understanding of the soft non-covalent interactions required for intelligent crystal engineering. Aspects on how polymorphism can be leveraged for targeted functional materials design will also be discussed.

1.1 Introduction to Polymorphism in Molecular Crystals

Aliments, paints, pharmaceuticals, magnets or electronics, many of the products in our everyday life contain active crystalline components.[1,3] Intermolecular interactions, which bring the molecules in order, determine the structure and properties of the crystalline solid.[1] The symmetry and periodicity of the crystalline material can be measured by X-ray diffraction.[1] Some crystal morphologies are even visible to the naked eye (*e.g.* crystallite shape, Figure 1.1).[1]

Molecular materials are challenging to study because the molecules in the material are *not* connected by covalent or ionic bonds with clear directionality, designated bond lengths and an-

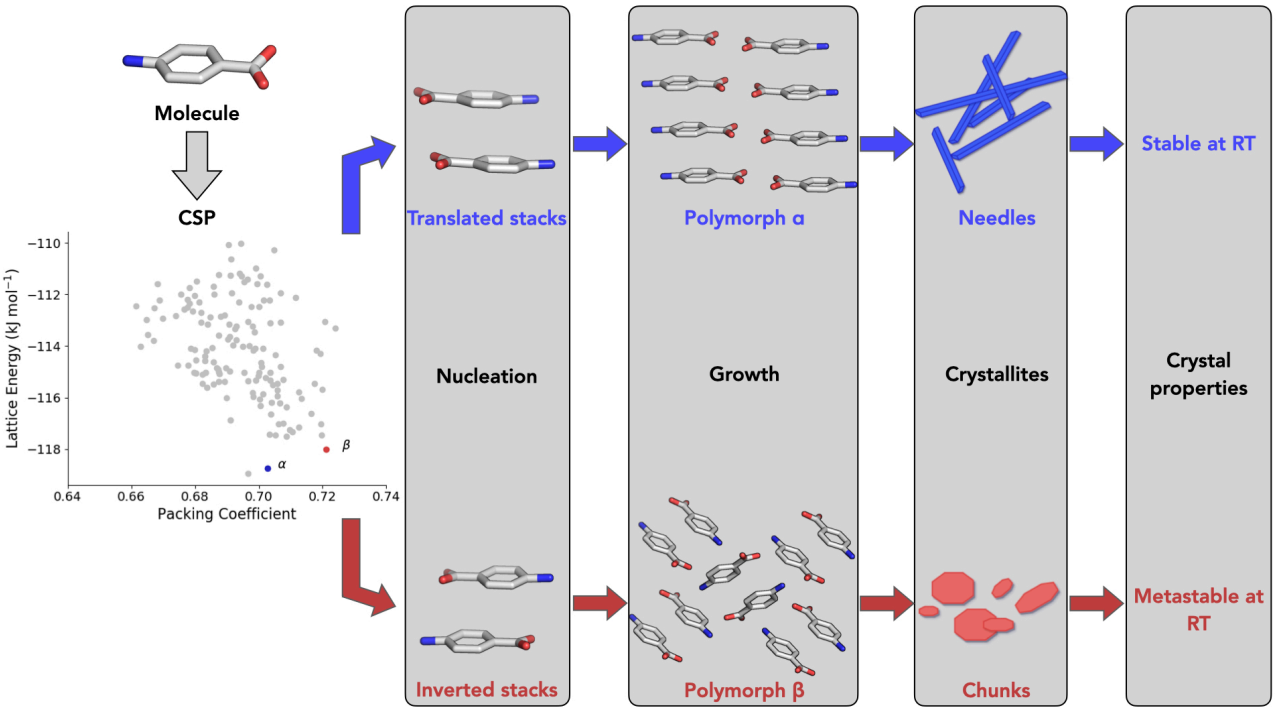


Figure 1.1: Schematic illustration of polymorphism in *p*-aminobenzoic acid (*p*ABA) adapted from References 1 and 2. Polymorphism, the ability of a molecule to (under distinct crystallisation conditions) assemble into multiple crystal structures results in different crystal shapes and properties. The *p*ABA α and β polymorphs are provided here. CSP computationally predicts the polymorphic landscape of all solid forms that could be experimentally observed. Hydrogens are omitted for clarity. Crystal-energy plot adapted with permission from Reference 2 Copyright 2021 Royal Society of Chemistry.

gles, molecular shapes and topologies. Instead, the subtle balance of the strongest possible intermolecular interactions and maximum space filling result in the molecular assembly of multiple solid-state structures.[1]

1.1.1 Polymorphism

Polymorphism (Greek: *poly* = many, *morph* = form) is having crystals with different structures and resultant properties, but the same molecular composition.[1] In other terms, polymorphism is the ability of a compound to crystallise in more than one crystal structure.[4,5] There have been numerous definitions of polymorphism in the literature, varying by whether allotropes, co-crystals (multicomponent crystals), salts or solvates are (not) included in the polymorphism definition.[6] Most definitions are based on McCrone (1965): "A polymorph is a solid crystalline phase of a given compound resulting from the possibility of at least two different arrangements of the molecules

of that compound in the solid state."[7] Some wider definitions allow interpretation of solvates as polymorphs: "When a substance can exist in more than one crystalline state it is said to exhibit polymorphism."[8] Sharma (1987) defined polymorphism as "different crystal forms, belonging to the same or different crystal systems, in which the identical units of the same element or the identical units of the same compound, or the identical ionic formulas or identical repeating units are packed differently."[9] Sharma clearly distinguishes between allotropism (*i.e.* the ability of a chemical element to arrange in two or more physical forms) and polymorphism, in contrast to Haleblian and McCrone (1969), who see allotropism as a special type of polymorphism.[5] Haleblian and McCrone (1969) give the element carbon, which can either pack in a diamond cubic structure or hexagonal graphite form, as an example of a special type of polymorphism.[5] Gavezzotti (2007) defines polymorphism based on three criteria: Polymorphs are crystals of (i) identical chemical composition, (ii) composed of molecules with the same connectivity (but allowing different molecular conformations, *e.g.*, rotamers) and (iii) with distinctly different periodic symmetry operations.[10] If there is no substantial change in three-dimensional symmetry operations, the structure should not be referred to as a polymorph, but as a modulation or phase of the same crystal.[10,11] Purohit and Venugopalan (2009) redefined polymorphism as the "ability of a substance to exist as two or more crystalline phases that have different arrangements or conformations of the molecules in the crystal lattice."[12] They also clearly distinguish between *packing polymorphism* (Figure 1.1), if the difference in solid form arises from the packing, and *conformational polymorphism*, if it arises from changes in the molecular conformation.

The interplay between the thermodynamics and kinetics during crystallisation gives rise to polymorphism.[1] Crystallisation, the process of forming a crystal from solution, can be divided into two stages (i) nucleation and (ii) growth.[1,13] The nucleation step in the formation of molecular crystals is the aggregation of molecules into initially reversible small clusters before they become stable nuclei.[6] The nucleation step is often governed by kinetics, which is why the form that crystallises tends to be the fastest crystallising form (the kinetic product) and not the thermodynamically most stable form.[6,13,14] Once nuclei have formed, the growth step is the incorporation of more molecules into the nuclei, growing into crystallites and crystals.[1,6] Crystallisation follows on one hand the interplay of soft intermolecular interactions to attain the strongest possible inter-

molecular interactions and on the other hand molecular shape with the goal of maximum space filling.[1] Thus, under different crystallisation conditions different crystal structures with the identical molecular composition and of similar crystal energies are obtained.[1] The initial nucleation step can provoke different molecular arrangements in the initial cluster, depending on how the initial seeding molecules of the crystal meet (Figure 1.1).[6]

Polymorphism was first reported in the context of crystallography by Mitscherlich in 1822, when he noted a difference in physical and chemical properties in different inorganic crystal structures of arsenate and phosphate salts.[15,16] In 1832, Wöhler and Liebig first reported polymorphism in an organic molecular crystal (benzamide).[17] The crystallographic complexity of benzamide poses open questions to date. Most recently, Fellah *et al.* proposed a new metastable polymorph of benzamide.[18] The area of polymorphism experienced a renaissance since its first discovery, due to technological advances (*e.g.* reduced time and cost to carry out crystal structure determinations, crystal structure databases), which allow easy comparison of obtained polymorphs[19] and the increasing realisation that a compound can have very different properties depending on its polymorphs, which is of major interest to the pharmaceutical industry.[4,11] Controlling crystal polymorphism has practical relevance to all chemical industries which deal with the organic solid state.[3,20] Only by knowingly forming the same polymorph, it can be assured that each batch of the product will have the same appearance, color, electrical conductivity, hygroscopicity, solubility, mechanical and thermal stability *etc.*, which will determine the tableting property of a pharmaceutical material and its processability under different mechanical procedures in the industrial production process.[1,6,21] Additionally, high-profile patent litigations, most famously the Zantac litigation involved questions of polymorphism, and resulted in a statistically significant increase in polymorphism citations.[22] To give an industry example of polymorphism, the commonly used drug paracetamol (*para*-acetylamino-phenol) has a monoclinic stable form I,[23] a metastable orthorhombic form II[24] and a highly metastable form III.[25] The packing arrangement of the drug determines its compressibility for drug delivery as tablets.[6] The most stable form I has hydrogen bonds which pack in waves and thus can be less easily compressed than the metastable form II consisting of flat hydrogen bonded sheets.[26] However, form II is hard to isolate, thus researchers tried to isolate form III to find out whether it would be compressible and discovered that form

III[27] converts into form II upon heating.[28,29] Paracetamol is a well-known drug, which has been commercially available in the US since the 1950s.[30] However, only in 2009 and thanks to high-quality temperature-dependent X-ray powder diffraction equipment, was it possible to measure the elusive form III and to shed some new light on its polymorphism.[27,30]

1.1.2 Is This Material Polymorphic?

Until today, it is not yet understood why some organic molecules are polymorphic whilst others are not.[1] Whether the phenomenon of polymorphism in organic molecules is pervasive[5] or elusive,[31] and only restricted to a few molecules, has been a long-lasting point of discussion. McCrone (1965) suggested that every compound is polymorphic and "the number of forms known for a given compound is proportional to the time and money spent on research of that compound."[7] Stahly found that 90% of the organic molecules screened had multiple crystalline (polymorphs, hydrates, solvates) and noncrystalline (amorphous) forms, and 48% were polymorphic.[32] An analysis on the occurrence of polymorphism in organic drug compounds in the European Pharmacopoeia by Braun (2008) yielded 42% of the compounds to exhibit polymorphism.[22] In the Cambridge Structural Database (CSD), polymorphism was observed for 34% of all unique compositions and 37±1% of single-component unique compositions.[22] In contrast to the large CSD, the smaller Eli Lilly set of compounds, which has been intensively screened for polymorphs, contained polymorphs for 25% of all unique compositions and 66±11% of single-component unique compositions.[22] Most recently, Bernstein (2020) stated that there may be a *possibility* for any compound to be polymorphic, but the crystallisation conditions required are difficult to attain due to the absence of any systematic empirical crystallisation design rules.[22]

Irrespective of whether polymorphism is a phenomenon pervasive to all organic molecular materials or not, the factors determining whether a compound is likely to have many different stable (or even reversibly interconvertible) solid forms (*e.g.*, 6,13-bis(triisopropylsilylethynyl)-pentacene (TIPS-Pn)) are not yet inherently understood.[33] Flexible molecules can give rise to *conformational polymorphism*, which refers to the crystalline structures not only differing by packing type, but also by the conformation of the single-molecule across different crystal structures.[11,34] In contrast, dynamic isomerism or tautomerism does not count as a type of

polymorphism, because it involves structural changes to the molecule.[22]

The tendency of flexible molecules to crystallise is reduced due to the presence of multiple conformers in solution, some of which are high-energy conformations.[14,35,36] The high conformational flexibility of the molecule slows down the crystallisation rate and defects can arise from the 'wrong' conformer being incorporated into the crystal.[36] Furthermore, it is hypothesised that the high conformational flexibility increases the likeliness of a high number of molecules in the asymmetric unit (Z'>1) of the crystal.[37] Most recently, in an attempt to find the molecular attributes that cause polymorphism, Cruz-Cabeza *et al.* found that flexible/large molecules are not less likely to be polymorphic than rigid/small molecules.[4] Furthermore, they observed chiral molecules to be less commonly polymorphic than achiral molecules.[4] If the molecule making up the molecular crystal is chiral, the number of theoretically possible polymorphs is increased further, because the molecule can then arrange in enantiopure (same handedness) and racemic polymorphs. More details on chirality affecting polymorphism will be provided in Section 1.3. Similarly, the hydrogen bonding ability of the molecule was found to slightly increase the material's probability of being polymorphic.[4]

1.1.3 Crystal Structure Prediction

Crystal structure prediction (CSP) is the computational prediction of multiple solid forms (polymorphs) for a given molecule starting from a two-dimensional chemical diagram only. CSP tries to predict the the set of most stable polymorphs (*i.e.*, the spatial arrangement of molecules with the lowest possible energy), often including experimentally observed polymorphs, by using the laws of physics, not human nor chemical intuition.[38] Thus, CSP can be divided into two problems, the *search* and the *ranking problem*. The *search problem* arises from the fact that the number of possible crystalline arrangements is astronomically large and the configurational space needs to be sampled in an efficient way.[38] Furthermore, the generated structures need to be relaxed to identify stable packing arrangements (local minima). The *ranking problem* is the challenge to accurately rank the generated crystal structures by their relative energies, even though the energetic differences tend to be extremely small.[38]

CSP has made huge progress within the last two decades and it is now often possible to pre-

dict the lowest-energy stable polymorph if accurate dispersion-corrected density functional theory (DFT-D) methods are used to rank the polymorphs obtained.[1,39,40] The accuracy of CSP methods is periodically objectively assessed in blind tests organised by the Cambridge Crystallographic Data Centre (CCDC).[41] In the blind test, research groups who develop CSP methods compete on the most accurate prediction of crystalline packings without knowing the experimentally observed polymorph. This way, the progress of an entire research area is regularly assessed to test any achievements and identify remaining challenges. Six blind tests have been held by the CCDC to date, in 1999,[42] 2001,[43] 2004,[44] 2007,[45] 2010[46] and 2015[47]. The seventh blind test (CSP2020) is forthcoming.

The fourth blind test[45] (CSP2007) was a major step forward compared to the third blind test[44] due to a higher success rate; correct predictions were submitted by 13 out of 14 participating research groups.[40,41] Here, 'successful participation' signifies that the experimental crystal structure was included among the three submitted predictions of a participant.[40] The number of correct predictions obtained for flexible molecules significantly increased compared to the three previous blind tests.[41,45] Given the progress in the field, the difficulty of the blind tests has continuously risen since the first blind test in 1999. Increases in difficulty come from providing no restrictions on the space group of the unknown polymorphs (*i.e.* the crystal structure one aims to predict),[44] to allowing Z'=2 crystal structures,[44] and to gradually predicting larger molecules. The largest molecule predicted in a blind test to date was a drug target with five solid forms, including a salt hydrate and a co-crystal in the sixth blind test.[47]

The list of computational approaches used by participating groups in the sixth blind test is long.[47] Among the 25 participating groups, a large variety of search algorithms ranging from evolutionary algorithms,[48–50] Monte Carlo simulated annealing (MC SA),[51,52] Monte Carlo parallel tempering,[53] CSD analogue searching,[54] random searches to quasi-random searching (based on Sobol sequences).[55] Evolutionary algorithms[49] and MC SA were overall less successful search algorithms in predicting the experimental structures, yet the computational search using Monte Carlo parallel tempering[53] generated the majority of the experimental crystal structures when combined with DFT-D methods with an empirical van der Waals dispersion correction (PBE-Neumann-Perrin) (12 structures predicted).[53,56] On the other hand, quasi-random search algorithms,[57–59] which search

for local minima and rank the obtained stable crystals by their lattice energies from distributed multipoles and the FIT intermolecular potential lead to similarly good results (10 structures predicted).[60] These CSP results were subsequently re-ranked to lower the crystal energy of the experimental crystal structure relative to the global minimum (the lowest energy structure) and to refine the crystal-energy landscape (the set of thermodynamically plausible solid forms).[61] As exemplified by the blind test, a typical CSP search algorithm can be divided into three parts:[40,62] (i) Firstly, space group searching and minimisation of tentative crystal structures at fast and medium quality; (ii) secondly, for the lowest-energy crystal structures, the second derivatives are computed with low-cost, medium quality methods to estimate the zero-point vibrational energies and finally, (iii) the set of lowest-energy crystal structures is relaxed at a more computationally expensive, high-quality level of theory, usually incorporating dispersion corrections at 0 K.[40,62,63] This three-step procedure will produce the final crystal-energy landscape.[40,62]

The fifth and sixth blind tests led to an enhancement in predictions via the usage of dispersion-corrected DFT methods (DFT-D) to obtain more accurate crystal-energy rankings at the final stage.[53] Refining crystal-energy landscapes with dispersion-corrected DFT has become the go-to approach in the field ever since.[62,64,65] Dispersion-corrected Hartree-Fock or DFT methods account for the long-range electron correlation, more commonly known as van der Waals (vdW) interactions. There are several dispersion correction methods currently in use: (i) semi-classical three-body dispersion correction (D3),[62] pairwise dispersion corrections *e.g.*, (ii) the exchange-hole dipole moment (XDM)[64] dispersion correction or (iii) the Tkatchenko and Scheffler (TS)[66] dispersion correction. However, the most computationally costly and most accurate dispersion correction is the many-body dispersion (MBD) interaction by Tkatchenko and Scheffler.[67,68] MBD reaches the lowest standard deviation (0.95 kcal mol^{-1}) in contrast to the XDM correction, which systematically underbinds hydrogen bonded systems, which would lead to a worse result if the three-body correction was used.[62] Overall, CSP has moved from its infancy to become a state-of-the-art approach to guide experimental polymorph screenings in the pharmaceutical industry and for use in materials discovery.[69,70]

Once the most accurate crystal-energy landscape has been calculated, the polymorphic landscape is plotted as energy-density plots. This is because the predicted crystal-energy landscapes

can be subsequently analysed via the rule of closest packing (density rule) for molecular crystals, according to which the most efficiently packed crystal structures should have the shortest intermolecular separations.[22,71] These separations are most commonly associated with hydrogen interactions, due to hydrogens marking the periphery of most molecules. The most energetically favourable crystalline packing should thus be the most efficiently packed form.[22,71] Likewise, the study of crystal-energy landscapes allows for quantification of the limits of molecular strain that can be imposed on a flexible molecule subject to conformational polymorphism imposed by the crystal lattice.[72] Thompson *et al.* find some molecules to be distorted by up to 20 kJ mol^{-1} by crystal packing forces.[72] In summary, as CSP becomes increasingly accurate and often experimentally observed crystal structures can be predicted successfully, our understanding of polymorphism is not sufficient to explain why some crystal structures are easily experimentally observed, whilst others seemingly similar structures, *i.e.* with similar packing motifs and lattice energies, are never observed.[73]

1.1.4 Crystal Packing Analysis

The set of lowest-energy crystal structures can be analysed for their packing similarity. The most fundamental approach for quantifying the (dis)similarity of two solid forms is to compute the Euclidean distance between their cartesian coordinates at the position of best possible atomic overlap. This is often referred to as the root mean square deviation (RMSD). A commonly used RMSD-based packing similarity tool is COMPACK,[74] which is available under the CCDC software license agreement. The basic approach of COMPACK can be described by the following steps: Firstly, the centre-of-mass (COM) distances between the central molecule and surrounding molecules in the crystal lattice are used to determine the N nearest neighbour molecules. The default number of nearest neighbours N is 14, but it can be increased for larger crystal structures. Once this cluster of nearest-neighbour molecules (of size N) has been determined it is used to represent the entire periodic crystal.[74] Secondly, the set of interatomic distances between neighbouring molecules is computed. Neighbouring molecules are pairs of molecules with at least one interatomic distance shorter or equal than the sum of the atomic van der Waals radii plus 2 Å.[74] At least as many interatomic distances as atoms in the largest molecular pair are computed and every atom has to

be used at least once.[74] The algorithm is biased towards shorter distances, *i.e.* first the shortest contact between two molecules is found, then the second shortest contact and so on. Two clusters representing to crystal structures are superimposed using the Kabsch algorithm to yield the best agreement between atomic positions of all atoms in the cluster.[74] For each matched atom in one cluster, the distance between its equivalent (matched) atom in the other cluster is used to derive a root-mean-square deviation (RMSD) value.[74] The RMSD is a quantitative measure of the similarity between two crystal structures and strongly increases with cluster size N, as the deviations for matching structures increases with increased intermolecular distance. However, this method has some shortcomings, if RMSDs of pairs of solid forms, which cannot be perfectly mapped on top of each other, are compared,[75] then, the RMSD drastically varies with the size of the two structures (number of atoms matched) that are being compared.[75] To solve this problem, a permutation-invariant distance between atomic configurations allowing for size-independent comparison of atomic configurations, without having to reduce the dimensionality of the data, has been introduced.[76] In contrast, Sadeghi *et al.* performed dimensionality reduction, by measuring the Euclidean distance in configuration space of two configurational fingerprint vectors, instead of directly comparing atomic positions.[77] With a focus on polymorphism and exploring the (dis)similarity of crystals across the crystal-energy landscape, Zhu *et al.* introduced a fingerprint based metric as a similarity measure for crystal structures.[78] Fingerprint based crystal similarity measures promise to aid faster exploration of the structure-energy landscape in high-throughput screenings, as part of swarm intelligence algorithms or minima hopping.[78,79] However, state-of-the-art fingerprint methods, such as RDkit molecular fingerprints, are unsuitable for the axially chiral molecules discussed in this thesis, since the three-dimensional axial chirality of the molecule cannot be represented by standard 2D fingerprint methods. Thus, one would be unable to determine the chirality of the helicene molecule, if represented as a connectivity-based RDkit fingerprint. The chirality of the helicenes is important to the polymorphism and the resulting charge-transport properties under investigation. Being able to reliably quantify the (dis)similarity of crystalline structures in the crystal-energy landscape paves the way for computationally efficient crystal structure prediction algorithms, such as evolutionary crystal structure prediction.[80] However, the chances of a random search algorithm finding the ground-state crystal

structure of a given system decrease with increasing system size and increasing number of degrees of freedom.[80] Being able to quantify the (dis)similarity across predicted crystal structures nonetheless allows for a more sophisticated exploration of the configurational space with greater computational efficiency.[80]

Machine learning (ML) has become increasingly prevalent in the field of molecular materials science and is used to anticipate a compound's properties prior to its synthesis in the laboratory.[81] For example, the (dis)similarity in molecular crystals has been previously quantified via similarity kernels, *e.g.*, the GRAPE,[82] SOAP,[83] or SOAP-REMatch[75] kernel.[84] A kernel is a function, which takes two input objects (*e.g.*, crystals) and implicitly (potentially nonlinearly) maps them into a higher dimensional feature space.[85–87] Then, the similarity of the two objects in the higher dimensional feature space can be computed from the inner product resulting in a positive-semidefinite value quantifying the similarity of the two objects.[86] The usage of positive-semidefinite similarity kernels is particularly appealing, because one obtains a quantifiable kernel-induced similarity distance between two crystal structures A and B, that can be used for further analysis.[84] Subsequent analysis techniques range from clustering or dimensionality reduction. One approach to using a kernel function is to first convert the raw crystal structure into graph representations.[88] Graph representations of crystal structures can be used as an indirect dimensionality reduction tool to quantify the (dis)similarity of crystalline structures. Ferré *et al.* used localised graph kernels, in particular the random walk graph kernel, to compute the similarity of two adjacency matrices, each representing one local atomic environment.[82] The computed graph kernel can then be used to predict chemical properties (*e.g.*, atomisation energies) using nonlinear regression algorithms such as kernel ridge regression (KRR) or gaussian process regression (GPR).[82,89,90] KRR, also known as kernel regularised least squares, implicitly projects the objects into a kernel space and finds a linear function in the kernel space which minimises the quadratic cost.[91] Gaussian process regression is a Bayesian nonparametric method for nonlinear regression.[92,93]

Graph representations of crystals have been previously employed in functional, optical and electronic materials discovery.[94,95] For instance, they have been used in searches for topological similarity (isotypism) of sulfates and binary inorganic compounds,[96] the computation of kernel-based crystal similarity maps[84] or graph convolutional neural networks to predict electronic prop-

erties.[95] Whilst high-level crystal-packing analysis methods have a high predictive power, they often lack interpretability. Being able to predict the properties using trained models or being able to match the crystals by their similarity, does not necessarily help identify the main intermolecular interactions driving molecular assembly nor why certain recurring structural motifs are prevalent over others. Low-level crystal packing analysis, such as distance or density based-methods are more interpretable.

Other crystal packing analyses aimed at linking the packing coefficients and the amount of short contacts,[97,98] *i.e.* how tightly the molecular crystals pack, to their respective cohesive energies.[99] This example and other crystal packing analyses are largely based on the assumption that the shortest interactions are the most dominant ones, in other words: "shorter is stronger" and these short interactions would drive molecular assembly. This assumption is debatable and often erroneous, because short interactions are destabilising due to the Pauli exclusion principle.[100] The destabilising power steeply increases with decreasing intermolecular distance.[100] A large number of stabilising contacts at longer distance can bring about a small number of destabilising short contacts; together they result in a net stabilisation of the crystalline structure.[99,100] Thus, crystal packing analyses must not stop at the sum of the atomic van der Waals radii (van der Waals distance), because the most stabilising contacts are found at a distance 0.4 Å larger than the van der Waals distance.[100]

1.1.5 Challenges

Even though there has been significant progress in CSP, being able to computationally predict crystal-energy landscapes does not mean that the predicted crystals can be obtained experimentally.[1,101] Instead, CSP often predicts many more stable polymorphs than can be observed experimentally.[61] This can have many reasons, such as computational approximations, *e.g.*, ignoring thermal effects[102] and the neglect of dynamic or static disorder, which is why the calculated crystal structure can represent more than one crystal structure.[61,101,103] Another reason is the fact that crystal-energy landscapes completely ignore kinetic effects, that in reality provoke rearrangements of the crystalline packing during the nucleation and growth step.[101,104,105] Thompson and Day reported that in the crystallisation of flexible molecules, high-energy conformers are often

selected if the high-energy molecular conformer is more extended (higher surface area) allowing for greater intermolecular stabilisation than low-energy solid forms.[72] Thus, the necessary crystallisation technique for synthesis of the computationally predicted crystal structure might not have been performed yet[61,106] or might not be possible.[31] Furthermore, the predicted structure could be metastable relative to other molecular compositions (stoichiometries).[107,108] Finally, the current limitations in the detection and characterisation techniques of polymorphs can limit the number of polymorphs observed.[101] Consequently, CSP either serves to eliminate or expand the experimental work required, *i.e.*, CSP can confirm that all low-energy solid forms have been found or suggest that many more might be feasible and should be screened for.[61]

Although CSP facilitates the prediction of tentative crystal structures with their relative thermodynamic stabilities, it cannot predict the rate constants required for the growth of each crystal.[1] Another major challenge is the fact that some materials do not crystallise at all and only form amorphous phases.[69] ML techniques, such as support vector machines and random forests have been used to predict the propensity of a material to be amorphous or crystalline.[109,110] Research efforts to understand the nucleation and growth of the crystals and to link nucleation rates to crystal structures have been made, yet many open questions remain.[104]

Understanding the interplay of the soft intermolecular interactions is key to determining the relative stability of polymorphs and why they undergo polymorphic transitions. However, since the soft intermolecular interactions are inherent to the specific material, conclusions about the importance of certain interactions are hampered by the fact that they cannot be looked at individually, only jointly, which hampers conclusions that can be generalised for a larger family of molecules. Attempts to assess the strength of intermolecular interactions include calculations of the mechanical stability under compression or shear stress. For example, the mechanical stability of multiple aspirin polymorphs has been computationally predicted to derive equilibrium conditions and polymorphic interconversion times.[111]

1.1.6 Future Outlook

CSP has the potential to lead to major advances in intelligent data-driven materials discovery.[112] Polymorphism enables computation of structure-property relationships for a molecular mate-

rial's set of crystals. These learned relationships can be used to make predictions on future unknown materials and their properties.[113] New 'hidden' porous materials and new solvent stabilised forms can be found by exploring structure-property relationships across CSP landscapes with high-throughput robotic crystallisation screens.[114] This hybrid computational–experimental approach by Cui *et al.* could be equally applied to other materials design challenges, such as organic electronics and drug discovery.[114]

Understanding polymorphism is key to many novel functional materials, especially in the field of organic semiconductors, where the charge transport properties depend on the frontier orbitals of the single molecule and the relative orientation of molecules in space - the crystal packing. Many benchmark organic semiconducting molecular materials have been reported to be polymorphic: pentacene,[115] rubrene,[116] TIPS-pentacene,[117] sexithiophene (6T), [1]benzothieno[3,2-b]benzothiophene (BTBT) derivatives.[118] In the same way, the structure-property relationship between crystal structure and its mechanical properties can be explored.[21] Polymorphism introduces extra flexibility to a material to be optimised for a tailored property, *e.g.*, solubility,[119] conductivity.[1,120] To give an example, TIPS-Pn has an eleven times increased hole mobility in its metastable form, than in its thermodynamically most stable form.[117] The factors determining charge transport in molecular crystals will be discussed in Section 1.2.

1.2 Organic Semiconductors

It might seem odd to study *organic* semiconductors (OSC), because *inorganic* materials are typically much better conductors, due to their free charge carriers.[121] However, what makes a material a semiconductor and differentiates it from an insulator is, that if free charge carriers are injected into the material, *e.g.* by high-energy radiation, the semiconducting material can reach charge-carrier mobilities comparable to covalently bound inorganic semiconductors or larger.[121] OSC are environmentally friendly alternatives to inorganic semiconductors with a range of new potential functionalities presented in the following.

1.2.1 Applications

Organic electronics can help manage the electronic waste (e-waste) problem of the 21^{st} century, because organic electronics processing does not require scarce elements, toxic materials or solvents, and the materials are fully biodegradable.[122] Another major advantage of organic electronics is the ability to solution process the material using a vast pool of printing technologies.[123] Printed electronics technologies are much more cost-effective, low-waste and faster alternatives to traditional processing techniques such as lithographic or chemical vapour deposition methods.[123,124] Organic electronics can be described as "cheap and cheerful",[125] because the material cost is low and they are environmentally friendly materials with added functionality (*e.g.*, flexibility). Organic molecular materials have promising applications as low-cost organic semiconductors (OSC) for use in organic light-emitting diodes (OLED), organic field-effect transistors (OFET) and organic photovoltaic (OPV) cells, sensors, photorefractive devices, which become increasingly important in the search for new and green technologies.[121,126,127] In OFETs, organic molecular materials could replace classical electronic devices with inorganic components based on silicon. The second main application is in light-harvesting devices (OPVs) and light-emitting devices (OLEDs).[128] New technologies only feasible with organic semiconductors include flexible computer screens,[125,129,130] ultra-thin displays, biodegradable electronics (*e.g.*, disposable phones),[122,123,125] and energy-harvesting smart materials (electronic (e-)skin, smart windows, smart clothing). Organic molecular materials are versatile due to their widely applicable electrical, magnetic and optical properties, and at the same time compatible with a wide range of other substrates, such as polymeric materials.[118]

1.2.2 Polymorphism in Organic Semiconductors

Device performance strongly depends on the ease with which charge-carriers move from one molecule to another, which in turn is determined by the solid-state spatial arrangement under ambient conditions.[118] Thus, polymorphism can be thought of as "an emerging design strategy for high-performance organic electronics".[118] OSC molecular materials tend to be polymorphic and able to exist in multiple crystalline packings under ambient conditions. Thus, different solid

forms, sometimes only with marginal changes to the crystal packing, can have considerably different performances.[118,121,131,132] The crystal packing of the single crystal within (semi-/poly-) crystalline layers or single crystals plays a central role for charge transport.[133] Thus, understanding the likelihood with which different solid forms are accessed and their respective organic semiconductor performance is vital to the search for high-performance organic semiconductors. Hence, materials scientists can leverage polymorphism to optimise for high-performance organic electronics via polymorphic transitions to a higher conducting metabstable polymorph with good processability.[120,132,134]

1.2.3 Modelling Charge Transport

There is a vast pool of computational methods available to describe charge transport in organic molecular materials in different regimes and at different levels of accuracy. There are two main paradigms for modelling charge transport in organic molecular crystals: hopping (weak-coupling) and band-like (strong-coupling) transport models, covering both far ends of the spectrum (Figure 1.2). Both models are still controversial at ambient temperatures.[135] Yavuz *et al.* analysed charge-carrier mobilities of ~60 organic molecules at room-temperature and observed that sometimes neither of the models could accurately predict a broad range of charge-carrier mobilities.[135] The hopping model reliably estimates low mobilities around ~1 $cm^2\ V^{-1}\ s^{-1}$ (weak-coupling).[135] Similarly, band-like transport tends to accurately predict mobilities above 50 $cm^2\ V^{-1}\ s^{-1}$ (strong-coupling).[135]

1.2.3 Hopping Transport

If charge transport is described as localised hopping or polaronic transport, charge carriers (electron or hole), hop from one site to another. The term polaron refers to a quasiparticle describing the interactions between excess charge carriers and a polarisable solid material.[137] Depending on the spatial extent of the polarisation cloud formed by the excess charge carrier localised within a potential well and causing a displacement of the neighbouring ions, one can distinguish between small and large polarons.[137] In the hopping regime, the charge carriers are fully localised at a given site (molecule) and there is an activation barrier between localisation sites.[128,131] The

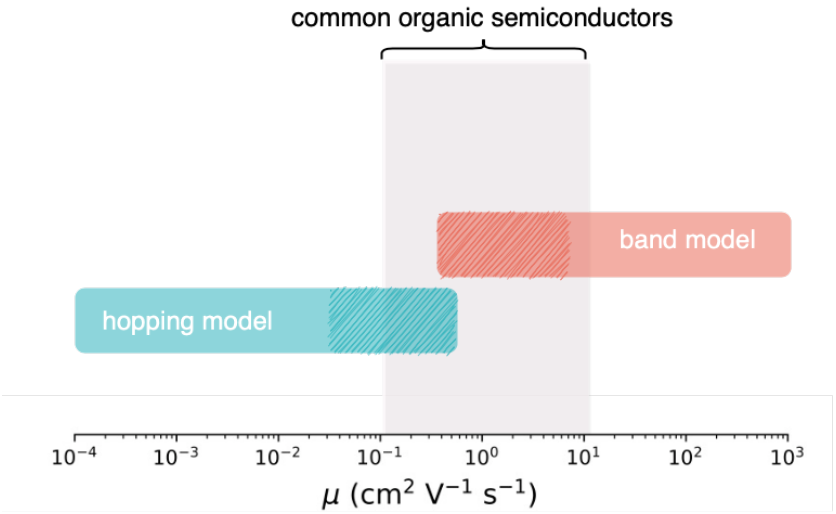


Figure 1.2: Illustration of the ranges of validity of hopping (teal) and band (red) models adapted from Reference 128. The grey box in the background denotes the range of mobilities observed for common organic semiconductors as determined by Oberhofer *et al.* based on Reference 136. The hatched areas describe ranges of where the validity of the model depends on the intrinsic materials' properties.

activation barrier depends on the transition state, the state in which the charge carrier is equally divided between the initial and final sites (molecules).[138] Hopping transport is observed in weakly coupled organic molecular crystals or in disordered polymer melts.[128] The hopping model tends to result in charge-carrier mobilities that agree qualitatively and quantitatively with experiments at room-temperature.[139] However, the hopping model ignores thermal fluctuations. As this "dynamic electronic disorder" can cause large variance in the transfer integrals, it can be the limiting factor to high charge-carrier mobility.[139,140]

1.2.3 Band Transport

Band-like transport is observed when transport is dominated by the transfer integral, J, *i.e.* J is considerably larger than the reorganisation energy, λ.[127] If transport is dominated by the transfer integral, carriers are highly delocalised and a propagating Bloch wave forms in a broad (valence or conduction) carrier band.[141] A Bloch wave is a family of periodic functions depending on the wave vector k. In the band structure, carriers propagate with an effective mass equivalent to the inverse of the band structure curvature[142–144] and charge transport is limited by crystal defects, impurities or collisions with phonons.[128,141] In simpler words, one tends to observe metal-like transport properties.[128] By definition, band transport assumes periodic systems and is thus re-

stricted to highly ordered and essentially defect-free systems. Hence, band theory cannot be used to describe charge transport in disordered polymers. In an ideal crystal, carriers are completely delocalised and carriers move without scattering. However, real crystals contain defects, that scatter carriers in different directions and thereby decrease mobility. Lowering the temperature reduces lattice vibrations and thus increases charge-carrier mobility in band-like transport.[141]

1.2.3 Intermediate Regime

The charge transport in most organic single crystals falls into the "intermediate regime", incorporating characteristics of both models,[128,145] *e.g.* as soon as high transfer integrals, J, are observed, the charges are no longer fully delocalised on separate molecules because the activation barrier, $\Delta G^{\dagger}$, separating the initial and final state can be overcome;[128] similarly large electronic bandwidths in organic molecular crystals can be strongly reduced by scattering.[145,146] There is no consensus in the literature as to what the exact limiting condition is at which the hopping regime breaks down, other than that the transfer integral, J, should be significantly smaller than the reorganisation energy, λ, within the hopping regime ($J << \lambda$). Thus, a potential limiting condition was attempted as $J \leq \frac{1}{2}\lambda$ (at $\Delta G^0 = 0$), where J is the transfer integral and λ is the reorganisation energy.[128] However, this condition is not good enough to serve as a condition for either transport because it ignores molecular vibration effects (including zero-point vibrations).[128] More precisely, whether hopping or band transport dominate the charge transport mechanism is determined by the ratio of relative contributions to the one-electron Hamiltonian describing the energy of the system.[127] There are five terms contributing to the Hamiltonian:[147]

$$\hat{H} = H_0 + H_1 + H_2 + H_3 + H_4, \tag{1.1}$$

where H_0 describes the electronic and vibrational excitation. H_1 stands for the electron transfer between two sites with the probability of transition being quantified by the transfer integral, J. Local (H_2) and nonlocal (H_3) electron-phonon coupling contributions to $\hat{H}$ represent the degree of dynamic disorder affecting charge transport. More precisely, they describe polaronic effects stemming from electronic excitation interactions and intra- and intermolecular vibrations.[127] The local

electron-phonon coupling (H_2) contribution implicitly includes intra- and intermolecular contributions to the reorganisation energy (λ).[127] H_4 describes the static disorder contribution, *i.e.* the structure of the molecular solid, and is thus thought of dominating charge transport in amorphous materials.[127] Now, if the local and nonlocal electron-phonon coupling contributions, H_2 and H_3, dominate compared to H_1 (assuming H_4 is minimal), a small polaron is formed which propagates via hopping.[127] In contrast, if H_1 dominates over H_2 and H_3, band transport is favoured. The competition between the different contributions to charge transport shapes the transport characteristics and renders a consistent description of charge transport troublesome.[127]

A generalised theory describing the smooth transition from hopping to band transport with decreasing temperature was proposed by Ortmann *et al.*[148] by theoretically describing phonon scattering (beyond the Holstein model[149]) to improve temperature and direction dependence (anisotropy) of charge mobility in wide-band and narrow-band crystals.[145] Charge mobilities based on the polaronic band model were able to reconcile both results based on the low-temperature Boltzmann equation and the Marcus theory for high temperature.[148,150] Furthermore, the results closely matched experimental measurements, as demonstrated for the naphthalene molecular crystal.[145] Nonetheless, the polaronic band model becomes problematic under ambient conditions.[151]

Large amplitude thermal molecular motions result in large fluctuations of intermolecular transfer integrals (dynamic disorder) leading to a 'transient localisation' of the carrier's eigenstates on timescales shorter than the period of molecular oscillations, thereby strongly restricting the carrier diffusion.[152,153] Thus, 'transient localisation' theory[152–156] can (in the limit where hopping theory leads to irrationally high transport rates) correctly describe a material's charge transport properties under ambient conditions. Molecular organic semiconductors with mobilities exceeding ~1 $cm^2\ V^{-1}\ s^{-1}$ are suggestive of being outside the hopping regime.[153] In addition to hopping theory, transient localisation theory requires calculations of the intermolecular vibrational modes to assess the disorder in J resulting from thermal fluctuations, the degree of anisotropy in these fluctuations, and the consequent effect on reducing mobility.[157] Fratini *et al.* proposed a map of transient localisation lengths of high-mobility OSCs to identify which pairs of nearest-neighbour transfer integrals prevent the OSC from dynamic disorder.[153] The dynamic disorder leads to a

reduced charge-carrier mobility.[153] The validity of the transient localisation theory has been proven via quantum localisation of charge carriers by solving the time-dependent Schrödinger equation.[158] Polaron size and charge mobility indicated that the charge propagates through the material, delocalised over 10-20 molecules, via diffusive jumps over several lattice spacings.[158]

1.2.4 Experimental and Computational Work to Maximise Charge Mobility of Molecular Semiconductors

Organic semiconducting molecular materials are used as the active layer in thin films of many organic electronics devices.[117] If a thin film form of a crystalline material is produced, usually via spin-coating or solvent vapour deposition, the crystalline packing of the thin film can differ from the single crystal, which can in turn affect electronic properties. Hence, understanding the polymorphic landscape can be of fundamental importance to obtaining materials with tailored properties. So far, the polymorphic landscape of only a few organic semiconductors has been experimentally explored.[117] Reasons for the lack of knowledge of structure-property relationships for multiple solid forms of a given OSC molecule are that most metastable forms are elusive and obtaining highly crystalline thin films of different solid forms can be challenging. Thus, Diao *et al.* proposed a material-independent technique to obtaining highly metastable polymorphs by enhancement of its kinetic stability via the nanoconfinement effect with flow-enhanced crystal engineering.[117] The method proved to be successful for TIPS-Pn and a BTBT derivative.[117] Jurchescu *et al.* reported two interconvertible fluorinated 5,11-bis(triethylsilyl-ethynyl) anthradithiophene polymorphs with small crystal packing differences leading to measurable differences in electronic properties in the single-crystal OFET device.[159] Mattheus *et al.* analysed the interlayer spacings and intralayer shifts to rationalise dominant interactions and derive structure-transport relationships for at least four different pentacene crystal structures observed.[160] Sorli *et al.* studied how atomistic substitution, *e.g.* changing silicon to germanium in TIPS-Pn to TIPGe-Pn, affects the molecular packing.[161] The germanium side group is similar in size, but higher in electron density. However, three different polymorphs were observed for the TIPGe-Pn single crystal, whilst only the brickwork polymorph was found for for TIPS-Pn.[161]

Computational approaches to optimising organic molecular crystals for charge transport in-

clude combining CSP with material property prediction to build energy-structure-function (ESF) maps for each molecule as proposed by Pulido *et al.*.[162] The ESF map approach aids the targeted discovery of functionalised materials by predicting the stable crystalline arrangements along with their functions or physical properties of interest.[163] This can be understood as a "cartography for materials discovery",[164] because it allows researchers to choose the most promising molecule based on its solid forms and desired target properties, prior to its synthesis.[162,164,165] Aitchison *et al.* used ESF maps to determine whether three given molecules were likely to form porous or π-stacked structures optimal for photocatalytic applications.[166] The ESF maps generated for 1,3,6,8-tetra(4'-carboxyphenyl)pyrene (TBAP) revealed a high propensity for π-stacked columns in the solid-state.[166] Furthermore, ESF maps highlighted that depending on the crystal packing, TBAP can be 20-200 times less photocatalytically active.[166] Musil *et al.* computationally investigated the variation of charge mobility across the crystal-energy landscape of pentacene.[84] Based on CSP results, the authors used the SOAP-ReMatch kernel to quantify the similarity across the crystal structures to identify clusters of highly similar packings.[84] Furthermore, they used ML techniques to predict charge-carrier mobilities and reduce the overall computational cost.[84] Shi *et al.* studied the intermolecular contacts in naphthalene and pentacene and found that extending the π-system will provide good contacts, boost transfer integrals and suppress unfavourable lattice vibrations.[167] Kunkel *et al.* screened the entire Cambridge Structural Database (CSD) to obtain statistically relevant structure-property relationships and identify molecular scaffolds (side groups), which lead to consistently improved charge-transport.[168] Likewise, Wang *et al.* observed that rigid extended π-systems and cyanation reduce the intramolecular reorganisation energy.[139,169] Atahan-Evrenk *et al.* demonstrated that (kernel) ridge regression and deep neural networks are able to predict intramolecular reorganisation energies using a dataset of 5631 OSC molecules.[170]

Pyzer-Knapp *et al.* successfully demonstrated that accurate multilayer neural networks can be trained from only a proportion of the Harvard Clean Energy Project dataset to accurately predict the frontier orbitals and power conversion efficiency from the molecule.[171] Then, if ML techniques are combined with high-throughput screening, much larger databases (than generated via *ab initio* calculations) can be obtained because the properties of a molecule are inferred from existing results and do not require explicit computation.[171] Larger and more diverse databases accelerate

electronic materials discovery, since a much larger chemical space is explored.[171] Overall, there has been substantial progress in computationally accelerated organic electronic materials discovery and in the development of electronic materials design principles, yet many open challenges remain.[172,173] With scientists embracing ML to obtain statistical solutions, sharing their data and open-source packages, the number of reported successful applications are rapidly growing and only the surface of possibilities has been scratched.[81,174] This is the advent of a new generation of computational science with the power to revolutionise the field.[81]

1.3 Chiral molecular materials

Chirality is omnipresent in biological systems (*e.g.*, DNA, proteins) and observed in many physical systems (*e.g.*, chiral fermions,[175] chiral superconductivity[176]).[177] Chiral functional materials design for use in chiral photonics could include applications in quantum optics or optical spintronics (*e.g.*, as spin filters[178]).[179] Chiral functional materials can be used as chiroptical switches with promising applications in data storage, optical displays, light modulators or other sectors of technological relevance.[180,181]

Polymorphism in chiral molecular materials is more complex than polymorphism described previously, because now instead of molecular conformers and crystalline packing arrangements, causing conformational and packing polymorphism, an additional degree of freedom is added to the polymorphic landscape: crystal structures can now be also divided by their chirality (racemic or enantiopure). If the crystal is enantiopure, only chiral isomers with the same handedness are found in the solid form, if it is racemic, equal amounts of both chiral isomers are found. The existence of several crystal structures for both the racemic and the enantiopure version of a compound is called 'double polymorphism' and has been reported for diprophylline (DPL)[14], *n*-hexylsuccinic acid[182] and the anti-sleepiness drug, modafinil.[183]

If discussing chirality in molecular crystals, one has to carefully distinguish between three different types of chirality: (i) The chirality of the molecular component, (ii) the chirality of the space group, and (iii) the chirality of the crystal structure.[184] A chiral crystal structure needs to crystallise in one of the 65 Sohncke space groups, because these space groups represent the possible symmetries of chiral objects.[184,185] However, only 22 of the 65 Sohncke space groups are chiral space

groups. A chiral space group only contains symmetry operations of the first kind (*i.e.* rotations, translations) and hence no improper rotations (rotoinversions).[184] The 22 chiral space groups can be grouped into 11 enantiomorphic pairs.[184] Helicenes are chiral molecules, thus the molecular components are chiral. Helicenes can pack in Sohncke space groups in chiral crystal structures. If the helicene crystallises into a non-Sohncke space group, the crystal structure will be achiral. If the helicene crystallises into a chiral structure and the crystal symmetry equals one of the 22 chiral space groups, the material is chiral by all three aspects (i-iii).[184]

Electronic circular dichroism (CD) is the differential *absorption* of left- and right-handed circularly polarised light (LCP and RCP) of intrinsically chiral chromophores or chromophores in a chiral environment in the ground electronic state.[177,186,187] Circular dichroism is measured using CD spectroscopy as the molar circular dichroism ($\Delta\epsilon$) with respect to the wavelength (λ):

$$\Delta\epsilon(\lambda) = \epsilon_L(\lambda) - \epsilon_R(\lambda), \tag{1.2}$$

where $\epsilon_L(\lambda)$ and $\epsilon_R(\lambda)$ denote the absorption intensities of the LCP and RCP components respectively.[186] Similarly, circularly polarised luminescence (CPL) is the circular dichroism (CD) lumninescence analogue and reflects the intensity difference in *emission* (ΔI) of left- and right-handed circularly polarised light (LCP and RCP) of chiral fluorophores or fluorophores in a chiral environment in the excited electronic state at a given wavelength (λ_{em}).[186,188] CPL is stated as a dissymmetry factor:

$$\Delta I(\lambda) = I_L(\lambda) - I_R(\lambda), \tag{1.3}$$

where $I_L(\lambda)$ and $I_R(\lambda)$ denote the emission intensities of the LCP and RCP components respectively.[186] Chiral materials with excellent CPL can be used in future optoelectronic bio-sensing devices, to for example monitor target binding events via a change in CPL signal.[186] Esposito *et al.* successfully probed triple-helical chiral nanowires to show efficient CPL with high signal-to-noise ratio and high optical activity sensitivity to structure orientation.[189]

1.3.1 Helicenes

Helicenes are a class of chiral compounds characterised by *ortho*-fused benzenoid or other aromatic rings arranged in a cylindrical helix.[190–192] Helicenes are axially chiral molecules with an either left- or right-handed helix (*P*- or *M*- chirality, Figure 1.3), meaning they either rotate circularly polarised light clockwise (dextorotatory) or counterclockwise (levorotatory). Helicenes were first named, synthesised and classified by Newman in 1955.[193,194] Helicene chemistry dates back over 100 years, [4]helicene and [5]helicene were the first reported helicenes in 1913 and 1918 respectively.[191,195] Helicenes have been called "molecules in distress", because the normally planar π-system architecture is distorted due to strong steric strain.[196] Helicenes have thus been classified as edge case molecules, because they violate the norm of planar aromatic architecture.[196] The strained geometry provides energetic stabilisation;[191] similar to the double helix in DNA, helicenes are intrinsically chiral by shape and the chirality extends along the entire polyaromatic molecule. Helicenes of different axial chain length, [5]-[9]helicenes, can thermally racemise at elevated temperatures (57-295 °C), *i.e.* convert from the left-handed to the right-handed helix and *vice versa*.[190] The chirality of the helicenes affects polymorphism differently depending on the chain length. If interconversion between the left- and right-handed helix is thermally feasible, racemic crystal structures can readily form (*e. g.* [6]helicene lowest-lying intergrowth structure).[197] However, as the axial chain length increases and thermal racemisation is no longer energetically feasible, one is more likely to grow only enantiopure forms of the chirality that initially formed. Thus, the chirality of the helicene and its axial chiral length directly affect the polymorphism observed.

The local aromaticity of helicenes is still high, given its strained geometry. Thus, helicenes can serve as good π-donors if used as chiral ligands in metal complexes.[198–201] The strained helicene geometry can be relaxed if a helicene is used as a 'proton sponge'. This is possible if the helicene molecule has exceptionally high basicity constants, *e.g.*, diazahelicenes, in which nitrogens on both ends of the helicene backbone can hydrogen bond to a hydrogen to relieve its strained geometry.[202–204] The basicity of 1,14-diaza[5]helicene is reduced by two orders of magnitude compared to the basicity of 1,12-diaza[4]helicene.[204] These are only a few examples of the chemical property diversity of chiral helicene molecules.

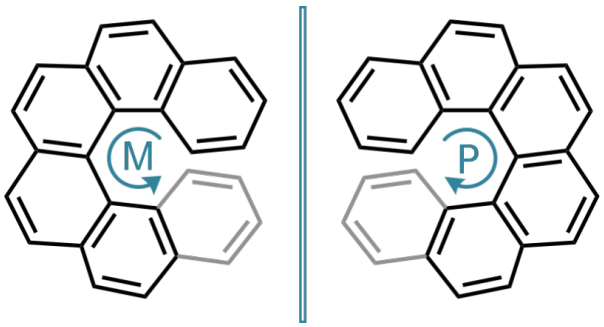


Figure 1.3: Axially chiral anticlockwise *M*- and clockwise *P*-[6]helicene enantiomers.

1.3.1 Applications

Helicenes are, due to their screw-like shape, chiroptical active materials whose photophysical response can be measured using molecular electronic circular dichroism (CD).[205,206] Helicenes have been tested as a new type of organic spin filter,[178] as antiglare filters in CP-OLED technology[207] or as a new type of chiroptical switches.[181,192,198] Helicene-based chiral molecular switches can be manipulated using a vast set of stimuli ranging from light-, redox- and pH-triggered helicene-based switches.[192] The range of external stimuli and its effects can be captured by different techniques, promising attractive applications as functional materials.[192] Helicenes are particularly valuable building blocks for chiral switches due to their large chiroptical response and thus very pronounced optical rotation values.[181] Most recently, [6]helicene derivatives, S-shaped double [6]helicenes, have been reported to form open-shell diradicals with a small singlet-triplet gap, giving rise to measurable magnetic susceptibility.[208]

Helicenes can be used in chiral molecular recognition[181,198,210,211] of nanomaterials, because they have chirality-dependent inter(nano)molecular attractive interactions, which result in the self-aggregation and precipitation of aggregates of different chirality.[212] For example, depending on whether helicenes form homochiral or heterochiral self-aggregates with gold nanoparticles, aggregates of different size and different interaction strength between chiral gold nanoparticles are observed (Figure 1.4).[212] In gold nanoparticle aggregates, the strongest interaction strength was observed between two homochiral gold particle aggregates of opposite handedness.[212] Helicenes have a tendency to form more stable complexes if pairs of the same configuration of helicenes are used, rather than pairs of enantiomeric helicenes (opposite handedness).[210,212] This observation does not only apply to self-aggregation, but also crystallisation, charge transfer complex-

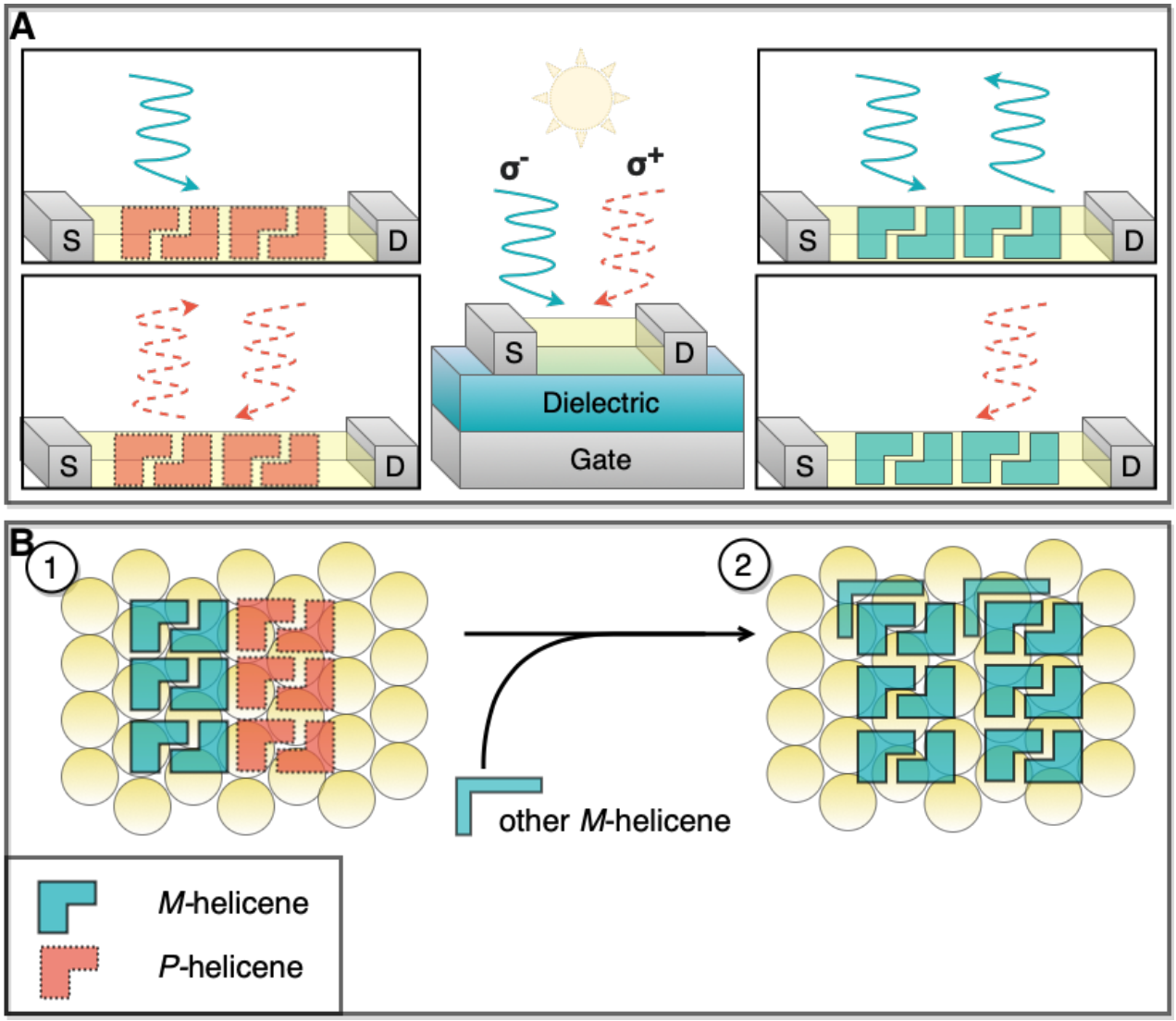


Figure 1.4: Illustrations of two helicenes applications: (A) Chiral optoelectronic application: bottom-gate, bottom-contact circularly polarised light-detecting OFET adapted from Reference 179. The active layer (yellow) is either *P*-(+)-1-aza[6]helicene (red) or *M*-(–)-1-aza[6]helicene (blue). OFETs with a *P*-(+)-1-aza[6]helicene layer show a response to right-handed (σ^+), but not left-handed (σ^-) circularly polarized light and the opposite is true for *M*-(–)-1-aza[6]helicene. (B) Stereoselective molecular recognition application[209]: Adsorption of [5,6, 9,10]-dibenzo-[5]helicene racemate on a Au(111) surface leads to homochiral domains of either chirality. Upon adding 26% of *M*-[7]helicene into the racemate monolayer, only homochiral domains of the *M*-helicenes are observed. Figure based on Reference 209.

ation, layer structure formation and double helix formation.[210] An area of molecular recognition helicenes are widely employed in is asymmetric catalysis for enantioselective synthesis.[213,214] Furthermore, (hetero)helicenes have found photovoltaic applications as building blocks for organic donor-acceptor π-conjugated (D-π-A) dyes on a TiO_2 surface for use in dye-sensitized solar cells (DSSCs).[215–217] Highly stable helicene cations can be used as organic optical active dyes.[218] Additionally, helicenes can be used as inorganic dyes in helical metallocenes.[198,219–221] Katz *et al.* synthesised helicenes which can give rise to oligomeric[222] or polymeric[223] metallocenes. Helicenes can further be employed as small monomolecular circularly polarised luminescence (CPL) dyes.[224] CPL is a promising property for new technologies, such as quantum computing[225], OLED displays with enhanced contrasts[226] or bioresponsive imaging technologies,[188] because CPL produces three types of signals (+,- and 0).[224] In chiral photonics (*e.g.* for use in quantum optics or optical spintronics applications), helicenes can serve as chiral OSC additives in a polymer based OFET devices to couple directly to circularly polarised light.[179] Thus, the device would have a chiral response, with a right-handed helical molecule responding to right-handed circularly polarized photons and *vice versa* (Figure 1.4A).[179] Most recently, Wan *et al.* managed to tune and invert the sign of the CP light as a function of the active layer thickness.[227] There are many more applications of helicenes, such as in liquid crystals,[228] chiral langmuir-blodgett films,[229] self-assembly[198] or chiral surface science.[230,231] Furthermore, helicenes are extensively used in asymmetric synthesis/catalysis, which is beyond the scope of this thesis.[191]

This thesis will focus on the application of helicenes in organic electronics, more precisely in organic semiconducting chiral molecular materials. A hole mobility of up to 2.1 $cm^2\ V^{-1}\ s^{-1}$ for double thiahelicene single-crystals has been recently measured in p-channel OFETs.[232,233] For the hole transporting (+)-1-aza[6]helicene OFET, a high on/off ratio of 1×10^3 and a mobility of $1 \times 10^{-4}\ cm^2\ V^{-1}\ s^{-1}$ were observed.[179] The (+)-1-aza[6]helicene mobility[179] is three orders of magnitude higher than the value reported for the racemic tetrathia-[7]-helicene OFET ($1.7 \times 10^{-7}\ cm^2\ V^{-1}\ s^{-1}$).[233] The high on/off ratio quantifies the average charge carrier drift velocity per unit electric field and drain-source current gain when device is turned on.[233] Previously, Rice *et al.* categorised the set of lowest-energy [6]helicene molecular crystals into recurring structural motifs (Figure 1.5) to determine motifs beneficial/detrimental to charge transport.[197] In summary, he-

licenes have great potential applications relevant to functional OSC materials design, because of their chemical property diversity, relatively rigid conformational space, large chiroptical response, and vast range of electronic properties controlled by the solid-state packing.

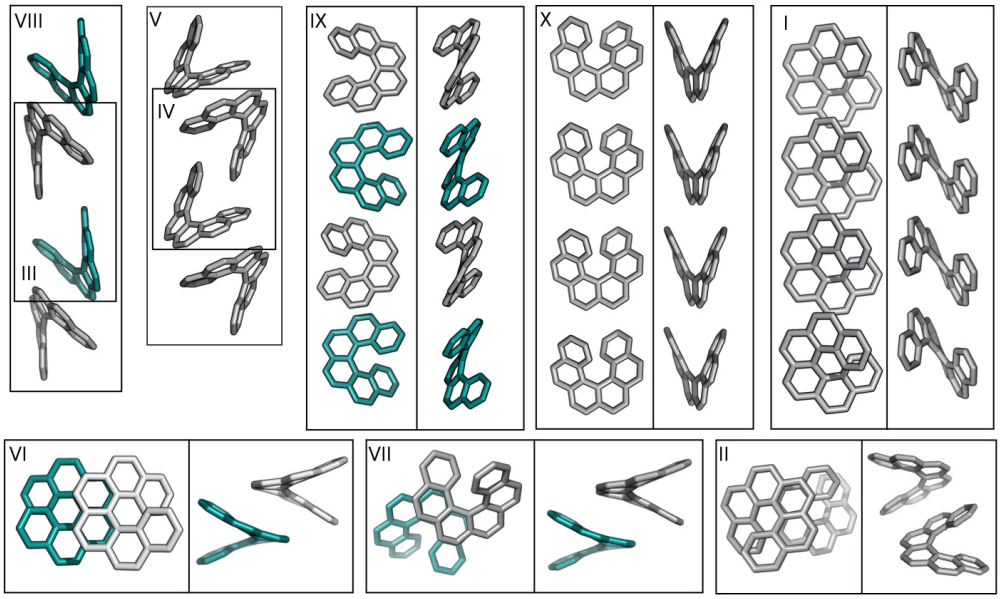


Figure 1.5: [6]helicene packing motifs adapted from Reference 197. Homochiral packing motifs are shown in grey, heterochiral motifs are grey and blue. The homochiral translational motif (I) was most frequently observed (63%).[197] The herringbone chain (X) and the homochiral back-to-back motif (II) and interlocked motif (IV) are other homochiral motifs. Both back-to-back (II, VI and VII) and interlocked (III and IV) motifs can either be homo- or heterochiral.

1.3.2 Columnar Packing in Chiral Materials

In the past decades during which the field of discotic liquid crystals (DLC) emerged, recurring packing assemblies have been categorised as phases of a material (nematic, hexagonal columnar, rectangular columnar).[234] One special type of columnar packing arrangement of a chiral discotic molecule is a *chiral column*.[235] The π-π conjugation between the aromatic cores of molecules arranged in a column allow 1-D charge transfer along the columnar axis, which enables DLC potential applications as nanowire in molecular electronic devices.[235] Discotic (meaning disk-like) liquid crystals of helicene molecules have been synthesised.[234,236] Due to a strong tendency of a chiral molecule to aggregate in a monodisperse fibre, instead of a chiral column, only a few examples of a columnar phase have been reported.[236–238]

Helical molecular materials can pack in *chiral columns* of the same handedness, *i.e.* each col-

umn in the crystal is homochiral.[237,239,240] There are either racemic or enantiopure crystals with *chiral columns*. An enantiopure crystal with chiral columnar packing is a crystalline columnar material with same chirality (homochiral). A racemic crystal consists of homochiral columns of both *P*- and *M*-helicenes adjacent to each other. Each column in the crystal carries one chirality (*M* or *P*) and thus all the dipole moments either point up or down, and the dipole moment in each column follows a one-way street from top to bottom (Figure 1.6).[240] If the crystal was enantiopure, one would not have any antiparallel dipole moments; they would be all parallel and the material would show a strong chiroptical response, *i.e.* strong circularly polarised fluorescence.[237,238,240]

Enantiopure and racemic crystals containing chiral helicene columns both pack in long "well-aligned fibres".[237] Time-of-flight measurements of racemic and enantiopure helicene crystals with chiral columnar packing revealed potential usage as *p*-type and *n*-type organic semiconductors respectively.[239] In the future, crystalline helicene materials with columnar packings could be used in bipolar junction transistors or morphology-controlled bulk heterojunction solar cells.[239] Whilst some solid-state forms of helicene molecular materials have a columnar packing, the driving forces determining whether certain helicenes adapt a columnar packing and others do not, are not yet inherently understood.[240] Nakano *et al.* suggested by testing helicenes with different phosphole oxide or a phosphole sulfide moieities, that if the helicene molecule has an internal dipole moment vector parallel to the helical axis, columnar packing is enhanced.[240]

1.4 Thesis Outline

In this thesis I will describe my research into the computation of structure-property relationships to relate the crystalline arrangement of chiral molecular helicene materials to the predicted charge-carrier mobilities. The prediction of structure-property relationships involves state-of-the-art crystal structure prediction methods[58,241] and the use of previously developed methods[242] to describe charge transport in periodic molecular materials. The eventual aim is to design new molecules with crystalline arrangements engineered for maximal charge transport. This would require rapid computational screening of candidate structures.

Chapter 3 will present a detailed analysis of the polymorphic landscape of two 2,2'-disubstituted [6]helicene molecules, the packing motifs accessed and the associated charge

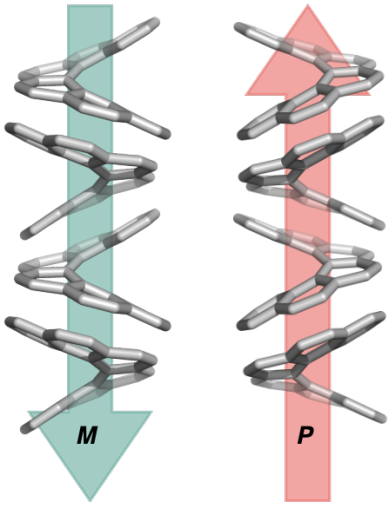


Figure 1.6: Illustration of two types of helical *chiral columns* of [6]helicenes (*M*- and *P*-[6]helicene). If both types of columns are present in a crystal it is *racemic*; if only one of the column types is present the crystal is *enantiopure*. The arrows are orthogonal to the plane of the helicene molecules.

transport properties. Two helicene molecules were predicted using CSP and their experimentally observed crystal structures were identified. Both molecules were synthesised by our experimental collaborators. For the set of lowest-lying crystal structures, the crystalline arrangement is analysed at DFT-XDM level of theory. A Python package is being presented to automatically analyse the intermolecular interactions within the crystal. Despite a crystal packing similarity analysis, the charge-carrier mobilities of the lowest-lying crystal structures is being discussed.

Chapter 4 will discuss the computational screening of ~1,300 of [6]helicene molecules for their charge transport parameters via the development of a dimer-based screening model. The choice of dimer is preceded by a CSD analysis and previous in-depth [6]helicene packing analyses to justify the choice of the dimer motif. A full computational workflow that (i) terminally substitutes molecules in all positions, (ii) performs a one-dimensional separation scan and (iii) calculates the electronic properties is developed. The charge transport screening parameters are combined in two distinct screening criteria to select the most promising candidates for an in-depth CSP study. For the top 4 candidates, a full CSP search is presented and the charge-carrier mobilities of the periodic crystals were calculated. This allows for comparison with the charge transport parameters predicted by the computational screening. To the best of my knowledge, this is the first computational screening study of helicenes.

Expanding upon the previous studies on [6]helicenes, Chapter 5 will discuss the CSP of helicene molecules of varying axial chain length ranging from [3]-[12]helicene. Motivated by the previously performed crystal packing analysis on [6]helicenes (Chapter 3), Chapter 5 presents the development of a molecule-based percentage score to assess the degree of π-π and C-H$\cdots\pi$ stacking for each crystal in the computationally predicted crystal-energy landscapes. Finally, a conclusion and an outlook of future work is provided in Chapter 6.

The main aims and objectives within the scope of this PhD thesis can be summarised as follows:

- Accelerate the discovery of new chiral molecular materials by determining which packing motifs are likely to dominate for a given molecule and identifying motifs most favourable for charge transport (Chapter 3).
- Provide an avenue for grouping crystal structures by packing motifs, crystal packing similarity (Chapter 3) and dominant crystal stacking interactions (Chapter 5)
- Develop a rapid computational screening strategy for side-group functionalised helicenes across various substitution positions and patterns to improve charge transport properties (Chapter 4).
- Quantify the degree of π-π and C-H$\cdots\pi$ stacking across the polymorphic landscapes of helicenes of different axial chain length (Chapter 5).
- Understand driving forces of double polymorphism in crystal-energy landscapes of chiral materials (Chapter 5).

Chapter 2

Theoretical Background

2.1 Force Fields

Force fields (FF) are a classical approach to deriving a potential energy surface for a given molecular system as a function of the system's nuclear coordinates.[243] In force fields, atoms are modelled according to classical mechanics and electrons are not treated explicitly. Instead, a force field considers a molecule being composed of recurring structural units, *e.g.*, C-H bonds, C-C bonds. The force field then provides atom- and bonding-specific parameters obtained from empirically fitting the force field to higher-level computational methods or experimental data.[243] With this representation, force field calculations enable the prediction of stable molecular conformers or the prediction of energy barriers within the single molecule and as part of molecular dynamics simulations of a material.[243] The force field energy (E_{FF}) is a sum of all intramolecular and intermolecular energy terms required:

$$E_{\mathrm{FF}} = E_{\mathrm{str}} + E_{\mathrm{bend}} + E_{\mathrm{tors}} + E_{\mathrm{el}} + E_{\mathrm{vdW}} \tag{2.1}$$

where the first three terms, E_{str}, E_{bend}, E_{tors} represent the respective stretching, bending and torsional energy terms. These contributions to the force field energy are all intramolecular terms. The electrostatic energy E_{el} and van der Waals interactions energy E_{vdW} terms describe intermolecular interactions. In some force fields, there is also a cross term, E_{cross}, that represents interdependen-

cies of the different intramolecular terms, *e.g.*, it can help describe bond elongation (E_{str}) upon angular compression (E_{bend}).[243] In a FF geometry optimisation, the nuclear coordinates for which the force field energy E_{FF} is minimised (Equation 2.1) are being searched for to obtain a stable molecule. The stretching energy between two atoms A and B can be described by a harmonic oscillators, *i.e.* the energy quadratically increases with the deviation of the A-B bond distance (R^{AB}) from its equilibrium bond length R_0^{AB}:[243]

$$E_{str}(R^{AB} - R_0^{AB}) = \sum_{bonds} \frac{k^{AB}}{2}(R^{AB} - R_0^{AB})^2 \tag{2.2}$$

where R^{AB} is the actual A-B distance, and k^{AB} is the force constant of the A-B bond, *i.e.* a measure for the softness/stiffness of the chemical bond depending on the interacting atom types.[243] Similar to the stretching energy, the bending energy E_{bend} is traditionally described by a harmonic approximation. Thus, E_{bend} describes the energy difference associated with the distortion of an angle A-B-C from its equilibrium bond angle (θ_0^{ABC}).[243]

$$E_{bend}(\theta^{ABC} - \theta_0^{ABC}) = \sum_{angles} \frac{k^{ABC}}{2}(\theta^{ABC} - \theta_0^{ABC})^2 \tag{2.3}$$

θ^{ABC} and θ_0^{ABC} represent the actual and the equilibrium bond angle respectively. k^{ABC} is the harmonic force constant of the bond angle.[243] The third term to the force field energy, the torsional energy, is the energy associated with a change in torsional (or dihedral) angle A-B-C-D. The torsional angle is quantified as the angle between the A-B and C-D bond, whilst looking down the B-C bond. Thus, the torsional angle can vary from 0° to 360°. The periodicity of the torsional energy can be described as:[243]

$$E_{tors}(\omega^{ABCD}) = \sum_{dihedrals} \frac{V_n}{2}[1 + cos(n\omega^{ABCD} - \gamma)] \tag{2.4}$$

where ω is the dihedral angle, V_n is a constant describing the barrier height to rotation, *i.e.* how easy/hard it is to change the dihedral angle and n relates to the number of minima within the energy function (multiplicity); n=1 signifies that the rotation is periodic by 360°. Thus, n=2 would signify that the rotation was periodic by 180° and so on. γ is the phase factor and describes the

position of the minimum across the energy function. In contrast to E_{str} and E_{bend}, the torsional energy E_{tors} tends to be low upon change of angle ω and cannot be described by a harmonic approximation. Furthermore, in contrast to the previous two energy terms, which were purely intramolecular, the torsional energy is also impacted by intermolecular interactions (*i.e.* van der Waals, electrostatic interactions).[243] The electrostatic energy E_{el} term describes the intermolecular interaction energy arising from an unequal distribution of the electrons, which in turn arises from differences in the electronegativities (ΔEN) of the interacting atoms. Thus, the simplest way to model these interactions is the use of partial charges assigned to each atom.[243] The interaction between two charges q_A, q_B is described by the Coulomb interaction potential:

$$E_{\text{el}}(R^{\text{AB}}) = \sum_{\substack{\text{non-bonded} \\ \text{pairs A,B}}} \frac{q_A q_B}{4\pi\epsilon_0 R^{AB}} \tag{2.5}$$

where R^{AB} is the distance between the non-bonded atoms A and B and ϵ_0 is the electrical permittivity in vacuum. The charges q_A, q_B are commonly assigned based on a quantum-mechanical molecular electrostatic potential (MEP) calculation.[243,244] A Coulomb potential is easy to calculate, but it does not provide a wholistic representation of the electrostatic energy E_{el}, because in many cases atomistic point charges are not sufficient to describe electrostatic interactions, *e.g.* the important quadrupole moment of molecular oxygen cannot be described by atom-centered point charges.[244] Furthermore, the fixed atom-centered charge distribution is a crude description of the electrostatic potential assuming all charges sit isotropically at the position of the nuclei. Instead the error in this electrostatic description could be reduced if partial charges at non-nuclear sites and higher order electric moments, *i.e.* dipoles, quadrupoles, were considered. One method to describe intermolecular electrostatic interactions beyond atom-centered point charges has been developed by Price *et al.*, who developed DMACRYS,[58] a distributed multipole electrostatic model for accurate determination of the relative lattice energies across hundreds of crystal structures, which will be discussed in more detail in Section 2.6.2.

The van der Waals dispersion energy represents the nonpolar non-bonded interaction energy arising from interacting groups on the same or different molecules. The van der Waals dispersion energy is zero at large distances and highly repulsive at short distances, because the electrons repel

each other at short distances.[243] However, at intermediate distances the van der Waals interactions are slightly attractive, because even a completely nonpolar molecule (no permanent dipole) will have an uneven distribution of electron density. This uneven electron density distribution or electron cloud on one molecule will polarise the charge density distribution on an adjacent molecule (or within the same molecule), thereby creating an attractive interaction and introducing a temporary dipole to the other molecule (part of the molecule).[243] This effect can be described by electron correlation. The attractive force at intermediate distances arising from van der Waals interactions is proportional to the inverse sixth power of the distance (R^{-6}, Equation 2.6).[243] The C^{AB} term is the van der Waals interaction constant for the two-atom system A and B.

$$E_{vdW}(R^{AB}) = \sum_{\substack{\text{non-bonded} \\ \text{pairs A,B}}} E_{repulsion}(R^{AB}) - \frac{C^{AB}}{(R^{AB})^6} \tag{2.6}$$

It is not possible to theoretically derive a function to describe the repulsive part of the van der Waals interaction, however many interaction potentials meet the requirements, *e.g.*, the Lennard-Jones potential is most famously used to describe van der Waals interactions (Equation 2.7).[245]

$$E_{LJ}(R^{AB}) = \sum_{\substack{\text{non-bonded} \\ \text{pairs A,B}}} 4\epsilon \left[\left(\frac{R_0^{AB}}{R^{AB}} \right)^{12} - \left(\frac{R_0^{AB}}{R^{AB}} \right)^{6} \right] \tag{2.7}$$

where ϵ denotes the depth of the potential well, R_0^{AB} the equilibrium distance of atom A and B.

2.1.1 The Williams W99 Intermolecular Force Field

Williams W99 potential parameters will be used throughout this thesis. The Williams W99[246] force field was specifically designed to model the intermolecular interactions of hydrocarbons in crystal structures. The W99 potential parameters were fitted to best match experimental crystal structures as well as heats of sublimation, and they were initially fitted to 134 organic molecules with 5-16 carbon atoms.[246] W99 models intermolecular electrostatic interactions by point charges on the nuclei plus off-nuclear partial charges to represent the anisotropy of the electron density arising from the surrounding atoms inside the molecule.[247]

The Williams W99 intermolecular force field is of the type exp-6-1 type.[246] The W99 model assumes that the crystal energy can be computed as the sum of pairwise atom-atom interactions of spherical atoms of the reference molecule with all surrounding molecules.[246] Thus, the crystal energy can be represented by

$$E = \frac{1}{2}\sum_{j}\sum_{k} E_{jk}, \tag{2.8}$$

where j runs over all atoms of the reference molecule, k runs over all atoms of surrounding molecules and the $\frac{1}{2}$ prevents double counting of atom-atom interactions.[246] The individual intermolecular atom-atom interaction can be described by three terms

$$E_{jk} = B\exp(-Cr_{jk}) - Ar_{jk}^{-6} + q_j q_k r_{jk}^{-1}, \tag{2.9}$$

where the first term is the exchange repulsion energy, the second term refers to the dispersion attraction energy and the third term represents Coulomb's law for the interaction between net atomic charges.[246] The dispersion interaction term has higher order terms (*e.g.*, r^{-8}, r^{-10}), which are not necessarily negligible.[246] However, since A is empirically fitted to the crystal structures, the force field parameters are designed to include an adjustment to A to account for these higher order contributions to the dispersion attraction term. The Coulomb term to the intermolecular atom-atom interaction energy is particularly important when describing molecular crystals of polar molecules.[246] Consequently, the derived A, B, C parameters are characteristic for certain element or subclass of an element. For example, there is a single set of A, B, C parameters to represent hydrogens, but carbon atoms connected to four C(4) or three C(3) other atoms were allocated different force field parameters.[246]

2.2 Density Functional Theory

The eventual aim of most quantum chemical methods is to yield an approximate solution of the time-independent non-relativistic Schrödinger equation:

$$\hat{H}\psi_i = E_i\psi_i \tag{2.10}$$

where $\hat{H}$ is the Hamiltonian differential operator of the system describing the energy of the N electrons and M nuclei in the system.[248] ψ_i is the wavefunction in the state i containing all information about the system, including the $3N$ spatial coordinates, N electron spin coordinates and 3M spatial coordinates of the nuclei.[248] Finally, the Hamiltonian operator, if acted on the wavefunction in the i'th state, yields the total energy E_i of the state described by ψ_i (Equation 2.10).[248] According to the Born-Oppenheimer approximation, the nuclei can be treated as stationary and noninteracting. Consequently, the kinetic energy of the nuclei part of $\hat{H}$ can be approximated to zero and the nuclei-nuclei repulsion part of $\hat{H}$ can be considered constant. Thus, $\hat{H}$ reduces to an electronic Hamiltonian $\hat{H}_{\text{elec}}$ with nuclear coordinates only parametrically and not explicitly included in ψ_{elec}. Hence, the result of $\hat{H}_{\text{elec}}$ (hereafter $\hat{H}$) on the electronic wavefunction, ψ_{elec}, is the electronic energy E_{elec}. $\hat{H}$ needs to be set up with all its system-specific parts in order to solve the Schrödinger equation for a given system.[248] The only system-dependent parts of $\hat{H}$ are: the number of electrons N and the external potential V_{ext}, which varies depending on the positions and charges of all nuclei present in the system.[248] With the adequate Hamiltonian set up, the eigenfunctions ψ_i and corresponding eigenvalues E_{i} of the system must be determined. Once they have been obtained, one can apply any other property operator to compute any physical property for the given system.[248] However, no strategy for exactly solving the Schrödinger equation and obtaining the exact solutions for a multi-electron system is known. As a consequence the variational principle is used to yield an approximate numerical solution of the ground state wavefunction ψ_0 corresponding to the lowest energy E_0.

According to the variational principle, the expectation value E_{trial} of the Hamiltonian for a normalised ψ_{trial} wavefunction (ψ^2= 1), which satisfies the boundary conditions, will correspond to the upper bound limit of the true ground state energy (Equation 2.11).[249] The upper bound limit implies that the lower in energy the approximate wavefunction, the more accurate the solution.[249] Hence the variational principle consists of minimising the functional $E[\psi]$ by testing all acceptable N-electron wavefunctions.[248] The acceptable (*i.e.* satisfying certain physical requirements) N-electron wavefunction which minimises the expectation value is the best possible estimate of

the exact ground state energy.[249]

$$\langle \psi_{\text{trial}} | \hat{H} | \psi_{\text{trial}} \rangle = E_{\text{trial}} \geq E_0 = \langle \psi_0 | \hat{H} | \psi_0 \rangle \tag{2.11}$$

Equation 2.11 taken from Ref. 248 is the variational principle. Equation 2.11 contains the main mathematical concept of density functional theory, because a 'rule' which assigns E_{trial} to a function ψ_{trial} is called a *functional*. A functional takes a function as its argument, in contrast to a function, which takes a variable as its argument. In density functional theory, the density functional takes electron density (a function) as its input and yields the electronic energy as its output.

Kohn-Sham density functional theory (DFT) is a method for approximately and indirectly solving the Schrödinger equation by describing the kinetic energy of a noninteracting reference system with the same density ρ as the real system as accurately as possible.[248] For a given external potential, $v_{\text{ext}}(r)$, the full ground-state DFT electronic energy functional is given by[250]

$$E_{v(r)}[\rho(r)] \equiv \int v_{\text{ext}}(r)\rho(r)dr + T_S[\rho(r)] + \frac{1}{2}\int \frac{\rho(r)\rho(r')}{|r-r'|} dr dr + E_{XC}[\rho(r)], \tag{2.12}$$

where the first term describes the external potential, the second the kinetic energy of a noninteracting system with density $\rho(r)$, the third term is the classical interaction energy expression and the final E_{XC} term refers to the exchange-correlation energy.[250] Since the kinetic energy of a noninteracting system is different from the interacting system despite the similar density, the additional exchange-correlation energy term to the density functional equation was introduced to account for the difference between the true functional and the functional of the electron density $F[\rho]$ derived from the noninteracting system.[248]If the exact exchange-correlation energy E_{XC} and the potential due to the exchange-correlation energy V_{XC} were known, the Kohn-Sham approach to density functional theory could provide the exact solution to the Schrödinger equation. All unknown parts of the exact true functional are collected in the exchange-correlation functional $E_{\text{XC}}[\rho]$.[248] Thus, the accuracy of the chosen density functional theory approach depends on the approximations chosen to represent the E_{XC} and V_{XC} terms. The exchange-correlation energy E_{XC} is expressed by a density functional approximation, such as the local density approximation (LDA), the generalised gradient approximation (GGA) and hybrid GGAs. GGA marked a significant im-

provement to the method compared to LDA (which assumes a uniform electron gas), because it describes exchange-correlation energy as a function of the electron density at a particular position $\rho(r)$ *and* the gradient of the density $\nabla\rho(\vec{r})$ at that position (Equation 2.13).[250] Hence, if the density gradient is zero, all GGA density functional approximations revert to the LDA.[251]

$$E_{\mathrm{XC}}^{\mathrm{GGA}} = E_{\mathrm{XC}}\left[\rho(r), \nabla\rho(r)\right] \tag{2.13}$$

The most widely-used parameter-free GGA functional is PBE, developed by Perdew, Burke and Ernzerhof.[252]

The B86bPBE GGA exchange-correlation functional paired with the exchange-hole dipole moment (XDM) dispersion model has been proven to accurately represent intermolecular interactions,[253,254] and perform better than the PBE functional paired with XDM.[255] The good performance of the B86bPBE functional is ascribed to the better behaviour of the B86b exchange enhancement factor in the large reduced-density gradient limit.[254–256] In the majority of this work, the B86bPBE[252,256] functional will be used to relax crystal structures.

All transfer integral calculations in this thesis use hybrid density functional theory. Hybrid density functionals are based on GGA functionals, but they are supplemented by a small fraction of the exact Hartree-Fock (HF) exchange energy, E_X^{exact}, in the exchange energy term E_{XC}. The exchange-correlation energy in B3LYP:[251,257]

$$E_{\mathrm{XC}} = E_{\mathrm{XC}}^{\mathrm{LDA}} + a(E_X^{\mathrm{exact}} - E_X^{\mathrm{LDA}}) + b\Delta E_X^{\mathrm{B88}} + c\Delta E_C^{\mathrm{PW91}}, \tag{2.14}$$

with $a = 0.20$, $b = 0.72$, $c = 0.81$; a is the fraction of exact HF exchange energy (20%) present in the B3LYP functional.[251]

2.3 Dispersion Corrected Electronic Structure Methods

Common density functionals (DFs) do not describe long-range dispersion interactions correctly.[258] Semilocal and conventional hybrid DFs (*e.g.,* BLYP, B3LYP) are purely repulsive and do not provide the correct $-C_6$ / R^6 dependence of the dispersion interaction energy on the intermolecular distance, R. However, van der Waals dispersion interactions in periodic crystalline

materials contribute significantly to the stability of the molecular crystal. Dispersion is a special type of electron correlation dominating at long-range scales.[258] Dispersion interactions are inherently nonlocal (and often anisotropic) and experimentally proven nonadditive interactions, which cannot be described by widely used semilocal density functionals.[259,260] Thus, whilst standard DFs are able to describe dispersion interactions at short length scales, *i.e.* short electron-electron distances due to its dependence on the electron density changes. The functionals fail at describing long-range dispersion interactions.[258]

If the exact solution to the Schrödinger equation was known, the exact contribution of the dispersion interactions to the total energy of the many-electron system at all length scales would be described in the electron energy term.[260] However, to be able to account for the short- and long-range asymptotic regions of the dispersion interaction energy, which are fairly well understood separately, one derived dispersion corrections to the DFs to provide an extension to the traditional Kohn-Sham density functional theory.[258,259] Dispersion corrections are either directly included self-consistently within the density of the functional (*e.g.* vdW-DF1, vdW-DF2[261]) or as a *post hoc* dispersion correction (DFT-D),[262]

$$E_{\mathrm{DFT-D}} = E_{\mathrm{KS-DFT}} + E_{\mathrm{disp}} \tag{2.15}$$

where $E_{\mathrm{KS-DFT}}$ is the usual self-consistent Kohn-Sham energy and E_{disp} is the dispersion correction. The dispersion correction is usually an atom pairwise sum over $-C_6$ / R^6 potentials.[258] State-of-the-art post hoc dispersion corrections include Grimme's (D3) dispersion correction,[62] Tkatchenko-Scheffler (TS) correction,[66] the many-body dispersion model (MBD)[263] or the exchange-hole dipole moment (XDM)[64] dispersion model. Dispersion corrected density functional theory (DFT-D) allows for significantly more accurate descriptions of large infinite systems such as molecular crystals, graphene sheets or for the description of surface interactions. Thus, for the purpose of this thesis, dispersion corrections should be used, where possible, to correctly model the intra- and intermolecular noncovalent interactions within a periodic crystal.[264] The dispersion correction employed throughout this work is the semiclassical density-dependent exchange-hole dipole moment (XDM) dispersion model. The approximation is not based on the long-range Coulomb correction.[265] Instead, Becke and Johnson noticed that if the exchange-hole

around an electron is not spherical and symmetric around the electron, the dispersion energy can be derived from the perturbation due to an electron and its exchange hole and the associated multipole moments.[265] The pairwise dispersion energy can be described as:

$$E_{\text{disp}}^{\text{XDM}} = -\sum_{\text{AB}} \sum_{n=6,8,10} \frac{C_n^{\text{AB}}}{R^n} f_{\text{damp}}^{(n)}(R) \tag{2.16}$$

where C_n^{AB} refers to the dispersion coefficients derived from the high-order multipole moments and $f_{\text{damp}}^{(n)}(R)$ refers to the Becke-Johnson damping function (Equation 2.17), which is also used in a variant of the D3 dispersion correction[265]

$$f_{\text{damp,BJ}}^{(n)}(R) = \frac{R^n}{R^n + (a_1 R_0 + a_2)^n}, \tag{2.17}$$

where $R_0 = \frac{1}{3}\left(\sqrt{C_8^{\text{AB}}/C_6^{AB}} + \sqrt{C_{10}^{\text{AB}}/C_8^{\text{AB}}} + \sqrt[4]{C_{10}^{\text{AB}}/C_6^{\text{AB}}}\right)$. Here, the dispersion energy is damped to a finite value at short distances, R.[265] The XDM model is routinely used in solid-state calculations, *e.g.* to accurately describe adsorption energy calculations of organic molecules[266] and for the analysis of noncovalent interactions in periodic molecular crystals.[267] Furthermore, the composite and low-cost DFT-XDM approaches for relaxing molecular crystals have been benchmarked for aza[6]helicene.[268,269]

2.4 GFN-xTB

The semiempirical Geometry, Frequency, Noncovalent (GFN) extended tight-binding (xTB) approach is designed for the fast calculation of structures and noncovalent interaction energies.[270] Typical density functional tight-binding (DFTB) methods are based on the (semi)local density approximation (LDA). Thus, the local electron correlation functional neglects long-range electron correlation effects responsible for dispersion interactions.[270] Hence, conventional DFTB approaches typically require the use of a post hoc dispersion correction to the DFTB energy. GFN-xTB on the other hand assumes a general density functional approximation, which directly includes a nonlocal (NL) VV10-type[271] correlation contribution.[270] Thus, the total Kohn-Sham en-

ergy can be described by

$$E[\rho] = \int \rho(r)\Big[T[\rho(r)] + V_{\text{ext}}(r) + \epsilon_{\text{XC}}^{\text{LDA}}[\rho(r)] + \frac{1}{2}\int\Big(\frac{1}{|r-r'|} + \phi_C^{\text{NL}}(r,r')\Big)\rho(r')dr'\Big]dr + E_{\text{nm}}, \quad (2.18)$$

where $T[\rho(r)]$ is the kinetic energy per particle, V_{ext} is the external potential arising from the nuclei, the remaining terms within the integral arise from the electronic contribution to the semilocal mean-field potential, where $\epsilon_{\text{XC}}^{\text{LDA}}[\rho(r)]$ is the (semi)local exchange correlation (E_{XC}), $\frac{1}{2}\int(\frac{1}{|r-r'|}$ is the Coulomb interaction term and the nonlocal electron correlation energy is given by $\phi_C^{\text{NL}}(r,r')$ and captures long-range electron correlation effects.[270] E_{mn} is the classical nuclear repulsion energy.[270] The GFN-xTB total energy, as initially described in the first release, is the sum of the electronic (E_{el}), atom pairwise repulsion (E_{rep}), dispersion (E_{disp}) and halogen-bonding interactions (E_{XB})[272]

$$E_{\text{GFN-xTB}} = E_{\text{el}} + E_{\text{rep}} + E_{\text{disp}} + E_{\text{XB}}. \quad (2.19)$$

The conventional GFN-xTB approach describes electrostatic (ES) and exchange-correlation (XC) contributions using isotropic approximations. The latest version of GFN-xTB, GFN2-xTB,[270] additionally incorporates anisotropic second order density fluctuation effects to the ES and XC terms, *i.e.* anisotropic electrostatic (AES) and anisotropic XC (AXC) energy terms as

$$E_{\text{GFN2-xTB}} = E_{\text{rep}} + E_{\text{disp}} + E_{\text{EHT}} + E_{\text{IES+IXC}} + E_{\text{AES}} + E_{\text{AXC}} + G_{\text{Fermi}}, \quad (2.20)$$

where $E_{\text{IES+IXC}}$ refers to the isotropic electrostatic (IES) and the isotropic XC (IXC) energy contributions. E_{EHT} refers to the extended Hückel-type energy, which describes covalent bonds.[270] The dispersion energy (E_{disp}) term is calculated using the D4[273] dispersion correction model. G_{Fermi} refers to the entropic contribution of electronic free energy at finite temperature (300 K) and is required to describe fractionally occupied orbitals.[270] The previous E_{XB} is no longer required, because this version of GFN-xTB does not require any classical halogen (XB) or hydrogen bonding corrections and is thus more physically sound than its precursor.[270] Chapter 4 of this thesis uses GFN-xTB[272] version 6.0.1, which includes the new GFN2-xTB parameters.[270] GFN2-xTB performs well on geometry optimisations of molecules, vibrational frequency calculations and noncovalent interaction calculations. However, the GFN2-xTB works less well on metallic systems and polar

inorganic clusters due to occasional SCF convergence problems. This deficiency is likely to arise from the short-range anisotropic electrostatic (AES) potential contribution to the total energy.[270]

2.5 DLPNO-CCSDT

As outlined in Section 2.2 on DFT, despite its many computational advantages, the performance of density functional theory can vary drastically depending on the system under investigation and the chosen approximation to represent the exchange-correlation energy term. This shortcoming of DFT motivates the use of highly accurate post Hartree-Fock wave function based methods, which rigorously achieve chemical accuracy and they do not require any calibration to experimental measurements.[274] The coupled cluster approximation with singles and doubles (CCSD) was first proposed by Pople and co-workers in 1985.[275] Coupled cluster with inclusion of doubles and triples excitations, CCSD(T), is the globally most widely-used wave function method, the "gold standard" of computational chemistry.[274,275] Due to the high computational cost of CCSD(T), only extremely small systems with less than 1000 basis functions could be described. Only when more approximations were derived, one of them being the domain based local pair-natural orbital (DLPNO) based singles and doubles coupled cluster method (DLPNO-CCSD) with triples correction (T), the computation of medium-sized molecular systems was rendered computationally feasible, whilst maintaining a high degree of accuracy in describing relative and absolute energies of test sets.[275] DLPNO signifies that instead of using canonical delocalised orbitals to build the wave function equation, localised pairs of natural orbitals are used, which allows sorting of them into domains and eases the selection of the most important excitations relevant to electron correlation.[274,276]

2.6 Crystal Structure Prediction

As discussed in Section 1.1.3, there is a vast pool of competing CSP methodologies to best determine the lowest-lying and experimentally observed polymorphs. This section will discuss the deterministic CRYSTALPREDICTOR/DMACRYS approach, to generate the set of lowest-energy tentative crystal structures, which will be applied throughout this thesis. The generalised workflow is shown in Figure 2.1.

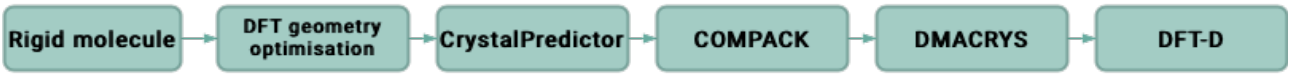


Figure 2.1: CSP workflow for a rigid molecule (from left to right): 1. single-molecule DFT geometry optimisation in GAUSSIAN16,[278] 2. CRYSTALPREDICTOR search for candidate structures, 3. COMPACK to remove duplicate structures, 4. DMACRYS relaxation and re-ranking of candidate structures using distributed multipoles, 5. DFT-D: post-processing of the set of lowest-energy structures using dispersion corrected DFT methods

2.6.1 CRYSTALPREDICTOR

The CRYSTALPREDICTOR[241,277] algorithm can be broken down into the following steps:

1. The candidate structures are generated based on the 59 most frequently observed space groups according to the CSD (Global Search Stage). Extremely rare space groups are not explicitly considered.[241] Once a space group for the candidate molecule is selected, the cell parameters (a, b, c, α, β, γ) are optimised, by obeying the space group specific constraints. The multivariable domain search is performed by the use of quasi-random Sobol sequences to more uniformly cover all configurational space than possible with a random sequence. Once a tentative crystal structure is successfully suggested, the structure is checked against certain thresholds *e.g.* density or lattice energy constraints.[241] If the tentative structure meets all requirements, a minimisation of the tentative structure is attempted (next step), otherwise this step (step 1) repeats.[241]

2. The minimisation of candidate structures with space group restrictions is attempted using a global objective function to determine all crystal variables. The Sobol sequences assure that crystal structures that are rejected at this stage, based on the limits defined in lattice energy minimisation problem, will not lead to low-energy crystal structures. Similarly, the Sobol sequences guarantee a wide enough sampling space, such that enough candidate structures will be sampled close enough to converge to a low-lying minimum energy structure upon optimisation.[241] Furthermore, an upper limit of the volume per molecule in the unit cell is introduced to limit the configurational space to be sampled, because the lowest-lying, empirically observed crystal structures tend to be the closest packed crystal structures (Kitaigorodsky's principle).

3. This step serves the purpose of identifying duplicate candidate structures, *i.e.* structures which are reasonably similar to previously generated structures in the search.[241] To identify duplicate structures, for each candidate structure, the density and energy as well as the shortest non-bonded pairwise distances between all pairs of atom types are assessed to determine, whether the structures are highly similar and should be assigned the same cluster. Lattice energy, density and non-bonded distance thresholds are applied to determine whether structures are considered to be similar (cutoffs are 0.05 kJ mol^{-1}, 0.50 kg m^3 and 0.1 Å respectively).[241]

4. Finally, the set of unique candidate structures is minimised again without any space group restrictions because some crystal structures might have only appeared as local minima due to the space group constraints put in place while they were actually saddle points. Thus, for each local minimum (relaxed without any space group restrictions) the Hessian matrix of the lattice energy equation is computed.[241] Only if all eigenvalues of the Hessian matrix of the lattice energy are positive values, the candidate structure is assured to lie in a true minimum (and not a saddle point). If the candidate structure is not a true minimum, it is used as an initial guess to another re-optimisation, which often returns a true local minimum structure, which is part of a lower symmetry space group than the one that served as the initial guess.

Once step 4 is complete, the true local minimum-energy structures are checked again for duplicates using the COMPACK[74] executable. COMPACK is a CCDC tool which builds a neighbouring shell of typically 15 molecules for each crystal structure, then each neighbouring shell is mapped onto all others. If all molecules were matched onto one another and the RMSD is reasonably low (RMSD $\approx$ 0). The structures are considered duplicates.

The complexity of crystal structure prediction (CSP) drastically increases with the number of molecules in the asymmetric unit (Z') and the number of flexible degrees of freedom within the molecule giving rise to the crystal. Concerning the conformational flexibility of the isolated molecule, the CRYSTALPREDICTORII algorithm was developed to perform CSP on flexible molecules, which requires accounting for intramolecular flexible degrees of freedom, whilst optimising intermolecular interactions. Whether a molecule is rigid or flexible is determined by the number of rotable bonds present in the molecule, *i.e.* the molecule is described as consisting

of a set of rigid fragments, which are only distinguishable by the torsional angles of their flexible degrees of freedom.[277] Some rotations may be less significant than others, for example, a rotating $-CH_3$ group tends to not give rise to different conformers, because the rotation of the methyl group does not generate multiple stable conformers of the same molecule. Thus, dihedral angle scans of the isolated molecule are performed at DFT level to identify multiple stable conformers, *e.g.* a dialdehyde side group can give rise to different conformers, depending on whether the oxygens on the carbonyl groups point towards each other or in opposite directions. This example will be discussed in detail in Chapter 4 of this thesis. Besides this single flexible example, this thesis will focus on the CSP of rigid molecules.

2.6.2 DMACRYS

In the most recent sixth blind test, Price *et al.* demonstrated the accuracy of the CRYSTALPREDICTOR/DMACRYS workflow by successfully predicting a large proportion of the target candidates.[47] The CRYSTALPREDICTOR workflow for generating candidate crystal structures has been presented in Section 2.6. The distributed multipoles analysis (DMA) of crystal structures software package, DMACRYS, is used on the tentative crystal structures (generated via CRYSTALPREDICTOR) to more accurately estimate relative lattice energy differences, ΔU_{latt}, across the polymorphic landscape, which tend to be only a few kJ mol^{-1}.[58] DMACRYS uses a distributed multipole representation of the molecule's charge density to yield a more accurate representation of the anisotropic atom-atom electrostatic potential parameters than possible with conventional isotropic atom-atom force field modelling using partial charges. Hence, the multipole moments on each atom (atomic multipoles) are extracted from a wavefunction calculation in GAUSSIAN16 by integration via the GDMA2[279] programme. Furthermore, DMACRYS includes a symmetry reduction functionality, which removes symmetry operations from a tentative crystal structure. Thus, a crystal structure which might be constrained by the space-group symmetry may transition to a significantly lower-energy polymorph upon cell relaxation at reduced symmetry.[58] Initially, the single molecule energy and charges are computed using GAUSSIAN16.[278] The single molecule charges are then extracted to derive the multipoles, polarisabilities and dispersion coefficients and the multipoles are in turn used to parametrise anisotropic atom-atom intermolecular potential pa-

rameters to describe the electrostatic contribution to U_{latt}.[58] The non-electrostatic contributions to U_{latt} are described using empirically fitted and isotropic force field parameters, such as W99[246] or FIT[73,280] potential parameters.[58]

2.7 Machine Learning in the Chemical Sciences

Conventionally, computers have been used in the chemical sciences to efficiently solve hard-coded problems similar to a powerful calculator. An example of such a powerful calculator in the chemical sciences was the development of the GAUSSIAN software package, which allows chemists with limited quantum-chemical knowledge to perform *ab initio* quantum-chemical calculations, *i.e.* solve the Schrodinger equation and predict the energy, behaviour, *etc.* of molecules of modest size purely from the laws of physics.[81] With the advent of scientific high-performance computing, large sets of chemical data could be computed using such conventional chemistry software packages. The large sets of chemical data obtained as well as the increase in high-performance computing power (*e.g.* GPUs) as well as the development of easy-to-use software packages (*e.g.* sklearn,[281] tensorflow[282]) enabled the use of ML in the chemical sciences to learn outcomes based on large sets of existing chemical and physical data.[81] In machine learning (ML), a computer program is said to learn from experience E with respect to some task T and performance measure P, if its performance P at executing task T increases with experience E (Tom M. Mitchell, 1997). To provide an example, the task of the computer could be to play checkers, the experience, the number of games of checkers already played and the performance, the probability of winning against the human opponent. Thus, in contrast to conventional computing, ML approaches learn the rules that underlie a dataset by assessing a proportion of the data and deriving a model to rationalise the data.[81] The performance of the model can then be assessed by how well it is able to predict the outcomes of the remaining data points, not used in the "fitting" of the model. The training of a ML model can be *supervised*, *semisupervised* and *unsupervised*, depending on the amount and availability of the data.[81] In simple terms, supervised learning is performed on labelled data, unsupervised on unlabelled data and semisupervised on data which is partially labelled. Here, the label means the target variable one tries to predict, which can either be a discrete (class label) or a continuous variable.

2.7.1 Supervised Learning

Supervised learning is a type of ML where a model is built based on a set of features x to predict a target variable y. The aim is to build a model which describes the maximal proportion (%) of variance of the data (x_{train}) and which at the same time remains generalisable to previously unseen examples (x_{test}), having learned important regularities of the input. If the target variable y is a continuous value (a discrete label/category), the supervised learning problem is a regression (classification) problem.[283]

In this thesis, kernel ridge regression is employed to build a model, which can predict the inner-sphere reorganisation energy λ_{i} from a molecular fingerprint. Kernel ridge regression can be used together with a variety of kernels, *e.g.*, the Exp-Sine-Squared kernel, which models periodic functions,

$$k(x,x') = \exp(-2(sin(\pi/p * d(x,x'))/l)^2), \tag{2.21}$$

where l describes the length-scale parameter, p the periodicity and $d(x,x')$ the Eucledian distance.[284]

Given enough chemical data and a discovery algorithm, machine learning uses the computing power available to determine all known and unknown physical rules without human input and then use the model derived to predict new properties for unseen examples.[81] In traditional computing, a computer is employed similar to a calculator, solving a hard-coded problem. Machine learning however enables the computer to flexibly adapt an algorithm to solve unseen problems.

2.7.1 Clustering

Clustering is a method of grouping a set of features x by their similarity into groups of increased similarity, so-called clusters. This strategy is extremely helpful for the subsequent downstream analysis of large datasets and identifying subsets in large pieces of data. However, clustering falls into the range of unsupervised learning techniques, thus there is no ground truth or target variable against which the performance of the clustering algorithm can be assessed. In order to be able to compare performances of clustering algorithms, quality functions must be defined to assess the "goodness of a partition", such as for graph clustering methods: consensus clustering or modular-

ity. Consensus clustering[285,286] in complex networks is finding a median partition amongst multiple partitions, *e.g.*, multiple runs of the clustering algorithm, which is in stark contrast to greedy global quality functions such as modularity.[287] The Louvain clustering algorithm is an example of a greedy community-detection algorithm with the modularity of the graphs as the optimisation function.[288]

2.7.2 Graph Representations

Graphs can be used to study complex network systems, such as social networks, interacting proteins, or to understand the similarity across crystal structures. A graph is a set of elements, *i.e.* a set of nodes connected by edges. A way of representing a graph is by means of an adjacency matrix, A_{ij}, a $n \times n$ square matrix, where n is the total number of nodes present in the graph. The adjacency matrix elements indicate whether two pairs of nodes (vertices) are adjacent or not in the graph. To give an example, methane (CH_4) could be converted into a simple *binary graph* with a 5 × 5 adjacency matrix. Now, every matrix element in the $A_{C,H}$ or $A_{H,C}$ position would contain a 1 to indicate the presence a C-H bond, whilst every $A_{H,H}$-position of the matrix would contain a 0, to indicate the absence of a H-H bond.

2.7.2 Molecular Fingerprints

Graph representations of molecules are widely employed in the Python chemical toolkit RDKit.[289] RDKit uses kernels to encode the molecule's connectivity and hash its features into a binary bit vector of finite length. Thus, each part of the bit vector can be attributed to a fragment of the molecule. The exact encoding strategy varies depending on the fingerprint chosen (*e.g.*, Morgan fingerprint); equally there are many algorithms to assess the similarity of the molecular graphs, *e.g.*, substructure searching, Tanimoto similarity.[290,291]

2.7.3 Kernels

Kernels are functions which compute the pairwise similarity between two set of features (*e.g.*, between two molecular fingerprints) by computing the dot (inner) product in high-dimensional fea-

ture space[292]

$$k(x, x') = \langle \phi(x), \phi(x') \rangle, \tag{2.22}$$

where x and x' are two input examples; ϕ is the function mapping each input into high-dimensional feature space, k is the kernel. Equation 2.22 is known as the *kernel trick*, because the computationally expensive transformation of explicitly mapping the input features x into the feature space $\phi(x)$ is avoided by instead directly computing the dot (inner) product between two input features in the high-dimensional feature space.[293,294] If the aim is to be able to predict the target variable y from an input x, the kernel is a similarity measure identifying whether the new input x' is highly similar to another example x and is thus likely to have the same target variable ($y' = y$). Thus, kernels are used to find higher-dimensional decision boundaries for nonlinear problems, which cannot be found within the initial dimensionality of the data. A kernel needs to be symmetric ($K(x, x') = K(x', x)$) and positive (semi)definite.[292] Common kernels include Gaussian (RBF) kernels, polynomial kernels or Laplacian kernels.

When computing kernel-based similarity matrix across hundreds of tentative crystal structures, the similarity matrix is a square matrix $n \times n$ where n is the total number of crystal structures analysed. Thus, the similarity matrix of hundreds of crystal structures can be very memory intensive. Memory efficient popular graph kernels include the Shortest Path (SP), Weisfeiler-Lehman (WL),[295,296] Vertex Histogram (VH), and the Pyramid Match (PM) kernel[297–299]. The Pyramid Match (PM) kernel[297–299] is widely used in computer vision for object recognition tasks. The PM kernel measures the similarity of two sets of features via the intersection of their respective multi-resolution histograms. It is computationally efficient and thus suitable for large unordered feature sets.[300] More specifically, the PM kernel maps sets of features, *i.e.* nodes of the graphs, "to multi-resolution histograms, and then compares the histograms with a weighted histogram intersection measure to approximate the similarity of the best partial matching between the feature sets."[300] Thus, features not found to 'match' at a high resolution level can be 'matched' at lower resolutions.[300] New matches are weighted according to the resolution level (size of that histogram's bin) at which they were made, *i.e.* matches made at larger bins are weighted less than matches made at smaller bin sizes.[300] Each vertex becomes a point in the d-dimensional unit of the hypercube. Thus, the two hyper-parameters of the PM kernel are d, the dimension of the hypercube,

and L, the number of levels of the pyramid histogram (default: L=4, d=6). Higher dimension hypercubes and more histogram levels were also tested: L=14, d=8; or L=10 and d=12. However, PM kernel similarity matrices computed with higher hyper-parameters (L=14, d=8) require too much memory (>96 GB) to be computationally feasible for use in the computation of a similarity matrix between ~600 crystal structures.

2.8 Charge-Carrier Mobility Calculations Using Semi-classical Marcus Theory

This section outlines the methodology used to predict charge-carrier mobility of periodic molecular materials. The charge transport methodology was previously developed as part of previous PhD theses in the group (References. 242, 301). Charge transfer is the transfer of an electron from molecule A to molecule B

$$A^- + B^0 \rightarrow A^0 + B^-, \tag{2.23}$$

where the superscript denotes the charge of the molecule; – and 0 for negative and neutral respectively. The ideal method for investigating charge transport (CT) in organic semiconductors is not straightforward and strongly depends on the system under investigation.[128] In molecular materials studied in this thesis, the transfer integrals J under investigation arise from through space coupling of the helicene molecules and no through-bond coupling is observed. For small-molecule organic semiconductors such as helicenes, the transfer integral tends to be significantly lower than the reorganisation energy; thus $J \ll \lambda$. Therefore, the hopping regime tends to be the most appropriate choice.

To compute the crystal charge-carrier mobilities, non-adiabatic semi-classical Marcus theory is used. This a popular and widely used theory, which was first developed by Marcus in 1964.[302] For the calculation of the transfer integrals, a 3×3×3 supercell with n×27 molecules is generated, where n is the number of molecules in the unit cell. The supercell is used to compute transfer integrals J across all possible molecular pairs within a 15 Å centre-of-mass to centre-of-mass cutoff.

2.8.1 Transfer Integrals

There are many ways of computing transfer integrals,[303–306] the transfer integrals discussed in this work are based on the transfer integral definition as found in the Marcus Equation 2.25. The transfer integrals, J, for each molecular pair are computed by projecting the computed orbitals of the dimer onto the unperturbed localised orbitals of the individual molecules (projective method).[197,242,301,307] The transfer integral is either a measure of the LUMO-LUMO (electron) or HOMO-HOMO (hole) coupling between two adjacent interacting molecules

$$J_{\mathrm{ij}} = \langle i|\hat{\mathbf{F}}|j\rangle, \tag{2.24}$$

where $\hat{\mathbf{F}}$ denotes the Fock matrix of the dimer system, while $|i\rangle$ and $|j\rangle$ stand for the orbitals localised on the molecules i and j, respectively. All transfer integrals were computed at B3LYP/6-31G(d) level in GAUSSIAN16.[278] The chosen level of theory has proven successful in previous bench-marking examples, *e.g.*, for [6]helicenes[197], aza-[6]helicene[308], and other organic semiconductors.[307]

2.8.2 Marcus Rate

Based on the transfer integrals, the charge hopping rate from molecule i to j is computed using the Marcus equation[309]

$$\Gamma_{\mathrm{ij}} = \frac{|J_{\mathrm{ij}}|^2}{\hbar}\sqrt{\frac{\pi}{\lambda k_{\mathrm{b}} T}}\exp\left(-\frac{(\Delta E_{\mathrm{ij}}+\lambda)^2}{4\lambda k_b T}\right), \tag{2.25}$$

where J_{ij} is the electron or hole transfer integral (see Section 2.8.1), ΔE_{ij} is the change in energy between molecules i and j, λ is the reorganisation energy, and unless stated otherwise, a temperature of 300 K was assumed.

The reorganisation energy, λ, is the sum of the energy cost associated with the rearrangement of the molecules involved in the hop (λ_{i}) and their surroundings (λ_{o}). The λ_{i} is calculated with the four-point method (Section 2.8.3).[310] The outer-sphere reorganisation energy, λ_{o}, cannot be calculated explicitly and is always (unless stated otherwise) estimated to λ_{o} = 0.3 eV. Since the

outer-sphere reorganisation energy should not vary strongly depending on morphology,[311] I estimate this value to be the same across all morphologies. Since λ is essentially a scaling parameter to the overall mobility, by employing the same value for λ_o, the obtained mobilities are uncertain to the same extent across all systems.

The methodology has been routinely used on helicenes in previous works.[197,312] The change in energy ΔE_{ij} between the sites is

$$\Delta E_{ij} = \epsilon_i - \epsilon_j - q_e F \cdot r_{ij}, \tag{2.26}$$

where q_e is the elementary charge of an electron, F the applied electric field vector and r_{ij} the distance vector between the two sites i and j. The energies of the frontier orbitals for molecule i and j are given by ϵ_i and ϵ_j. Since we ignore energetic disorder, $\epsilon_i = \epsilon_j$ and thus, ΔE_{ij} simplifies to

$$\Delta E_{ij} = -qF \cdot r_{ij}. \tag{2.27}$$

The angular dependent mobility was computed, unless stated otherwise, by applying a field strength F of 10^4 V cm^{-1} at 10° intervals in each of the three planes of the crystal (*ab*, *bc*, *ac*) by solving the Master equation

$$\frac{dP_i}{dt} = \sum_{i \neq j} [\Gamma_{ij} P_i(t) - \Gamma_{ij} P_j(t)], \tag{2.28}$$

where Γ_{ij} is the hopping rate between sites i and j, P_i and P_j are the occupation probabilities of the respective sites i and j being occupied. The Master equation is solved in the steady state, $AP = 0$, where A is defined as

$$A_{ij} = \begin{cases} \Gamma_{ij}, i \neq j \\ -\sum_{i \neq j} \Gamma_{ij}, i = j. \end{cases} \tag{2.29}$$

Singular value decomposition (SVD) was used to find the null space of the A_{ij} matrix. The null space corresponds to the steady state probabilities P. Therefore, the average velocity of charges is found according to

$$\vec{v} = \sum_{i \neq j} \vec{r}_{ij} A_{ij} P_{ij}. \tag{2.30}$$

Finally, the velocities, $\vec{v}$, are used to compute the charge-carrier mobility,

$$\mu = \frac{\langle v \rangle}{F}, \tag{2.31}$$

where $\langle v \rangle$ is the average velocity of the charges in the direction of the field and F the applied external electric field.

2.8.3 Inner-sphere Reorganisation Energy

The reorganisation energy, λ, is a key parameter in the hopping transport process determining the the rate of the activated hopping process.[313] The inner-sphere reorganisation energy, λ_i, quantifies the energy associated with the change in molecular geometry of a molecule when moving between the charged and neutral state.[313] More precisely, for hole transporting materials λ_i reflects the geometric relaxation energy associated with accepting the hole, *i.e.* transitioning from the neutral geometry to the positively charged (+) geometry and the adjacent molecule, which accepts the hole, transitioning in reverse fashion.[314] Thus, there are two distinct inner-sphere reorganisation energies, which can be computed either for the hole and the neutral geometry or the electron and neutral geometry. Da Silva Filho *et al.* computed the contribution of vibrational modes of rubrene to λ_i and find that λ_i is expected to be larger in rubrene than for other common organic semi-conducting molecular crystals, such as pentacene and tetracene, due to the presence of low-frequency motions of the phenyl groups.[314]

λ_i was computed for each individual molecule using four-point-method[310]

$$\lambda_\mathrm{i} = E_\mathrm{n}^{-/+} - E_\mathrm{c}^{-/+} + E_\mathrm{n}^{0} - E_\mathrm{c}^{0}. \tag{2.32}$$

Here, single-point energies were computed for the charged (-/+) or neutral states (0) at either the charged or neutral optimised geometries (c or n, as denoted by the subscripts). The sum of the two differences gives rise to λ_i (Eq. 2.32).

Thus, the inner-sphere electronic reorganisation energy, λ_i^{-}, is a quantitative measure of the ability of a single molecule to accommodate a negative charge, and the energetic penalty for conversion between the negatively charged to the neutral geometries (Fig. 2.2). When designing

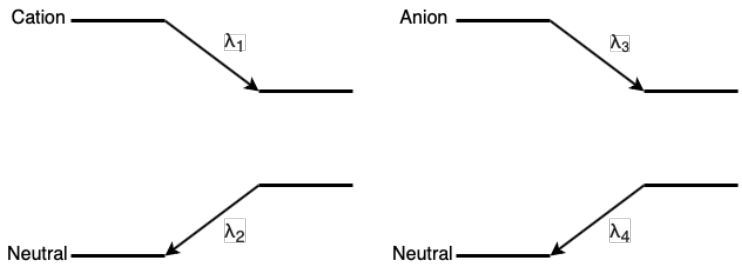


Figure 2.2: Energy levels of a molecule involved in vertical transitions. λ_i^+ and λ_i^- are given by $(\lambda_1 + \lambda_2)$ and $(\lambda_3 + \lambda_4)$, respectively. The Figure is adapted from Ref. 315.

small-molecule OSC (with minimal λ_i) one has to keep in mind that the electronic inner-sphere reorganisation energy, λ_i^-, is largely determined by the moiety that contributes to the LUMO.[315] Conversely, λ_i^+ is mostly dominated by the moiety that determines the HOMO.[315]

2.8.4 Mobility Anisotropy

If a material is isotropic, its properties are the same in all directions. However, single crystalline materials are predominantly anisotropic, *i.e.* their properties are direction dependent. For example, single crystalline materials can exhibit elastic anisotropy,[316] anisotropy of superconductivity[317] or anisotropy of charge-carrier mobility.[318,319] The field-effect angular-dependent mobility gives rise to anisotropy, meaning the measured and calculated mobility is orientation dependent and not isotropic. Charge-carrier mobility anisotropy has been experimentally measured for organic single crystals[319] and thin-films of TIPS-Pn and TES-Pn via control of the orientation of the crystal domains.[320] The crystal anisotropy will be an important indicator to compare the relative suitability of packing motifs for charge transport. Measures to compute the mobility anisotropy will be provided alongside the results where they appear in Chapter 3.

Chapter 3

Structure-property Relationships of 2,2'-Diethynyl- and 2,2'-Dicyano-[6]helicenes: Do Chiral Columns Promote Charge Transport?

3.1 Introduction

Computational studies of the crystal-energy landscape of [6]helicene and 1-aza[6]helicene showed that variations in crystal packing strongly effect the electronic properties of [6]helicenes.[197,308] In-depth analysis of the solid forms obtained from crystal structure prediction (CSP) revealed that the crystal packings required for optimal hole or electron transport differ significantly and there is no one packing motif responsible for consistently high charge transport.[197] Besides the variation of electronic properties with packing motifs, how the chirality of the material affects charge transport is not yet inherently understood. Across the polymorphic landscape of [6]helicene, the highest electron-transporting crystal was the experimentally observed and enantiopure crystal structure (Cambridge Structural Database (CSD) reference code: HEXHEL).[197] In contrast, the highest hole-transporting crystal structure was a racemic crystal structure.[197] Similarly, in the polymorphic landscape of 1-aza[6]helicene, the racemic 1-aza[6]helicene polymorph was experimentally found to lead to an 80-fold increase in hole carrier-mobility compared to the enantiopure polymorph.[308] These findings and previous research in the field of polymorphism in organic semicon-

ductors (see Section 1.2.2) demonstrate the strong dependence of high organic semiconductor (OSC) performance on the adequate spatial arrangement of molecules in the solid-state. In light of this, the impact of molecular structure on charge transport was investigated independently of crystal packing effects.[312] For aza[6]helicenes, the position of the nitrogen atom was changed, assuming the same crystal packing in each case, and large fluctuations in the predicted charge mobilities were observed.[312]

Based on the work by Salerno *et al.*, which focused on how nitrogen-substitution of the aromatic [6]helicene backbone affects charge transport in helicene materials, the next natural step was to determine how terminal functionalisation influences charge transport. Whilst for the investigation of the nitrogen-position in the [6]helicene backbone the approximation of similar crystalline packings mainly held true, terminal functionalisation of the [6]helicene backbone alters the shape and size of the molecule. These new molecular geometries can access novel packing motifs with altered intermolecular interaction strengths.[321,322]

A search of the Cambridge Structural Database (CSD) for terminally substituted [6]helicenes was performed to assess the complexity of terminally substituted [6]helicenes already synthesised and to identify potentially interesting crystal packings for charge transport. The CSD search resulted in a total of 330 helicene type crystal structures (including any helicene type structure, *e.g.,* helicenes of any length, helicene substructures in a larger coronene type molecule, helicenes as ligands in inorganic complexes, or helicene co-crystals). Following the CSD search and inspection of pure [6]helicene molecular crystals, [a] three terminally 2,2'-disubstituted [6]helicenes with exceptional crystal packings were identified: UJOCOL (2003),[323] NEMKUO (2017),[324] and UJOCUR (2003)[323] (Figure 3.1). All 2,2'-disubstituted [6]helicenes have crystal structures with molecules oriented in "chiral columns". In a "chiral column", helicenes are stacked on top of each other along their helical axis and each helicene is rotated along its helical axis by some degree with respect to the helicene situated above and below until a full 360° rotation and the pattern repeats (see Section 1.3.2). None of the [6]helicene polymorphs was reported to contain chiral columnar packing; instead the adjachent homochiral translational chain or the herringbone chain were frequently observed packing motifs. The chiral columns observed across the three molecules differ: Chiral

[a] work carried out by Francesco Salerno, PhD student Imperial College 2018

columns in 2,2'-dicyano- and 2,2'-diethynyl-[6]helicenes are enantiopure (chiral) crystals with a chiral space group, whereas the 2,2'-dimethyl ester-[6]helicene crystal structure is a racemic crystal (achiral, $Pna2_1$).[323] Furthermore, the ester functionalised 2,2'-dimethyl ester-[6]helicene is a Z'=2 crystal, whilst the other two are Z'=1.

The chiral columnar packing observed for 2,2'-dicyano-[6]helicene is particularly interesting, because the material organises in an extremely rare space group ($P6_122$) with an appearance frequency lower than 0.05%, which is why this space group was not included in the initial release of CrystalPredictor.[241] The $P6_122$ space group consists of 12 symmetry operations. Thus, the unit cell is relatively large, as it contains 12 molecules. Conventionally, the chiral column is arranged along the *c*-axis of the unit cell, resulting in an exceptionally long cell length (70.7 Å). Conversely, the chiral column formed by 2,2'-diethynyl-[6]helicene contains only three helicene molecules and is found in the much more typical space group $P3_1$. Based on the CSD search, 2,2'-diethynyl-[6]helicene was found to exist as a second, non columnar polymorph (Figure 3.1).[324] The non-columnar packing resembles the [6]helicene packing (CSD reference code: HEXHEL) as it shares the same adjacent translational chain. Comparing these packing motifs to those seen in the previous work on [6]helicene[197] and 1-aza[6]helicene,[308] for which there is no observed chiral columnar packing, the question arose as to whether chiral columnar packing is more suitable for charge transport than previously investigated solid-state arrangements. Chiral columns have a high degree of π-π stacking along the chiral axis, which makes them an interesting study case for maximising charge transport along a crystal packing motif.[239] Having studied how the position of the nitrogen atom in aza[6]helicenes affects charge transport given the same crystalline packing,[312] it was a natural step forward to now investigate how terminally substituted [6]helicenes compare in terms of electronic properties *and* the set of thermodynamically accessible crystalline packings. Thus, this computational study of the two 2,2'-disubstituted [6]helicenes aims to address how i) molecular structure and ii) crystal packing affect charge transport. The two molecules with linear triple bond containing side-groups, 2,2'-dicyano- and 2,2'-diethynyl-[6]helicene, were chosen for further investigation, because terminal functionalisation involving the addition of triple bonds is expected to i) enhance electronic properties by increasing the π-system size of the [6]helicene backbone,[325] and ii) to direct molecular assembly in organic semiconductors (*e.g.*, 6,13-

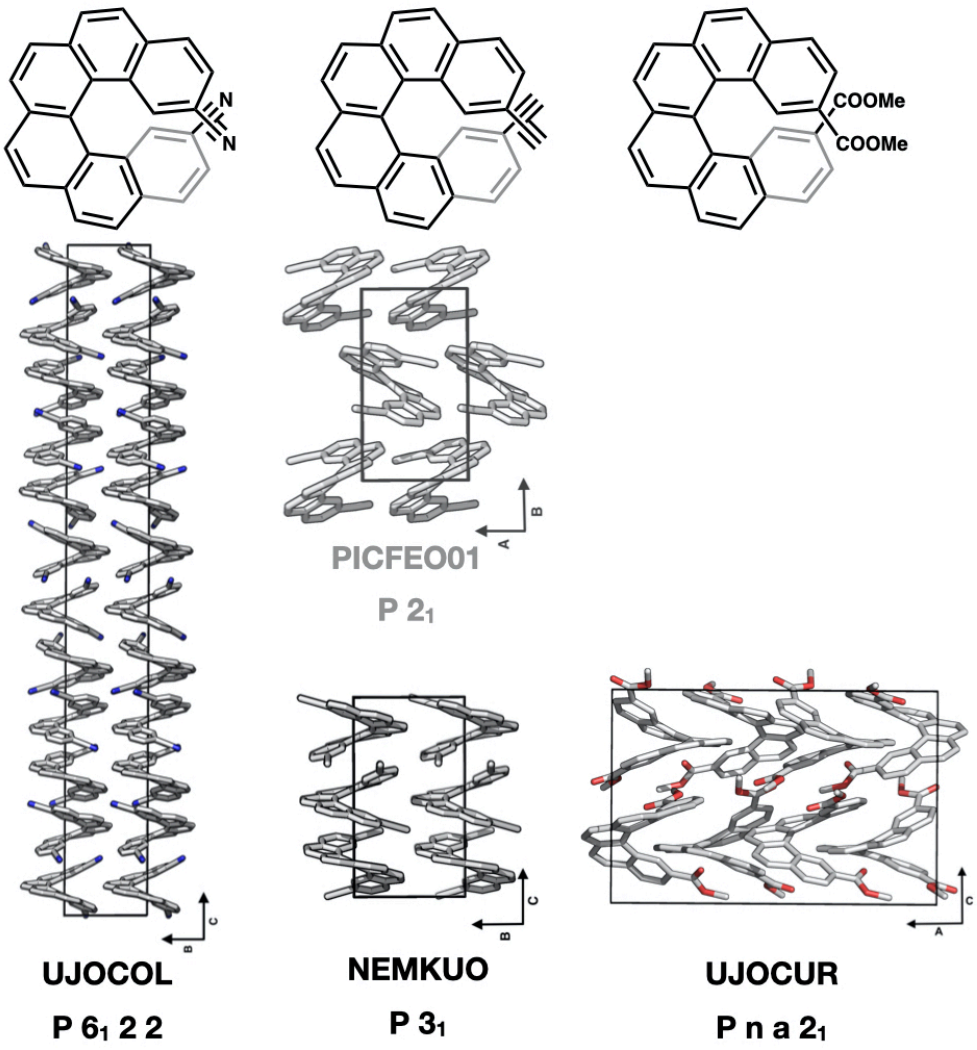


Figure 3.1: 2,2'-disubstituted-[6]helicene crystal structures with columnar packing along the *c*-axis.[323,324] Left: 2,2'-dicyano-[6]helicene (UJOCOL), middle: 2,2'-diethynyl-[6]helicene (NEMKUO, PICFEO01), right: 2,2'-diemethyl ester-[6]helicene (UJOCUR). 2,2'-diethynyl-[6]helicene has two experimental polymorphs deposited to the CSD; one of which packs in a chiral column (NEMKUO), the other does not (PICFEO01).[324]

bis(triisopropylsilylethynyl) (TIPS) pentacene[326], dinaphtho-fused *s*-indacenes[327]).

The 2,2'-dimethyl ester derivative was not selected for further investigation, because of the flexibility of the side group. This added flexibility would likely require the use of multiple adaptive local approximate models (LAMs) to consider many different conformers as part of a computationally costly flexible CSP search in order to sample the full configuration space. Furthermore, the increased molecular flexibility renders crystal-lattice vibrations likely to be increasingly important and limiting charge-carrier mobility.[328] Moreover, the 2,2'-dimethyl ester-[6]helicene experimental structure was found with Z'=2, which means that there are two chemically identical molecules that are not related by symmetry and arrange in distinct positions.[329] In order to predict this crystal structure, computationally expensive higher dimensionality CSP searches with more than a single molecule as the repeating building block of the crystal would be required.

The computational results presented in the following have been performed as part of a systematic interdisciplinary research study in collaboration with the Fuchter group (Imperial College London) and the Johnson group (Dalhousie University). Francesco Salerno (Fuchter group) synthesised the selected helicenes with the goal of obtaining multiple polymorphs, including chiral columns, whose chiroptical response and electronic properties could be measured. The experimentally obtained polymorphs could then be compared to the crystal-energy landscape obtained via CSP. Similarly, mobility measurements of multiple polymorphs and along different axes of the crystal would be comparable to computational predictions. High-level periodic DFT-XDM relaxations of the lowest-energy crystals have been performed by Dr Luc LeBlanc and Joseph Weatherby in the Johnson group. To the best of my knowledge, no systematic interdisciplinary study of chiral columnar packing in helicene molecular crystals has been previously reported.

3.1.1 Chapter Outline

The following chapter will be divided into two main sections: (i) the crystalline packings obtained via CSP, *i.e.* the crystal-energy landscape and (ii) the crystal-mobility landscape, *i.e.* the fluctuations in charge-carrier mobility depending on the solid-state arrangements. The first section (i) crystal-energy part can be further divided into two subsections: (a) a detailed discussion of the predicted crystal-energy landscape (including experimentally observed polymorphs, lattice en-

ergy rankings, *etc.*) and (b) a packing analysis of the obtained crystal structures. The packing analysis contains both a discussion of recurring structural motifs and a packing similarity analysis based on a similarity matrix.

Research questions include:

- Will the crystal packing motifs observed vary significantly from motifs previously classified in the [6]helicene crystal-energy landscape?
- Does symmetrical terminal disubstitution lead to the formation of more solid forms with chiral columns?
- Are chiral columns better charge hopping pathways than other packings?

3.2 Methods

The methods described in the following can be divided into three parts: (i) *crystal structure prediction*, methodology related to the computational prediction of crystal structures via CSP, (ii) *crystal packing analysis*, methodology related to the analysis of the crystalline packings of the crystal structures predicted and (iii) *charge-carrier mobility calculations*, methodology related to the prediction of charge-carrier mobility using semi-classical Marcus theory. The methods will be presented in the same order.

3.2.1 Crystal Structure Prediction

This section describes the process of generating hundreds of tentative crystal structures based on the intramolecular connectivity of a single molecule only. 2,2'-dicyano- and 2,2'-diethynyl-[6]helicene were both geometry optimised using GAUSSIAN 16[278] at B3LYP[330]/6-31G(d,p) level of theory, tight convergence criteria and assuming no symmetry. Based on the optimised geometry, the charges from electrostatic potentials using a grid-based method (ChelpG)[331] were computed to represent the atomic charges fitted to an electrostatic potential. The CRYSTALPREDICTOR II[55,332] software package was used to generate hypothetical crystal structures. The search was restricted to only one molecule in the asymmetric unit (Z'=1) and treating the [6]helicene molecules as rigid

molecules. For both molecules, the crystal-energy landscape was searched in a polymorphic region of 20 kJ mol^{-1} and using the pairwise van der Waals parameters. 500 000 minimisations were performed across the most common space groups: $P1$, $P\bar{1}$, $P2_1$, $P2_1/c$, $P2_12_12$, $P2_12_12_1$, $Pna2_1$, $Pca2_1$, $Pbca$, $Pbcn$, C_2/c, Cc, $C2$, Pc, Cm, $P2_1/m$, $C2/m$, $P2/c$, $C222_1$, $Pmn2_1$, $Cmc2_1$, $Aba2$, $Fdd2$, $Iba2$, $Pnna$, $Pccn$, $Pbcm$, $Pnnm$, $Pmmn$, $Pnma$, $Cmcm$, $Cmca$, $Fddd$, $Ibam$, $P4_1$, $P4_3$, $I\bar{4}$, $P4/n$, $P4_2/n$, $I4/m$, $I4_1/a$, $P4_12_12$, $P4_32_12$, $P\bar{4}2_1c$, $I\bar{4}2d$, $P3_1$, $P3_2$, $R3$, $P\bar{3}$, $R\bar{3}$, $P3_121$, $P3_221$, $R3c$, $R\bar{3}c$, $P6_1$, $P6_3$, $P6_3/m$, $P2_13$, $Pa\bar{3}$, $P222_1$, $Pba2$. Then all space groups are searched in proportion according to their relative abundance in the CSD.[241] The extremely rare space group of the experimental 2,2'-dicyano-[6]helicene ($P6_122$) was added to the CRYSTALPREDICTOR[55,332] package and searched for in an additional CSP search under the same conditions. Subsequently, all crystal structures were clustered to remove duplicates from the search using COMPACK.[74]

The tentative crystal structures obtained from the CRYSTALPREDICTOR stage were relaxed using distributed-multipoles and the repulsion-dispersion potential to the W99 parameters (DMACRYS).[58] The CSP results were re-ranked and relaxed using the exchange-hole dipole moment (XDM) dispersion model to yield a more accurate representation of the CSP landscape.[47,64] All unique structures within the crystal-energy landscape, which were lower in W99 lattice energy than the experimentally observed packing (including the experimentally observed structure, *i.e.* a set of 38 structures) were relaxed using periodic boundary DFT calculations in Quantum ESPRESSO[333] with the B86bPBE exchange-correlation functional[252,256] and the XDM dispersion correction.[253,254] The projector augmented-wave (PAW) method was employed.[334] Due to the high computational cost of these calculations, only the set of 50 lowest-energy 2,2'-diethynyl-[6]helicene crystal structures (at force field level) was periodically relaxed (unit-cell parameters and atomic positions) with CRYSTAL17.[335] The sHF-3c[336] method, a variant of the HF-3c[62] model except a modified s8 parameter of 0.6143, was used including the three body Axilrod-Teller-Muto term. A 4×4×4 Monkhorst-Pack[337] k-point grid was used to sample the Brillouin zone for all crystals. All other parameters such as convergence thresholds in the energy and geometry optimisations were used at the CRYSTAL17 defaults.

Please note that an enantiopure crystal structure crystallises in a Sohnke space group with a chiral crystal structure and they cannot form achiral crystal structures.[184] Racemic crystals tend to

pack in achiral crystal structures. It is extremely rare (~50 of 25,000 crystals in CSD), but a racemic molecular material can crystallise in a chiral crystal structure.[184] For instance, enantiomeric excess (ee) measurements of the [6]helicene molecular crystal revealed, that despite the Sohncke space group ($P2_12_12_1$), the material is racemic with an ee of a few percent.[338] In this edge case, the enantiopure crystal is hypothesised to be an intergrowth crystal, containing multiple enantiopure regions of opposite chirality, which overall balance out to a racemic crystal.[338] In the following, only computationally predicted Z'=1 crystal structures will be discussed and a chiral (achiral) crystal structure will be assumed to be an enantiopure (racemic) crystal structure.

3.2.2 Crystal Packing Analysis

This section describes two different approaches in analysing the various crystal packings (solid forms) obtained from CSP. Firstly, a python module developed here, CRYSTACK, is presented. The module pinpoints all intermolecular interactions between nearest neighbour molecules to the central molecule in the asymmetric unit. Based on the absence/presence of a set of intermolecular contacts, and the orientation of the molecules to each other, three crystal packing motifs can be identified. However, due to the limited functionality of a motif classification approach based on heuristic rules, in a second step, all crystals are being converted to graphs; graph similarity kernels are being computed, and clusters of higher similarity are detected.

3.2.2 CRYSTACK

The intermolecular interactions between each of the helicene molecules in the crystal were assessed by building a python module CRYSTACK presented here. CRYSTACK adapts intermolecular interactions functionality from PLIP,[339,340] a Python-based Protein-Ligand Interaction Profiler module, to identify interactions in molecular crystals instead of proteins and their ligands. The main feature of CRYSTACK, a crystal stacking analysis module, is to identify parallel π-π stacking interactions of aromatic rings and C-H$\cdot\cdot\pi$ stacking interactions of aromatic rings orthogonal to each other. The intermolecular interactions data extracted using CRYSTACK is used to classify molecular pairs of the crystal into recurring structural motifs. The three most commonly encountered recurring structural motifs for [6]helicene are illustrated in Figure 3.2.

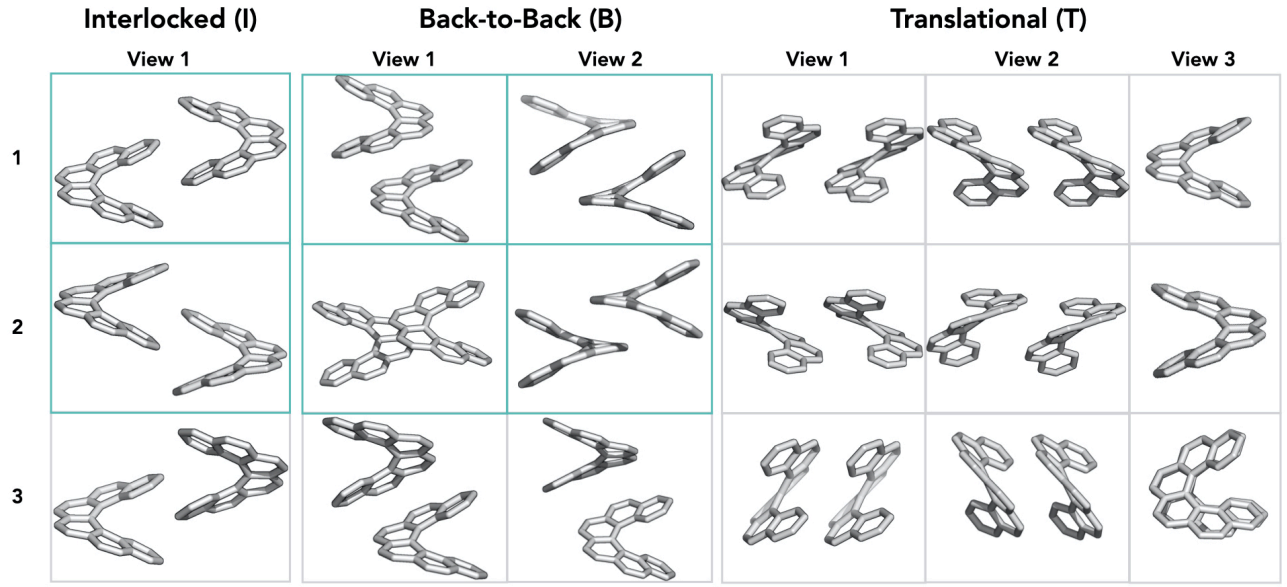


Figure 3.2: Illustration of the three motifs of interest, interlocked (I), back-to-back (B) and homochiral translational (T), previously identified for [6]helicenes.[197] Homochiral dimer motifs are shown in a grey box, heterochiral motifs are shown in turquoise. For each motif three examples are given to highlight its diversity. The intermolecular information obtained via CRYSTACK is being interpreted to identify whether these motifs are present in the crystalline structure.

The homochiral translational dimer motif (T) is, as the name suggests a purely homochiral (*M*- or *P*-handedness only) translational motif and characterised by one molecule being translated with respect to the other along a vector perpendicular to the plane of the helix. The T motif has good π-π overlap of the adjacent helicene backbones. If the separation is small enough, a characteristic parallel π-stack is observed across the aromatic rings near the centre-of-mass of the [6]helicene backbones. This is in stark contrast to the interlocked (I) and back-to-back (B) motifs, which are rotational-translational. All π-π stacking interactions within the T motif are parallel/slipped-parallel. There is no orthogonal CH$\cdot\cdot\pi$-stacking interaction. Furthermore, a symmetry condition has to be applied to not falsely classify other purely translational motifs, *e.g.* herringbone or side-translational as translational.

The [6]helicene backbones in the interlocked dimer (I) are characterised by being nearly antiparallel to each other, with the terminal rings pointing towards each other. The terminal ring of one [6]helicene backbone characteristically lays in between the terminal rings of the other [6]helicene backbone, sometimes allowing for a characteristic parallel π-π stacking interaction with one terminal ring and a C-H$\cdot\cdot\pi$-π interaction with the other terminal ring. There are many variations of interlocked motifs, sometimes with one aromatic ring being surrounded by the terminal aro-

matic rings of the other molecule in the dimer (similar to a tweezer, Figure 3.2 I). Thus, to make the classification comparable across landscapes, all dimers categorised as having the interlocked dimer either had at least 1 π-π stacking interaction. In case of the sterically more hindered 2,2'-dicyano-[6]helicene with its nitrogen atom, nitrogen-carbon short contacts were also accepted as a criterium.

The back-to-back motif (B) with its molecules antiparallel to each other (molecules pointing in opposite directions) show characteristic parallel π-π stacking interactions of the adjacent branches, varying with the chirality of the type of back-to-back dimer. Consequently, there are two major types of B motifs, the heterochiral and the homochiral which, due to the axial chirality, only the former type of B motif exhibits perfectly parallel π-π-stacking.

The stepwise procedure for detecting the three [6]helicene motifs is illustrated in Figure 3.3 (bottom path). Firstly, the W99 force-field optimised residual file is converted into a CIF file, then the periodic fractional coordinates are used to generate a 4×4×4 supercell using pymatgen.[341] Subsequently, the most central molecule in the unit cell is found (assuming Z'=1 only). Around the most central molecule, the centre-of-mass to centre-of-mass distance to all surrounding molecules is computed. The five unique dimers with the shortest centre-of-mass to centre-of-mass distance are extracted from the shell (ignoring duplicate dimers). For each dimer, the CRYSTACK module computes all molecule-molecule interactions. In addition to intermolecular interactions, such as orthogonal C-H$\cdots\pi$-π interactions between aromatic rings and (slipped)-parallel π-π stacking interactions between aromatic rings, the offset of the slipped-stacked aromatic rings is quantified and whether the interaction occurs between terminal or non-terminal rings is known. Furthermore, the relative orientation of the molecules with respect to each other (parallel, antiparallel), hydrophobic interactions C··C interactions between aromatic carbons and short contacts are measured. As much information regarding the solid-state arrangement within each dimer (*e.g.* relative orientation, *etc.*) as possible is extracted to determine whether the dimer forms part of any of the three commonly observed motifs. The heuristic rules for the three motifs were derived and fitted such that they would correctly classify 24 2,2'-diethynyl-[6]helicene and 27 2,2'-dicyano-[6]helicene polymorphs. The same set of heuristic rules was applied to both molecules.

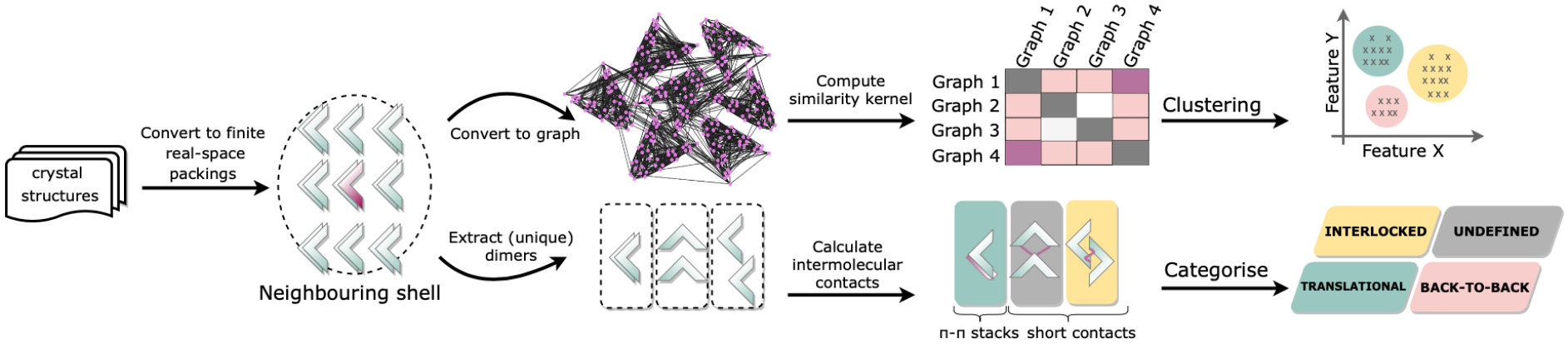


Figure 3.3: Computational workflow of crystal packing analysis. The top path illustrates the graph-kernel similarity comparison and the bottom path the discrete three-packing structural classification. Initially crystal structures in Crystallographic Information File (CIF) format are used to generate the neighbouring shell of molecules in cartesian coordinates. *Top:* The neighbouring shell is converted into a graph representation. Across all neighbouring shells as graphs the similarity amongst graphs is computed via a graph similarity kernel. The graph similarity kernel is then used for clustering to identify regions of higher similarity. *Bottom:* From the neighbouring shell a subset of unique molecule pairs (dimers) is being extracted. For each unique dimer the intermolecular contacts are being computed using CRYSTACK. Based on features associated to certain packing motifs (`I`, `B`, `T`) heuristic rules have been defined to detect whether any of these packing motifs is present.

3.2.2 Graph-based Similarity Map of Crystal Structures

In view of the limited adaptability of rigidly defined 'motifs' to analyse the solid-state arrangement in unseen crystals and for new molecules, an alternative approach was developed. The crystal structures were instead compared using graph kernels and the similarity was illustrated in a graph kernel similarity map, because kernels can effectively describe complex decision boundaries and generalise well to unseen data.[300,342] The approach is illustrated in Figure 3.3 (top path). Each crystal structure is represented by a neighbouring shell consisting of 15 molecules with 42 (44, 46) atoms per molecule. Then, each neighbouring shell is converted into an undirected graph with a 630×630 (660×660, 690×690) adjacency matrix A_{ij}. If the separation of two atoms (nodes) was lower than the sum of their van der Waals (vdW) radii plus 2 Å then the two nodes would be linked by an edge. The weight of the edge was simply the Euclidean distance between the two atoms (calculated given their cartesian coordinates).

Based on a feature importance analysis performed on a range of [6]helicene dimers via HCGA,[343] the *edge betweenness centrality* feature was identified as an adequate node label feature. The *edge betweenness centrality* is defined as the number of the shortest paths passing through an edge in a graph or network.[344] The shortest-path *edge betweenness centrality* of a node v is the sum of the fraction of all-pairs shortest paths that pass through v.[345,346] In simpler terms, the *edge betweenness centrality* is a measure for whether a node contains edges acting as bridge-like connectors between two parts of a network, and whether the removal of edges to this node would affect the communication between many pairs of nodes through the shortest paths between them.[345–347] Based on the crystal graphs, multiple graph kernels (SP, WL, PM kernel) were tested using GraKel,[348] a Python-based graph kernel library. Finally, the Pyramid Match (PM) kernel was selected, because it is memory efficient, lead to a diverse range of similarity scores and correctly identified duplicate crystal structures.

3.2.2 Clustering Based on Crystal Packing Similarity Kernels

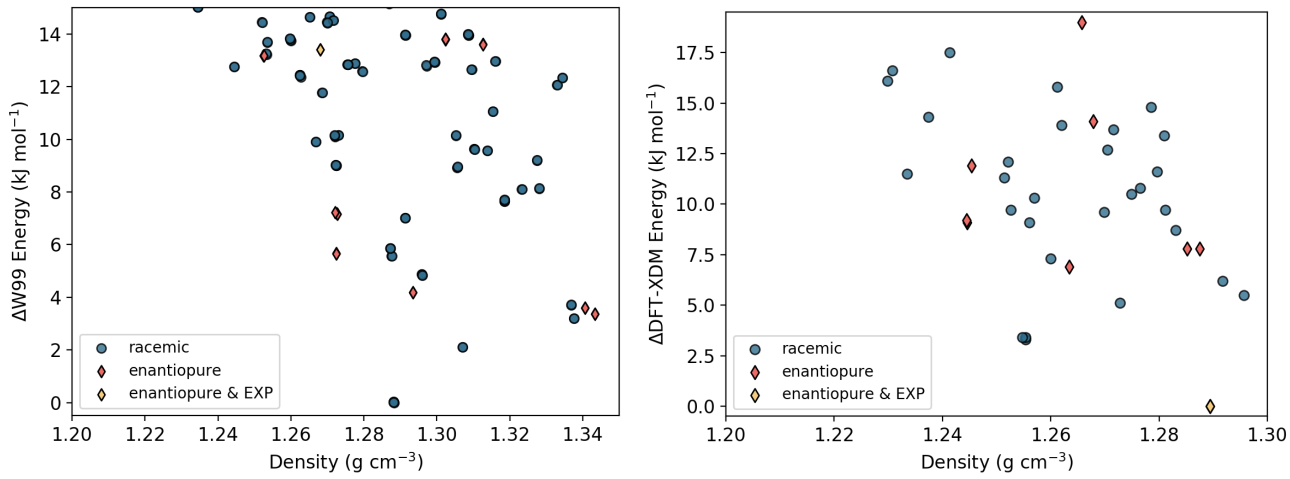
The similarity kernels were used for clustering of the 642 2,2'-dicyano- and 654 2,2'-diethynyl-[6]helicene polymorphs. Different clustering algorithms were tested, ranging from clustering al-

gorithms with pre-defined cluster size (*e.g.* Agglomerative, K-Means) to clustering techniques without a predefined number of clusters (Louvain[288], DBSCAN). The similarity scores obtained for duplicate structures were clipped to the highest non-duplicate similarity score to avoid the noise from exceptionally high scores, four times larger than the remaining similarity scores, would corrupt the clustering.[349,350]

Different clustering algorithms were applied to the 2,2'-dicyano- and the 2,2'-diethynyl-[6]helicene graph kernel similarity matrix, either directly (Louvain) or on the first 2, 3 or 6 principal components. Several features to assess the best number of clusters, such as minimum, maximum and average distances between clusters, Davies Bouldin score and Dunn index were tested. The best cluster separation for 2,2'-dicyano-[6]helicene was observed for agglomerative clustering using the *ward* linkage criterium and 6 clusters, if clustered based on the first two principal components. This resulted in the lowest Davies Bouldin score and the largest splitting in the dendrogram. The second best cluster separation for 2,2'-dicyano-[6]helicene was observed for agglomerative clustering using the *ward* linkage criterium and 4 clusters, if clustered based on the first three principal components. The Louvain clustering of the 2,2'-dicyano-[6]helicene resulted in two large clusters comprising the majority of the data and a large quantity of extremely small clusters. The best cluster separation for 2,2'-diethynyl-[6]helicene was observed for Louvain clustering on the PM kernel (for both (d=5, L=7) and (d=4 , L=6)). Three clusters were identified. This is consistent with the largest split of the data being found for three clusters according to the agglomerative clustering dendrogram based on the first 3 principal components and using the *ward* algorithm.

3.2.3 Crystal Charge-carrier Mobilities

The crystal charge-carrier mobilities were computed for both carrier types (hole and electron) for the set of lowest-lying crystal structures of the 2,2'-dicyano-, 2,2'-diethynyl- and [6]helicene crystal structures according to the exact details provided in Section 2.8. Despite one exception (1.29 $cm^2\,V^{-1}\,s^{-1}$), all charge-carrier mobilities are well below 1 $cm^2\,V^{-1}\,s^{-1}$ and thus clearly inside the hopping regime. Alongside the angular-dependence of mobility, the mobility anisotropy is provided. The mobility anisotropy is quantified using two types of anisotropy measures (A_1, A_2),



(a) (b)

Figure 3.4: 2,2'-dicyano-[6]helicene crystal-energy landscapes at the W99 (left) and the DFT-XDM (right) level of theory. Racemic (enantiopure) crystal structures are shown with blue dots (red diamonds). The experimentally observed polymorphs are highlighted in yellow.

discussed where they appear. Both anisotropy measures have been previously used for the crystal-mobility analysis of [6]helicene.[197]

3.3 Results and Discussion

First, the *crystal-energy* landscapes for both molecules, 2,2'-dicyano- and 2,2'-diethynyl-[6]helicene will be presented sequentially. Then, the *crystal packing analysis* results will be presented. Finally, the *crystal-mobility* results will be discussed.

3.3.1 Prediction of Experimentally Observed Polymorphs with CSP

3.3.1 2,2'-Dicyano-[6]helicene CSP

In the entire polymorphic landscape of 2,2'-dicyano-[6]helicene generated via a global search with a polymorphic search range of 20 kJ mol^{-1}, 9.7% of the structures predicted were enantiopure. Thirty-eight 2,2'-dicyano-[6]helicene crystal structures within 15 kJ mol^{-1} of the global minimum at force-field level were relaxed at the DFT-XDM level of theory (Figure 3.4b), nine of which were enantiopure (24%). Among the 9 enantiopure crystals, the experimentally observed $P6_122$ crystal structure (CSD reference: UJOCOL) and structure 13 (*i.e.* the 13$^{\text{th}}$ lowest-lying structure at the DFT-XDM level) are the only crystals with (homo)chiral columnar packing. Ten racemic crystals

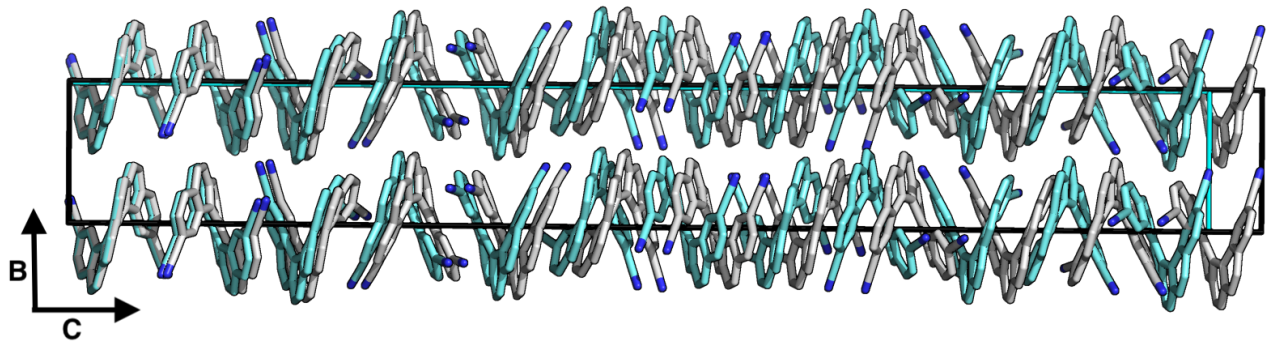


Figure 3.5: An overlay of the CSD (UJOCOL) structure (grey) and the CSP predicted 2,2'-dicyano-[6]helicene crystal structure at the DFT-XDM level (blue) in the *bc*-plane. The black (blue) cell shown belongs to the CSD (DFT-XDM) structure.

contain chiral columns. The experimental crystal with its extremely rare $P6_122$ space group and large unit cell size (c-axis >70 Å), was successfully predicted; the experimental structure was found to be the lowest-energy crystal structure (global minimum) at the DFT-XDM level of theory (Figure 3.4b). The $P6_122$ crystal was synthesised by Francesco Salerno and it forms a crystallite of hexagonal morphology along the c-stacking axis of the chiral columnar packing. The crystal packing similarity between the predicted and the CSD crystal structure was assessed using COMPACK[74] in MERCURY.[351] The predicted crystal structure is condensed by 4.4% along the c-axis compared to the CSD structure (Figure 3.5). Due to the more condensed unit cell size at DFT-XDM level of theory, overlaying all molecules inside the unit cell with minimal RMSD becomes poorer the further one moves away from the origin along the unit cell c-axis. Nonetheless, comparison of the predicted crystal structure with the CSD structure with a distance tolerance of 20% and an angle tolerance of 20° revealed identical crystal packings with 48 of 48 molecules within the neighbouring shell mapped onto one another with a root mean square deviation (RMSD) of 0.85 Å.

Next, the energy ranking of each crystal on the crystal-energy landscapes was analysed and the energy difference *i.e.* where a structure is relatively situated in the crystal-energy landscape across two levels of theory was compared between (i) crystals relaxed at force-field level using the Williams99 (W99) potential with distributed multipole analysis (DMA)[58] to obtain atomic multipole moments (hereafter W99) and (ii) crystals relaxed at the B86bPBE-XDM/DZP level of theory, with single point energies computed at B86bPBE-XDM/PAW (hereafter DFT-XDM). The energy re-ranking error E_{rank} is quantified as the energy difference for each crystal in relative energy to the respective global minimum across the two crystal-energy landscapes. The change in energy dif-

ference to the global minimum was analysed, because the correct ordering of the lowest-energy crystals (especially identifying the global minimum) is an important criterium for successful participation in the CSP blind tests. Thus, the 'energetic re-ranking error' across the two levels of theory was computed using

$$\Delta E_{\text{rank}} = (E_{\text{cryst}}^{\text{W99}} - E_{\text{glomin}}^{\text{W99}}) - (E_{\text{cryst}}^{\text{DFT-XDM}} - E_{\text{glomin}}^{\text{DFT-XDM}}) \tag{3.1}$$

where $E_{\text{cryst}}^{\text{W99}}$ and $E_{\text{cryst}}^{\text{DFT-XDM}}$ are the energies of the crystal and $E_{\text{glomin}}^{\text{W99}}$ and $E_{\text{glomin}}^{\text{DFT-XDM}}$ are the respective energies of the lowest-energy structure (global minimum) of each crystal-energy landscape. Whilst the experimentally observed $P6_122$ structure was the global minimum at the DFT-XDM

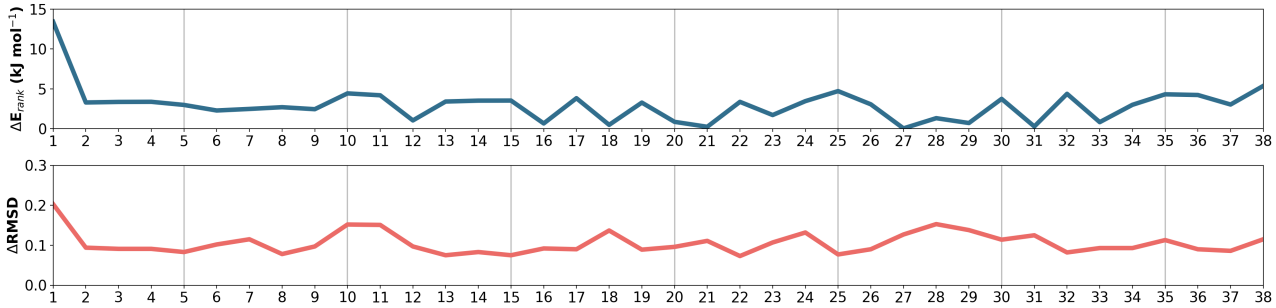


Figure 3.6: The lowest 38 2,2'-dicyano-[6]helicene crystals at the DFT-XDM level of theory are plotted in ascending order from the global minimum (crystal 1) with respect to (i) the relative energetic reranking error going from W99 to DFT-XDM (blue) and (ii) the relative change in RMSD going from W99 to DFT-XDM (red).

level, it was found 13.4 kJ mol^{-1} away from the global minimum at W99 level (Figure 3.4a). Hence, the largest 'energetic re-ranking error' across the entire landscape is observed for the $P6_122$ crystal structure. The $P6_122$ crystal structure (Z=12) also has the largest change in crystal geometry upon DFT-XDM relaxation quantified by the root mean square deviation (RMSD) (Figure 3.6). Upon DFT-XDM relaxation, the $P6_122$ crystal density increases by +0.02 g cm^{-3}.

Comparing the relative changes in ΔRMSD, there is no correlation between a large re-ranking error, ΔE_{rank}, and a large change in crystal geometry ΔRMSD (Figures 3.6, 3.7). Two other enantiopure crystal structures (structures 10, 11) have a slightly increased re-ranking error and a coinciding moderately larger change in crystal geometry, ΔRMSD (Figure 3.6). These two enantiopure structures do not display any (homo)chiral columnar packing. For the rigid helicenes discussed here, a large re-ranking error does not appear to be directly correlated to a large change in crystal

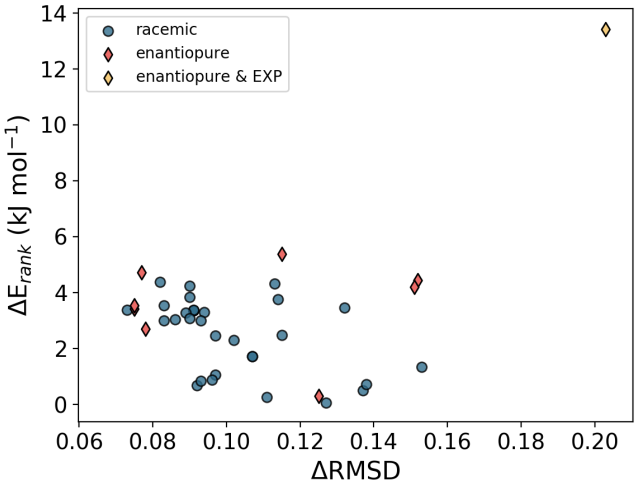


Figure 3.7: The relative change in crystal geometry upon periodic relaxation is not directly associated with a higher energetic reranking error.

geometry, *i.e.*, a relaxation of the periodic cell. Instead ΔE_{rank} is more likely to be an effect of better incorporating non-covalent interactions via the use of dispersion correction. There is no clear correlation between ΔE_{rank} and the space groups of the crystals, nor the number of molecules inside the unit cell, *i.e.*, unit cell size.

Given the high ΔE_{rank}, the lowest-energy structure, which is observed experimentally, would have been missed if only crystals with a maximum lattice energy difference of +7.2 kJ mol^{-1} (valid in 95% of cases) would have taken from the set of W99 crystals for further DFT-XDM minimisation.[352] Based on the high ΔE_{rank} observed uniquely for the $P6_122$ crystal structure with columnar packing, the question arose whether the increased π-π stacking in the chiral column could be responsible for the error at the W99 level. The $P6_122$ crystal structure is the crystal structure with the longest chiral column per unit cell. However, inspecting the crystal-energy landscape, other crystals displaying chiral columnar packing were identified (structures 1-5, 9, 13, 16, 19, 22, 23, 29). These crystals with columnar packing do not show an increased re-ranking error, which is why the hypothesis was disproved. Hirschfeld charge analysis of the periodically relaxed low-lying crystal structures (Figure 3.8) demonstrates that the $P6_122$ crystal structure, compared to the other low-lying crystal structures, is characterised by a very high degree of π-π stacking at ~1.7–1.8 Å, quanitfied here as C··C stacking, and a high degree of C-H··π stacking. The extremely pronounced light-blue π-π stacking at ~1.7–1.8 Å, which no other crystal contains, can be rationalised by the very long chiral column along the crystal's *c*-axis. Other structures containing chiral

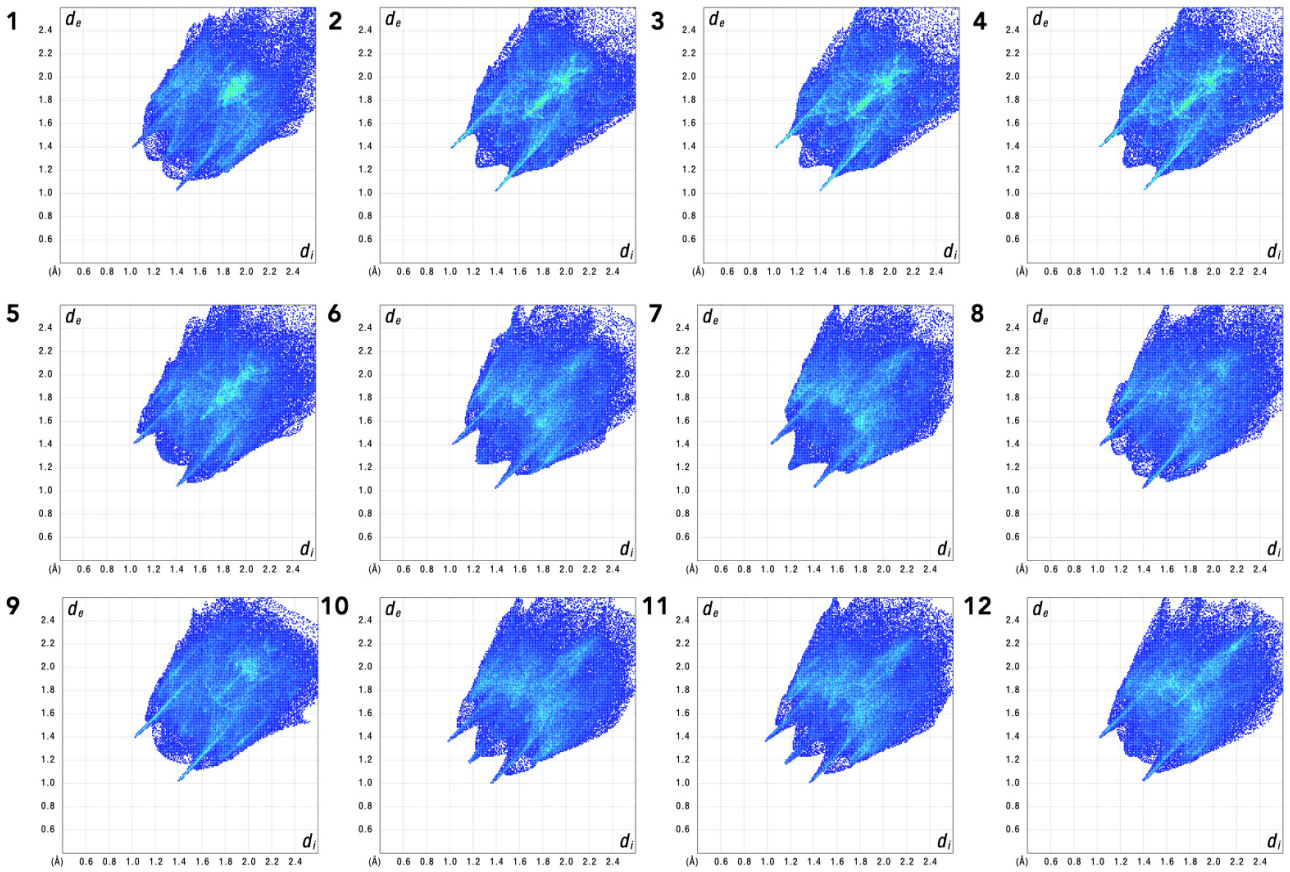
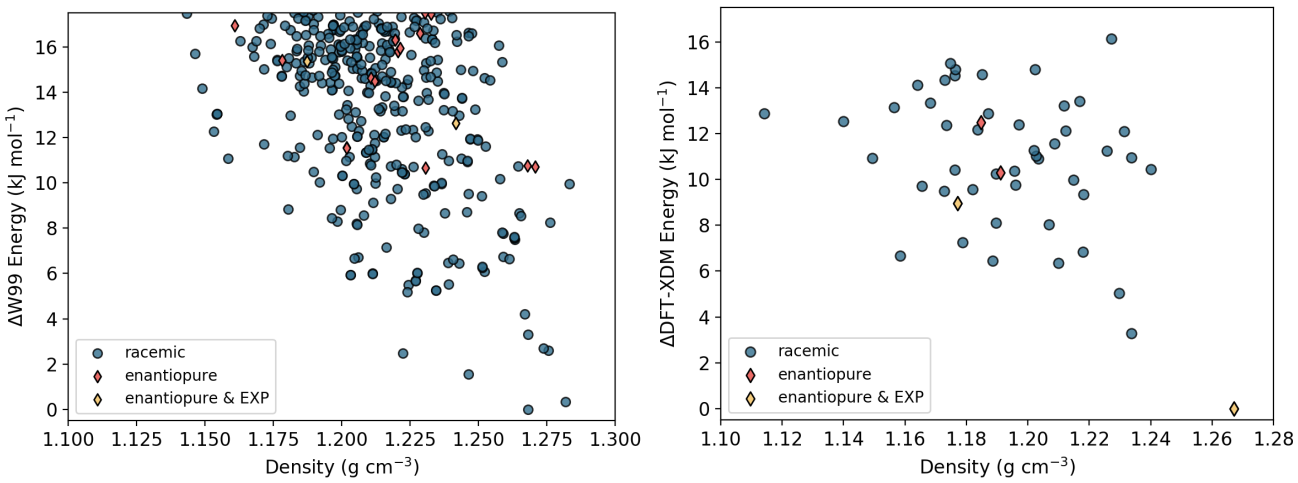


Figure 3.8: Two-dimensional (2D) hirschfeld fingerprint plots for the 12 lowest lying crystal structures for 2,2'-dicyano-[6]helicene. All plots were generated showing the region of interacting atoms between 0.4–2.5 Å. The distance to external (surface) atoms is denoted d_e (y-axis). The distance to the internal (surface) atoms, d_i (x-axis), is the distance from the helicene molecule's surface to the nearest atom on the same molecule. The colouring indicates the frequency of the (d_i, d_e) pair. Low frequency is denoted with dark blue, brighter areas indicate higher frequency of a (d_i, d_e). π-π interactions between aromatic rings show as brighter spots on the diagonal at ~1.7-1.8 Å (as found in structure 1 and less pronounced in structures 2-5), whilst C-H$\cdots\pi$ interactions would show as 'wings' on the structure. N··H contacts are the two 'tips' at (1.4, 1.0) and vice versa. H··H contacts are present if a third 'tip' is shown at ~1.2 Å (as found in structures 10 and 11).

columns (structures 2-5, 9) also show a continuous light blue area at ~1.8–2.0 Å.

3.3.1 2,2'-Diethynyl-[6]helicene CSP

Fifty unique crystal structures within 16 kJ mol^{-1} of the global minimum were relaxed at the DFT-XDM level of theory (Fig. 3.9b) and 4 were enantiopure (4.3%). 2 of the 4 enantiopure structures are the experimentally observed crystal structures (CSD references: NEMKUO, PICFEO01). The experimental $P3_1$ polymorph with columnar packing (CSD reference: NEMKUO) is the thermodynamically most stable solid-state structure, *i.e.* the global minimum. Comparing the crystal



(a) (b)

Figure 3.9: 2,2'-diethynyl-[6]helicene crystal-energy landscapes at the W99 (left) and DFT-XDM level (right). Racemic (enantiopure) crystal structures are shown with blue dots (red diamonds). The 25 lowest-energy crystals for which the charge-carrier mobilities have been computed are annotated. The experimentally observed polymorphs are highlighted in bold.

packing of this polymorph at DFT-XDM level of theory with the CSD deposited structure (CSD reference: NEMKUO); the entire neighbouring shell of 15 molecules around the molecule in the asymmetric unit could be mapped onto one another with a root mean square deviation (RMSD) of 0.329 Å. Similarly, the experimentally observed $P2_1$ polymorph with no columnar packing (CSD reference: PICFEO01), was predicted via CSP. The neighbouring shell, representative of the entire crystal environment, was compared between the DFT-XDM structure and the CSD structure. The difference in atomic coordinates resulted in a RMSD of 0.188 Å. Thus, the PICFEO01 crystal structure was more accurately predicted by CSP than the NEMKUO crystal structure, with a significantly lower RMSD (Fig. 3.10).

The set of lowest-energy crystals was numbered in order of increasing relative energy starting from the global minimum (Fig. 3.11). The trends in energetic re-ranking errors, ΔE_{rank}, were similar to 2,2'-dicyano-[6]helicene, with the largest error being observed for the experimental $P3_1$ polymorph with columnar packing (NEMKUO). The structure was initially found +12.6 kJ mol^{-1} away from the global minimum at W99 level (Table 3.1). The other experimentally observed, non-columnar, $P2_1$ crystal structure (PICFEO01) had a reduced energetic re-ranking error, ΔE_{rank}, because the structure remained +8.9 kJ mol^{-1} away from the global minimum upon DFT-XDM relaxation (initially +15.4 kJ mol^{-1} at W99 level). However, neither of the two structures were found in

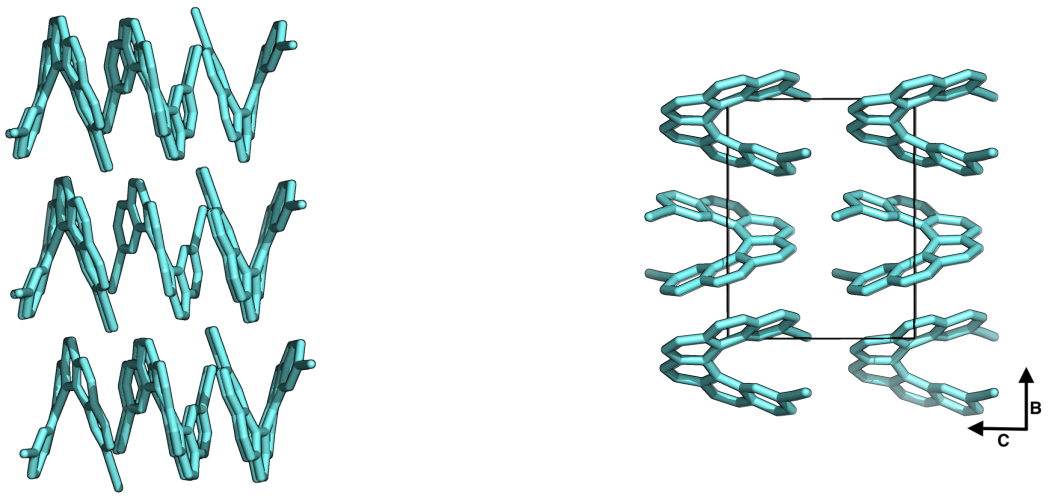


Figure 3.10: The 2,2'-diethynyl-[6]helicene CSD structures overlaid with computationally predicted structures (RMSD = 0.329 Å). Both structures are shown in blue. Left: An overlay of NEMKUO (blue) down the *b*-axis and the DFT-XDM structure (blue) down the *a*-axis (RMSD = 0.188 Å). Right: An overlay of PICFEO01 down the *a*-axis (grey) and the predicted crystal structure at B86bPBE-XDM level (blue) down the *c*-axis. Hydrogens are omitted for clarity.

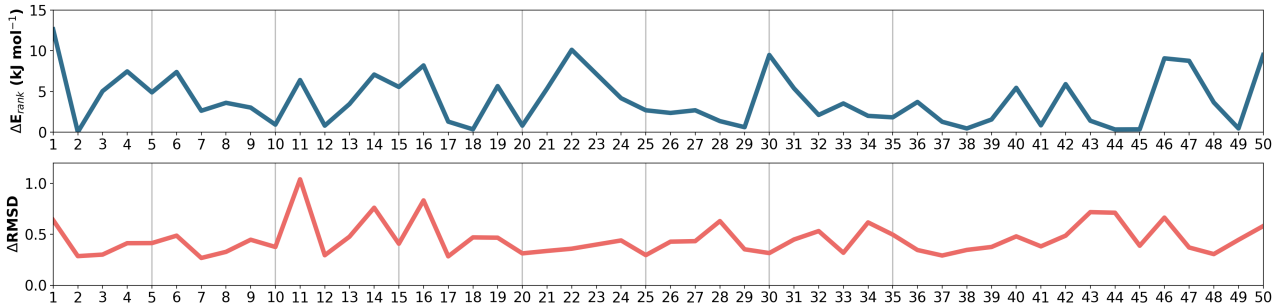


Figure 3.11: The 2,2'-diethynyl-[6]helicene: The lowest 50 crystals at the highest level of theory are plotted in ascending order from the global minimum (crystal 1) with respect to (i) the relative energetic reranking error going from the W99 to DFT-XDM level (blue) and (ii) the relative change in RMSD going from the W99 to DFT-XDM level (orange). The ΔRMSD changes upon DFT-XDM minimisation vary more than for 2,2'-dicyano-[6]helicene.

the polymorphic lattice energy range of +7.2 kJ mol^{-1} from the global minimum in which 95% of crystal structures are obtained.[352] Figure 3.11 demonstrates that there are more structures with an increased ΔE_{rank} >8 kJ mol^{-1} (structures 16, 22, 30, 46, 47, 50), but not as large as for the experimental $P3_1$ polymorph. The data shows a weak positive correlation between ΔE_{rank} and ΔRMSD with a Pearson coefficient of 0.30. The energetic re-ranking error, ΔE_{rank}, plotted against the change in crystal geometry, ΔRMSD, is illustrated in Figure 3.12. Changes in crystal geometry with ΔRMSD >0.55Å tend to result in some energetic re-ranking error, either of the range of 1-2 kJ mol^{-1} or >6 kJ mol^{-1}. The experimentally observed $P2_1$ crystal (PICFEO01) is the 11th

lowest-energy crystal at DFT-XDM level and the crystal with the highest change in crystal geometry, RMSD, upon DFT-XDM relaxation (RMSD=1.04 Å). Thus, whilst CrystalPredictorII is capable of predicting all experimentally observed polymorphs, *i.e.* the crystal geometries, predicted and CSD X-ray crystal structure, varied significantly (Fig. 3.10).

In the 2,2'-diethynyl-[6]helicene polymorphic landscape, there are two other low-lying crystal structures with columnar packing, besides the $P3_1$ global minimum structure. The fingerprints (Figures 3.13) highlight that chiral columnar packing can provoke a high frequency of π–π stacking (light blue area at ~1.8 Å). Figure 3.13 suggests that the $P3_1$ crystal structure (structure 1) and structure 12 are characterised by an exceptionally high degree of π-π stacking at ~1.7–1.8 Å which is more pronounced than for any other of the low-lying crystal structures shown. Among the set of 25 lowest-energy 2,2'-diethynyl-[6]helicene crystal structures, structures 1, 9, 12, 17, 19, 25 are arranged in chiral columns. Structure 9 is the outlier, containing chiral columnar packing, but no measurable π–π stacking interaction is observed for the fingerprint, which might be due to a potentially larger molecular spacing. Figure 3.14 illustrates the intermolecular distances in two structures with chiral columns (structure 19, 35); despite having high π-π stacking in chiral columns, they have a moderate to small re-ranking error, which disproves the hypothesis of the chiral column being responsible for the larger re-ranking error.

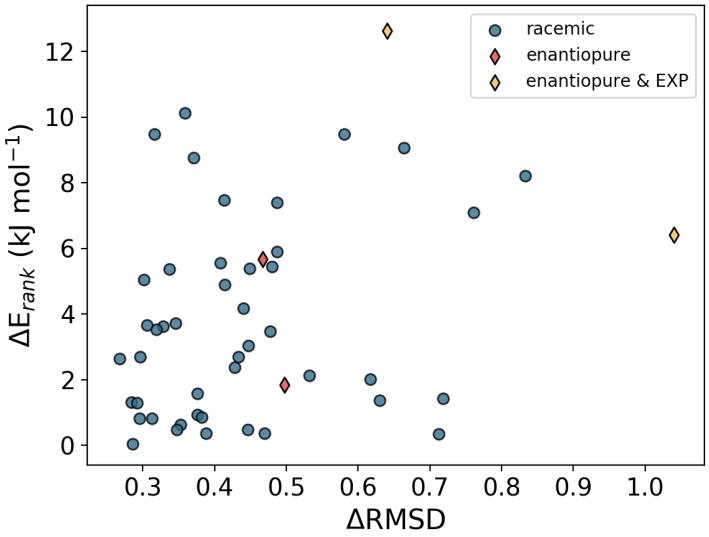


Figure 3.12: The relative change in 2,2'-diethynyl-[6]helicene crystal geometry upon periodic relaxation is not directly associated with a higher energetic reranking error.

In both CSP searches, the experimentally observed crystal structures, which were initially encountered high in energy (Table 3.1), were associated with the highest energetic re-ranking errors,

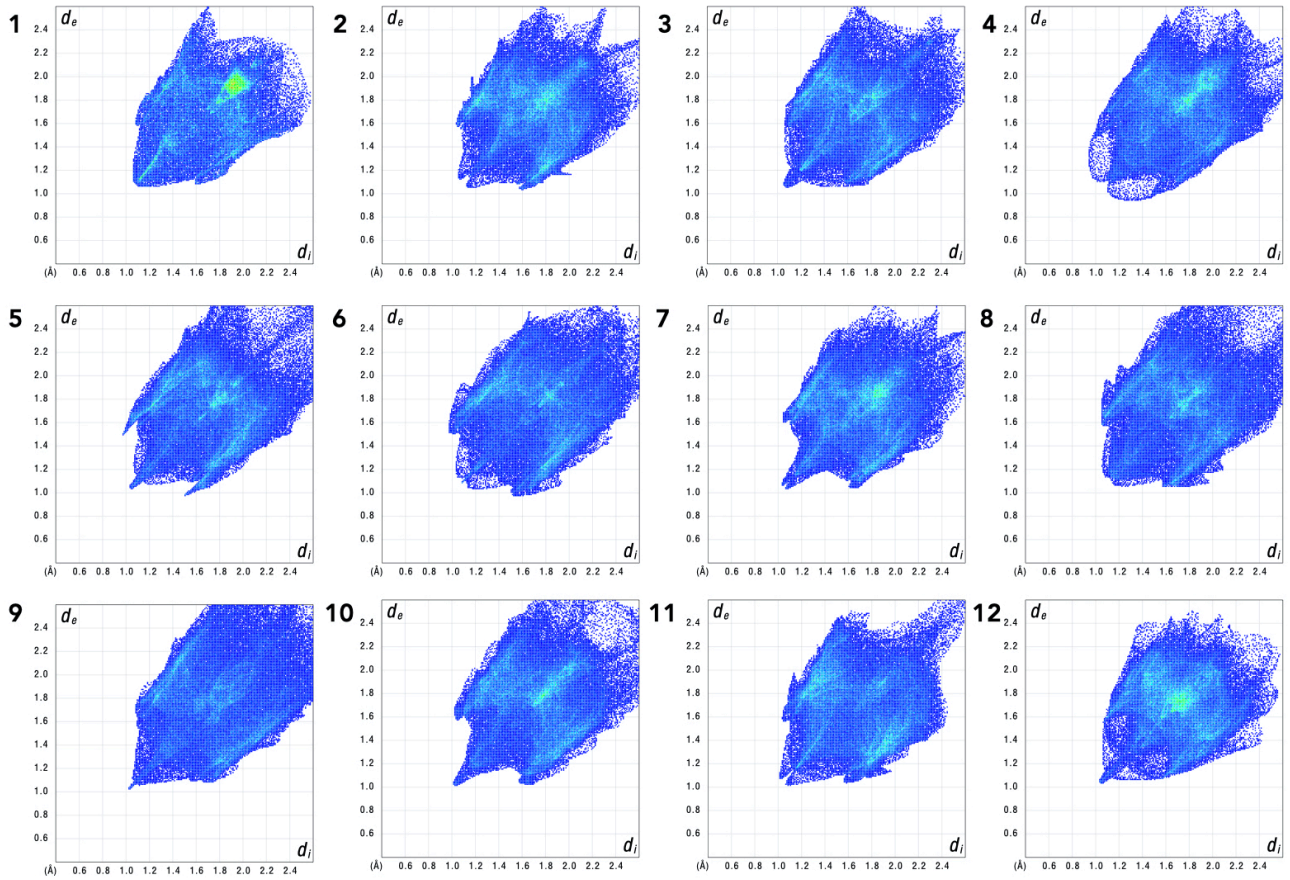


Figure 3.13: Two-dimensional (2D) hirschfeld fingerprint plots for the 12 lowest lying crystal structures for 2,2'-diethynyl-[6]helicene. All plots were generated showing the region of interacting atoms between 0.4–2.5 Å. The distance to external (surface) atoms is denoted d_e (y-axis). The distance to the internal (surface) atoms, d_i (x-axis), is the distance from the helicene molecule's surface to the nearest atom on the same molecule. The colouring indicates the frequency of the (d_i, d_e) pair. Low frequency is denoted with dark blue, brighter areas indicate higher frequency of a (d_i, d_e). π-π interactions between aromatic rings show as brighter spots on the diagonal at ~1.7-1.8 Å (as found in structure 1 and less pronounced in structures 4,7,12), whilst C-H$\cdots\pi$ interactions would show as 'wings' on the structure. H··H contacts are present if a third 'tip' is shown in between at ~1.2 Å (as found in structure 11).

i.e., the experimental crystals with columnar packing (NEMKUO, UJOCOL) were initially found high in energy (+12.6 and +13.4 kJ mol^{-1}) at the W99 level. It was disproven that the high ΔE_{rank} > could be arising from the crystal being arranged in chiral columns. Both molecules contain multiple low-lying crystal structures arranged in (homo)chiral columns. For 2,2'-dicyano-[6]helicene, the 5 lowest-energy crystal structures contain chiral columns. In the next section, the crystal-packing (dis)similarity across the two crystal-energy landscapes will be discussed (with a focus on chiral columnar packing).

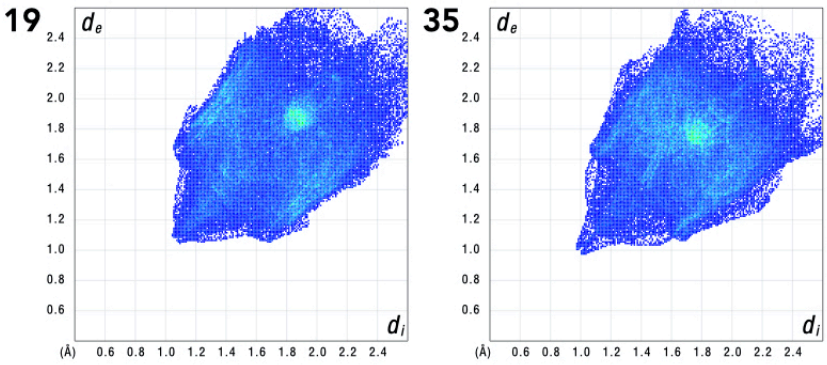


Figure 3.14: Two-dimensional (2D) Hirschfeld fingerprint plots for the two other low-lying 2,2'-diethynyl-[6]helicene crystal structures containing chiral columnar packing. All plots were generated showing the region of interacting atoms between 0.4–2.5 Å. The distance to external (surface) atoms is denoted d_e (y-axis). The distance to the internal (surface) atoms, d_i (x-axis), is the distance from the helicene molecule's surface to the nearest atom on the same molecule. The colouring indicates the frequency of the (d_i, d_e) pair. Low frequency is denoted with dark blue, brighter areas indicate higher frequency of a (d_i, d_e). π-π interactions between aromatic rings show as brighter spots on the diagonal at ~1.7-1.8 Å (as found in both structures), whilst C-H$\cdots\pi$ interactions show as 'wings' on the structure.

			W99		DFT-XDM	
crystal	EXP	space group	density	Δ E	density	Δ E
UJOCOL	1	$P6_122$	1.27	13.4	1.29	0.0
NEMKUO	1	$P3_1$	1.24	12.6	1.27	0.0
PICFEO01	2	$P2_1$	1.19	15.4	1.18	8.9

Table 3.1: Table highlighting the density and relative energies in the CSP landscape associated with the experimentally observed polymorphs at W99 and DFT-XDM level of theory. Densities are given in g cm^3. Energy differences are given in kJ mol^{-1}.

3.3.2 Identifying Recurring Structural Motifs Across the 2,2'-Dicyano- and 2,2'-Diethynyl-[6]helicene Crystal-energy Landscapes

This section will provide an analysis of the variation in solid-state packings across the two computed crystal-energy landscapes by using CRYSTACK intermolecular crystal stacking information. The motif search is intended to find the three motifs, interlocked (I), back-to-back (B) and translational (T), and is based on the assumption that the solid-state arrangements of the two terminally functionalised [6]helicenes will resemble the unsubstituted reference. The heuristic rules, *i.e.* the conditions based on which a motif was classified were empirically fitted according to unique features present in either of the three motifs. The rules were derived with the intention of minimising false positives at the expense of risking a few false negatives, which would not strictly coincide

with the conditions defined. The following heuristic rules were applied:

(i) An interlocked I dimer motif, despite containing some symmetry requirements (*i.e.* the two molecules need to lie in the same plane and point in opposite directions), tends to contain 1-2 π-π stacks per dimer. If two π-π stacking interactions are found, they must both occur between terminal aromatic rings only, and one π-π stack must be a parallel π-stack, whilst the other must be an orthogonal C-H$\cdot\cdot\pi$ stacking interaction. If only one π-π stacking interaction is observed, it must also occur between the two terminal rings of the two helicenes. An interlocked motif without any π-π stacks was only classified as interlocked, if it instead possessed a N$\cdot\cdot$C short contact between the nitrogen and the carbon on the other molecule. This scenario was added for the cyano side group where π-π stacking is sterically hindered. Some 2,2'-diethynyl-[6]helicene I dimers showed a similar C$\cdot\cdot$C short contact based on its side group.

(ii) The B motifs tend to contain ≥3 π-π stacking interactions per dimer. If 3–4 π-π stacks were found, they must all be parallel π-π stacks. This type of condition usually applied to the heterochiral dimers with backbones of opposite helical pitch stacked neatly parallel on top of one another. However, for the homochiral B motif, the π-π stacking interactions are not as neat. Therefore, the homochiral B motif is characterised by fewer (0-1) π-π stacks and it tends to instead occur with many ≥10 hydrophobic interactions (many aromatic carbon atoms of close contact). This homochiral B dimer is frequently observed for helicenes arranged in homochiral columnar packing.

(iii) The T motifs are (unlike B and I) purely translational motifs. The T motif occurs with ≤3 π-π stacking interactions. The motif is present if 1-3 parallel π-π stacks are found between non-terminal rings and the two molecules are nearly perfectly parallel to each other. If no π-π stacking interaction is found, the angle between the centroid-centroid vector and the centroid-terminal rings vector is computed and must lie around 80-110°. This is to avoid misclassification of other motifs (*e.g.* herringbone (Figure 1.5, motif X), side-translational[312]) as translational.

Recurring Structural Motifs in 2,2'-Dicyano-[6]helicene Crystals

Figure 3.15a suggests that many of the low-energy, *low-density* crystals contain the T and B motifs (2/3), whilst low-energy, *high-density* crystals tend to contain all three (T, B, I) motifs (3/3). This suggests that if all three motifs are present in one crystal; I and B in one plane and the T in the plane

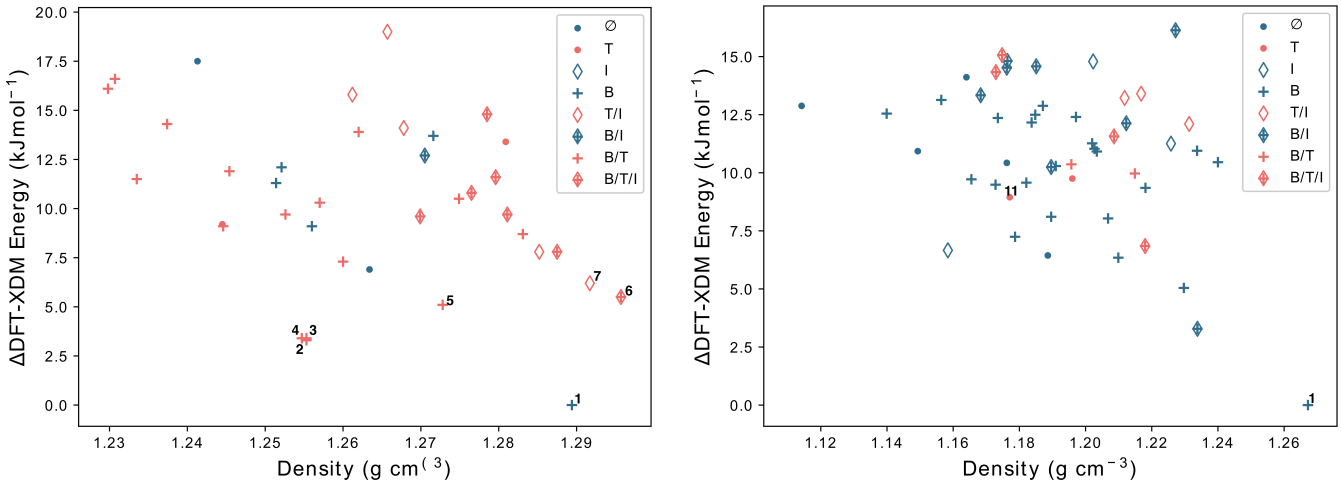
orthogonal to it, this might give rise to a more density-efficient packing than without the I motif.

Figure 3.15a illustrates that the experimentally observed crystal structure (global minimum) only contains the B motif. There were multiple B motifs found for this structure due to the chiral columnar packing. However, the T was not found in the set of the five closest unique dimers around the central molecule. Given that there are 12 molecules in the unit cell and this workflow focusses on the five closest unique nearest-neighbour molecules to identify key interactions, the T dimer was not classified, because other dimers *e.g.* herringbone and B were found at shorter distances. The largest energetic reranking error was found for crystal structure **1**. For this crystal structure two types of back-to-back dimers (part of the chiral column) are the most important motifs. The set of second lowest-lying crystals (structures **2**, **3**, **4**) also contained two types of B motifs (Figure 3.15a, red crosses). The next set of crystals slightly higher up in energy (structures **5**, **6**, **7**) did not contain two types of B motifs. Crystals 1-5 all contain some B motifs because they show some type of chiral columnar packing (Figure 3.17), which is also the reason why for some crystals two types of B motifs are observed within the set of nearest neighbour dimers.

The interlocked dimer was less commonly found in the 2,2'-dicyano-[6]helicene, which might be in part caused by our rigid definition of the interlocked dimer as either containing terminal π-π stacking interactions or N··C short contacts between the cyano group of one molecule and the backbone of the other; but because terminal difunctionalisation in the 2,2'-position affects this motif the most with respect to intermolecular distances and geometry. In T and B motifs, the side groups are not involved in any key interactions. The three motifs classification on both crystal-energy landscapes was performed using the exact same criteria.

Recurring Structural Motifs in 2,2'-Diethynyl-[6]helicene Crystals

Figure 3.15b of the 2,2'-diethynyl-[6]helicene polymorphic landscapes shows that the B motif was most frequently observed. The I motif is still more frequently found than for 2,2'-dicyano-[6]helicene. However, the T motif is significantly less frequently found than for 2,2'-dicyano-[6]helicene (fewer red data points). Similar to the 2,2'-dicyano-[6]helicene landscape, the majority of crystals containing the I motif tend to be higher density crystals of a given relative energy, whilst lower density crystals less frequently arrange in this motif.



(a) 2,2'-dicyano[6]helicene

(b) 2,2'-diethynyl[6]helicene

Figure 3.15: Crystal-energy landscapes illustrating whether the T, B, I motifs are found within the closest nearest-neighbour dimers of a crystal.

The lowest-lying and experimentally observed polymorph, which arranges in a chiral column, contains the B motif. The overlap angle of the T is not parallel, *i.e.* the planes of the molecules are not oriented parallel to each other, which is why this crystal was not classified as containing the T motif. The 11th lowest-lying and experimentally observed polymorph (CSD reference code: PICFEO01) only contained the T motif. Whilst Figure 3.10 might suggest that there are homochiral B motifs, the two molecules are not in the same plane, *i.e.* the molecules are orthogonal to each other, which is why it was not classified as a typical B motif.

One way to tackle this problem is to develop a software tool targeted for solid-state pattern recognition specific to the dimer-based motifs seen in this small set of helicene molecules. However, this approach would only be of use to our small data set and be extremely time-intensive to develop because it requires a lot of unit testing of the derived heuristic rules, which need to be distinctive enough of a motif to not identify false positives. This approach is highly interpretable for the user and extracts a lot of high-level information about the solid-state environment of a crystal. However, the actual classification based on heuristic rules has very limited functionality. A rigid set of heuristic rules is not flexible enough to adapt to the vast set of crystalline packings present in the polymorphic landscape. Instead, more sophisticated solutions, maybe incorporating machine-learning (ML) techniques that learn dimer-based motifs, to classify new unseen motifs with a high degree of similarity as belonging to the same class of motifs would be much

more useful. Future work could also include moving from a rigid classification to a percentage likeliness score to better account for the fluid reality of solid-state arrangements, which cannot be accounted for by discrete 'motifs'.

3.3.3 Graph Kernel-based Crystal Similarity Map

The (dis)similarity of solid-state arrangements is hard to quantify, because the dominant non-covalent contacts are not straightforward to pinpoint. The crystal packing similarity was computed via a Pyramid Match (PM) kernel similarity matrix of the crystal graph representations. The PM kernel successfully identified duplicate crystal structures, not already detected by COMPACK. For the 2,2'-dicyano-[6]helicene (2,2'-diethynyl-[6]helicene) polymorphic landscape, 11 (1) duplicate crystal structures were identified. For example, the kernel detected two structures (ID: 2, 12) of the 2,2'-diethynyl-[6]helicene polymorphic landscape to have the same crystalline packing. However, the structures were not initially identified as duplicates by conventional tools, because COMPACK only overlaid 9/15 molecules of the neighbouring shell with a RMSD of 0.809 Å at the default distance and angle tolerance cutoffs (20% tolerance). This is due to the fact that the molecular assembly is the same, but the actual non-covalent distances are altered. Similarly, for 2,2'-dicyano-[6]helicene various duplicate structures were identified during graph kernel-based clustering, which were not identified before (crystal IDs: (325,334), (213,214), (78,90), (582,585), (561,566), (67,71)).

2,2'-Dicyano-[6]helicene

Various clustering algorithms and hyper-parameters were tested to best cluster the 2,2'-dicyano-[6]helicene crystal-energy landscape. Good clustering algorithms usually lead to a low "within-cluster" variation, *i.e.* a good compactness of clusters, and a high "between-cluster" variation, *i.e.* clusters are well separated. The density-based clustering algorithm, DBSCAN, and the Louvain method for community detection algorithm did not perform as well on the given data (Figure A.1). A grid search using both the PM (4,6) and PM (5,7) graph kernel and varying between (i) K-Means and agglomerative clustering, (ii) the number of principal components (2,3,6), and (iii) the number of distinct clusters from 2-11, was performed. Within the agglomerative clustering algorithm,

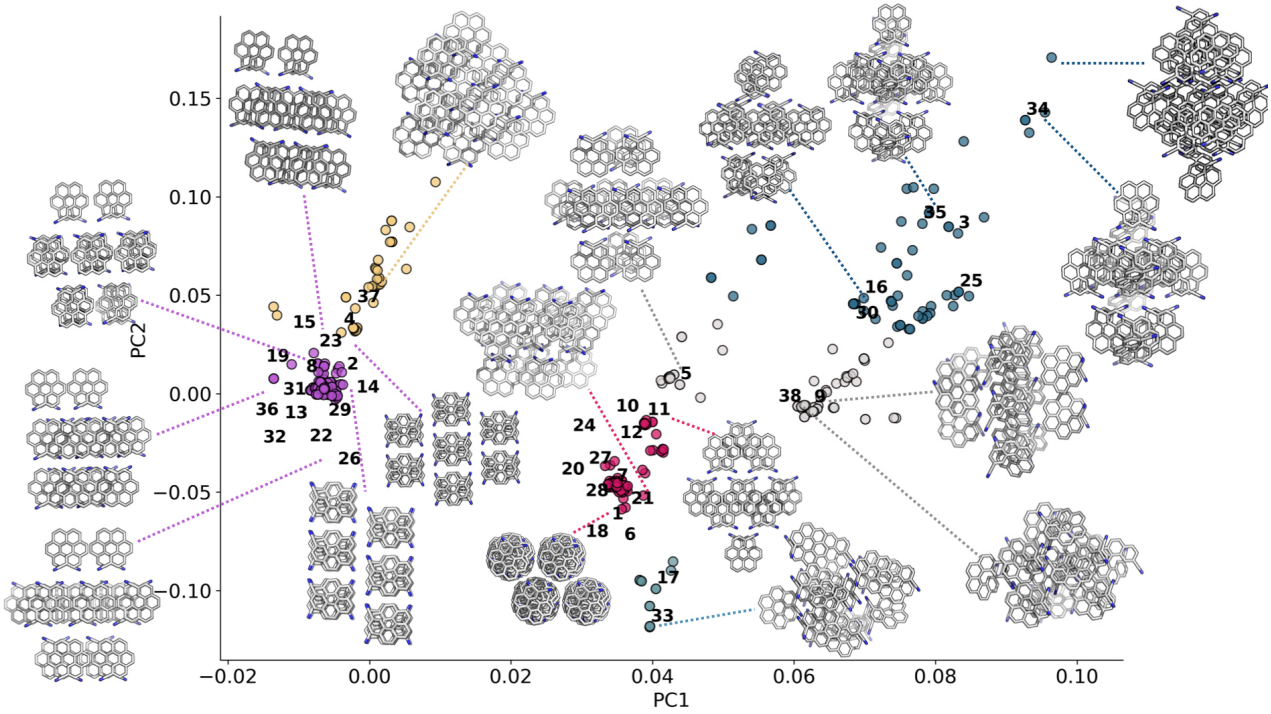


Figure 3.16: Graph PM kernel-based agglomerative cluster similarity map of 2,2'-dicyano-[6]helicene of the first two principal components and using six clusters. The six different clusters are highlighted in grey, blue, turquoise, red, yellow and purple. For each cluster, crystal packing examples are shown.

different algorithms upon which two points would be merged into the same cluster were assessed, *i.e.* the performance of the *ward*, *single* and *complete* linkage algorithms were compared. Generally, agglomerative clustering performed better than K-means clustering. The optimal linkage algorithm, which maximises the distance between the clusters in hierarchical agglomerative clustering was found to be the *ward* algorithm. The *ward* algorithm aims to minimise the increase in error sum of squares (ESS) upon merging two clusters together. Based on our analysis, the agglomerative clustering algorithm, using the *ward* linkage criterium, six clusters, and the first two principal components yielded the best results and the lowest Davies Bouldin score 0.45. The lower the Davies Bouldin index, the better the clustering and the more pronounced the ratio between "within-cluster" and "between-cluster" distances. For comparison, the best split of the data using agglomerative clustering on the first three principal components of the data was found for a cluster size of 4 with a Davies Bouldin score of 0.65. The dendrogram and the distances between the cluster centres supported the selection of 6 clusters (Figure 3.16).

There was no direct correlation between the cluster-based similarity map and the crystal lattice

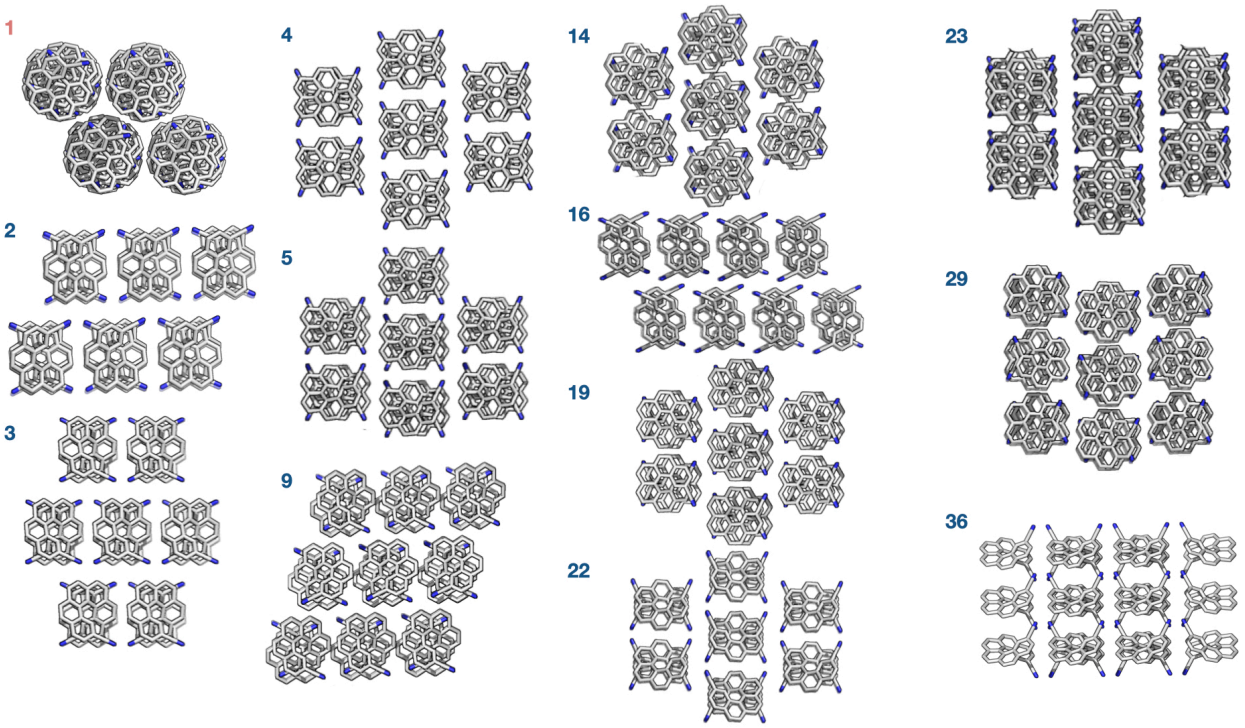


Figure 3.17: 2,2'-Dicyano-[6]helicene crystal structures that arrange in chiral columns. The structures are numbered by their relative position in the DFT-XDM energy landscape. Hydrogens are omitted for clarity. Homochiral columns are labelled in red (structure 1), heterochiral columns in blue.

energy or density. There was not sufficient mobility data to colour a large proportion of the crystal-energy landscape by its mobility. However, some packing types tend to be correlated with higher charge transport (*e.g.* purple cluster) due to their high degree of parallel π-π stacking, whilst others tend to have negligibly small charge transport and a high degree of disorder (blue cluster). For the few crystals computed in the landscape, all crystals with good to very good mobilities (>0.3 $cm^2\ V^{-1}\ s^{-1}$) possess a high degree of parallel π-π stacking. The interplay between structure and electronic properties will be discussed in the next section (Section 3.3.4). 2,2'-dicyano-[6]helicene has many crystals with many different types of chiral columnar packing (Figure 3.17), which are not all grouped into the same cluster.

2,2'-Diethynyl-[6]helicene

The crystal-energy landscape of 2,2'-diethynyl-[6]helicene is clustered using the Louvain local search algorithm, which naturally finds the best cluster size based on optimising the modularity. The modularity is a quality index for clusterings, based on the number of edges falling within

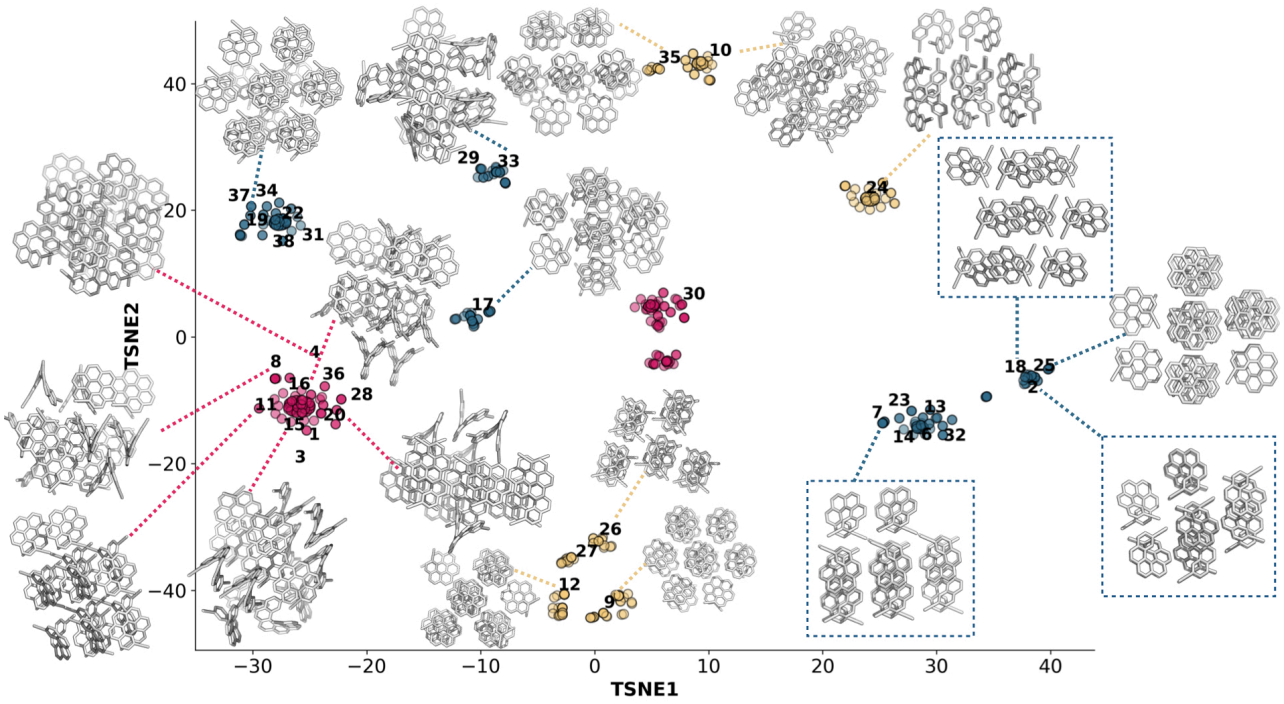


Figure 3.18: Graph kernel-based similarity map of the 2,2'-diethynyl-[6]helicene polymorphic landscape obtained via Louvain community detection. The three different clusters are highlighted in blue, red and yellow. For each cluster, crystal packing examples are provided. Examples of a new packing motif, the 'sparse double-stranded chiral columns', are highlighted with dashed boxes surrounding the neighbouring shells.

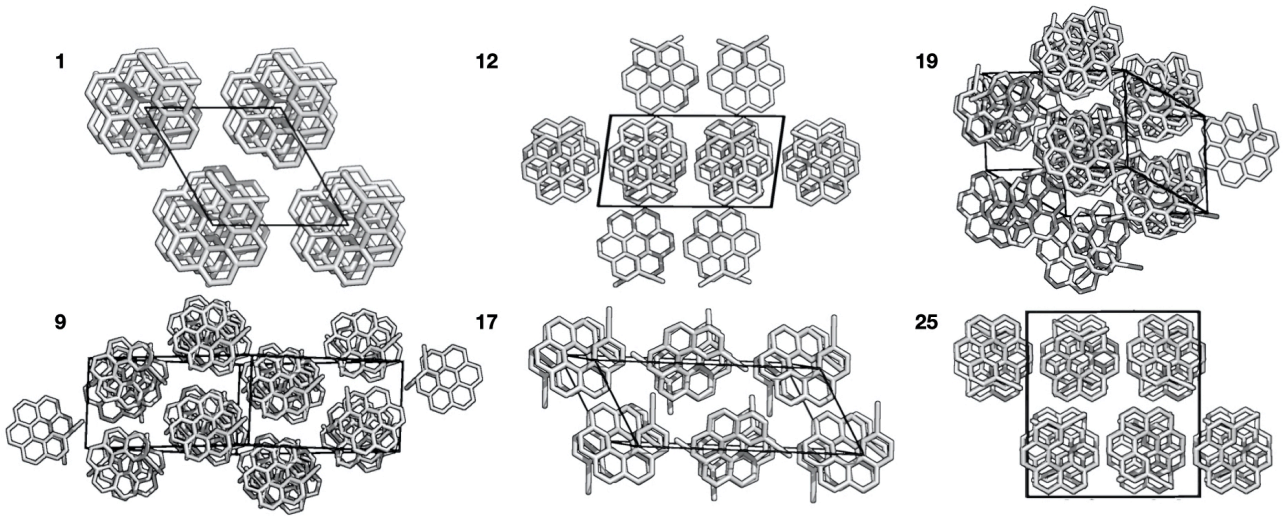


Figure 3.19: Illustration of the 2,2'-diethynyl-[6]helicene crystal structures that arrange in chiral columns. The structures are numbered by their relative position in the DFT-XDM energy landscape. Hydrogens are omitted for clarity.

groups minus the expected number in an equivalent network with edges placed at random.[353] The Louvain[288] clustering algorithm is a popular clustering algorithm for graph similarity clustering and has received much attention over the last decade because of its very fast and very good clustering results, outperforming most other clustering algorithms.[354] The Louvain clustering algorithm identified three distinct clusters based on optimising modularity (Figure 3.18). The pink cluster is characterised by sheets of parallel π-π, which are orthogonal to adjacent strips of π-π stacking. The yellow cluster contained chiral columnar type packing in one region (structures 9, 12, 26) and more disordered types of packings (structures 10, 24). The blue cluster was formed by crystals showing columnar packing, *e.g.* chiral columns (structures 19, 25) or double-stranded chiral columns (structures 2,7,18) and some disordered structures (structure 33). The two experimentally observed crystal structures form part of the same cluster. The resulting clustering maps were coloured by structure-dependent properties, *e.g.* density, W99 potential energy and mobilities, for which only limited examples were computed (25), but no consistent trend was observed. Based on visual inspection, six crystal structures were identified as containing some type of chiral columnar packing (Figure 3.19), but they did not end up in the same cluster on the similarity map. Nonetheless, the successful identification of duplicate structures based on this approach was further supported by molecular overlays of structures very close to each other on the graph kernel-based similarity map; *e.g.* structure 2 and 18 are found in the same cluster and very close

on the map. If superimposed using MERCURY, they have 7/15 molecules in common. Similarly, structure 33 and 29 were found reasonably close on the map and they have a shared heterochiral back-to-back motif.

In conclusion, graph kernel-based clustering is far from being a perfect tool, because, for example, there are not enough features extracted for each crystal to allow a sparser similarity map for 2,2'-diethynyl-[6]helicene. In the future, instead of building the graph representations based on full neighbouring shells, just the neighbouring atoms within a radial cutoff could be used to characterise the differences in the Z'=1 crystalline environments. Whilst several kernels were tried to yield a best graph similarity measure, ideally one would want to derive their own kernel, specifically tailored towards organic molecular crystals, to yield the best analysis of local and more global features, such as chiral columnar packing.

3.3.4 Charge-carrier Mobilities Across the Crystal-energy Landscape

This section will discuss the charge mobilities observed across the polymorphs of both 2,2'-disubstituted [6]helicenes and the unsubstituted reference. The charge-carrier mobilities of the 25 lowest-lying [6]helicene crystal structures of the CSP search reported in Ref. 197 were recomputed with an outer reorganisation energy λ_o of 0.3 eV and a hole and electron inner-sphere reorganisation energy λ_i of 0.156 eV. For all results discussed in the following, the transfer integrals J are significantly lower than the reorganisation energy λ, in consequence, the values can be interpreted qualitatively to describe charge transport in the material. Table 3.2 and 3.3 list the maximum hole and electron mobilities across multiple polymorphs of each of the three molecules under investigation.

To best assess the charge transport properties observed across the 2,2'-diethynyl- and 2,2'-dicyano-[6]helicene here, the highlights of the [6]helicene structure-mobility landscape are summarised. In contrast to the polymorphic landscapes observed for the 2,2'-disubstituted [6]helicene molecules, none of the [6]helicene crystals arrange in chiral columns. Solid-state arrangements of [6]helicene crystals with the highest hole μ_{max}, 0.34-0.36 $cm^2\ V^{-1}\ s^{-1}$ include structure 3, 12 and 23 (Table 3.2, Figure 3.20). The exceptionally high anisotropy of structure 3 compared to *e.g.* structure 23, can be rationalised by the former only having a single plane of high hole transport, namely

molecule	E_{rank}	J_{max} (meV)	chiral column?	μ_{max}	plane(μ_{max})	A1	A2
[6]H	3	55	-	0.35	AC	0.80	157.35
	12	110	- (B)	0.36	BC	0.53	12.45
	23	98	- (B)	0.34	BC	0.42	5.32
2,2'-dicyano-[6]helicene	1	105	Y	0.42	BC	1.03	38.51
	6	151	N	1.26	AB	0.98	658.54
	14	37	Y	0.30	AC	0.63	20.36
2,2'-diethynyl-[6]helicene	9	93	Y	0.36	AC	0.97	34.70
	12	64	Y	0.33	AC	1.00	76.43
	13	90	N	0.31	AB	0.48	5.46
	16	102	N	0.45	BC	1.00	43.50
	17	43	Y	0.52	AC	0.61	14.57
	19	124	Y	0.59	BC	1.02	43.94
	25	70	Y	0.33	BC	0.18	2.00

Table 3.2: Table illustrating the highest hole mobilities obtained within the polymorphic landscapes of [6]helicene, 2,2'-dicyano- and 2,2'-diethynyl-[6]helicene. The structures are numbered according to their relative position in the DFT-XDM crystal-energy landscape. The 'chiral column?' column indicates, whether the crystal arranges in a chiral column, if so (Y), otherwise (-). For [6]helicene, which does not contain any chiral columnar packing, the motif giving rise to the J_{max} is given. Mobilities are given in $cm^2\ V^{-1}\ s^{-1}$ and only crystal mobilities exceeding 0.30 $cm^2\ V^{-1}\ s^{-1}$ are shown.

along the herringbone chain (J_{max} = 55 meV). The highest hole-transporting dimer in structures 12 and 23 is the back-to-back (B) pair with a J_{max} of 98 and 110 meV respectively. These findings are consistent with previous results by Rice *et al.*.[197] Across the lowest-lying [6]helicene polymorphs, six polymorphs have electron mobilities exceeding 0.40 $cm^2\ V^{-1}\ s^{-1}$, *i.e.* the maximum electron mobilities obtained slightly exceed the hole mobilities with a maximum electron mobility of 0.54 $cm^2\ V^{-1}\ s^{-1}$, compared to a maximum hole mobility of 0.36 $cm^2\ V^{-1}\ s^{-1}$ (Figure 3.21a). Across the set of highest electron-transporting crystals, despite structures 12 and 25, high electron transport, μ_{max}, is associated with the presence of the homochiral translational dimer motif (Table 3.3). Hence, the presence of a homochiral translational motif at a good overlap angle and separation, with J = 69-72 meV, and/or a back-to-back dimer with J = 91-123 meV are indicators for good electron transport in [6]helicene polymorphs. However, structure 23, illustrated in Figure 3.20a is an example of high J_{max} not necessarily guaranteeing high transport; the highest electron-transporting back-to-back dimer (95 meV) is disconnected from other high-transporting dimers. Consequently, maximum electron transport for structure 23 is only 0.17 $cm^2\ V^{-1}\ s^{-1}$.



(a) Structure 23 (b) Structure 12

Figure 3.20: [6]helicene neighbouring shells, *i.e.* crystal packing excerpts. Both structures have a high-transporting back-to-back dimer, but only structure 12 has exceptionally high mobility.

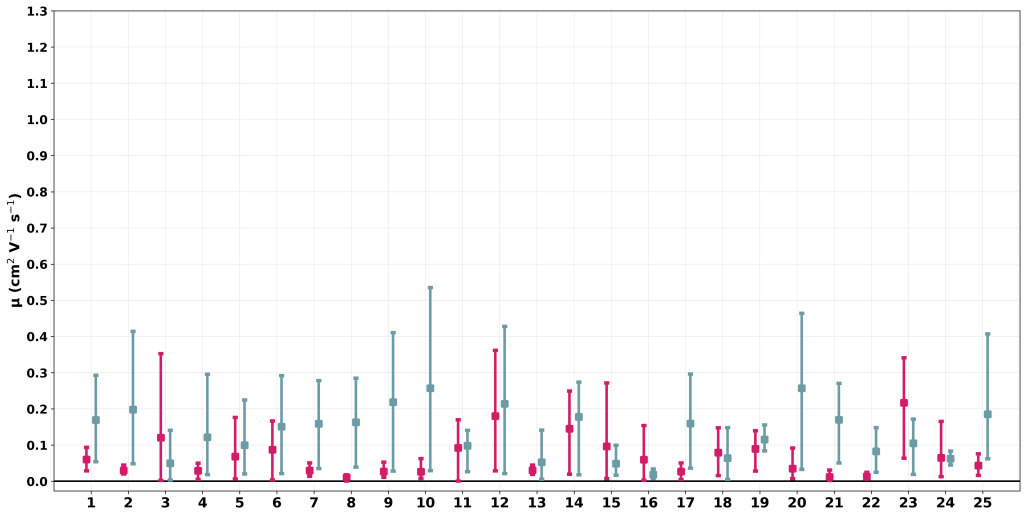
For the set of 38 lowest-lying crystal structures of the polymorphic landscape of 2,2'-dicyano-[6]helicene, the charge-carrier mobilities were computed. The polymorph with the highest predicted hole transport, μ_{max}, is structure 6. Hole μ_{max} is highest along the back-to-back dimers stacked along the *b*-axis, 1.26 $cm^2\ V^{-1}\ s^{-1}$ (Figure 3.22). None of the electron mobilities are of similar scale to the highest hole-transporting polymorph, but far more polymorphs showed μ_{max} exceeding 0.30 $cm^2\ V^{-1}\ s^{-1}$ than for hole transport (Tables 3.3 and 3.2). The highest-predicted electron mobility, μ_{max}, was less than half the magnitude of the hole, 0.49 $cm^2\ V^{-1}\ s^{-1}$, and observed for structure 5 with the translational motif as the highest-transporting dimer (Figure 3.23, Table 3.3). The experimentally observed $P6_122$ crystal (structure 1) is amongst the top hole-transporting crystals (μ_{max} = 0.42 $cm^2\ V^{-1}\ s^{-1}$), and somewhat ambipolar with an electron mobility, μ_{max}, of 0.26 $cm^2\ V^{-1}\ s^{-1}$. For either carrier type, charge transport is highest along the (homo)chiral column, along the *c*-axis of the crystal.

Across the set of 2,2'-dicyano-[6]helicene polymorphs, for both hole and electron transport, the highest transporting dimer motif is the translational and back-to-back dimer motif respectively. Back-to-back dimer motifs are generally frequently observed in high-transporting crystal structures. The π-π stacking across back-to-back dimers can vary significantly. Especially if the back-to-back dimer occurs as part of a 'chiral column', depending on the helical pitch, *i.e.* the number of molecules required for a full 360° rotation, the motif is altered. The longest helical pitch is observed for the experimentally observed crystal structure (structure 1) with 12 molecules required for a full 360° rotation, giving rise to the second-highest hole-transporting pathway along the chiral column (Figure 3.25). Comparing structure 1 to other high-transporting structures with

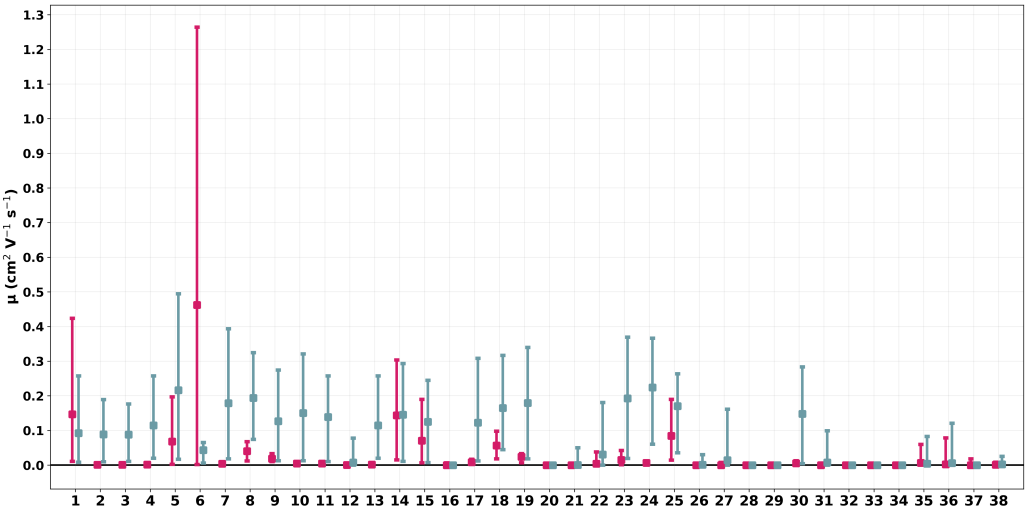
chiral columnar packing, structure 5, the highest electron-transporting 2,2'-dicyano structure, only requires 2 molecules for a full rotation. However, the highest charge transport pathway does not occur along the chiral column (Figure 3.23). Back-to-back dimers appear as the highest-transporting dimers as part of chiral columns in structures 1 (Figures 3.24-3.25), 19, 23 and 29. Out of the four lowest-energy polymorphs with chiral columnar packing, one polymorph with columnar packing is the second highest hole-transporting crystal (structure 1) and two polymorphs with columnar packing are amongst the five highest electron-transporting crystals (structures 19 and 23). Merely, polymorph 29 demonstrates that containing a chiral column with a high-transporting dimer does not guarantee high transport, because not every back-to-back dimer in the chiral column displays an equally high J and thus kinetically traps the charge. Consequently, neither chiral columnar packing nor the presence of high-transporting back-to-back dimers do *per se* assure a high-performance OSC molecular material.

The charge-carrier mobilities for the 25 lowest-energy crystal structures of the 2,2'-diethynyl-[6]helicene polymorphic landscape were computed. Across the set of 2,2'-diethynyl-[6]helicene polymorphs, no predisposition towards hole or electron transport is observed, maximum mobilities predicted for both carriers are of similar magnitude (0.59-0.60 $cm^2\ V^{-1}\ s^{-1}$) and occur at the same relative frequency, *i.e.* in either case three polymorphs have hole and electron μ_{max} >0.40 $cm^2\ V^{-1}\ s^{-1}$. Ambipolar carrier-transporting abilities were observed for structures 12 and 13, with maximum hole and electron mobilities of 0.33, 0.31 and 0.43, 0.60 $cm^2\ V^{-1}\ s^{-1}$ respectively. If one was searching for polymorphs with ambipolar transport, future research could include testing whether the ambipolar type transport persists at different applied external field strengths, F, and temperatures, T.[355]

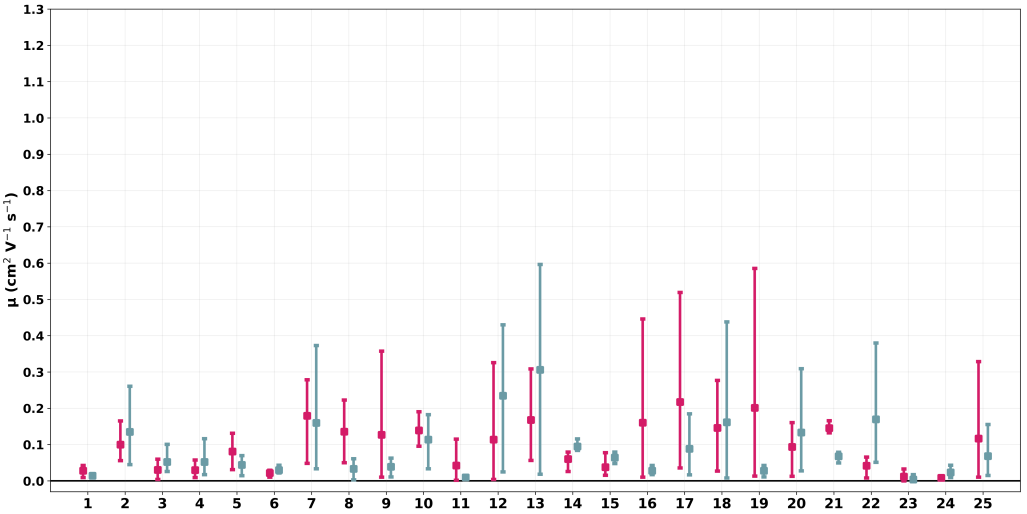
Among the 25 polymorphs for which the charge-carrier mobility was computed, 5 out of 7 highest hole-transporting polymorphs arrange in chiral columns. 4 out of 7 highest electron-transporting polymorphs show chiral columnar packing, 3 of which are 'sparse heterochiral double stranded columns', *e.g.*, two back-to-back helicenes stacked on top of each other at every level along the chiral column (structures 2, 7 and 18, highlighted in Figure 3.18). The experimentally observed crystal structures 1 and 11 ($P3_1$, chiral column and $P2_1$ respectively) are poor hole-transporting molecular materials. The highest hole transport was predicted for structure 19, il-



(a) [6]helicene



(b) 2,2'-dicyano-[6]helicene



(c) 2,2'-diethynyl-[6]helicene

Figure 3.21: Minimum, average and maximum mobilities per lowest-energy crystals. All hole (pink) and electron (blue) mobilities are drawn to the same scale.

molecule	Structure	J_{max} (meV)	chiral column?	μ_{max}	plane(μ_{max})	A1	A2
[6]H	2	69	- (T)	0.41	BC	0.62	8.55
	9	72	- (T)	0.41	AB	0.56	14.51
	10	72	- (T)	0.54	AC	0.65	18.13
	12	123	- (B)	0.43	BC	0.55	19.73
	20	71	- (T)	0.46	AB	0.52	13.99
	25	91	- (B)	0.41	BC	0.64	6.54
	23	*95*	*- (B)*	*0.17*	*AC*	*0.39*	*8.96*
2,2'-dicyano-[6]helicene	5	95	Y	0.49	AC	0.70	29.33
	7	58	-	0.39	BC	0.70	22.02
	8	55	-	0.32	AC	0.43	4.35
	10	58	-	0.32	AC	0.67	24.76
	17	58	-	0.32	BC	0.85	25.47
	18	54	-	0.32	BC	0.53	7.07
	19	68	Y	0.34	BC	0.57	18.79
	23	92	Y	0.37	BC	0.59	19.22
	24	75	-	0.37	BC	1.15	6.08
	1	*77*	*Y*	*0.26*	*BC*	*0.96*	*31.01*
	29	*118*	*Y*	*0.0*	-	-	-
2,2'-diethynyl-[6]helicene	7	124	sparse (Y)	0.37	BC	0.58	11.12
	12	67	Y	0.43	BC	0.55	17.46
	13	104	-	0.60	AB	0.56	32.22
	18	102	sparse (Y)	0.44	AC	0.91	54.22
	20	63	-	0.31	AC	0.73	11.12
	22	69	-	0.38	AB	0.70	7.37
	25	*109*	*Y*	*0.16*	*AB/AC*	*0.73*	*10.34*

Table 3.3: Table illustrating the highest electron mobilities obtained within the polymorphic landscapes of [6]helicene, 2,2'-dicyano- and 2,2'-diethynyl-[6]helicene. The structures are numbered according to their relative position in the DFT-XDM crystal-energy landscape. The 'chiral column?' column indicates, whether the crystal arranges in a chiral column, if so (Y), otherwise (-). For [6]helicene, which does not contain any chiral columnar packing, the motif giving rise to the J_{max} is given. Mobilities are given in $cm^2\ V^{-1}\ s^{-1}$ and for [6]helicene and the 2,2'-disubstituted [6]helicenes only crystals with maximum mobilities exceeding 0.40 and 0.30 $cm^2\ V^{-1}\ s^{-1}$ are shown. Additionally, for each molecule, 1-2 examples of high J_{max} and low maximum electron mobility are shown in italics at the bottom of each example. The plane of maximum transport and the respective anisotropies are shown. $A1$ is the crystal anisotropy computed using the normalised standard deviation of the angular-dependent mobility. $A2$ is a second anisotropy computed as the ratio between maximum and minimum angular-dependent mobility.

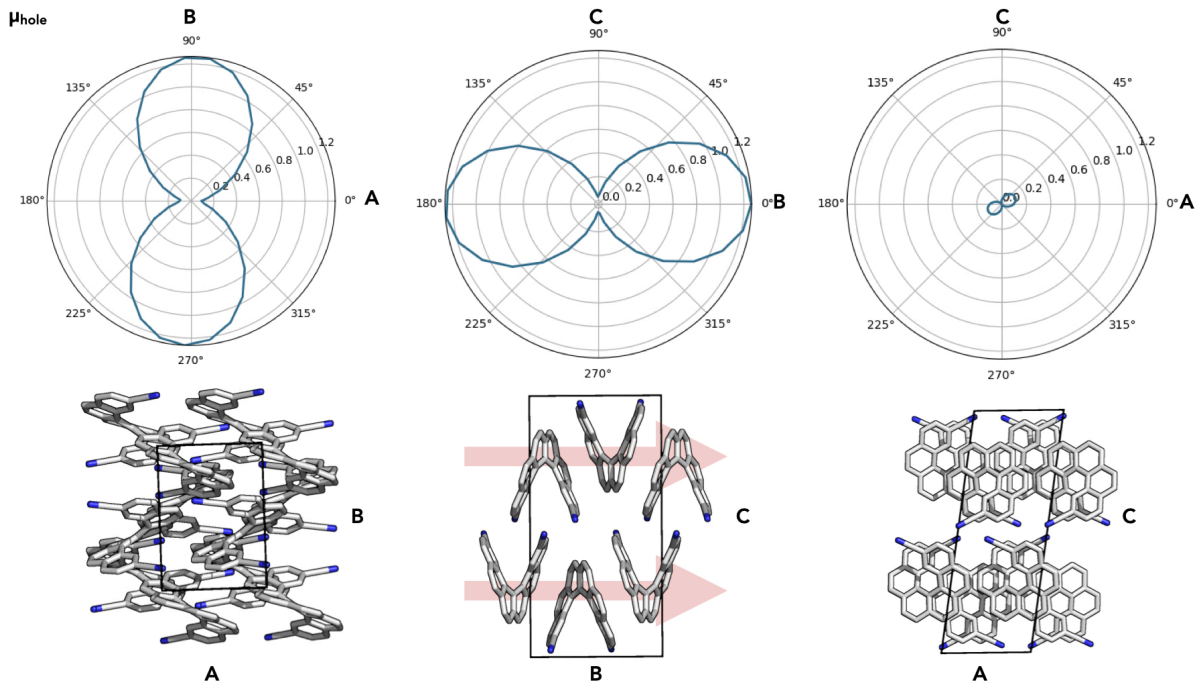


Figure 3.22: Polar plots of the highest hole-transporting 2,2'-dicyano-[6]helicene polymorph (structure 6) in $cm^2\ V^{-1}\ s^{-1}$. Charge transport is highest along the back-to-back dimers of the chiral column, as indicated by the arrow in the *bc*-plane. There is no transport along the *ac*-plane.

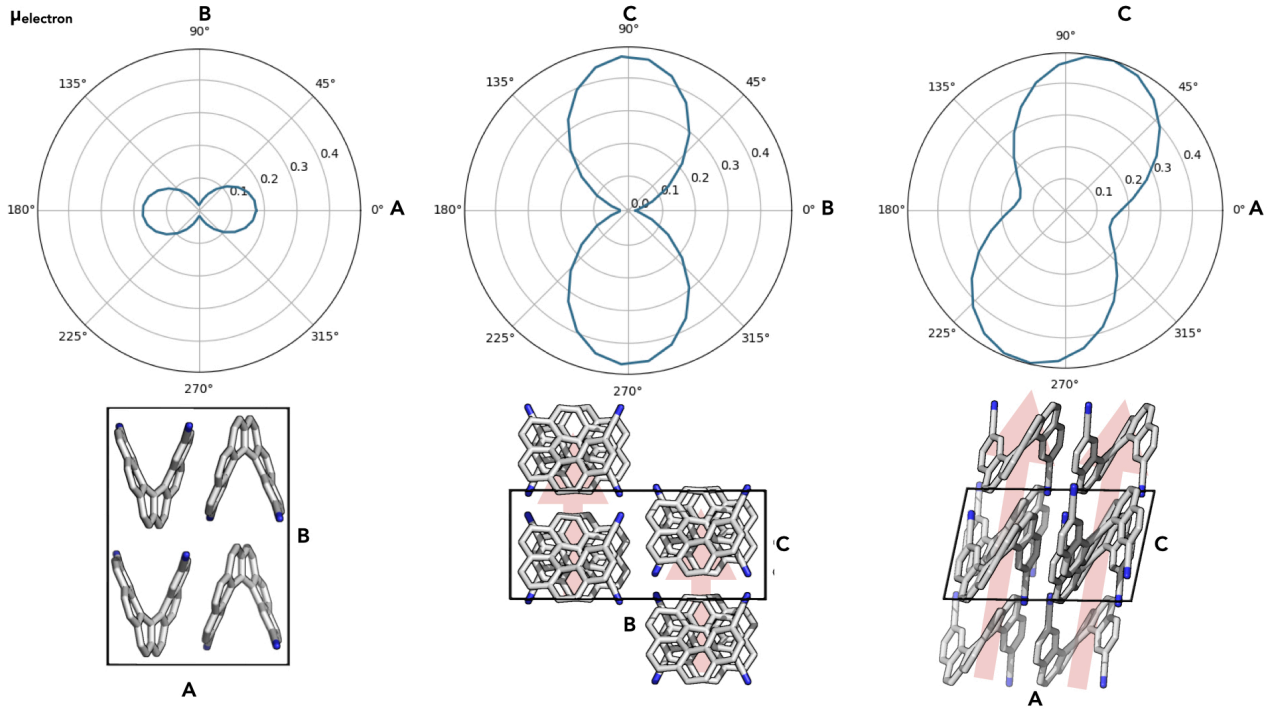


Figure 3.23: Polar plots of the highest electron-transporting 2,2'-dicyano-[6]helicene polymorph (structure 5) in $cm^2\ V^{-1}\ s^{-1}$. Charge transport is highest along the homochiral translational chain, as indicated by the arrows in the *ac*- and *bc*-plane.

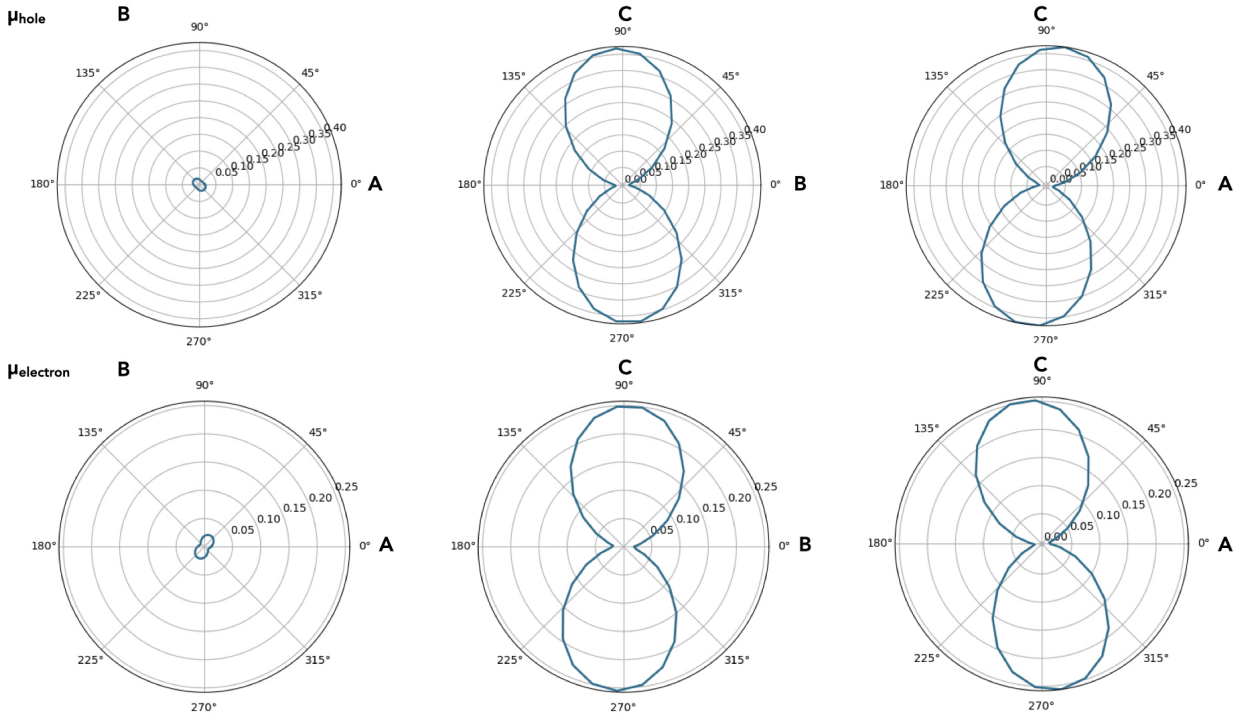


Figure 3.24: Polar plots of 2,2'-dicyano-[6]helicene structure 1 (UJOCOL) hole and electron mobilities in $cm^2\ V^{-1}\ s^{-1}$.

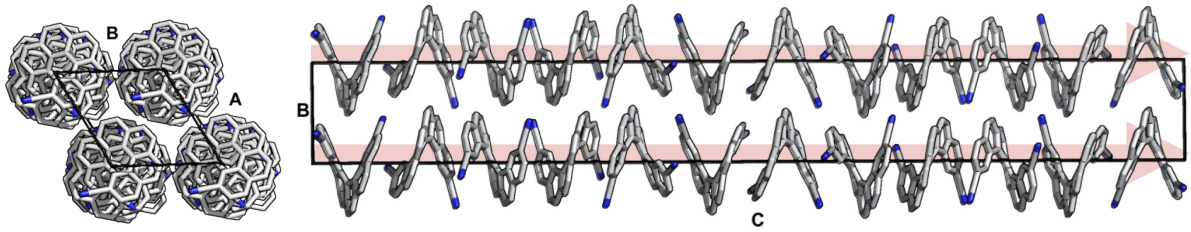


Figure 3.25: Structure 1 (CSD reference: UJOCOL) along the *ab*- (left) and *bc*-plane (right). There is no transport along the *ab*-plane, but along the chiral column in the *bc*- and *ac*-plane.

molecule	CSD reference	E_{rank}	J_{max}^{hole}	J_{max}^{elec}	chiral column?	μ_{max}^{hole}	μ_{max}^{elec}
carbo[6]H	HEXHEL	1	33	65	N	0.09	0.29
2,2'-dicyano[6]H	UJOCOL	1	105	77	Y	0.42	0.26
2,2'-diethynyl[6]H	NEMKUO	1	29	12	Y	0.04	0.02
2,2'-diethynyl[6]H	PICFEO01	11	6	10	N	0.12	0.01

Table 3.4: Maximum hole and electron mobilities for all yet experimentally observed crystal structures. The highest hole-transporting crystalline material is arranged in a chiral column, whilst the highest electron-transporting material is [6]helicene and shows no chiral columnar packing. Maximum transfer integrals J_{max} are shown in meV, charge-carrier mobilities in $cm^2\ V^{-1}\ s^{-1}$.

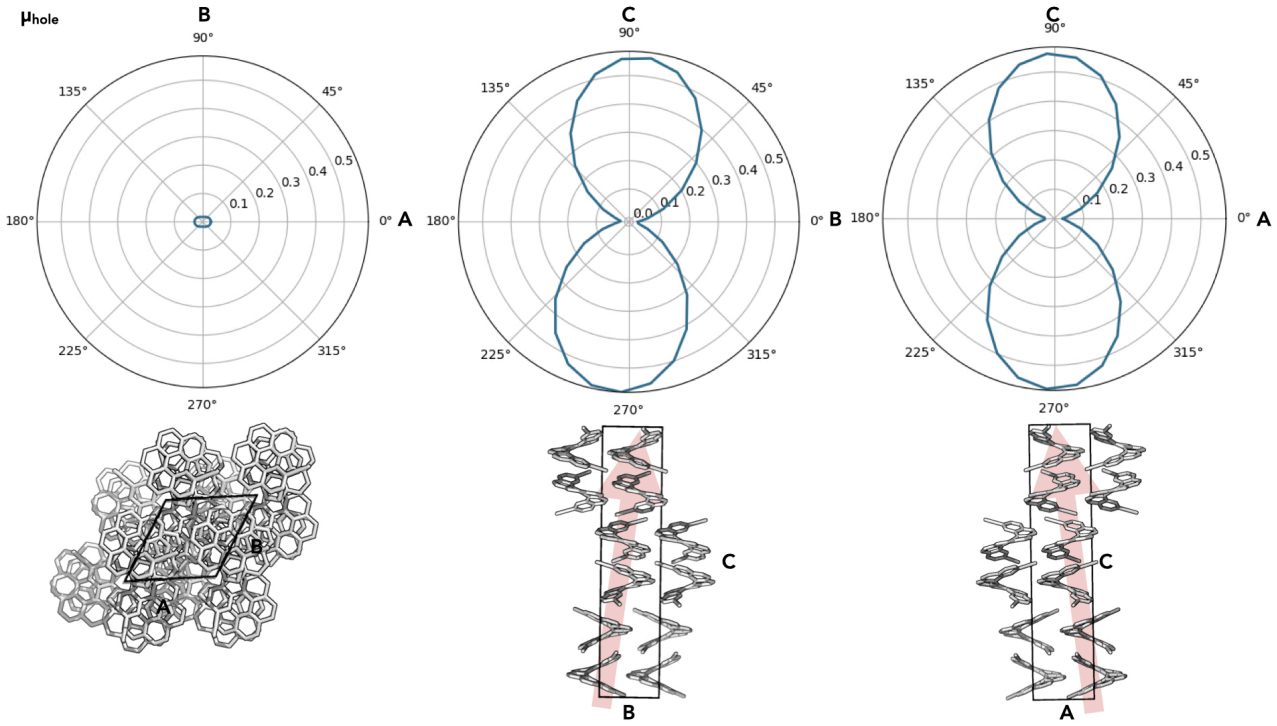


Figure 3.26: Polar plots of the highest hole-transporting 2,2'-diethynyl-[6]helicene polymorph (structure 19) in $cm^2\ V^{-1}\ s^{-1}$. Charge transport is highest along the (homo)chiral column, as indicated by the arrows in the *ac*- and *bc*-plane.

lustrated in Figure 3.26, with a predicted maximum hole mobility of 0.59 $cm^2\ V^{-1}\ s^{-1}$ along its (homo)chiral column. The chiral column helical pitch consists of 6 molecules for every 360° rotation. The angular-dependent electron mobility of the highest electron-transporting structure, 13, is depicted in Figure 3.27 and was not reported along a chiral column. Whilst chiral columns (and 'sparse heterochiral columns') are frequently observed along the highest transport planes and crystals, this packing does not guarantee a high hole (electron) mobility. In contrast, the experimentally observed crystal structure 1 (CSD reference: NEMKUO) arranges in chiral columns and does not exhibit high charge transport (Table 3.4).

In summary, [6]helicene shows slightly higher maximum electron mobilities than hole mobilities ($\mu_{max}^{elec} > \mu_{max}^{hole}$). In most polymorphic forms, both 2,2'-diethynyl and 2,2'-dicyano have similar μ_{max} to [6]helicene. 2,2'-dicyano has one outstanding structure with chiral columnar packing (μ_{max} = 1.26 $cm^2\ V^{-1}\ s^{-1}$). The best 2,2'-diethynyl structure also contains chiral columns but has a markedly lower mobility. Quantitative comparison of all charge-mobilities measured across each of the three landscapes shows that 2,2'-diethynyl-[6]helicene, with many polymorphs with columnar packing, has slightly improved average hole mobilities compared to the unsubstituted

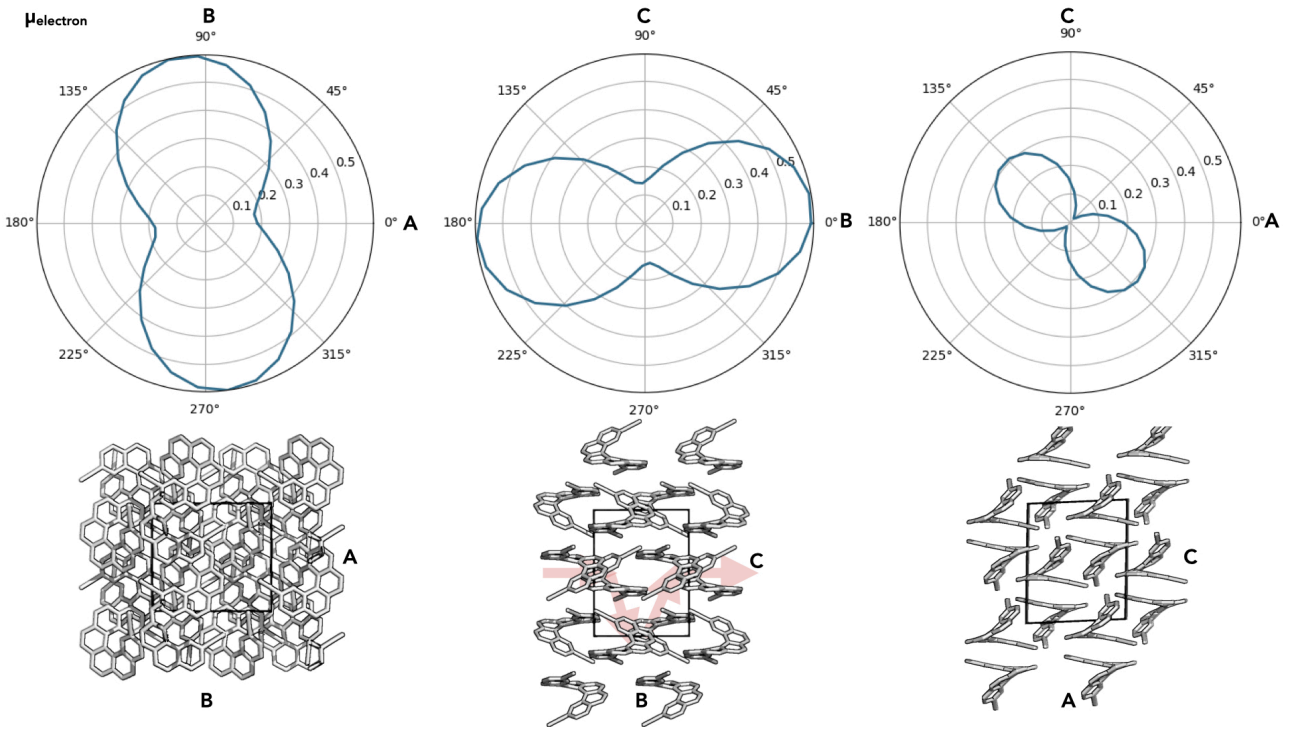


Figure 3.27: Polar plots of the highest electron-transporting 2,2'-diethynyl-[6]helicene polymorph (structure 13) in $cm^2\ V^{-1}\ s^{-1}$. Charge transport is highest along *b*-axis and along the back-to-back dimers as indicated by the charge-transport route in the *bc*-plane. The two outer arrows indicate the back-to-back dimer with highest charge transport (104 meV). Direct charge hopping across the interlocked dimer in between these highest-transporting back-to-back dimers is possible (15 meV), but decreased compared to the back-to-back dimers only route (23 meV) route.

[6]helicene reference, whilst 2,2'-dicyano-[6]helicene has a lower mean hole mobility. The unsubstituted [6]helicene reference has on average higher electron mobilities than both terminally difunctionalised molecules. Nonetheless, the highest electron mobility is measured across a 2,2'-diethynyl-[6]helicene polymorph with 0.60 $cm^2\ V^{-1}\ s^{-1}$. This polymorph is not arranged in a chiral column, but the best transport pathway is along back-to-back dimers. Whilst 2,2'-terminally difunctionalised [6]helicenes tend to more frequently arrange in chiral columns, chiral columnar packing is not a guarantee for high charge transport. In fact, both experimentally observed polymorphs with chiral columnar packing show poor electron transport properties. Neither of the two 2,2'-difunctionalised [6]helicene highest electron-transporting polymorphs arranged in chiral columns. However, the highest hole-transporting 2,2'-diethynyl-[6]helicene polymorph and the second highest hole-transporting 2,2'-dicyano-[6]helicene polymorph are both arranged in chiral columns. For the highest hole-transporting 2,2'-dicyano-[6]helicene polymorph, the molecules did not arrange in chiral columns, but the best charge transport route

was equally found across alternating back-to-back dimers, highly similar to a chiral column. The second highest 2,2'-diethynyl-[6]helicene hole mobilities were found for the experimentally observed crystal structure with long chiral columnar packing. Thus, whilst chiral columnar packing remains far from being a simple giveaway feature for high charge transport, homochiral columnar packing is more frequently observed in terminally functionalised [6]helicenes and if present, it can lead to significantly improved hole mobilities (*e.g.* 2,2'-dicyano-[6]helicene) than for the unsubstituted reference.

3.4 Conclusions and Future Work

For both 2,2'-disubstituted [6]helicenes, the experimentally observed enantiopure polymorphs arranged in homochiral columns and were successfully predicted as the lowest-lying polymorphs. The second experimentally observed 2,2'-diethynyl-[6]helicene polymorph, without homochiral columnar packing, was predicted to be the 11^{th} lowest-lying polymorph in the crystal-energy landscape. The experimental collaborators have successfully synthesised the lowest-lying 2,2'-diethynyl-[6]helicene crystal structures. However, targeted polymorph screening approaches are required to obtain more than one polymorph.The chiral columnar packing was found to be associated with characteristic intermolecular C··C interactions quantifiable in Hirschfeld plots. None of the experimentally observed 2,2'-diethynyl-[6]helicene structures proved particularly promising as a hole- or electron-transporting material. 2,2'-dicyano-[6]helicene has two polymorphs, both arranged in chiral columns (structures 1 and 6), which are suitable for hole transport. One of the two polymorphs is the experimentally observed structure (structure 1). Our experimental collaborators have synthesised both molecules, 2,2'-dicyano-[6]helicene and 2,2'-diethynyl-[6]helicene and single crystals of the material were grown. Further experimental analysis of the material and its chiroptical properties is forthcoming.

The crystal packing analysis using CRYSTACK, which was developed for this purpose, revealed that the interlocked dimer motif (with characteristic interactions between the terminal aromatic rings) tends to be present at higher density crystals, whilst low-density crystals are less likely to show the motif. The back-to-back dimer motif is generally most frequently observed. This is consistent with the observation that both 2,2'-difunctionalised molecules often arrange in chiral

columns. Future work on the CRYSTACK module could include capturing the axial chirality of the dimers: hetero vs. homochiral interlocked/back-to-back. Other motifs, such as the herringbone motif, could also added. There are many more motifs that in theory could be classified, because with every new [6]helicene crystal-energy landscape, new motifs are observed. Too many classes could end up hampering the usefulness of the classification, by providing a vast pool of motifs with not very distinctive features (*e.g.* subtleties of the back-to-back dimer motifs, *side* vs. *edge*) Thus, whilst this approach, once fitted to the group of molecules, provides an automised classification tool to analyse the crystalline packing across the polymorphic landscape, it lacks generalisability and there could be more informative solutions to the crystal packing analysis problem.

A graph-based similarity kernel-based approach has been developed to analyse the similarity of the solid-state arrangements observed for a given molecule. The clustering of crystal structure graph representations is efficient at picking up duplicate packings present in the crystal data, which were not identified by the COMPACK algorithm. In some cases, this was due to the fact that the two packings were identical, but the intermolecular distances between the molecules were slightly altered. However, if the differences in intermolecular distances are marginal, the two structures would converge to the same structure upon DFT-XDM relaxation. If the clustering was used prior to selecting the candidate crystal structures for DFT-XDM minimisation, it could help avoid costly repeated relaxations of essentially identical crystalline packings or having to visually check for duplicate structures. The presented approach clearly identified crystalline environments with a similar packing to have a high similarity score. Whilst the approach was successful at differentiating between crystalline environments with only π-π stacking and some with C-H$\cdots\pi$ stacking between layers, *i.e.* at picking up local features. The presented approach remains unable to identify more sophisticated and less localised aspects of the crystalline environment, such as chiral columnar packing. However, being able to identify clusters of high degree of parallel π-π stacking can help detect polymorphs with promising charge transport properties. The in-depth analysis of the charge-carrier mobilities across the polymorphic landscapes revealed that chiral columnar packing can give rise to good hole transport in some solid forms, yet it may not be the case for all.

Chapter 4

Computational Screening of Chiral Organic Semiconductors: Exploring Side-group Functionalisation and Assembly to Optimise Charge Transport

The work in this chapter has been published in the following contributions:

- **Schmidt, J. A.**, Weatherby, J. A., Sugden, I., Santana-Bonilla, A., Salerno, F., Fuchter, M., Johnson, E. R., Nelson, J. & Jelfs, K. E., Computational Screening of Organic Semiconductors: Exploring Side-Group Functionalization and Assembly to Optimize Charge Transport in Chiral Molecules, Crystal Growth & Design, 2021, Accepted Article. https://doi.org/10.1021/acs.cgd.1c00473.

- Peach, R. L., Arnaudon, A., **Schmidt, J. A.**, Palasciano, H. A., Bernier, N. R., Jelfs, K. E., Yaliraki, S. N. & Barahona, M., HCGA: Highly Comparative Graph Analysis for Network Phenotyping, Patterns, 2021, 2, 4, 100227. https://dx.doi.org/10.2139/ssrn.3720300

4.1 Introduction

The interdependence of molecular structure, crystalline packing, low-frequency vibrations and charge transport poses many challenges to organic electronic materials discovery. Therefore, developing a framework to design these materials would be of significant value to the advancement of the field. Crystal Structure Prediction (CSP) can be used in organic semiconductor (OSC) materials research to bridge the gap between molecular structure and solid-state arrangement. However, CSP is not computationally efficient enough to perform large high-throughput screenings, whereby a large number of candidate molecules are tested and poorly performing candidates are removed. Thus, explicitly considering the effect of solid-state arrangements in a high-throughput screening, as presented in the following, is only feasible with an approximate and simplified representation of crystalline packing. In the following chapter, a dimer-based screening approach for the search for high-performance organic semiconductors across ~1,300 side-group functionalised [6]helicenes is presented. The dimer-based screening model leverages the ubiquity of a particular translational motif of [6]helicenes. Since CSP is too computationally expensive to be included at the screening stage, yet serves as a reliable predictor for the set of lowest-lying crystal structures; CSP will be used in the refinement stage to assess the charge-transport across the most promising molecules obtained from the screening stage. The overarching goal of this study is to detect structure-property relationships that will accelerate the design of new chiral organic electronic materials with tailored properties. New chiral helicenes are generated and tested for their OSC suitability, because small-molecule chirality can provide additional functionality to the material over conventional achiral small-molecule OSCs (see Section 1.3.1.1).[178,192,207,356,357]

For the computational screening of hundreds of functionalised [6]helicenes by a dimer-based approach, a single homochiral translational motif was selected (Figure 4.1). This translational motif was the most abundant across the set of lowest-lying [6]helicene crystalline packing arrangements, it tends to occur in a plane orthogonal to the interlocked or back-to-back dimer motifs (Figure 4.1b). The translational dimer motif has been introduced in the previous chapter (Section 3.2.2.1). It is characterised by one molecule being translated with respect to the other along a vector perpendicular to the plane of the helix (Figure 4.1c). The translational motif was chosen based

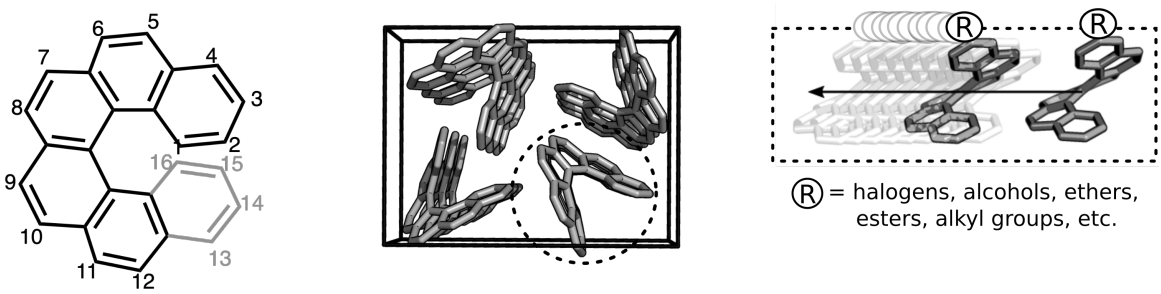


Figure 4.1: Figure introducing the homochiral translational motif: (a) the *M*-[6]helicene backbone, which will be mono- and symmetrically disubstituted with side-groups during the screening. (b) Front view of the homochiral translational motif (dashed circle) as observed in the experimental [6]helicene structure (CSD reference code: HEXHEL). The motif propagates going into the plane of the page. (c) Side view of the homochiral translational motif observed for a hypothetical side-group functionalised [6]helicene pair, here the particular case of monosubstitution in the 3-position is shown. Reprinted with permission from Reference 358. Copyright 2021 American Chemical Society.

on previous [6]helicenes research results:

- The previous chapter discussed what packing arrangements are needed in order to optimise the molecular material (*i.e.* the two 2,2'-difunctionalised [6]helicenes) for high *p*-type and *n*-type transport. Back-to-back dimers, as part of chiral columns, dominate the highest electron-transporting J_{max} dimers. However, the homochiral translational motif gave rise to the overall highest electron mobility across the lowest-lying 2,2'-dicyano-[6]helicene polymorphs.

- Previous analysis of the [6]helicene crystal-energy landscape by Rice *et al.* revealed the homochiral translational chain to be the highest electron-transporting carrier path across all polymorphs.[197] Furthermore, Rice *et al.* found the homochiral translational motif to be the most abundant motif across all motifs identified and observed in 63% of hypothetical lowest-lying polymorphs.

- More recently, the translational motif was found to give rise to the highest electron transfer integrals across the majority of aza[6]helicenes isomers.[312]

Finally, to investigate whether this packing motif also dominates in solid-state arrangements of terminally substituted [6]helicenes (beyond the molecules presented in Chapter 3), a CSD search

is presented as part of this chapter, which assesses the suitability of this motif for a screening-based approach.

Based on the homochiral translational motif (hereafter translational motif), a computational screening of ~1,300 terminally substituted [6]helicenes toward optimised charge transport and molecular assembly for *n*-type (electron) transport is being presented. The focus on *n*-type transport was chosen because the translational motif was consistently found to be associated with high *n*-type transport over *p*-type transport.[197,312] Besides, *n*-type OSC performance is less studied and still lags behind *p*-type OSC performance.[359] The screening of hundreds of differently functionalised [6]helicenes for good OSC properties provides insights into the following aspects of terminal functionalisation and molecular assembly: (i) the lowest-energy separation of each screened translational dimer motif will highlight to what extent the particular side-group functionalisation/molecular geometry affects the molecule's tendency to adapt the motif; (ii) the magnitude of the transfer integral, J, and its inner-sphere reorganisation energy, λ_i, will serve as indicators for the molecule's charge-transport properties; (iii) testing each side group for many different substitution positions (and patterns) sheds light on the effect of identifying the best functionalisation position(s) on charge transport. A better understanding of the impact of specific substitution patterns on charge transport is particularly useful to synthetic chemists. As the synthesis of functionalised (and particularly asymmetrically functionalised) helicenes is a time-consuming effort, by knowing exactly which substitution pattern to target the synthesis of functionalised helicenes with bespoke properties can be accelerated.

4.1.1 Chapter Outline

The following chapter, will be divided into two main parts: (i) The first part consists of the translational dimer-based screening of ~1300 new terminally functionalised helicene molecules to discover molecules suitable as high-performance OSCs (ii) The second part contains a full CSP analysis and periodical crystal mobilities of the most promising side-group functionalised [6]helicenes. The second part aims to confirm whether the lowest-lying polymorphs obtained from CSP actually contain the translational motif and whether crystal mobilities of the lowest-lying polymorphs can outperform the [6]helicene reference. Thus, the OSC performance of all candidate molecules

will be primarily compared to the [6]helicene crystal-energy landscape and electronic properties. The periodic DFT-XDM crystal energy calculations of the tentative polymorphs obtained from CSP were all performed by Joseph Weatherby, PhD student in the Johnson group at Dalhousie University.

4.2 Methods

The methods are structured according to the general chapter outline. Thus, at first all methods concerning the dimer-based computational screening part (i) are provided. Then, all methods used for the CSP and crystal charge-carrier mobility part (ii) are given. The methods given as part of the computational screening (i) further include: a CSD search for the frequency of the translational motif, a derivation of the mobility in the low-field limit and regression techniques to predict the inner-sphere reorganisation energies. The derivation of the mobility in the low-field limit was performed by Prof Jenny Nelson (Imperial College London).

4.2.1 Compound Screening

4.2.1 CSD Search

A CONQUEST[360] search was performed using the two-dimensional chemical diagram of the [6]helicene backbone without any hydrogens attached. 201 crystal structures were found (04.10.2019). Some structures were subsequently discarded for various reasons, *e.g.* (i) the crystal structure was not sufficiently well resolved; (ii) the structure was too different to resemble a conventional helicene molecular crystal; (iii) the helicene was a ligand to a metal-centre. Crystal structures which were omitted because of (ii) were mostly graphene-type sheets with the helix twist lost, helicenes with large parts of fused rings attached to ether end (*e.g.* "superhelicenes"). Helicenes of axial chain length other than six and crystals with bridged helicene dimers were omitted. Crystal structures containing unsubstituted [6]helicenes and/or [6]helicenes as solvent molecules were excluded from the set. Some substituted [6]helicene crystals may contain traces of solvent molecules. Finally, a total of 87 terminally substituted [6]helicenes, out of which 29 contained the translational dimer (33%, Table B.1), were obtained. The relative proportion with which each

recurring structural motif was observed is shown in Figure 4.2.

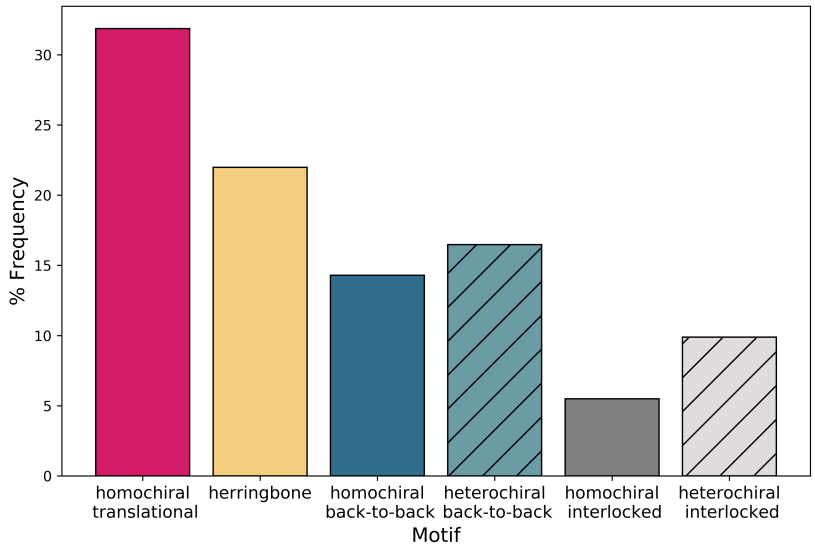


Figure 4.2: The relative abundance of the most common recurring structural [6]helicene dimer motifs across the terminally substituted [6]helicenes obtained from the CSD. The motifs are shown in Figure 1.5. The hatched black bars indicate heterochiral motifs (racemic crystals) compared to homochiral motifs (enantiopure/racemic crystals). The homochiral translational dimer motif is most frequently observed. Reprinted with permission from Reference 358. Copyright 2021 American Chemical Society.

4.2.1 Compound Screening Approach

The *M*-[6]helicene backbone, as obtained from the HEXHEL [6]helicene crystal structure, was either mono-substituted (positions 1-16) or symmetrically disubstituted (2,2'-, 3,3'-, 4,4'-, 5,5'-, 6,6'-, 7,7'-, 8,8'-) to investigate the effect of terminal side-group functionalisation on *n*-type transport. The functional groups tested range from alcohols, halogens, esters, ethers, acids to alkyl and aromatic side groups. The side-groups screened for are illustrated in Figures 4.3 and 4.4. Side-groups significantly larger or of similar size to the [6]helicene backbone were not tested because the dimer-based screening requires the intermolecular interactions of the [6]helicene backbone to dominate the molecular assembly for the translational motif to be viable. For each of the 56 side-groups selected, 24 substitution patterns were tested (16 monosubstitutions and 8 disubstitutions), resulting in a total of 1344 (24×56) functionalised [6]helicenes considered in the screening.

The side-group substitutions of the [6]helicene backbone were performed via addition of the

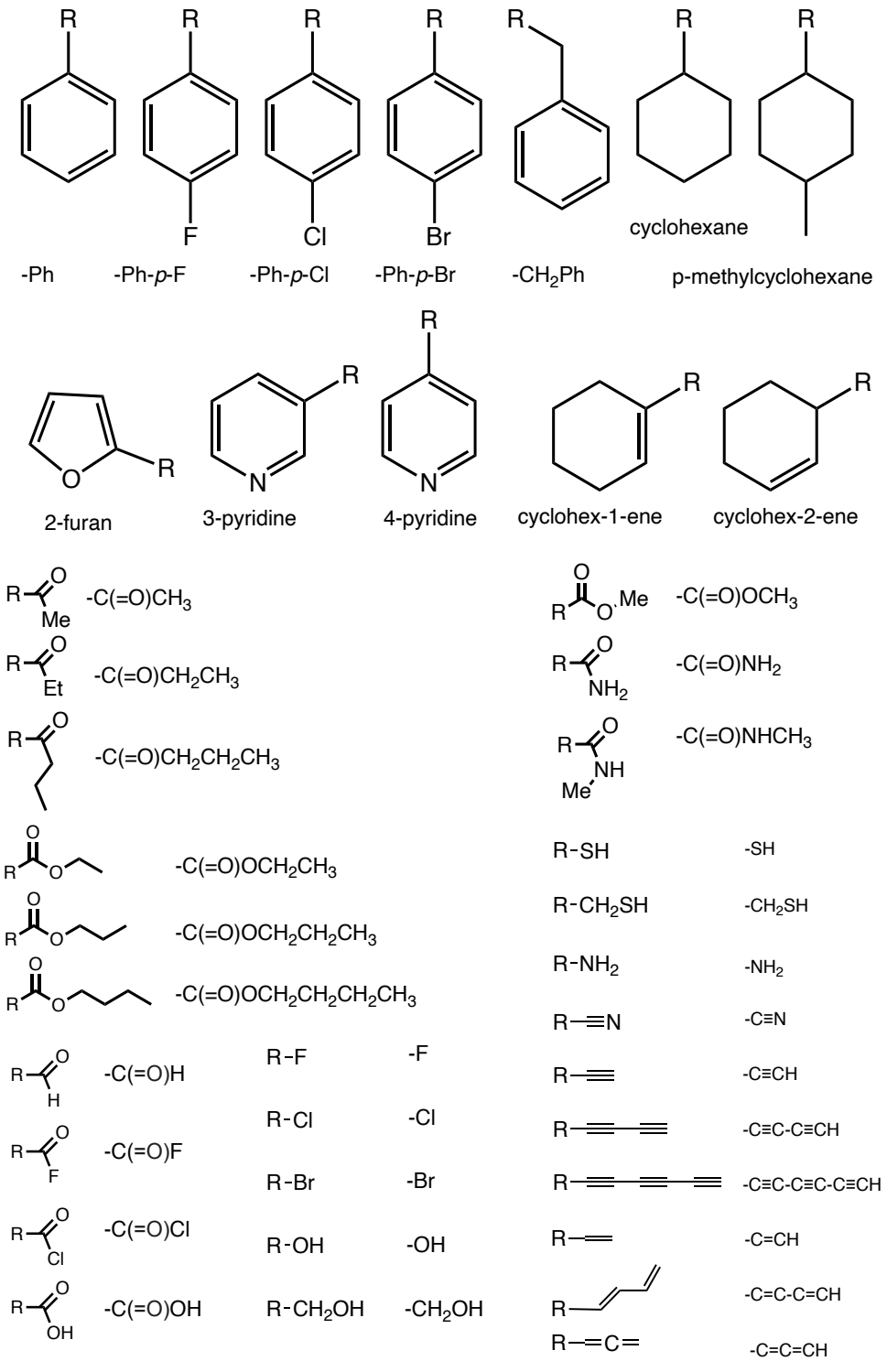


Figure 4.3: The functional groups considered in the computational screening (part I). Reprinted with permission from Reference 358. Copyright 2021 American Chemical Society.

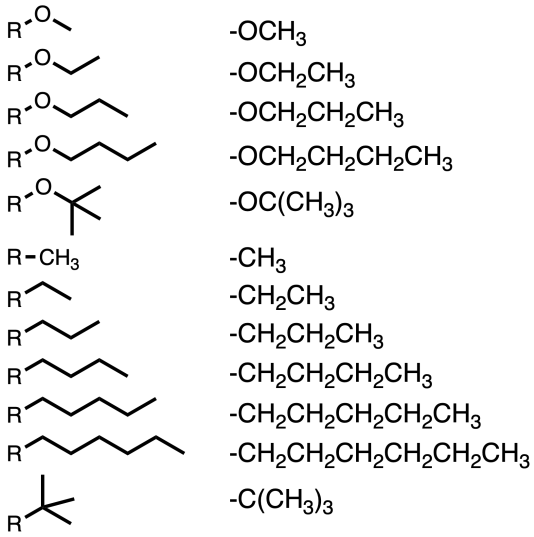


Figure 4.4: The functional groups considered in the computational screening (part II). Reprinted with permission from Reference 358. Copyright 2021 American Chemical Society.

SMARTS representation to Lukas Turcani's `chem_tools` substitution code.[a]. Each substituted [6]helicene molecule was force field optimised using UFF[361] in a python open-source toolkit for cheminformatics, RDkit.[289] Then, a short molecular dynamics (MD) simulation at 700 K was performed using OPLS3e[362] force field as implemented in the Maestro Schrödinger software (timestep = 1.0 fs, maximum iterations = 2500, number of conformers sampled = 50) in *stk*[363] to confirm whether the force field optimised geometry was also the lowest-energy conformer and not a local minimum. This step reduced the computation time required for DFT geometry optimisation at B3LYP/6-31g(d) level of theory using GAUSSIAN16.[278] The default tight SCF convergence criteria, which use the EDIIS+ CDIIS[364] algorithm is employed. The geometry optimisation convergence criteria were a maximum RMS change in density matrix and RMS change in forces smaller than a threshold of 10^{-6} Ha/Bohr.

Once the geometry optimised substituted [6]helicene was obtained, the molecule was superimposed with the [6]helicene in the asymmetric unit of the HEXHEL crystal structure using RDkit[289] and the translational dimer was extracted. For each translational dimer, the GFN-xTB[270] lowest-energy separation was computed via a scan from 6 Å to 14 Å, in 0.1 Å increments. For each lowest-energy distance translational dimer, the target property, *i.e.* the coupling integral J was predicted. For an overview of the generalised workflow for the screening procedure, refer to Figure 4.5. Prior to the screening, the accuracy of dimer separations predicted by the GFN-xTB scans and

[a] `https://github.com/lukasturcani/chem_tools`, last accessed 08.04.2020

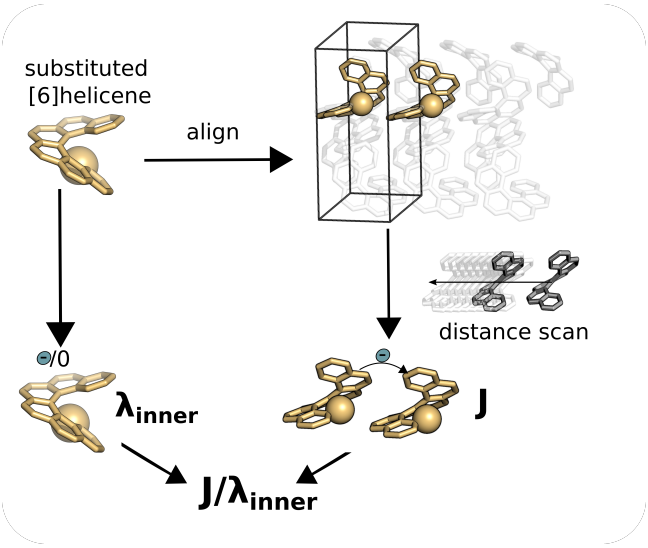


Figure 4.5: Overview of the screening procedure. A side group of choice is attached to the helicene backbone (top left) and the molecule is geometry optimised. The functionalised [6]helicene is then mapped back onto the unit cell, with the translational dimer shown along the c-axis, in the plane of the page. An intermolecular distance scan for this dimer is performed by modifying the c-axis length. The transfer integral (J) is computed at the minimum-energy distance for the dimer, while the inner reorganisation energy (λ_i) is computed from only the optimised single-molecule geometries for the neutral and anionic states. The J/λ_i ratio is calculated as a measure for OSC suitability. Reprinted with permission from Reference 358. Copyright 2021 American Chemical Society.

OPLS3e scans was tested against DLPNO-CCSD(T).[365]

Shortcomings of this workflow include three common breaking points (i) the RDkit alignment of the functionalised helicene with the reference inside the HEXHEL unit cell fails and (ii) the intermolecular distance scan does not find a minimum-energy distance between 6-14 Å, indicative of the functionalisation being incompatible with the translational dimer motif and (iii) some λ_i calculations only converged at larger basis set sizes. Others did not converge at higher maximum convergence cycles and basis set sizes (*e.g.* *o*-, *m*-, *p*-phenol). Where convergence was more computationally expensive, the inner-sphere reorganisation energy λ_i could be predicted using regression techniques (see Section 4.2.1.6). Furthermore, 18 molecules were discarded from the screening, because iv) the GFN-xTB energy-distance scan resulted in a minimum-energy dimer separation at covalent bonding distances. Thus, a dimer at such short covalent bonding distances would essentially not be a molecular material anymore. Theoretically, a total of 1344 minimum-energy translational dimers and their J and λ_i could have been computed. Due to the shortcomings described previously, only 1245 data points contained both screening metrics (J and λ_i). For

1301 molecules the transfer integral could be computed, the remaining 43 failed due to reasons i), ii) and iv).

4.2.1 DLPNO-CCSD(T) Benchmark Intermolecular Distance Scans

Prior to any screening, the semi-empirical GFN-xTB[270] and the OPLS3e[362] force field were tested against domain based local pair natural orbital-coupled-cluster single double (DLPNO-CCSD).[366,367] "Domain based LPNO-CCSD" enables computations as accurate as LPNO-CCSD[366] with unprecedented efficiency for large molecules.[367] The DLPNO-CCSD(T)/def2-TZVPP method with the RIJCOSX approximation, the computationally costly grid6 and the default auxiliary basis sets was used as implemented in the ORCA software package.[365] DLPNO-CCSD(T) benchmark scans were performed for three benchmark molecules: [6]helicene, 2,2'-dicyano-[6]helicene and 2,2'-diethynyl-[6]helicene. In addition to the intermolecular distance scans, the variation of transfer integral, J, near the minimum-energy distance was computed. The results are shown in Figures 4.6-4.8. Based on these benchmark scans, GFN-xTB was selected for screening, because it is computationally cheap and the lowest-energy separations, d_{min}, at DLPNO-CCSD(T) and GFN-xTB level consistently varied by 0.1 Å, encouraging the use of GFN-xTB. The OPLS3e lowest-energy distances were consistently overestimated by 0.1 Å for [6]helicene and 0.3 Å for the two 2,2'-disubstituted [6]helicenes. In 'transient localisation theory', variations of transfer integral J and mobility μ are often linked to changes to the π-π and long-axis intermolecular distances and the corresponding transient localisation lengths.[154,155] The hopping model employed in this work does not take these fluctuations into account. With the separation dependence of J, the error estimates arising from incorrect minimum-energy predictions can be quantified.

4.2.1 Property Prediction

The transfer integral J and inner-sphere reorganisation energy λ_i calculation procedures were performed exactly as described in Section 2.8.1 and 2.8.3. The higher the coupling integral J and the lower the molecule's inner-sphere reorganisation energy λ_i, the higher the OSC suitability of the candidate molecule. At first, both features will be discussed separately. Then, the inter- and intramolecular aspects to charge transport will be combined into a single parameter, the J/λ_i ra-

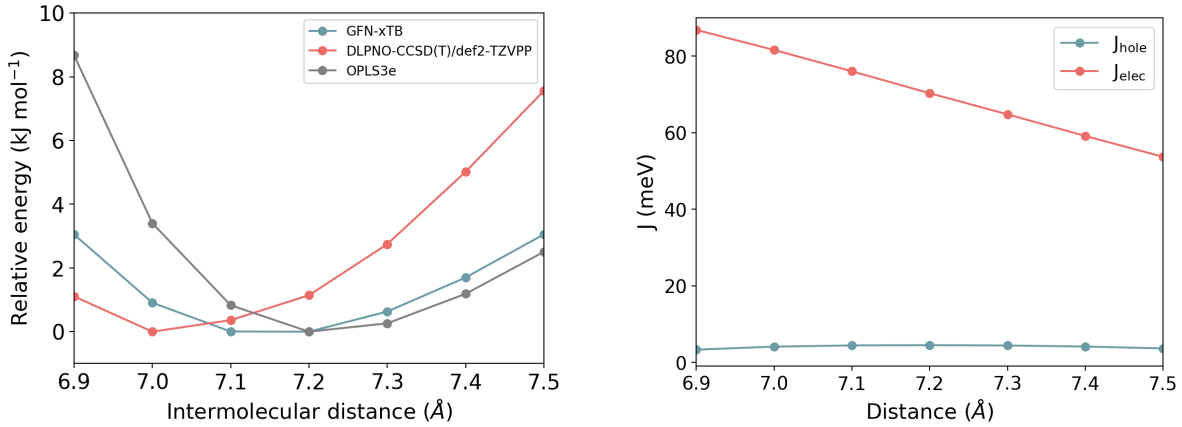


Figure 4.6: [6]helicene translational dimer separation scan (left) and the transfer integral computed at each intermolecular separation (right). Reprinted with permission from Reference 358. Copyright 2021 American Chemical Society.

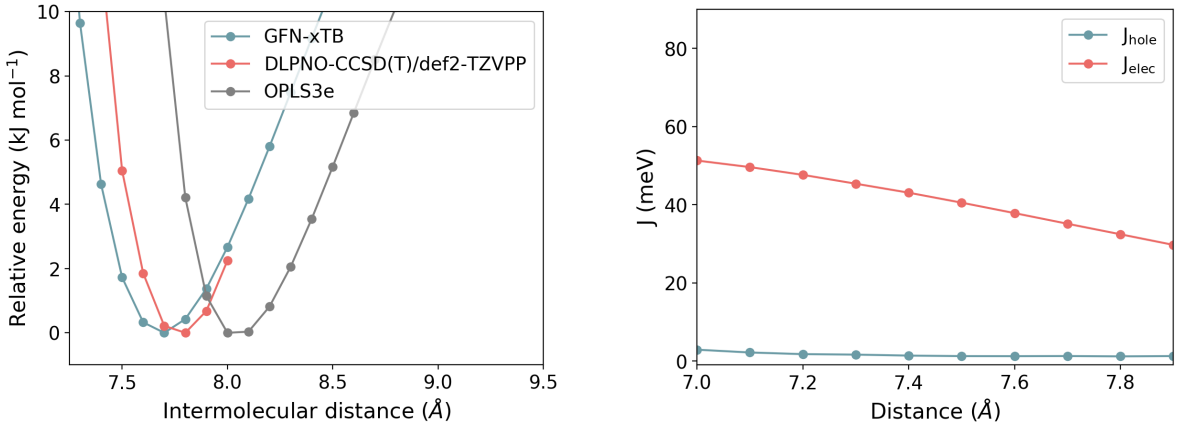


Figure 4.7: 2,2'-dicyano-[6]helicene translational dimer separation scan (left) and the transfer integral computed at each separation (right). The *y*-axes are drawn to the same scale as in Figure 4.6 for direct comparison. Reprinted with permission from Reference 358. Copyright 2021 American Chemical Society.

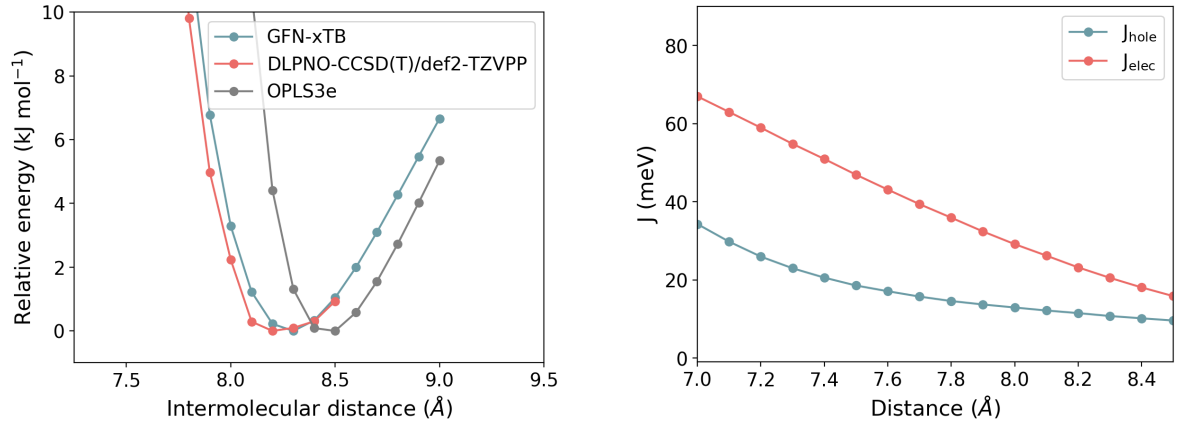


Figure 4.8: 2,2'-diethynyl-[6]helicene translational dimer separation scan (left) and the transfer integral computed at each intermolecular separation (right). The *y*-axes are drawn to the same scale as in Figure 4.6 for direct comparison. Reprinted with permission from Reference 358. Copyright 2021 American Chemical Society.

tio (hereafter J/λ) to quantify the relative trends and ease identification of promising candidate molecules. The J/λ ratio is most frequently used as the limiting condition of when the hopping regime breaks down.[128] The larger the J/λ ratio, the more suitable for charge transport but if the ratio is too high, hopping theory would become invalid. In addition to the J/λ ratio, a second screening metric, the electron mobility μ from non-adiabatic Marcus theory in the low-field limit was computed. For both screening metrics J/λ_i and μ, the outer-sphere reorganisation energy λ_o was disregarded ($\lambda = \lambda_{\mathrm{i}}^{\mathrm{elec}}$). Hence, both screening metrics will be overestimates of the true values. Thus, one cannot expect absolute values to be correct, yet the screening metrics should help identify relative trends between systems.

While the assumption of weak-coupling (hopping regime) may not be valid in *every* case, it is valid for relatively low J and hence, low J/λ. All screening metrics are compared to the [6]helicene reference, which, based on charge-carrier mobilities computed for all low-lying polymorphs, falls into the hopping regime.[197] In other words, the screening should be good at eliminating unpromising structures, and systems with unexceptionally high screening metrics, which may not be adequately described by the hopping regime, will be highlighted. Beyond the validity of the hopping regime, high J values remain a necessary condition for high charge mobility.

4.2.1 Charge-carrier Mobilities in the Low-field Limit

The derivation presented in the following was derived by Prof Jenny Nelson and reproduced here with her permission. The charge-carrier mobility derived as a screening metric is similar to the crystalline charge-carrier mobility described in Section 2.8. Here, a periodic one-dimensional (1-D) adjacent homochiral translational chain of helicenes is assumed, *i.e.* helicenes are equally spaced along the axis that defines the translational motif. The charge-carrier mobility along the 1-D chain can be predicted as an approximate value representative of charge transport in the bulk material. Hence, a periodic one-dimensional (1-D) adjacent homochiral translational chain of helicenes is assumed; helicenes are equally spaced along the axis that defines the translational motif. The molecular spacing a, λ_i and J are properties directly obtained from the screening, and can be used to compute the mobility in the low-field limit when $eaF << k_BT$, where e is the elementary charge, k_B the Boltzmann constant, T the temperature, and F the applied electric field. In

the low-field limit and at low temperature, the Marcus expression for non-adiabatic charge transfer rate (Eq. 2.25) reduces to Eq. 4.1 for transfer along the direction of the field, Γ_+, and transfer against the direction of the field, Γ_-. Again, a is the lattice constant along the field direction.

$$\Gamma_+ = \frac{|J|^2}{\hbar}\sqrt{\frac{\pi}{\lambda k_B T}} e^{\frac{-\lambda}{4k_B T}} (1 + \frac{eaF}{2k_B T}) \tag{4.1}$$
$$\Gamma_- = \frac{|J|^2}{\hbar}\sqrt{\frac{\pi}{\lambda k_B T}} e^{\frac{-\lambda}{4k_B T}} (1 - \frac{eaF}{2k_B T})$$

Based on Eq. 4.1 and assuming the occupation probability is the same at each site, the net charge flux in the direction of the field can be expressed using the charge transfer rates in the forward and reverse directions. The net electron velocity, $\bar{v}$, can then be defined by the lattice constant a and the difference between rates of the carrier moving along and against the applied external electric field F as

$$\bar{v} = a[\Gamma_+ - \Gamma_-] = \frac{|J|^2}{\hbar}\sqrt{\frac{\pi}{\lambda k_B T}} e^{\frac{-\lambda}{4k_B T}} (\frac{ea^2 F}{k_B T}). \tag{4.2}$$

Because $\mu = \frac{\bar{v}}{F}$, the electron mobility μ in the limit of low field (for charge-transport along a regular 1-D array of sites spaced by a), simplifies to

$$\mu = \frac{|J|^2}{\hbar}\sqrt{\frac{\pi}{\lambda k_B T}} \left(\frac{ea^2}{k_B T}\right) e^{-\frac{\lambda}{4k_B T}}. \tag{4.3}$$

This expression for the electron mobility, μ, is independent of the exact low electric field (F) applied.

4.2.1 Inner-sphere Reorganisation Energy Data Augmentation Using Regression Techniques

For 4.3% of the λ_i computed for 1,344 molecules, the single-point energy calculations required to compute λ_i did not reach convergence and these molecules were excluded from the results data set. In order to ensure that no promising candidate would be omitted, the 'missing' λ_i were predicted using machine-learning regression techniques. The prediction of λ_i was tested using eX-

	R^2	**MAE** (meV)	**RMSE**
KRR	0.90	10.61	21.08
RF	0.85	18.14	26.23

Table 4.1: Comparison of XGBoost regression vs. kernel ridge regression (KRR). The Mean Absolute Error (MAE) is given in meV. Reprinted with permission from Reference 358. Copyright 2021 American Chemical Society.

treme Gradient Boosting, XGBoost,[368] and kernel ridge regression, KRR,[369] based on the molecule data converted into Morgan radial fingerprints.[370] Kernel ridge regression performed slightly better than the gradient boosting, with R^2=0.90 on the test set and a mean squared error (MSE) of 10.61 meV (Table 4.1).

Inner-sphere reorganisation energies, λ_i, have been previously predicted based on Morgan Circular Fingerprints using KRR (R^2 = 0.88) and deep neural networks (DNN, R^2 = 0.92).[170] The predictive power of the model was improved with more training data, *i.e.* the total number of λ_i values was increased from 1257 screening values to 1533 λ_i values. Each computationally generated [6]helicene molecule was converted into a Morgan fingerprint (FP) representation, *i.e.* a vector of certain bit-length containing connectivity information within a given radius, but no molecular geometry. An FP radius of 4 with 1024 bits length was used. The data set was cleaned from the methylperoxy side-group, because the 15 λ_i computed for this type of functionalised [6]helicenes yielded $\lambda_i > 2500$ meV. The remaining λ_i did not exceed 550 meV. Subsequently, the data set was split into a train and test set using a 80:20 split (1226:307). The KRR hyper parameters were selected based on a parallelised 10-fold grid search performing 4000 fits (computation time of 93 mins) on the training set. A regularisation parameter of 0.001 (α) and the Exp-Sine-Squared kernel (Eq. 2.21) gave rise to the best fit. The Exp-Sine-Squared kernel models periodic functions, which can vary by their length-scale parameter, l, and periodicity p. The best parameters were thus found to be $l = 4.64$ and $p = 21.5$. Similarly the hyper parameters for regression via XGBoost,[368] a decision tree based algorithm, were obtained via a 10-fold grid search on the training set. The hyper parameters were tuned to 100 trees with a maximum depth of 4. The minimum loss reduction γ and learning rate η were set to 0.3 and 0.1 respectively. The minimum sum of instance weight required in a child was 4. As shown in Table 4.1, KRR performed slightly better than XGBoost. Thus, all λ_i were predicted using KRR.

4.2.2 CSP Analysis of Screening Candidates

4.2.2 Crystal Structure Prediction on Top Candidates

The CRYSTALPREDICTOR search details for the rigid molecules 6-triacetylene-[6]helicene, 4-fluoro-[6]helicene and 4,13-difluoro-[6]helicene are as detailed in Section 3.2.1. However, at the DMACRYS stage, the charge density computed using distributed multipoles at RHF/6-31G(d,p) level was combined with FIT potentials.[371] Additionally, for the large 6-triacetylene side group, the default cut-off of 15 Å was extended to 30 Å to account for larger intermolecular distances.

For the CSP of configurationally flexible 7,10-dialdehyde-[6]helicene, 4 possible potential energetic minima were analysed. The 4 conformers differ by the possible orientations of the aldehyde side-groups, *i.e.* both oxygens pointing up, towards each other, away from each other, or both down. The *Z*-matrix to represent the molecule was defined such that the orientation of the aldehyde hydrogen would be dependent on the position of the aldehyde oxygen (see scheme in Figure 4.9). Based on the *Z*-matrix, local approximate models (LAMs) are being generated. LAMs are used to efficiently and accurately compute the intramolecular energy as a function of flexible torsion angles. Each LAM was geometry optimised using GAUSSIAN16 at PBEPBE[252]/6-311G(d,p) level of theory, tight convergence criteria, assuming no symmetry and with force constants being computed at the last step. Similarly, the Hu, Lu and Yang charge fitting method was used to obtain atomic charges for every geometry optimised LAM.[372] According to Figure 4.9, multiple stable minima exist requiring the need for a flexible CSP search. The CSP search was performed with two flexible degrees of freedom (θ_1, θ_2). All remaining search criteria space group, polymorphic range, *etc.* were used as stated in Section 3.2.1. CrystalOptimizer was used for flexible energy minimisations of the hypothetical crystal structures from the global search.

All tentative crystal structures within 15 kJ mol^{-1} were energetically re-ranked by single-point energy calculations of the force-field optimised structure, employing density-functional theory (DFT) with exchange-hole dipole (XDM) dispersion correction. The XDM dispersion correction has been previously proven accurate for the relative energy rankings of chiral molecular crystals.[63,373] All DFT-XDM calculations were performed by Joseph Weatherby (Dalhousie University) and used their in-house version of the SIESTA code (4.0b-485),[374] the B86bPBE

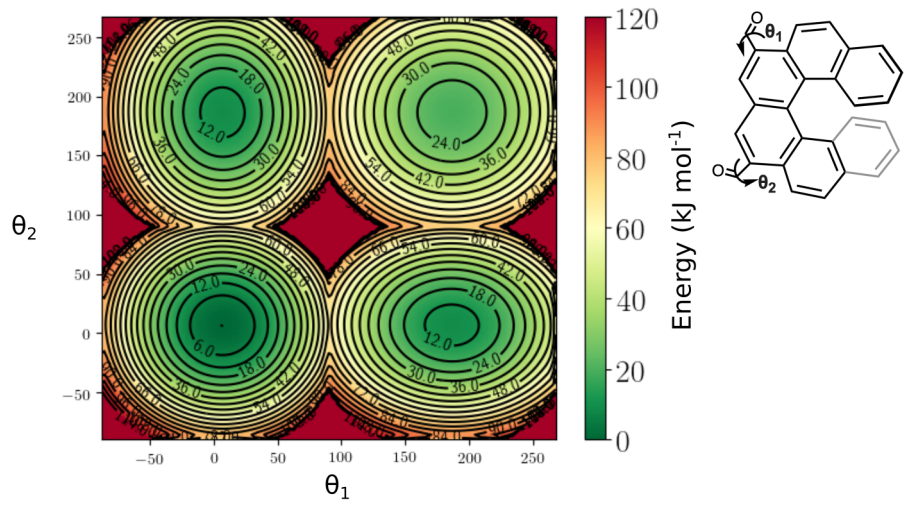


Figure 4.9: The 7,10-dialdehyde-[6]helicene contour plot of 2D potential energy surface (PES) of the dihedral angle scan (2° increments) with respect to two dihedral angles (θ_1, θ_2). Reprinted with permission from Reference 358. Copyright 2021 American Chemical Society.

density functional,[375,376] and the XDM dispersion correction, with the damping parameters set to a_1 and a_2 set to 0.5000 and 2.5556 Å respectively.[374] Troullier-Martins[377] norm-conserving pseudopotentials[378,379] were used together with a double-zeta plus polarization (DZP) basis set of atom-centered numerical orbitals. The confinement radius set to 0.001 Ry.[374,380] A 4 × 4 × 4 Monkhorst-Pack[381] k-point sampling of the Brillioun zone and a real-space integration grid cut-off value of 400 Ry were used. Periodic crystal relaxations at DFT-XDM level were omitted, due to their high computational cost and time constraints. Further studies based on interesting polymorphs identified should contain full relaxations.

4.2.2 Crystal Charge-carrier Mobilities

The molecular crystal charge-carrier mobilities were computed for the 10 lowest-lying polymorphs in each crystal-energy landscape according to the general methodology for the prediction of charge-carrier mobility in molecular crystals using semi-classical Marcus theory outlined in Section 2.8. The computed electronic inner-sphere reorganisation energy λ_i used to calculate the Marcus rates for 4,13-difluoro-, 13-fluoro-, 6-triacetylene- and 7,10-dialdehyde-[6]helicene were 0.1743 eV, 0.1704 eV, 0.2265 eV and 0.2365 eV respectively.

4.3 Results and Discussion

4.3.1 Screening of Side-group Functionalised Molecules for Maximal Transfer Integrals

This section discusses the results obtained from the compound screening of 1344 helicene molecules for their OSC suitability. Aspects of the computational screening procedure developed for this purpose will also be discussed. The key transport screening parameters, λ_i and J, are initially discussed individually and then as a combined score (J/λ and low-field mobility). The computational screening data is publicly available for download at https://data.hpc.imperial.ac.uk/resolve/?doi=7858.

4.3.1 Separation Dependence of Transfer Integrals

Figure 4.10 illustrates the exponential separation dependent decay of transfer integrals. The distance dependance helps identify trends of particularly suitable side-group functionalisations independent of their molecular size. For some side-group functionalised helicenes with unsaturated C-C bonds, the transfer integrals are higher in magnitude than predicted by the exponential decay (yellow points). Across disubstitution, the fluorine and aldehyde side groups gave rise to the respective shortest and second-shortest separations at GFN-xTB level of theory (Figure 4.11). Helicenes with this type of side-group functionalisation also show high transfer integrals (78 and 191 meV respectively, Table 4.2). In the case of monosubstitution, the side-group functionalisation with the generally closest packing for the translational dimer was chlorination, which is listed among the top-ranked molecules across the screening metrics discussed in the following. The large number of side-group functionalisations resulting in transfer integrals near zero highlights the fact that the chemical space one has to screen is extremely vast and only a high-throughput computational screening approach with an approximate representation of the crystalline environment is capable of screening this extremely vast chemical space. Across mono- and disubstituted [6]helicenes, fluorinated helicenes crop up at short separations at the top of the exponential decay distribution (Figure 4.10). This might be the reason why fluorination gives rise to excellent coupling integrals. Fluorination has been widely used in crystal engineering to reduce the inter-

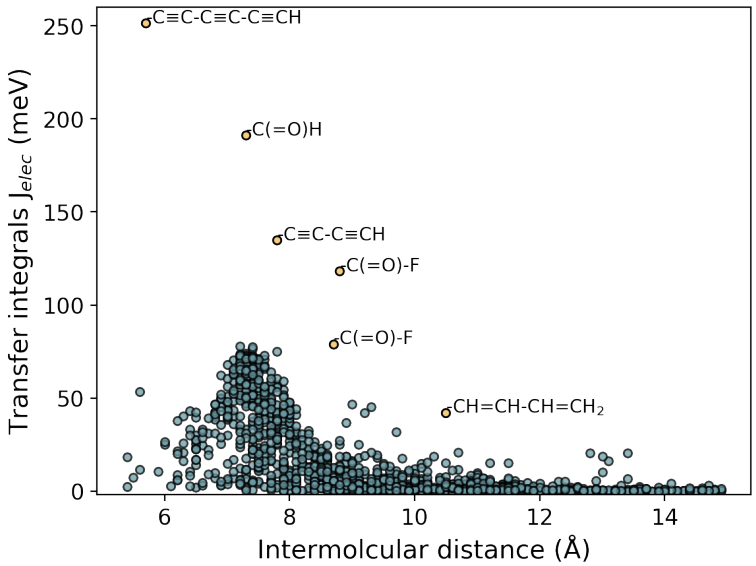


Figure 4.10: Distribution of transfer integrals (J) obtained at different minimum energy separations (d_{min}) within the translational dimer model for both mono- and disubstituted helicenes.

molecular spacing in electronic materials.[97,382]

4.3.1 Can terminal substitution improve transfer integrals?

A distribution of transfer integrals grouped by side-groups is given in Figure 4.12. Comparing the transfer integrals, the fluorine side group performs best across all side groups screened for and across mono- and disubstituted [6]helicenes. However, none of the mean or median transfer integrals obtained per side group exceed the unsubstituted reference. This might be due to any side-group functionalisation only enhancing charge transport in some substitution positions and not across all substitution patterns. This observation is a consequence of the translational motif, which has some substitution positions which are unfavourable for charge transport independent of the side group attached. For these positions, terminal side-group functionalisation results in increased intermolecular separations and a loss of π–π-overlap of the [6]helicene backbones. Thus, whilst there is no side group improving charge transport across all substitution patterns, some side-group functionalisations reach transfer integrals surpassing the unsubstituted reference. Hence, one can synthetically target a specific substitution pattern and side group that is likely to surpass the [6]helicene reference with respect to OSC performance. Whilst ideally, one would want to find a side group which consistently outperforms the unsubstituted reference molecule using this screening procedure. This side group could then be attached anywhere

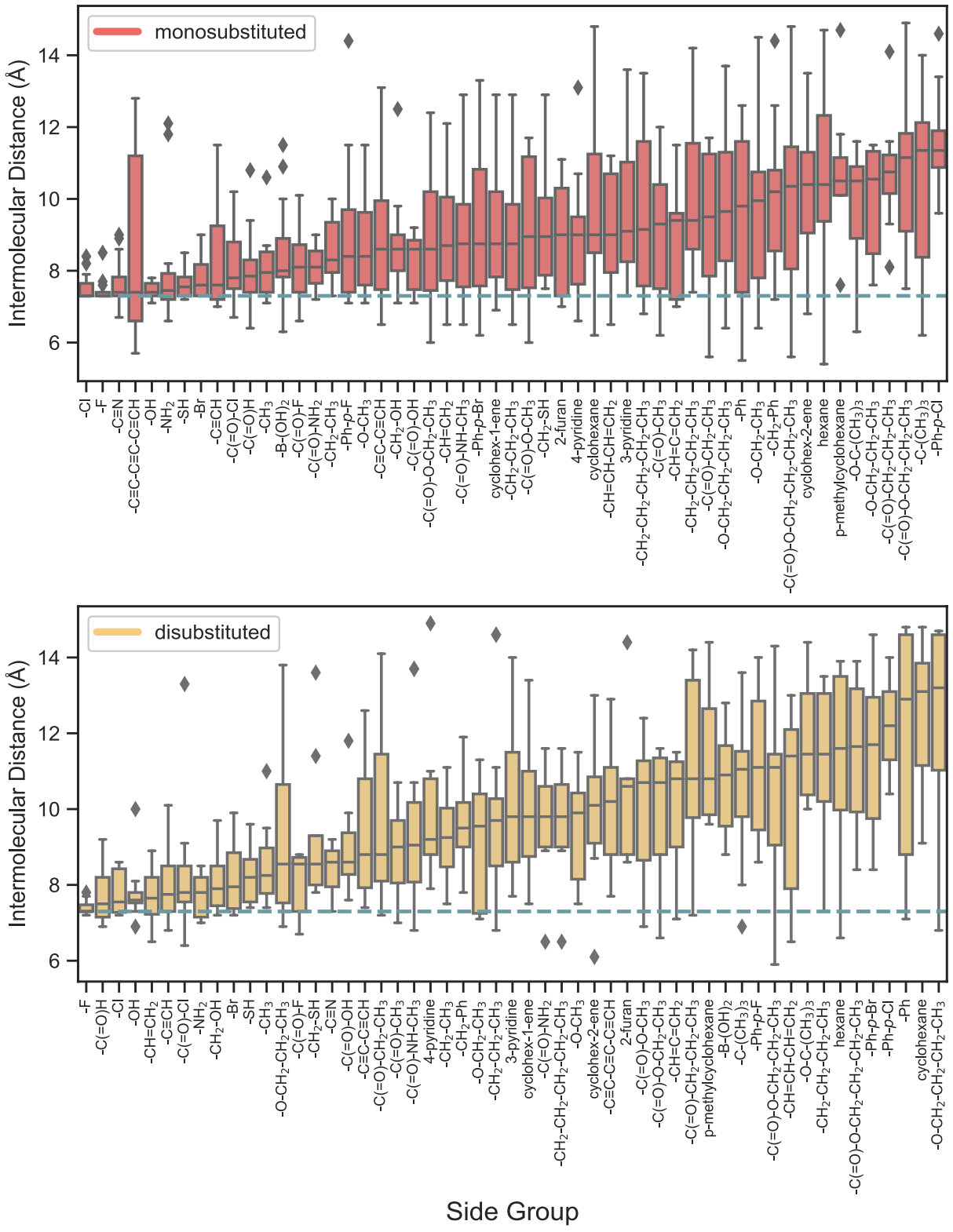


Figure 4.11: Boxplot showing the minimum energy intermolecular distances found in the translational dimer model. The dashed line is the intermolecular distance of the unsubstituted [6]helicene backbone. The size of the box is the interquartile range, containing 50% of the ratios, and the horizontal line inside each box is the median value. The whiskers indicate the range of ratios obtained outside the middle 50% of data points, with the exception of some large outliers, shown as grey diamonds. Reprinted with permission from Reference 358. Copyright 2021 American Chemical Society.

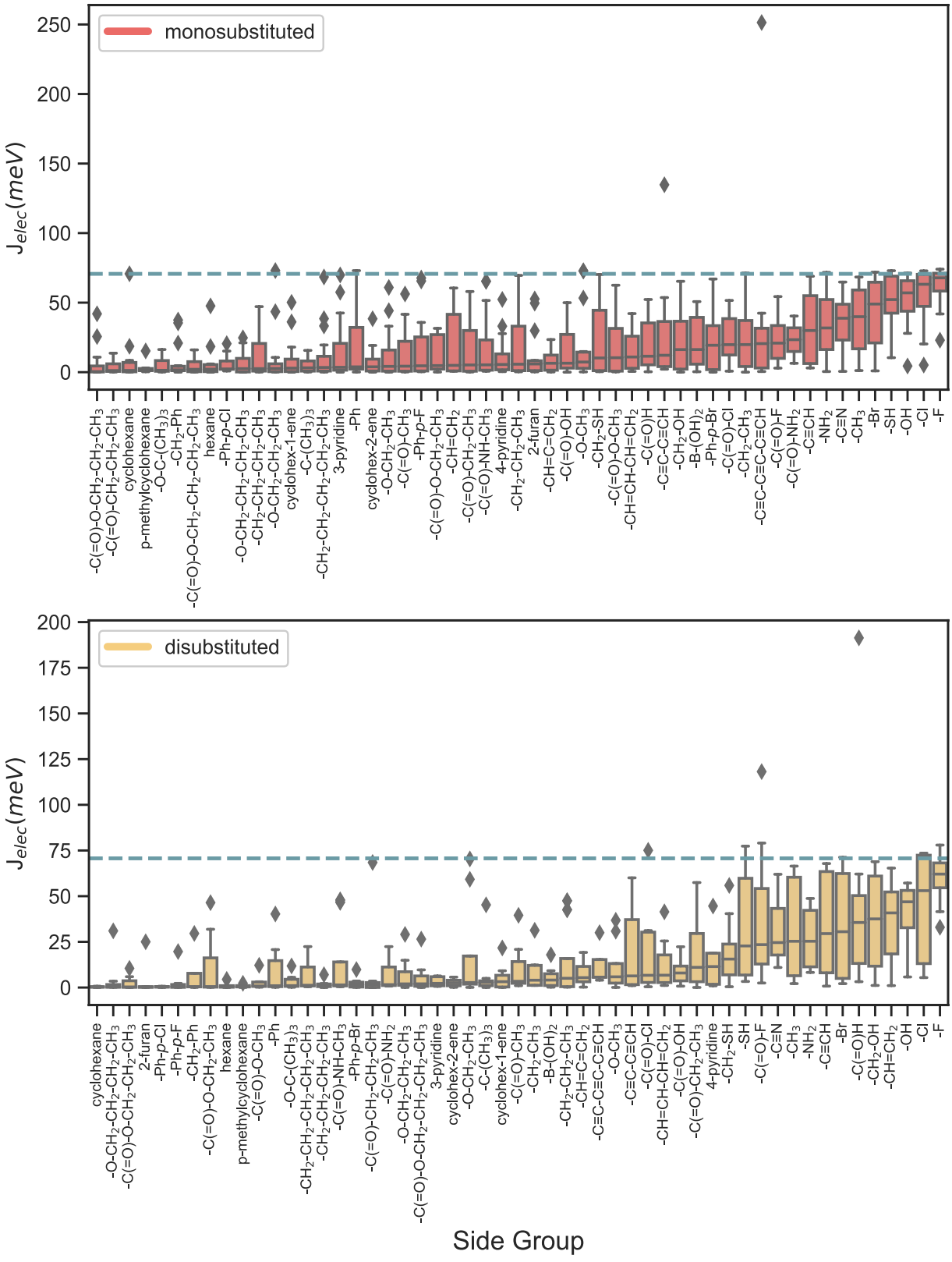


Figure 4.12: Boxplot of the electronic transfer integrals J_{elec} per side group computed across the translational dimer model and sorted by median. The size of the box is the interquartile range, containing 50% of the ratios, and the horizontal line inside each box is the median value. The whiskers indicate the range of ratios obtained outside the middle 50% of data points, with the exception of some large outliers, shown as grey diamonds. Reprinted with permission from Reference 358. Copyright 2021 American Chemical Society.

to the helicene backbone, at varying helicene length or to other OSC backbones (*e.g.* pentacene, *etc.*). However, it is much more likely to identify side groups, which enhance transport under some circumstances: a specific substitution position relative to the π-system, a specific recurring crystalline packing arrangement. Therefore, besides the side-group performance comparisons, one has to pay attention to specific outliers that outperform the reference.

The closest packed translational dimer as predicted by the GFN-xTB intermolecular distance scan was the 6-triacetylene-[6]helicene with a separation of d_{min} = 5.7 Å. The substitution in the 6-position does not impede π-π overlap of the [6]helicene backbone and provokes a very high transfer integral of J_{elec} = 250 meV, raising the question: is this molecule more adequately described in the transient localisation regime?[383] In the 6-triacetylene functionalised translational dimer, the side groups are aligned perfectly parallel and at short intermolecular distances, which allows for good π-π stacking across the side groups (Figure 4.13). In the 6-triacetylene dimer the lowest unoccupied molecular orbital (LUMO) is mainly localised on the side-groups of the two molecules and only on half of the [6]helicene backbones. The LUMO is clearly delocalised over the two molecules via their π-conjugated side groups, allowing for good electron transport. For the 7-diacetylene-[6]helicene molecule a very high transfer integral (J_{elec}=134 meV, d_{min}= 7.3 Å) is observed uniquely for the 7-position and not observed for any other monosubstituted position. The high transfer integrals observed for triacetylene or diacetylene side groups suggest that, if placed in the adequate position (6-position for triacetylene, 7-position for diacetylene), these side groups can enhance charge hopping via perfectly π-π stacked side groups and helicene backbones. Given that the ideal substitution position varies with side group, a computational screening such as the one presented here is necessary.



Figure 4.13: The LUMO frontier orbital for 6-triacetylene-[6]helicene in the ground state (isosurface 0.02 au). The LUMO was computed based on the lowest-energy separation of the dimer. Reprinted with permission from Reference 358. Copyright 2021 American Chemical Society.

4.3.2 Screening of Side-group Functionalised Molecules for Minimal Inner-sphere Reorganisation Energies

In addition to screening for high transfer integrals J, low inner-sphere reorganisation energies, λ_i, are desirable for high charge transport and hence were also included in the screening. λ_i is a transport property, which describes how easily a molecule accepts or donates an electron without a high energetic barrier required for the geometry change between the charged and neutral geometry and *vice versa*. Owing to the fact that the computational compound screening is targeting side-group functionalisations to the identical [6]helicene backbone, any trend in λ_i can be directly attributed to the presence of the side-group attached.

Some inner-sphere reorganisation energy λ_i calculations (4.3%) are more computationally complex than others and only reach convergence at higher basis set sizes and increased computational effort. The set of computed λ_i was used to predict λ_i on a test set with 90% (85%) accuracy using KRR (XGBoost). This suggests that potentially all inner-sphere reorganisation energies could have been predicted using ML. Inclusion of the additional ML-predicted λ_i does not affect the relative side-group ranking with respect to λ_i, nor with respect to the J/λ screening metric, which will be further discussed in the next section. Thus, prediction of the pending reorganisation energies does not result in any additional promising candidate molecules than the ones conventionally computed. As a result, one can confidently exclude these examples from the following discussions.

Comparison of the median and mean λ_i obtained depending on mono- and disubstitution of a certain side group gives rise to structure-property relationships, *i.e.* having tested a side group in all substitution positions and patterns allows for a generalised conclusion on whether the side-group functionalisation impedes or promotes charge transport by raising or lowering λ_i relative to the [6]helicene reference. This side-group comparison of λ_i is provided in Figure 4.14. The monosubstitutions do not result in a side group with a median λ_i lower than the unsubstituted reference. Across the disubstitutions screened for, there are four side-group functionalisations that stand out, due to them being the only ones with a median λ_i lower than the unsubstituted reference. The four disubstituted side groups are triacetylene, diacetylene cyano and ethynyl. All

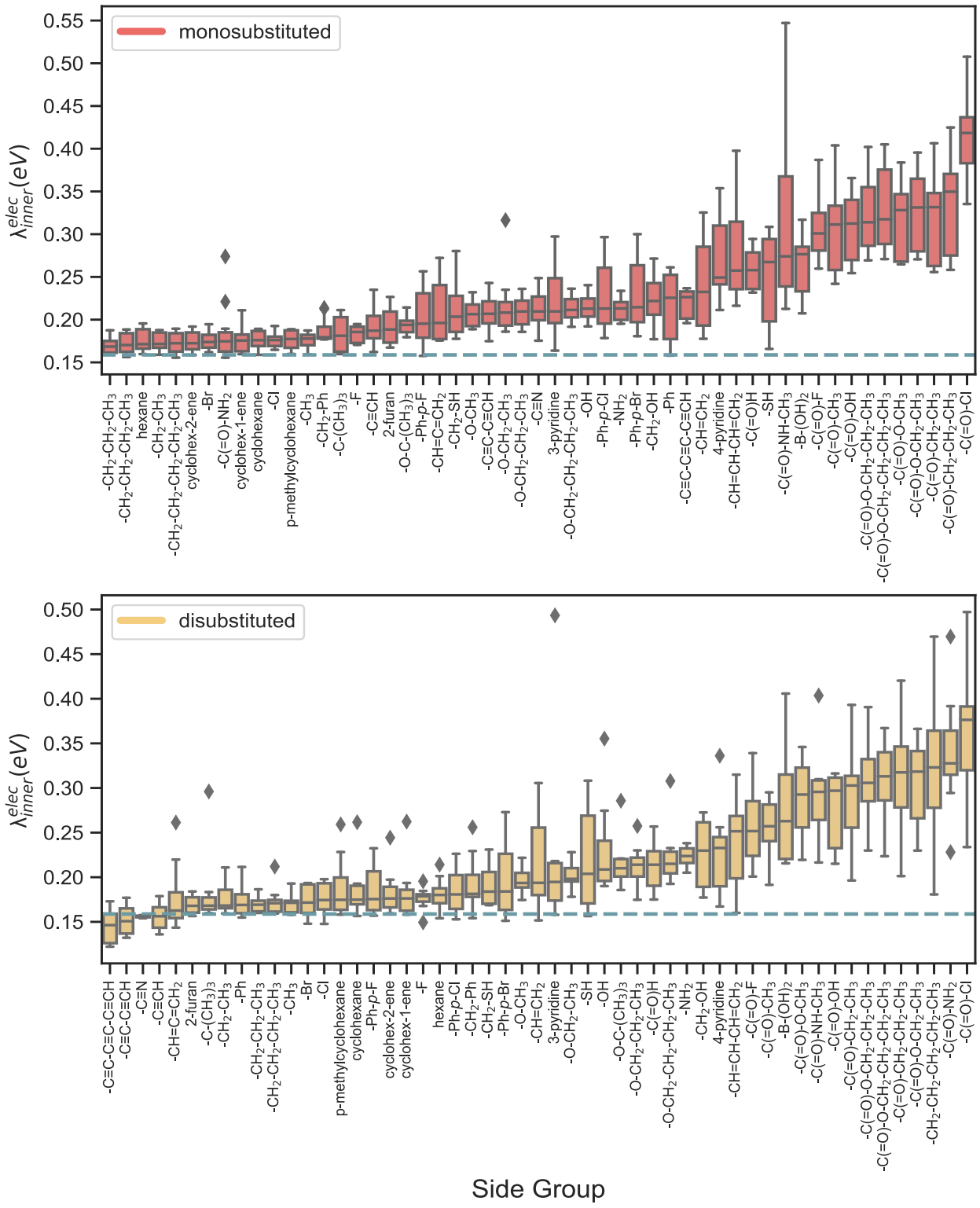


Figure 4.14: Boxplot showing the trends of inner electronic reorganisation energies λ_i^{elec} per side group. The dashed line is the inner electronic reorganisation energy of the unsubstituted [6]helicene backbone. This is the raw data prior to any learning on the missing reorganisation energies. The size of the box is the interquartile range, containing 50% of the ratios, and the horizontal line inside each box is the median value. The whiskers indicate the range of ratios obtained outside the middle 50% of data points, with the exception of some large outliers, shown as grey diamonds. Reprinted with permission from Reference 358. Copyright 2021 American Chemical Society.

promising electron-accepting side groups contain triple bonds. The π-system of the side-group delocalises over both side-group and backbone, thereby augmenting the total π-system size.

Figure 4.15 shows that for symmetrical disubstitution, λ_i decreases with increasing π-system size, particularly with the number of triple bonds (TB), *i.e.* λ_i decays in the order: ethynyl (1TB, 34 π electrons) > diacetylene (2TB, 42 π electrons) > triacetylene (3TB, 50 π electrons). The observation presented in Figure 4.15 is in agreement with experimental measurements of λ, which found λ to be inversely proportional to the inverse participation ratio (IPR) of the electron wavefunction that evaluates charge delocalisation.[325] Uniquely, the diphenyl- and the trans,trans-1,3-butadiene-substituted side groups disagree with this observation. However, these side groups are also not linear, and depending on the orientation of the side group relative to the backbone in the optimised gas phase geometry of the molecule, the two π-systems might wind up disconnected from each other.

The inverse relationship between π-system size and λ_i could not be witnessed in the case of monosubstitution. This effect may not be noticeable for monosubstitution because the increase in π-system size for monosubstitution is not as big as for disubstitution. Another reason why this is not observed for monosubstitution might be that the electron density is not as delocalised in the case of asymmetric monosubstitution as for symmetric disubstitution.

4.3.3 Identifying OSC Suitability Trends from Screening

With the goal of designing functionalised [6]helicenes with improved OSC suitability in mind, the transport properties (λ_i, J) collected during the screening of 1344 mono- and disubstituted [6]helicene molecules were expressed as two distinct OSC screening criteria: (i) the J/λ_i (hereafter J/λ) ratio and (ii) the charge-carrier mobility in the low-field limit (μ). For high-performance OSC molecules, both OSC criteria (J/λ, μ) would be maximal. Thus, the transfer integral J will need to be maximised and the inner-sphere reorganisation energy λ_i minimised. However, if the J/λ ratio becomes too high, hopping theory would become invalid. Uniquely, the highest transport integral (J = 251 meV) observed for the 6-triacetylene-[6]helicene molecule raised the question of whether it would fall outside the hopping regime. However, one has to keep in mind that all J/λ ratios are overestimates, because they only contain the inner-sphere reorganisation energy, *i.e.* $\lambda = \lambda_i$, and

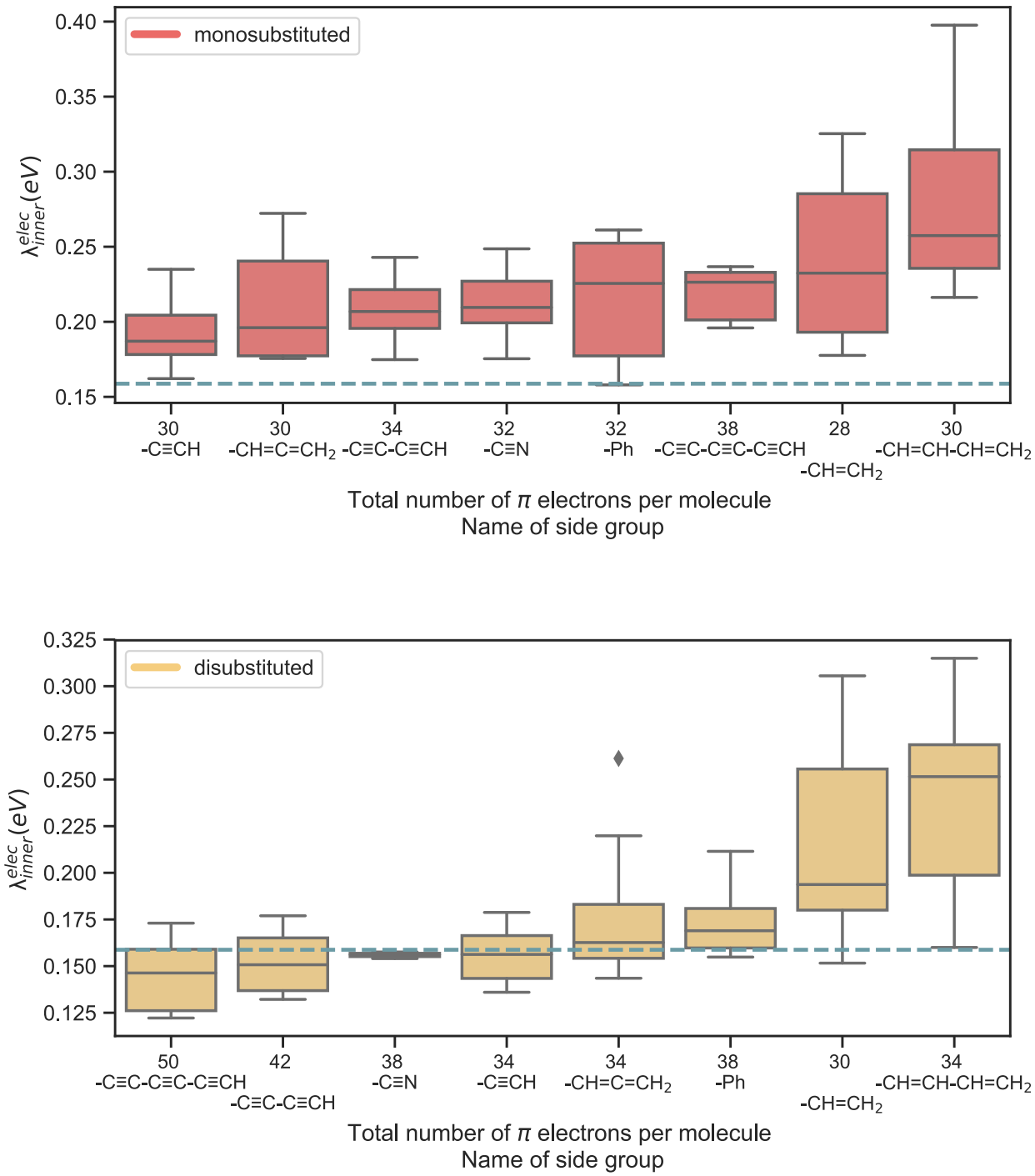


Figure 4.15: Boxplot showing the trends of inner electronic reorganisation energies λ_i^{elec} per π-system containing side groups, monosubstituted (red, top) and disubstituted (yellow, bottom). The dashed line is the λ_i^i of [6]helicene. This subset of the data only includes side groups which increase the total π-system size. For symmetrically disubstituted triple bonds the median λ_i^{elec} is lower than for our [6]helicene reference (yellow, bottom). The size of the box is the interquartile range, containing 50% of the ratios, and the horizontal line inside each box is the median value. The whiskers indicate the range of ratios obtained outside the middle 50% of data points, with the exception of some large outliers, shown as grey diamonds. Reprinted with permission from Reference 358. Copyright 2021 American Chemical Society.

no outer-sphere reorganisation energy. Even if the system was not in the hopping regime, and was more accurately described by the 'transient localisation' regime, the charge mobility would still increase with the power of J. Thus, J would remain a powerful screening parameter for good transport.[155]

The effect of side-group functionalisation on the dimer charge transport, molecular assembly and crystalline charge transport of the aromatic [6]helicene backbone was studied by exploring many side groups and substitution positions. The ranges of J/λ screening ratios obtained for each side-group are shown in Figures 4.16 and 4.17. Figure 4.16 shows that no side group consistently outperforms the unsubstituted reference. This is due to the translational dimer motif being disfavoured for some substitution positions due to steric hindrance introduced by the side-group. This interpretation is supported by the increase in spread (*i.e.* the vertical length between minimum and maximum scores obtained per side group) with increasing J/λ ratio. Thus, some substitution positions, independent of the side group attached, are consistently disadvantageous for charge transport in this motif. To give an example, 1-,16-monosubstitutions and 1,16-disubstitutions consistently result in very low OSC performance scores, because side-group functionalisation in these positions causes the side groups to lay in between the π-π systems of the aromatic backbones, sterically hindering the translational motif. Whilst for other substitution positions, if a suitable side group is attached, a large transfer integral is observed for some substitution position, but not all, resulting in a larger range of J/λ ratios for this side group.

Only a marginal proportion of J/λ ratios (0.4%) exceeds the unsubstituted [6]helicene reference. This result shows that side-group functionalisation of [6]helicenes to improve electron transport might not be as straightforward as assumed. Contrarily, this result reiterates the fact that high-performance OSC screening is a complex multivariable problem, requiring sophisticated solutions to account for molecular assembly and limitations of charge transport models.

Generally, the side groups can be grouped into three broadly similar classes of substituents: small electronegative side groups, moderately sized acyclic side groups with unsaturated bonds and bulky (a)cyclic side groups. Small electronegative side groups, *e.g.* halogens or the cyano group, tend to give rise to the best J/λ ratios. Within the range of acyclic groups with unsaturated bonds (*e.g.* triacetylene, diacetylene, or the ethynyl group presented in the previous chapter) some

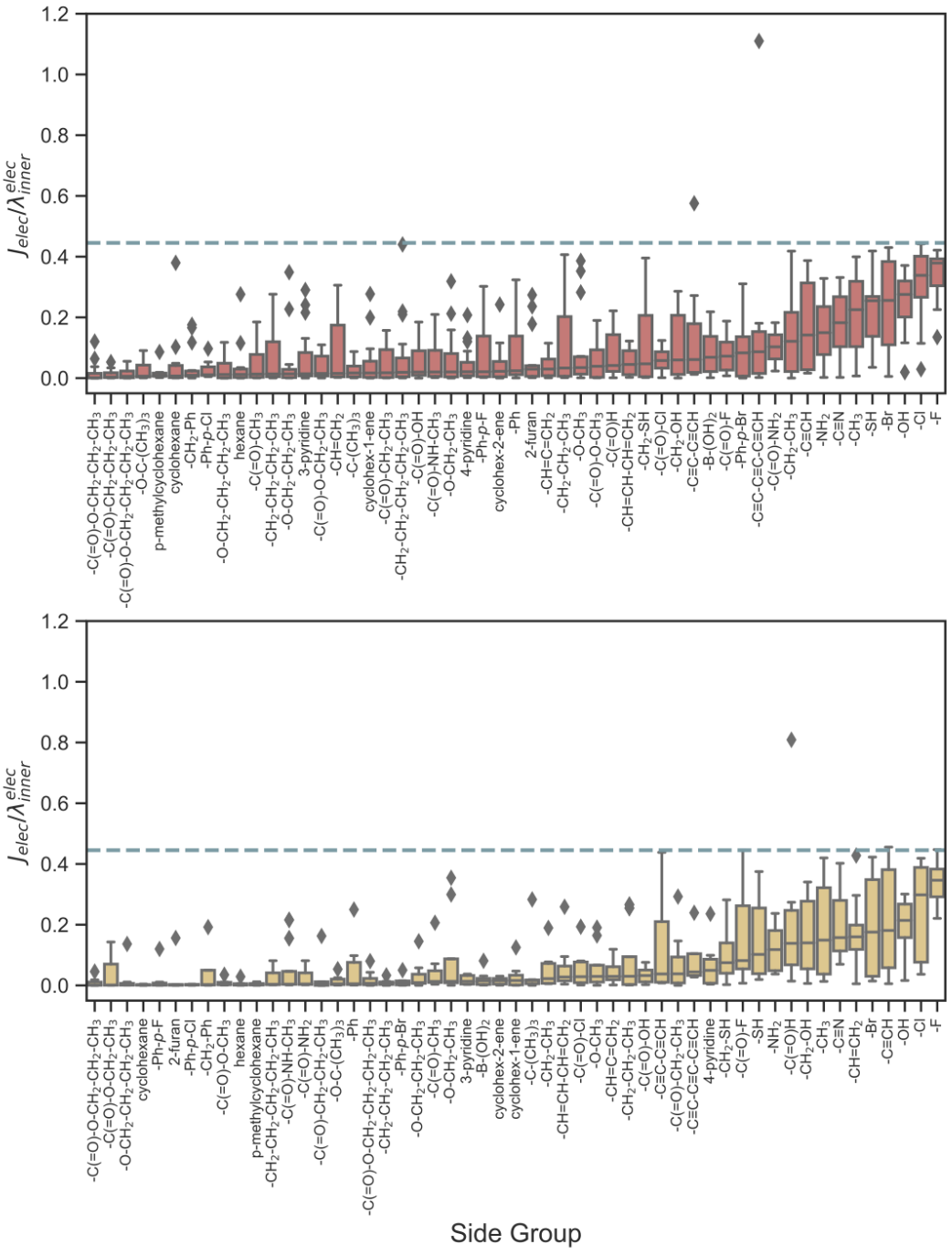


Figure 4.16: Boxplot illustrating the OSC suitability of functionalised [6]helicene molecules, using J/λ_i as the target property to be maximised. Results are shown for monosubstitution (top, red) or symmetric disubstitution (bottom, yellow) of the [6]helicene backbone. For each side group, 16 (monosubstituted) or 8 (disubstituted) data points were obtained, corresponding to all possible backbone sites. The size of the box is the interquartile range, containing 50% of the ratios, and the horizontal line inside each box is the median value. The whiskers indicate the range of ratios obtained outside the middle 50% of data points, with the exception of some large outliers, shown as grey diamonds. The dashed line is the unsubstituted [6]helicene reference. Figures 4.14 and 4.12 show the different components (λ, J) of this data separately. Reprinted with permission from Reference 358. Copyright 2021 American Chemical Society.

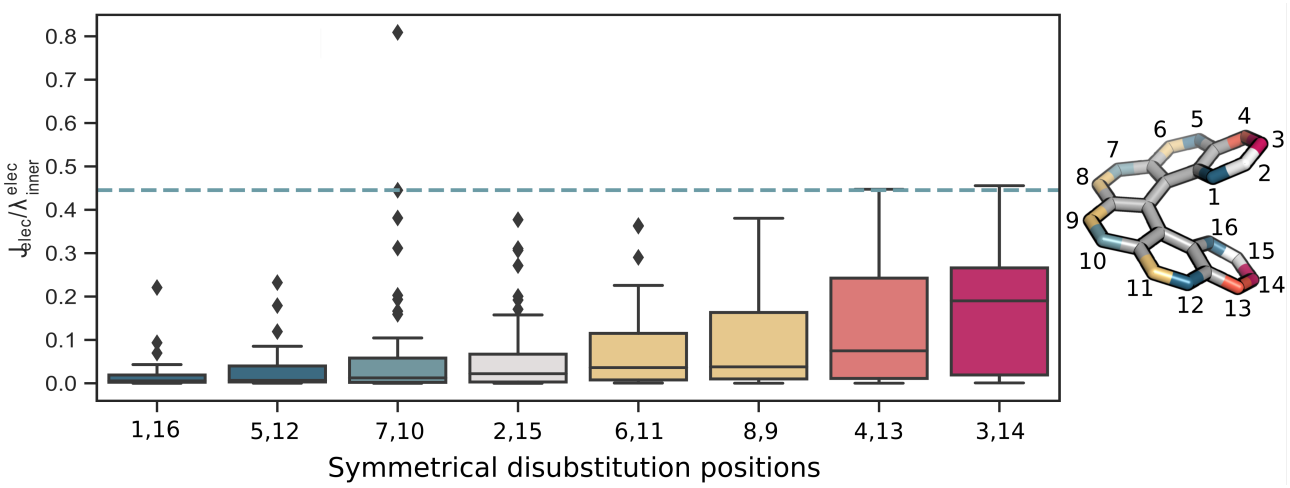


Figure 4.17: Boxplot illustrating ranges of the computed $J_{elec}/\lambda_i^{elec}$ ratios spanning all 56 side groups. Results are shown for the 8 symmetrical disubstitution patterns of the functionalised [6]helicene molecules, sorted in order of ascending median, and the colouring ranges from good charge transport (red) to no charge transport (blue). Significant outliers are shown as grey diamonds and the dashed line is the unsubstituted [6]helicene reference. Analogous results for the monosubstituted [6]helicenes are shown in Figure 4.18. Reprinted with permission from Reference 358. Copyright 2021 American Chemical Society.

outlier examples of high J/λ ratios for some substitution positions are observed. The last bulky substituent group tends to yield poor transport. Overall, smaller side groups tend to perform better than large bulky side groups, because the latter can affect electron delocalisation along the [6]helicene backbone by changing its inner strain and interplanar angle (the dihedral angle between the two terminal rings). Similar trends are observed for μ in the low-field limit, but they are less pronounced.

Triple bonds lead to significantly lowered λ_i, supposedly due to an increased propensity to accommodate a negative charge on this type of functionalised [6]helicene due to its increased π-system size. For disubstitution, the inner-sphere reorganisation energy λ_i decreases with the number of triple bonds attached per side group (Figure 4.15 bottom). Triple bond containing side groups have been previously engineered, most famously the triple-bonded 6,13-bis(triisopropylsilylethynyl) (TIPS) side group, to fine tune transport properties for molecular cores (*e.g.* pentacene,[326] *s*-indacenes[327]). The TIPS side group is a prime example, because it increases the molecule's solubility and directs the crystal packing, without reducing the degree of intermolecular electronic coupling. Depending on the size and shape of the triple-bond containing side group on a backbone, such as pentacene, different crystalline packings with varying intermolecular in-

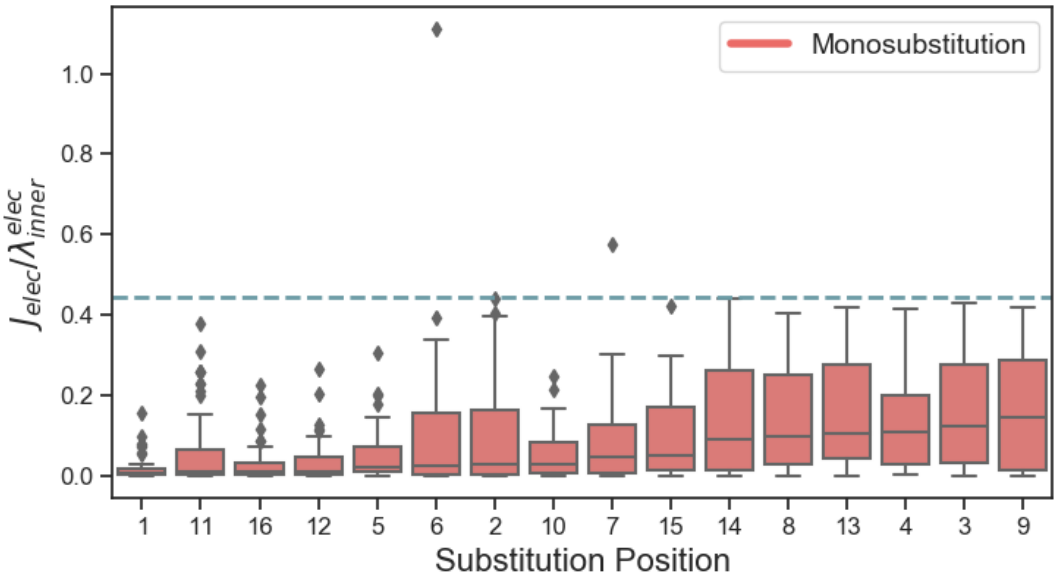


Figure 4.18: $J_{\text{elec}}/\lambda_{\text{elec}}$ ratios per functional group position in increasing order of median $J_{\text{elec}}/\lambda_{\text{elec}}$ from left to right. This Figure compares $J_{\text{elec}}/\lambda_{\text{elec}}$ ratios obtained from monosubstituted helicenes only. Reprinted with permission from Reference 358. Copyright 2021 American Chemical Society.

teraction strengths are observed, *i.e.* the crystalline packing can be controlled.[321,322] Furthermore, for TIPS-Pn, TMTES-Pn and diF-TESADT (triple-bond containing side groups), the distribution of vibrational modes is controlled by the side group attached.[384] Thus, a side group can be selected to additionally minimise nonlocal electron-phonon coupling processes along the π-π stacking direction (translational motif), which otherwise limit the crystal mobility.[384]

Another promising side-group functionalisation suggested by the screening is fluorination, because for both mono- and disubstituted [6]helicenes fluorination performed better than any other side group. Electron-withdrawing halogens have been previously reported to enhance charge transport by lowering the LUMO energy and the electron-injection barrier, such that hole and electron transport can occur simultaneously.[385–387] In particular, fluorination has been a prominent modification for other state-of-the art OSC, ranging from arylene diimides,[388,389] naphthalene,[390] to functionalised acene derivatives[321], such as pentacene[391–393] or TIPS-pentacene.[393] Besides the direct impact of fluorination on charge transport properties, fluorination also indirectly influences charge transport by decreasing the intermolecular spacing within the molecular crystal. Reduced intermolecular distances upon fluorination have been observed for carbon diimides[97], naphthalene,[390] and benzene.[394]

The set of monosubstituted [6]helicenes (Figure 4.12) showed higher mean and maximum

Rank by $J_{elec}/\lambda_{\mathrm{i}}^{\mathrm{elec}}$	Monosubstition	$J_{\mathrm{elec}}/\lambda_{\mathrm{i}}^{\mathrm{elec}}$	Disubstitution	$J_{\mathrm{elec}}/\lambda_{\mathrm{i}}^{\mathrm{elec}}$
1	6-triacetylene	1.11	7,10-dialdehyde	0.81
2	7-diacetylene	0.58	3,14-diethynyl	0.46
3	2-chloro	0.44	4,13-difluoro	0.45
6	15-fluoro	0.42		
12	13-fluoro	0.42		
Rank by μ_{elec}	Monosubstitution	μ_{elec}	Disubstitution	μ_{elec}
1	6-triacetylene	1.46	7,10-dialdehyde	0.75
2	7-diacetylene	0.38	3,14-diethynyl	0.28
3	2-chloro	0.26	4,13-difluoro	0.27
8	15-fluoro	0.24		
14	13-fluoro	0.24		
Rank by J_{elec}	Monosubstitution	J_{elec}	Disubstitution	J_{elec}
1	6-triacetylene	251	7,10-dialdehyde	191
2	7-diacetylene	135	7,10-diacylfluoride	118
3	15-fluoro	74	5,11-diacylfluoride	79
4	14-propylether	73	4,13-difluoro	78
29	13-fluoro	71		

Table 4.2: Top candidates for the mono- and disubstituted [6]helicenes according to both OSC measures: $J_{\mathrm{elec}}/\lambda_{\mathrm{i}}^{\mathrm{elec}}$ ratio and low-field limit charge-carrier mobility. The electron mobility μ_{elec} is given in $\times 10^{-3} cm^2\ V^{-1}\ s^{-1}$. The relevant J_{elec} values are stated in meV. Reprinted with permission from Reference 358. Copyright 2021 American Chemical Society.

transfer integrals than observed for the set of symmetrically disubstituted molecules (Mono: $\bar{\mu}_{\max}$ = 21 meV, $J_{\max}$ = 251 meV; Di: $\bar{\mu}_{\max}$ = 17 meV, $J_{\max}$ = 191 meV). Equally, the mean J/λ ratio was slightly higher for mono- than for disubstituted helicenes (0.10 vs. 0.08). Given that the quickest and most common synthetic route to [6]helicenes is a photocyclisation,[198] monosubstituted [6]helicenes have been less experimentally investigated. Figure 4.17 shows the best symmetrical disubstitution positions are the 3,14- and 4,13- positions (3,3'- and 4,4'- respectively). Similarly, the best monosubstitution positions according to Figure 4.18 were 3-, 9-, and 13- positions (3-,8'-,4'-, respectively).

The most promising side-group functionalised [6]helicene molecules, *i.e.*, among the top-ranked in either of the two screening metrics applied, are listed in Table 4.2. Comparing the performances of all side groups screened for, the highest OSC performance scores, J/λ and μ, were consistently observed for the fluorine side group (Figures 4.16 and B.1). Fluorine was by far the highest-performing side group. Across all possible fluorine substitution patterns,

the 4,13-difluoro-[6]helicene molecule resulted in the highest scores (both J/λ and μ). Since fluorinated [6]helicene are considered promising based on the screening, the highest-performing 4,13-difluoro-[6]helicene molecule was selected for further CSP analysis. Furthermore, the structurally related and second best monofluorinated molecule (scores are nearly equivalent), 13-fluoro-[6]helicene, was selected for CSP analysis (Table 4.2). Selection of the structurally related 13-fluoro-[6]helicene molecule over the highest performing monosubstituted 15-fluoro-[6]helicene facilitates direct comparison of the CSP results between the difluorinated and monofluorinated molecules. This may result in a deeper understanding of how polyfluorination over monofluorination shapes the crystal-energy landscape and the resulting charge transport properties.

Besides the generally best performing side group, Table 4.2 highlights that there are large outliers with significantly higher screening metrics than observed for fluorination. For both screening metrics, J/λ and μ, the top three highest-performing disubstituted [6]helicenes are 7,10-dialdehyde-, 3,14-diethynyl-, and 4,13-difluoro-[6]helicene. Equally, for monosubstitution the trends remain the same according to both metrics with the top three molecules being 6-triacetylene-, 7-diacetylene-, and 2-chloro-[6]helicene. The overall highest performing monosubstituted and disubstituted molecules, 6-triacetylene- and 7,10-dialdehyde-[6]helicene, have J/λ scores 164% and 80% higher than the highest mono- or difluorinated [6]helicene respectively. Therefore, besides selecting two fluorinated helicenes for CSP analysis, the highest performing mono- and disubstituted molecules were also chosen for further CSP analysis.

4.3.4 Analysis of CSP Results

This section serves as a summary of the general CSP results obtained. Then, in the following subsections, which are named by the respective molecule, the polymorphism predicted as well as the associated structure-mobility relationships (*i.e.* the crystal mobility in the lowest-lying polymorphs within each landscape) and the role of the translational motif will be discussed in more detail.

The computational screening of ~1300 helicene molecules for high OSC suitability revealed the fluorine side group as the best side group of choice, independent of the substitution po-

sition and pattern. The highest-performing 4,13-difluoro-[6]helicene and the second highest monofluorinated [6]helicene, 13-fluoro-[6]helicene, were selected for CSP. The two molecules with outstandingly high OSC screening criteria (Table 4.2), 7,10-dialdehyde- and 6-triacetylene-[6]helicene, were also selected for further CSP analysis. The four selected molecules for CSP are shown in Figure 4.19. None of these molecules have been synthesised, nor have their charge mobilities been measured. To validate the effectiveness of the presented screening approach in identifying promising molecules, the lowest-energy polymorphs of the four candidate molecules were generated using CSP (Section 4.2.2.1). The relative crystal-energy landscapes are shown in Figure 4.20. Broadly speaking, the lower in energy the polymorph, the more likely it is to be observed experimentally. Then, for the 10 lowest-lying polymorphs, the charge mobilities were computed to allow a comparison across systems.

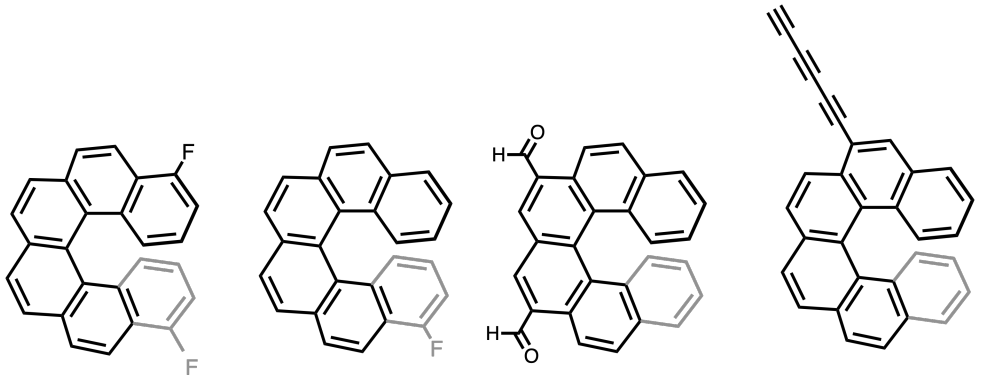


Figure 4.19: The four promising OSC molecular material candidates were identified by the screening. From left to right: 4,13-difluoro-[6]helicene, 13-fluoro-[6]helicene, 7,10-dialdehyde-[6]helicene, and 6-triacetylene-[6]helicene. Reprinted with permission from Reference 358. Copyright 2021 American Chemical Society.

Figure 4.20 illustrates the frequency with which the translational dimer motif was found (blue dots). Across the lowest-energy crystals shown in the four polymorphic landscapes, the translational dimer motif was found with 68% and 46% frequency for 13-fluoro-[6]helicene and 4,13-difluoro-[6]helicene, respectively. For 7,10-dialdehyde and 6-triacetylene-[6]helicene, the two molecules with outstandingly high OSC rankings, the translational dimer motif was found less, with 18% and 22% frequency, respectively. In the [6]helicene polymorphic 'reference' landscape, calculated using the same computational workflow, the motif was observed in 68% of the crystals. However, if the frequency of the motif is assessed across the 10 lowest-lying crystals for each landscape, the motif is more prevalent and observed in 100% (4,13-difluoro),

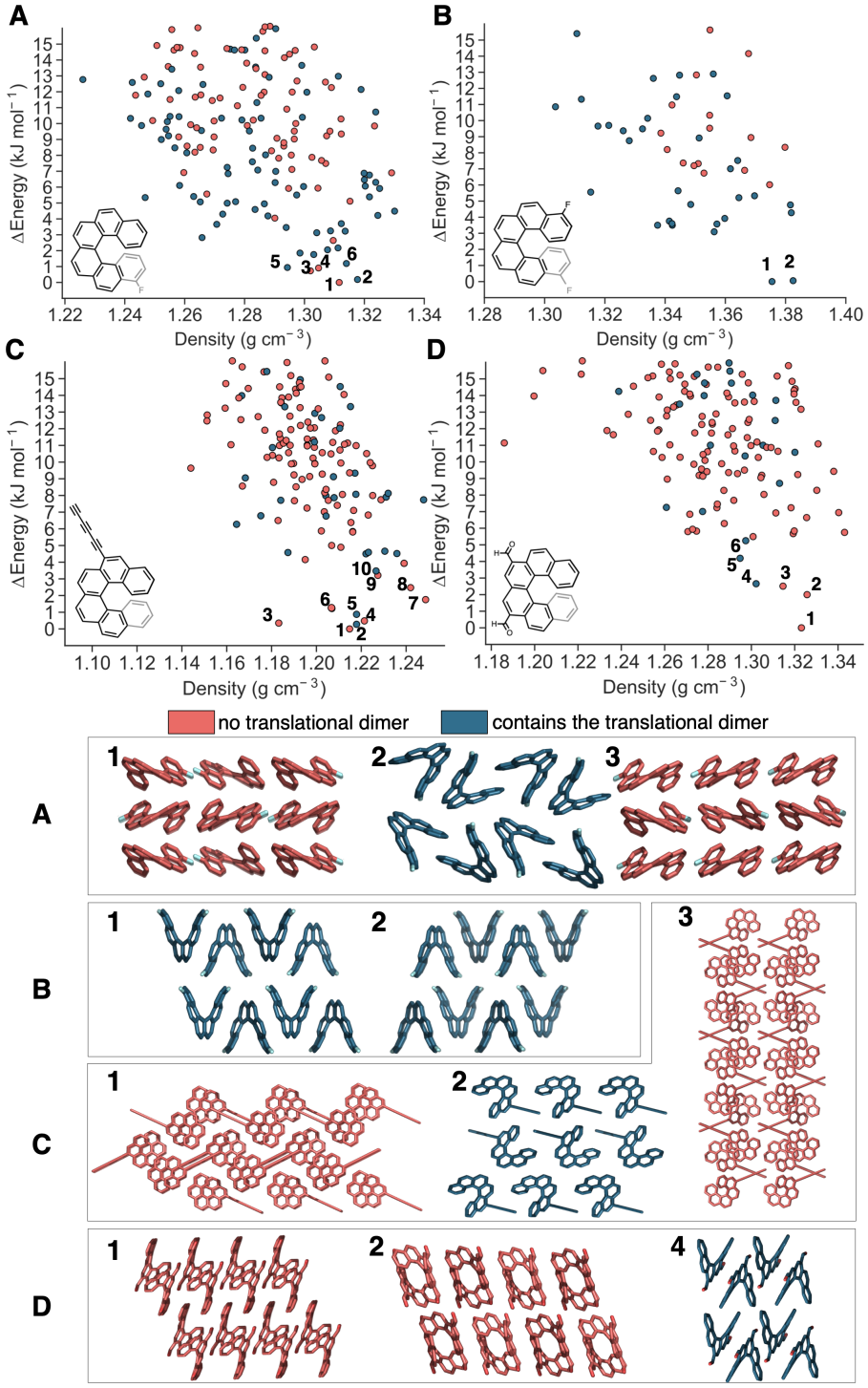


Figure 4.20: Top: Crystal-energy landscapes for 13-fluoro-[6]helicene (**A**), 4,13-difluoro-[6]helicene (**B**), 6-triacetylene-[6]helicene (**C**) and 7,10-dialdehyde-[6]helicene (**D**). Data points are shown in blue (red) if they do (do not) contain the translational dimer motif. Low-energy crystals are numbered in order of ascending relative energy from the global minimum. The electron mobilities for the 10 lowest-energy polymorphs are reported in Figures 4.21 and 4.22. Bottom: low-energy crystal packings, corresponding to the labeled structures above. Oxygen atoms are shown in red, and fluorine in light blue. Hydrogen atoms are omitted for clarity. Reprinted with permission from Reference 358. Copyright 2021 American Chemical Society.

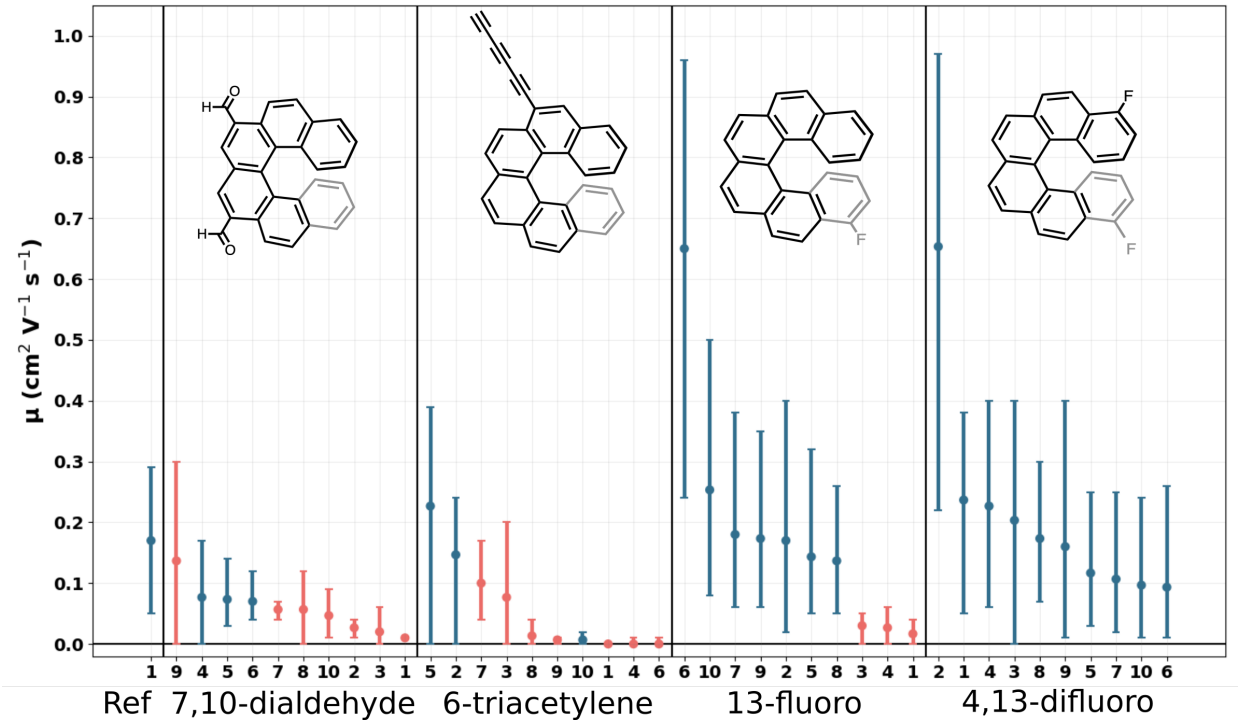


Figure 4.21: Minimum, maximum, and average electron mobilities, μ, computed for each of the 10 lowest-energy polymorphs identified from the CSP searches. Data for structures with (without) the translational motif are shown in blue (red). Results for the reference [6]helicene crystal (HEXHEL) are shown on the far left. Reprinted with permission from Reference 358. Copyright 2021 American Chemical Society.

70% (13-fluoro), 30% (6-triacetylene) and 30% (7,10-dialdehyde) of the 10 lowest-lying crystals ([6]helicene: 90%). The low proportion of translational dimer motif found for 6-triacetylene-[6]helicene can be attributed to the bulky character of the side group, largely affecting molecular shape and size, and thereby accessing unseen and a wider range of packing motifs. Similarly, the flexible 7,10-dialdehye side-group functionalisation results in alterations to the molecular shape and size. The fluorine side-group functionalisation does not affect the molecular shape and similar structural motifs to [6]helicene are observed, *e.g.* homochiral translational chain in one direction and back-to-back dimers in the orthogonal plane. Owing to the altered single-molecule geometry of 6-triacetylene- and 7,10-dialdehyde-[6]helicene, the translational dimer motif is unfortunately not the most dominant motif and hence less likely to be found. However, it is plausible that a solid form containing the motif can be obtained via a targeted experimental polymorph screening. Furthermore, despite side-group functionalisation, the translational motif does remain a commonly observed 1-dimensional packing motif in these functionalised helicene systems.

For the 10 lowest-energy crystal structures, the electron mobilities were computed (along all three planes of the crystal) to acquire structure-property relationships and assess the OSC performance of the four candidate molecules (Figure 4.21). In both fluorination examples, where the translational motif is prevalent in the lowest-energy crystals, it gave rise to the highest or second-highest transfer integrals in the crystal polymorphs and the transfer integrals were in good agreement with the screening prediction. Across all 4 molecules, the occurrence of the translational motif is associated with high electron mobilities relative to other solid forms in the polymorphic landscape.

A distribution of all mobilities computed for any of the planes of the crystal and any of the solid forms is presented in Figure 4.22. This quantitative comparison reveals that 6-triacetylene- and 7,10-dialdehyde-[6]helicene both have mean and maximum mobilities significantly lower than the [6]helicene backbone. Hence, the initial doubt of 6-triacetylene potentially being outside the hopping regime was not confirmed by high electronic transfer integrals and charge-hopping rates. In contrast, this observation indicates that neither of the two molecules are particularly likely to enhance charge transport over the unsubstituted [6]helicene reference. Nonetheless, for both fluorinated systems, higher mean and maximum electron mobilities than for the unsubstituted reference were observed. Thus, the screening result of fluorination provoking improved charge transport could be confirmed. Consequently, one would be most hopeful for the synthesis of fluorinated [6]helicenes to yield polymorphs with high charge-carrier mobility.

4,13-Difluoro-[6]helicene

The CSP of the symmetrically disubstituted 4,13-difluoro-[6]helicene generated two crystal structures with relatively similar crystalline packings, both significantly lower in energy than the remaining polymorphs (structures 1 and 2 in Figure 4.20B). All 10 lowest-lying polymorphs contain the translational motif (100% frequency). The dimer-based screening for this functionalised helicene predicted a dimer separation of 7.2 Å with an electronic transfer integral of J_{elec} = 78 meV. In the two lowest-lying polymorphs, the translational motif was observed at 7.3 Å with a transfer integral of J_{elec} = 77 meV. All transfer integrals are similar to the unsubstituted reference dimer (78 meV).

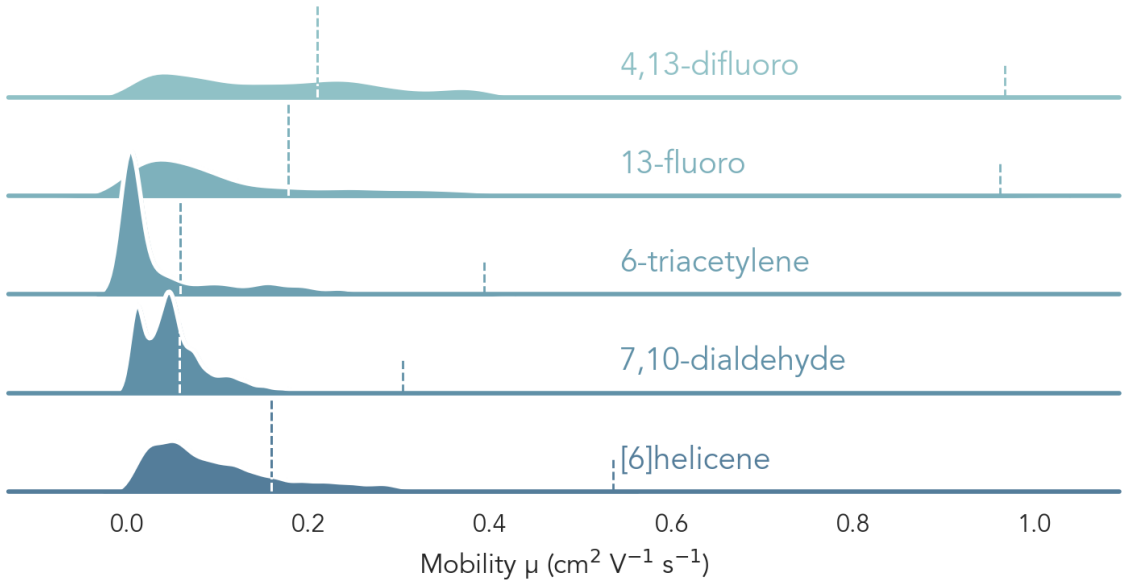


Figure 4.22: Distribution of electron mobilities observed across the ten lowest-energy crystals for each of the top four candidate molecules for which a CSP was carried out ([6]helicene results were recomputed based on crystals previously reported in Ref. 197). Maximum and average electron mobilities are indicated with dashed lines. The mobility plots for individual crystals can be found in Figure 4.21. Reprinted with permission from Reference 358. Copyright 2021 American Chemical Society.

Inspecting the translational motif's role in the hopping transport, in the lowest-lying polymorph (structure 1), the translational motif was the highest transporting (J_{max}) dimer in the crystal structure. In the second-lowest polymorph (structure 2), the back-to-back dimer (Figure 1.5 VI,VII) was the highest transporting (J_{max}) dimer, with a J of 133 meV. This is the reason why the mobility in this particular polymorph is significantly higher compared to the remaining crystals (Figure 4.21). Thus, the best performing motif in this particular polymorph (structure 2) is the back-to-back dimer. Whilst the screening predictions could be successfully confirmed by CSP, it is not ideal that the back-to-back dimer is the highest-transporting dimer in the highest-transporting crystal structure. This means that, whilst the general trend of improving charge transport with this side-group substitution pattern was correctly predicted by the screening, a second dimer (included at the screening stage) would be ideal to capture the full length of the CSP results in the computational screenings. Generally, the highest-performing fluorine functionalised molecule predicted by the dimer-based screening remains the most promising OSC candidate upon CSP

analysis, because the predicted charge mobilities of the two low-energy crystals both outperform the unsubstituted [6]helicene. At the chosen level of theory, both solid forms are nearly degenerate, with structure 2 having significantly higher charge mobility (due to the back-to-back dimer). Further periodic relaxations of the crystal structures could potentially flip the energy ranking and the highest-transporting structure could become the most stable polymorph, *i.e.* the most likely experimentally observed structure.

13-Fluoro-[6]helicene

CSP of 13-fluoro-[6]helicene produced in 176 low-energy crystal structures within 15 kJ mol^{-1} of the global minimum (Figure 4.20A). This is three times as many as obtained for the disubstituted, structurally related 4,13-difluoro-[6]helicene. However, the asymmetric 13-fluoro- sidegroup functionalisation stochastically increases the number of homo and heterochiral packing motifs possible (compared to the symmetrically disubstituted 4,13-difluoro-[6]helicene). 68% of the 13-fluoro-[6]helicene polymorphs within 15 kJ mol^{-1} arranged in the translational motif (Figure 4.20A). Seven out of the ten lowest-lying polymorphs organised in the translational dimer motif with an average dimer separation of $\bar{d}$ = 7.35 Å. The screening predicted a J of 71 meV at a separation of d_{min}=7.3 Å. Again, the two lowest-lying solid forms forming the translational motif (structures 2 and 5) at a separation of 7.3 Å and a J = 72 meV are in excellent agreement with the screening.

Comparison of the charge mobilities predicted across the 10 lowest-lying polymorphs strongly suggests a structure-property relationship between the occurrence of the translational dimer motif and increased electron mobilities. For polymorphs forming the translational motif (structures 2, 5-10), the mean maximum transfer integral is J = 80 meV. On the other hand, polymorphs *not* containing the translational motif (structures 1, 3 and 4) showed a mean J of 24 meV. Not only does the average transfer integral improve upon the presence of the translational motif, the mean maximum electron mobilities also significantly increase when the motif is present. The mean maximum mobility, $\bar{\mu}_{\text{max}}$, is 0.452 and 0.05 cm^2 V^{-1} s^{-1} across polymorphs with and without the motif, respectively. The highest-transporting solid form (structure 6) has a J_{max} = 128 meV and μ_{max} = 0.96 cm^2 V^{-1} s^{-1}, provoking a 231% performance increase in the electron mobility com-

pared to the unsubstituted reference (0.29 $cm^2\ V^{-1}\ s^{-1}$). However, structure 6 is not among the lowest-lying crystal structures, but within <2 kJ mol^{-1} of the global minimum and thus could still be observed experimentally. Similarly to the difluorinated case, higher level periodic relaxations or consideration of thermal corrections to the free-energy could lead to energetic re-ranking of the crystal structures.

6-Triacetylene-[6]helicene

The 6-triacetylene case is the sterically largest side-group functionalisation considered for CSP analysis and hence a challenging test case for the translational dimer-based screening approach. The dimer-based screening predicted the 6-triacetylene-[6]helicene molecule to align in the translational dimer at the shortest seen separation of 5.7 Å of all helicene molecules considered. The CSP analysis endorses the screening prediction with the translational dimer being found at a separation of 5.5 Å for forms 2 and 5 (Table B.2). Despite the short separation being in agreement with our dimer-based screening, the translational dimer was found more scarcely, in 22% of the crystal structures. However, presence of the translational motif was again associated with high electron mobility reaching outstanding transfer integrals of J_{elec} = 125 meV (structure 2) and J_{elec} = 126 meV (structure 5). These transfer integrals are 76-77% higher than the [6]helicene reference (65 meV). The high electronic coupling integral potentially arises from parallel π–π stacking of the triacetylene side groups and the chiral backbone also being found at short distances. Figure 4.22 illustrates that structures 2 and 5, containing the translational dimer with outstanding J, generate the highest mobilities across the set of lowest-lying polymorphs; structure 5 shows a maximum mobility (0.39 $cm^2\ V^{-1}\ s^{-1}$) exceeding the [6]helicene reference. This mobility is still low compared to the fluorinated systems. From a crystal engineering perspective, it shall be mentioned that structure 6 contains a π-π stacked co-facial dimer, highly similar to the motif, except for it being a heterochiral stacked dimer, with the side group being attached to opposite ends of the backbone. This small alteration of the motif, despite sharing similar π-π overlap of the helicene backbone, reduces the mobility. On the whole, mobilities of crystals containing the translational motif perform better relative to crystals not containing this motif and the translational dimer was found at short separations as predicted by the screening (Table B.2). Nonetheless, the crystal mobilities do not

outperform [6]helicene crystal mobilities nor the fluorinated systems. Hence, 6-triacetylene side-group functionalisation does not remain as promising as intuitively suggested by the screening.

The lowest-lying structure does not contain the translational motif and shows extremely poor mobility. However, within just 1 kJ mol^{-1} of the global minimum, five structures are found, two of which contain the translational motif, and they are both the highest two electron-transporting polymorph (structures 2 and 5).

Furthermore, the high J/λ screening ratios suggest that this helicene might no longer operate in the hopping regime and could be more accurately described in the 'transient localisation' regime. Ideally, one would calculate the intermolecular vibrational modes to elucidate charge transport in this OSC by studying the thermal fluctuations, which might impede it. At this point, the material will be somewhat electron transporting, if the most promising solid forms could be formed. At this level of accuracy of the crystal-energy landscape, the charge transport of multiple solid forms would need to be computed at a more computationally demanding level, which would not be a worthwhile effort, nor within the scope of this dimer-screening based study.

7,10-Dialdehyde-[6]helicene

The second 'outlier' molecule, 7,10-dialdehyde-[6]helicene predicted by the screening was studied via a flexible CSP search, allowing for the rotation of the aldehyde groups inside the crystal structure. The CSP generated new packing motifs, which had not been previously observed in [6]helicene, *e.g.* back-to-back dimers with the side-arms overlapping to generate more of a network-type architecture (structure 1 in Figure 4.20D). Comparison of the tentative crystal structures with the screening dimers reveals that the orientation of the carbonyl oxygens with respect to the initial gas phase dimers has changed. Thus, the translational motif observed in the screening is not exactly the same as the translational dimer observed in the bulk crystal, but still found 18% of the time, or 30% of the time across the set of 10 lowest-energy crystals. Additionally, none of the crystal mobilities outperform the reference. This molecule has the poorest electron transporting capabilities across all four candidate molecules investigated (Figure 4.22). The majority of the polymorphs hardly exhibits any electron transport. The translational dimer motif was found in solid forms with weak electron transport, *e.g.* structures 4 and 5 with maximum mobilities of 0.17 and 0.14 $cm^2\ V^{-1}$

s^{-1}, respectively, compared to crystals without the motif (structures 1-3) and with extremely poor charge transport (0.01–0.06 $cm^2\ V^{-1}\ s^{-1}$).

4.4 Conclusions

In this chapter, a high-throughput dimer-based screening approach was developed and tested using 1344 side-group functionalised helicene molecules that were screened for their potential suitability as molecular OSC materials. The translational dimer-based screening leveraged knowledge regarding the ubiquity of the translational motif in previously studied [6]helicene systems to allow for an informed guess of this motif being observed in other [6]helicene crystal structures. The translational motif was the most frequently observed motif in the [6]helicene crystal-energy landscape and across the terminally substituted [6]helicene systems collected from the CSD.

The best side group predicted by the screening with the highest screening metrics overall (fluorine), equally resulted in the highest calculated electron mobilities in the solid-state, significantly surpassing the unsubstituted reference (μ increase by over 200%) and with the translational motif being frequently observed across the crystal-energy landscapes (70-100%). This was observed for some 13-fluoro and 4,13-difluoro-[6]helicene polymorphs with maximum electron mobilities μ of 0.96, and 0.97 $cm^2\ V^{-1}\ s^{-1}$, respectively ([6]helicene: 0.30 $cm^2\ V^{-1}\ s^{-1}$). The screening approach in conjunction with CSP was able to demonstrate that terminal functionalisation increases the helicene's suitability for *n*-type transport, in contrast to heteroaromatic substitution (*e.g.* aza[6]helicenes), which has shown to favour *p*-type transport.[312] Out of the four most promising molecules nominated by the screening, the 4,13-difluoro-[6]helicene manifested itself as the most promising electron transporting molecular material. It presented solid forms, especially the two lowest-energy forms, with crystalline packings with large charge mobilities consistently higher than the [6]helicene reference molecule. Moreover, the synthesis of this symmetrically 4,4'-difluorinated helicene should be straightforward via a Wittig reaction starting from 2-fluorobenzaldehyde. Owing to the two fluorinated helicene systems exhibiting the highest electron transport properties and the fact that the difluorinated helicene resulted in higher charge transport than the monofluorinated molecule for both the screening and the crystal mobilities, further research could consist of raising the degree of fluorination (tetra-, hexa- and octafluorina-

tion, *etc.*) to verify if this would further augment charge transport properties, as has been previously investigated for pentacene molecular materials.[392]

Across the polymorphic landscapes of the four most promising screening candidates, the translational dimer motif was observed with a relative proportion of 100% (4,13-difluoro), 70% (13-fluoro), 30% (6-triacetylene) and 30% (7,10-dialdehyde) across the 10 lowest-lying polymorphs. This observation encourages our translational dimer-based packing simplification, because targeted experimental polymorph screening could lead to the motif. The motif was observed in the new systems suggested and the separations at which the motif was found matched the separations predicted by the screening. This was true for even the closest packed 6-triacetylene dimer (5.7 Å). Even though the translational dimer motif did not naturally provide the maximum J in the bulk crystal, its occurrence was strongly associated with improved transport properties relative to other polymorphs, which did not align in this motif.

Nonetheless, one has to acknowledge that simplifications made regarding the crystalline packing (required in a screening approach) for a complex multivariable problem such as this risk the misclassification of systems that do not present with the assumed packing motif, or the screening may lead to incorrect OSC rankings of systems that lie outside the hopping regime. Regarding the latter, J/λ has to be used with care for strongly coupled systems and only serves as a first indicator for promising electronic properties. For full characterisation of strongly coupled systems, a phonon calculation of the relevant molecular crystal would be required to confirm less sophisticated screening metrics.

The dimer-based screening relies on a large assumption, but it is a computationally fast and cheap method to circumvent CSP, which is computationally infeasible at the screening stage. The presented screening approach allows for some consideration of molecular assembly, unlike most single-molecule screening approaches focussing uniquely on single-molecule properties (*e.g.* dipole, HOMO-LUMO gap, λ_i). However, the diverging relative abundance of the motif in the polymorphic landscapes highlights that there is still further work required to yield screening approaches for which screening more than ~1300 structures is feasible with a more representative approximation for the crystalline environment. The presented approach remains an improvement upon single-molecule and database screening approaches, which either completely ignore

the crystalline packing or are limited by molecules already known and/or previously synthesised.

In summary, the dimer-based screening approach as driver for the discovery of new chiral organic electronic materials with promising charge transport properties is short of a 'universal solution' in OSC research. Nonetheless, the motif-based screening is a way to predict observable structure-mobility relationships and to thereby guide the design of new functionalised electronic materials with tailored properties. In the future, this packing-based screening approach could be developed further to incorporate more crystalline packings and thus be more representative of the local crystal environment. An intelligent screening based approach leveraging polymorphism has the potential to accelerate the discovery of new chiral organic electronic molecular materials with bespoke properties.

Chapter 5

On the Effect of [n]helicene Length on Crystal Packing Interactions and Polymorphism

5.1 Introduction

Previous chapters revealed that terminally substituted [6]helicenes do not necessarily result in improved charge transport properties and large screenings for high charge transport were limited by the translational dimer-based model as a representation of the crystal packing. The study of terminal side-group functionalisation was in turn inspired by the earlier study on heteroaromatic (nitrogen) substitutions to the [6]helicene backbone. Thus, the next natural step was to extend the helicene research to helicenes of different helicene lengths, n, and to consider the polymorphic landscapes for each [n]helicene molecule. Therefore, the crystal-energy landscapes for naphthalene and [3]-[12]helicenes were predicted using crystal structure prediction (CSP). As a benchmark molecule with no axial chirality, naphthalene, consisting of two fused benzene rings, was chosen. The technically shortest helicene, [3]helicene, is more commonly known as phenanthrene. According to the X-ray crystal structure measurements of [3]helicene, the molecule shows small deviations from complete planarity, but the molecular geometry does not significantly differ from C_{2v} symmetry.[395] The chiral twist of the [n]helicenes is illustrated in Figure 5.1. The axial chirality is characterised by having a non-linear asymmetric dipole around the chiral axis (*i.e.* in a third dimension) and not around a chiral centre.[396] In other words, the spatial arrangement of the

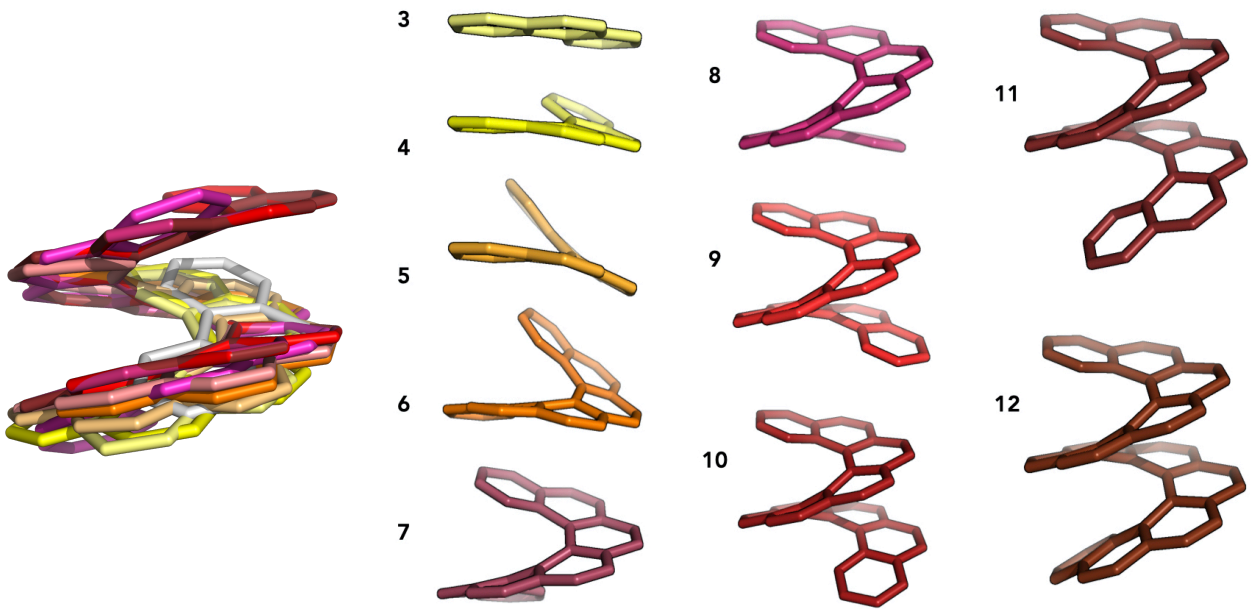


Figure 5.1: Molecular overlay and overview of [n]helicenes of varying helicene length n, here n=3-12. Helicenes are coloured darker with increasing n, for better visibility, the [3]helicene is highlighted in white in the molecular overlay. Hydrogens are omitted.

molecule around the chiral axis is not superimposable on its mirror image. Nakai *et al.* studied the chiroptical response of [n]helicenes (n = 4-10) as a function of the number of benzene rings (n).[206] They found that the leap of helical pitch upon moving from [5] to [6]helicene, *i.e.* the completion of the first helix, results in linearly correlated chiroptical properties.[206]

The naphthalene crystal structure was chosen as a benchmark case, because it has been extensively studied, with 55 naphthalene crystal structures deposited to the Cambridge Structural Database (CSD) and 15 publications associated with it. Naphthalene has been used as a model system for CSP using genetic algorithms,[397] CSP with molecular dynamics based sampling approaches[398] or the computation of the phonon frequencies.[399,400] Naphthalene molecular crystals, which have been synthesised many times, have not been reported to be polymorphic.[22] Naphthalene molecular crystals are organic semiconducting; using an electron-phonon scattering model, charge transport in naphthalene crystals follows band transport at low temperatures, whilst hopping transport prevails at room temperature.[145]

Out of the [3]-[12]-helicenes discussed in this chapter, [5]-,[6]-,[7]- and [9]-,[10]-,[11]-helicenes molecular crystals have been synthesised and X-ray crystal structures were deposited to the CSD.[401–411] For [5]-helicene[401–404] and [7]-helicene,[407,408] more than one crystal polymorph

has been obtained. The [3]helicene crystal structure could be regenerated from Reference 395 by means of the crystal parameters and the orientation of the carbon atoms provided. Even for helicenes of longer chain lengths, the synthesis can be performed via a one-step synthesis as demonstrated for [16]-helicene.[409]

As the n helicene length increases, the interplanar angle, θ, which is the angle between the two terminal rings of the helicene backbone and the distances r_1 and r_2 between the respective 1,1'- and 2,2'-disubstitution positions, as well as the distance between the terminal rings of each backbone increase (Table 5.1). The geometric differences based on geometry optimisations with and without dispersion corrections were assessed. Initially, at $n = 3$ the helicene optimises to a planar geometry, but at $n = 4$ a pronounced interplanar angle of 30-31° is observed at B3LYP theory (with and without using D3 dispersion correction). Table 5.1 shows the optimised geometric parameters using the B3LYP functional with and without D3 dispersion correction. Where possible, the geometric parameters of the experimentally observed X-ray structures are provided for comparison. For example, in the case of [6]helicene the r_1 and r_2 distances of the B3LYP-optimised geometries more closely match the experimental structure (CSD reference code: HEXHEL), whilst the interplanar angle (61.2°) is larger than observed for the X-ray geometry (57.3°) and slightly larger than for the B3LYP-D3 optimised geometry (61.0°). Contrarily, for [7]helicene the interplanar angle is more closely described (to the experimental structure) if the B3LYP optimised structure was used (than the B3LYP-D3 optimised structure), but the distances r_1, r_2 are better predicted by the B3LYP-D3 optimised structures. Comparing the [7]helicene optimised geometries, the experimentally observed interplanar angle is most accurately described if the structure is B3LYP optimised. However, the distances between the terminal ring positions (r_1, r_2) are better described by the B3LYP-D3 geometry. From the geometric parameters listed in Table 5.1 it is not clear whether dispersion correction will in fact consistently improve the overall geometries compared to B3LYP only, which is why the geometry optimisations without dispersion correction were used for CSP.

Research questions include:

- How does the crystal packing vary with varying helicene length n?
- How does the helicene length n affect the noncovalent interactions in the solid-state (*e.g.*, the degree of π-π and C-H$\cdots\pi$ stacking, presence of chiral columns)?

helicene length n	Method	r_1	r_2	θ
3	B3LYP	3.0	5.8	0.0
	B3LYP-D3	3.0	5.8	0.0
4	B3LYP	3.0	5.5	30.5
	B3LYP-D3	3.0	5.5	31.4
5	B3LYP	3.0	4.6	66.9
	B3LYP-D3	2.9	4.6	67.6
	DBPHEN02	2.9	4.6	67.1
	DBPHEN03	3.0, 2.9	4.7, 4.6	68.4, 69.3
	DBPHEN05	3.0	4.7	66.6
6	B3LYP	3.2	4.4	61.2
	B3LYP-D3	3.1	4.3	61.0
	HEXHEL	3.2	4.6	57.3
7	B3LYP	4.0	5.5	0.8
	B3LYP-D3	3.9	5.2	1.2
	HPTHEL	3.8	5.2	0.9
8	B3LYP	4.9	7.1	61.5
	B3LYP-D3	4.8	6.9	62.7
9	B3LYP	5.6	8.0	46.0
	B3LYP-D3	5.5	7.8	42.0
	1051158	5.4, 5.4, 5.4	7.8, 7.8, 7.8	42.2, 39.3, 46.1
10	B3LYP	5.9	8.0	15.8
	B3LYP-D3	5.8	7.7	19.8
	THELIC	5.7	7.7	19.8
11	B3LYP	6.2	7.6	69.0
	B3LYP-D3	6.0	7.3	72.4
	UHELIC	5.8	7.1	64.2
12	B3LYP	6.6	7.7	55.1
	B3LYP-D3	6.4	7.3	51.8

Table 5.1: Table comparing geometric parameters of [n]helicenes geometry optimised with and without Grimme's D3 dispersion correction, and as observed in the relevant X-ray crystal structures (CSD reference code or number are given where found). r_1 and r_2 are the distances between the carbon atoms in the 1,1'- and 2,2'-positions respectively. θ is the interplanar angle between the two terminal rings.

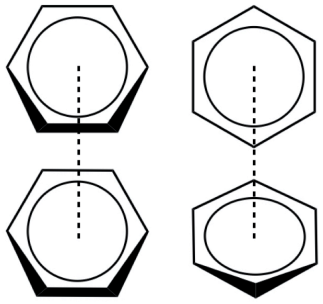


Figure 5.2: The two dominant types of crystalline π-stacking: π-π stacking and C-H$\cdots\pi$ stacking.

- Does the length of the helicene backbone impact the relative proportion of enantiopure and racemic crystals in the low-energy crystal-energy landscape?

This chapter does not aim to predict all possible crystalline packings ($Z'>1$). Instead, it will focus on Z'=1 crystals, even if in some cases higher order crystals have been experimentally observed.[403,409] Furthermore, the purpose of this chapter is not to provide accurate crystal-energy rankings for each of the crystal-energy landscapes. From the work presented in previous chapters, it is clear that the accurate crystal-energy ranking of helicenes in the CSP landscape is not a trivial problem and requires extensive dispersion-corrected DFT cell optimisations to yield an accurate crystal-energy ranking.

5.2 Methods

5.2.1 Crystal Structure Prediction

CSP searches were performed as described in Section 3.2.1 using the CRYSTALPREDICTOR/DMACRYS workflow, but without any DFT-D post-processing of the tentative crystal structures.

5.2.2 Crystal Packing Analysis

The crystal packing analysis was performed using CRYSTACK, a python module, developed here for the analysis of crystal stacking interactions. For each Z'=1 crystal in the 11 CSP landscapes, the degree to which each molecule was involved in parallel π-π and C-H$\cdots\pi$ stacking interactions was quantified on a per molecule percentage basis. The two types of stacking interactions are illus-

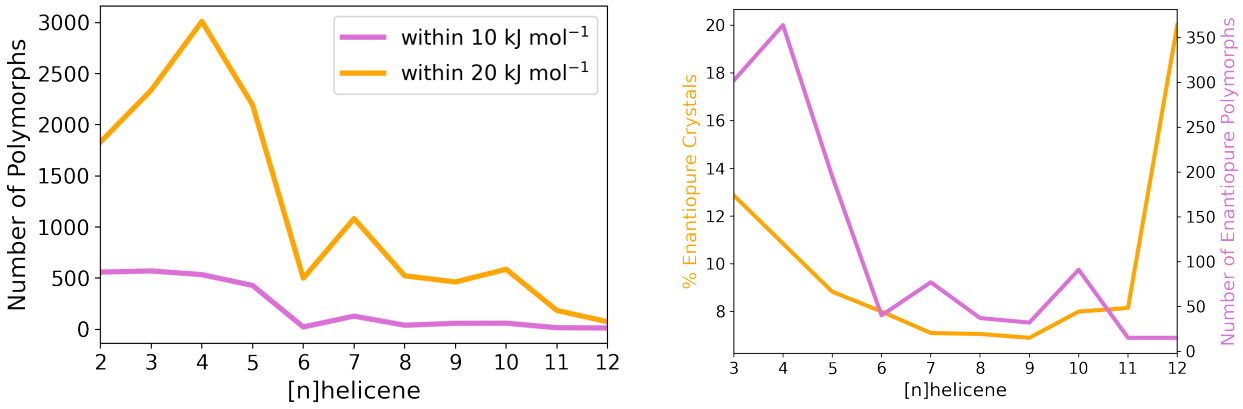
trated in Figure 5.2. To give an example, if the [6]helicene crystal-energy landscape was analysed and all six aromatic rings on the [6]helicene backbone are involved in parallel π-π stacks to neighbouring molecules, the parallel π-π stacking degree would be 100%. Similarly, if only 4 aromatic rings out of 6 would be involved in parallel π-π stacks, the parallel π-π stacking degree would be 66.6% (4/6) and so on. The same logic was applied to quantifying the degree of per molecule C-H$\cdot\cdot\pi$ stacking. Parallel π-π stacking (with and without offset) was only considered π-stacked if the intermolecular distance between the centre of the benzene rings did not exceed 5.5 Å and if the angle between the benzene rings was within 0° to 30°.. Rarely the two benzene rings are stacked perfectly on top of each other, meaning with their ring centres perfectly on top of one another. In reality, the centroid of one benzene ring tends to be slightly shifted and not perfectly underneath the other. This shift between the two parallel π-stacked benzene rings cannot exceed 2.0 Å. Otherwise it was not considered a π-π stacking interaction.

In addition to the degree of π-π and C-H$\cdot\cdot\pi$-π stacking, each crystal was tested for its degree of translational symmetry. It was measured whether all molecules in the neighbouring shell are parallel to each other and simply follow translational symmetry. The degree of parallel packing within each crystalline solid was computed using the ring centres of the terminal rings of each backbone. A vector connecting the terminal ring centres was computed for each molecule, then the dot product of the unit vectors of two molecular orientation vectors was computed. If the dot product was -1 or 1, the molecules were oriented antiparallel or parallel. If out of the 15 molecules in the neighbouring shell all molecules were oriented either parallel or antiparallel to the central molecule, the proportion of parallel molecular orientation was 100%.

5.3 Results and Discussion

5.3.1 General CSP Results

The number of stable crystal structures (polymorphs) obtained per [n]helicene molecule decreases with increasing helicene length n (Figure 5.3a). I hypothesise this is due to the increased spiral shape of the helicene with increasing number of benzene rings n. As the helicene length increases, less benzene rings remain accessible for stabilising interactions. Hence, the number



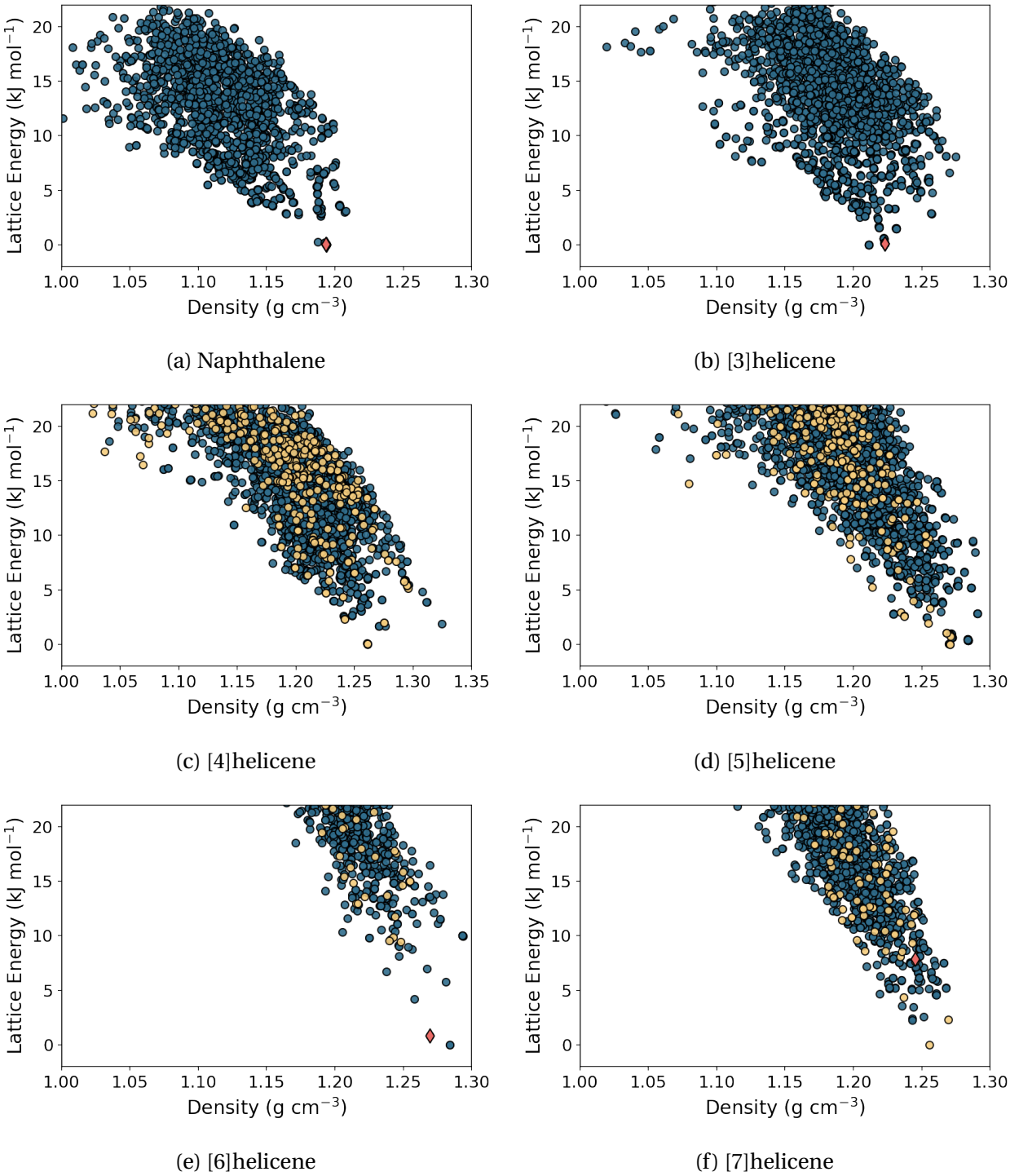
(a) (b)

Figure 5.3: (a) The number of polymorphs obtained for each helicene CSP search within the lowest 10 kJ mol^{-1} and 20 kJ mol^{-1} from the global minimum respectively. (b) The relative proportion of enantiopure crystals (orange) and absolute number of enantiopure polymorphs (pink) obtained for each helicene CSP search within the lowest 20 kJ mol^{-1} from the global minimum structure.

Helicene Length n	Total	Enantiopure	Racemic
Naphthalene (2)	1.	-	-
3	2.	2.	-
6	2.	1.	-
7	47.	-	44.
10	2.	2.	-
11	6.	1.	-

Table 5.2: Relative energy rankings of the experimentally observed structure according to the respective crystal-energy landscape. The relative position across the entire landscape (Total), among all enantiopure and racemic structures is shown.

of possible stabilising noncovalent interactions is reduced. If the helices are arranged parallel to each other (*e.g.*, in chiral columns), favourable π-π interactions across terminal rings could be enhanced. Conversely, C-H$\cdots\pi$ interactions should become scarce due to the increased steric hindrance. The individual energy-density plots for each CSP landscape are provided in Figures 5.4 and 5.5. Figures 5.4 and 5.5 are coloured by whether the crystals are enantiopure or racemic (except for naphthalene and [3]helicene), where possible, experimentally observed structures are highlighted too. Figures 5.4 and 5.5 depict the crystal-energy landscapes for naphthalene and [3]-[9]helicenes and [10]-[11]helicenes, respectively. Each of the CSP landscapes has the characteristic crystal-energy shape which follows the rule of closest packing (density rule) for molecular crystals. According to the density rule, the most densely packed forms tend to be the lowest-energy forms.[71] Hence, most CSP plots follow the diagonal relationship of higher density



(a) Naphthalene

(b) [3]helicene

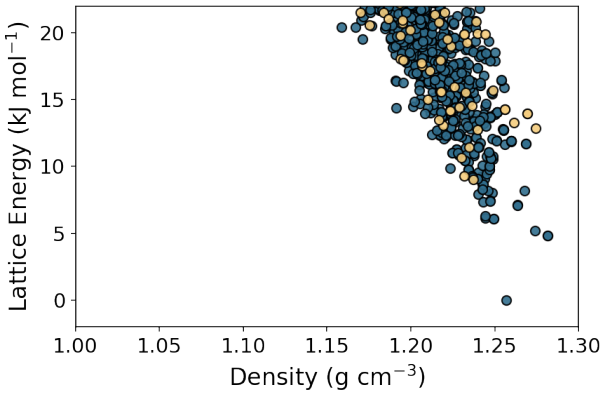
(c) [4]helicene

(d) [5]helicene

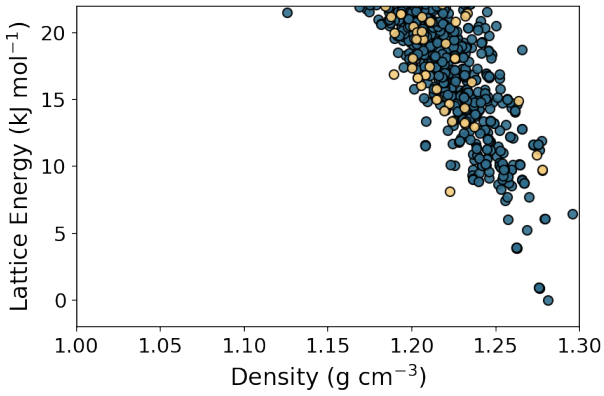
(e) [6]helicene

(f) [7]helicene

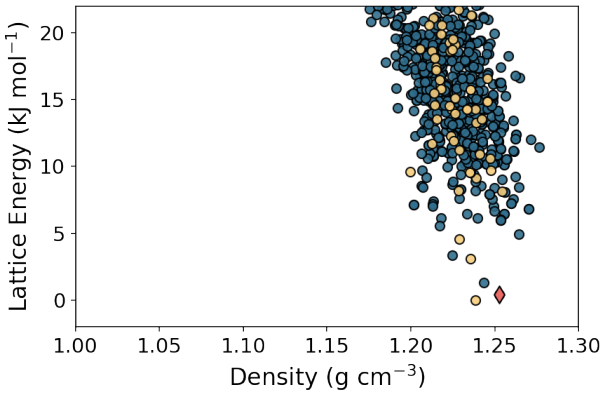
Figure 5.4: CSP landscapes for naphthalene and [3]-[7]helicenes of the lowest-lying polymorphs. Each crystal is coloured by the chirality of the crystal structure, either in yellow (enantiopure) or blue (racemic). The red diamonds indicate experimentally observed polymorphs, which have been predicted by CSP.



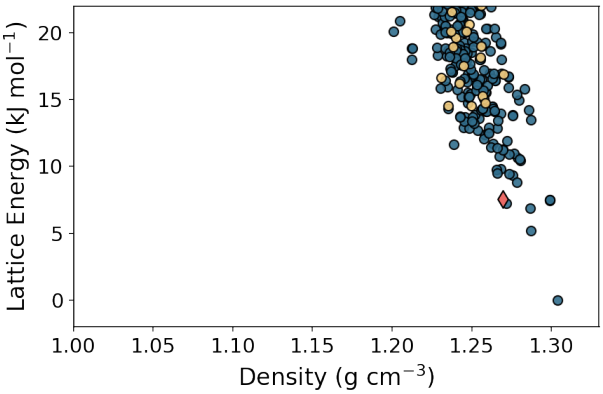
(a) [8]helicene



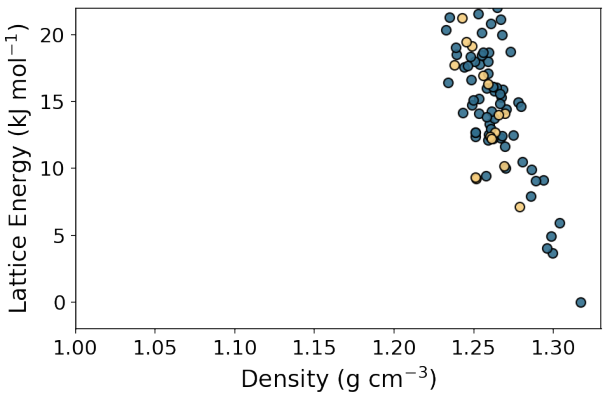
(b) [9]helicene]



(c) [10]helicene



(d) [11]helicene



(e) [12]helicene

Figure 5.5: CSP landscapes for [8]-[12]helicenes of the lowest-lying polymorphs. Each crystal is coloured by the chirality of the crystal structure, either in yellow (enantiopure) or blue (racemic). The red diamonds indicate experimentally observed polymorphs, which have been predicted by CSP.

leading to broadly lower relative energy. In the case of the [11]- and [12]helicene crystal-energy landscapes, the most dense form is the actual lowest-energy structure (global minimum).

Experimental structures were successfully predicted for naphthalene, [3]-, [6]-,[7]-,[10]- and

[11]helicene. Each of these successfully predicted helicene crystal structures (except for [7]helicene) was enantiopure (Z'=1) and found within the two lowest-lying enantiopure crystal structures (Table 5.2). All Z'=1 experimental crystal structures have been found, except for [5]helicene.

Naphthalene and [3]helicene: The naphthalene and the planar and shortest helicene, [3]helicene, CSP landscape show extremely wide crystal-energy landscapes (Figure 5.4a, 5.4b), which is suggestive of the fact that due to their planar architecture many potential crystalline packings are observed. The naphthalene molecule is planar and achiral. Hence it could not be analysed with respect to the chirality of the crystalline forms obtained. The experimentally observed naphthalene crystal structure was successfully predicted via CSP (Figure 5.6a). The predicted lowest-energy structure was the experimentally observed structure (RMSD = 0.63 Å, 15/15 molecules overlaid, Figure 5.4a). The experimental structure was predicted to be 0.10 kJ mol^{-1} lower in lattice energy than the second-lowest lying polymorph. This energy difference might increase upon DFT-D relaxation of the crystal structures. Despite many synthetic attempts, only one crystalline form of the naphthalene molecular crystal, *i.e.* the global minimum structure, has been previously grown.[412]

Since the molecular symmetry of [3]helicene largely follows C_{2v} symmetry, it can be considered achiral and hence the crystal-energy landscape has not been coloured by chirality. If the [3]helicene in the crystal structure is considered chiral, the relative proportion of enantiopure (chiral) structures observed in the CSP landscape is given in Figure 5.3b. The experimentally observed crystal structure parameters have been reported by Trotter in 1963 to be of $P2_1$ space group symmetry together with the carbon positions of the molecule in the asymmetric unit.[395] The crystal structure data provided was used to generate a CIF file, which was then overlaid with predicted structures of similar space group and crystal parameters. The molecular overlay of the experimental crystal structure with the second-lowest structure (structure 41) resulted in a match of all 15 molecules in the neighbouring shell with a RMSD = 0.342 Å (Figure 5.6b). The experimentally reported structure shows a dihedral angle along the inner carbons of 1.11°,[395] and is thus more pronounced than the dihedral angle measured for the geometry optimised molecule (0.02°). The experimentally observed $P2_1$ structure is the second lowest-lying crystal structure in the crystal-energy landscape and 0.10 kJ mol^{-1} away from the global minimum. The lowest-lying structure is of $P2_12_12_1$ symmetry. Part of the $P2_12_12_1$ crystalline packing resembles the $P2_1$ structure, *i.e.* 5

out of 15 molecules in the neighbouring shell can be superimposed.

[4]- and [5]helicene: With 3011 solid forms, the [4]helicene crystal-energy landscape has the most solid forms predicted within the lowest 20 kJ mol^{-1} energy window compared to any other helicene studied here (Figure 5.3a). The lowest-lying structure is enantiopure and 11.2% of the structures within the lowest 20 kJ mol^{-1} are enantiopure (Figure 5.3b). The density range is similar to the [3]- and [5]helicene crystal-energy landscapes. However, [4]helicene has some crystalline forms with particularly high densities (1.29-1.32 g cm^{-3}). These extremely efficiently packed crystalline forms are also found within the lowest 7.6 kJ mol^{-1} relative to the global minimum. There is no experimentally observed [4]helicene molecular crystal.

The [5]helicene crystal-energy landscape has 2194 polymorphs within the lowest 20 kJ mol^{-1} energy window (Figure 5.3a). Of the 429 polymorphs within the lowest 20 kJ mol^{-1} energy window, 9.3% were enantiopure. The energy-density trend largely follows the [4]helicene CSP landscape with higher density signifying lower relative energy. [5]helicene has three (unique) experimentally observed crystal structures (CSD reference codes: DBPHEN02 , DBPHEN03 and DBPHEN05), one of which is a Z'=2 crystal structure and hence could not be predicted within the Z'=1 CSP search.[401] The other two Z'=1 crystal structures were not completely matched with predicted structures. The highest similarity overlays between predicted and experimentally observed structures included matching 9/15 and 11/15 molecules within the neighbouring shell of each crystalline environment.[401]

[6]- and [7]helicene: The experimentally observed crystal structure (CSD reference code: HEXHEL) was successfully predicted (Figure 5.7a, 5.4e). The [6]helicene experimentally observed solid form is enantiopure ($P2_1$ symmetry). It was found at a density of 1.37 g cm^{-3} and a relative energy of 0.85 kJ mol^{-1} away from the global minimum. The experimentally observed crystal structure was found to be the second lowest-lying structure. The presented CSP results are consistent with the CSP results previously performed by Rice *et al.*.[197] The percentage of enantiopure crystals within 20 kJ mol^{-1} (5.8%) follows the global trend of decreasing percentage of enantiopure forms with increasing chain length n within the first full rotation from [3]- to [6]helicene (Figure 5.3b). Since the CSP landscape has been previously explored, only the 700 lowest-lying crystals (including the predicted experimental match) were carried onwards for further crystal packing analysis.

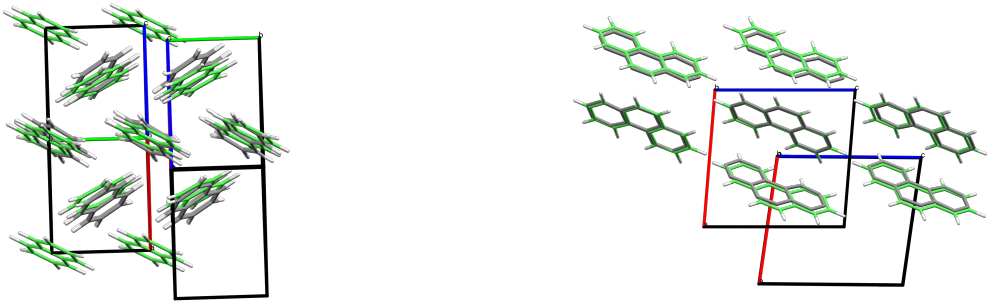
Despite the fewer number of polymorphs, the general shape of the crystal-energy landscape obeys the density rule (Figure 5.4e).

Likewise, the [7]helicene crystal-energy landscape follows the density rule. The two lowest-lying [7]helicene structures are enantiopure. The set of crystals within 20 kJ mol^{-1} is 6.3% and similar to the [6]helicene landscape. For the [7]helicene crystal-energy landscape, the experimentally observed structure (CSD reference code: HPTHEL01) was successfully predicted with a density of 1.25 g cm^{-3} and a relative energy difference from the global minimum of 7.86 kJ mol^{-1}. The experimentally observed structure is of $P2_1/c$ symmetry and thus racemic. The higher order Z'=2 experimentally observed structure (CSD reference code: IMEJIW) could not be predicted via the CSP search. The X-ray structure is of $P2_1$ symmetry and hence enantiopure.

[8]-, [9]- and [10]helicene: The crystal-energy landscape of [8]helicene follows on the trends of [6]- and [7]helicene with 6.1% of the crystals within 20 kJ mol^{-1} being enantiopure. Furthermore, the number of polymorphs observed continues to decrease with increasing helicene length n. The general shape follows the density rule, however here in contrast to previous landscapes, the set of lowest-lying crystals within 5 kJ mol^{-1} of the global minimum consists of only racemic structures. Unfortunately, there are no reported X-ray structures to compare the CSP results with.

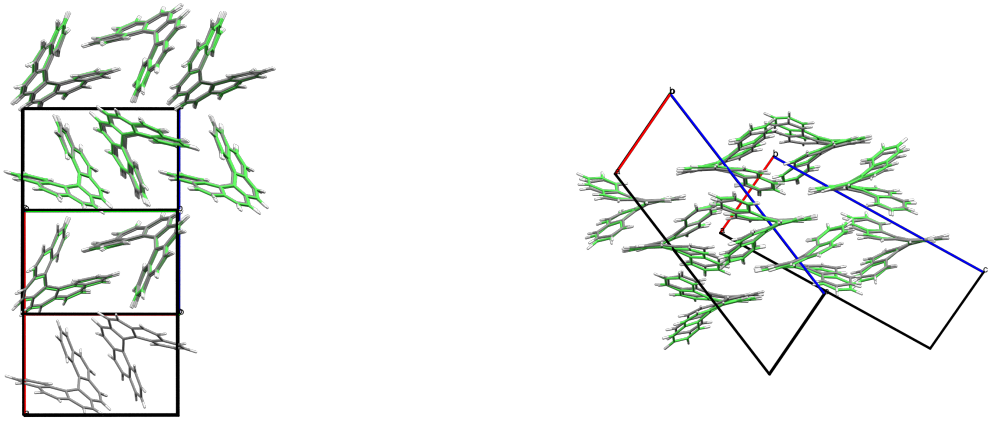
Equally, the set of lowest-lying [9]helicene crystal structures within 5 kJ mol^{-1} from the global minimum only comprises racemic crystalline forms. However, the proportion of enantiopure crystals within 20 kJ mol^{-1} moderately increases (7.5%). The racemic global minimum structure has a higher density than the racemic [8]helicene global minimum structure. There has been one experimentally observed X-ray structure deposited to the CSD (CSD reference code: QUJNEQ). However, this X-ray structure has three molecules in its asymmetric unit (Z'=3) and thus requires a much more challenging and computationally expensive higher dimensional CSP search, which is beyond of the scope of this work. The reported experimental structure of Z'=3 ($P2_1$ symmetry) is enantiopure.

The [10]helicene crystal-energy landscape appears particularly wide, meaning more degenerate crystalline forms are observed for a given lattice energy. The proportion of enantiopure crystalline forms within 20 kJ mol^{-1} slightly increases with 8.3% of the forms being enantiopure. Furthermore, the global minimum energy structure is enantiopure. The experimentally observed



(a) (b)

Figure 5.6: (a) Molecular overlay of the naphthalene X-ray (CSD reference code: NAPTHA18) and the lowest-lying predicted structure. For an overlay of 15 molecules, a RMSD of 0.609 Å was observed. (b) Molecular overlay of the [3]helicene X-ray and the second-lowest predicted structure. For an overlay of 15 molecules, a RMSD = 0.210 Å was observed. The experimental (predicted) structure is shown in grey (green).



(a)

Figure 5.7: (a) Molecular overlay of the [6]helicene X-ray (CSD reference code: HEXHEL) and the second-lowest predicted structure. For an overlay of 15 molecules, a RMSD = 0.210 Å was observed. (b) Molecular overlay of the [7]helicene X-ray (CSD reference code: HPTHEL) and the 47th lowest-lying predicted structure. For an overlay of 15 molecules, a RMSD of 0.315 Å was observed. The experimental (predicted) structure is shown in grey (green).



(a) (b)

Figure 5.8: (a) Molecular overlay of the [10]helicene X-ray (CSD reference code: THELIC) and the second-lowest predicted structure. For an overlay of 15 molecules, a RMSD of 0.308 Å was observed. (b) Molecular overlay of the [11]helicene X-ray (CSD reference code: UHELIC) and the sixth-lowest predicted structure. For an overlay of 15 molecules (at an increased cutoff of 30%), a RMSD of 0.687 Å was observed. The experimental (predicted) structure is shown in grey (green).

[10]helicene crystal structure (CSD reference code: THELIC) was successfully predicted as the second lowest-lying structure with a relative energy difference to the global minimum of 0.44 kJ mol^{-1} and a density of 1.25 g cm^{-3}. Similar to the [6]helicene crystal-energy landscape, the experimental polymorph is enantiopure, of $P2_1$ symmetry, and observed with a small energy difference to the global minimum structure.

[11]- and [12]helicene: For [11]helicene and [12]helicene, significantly fewer polymorphs are found (184 and 74 forms respectively, Figure 5.3a), *i.e.* the crystal-energy landscape is quite narrow. This could be due to their long helicene chain length, which might promote assembly into columns with the helices stacked on top of each other (along the central chiral axes) similar to a fibre or tubular arrangement. If there is a stronger direction towards a specific crystalline packing, this usually means that the potential energy surface (PES) is less flat. With respect to helicenes, this would signify that as the helicene approaches a longer spiral shape (n = 11, 12), the PES becomes steeper towards an all-parallel columnar-type packing.

The lowest-lying [11]helicene crystal structure is racemic and 5.2 kJ mol^{-1} away from the second-lowest lying structure, which is also racemic. The experimentally observed [11]helicene crystal structure (CSD reference code: UHELIC) was found to be the 6^{th} lowest-lying crystal structure (Table 5.2, ΔE = 8.02 kJ mol^{-1}, ρ = 1.27 g cm^{-3}). The experimentally observed X-ray structure is enantiopure ($P2_1$ symmetry). Likewise to the [6]- and [10]helicene crystal-energy landscapes, the experimental structure is enantiopure and of $P2_1$ symmetry. However, in the case of the [11]helicene crystal-energy landscape, the experimentally observed polymorph is not obtained with a low relative energy to the global minimum structure. However, comparing only enantiopure crystal structures across the CSP landscape, the experimentally observed crystal structure ends up as the lowest-lying enantiopure polymorph (Table 5.2). As the helicene length n increases, the energy barrier to interconversion from one enantiomer to another increases. Hence, the predicted crystals can be ranked separately with respect to their chirality within each CSP landscape. Thus, actually this is a completely successful prediction given that the molecules cannot invert chirality.

The [12]helicene crystal-energy landscape resembles the [11]helicene crystal-energy landscape with relatively few solid forms. The lowest-lying crystalline forms within 5 kJ mol^{-1} from

the global minimum are racemic and the energy gap between the lowest-lying form and the second-lowest form, which is also racemic is relatively high with 3.7 kJ mol^{-1}. The proportion of enantiopure crystals observed within 20 kJ mol^{-1} increases compared to [11]helicene (10.8%) with 15% of crystals being enantiopure.

5.3.1 Chirality Global Trends

Enantiomerisation describes the ability of the helicene to flip its molecular geometry, *e.g.* from the *M*-enantiomer to the *P*-enantiomer and *vice versa*. The activation energy barrier for the enantiomerisation of [n]helicenes, $\Delta G^{\dagger}$, increases exponentially with n from 4-9.[413] The enantiomerisation mechanism of [n]helicenes can range from a single-step mechanism (n = 3-7) to a complex multistep process ($n \geq 8$).[413] Hence, if one has an enantiopure experimental sample, one cannot necessarily obtain a racemic crystal structure. In other words, which polymorph has been obtained does not only relate to the CSP landscape, but rather what experimental techniques have been tested to growth crystals of multiple polymorphs. At high helicene length n (n = 11-12), the racemic structures tend to dominate the lowest-lying crystal structures (within 5 kJ mol^{-1}), whilst at shorter helicene length n (n = 4-7) enantiopure crystals are equally abundant within the set of lowest-lying crystal structures. Figure 5.3b shows that going from an almost planar structure ([3]helicene) to a helical shape with almost one full rotation ([6]helicene) the relative proportion of enantiopure crystals found within each landscape decreases with increasing helical pitch from [3]helicene to [6]helicene (Figure 5.3b). Then, the proportion stagnates from [6]- to [8]helicene and increases upon further increase of the n helicene length. The largest proportion of enantiopure crystals is observed for [12]helicene (15%). At [6]helicene, the aromatic terminal rings start overlapping, at [7]helicene the terminal rings lay on top of each other and the helicene length exceeds a full rotation. For [12]helicene nearly two full rotations are observed. An increase of the relative proportion of enantiopure crystals is observed when increasing the n helicene length beyond [8]helicene (Figure 5.3b). The increased proportion of enantiopure forms with increasing helicene length does not arise from chain-length specific changes of the molecular surface area. The single-molecule molecular surface area increases linearly with increasing helicene length n (Figure C.1). Thus, the molecular surface area is not a critical chain-length dependent property.[414]

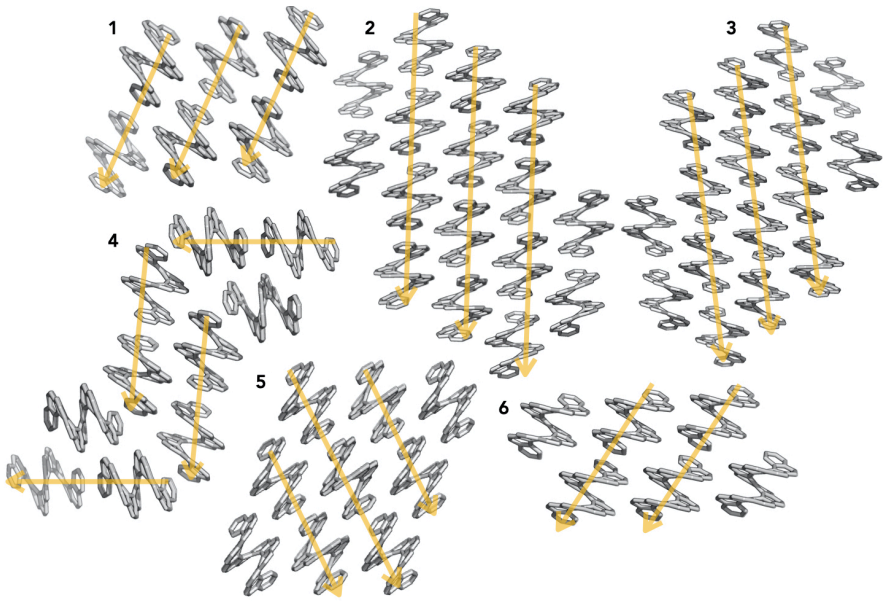


Figure 5.9: The lowest-lying structures within 7.2 kJ mol^{-1} from the global minimum all with their long helical axes arranged in parallel to one another. Uniquely, structure 6 is enantiopure and has homochiral columnar packing. All other packings are arranged in heterochiral staggered head-to-tail arrangements.

The fact that enantiopure forms are increasingly observed at higher helicene lengths might be attributed to the increasing spiral shape of the molecule exceeding the first full rotation, which might promote the helices having tubular head-to-tail solid-state arrangements. For example, the [11]helicene experimentally observed structure shows homochiral columnar packing. Furthermore, all the [12]helicene low-energy crystals within 7.2 kJ mol^{-1} of the global minimum have the helicenes oriented with their long helical axes arranged in parallel to one another (Figure 5.9). These structures have some local order maximising the potential chiroptical response of the material by having the molecular dipoles arranged in parallel. Only one packing has local columnar ordering (structure 4), but no global parallel arrangement.

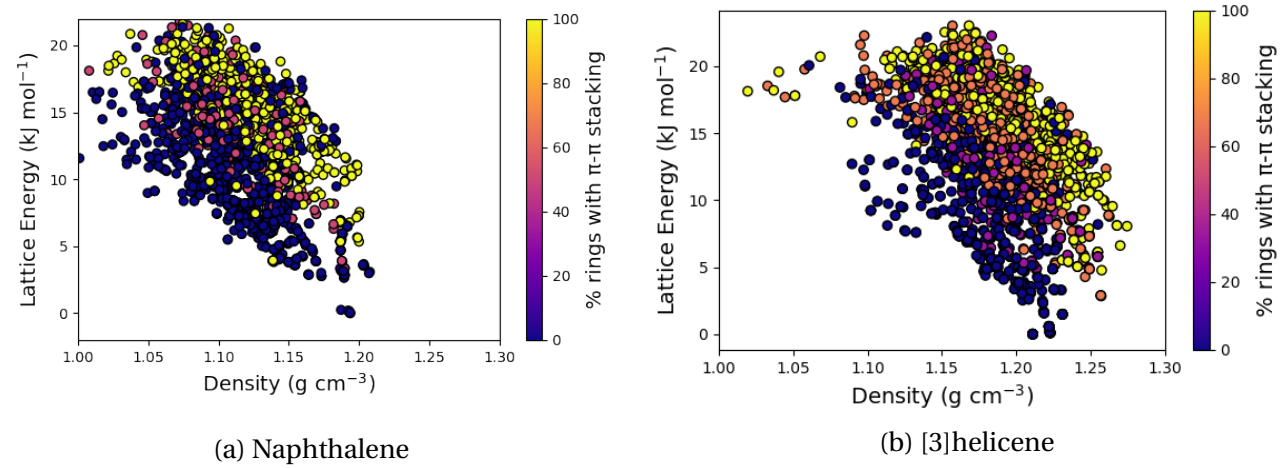
5.3.2 Crystal Stacking Interactions Behaviour Across the [n]helicene Polymorphic Landscapes

Figures 5.10 and 5.11 highlight that the degree of π-π stacking, quantified by the percentage of aromatic rings of the helicene backbone involved in π-π stacking interactions, is directly correlated to the crystal lattice energy. Hence, π-π stacking interactions are predominantly observed across high-energy crystals and more rarely observed for low-energy crystals within each crystal-energy

n	corr(π-π, C-H$\cdots\pi$)	corr(Energy, π-π)	corr(Energy, C-H$\cdots\pi$)	corr(ρ, π-π)	corr(ρ, C-H$\cdots\pi$)
2	-0.19	0.40	-0.48	0.14	0.14
3	-0.44	0.53	-0.63	0.12	0.04
4	-0.21	0.31	-0.36	-0.02	0.05
5	-0.20	0.20	-0.38	-0.04	0.08
6	-0.26	0.28	-0.29	0.07	0.08
7	-0.26	0.19	-0.18	0.03	0.00
8	-0.34	0.19	-0.13	0.05	0.01
9	-0.43	0.22	-0.09	0.03	-0.06
10	-0.41	0.44	-0.34	0.23	-0.02
11	-0.26	0.30	0.12	0.05	-0.29
12	-0.25	0.18	-0.09	0.08	-0.07

Table 5.3: Table listing the Pearson correlation scores (corr) between the degree of π-π stacking and C-H$\cdots\pi$ stacking, crystal lattice energy and density for varying helicene length n. Correlations from 0.30 to 0.50 (-0.30 to -0.50) are usually considered low, 0.50 to 0.70 (-0.50 to -0.70) moderate, and 0.70 to 1.00 (-0.70 to -1.00) high.[415]

landscape. Likewise, low-energy crystals within each crystal-energy landscape tend to more frequently exhibit C-H$\cdots\pi$ stacking than high-energy crystals of the same landscape. This weak to moderate correlation could be quantified using Pearson correlation coefficients, which are presented in Table 5.3. The energy and the degree of π-π stacking and the energy and the degree of C-H$\cdots\pi$ stacking correlation is strongest within the [3]helicene crystal-energy landscape with Pearson correlation coefficients of 0.53 and -0.63 respectively. Thus, the correlation between the lattice energy and the degree of crystal stacking interactions is best illustrated for the naphthalene and [3]helicene molecular crystal-energy landscapes (Figure 5.12a). The naphthalene and the [3]helicene crystal-energy landscape show that the C-H$\cdots\pi$ stacking proportion decreases as the π-π stacking proportion increases as one moves away from the global minimum (*i.e.* from left to right in Figure 5.12a). Naphthalene and [3]helicene are planar molecules. This could be the reason why this relationship is best observed here. As the helicene length increases, the degree of C-H$\cdots\pi$ stacking generally decreases (except for [10]helicene), rendering this correlation less obvious (Figure 5.12b). Due to increased steric hindrance in helicenes with a larger number of *ortho*-fused benzene rings, by nature only a smaller fraction of the benzene rings can be simultaneously involved in either type of stacking interactions. This hypothesis is supported by the maximum percentage of rings involved in either stacking interaction (%) decreasing from 100% ([5]helicene) to 67% ([12]helicene, Figure 5.12b).



(a) Naphthalene

(b) [3]helicene

(c) [4]helicene

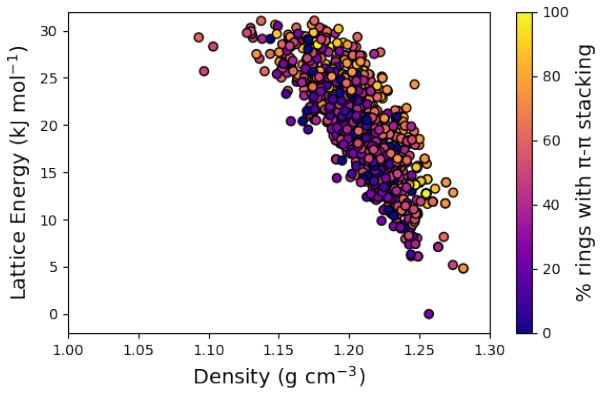
(d) [5]helicene

(e) [6]helicene

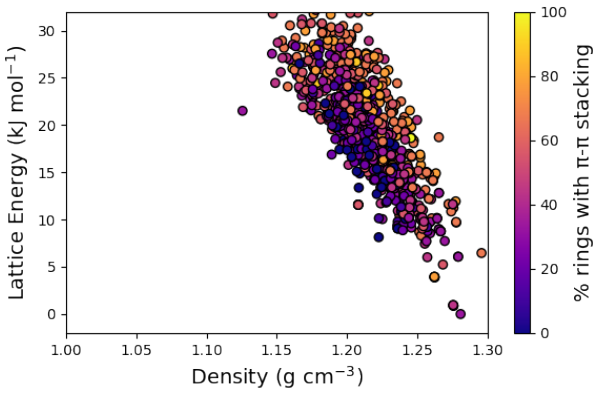
(f) [7]helicene

Figure 5.10: Crystal-energy landscapes for naphthalene and [3]-[7]helicenes. Each crystal is coloured by the degree of π-π stacking of the single molecule in the asymmetric unit. 100% (0%) denotes every (no) benzene ring of the respective backbone is involved in π-π stacking. The lattice energy is given in relative terms with respect to the global minimum to allow for easy comparison.

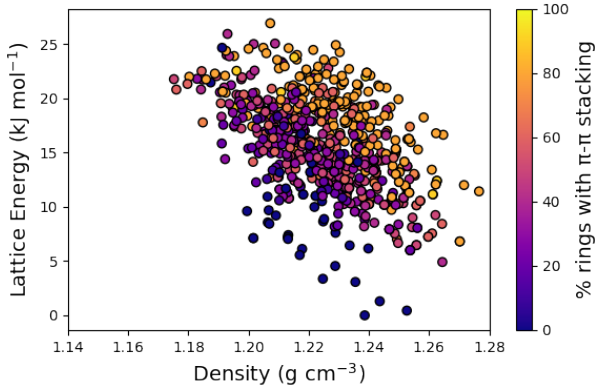
For a subset of the CSP landscapes, an additional positive correlation between density and the degree of parallel π-π stacking was found. In particular, the density and degree of π-π stacking across the [10]helicene crystal-energy landscape were found to be very weakly positively corre-



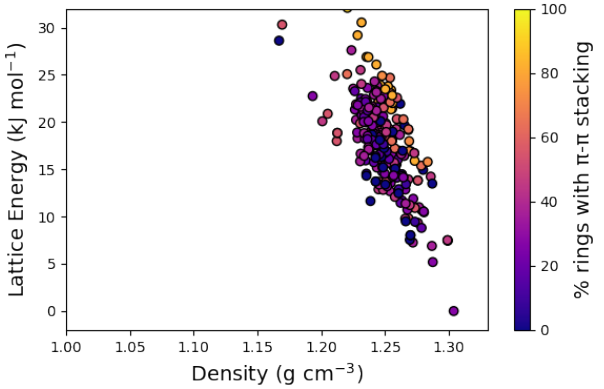
(a) [8]helicene]



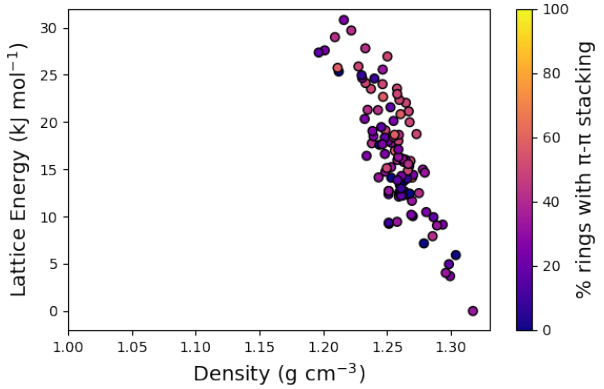
(b) [9]helicene



(c) [10]helicene



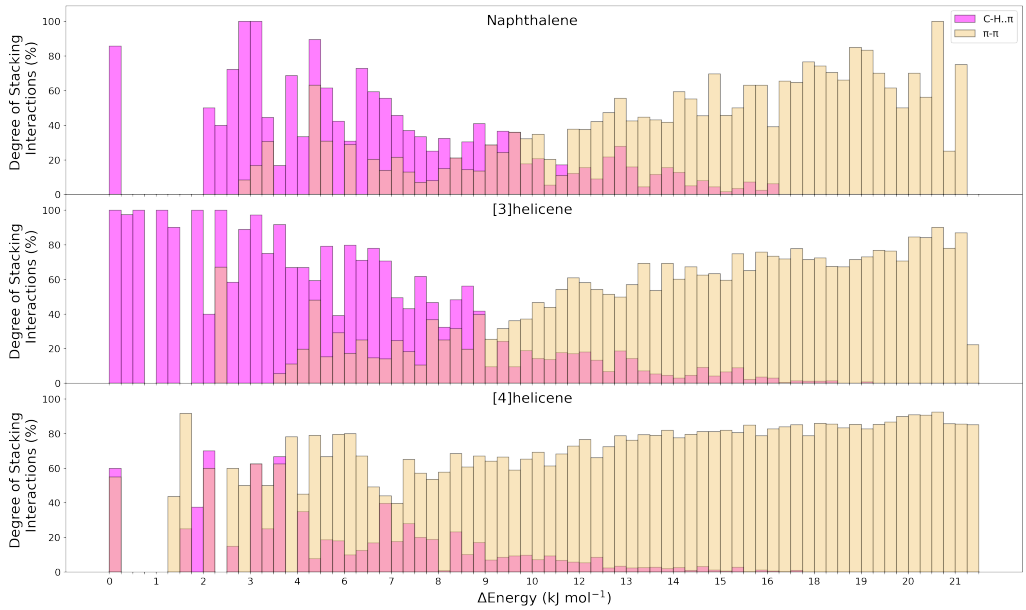
(d) [11]helicene



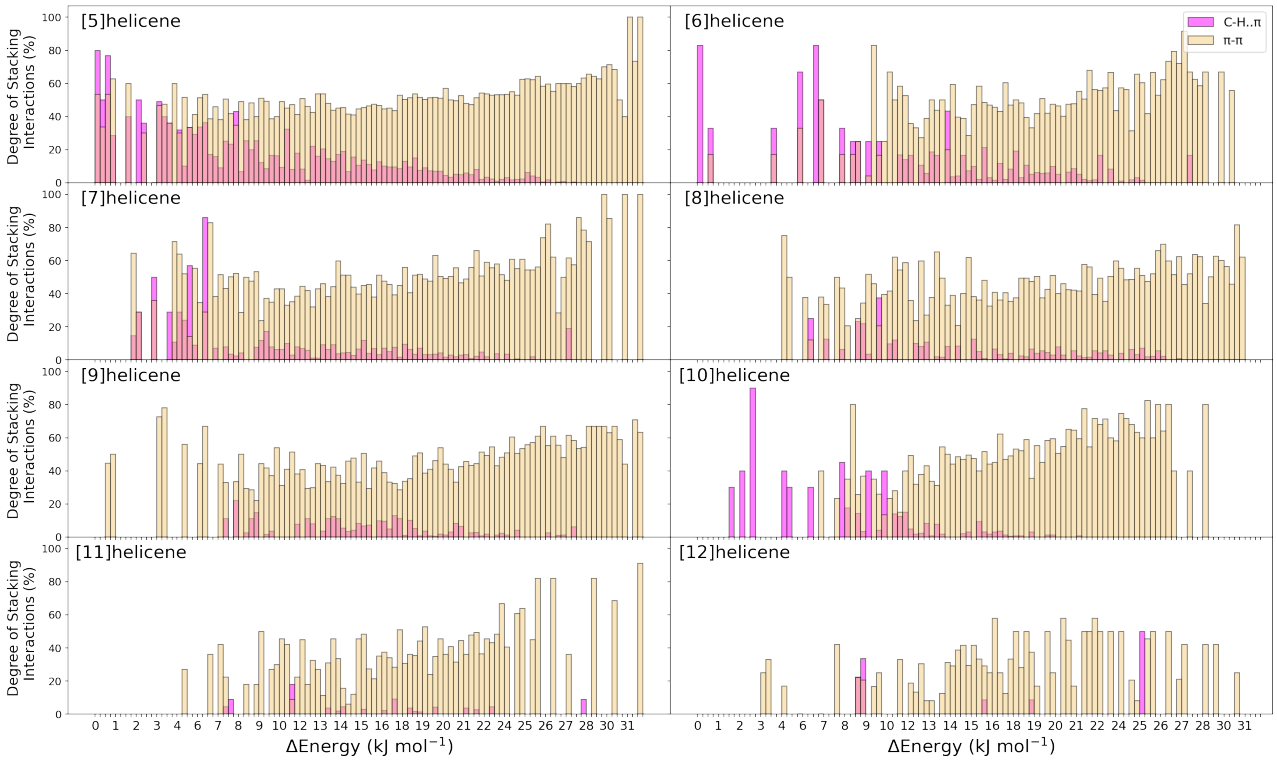
(e) [12]helicene

Figure 5.11: Crystal-energy landscapes for [8]-[12]helicenes. Each crystal is coloured by the degree of π-π stacking of the single molecule in the asymmetric unit. 100% (0%) denotes every (no) benzene ring of the respective backbone is involved in π-π stacking. The lattice energy is given in relative terms with respect to the global minimum to allow for easy comparison.

lated (Pearson correlation coefficient: 0.23). Similarly, the density and degree of C-H$\cdots\pi$ stacking within the [11]helicene crystal-energy landscape were found to be weakly negatively correlated (Pearson correlation coefficient: -0.29). If one examines the [10]helicene crystal-energy landscape



(a) Histogram of the degree of π-π and C-H$\cdots\pi$ stacking interactions for naphthalene, [3]- and [4]helicene binned by the relative energy difference to the global minimum energy structure at constant 0.30 kJ mol^{-1} binwidth.



(b) Histogram of the degree of π-π and C-H$\cdots\pi$ stacking interactions for [5]-[12]helicenes binned by the relative energy difference to the global minimum energy structure at constant 0.30 kJ mol^{-1} binwidth.

Figure 5.12

(Figure 5.11c), for a set of degenerate crystals, crystals with a high degree of π-π stacking would more likely be found at higher density and crystals with no π-π stacking would be found at lower density. This density right-shift with respect to the degree of π-π stacking as observed in the case of [10]helicene can be rationalised by the need for close contacts between adjacent aromatic rings to form π-π stacking interactions. However, the degree of π-π stacking is not as correlated to the crystal density as the crystal lattice energy and hence this correlation is mainly observed for [10]helicene and [11]helicene.

As the helical length is increased from [4]helicene to [5]helicene, there is a leap in helical pitch of the backbone, which provokes a decrease of the percentage of π-π stacking interactions. Whilst for the [4]helicene crystal-energy landscape, the majority of its polymorphs have all benzene rings on the backbone involved in π-π stacking (100%, yellow points), the [5]helicene crystal-energy landscape has very few examples of polymorphs with the entire backbone of the molecule involved in π-π stacking interactions (Figure 5.10).

The propensity for π-π stacking interactions has been found to be inversely correlated to the propensity for C-H$\cdots\pi$ stacking interactions (largest Pearson correlation coefficient: -0.44). Thus, crystals that have the lowest degree of π-π stacking tend to be the most likely to display a high degree of C-H$\cdots\pi$ stacking and *vice versa*. If one examines regions of highest degree of C-H$\cdots\pi$ stacking (Figure 5.13 and 5.14), they are the regions, which show the lowest π-π stacking in Figures 5.10-5.11. Consequently, the highest degree of C-H$\cdots\pi$ stacking is found for low-lying crystal structures and for a set of degenerate crystals, the crystal at lower density is more prone to displaying C-H$\cdots\pi$ stacking compared to the higher density structure. High-accuracy quantum mechanical studies of π-π interaction energies in benzene dimers at CCSD(T)/aug-cc-pVDZ* and estimated CCSD(T)/aug-cc-pVQZ* level in the complete basis set (CBS) limit revealed that the C-H$\cdots\pi$ (T-shaped) and the parallel displaced π-π interactions are nearly equivalent, *i.e.* the interaction energies for C-H$\cdots\pi$-stacking and parallel displaced π-π stacking are -2.24 kcal mol^{-1} and -2.22 kcal mol^{-1} (CCSD(T)/aug-cc-pVDZ*) and -2.74 kcal mol^{-1} and -2.78 kcal mol^{-1} (estimated CCSD(T)/aug-cc-pVQZ* CBS limit) respectively.[416] These calculations were made for benzene dimers, which are not identical to nonplanar π-π interactions as observed in the helicenes. Amemiya *et al.* reported that nonplanar π-π interactions in helicenes appeared stronger than

those across planar π-systems.[210] In conclusion, the density and energy conditions, *e.g.* low-energy and low-density, which are most suitable for having a high degree of C-H$\cdots\pi$ stacking are at the same time extremely unsuitable for having a high degree of π-π stacking and *vice versa*. Presumably, this is because both stacking interactions compete against each other, *i.e.* once one stacking interaction is present, it is more difficult (if not impossible) to get the other in these rigid molecules. However, if the crystal lattice energies are similar, it implies that something else favours one crystal stacking interaction being dominant over another at lower energy structures. It could potentially be that the packing type dominating the lower-lying polymorphs allows an overall denser packing and thus likely picks up higher intermolecular interactions.

The observation that π-π and C-H$\cdots\pi$ stacking are correlated to the lattice energy and inversely correlated to each other is independent of the chirality of the crystal, *i.e.* there is no correlation between the degree of π-π or C-H$\cdots\pi$ stacking and the chirality of the crystal structure. Thus, all observations remain the same regardless of whether racemic and enantiopure crystal structures within each CSP landscape are analysed separately.

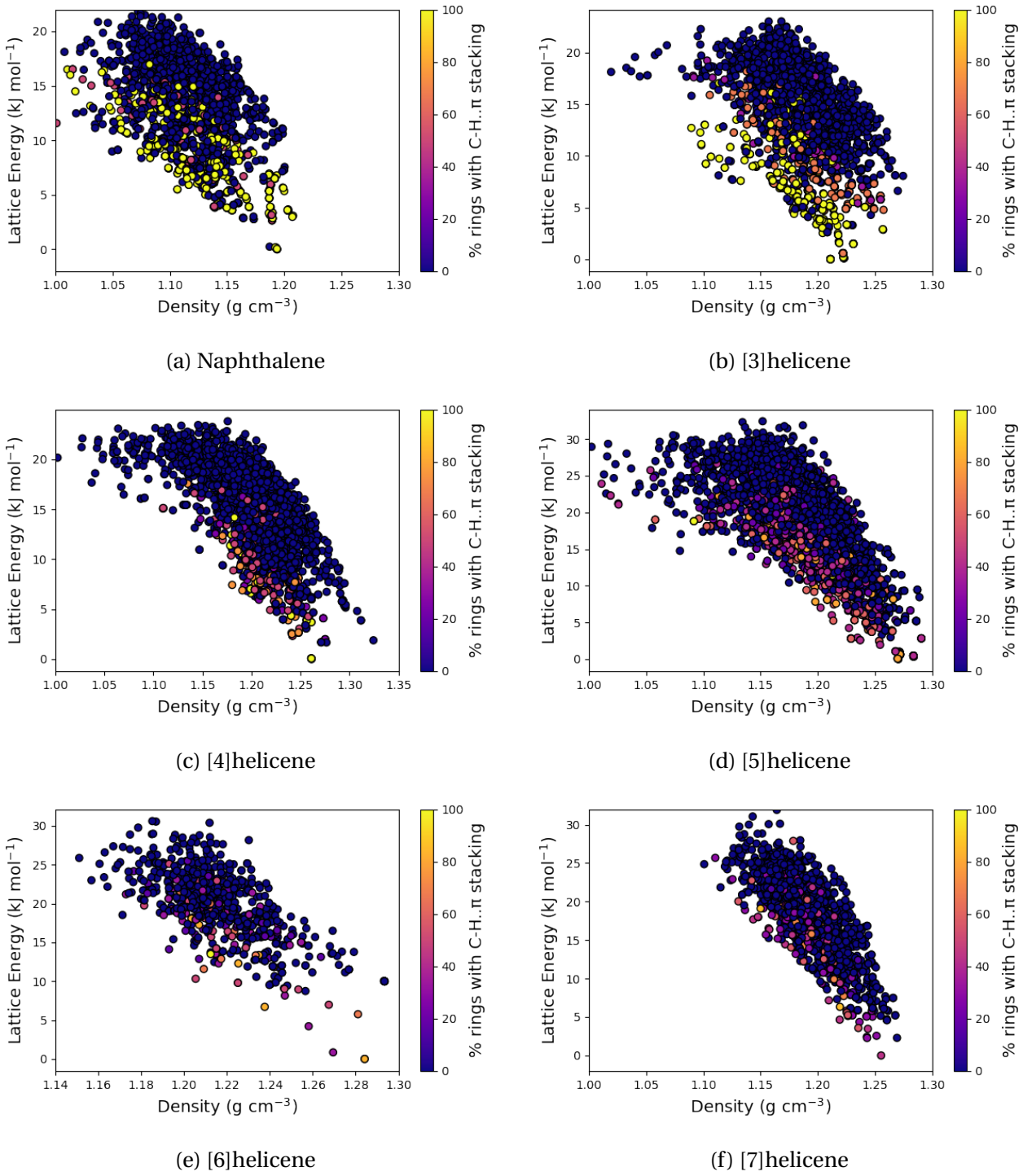
Furthermore, as the helical chain length increases, the degree of C-H$\cdots\pi$ stacking decreases across the CSP landscapes. A reason for this decrease in C-H$\cdots\pi$ stacking with increasing helicene length might be steric hindrance. As the helicene length increases, benzene rings near the centre of the helicene backbone will become increasingly shielded by benzene rings closer to the terminal rings of the helix. Thus, as a result of the helical shape of helicenes exceeding a full rotation (n>6), some 'inner' benzene rings become increasingly hindered by 'outer' benzene rings, naturally impeding every ring from being involved in C-H$\cdots\pi$ stacking.

Another reason for the decrease in C-H$\cdots\pi$ stacking interactions with increasing helicene length is that at higher helicene lengths (*e.g.* [11]-,[12]helicene) the lowest-lying crystals are dominated by the helicenes all being oriented parallel to each other, sometimes in chiral columnar packing or other types of parallel packing with helicenes not head-to-tail to each other, but still pointing in the same direction (see previous example of parallel columnar-type packing, Figure 5.9). Thus, the hypothesis is that as the number of *ortho*-fused benzene rings reaches a certain length and a screw-like spiral shape, it is more energetically favourable to have all helices in the molecular crystal oriented parallel to each other.

This hypothesis is supported by the observation that in both crystal-energy landscapes, [11]- and [12]helicene, out of the 9 lowest-lying crystal structures in each landscape, 8 are characterised by parallel fibre-like columnar packing. The only structure which is not all parallel (structure 4 in the case of [12]helicene) still has 11 out of 15 molecules within the neighbouring shell arranged parallel to each other. Therefore, in order to quantify the initial hypothesis that with increasing helicene length, beyond a first full rotation, the parallel packing of the helices dominates the crystal-energy landscape, a feature was designed to quantify the degree of parallel packing of the molecules in the crystalline solid. Hence, for each crystal structure the molecular orientation of the central molecule with all its 15 nearest neighbour molecules was computed. If all 15 nearest molecules are parallel to the central molecule (15/15), the degree of parallel molecular orientation was 100%. Likewise, if 11/15 molecules were arranged in parallel, the degree of parallel molecular orientation was set to 73% and so on. The results in Figures 5.15 and 5.16 clearly indicate that for the longest helicenes (n = 11, 12) the lowest-lying crystal structures (including the global minimum) have an extremely high degree (100%) of molecules oriented in parallel. This further supports the hypothesis that if the helical shape elongates, the parallel columnar packing becomes increasingly stable compared to any other type of packing. At the same time, for the perfectly parallel naphthalene molecule (n = 2), an all parallel molecular orientation (100%) is predominantly associated with high-energy crystals. For a helicene of intermediate length, *e.g.*, n = 6, an all parallel molecular orientation (100%) is equally not associated with the lowest-lying crystal structures. As illustrated in Figure 5.7a and discussed in previous chapters, the experimentally observed crystal structure (CSD reference code: HEXHEL) does not have all molecules oriented in parallel to each other. Looking at the occurrence of all parallel molecular orientations with respect to density, it is worth noting that across the [4]-,[6]-,[7],[10]-,[11]- and [12]helicene crystal-energy landscapes, the most dense crystal structures are characterised by all molecules oriented in an all parallel fashion (Figures 5.15 and 5.16).

In summary, one can conclude that the longer helicene length drives columnar packing, *i.e.* crystalline packings with molecules oriented in an all-parallel fashion (*e.g.*, chiral columnar, *etc.*) are increasingly observed across low-energy crystal structures. In chiral electronics, a major research goal is to obtain chiral molecular crystals, which maximise the chiroptical response and

is often associated with chiral columnar packing. Thus, knowing that an increased axial helicene length does promote the likelihood of observing chiral columnar crystals can be of major interest to chiroptical research.



(a) Naphthalene

(b) [3]helicene

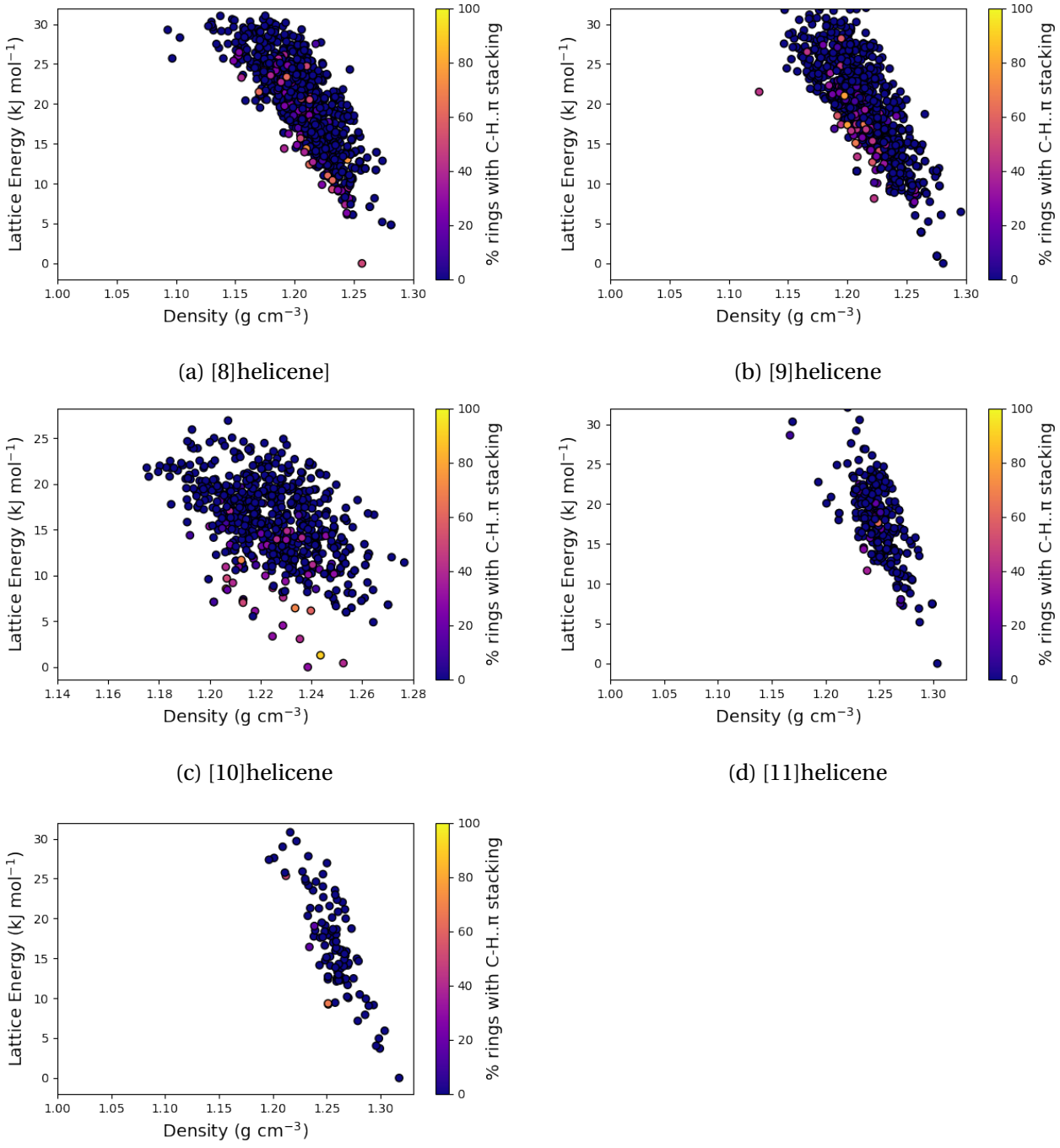
(c) [4]helicene

(d) [5]helicene

(e) [6]helicene

(f) [7]helicene

Figure 5.13: Crystal-energy landscapes for Naphthalene and [3]-[6]helicenes. Each crystal is coloured by the degree of C-H$\cdots\pi$ stacking of the single molecule in the asymmetric unit. 100% (0%) denotes every (no) benzene ring of the respective backbone is involved in C-H$\cdots\pi$ stacking. The lattice energy is given in relative terms with respect to the global minimum to allow for easy comparison.



(a) [8]helicene]

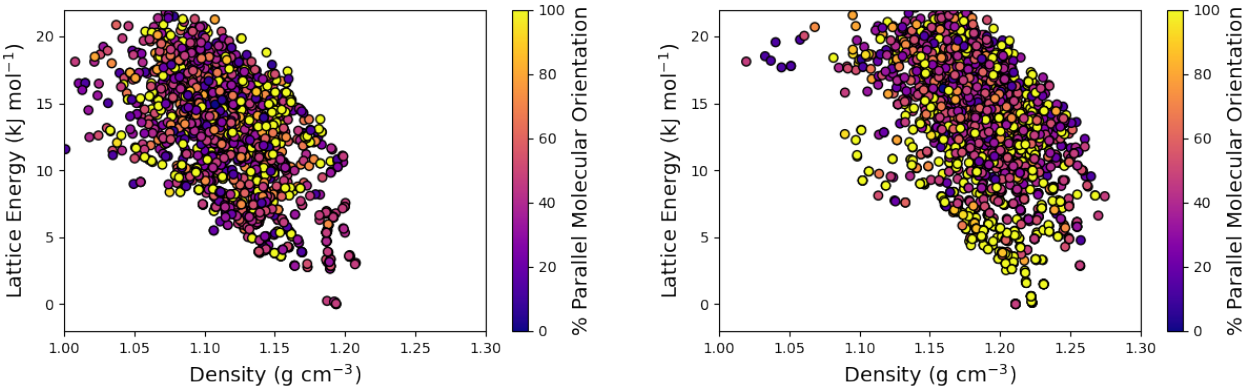
(b) [9]helicene

(c) [10]helicene

(d) [11]helicene

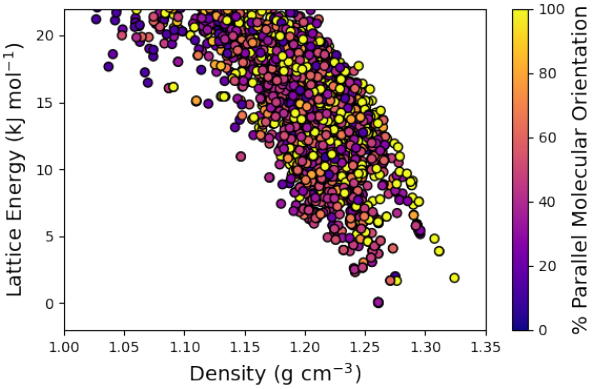
(e) [12]helicene

Figure 5.14: Crystal-energy landscapes for [8]-[12]helicene. Each crystal is coloured by the degree of C-H$\cdots\pi$ stacking of the single molecule in the asymmetric unit. 100% (0%) denotes every (no) benzene ring of the respective backbone is involved in C-H$\cdots\pi$ stacking. The lattice energy is given in relative terms with respect to the global minimum to allow for easy comparison.

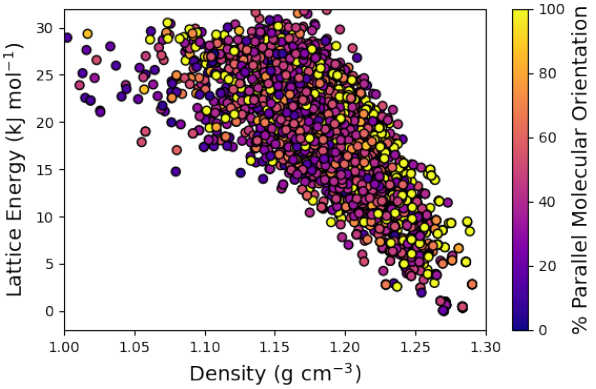


(a) Naphthalene

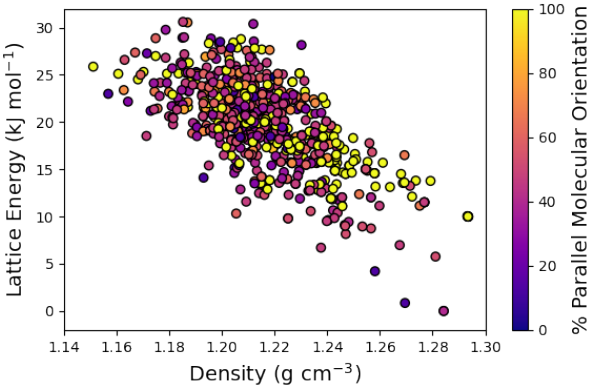
(b) [3]helicene



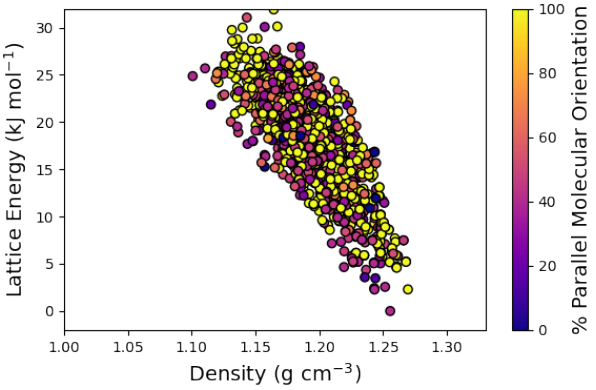
(c) [4]helicene



(d) [5]helicene

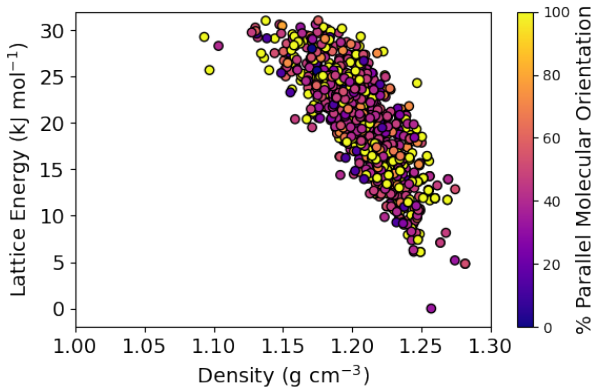


(e) [6]helicene

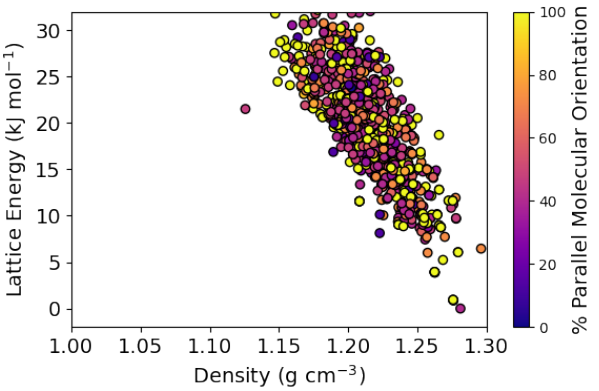


(f) [7]helicene

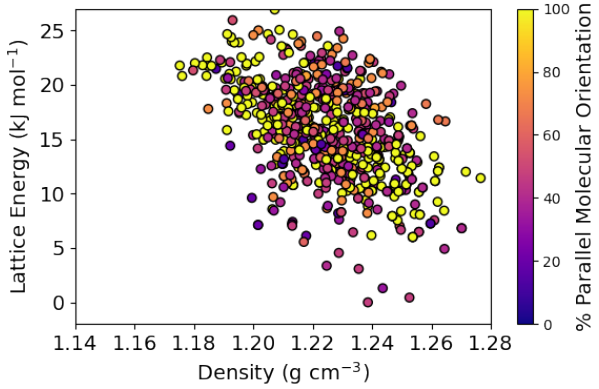
Figure 5.15: Crystal-energy landscapes for naphthalene and [3]-[7]helicene. Each crystal is coloured by the proportion of molecules oriented parallel to the molecule in the asymmetric unit. 100% (0%) denotes every (no) molecule is oriented in parallel to the central molecule in the neighbouring shell. The lattice energy is given in relative terms with respect to the global minimum to allow for easy comparison.



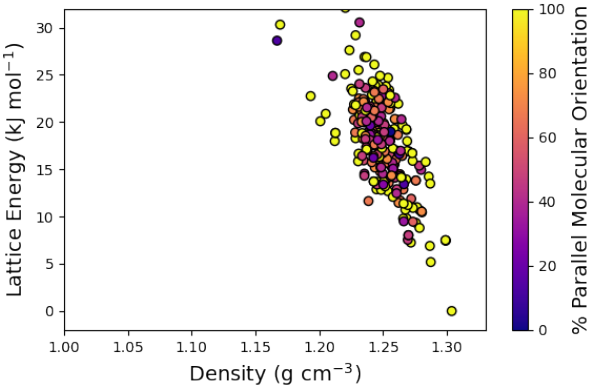
(a) [8]helicene]



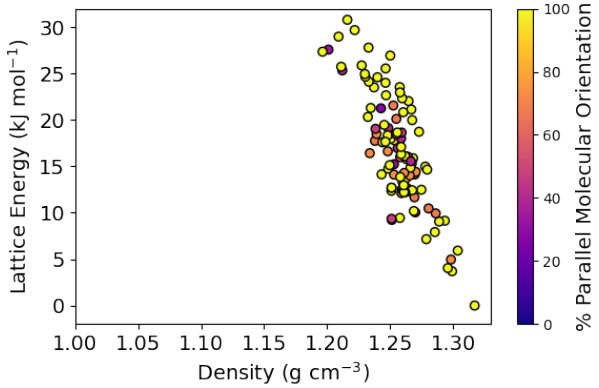
(b) [9]helicene



(c) [10]helicene



(d) [11]helicene



(e) [12]helicene

Figure 5.16: Crystal-energy landscapes for [8]-[12]helicene. Each crystal is coloured by the proportion of molecules oriented parallel to the molecule in the asymmetric unit. 100% (0%) denotes every (no) molecule is oriented in parallel to the central molecule in the neighbouring shell. The lattice energy is given in relative terms with respect to the global minimum to allow for easy comparison.

5.4 Conclusions

The change in crystal stacking interactions with varying [n]helicene length, n, has been successfully demonstrated by the use of a per benzene ring derived score to describe the involvement of the molecule in the asymmetric unit in crystal stacking interactions. Furthermore, the degree of parallel packing across the molecules in the crystalline solid was quantified across all crystal-energy landscapes. It was found that the frequency of crystals containing C-H$\cdots\pi$ stacking decreases with increasing helicene length. The relative proportion of benzene rings per molecule involved in this type of interaction equally decreases with increasing helicene length. C-H$\cdots\pi$ stacking is generally less observed than π-π stacking. Furthermore, it was found that molecular crystals with a high relative lattice energy, sometimes also an increased density, tend to exhibit the highest degree of π-π stacking for across the respective CSP landscape. As the helicene length increases beyond a full rotation, the benzene rings near the ends of the molecule increasingly shield the benzene rings near the centre of the helicene molecule and full involvement of all benzene rings in shifted parallel π-π stacking becomes increasingly impossible. Thus, the maximum degree of π-π stacking is lowered as the chain length increases from initially 100% (bright yellow) to ~80% (orange) and finally ~60% for [12]helicene.

It was further demonstrated that both types of crystal stackings are inversely correlated, meaning that if conditions for one stacking type are enhanced, the other stacking type is disfavoured. This is probably due to steric constraints rendering the interactions mutually exclusive. Either the molecules are oriented parallel to each other, allowing π-π stacking or rather orthogonal allowing C-H$\cdots\pi$ interactions.

Notably less polymorphs were found with increasing helicene length indicative of the fact that the longer the spiral shape of the molecule, the steeper the potential energy surface, *i.e.* the smaller the configurational space of possible stable solid-state arrangements. Moreover, it was shown that for long helicenes (n = 11, 12) the all-parallel columnar packing dominates the configurational space and especially across the lowest-lying crystal structures. This can probably be in part ascribed to the nature of the two stacking interactions. More π-π stacking interactions are simultaneously sterically feasible if all molecules are oriented parallel to each other than C-H$\cdots\pi$ interac-

tions sterically possible, if the helicenes are oriented orthogonal to each other. Furthermore, experimentalists have suggested that nonplanar π-π interactions across helicenes could be stronger than π-π interactions across planar systems.[210]

It was found that the relative proportion of enantiopure crystal structures within each landscape initially shrinks going from [3]helicene to [6]helicene, mainly stagnates from [6]- to [9]helicene and it increases thereafter up to [12]helicene. The increase in relative proportion of enantiopure crystals beyond $n > 9$ might be associated with the increasingly spiral molecular geometry, which favours parallel columnar-type stackings, one of which is the homochiral columnar packing encountered in enantiopure crystals. This hypothesis is further supported by experimental findings, which have found a tendency for homochiral helicene pairs to form more stable dimers than heterochiral pairs.[210] Previous work in Chapter 3 on the [6]helicene recurring packing motifs has shown that for a helicene of intermediate length such as [6]helicene, other highly stable configurations with a high degree of π-stacking (*e.g.*, heterochiral back-to-back dimer) can be formed, which is why the proportion of enantiopure crystals might be reduced. However, as the helicene length increases, the spiral columnar shape of the molecule reduces the number of stable recurring packing motifs and enhances a parallel pairwise arrangement.

Now that the general crystal packing trends across a family of polyaromatic molecules have been analysed and crystal packing features for thousands of crystal structures have been computed, this data can be used to guide further research into meaningful crystal similarity measures and clustering algorithms to more intelligently screen the polymorphic space for crystalline forms with desired features, such as a high degree of parallel packing, π-π stacking or C-H$\cdots\pi$ stacking. For example, a high degree of π-π stacking can give rise to improved charge-transport properties,[417–420] enhanced magnetic responses of the molecular material[421] and drive molecular assembly, *e.g.*, the formation of one-dimensional nano-structures.[422,423] Previous chapters (Chapter 3 and 4) have illustrated how packing motifs with a high degree of π-π stacking interactions, *e.g.*, the translational, herringbone or back-to-back motif tremendously influence charge-transport properties. Some packing motifs with no degree of π-stacking act as kinetic traps in the charge hopping process and can impede transport. Hence, improving the understanding of the underlying factors driving the occurrence of certain stacking interactions can direct electronic molecular materials

design. Furthermore, the degree of π-π stacking can influence optical and electrochemical properties.[424] For example, Hong *et al.* reported a strong chiroptical response for helicenes with homochiral π-stacked molecular assembly.[425] Another promising route would be to use the derived features for generative models or inverse design approaches to more intelligently leverage polymorphism for the development of bespoke functional materials. For example, the crystal-energy landscape could be intelligently screened for a solid form with specific crystal packing interactions which maximise the material's target property (*e.g.*, electronic, magnetic or chiroptical).

The comprehensive study of a series of [n]helicenes with potential applications as OSC is particularly challenging, because the measurable target property, *e.g.*, mobility, is extremely sensitive to molecular assembly, which again is hyper-sensitive to changes to the single molecule. However, here consistent trends in crystal packing interactions dominating molecular assembly across molecular-level changes have been identified. Being able to understand the subtleties of molecular assembly a little more, especially how paths of π-π stacking across crystalline solids are likely to be found, is of major interest to the OSC community and to understanding how polymorphism can be leveraged for high-performance OSC material design. Furthermore, being able to group the crystal-energy landscape into crystalline solids of similar crystal packing interactions allows for efficient approximations of the vast configurational space into major groups, which can accelerate future screening-based approaches for high-performance OSC.

Chapter 6

Conclusions

The work presented consisted of exploring polymorphism, chirality, and charge transport in chiral organic semiconducting molecular materials, using established methods as well as using custom-developed analysis routes.

In Chapter 3, crystal structure prediction and Marcus theory were used to predict microscopic charge-transport hopping rates for each molecular crystal across all lowest-lying polymorphs of each of the two symmetrically difunctionalised [6]helicenes (2,2'-dicyano- and 2,2'-diethynyl-[6]helicene). All experimentally observed crystal structures were successfully predicted using crystal structure prediction (CSP). Each 2,2'-difunctionalised helicene can arrange in homochiral columnar packing, which have not been previously studied for their suitability for charge transport. The range of solid-state arrangements were analysed on the basis of previous research done for the unsubstituted [6]helicene. The crystal packing was analysed using discrete motifs to identify recurring structural motifs initially observed in the [6]helicene polymorphic landscape. Given the limitations of this approach, a kernel-based approach for assessing the similarity of crystalline solids was presented. The approach proved successful at identifying duplicate packings and at picking up local features (*e.g.*, high degree of parallel π-π packing), yet unsuccessful at identifying more subjective features such as chiral columnar packing. It was found that 2,2'-diethynyl-[6]helicene appears largely unsuitable for both hole and electron transport and that the terminal difunctionalisation marks no significant improvement in charge transport properties compared to [6]helicene. However, in the case of 2,2'-dicyano-[6]helicene, two solid forms with significantly

improved hole transport properties were found, one of which was the experimentally observed form with homochiral columnar packing. Chiral columns did not provide consistently improved charge transport hopping pathways. However, often high-transporting crystals arranged in chiral columns. In the case of 2,2'-dicyano-[6]helicene, it was computationally predicted that hole transport could be enhanced by over 200% if an additional solid form (structure 6) to the experimentally observed structure (CSD reference code: UJOCOL) was obtained via polymorph screening procedures.

In Chapter 4, a high-throughput dimer-based screening approach was developed and applied to 1344 side-group functionalised [6]helicene molecules which were screened for their potential suitability as high-performance molecular organic semiconductor (OSC) materials. The dimer-based screening approach leveraged knowledge regarding the ubiquity of the translational motif in previously studied [6]helicene systems and terminally functionalised [6]helicenes observed in the Cambridge Structural Database (CSD) to make a simplified one-dimensional (1D) informed guess about the crystalline environment.

The computational screening predicted the fluorine side-group functionalisation to be best for maximising electron transport across the material. Furthermore, two outlier molecules with surprisingly high charge-transport parameters, 6-tricetylene- and 7,10-dialdehyde-[6]helicene, were identified by the screening. Moreover, the screening revealed that particular side-group positions are consistently unfavourable for charge-transport across the translational motif (*e.g.*, 1,1'-disubstitution). For two fluorine molecules (mono- and disubstituted) as well as the two outlier molecules, a full CSP search followed by prediction of charge-transfer rates using Marcus theory across all lowest-lying crystal structures within 10 kJ mol^{-1} from the global minimum was performed. This in-depth analysis allowed for validation of the dimer-based screening approach, confirming the hypothesis that the translational motif remained a frequently observed motif across all landscapes, observed in 30-100% of the lowest-lying structures. Furthermore, it was shown that the intermolecular separations and orientations at which the translational motif was predicted to be observed matched the crystalline environment. The CSP of 13-fluoro- and 4,13-difluoro-[6]helicene following the dimer-based screening results revealed that fluorination can yield to consistently improved transfer integrals and up to 200% increased charge-carrier mobilities (4,13-

difluoro-[6]helicene) across all crystalline packings containing the translational motif. Nonetheless, the use of a frequently occurring dimer to represent the entire crystalline environment as part of a screening remains a large simplification understating the broad configurational space and the complex subtleties of molecular assembly.

In Chapter 5, helicene research expands beyond the [6]helicene backbone and helicenes with varying chain length n (n = 3-12) were computationally investigated. For each helicene, geometric features were analysed, the crystal-energy landscape was predicted and experimentally observed polymorphs were matched where possible. Using a PYTHON module specifically designed as part of this thesis, CRYSTACK, the crystal stacking interactions were quantified to assess the degree of π-π and C-H$\cdots\pi$ stacking per molecule across each of the hundreds to thousands of crystal structures in each CSP landscape. It was found that π-π stacking interactions are frequently observed in higher-energy crystals, whilst C-H$\cdots\pi$ stacking is frequently observed across low-lying crystal structures within each landscape. For some CSP landscapes, even a tendency for the degree of π-π stacking to increase with increasing density (close packing, [10]helicene) was observed. The maximum proportion of benzene rings involved in π-π stacking decreases with increasing helicene length and chiral pitch from planar naphthalene (100%) to [10]helicene (80%) and [12]helicene (50%). Furthermore, it was found that the proportion of helicenes stacked with their long axes parallel to each other across the lowest-lying crystal structures increases upon increasing helicene length. Identifying helicenes which assemble in helicoidal fibres is of fundamental importance to experimental chemists due to their enhanced and unique chiroptical properties.[191] In particular, helicenes arranged in homochiral columns show a very intense circular dichroism response, a huge optical rotation, and a large second-order nonlinear optical response.[191] Therefore, for planar naphthalene the lowest-lying structures have a low proportion of parallel packing, whilst for [11]- and [12]helicene nearly all molecules across the lowest-lying structures are oriented perfectly parallel, presumably favouring chiral columnar packing arrangements due to their spiral shape. The here-fore developed CRYSTACK analysis tool is generalisable and could be applied to any other group of polyaromatic compounds (*e.g.*, acenes) to analyse the crystal stacking interactions. Being able to group the configurational space into degrees of stacking interactions as performed here allows for the development of future intelligent crystal-search algorithms with tar-

geted crystal packing interactions and thus bespoke supramolecular properties (*e.g.*, electronic, chiroptical). The charge-mobility results presented in this thesis are extremely sensitive to the interatomic distances. Hence, areas for improvement include the computation of charge-mobility on more accurate DFT-D relaxed crystal structures, to dynamically model the change in transfer integral and to account for the degree of electron-phonon coupling by assessing the dominant phonon modes. I have forwarded CSP results from Chapter 3 to the Ruggiero group (University of Vermont) for ongoing analysis on the degree of electron-phonon coupling. Furthermore, the 2,2'-dicyano-[6]helicene results from Chapter 3 have been combined with thin-film experiments by Dr Jess Wade based on the 2,2'-dicyano-[6]helicene crystals grown by Francesco Salerno (Fuchter group, ICL) to understand how the chiral small molecule orientation can be controlled by structural templating for maximal chiroptical and electronic responses. The set of lowest-lying tentative crystal structures of the [n]helicenes presented in 5 have been forwarded to the Johnson group (Dalhousie University) for more accurate crystal-energy calculations using DFT-XDM. The CRYSTACK analysis tool is planned to be made publicly available to aid the development of new tools for the analysis of crystal interactions of [n]helicenes and other molecules, within the group and beyond.

Helicenes computational research is still in its infancy and a comprehensive analysis of a wide range of terminally functionalised [6]helicenes, tested for their OSC potential (Chapters 3 and 4), should inspire synthetic chemists to attempt their synthesis and to grow multiple crystalline forms via polymorph screenings to yield new functionalised helicene crystals with enhanced charge-transport properties. In the near future, I expect phonon calculations to shed more light on helicene charge-transport properties. The research on helicenes of varying axial chain length first aimed to quantify the degree of involvement in noncovalent interactions across nonplanar π-systems. Unlike planar π-systems, which have been extensively studied in computational materials discovery via SMILES,[a] the computational electronic materials discovery of nonplanar π-systems as attempted in this thesis is more computationally challenging, yet hardly explored. The crystal packing interactions results (Chapter 5) have been presented independent of the target physical properties. Future research on helicenes in the solid state could consist of studying the

[a] Simplified Molecular Input Line Entry System (SMILES) is an easy and learnable notation to represent chemical structure by a computer, yet unable to represent axial chirality.

chiroptical responses of helicenes of long axial chain length arranged in homochiral columns and potential applications in chiral molecular recognition. Despite all the findings that have been presented in this thesis, the area of computational chirality-driven organic electronics research is still at a very early stage and can therefore be considered as a very promising area for future research. While helicenes have been synthesised for over a hundred years,[191] applying them in chirality-driven organic electronics is still an emerging field of research. This thesis has touched on the surface of the vast potential that lies in chirality-driven electronics and gave pointers to fruitful avenues of future research in this exciting field.

Bibliography

[1] Cruz-Cabeza, A. J.; Feeder, N.; Davey, R. J. *Communications Chemistry* **2020**, *3*, 1–4.

[2] Cruz-Cabeza, A. J.; Davey, R. J.; Oswald, I. D.; Ward, M. R.; Sugden, I. J. *CrystEngComm* **2019**, *21*, 2034–2042.

[3] Gavezzotti, A. *Accounts of Chemical Research* **1994**, *27*, 309–314.

[4] Cruz-Cabeza, A. J.; Reutzel-Edens, S. M.; Bernstein, J. *Chemical Society Reviews* **2015**, *44*, 8619–8635.

[5] Haleblian, J.; McCrone, W. *Journal of pharmaceutical sciences* **1969**, *58*, 911–929.

[6] Brog, J.-P.; Chanez, C.-L.; Crochet, A.; Fromm, K. M. *Rsc Advances* **2013**, *3*, 16905–16931.

[7] McCrone, W. C. *Physics and chemistry of the organic solid state* **1965**, *2*, 725–767.

[8] Rosenstein, S.; Lamy, P. P. *American Journal of Health-System Pharmacy* **1969**, *26*, 598–601.

[9] Sharma, B. *Journal of Chemical Education* **1987**, *64*, 404.

[10] Gavezzotti, A. *Molecular aggregation: structure analysis and molecular simulation of crystals and liquids*; OUP Oxford, 2007; Vol. 19.

[11] Cruz-Cabeza, A. J.; Bernstein, J. *Chemical reviews* **2014**, *114*, 2170–2191.

[12] Purohit, R.; Venugopalan, P. *Resonance* **2009**, *14*, 882.

[13] Black, J.; Cardew, P.; Cruz-Cabeza, A.; Davey, R.; Gilks, S.; Sullivan, R. *CrystEngComm* **2018**, *20*, 768–776.

[14] Brandel, C.; Amharar, Y.; Rollinger, J. M.; Griesser, U. J.; Cartigny, Y.; Petit, S.; Coquerel, G. *Molecular Pharmaceutics* **2013**, *10*, 3850–3861.

[15] Mitscherlich, E. *Ann. Chim. Phys* **1822**, *19*, 350–419.

[16] Mitscherlich, E. Sur la rapport qui existe entre les proportions chimiques et la forme crystalline. 3. Memoire sur les corps qui affectent deux forms cristallines differentes. Annales de Chimie et de Physique. 1823; p 264.

[17] Liebig, J.; Wohler, F. *Pharm Anntiinflam* **1832**, *249*, 514.

[18] Fellah, N.; Shtukenberg, A. G.; Chan, E. J.; Vogt-Maranto, L.; Xu, W.; Li, C.; Tuckerman, M. E.; Kahr, B.; Ward, M. D. *Crystal Growth & Design* **2020**, *20*, 2670–2682.

[19] Bond, A. D. *Current opinion in solid state and materials science* **2009**, *13*, 91–97.

[20] Gavezzotti, A.; Filippini, G. *Journal of the American Chemical Society* **1995**, *117*, 12299–12305.

[21] Varughese, S.; Kiran, M.; Ramamurty, U.; Desiraju, G. R. *Angewandte Chemie International Edition* **2013**, *52*, 2701–2712.

[22] Bernstein, J. *Polymorphism in Molecular Crystals 2e*; International Union of Crystal, 2020; Vol. 30.

[23] Haisa, M.; Kashino, S.; Kawai, R.; Maeda, H. *Acta Crystallographica Section B: Structural Crystallography and Crystal Chemistry* **1976**, *32*, 1283–1285.

[24] Drebushchak, T. N.; Boldyreva, E. V. *Zeitschrift für Kristallographie-Crystalline Materials* **2004**, *219*, 506–512.

[25] Burger, A. *Acta Pharm. Technol.* **1982**, *28*, 1–20.

[26] Joiris, E.; Di Martino, P.; Berneron, C.; Guyot-Hermann, A.-M.; Guyot, J.-C. *Pharmaceutical research* **1998**, *15*, 1122–1130.

[27] Perrin, M.-A.; Neumann, M. A.; Elmaleh, H.; Zaske, L. *Chemical communications* **2009**, 3181–3183.

[28] Garekani, H. A.; Ford, J. L.; Rubinstein, M. H.; Rajabi-Siahboomi, A. R. *International journal of pharmaceutics* **1999**, *187*, 77–89.

[29] Garekani, H. A.; Ford, J. L.; Rubinstein, M. H.; Rajabi-Siahboomi, A. R. *Drug development and industrial pharmacy* **2001**, *27*, 935–942.

[30] Prescott, L. F. *American journal of therapeutics* **2000**, *7*, 143–147.

[31] Dunitz, J. D.; Bernstein, J. *Accounts of chemical research* **1995**, *28*, 193–200.

[32] Stahly, G. P. *Crystal growth & design* **2007**, *7*, 1007–1026.

[33] Chen, J.; Anthony, J.; Martin, D. C. *The Journal of Physical Chemistry B* **2006**, *110*, 16397–16403.

[34] Bernstein, J.; Hagler, A. *Journal of the American Chemical Society* **1978**, *100*, 673–681.

[35] Yu, L.; Reutzel-Edens, S. M.; Mitchell, C. A. *Organic Process Research & Development* **2000**, *4*, 396–402.

[36] Hursthouse, M. B.; Huth, L. S.; Threlfall, T. L. *Organic Process Research & Development* **2009**, *13*, 1231–1240.

[37] Bernstein, J.; Dunitz, J.; Gavezzotti, A. *Crystal Growth and Design* **2008**, *8*, 2011–2018.

[38] Oganov, A. R. *Faraday Discuss.* **2018**, *211*, 643–660.

[39] Hoja, J.; Ko, H. Y.; Neumann, M. A.; Car, R.; DiStasio, R. A.; Tkatchenko, A. *Acta Hortic. Sin.* **2019**, *45*, eaau3338.

[40] Neumann, M. A.; Leusen, F. J.; Kendrick, J. *Angewandte Chemie* **2008**, *120*, 2461–2464.

[41] Day, G. M.; Motherwell, W. D.; Jones, W. *Phys. Chem. Chem. Phys.* **2007**, *9*, 1693–1704.

[42] Lommerse, J. P. M.; Motherwell, W. D. S.; Ammon, H. L.; Dunitz, J. D.; Gavezzotti, A.; Hofmann, D. W. M.; Leusen, F. J. J.; Mooij, W. T. M.; Price, S. L.; Schweizer, B.; Schmidt, M. U.; van Eijck, B. P.; Verwer, P.; Williams, D. E. *Acta Crystallographica Section B* **2000**, *56*, 697–714.

[43] Motherwell, W. D. et al. *Acta Crystallogr. Sect. B Struct. Sci.* **2002**, *58*, 647–661.

[44] Day, G. M. et al. *Acta Crystallogr. Sect. B Struct. Sci.* **2005**, *61*, 511–527.

[45] Day, G. M. et al. *Acta Crystallogr. Sect. B Struct. Sci.* **2009**, *65*, 107–125.

[46] Bardwell, D. A. et al. *Acta Crystallogr. Sect. B Struct. Sci.* **2011**, *67*, 535–551.

[47] Reilly, A. M. et al. *Acta Crystallographica Section B* **2016**, *72*, 439–459.

[48] Lund, A. M.; Pagola, G. I.; Orendt, A. M.; Ferraro, M. B.; Facelli, J. C. *Chemical physics letters* **2015**, *626*, 20–24.

[49] Curtis, F.; Li, X.; Rose, T.; Vazquez-Mayagoitia, A.; Bhattacharya, S.; Ghiringhelli, L. M.; Marom, N. *Journal of chemical theory and computation* **2018**, *14*, 2246–2264.

[50] Zhu, Q.; Oganov, A. R.; Glass, C. W.; Stokes, H. T. *Acta Crystallographica Section B: Structural Science* **2012**, *68*, 215–226.

[51] Karfunkel, H.; Gdanitz, R. *Journal of computational chemistry* **1992**, *13*, 1171–1183.

[52] Akkermans, R. L.; Spenley, N. A.; Robertson, S. H. *Molecular Simulation* **2013**, *39*, 1153–1164.

[53] Kendrick, J.; Leusen, F. J.; Neumann, M. A.; van de Streek, J. *Chemistry–A European Journal* **2011**, *17*, 10736–10744.

[54] Cole, J. C.; Groom, C. R.; Read, M. G.; Giangreco, I.; McCabe, P.; Reilly, A. M.; Shields, G. P. *Acta Crystallographica Section B: Structural Science, Crystal Engineering and Materials* **2016**, *72*, 530–541.

[55] Karamertzanis, P. G.; Pantelides, C. C. *Mol. Phys.* **2007**, *105*, 273–291.

[56] Neumann, M. A.; Perrin, M.-A. *The Journal of Physical Chemistry B* **2005**, *109*, 15531–15541.

[57] Kazantsev, A. V.; Karamertzanis, P. G.; Adjiman, C. S.; Pantelides, C. C. *J. Chem. Theory Comput.* **2011**, *7*, 1998–2016.

[58] Price, S. L.; Leslie, M.; Welch, G. W.; Habgood, M.; Price, L. S.; Karamertzanis, P. G.; Day, G. M. *Phys. Chem. Chem. Phys.* **2010**, *12*, 8478–8490.

[59] Willock, D. J.; Price, S. L.; Leslie, M.; Catlow, C. R. A. *Journal of Computational Chemistry* **1995**, *16*, 628–647.

[60] Beyer, T.; Lewis, T.; Price, S. L. *CrystEngComm* **2001**, *3*, 178–212.

[61] Price, S. L. *Chemical Society Reviews* **2014**, *43*, 2098–2111.

[62] Brandenburg, J. G.; Grimme, S. *Top. Curr. Chem.*; Springer, 2014; Vol. 345; pp 1–24.

[63] LeBlanc, L. M.; Johnson, E. R. *CrystEngComm* **2019**, *21*, 5995–6009.

[64] Whittleton, S. R.; Otero-De-La-Roza, A.; Johnson, E. R. *Journal of chemical theory and computation* **2017**, *13*, 441–450.

[65] Brandenburg, J. G.; Grimme, S. *The Journal of Physical Chemistry Letters* **2014**, *5*, 1785–1789.

[66] Tkatchenko, A.; Scheffler, M. *Physical review letters* **2009**, *102*, 073005.

[67] Reilly, A. M.; Tkatchenko, A. *The journal of physical chemistry letters* **2013**, *4*, 1028–1033.

[68] Reilly, A. M.; Tkatchenko, A. *The Journal of chemical physics* **2013**, *139*, 024705.

[69] Price, S. L.; Braun, D. E.; Reutzel-Edens, S. M. *Chemical Communications* **2016**, *52*, 7065–7077.

[70] Braun, D. E.; McMahon, J. A.; Koztecki, L. H.; Price, S. L.; Reutzel-Edens, S. M. *Crystal growth & design* **2014**, *14*, 2056–2072.

[71] Kitaigorodskii, A. *Organic Chemical Crystallography* **1961**, 113–240.

[72] Thompson, H. P.; Day, G. M. *Chemical Science* **2014**, *5*, 3173–3182.

[73] Beyer, T.; Price, S. L. *The Journal of Physical Chemistry B* **2000**, *104*, 2647–2655.

[74] Chisholm, J. A.; Motherwell, S. *Journal of Applied Crystallography* **2005**, *38*, 228–231.

[75] De, S.; Bartók, A. P.; Csányi, G.; Ceriotti, M. *Physical Chemistry Chemical Physics* **2016**, *18*, 13754–13769.

[76] Ferré, G.; Maillet, J.-B.; Stoltz, G. *The Journal of chemical physics* **2015**, *143*, 104114.

[77] Sadeghi, A.; Ghasemi, S. A.; Schaefer, B.; Mohr, S.; Lill, M. A.; Goedecker, S. *The Journal of chemical physics* **2013**, *139*, 184118.

[78] Zhu, L.; Amsler, M.; Fuhrer, T.; Schaefer, B.; Faraji, S.; Rostami, S.; Ghasemi, S. A.; Sadeghi, A.; Grauzinyte, M.; Wolverton, C.; Goedecker, S. *The Journal of Chemical Physics* **2016**, *144*, 034203.

[79] Schaefer, B.; Mohr, S.; Amsler, M.; Goedecker, S. *The Journal of chemical physics* **2014**, *140*, 214102.

[80] Oganov, A. R.; Lyakhov, A. O.; Valle, M. *Accounts of chemical research* **2011**, *44*, 227–237.

[81] Butler, K. T.; Davies, D. W.; Cartwright, H.; Isayev, O.; Walsh, A. *Nature* **2018**, *559*, 547–555.

[82] Ferré, G.; Haut, T.; Barros, K. *The Journal of chemical physics* **2017**, *146*, 114107.

[83] Bartók, A. P.; De, S.; Poelking, C.; Bernstein, N.; Kermode, J. R.; Csányi, G.; Ceriotti, M. *Science advances* **2017**, *3*, e1701816.

[84] Musil, F.; De, S.; Yang, J.; Campbell, J. E.; Day, G. M.; Ceriotti, M. *Chem. Sci.* **2018**, *9*, 1289–1300.

[85] Kang, Z.; Peng, C.; Cheng, Q. *Neurocomputing* **2017**, *267*, 210–219.

[86] Balcan, M.-F.; Blum, A.; Vempala, S. *Machine Learning* **2006**, *65*, 79–94.

[87] Balcan, M.-F.; Blum, A. On a theory of kernels as similarity functions. 2006.

[88] Bonchev, D. *Chemical graph theory: introduction and fundamentals*; CRC Press, 1991; Vol. 1.

[89] Rasmussen, C. *Cambridge, MA* **2006**,

[90] Tang, Y.-H.; de Jong, W. A. *The Journal of chemical physics* **2019**, *150*, 044107.

[91] Stulp, F.; Sigaud, O. *Neural Networks* **2015**, *69*, 60–79.

[92] Park, C.; Apley, D. *The Journal of Machine Learning Research* **2018**, *19*, 269–311.

[93] Rasmussen, C. E.; Williams, C. Gaussian processes for machine learning, vol. 1. 2006.

[94] Yang, G.; Wu, K. *Molecular Systems Design & Engineering* **2019**, *4*, 586–596.

[95] Xie, T.; Grossman, J. C. *Physical review letters* **2018**, *120*, 145301.

[96] Blatov, V. *Acta Crystallographica Section A: Foundations of Crystallography* **2000**, *56*, 178–188.

[97] Purdum, G. E. et al. *J. Am. Chem. Soc.* **2018**, *140*, 7519–7525.

[98] Ono, K.; Tanaka, S.; Yamashita, Y. *Angewandte Chemie International Edition in English* **1994**, *33*, 1977–1979.

[99] Dunitz, J. D.; Gavezzotti, A. *Accounts of chemical research* **1999**, *32*, 677–684.

[100] Dance, I. *New Journal of Chemistry* **2003**, *27*, 22–27.

[101] Price, S. L. *Acta Crystallographica Section B: Structural Science, Crystal Engineering and Materials* **2013**, *69*, 313–328.

[102] Hoja, J.; Reilly, A. M.; Tkatchenko, A. *Wiley Interdisciplinary Reviews: Computational Molecular Science* **2017**, *7*, e1294.

[103] Habgood, M.; Grau-Crespo, R.; Price, S. L. *Physical Chemistry Chemical Physics* **2011**, *13*, 9590–9600.

[104] Cruz-Cabeza, A. J.; Davey, R. J.; Sachithananthan, S. S.; Smith, R.; Tang, S. K.; Vetter, T.; Xiao, Y. *Chemical Communications* **2017**, *53*, 7905–7908.

[105] Hammond, R. B.; Pencheva, K.; Roberts, K. J. *CrystEngComm* **2012**, *14*, 1069–1082.

[106] Kendrick, J.; Stephenson, G. A.; Neumann, M. A.; Leusen, F. J. *Crystal growth & design* **2013**, *13*, 581–589.

[107] Cruz-Cabeza, A. J.; Karki, S.; Fabian, L.; Friščić, T.; Day, G. M.; Jones, W. *Chemical communications* **2010**, *46*, 2224–2226.

[108] Cruz-Cabeza, A. J.; Day, G. M.; Jones, W. *Chemistry–A European Journal* **2008**, *14*, 8830–8836.

[109] Wicker, J. G.; Cooper, R. I. *CrystEngComm* **2015**, *17*, 1927–1934.

[110] Bhardwaj, R. M.; Johnston, A.; Johnston, B. F.; Florence, A. J. *CrystEngComm* **2015**, *17*, 4272–4275.

[111] LeBlanc, L. M.; Otero-de-la Roza, A.; Johnson, E. R. *Crystal Growth & Design* **2016**, *16*, 6867–6873.

[112] Oganov, A. R.; Pickard, C. J.; Zhu, Q.; Needs, R. J. *Nature Reviews Materials* **2019**, *4*, 331–348.

[113] Nyman, J.; Day, G. M. *Physical Chemistry Chemical Physics* **2016**, *18*, 31132–31143.

[114] Cui, P.; McMahon, D. P.; Spackman, P. R.; Alston, B. M.; Little, M. A.; Day, G. M.; Cooper, A. I. *Chemical Science* **2019**, *10*, 9988–9997.

[115] Mattheus, C. C.; Dros, A. B.; Baas, J.; Meetsma, A.; De Boer, J. L.; Palstra, T. T. *Acta Crystallographica Section C: Crystal Structure Communications* **2001**, *57*, 939–941.

[116] Matsukawa, T.; Yoshimura, M.; Uchiyama, M.; Yamagishi, M.; Nakao, A.; Takahashi, Y.; Takeya, J.; Kitaoka, Y.; Mori, Y.; Sasaki, T. *Japanese Journal of Applied Physics* **2010**, *49*, 085502.

[117] Diao, Y.; Lenn, K. M.; Lee, W.-Y.; Blood-Forsythe, M. A.; Xu, J.; Mao, Y.; Kim, Y.; Reinspach, J. A.; Park, S.; Aspuru-Guzik, A.; Xue, G.; Clancy, P.; Bao, Z.; Mannsfeld, S. C. B. *Journal of the American Chemical Society* **2014**, *136*, 17046–17057, PMID: 25333565.

[118] Chung, H.; Diao, Y. *Journal of Materials Chemistry C* **2016**, *4*, 3915–3933.

[119] Pudipeddi, M.; Serajuddin, A. T. *Journal of pharmaceutical sciences* **2005**, *94*, 929–939.

[120] Gentili, D.; Gazzano, M.; Melucci, M.; Jones, D.; Cavallini, M. *Chemical Society Reviews* **2019**, *48*, 2502–2517.

[121] Bässler, H.; Köhler, A. *Unimolecular and supramolecular electronics I*; Springer, 2011; pp 1–65.

[122] Li, W.; Liu, Q.; Zhang, Y.; Li, C.; He, Z.; Choy, W. C.; Low, P. J.; Sonar, P.; Kyaw, A. K. K. *Advanced Materials* **2020**, *32*, 2001591.

[123] Tan, M. J.; Owh, C.; Chee, P. L.; Kyaw, A. K. K.; Kai, D.; Loh, X. J. *Journal of Materials Chemistry C* **2016**, *4*, 5531–5558.

[124] Lin, X.; Subbaraman, H.; Pan, Z.; Hosseini, A.; Longe, C.; Kubena, K.; Schleicher, P.; Foster, P.; Brickey, S.; Chen, R. T. *Electronics* **2014**, *3*, 624–635.

[125] Voss, D. Cheap and cheerful circuits. 2000.

[126] Mei, J.; Diao, Y.; Appleton, A. L.; Fang, L.; Bao, Z. *Journal of the American Chemical Society* **2013**, *135*, 6724–6746.

[127] Ostroverkhova, O. *Chemical reviews* **2016**, *116*, 13279–13412.

[128] Oberhofer, H.; Reuter, K.; Blumberger, J. *Chemical reviews* **2017**, *117*, 10319–10357.

[129] Sekitani, T.; Nakajima, H.; Maeda, H.; Fukushima, T.; Aida, T.; Hata, K.; Someya, T. *Nature materials* **2009**, *8*, 494–499.

[130] Gelinck, G.; Heremans, P.; Nomoto, K.; Anthopoulos, T. D. *Advanced materials* **2010**, *22*, 3778–3798.

[131] Coropceanu, V.; Cornil, J.; da Silva Filho, D. A.; Olivier, Y.; Silbey, R.; Brédas, J.-L. *Chemical reviews* **2007**, *107*, 926–952.

[132] Petty, A. J.; Ai, Q.; Sorli, J. C.; Haneef, H. F.; Purdum, G. E.; Boehm, A.; Granger, D. B.; Gu, K.; Rubinger, C. P. L.; Parkin, S. R.; Graham, K. R.; Jurchescu, O. D.; Loo, Y. L.; Risko, C.; Anthony, J. E. *Chem. Sci.* **2019**, *10*, 10543–10549.

[133] He, T.; Leowanawat, P.; Burschka, C.; Stepanenko, V.; Stolte, M.; Würthner, F. *Advanced Materials* **2018**, *30*, 1804032.

[134] Ruiz, C.; García-Frutos, E. M.; Hennrich, G.; Gómez-Lor, B. *J. Phys. Chem. Lett.* **2012**, *3*, 1428–1436.

[135] Yavuz, I. *Physical Chemistry Chemical Physics* **2017**, *19*, 25819–25828.

[136] Dong, H.; Fu, X.; Liu, J.; Wang, Z.; Hu, W. *Advanced materials* **2013**, *25*, 6158–6183.

[137] Franchini, C.; Reticcioli, M.; Setvin, M.; Diebold, U. *Nature Reviews Materials* **2021**, 1–27.

[138] Stafström, S. *Chemical Society Reviews* **2010**, *39*, 2484–2499.

[139] Wang, L.; Nan, G.; Yang, X.; Peng, Q.; Li, Q.; Shuai, Z. *Chemical Society Reviews* **2010**, *39*, 423–434.

[140] Troisi, A.; Orlandi, G. *The Journal of Physical Chemistry A* **2006**, *110*, 4065–4070.

[141] Lu, N.; Li, L.; Geng, D.; Liu, M. *Organic Electronics* **2018**, *61*, 223–234.

[142] LeBlanc Jr, O. H. *The Journal of Chemical Physics* **1961**, *35*, 1275–1280.

[143] Cheng, Y.; Silbey, R.; da Silva Filho, D. A.; Calbert, J. P.; Cornil, J.; Brédas, J.-L. *The Journal of chemical physics* **2003**, *118*, 3764–3774.

[144] Troisi, A.; Orlandi, G. *The Journal of Physical Chemistry B* **2005**, *109*, 1849–1856.

[145] Ortmann, F.; Bechstedt, F.; Hannewald, K. *New Journal of Physics* **2010**, *12*, 023011.

[146] Troisi, A.; Orlandi, G. *Physical review letters* **2006**, *96*, 086601.

[147] Köhler, A.; Bässler, H. *Electronic processes in organic semiconductors: An introduction*; John Wiley & Sons, 2015.

[148] Ortmann, F.; Bechstedt, F.; Hannewald, K. *Physical Review B* **2009**, *79*, 235206.

[149] Holstein, T. *Annals of physics* **1959**, *8*, 325–342.

[150] Ortmann, F.; Bechstedt, F.; Hannewald, K. *physica status solidi (b)* **2011**, *248*, 511–525.

[151] Troisi, A. *Chemical Society Reviews* **2011**, *40*, 2347–2358.

[152] Nematiaram, T.; Ciuchi, S.; Xie, X.; Fratini, S.; Troisi, A. *The Journal of Physical Chemistry C* **2019**, *123*, 6989–6997.

[153] Fratini, S.; Ciuchi, S.; Mayou, D.; De Laissardière, G. T.; Troisi, A. *Nature materials* **2017**, *16*, 998–1002.

[154] Ciuchi, S.; Fratini, S.; Mayou, D. *Physical Review B* **2011**, *83*, 081202.

[155] Fratini, S.; Mayou, D.; Ciuchi, S. *Advanced Functional Materials* **2016**, *26*, 2292–2315.

[156] Fratini, S.; Nikolka, M.; Salleo, A.; Schweicher, G.; Sirringhaus, H. *Nature materials* **2020**, *19*, 491–502.

[157] Nematiaram, T.; Troisi, A. *Materials Horizons* **2020**, *7*, 2922–2928.

[158] Giannini, S.; Carof, A.; Ellis, M.; Yang, H.; Ziogos, O. G.; Ghosh, S.; Blumberger, J. *Nat. Commun.* **2019**, *10*, 1–12.

[159] Jurchescu, O. D.; Mourey, D. A.; Subramanian, S.; Parkin, S. R.; Vogel, B. M.; Anthony, J. E.; Jackson, T. N.; Gundlach, D. J. *Physical Review B* **2009**, *80*, 085201.

[160] Mattheus, C. C.; De Wijs, G. A.; De Groot, R. A.; Palstra, T. T. *Journal of the American Chemical Society* **2003**, *125*, 6323–6330.

[161] Sorli, J. C.; Ai, Q.; Granger, D. B.; Gu, K.; Parkin, S.; Jarolimek, K.; Telesz, N.; Anthony, J. E.; Risko, C.; Loo, Y. L. *Chem. Mater.* **2019**, *31*, 6615–6623.

[162] Pulido, A. et al. *Nature* **2017**, *543*, 657–664.

[163] Zhao, C.; Chen, L.; Che, Y.; Pang, Z.; Wu, X.; Lu, Y.; Liu, H.; Day, G. M.; Cooper, A. I. *Nature communications* **2021**, *12*, 1–11.

[164] Day, G. M.; Cooper, A. I. *Advanced Materials* **2018**, *30*, 1704944.

[165] Campbell, J. E.; Yang, J.; Day, G. M. *J. Mater. Chem. C* **2017**, *5*, 7574–7584.

[166] Aitchison, C. M.; Kane, C. M.; McMahon, D. P.; Spackman, P. R.; Pulido, A.; Wang, X.; Wilbraham, L.; Chen, L.; Clowes, R.; Zwijnenburg, M. A.; Sprick, R. S.; Little, M. A.; Day, G. M.; Cooper, A. I. *J. Mater. Chem. A* **2020**, *8*, 7158–7170.

[167] Shi, W.; Wong, Z. M.; Deng, T.; Wu, G.; Yang, S.-W. *Advanced Functional Materials* **2020**, 2007438.

[168] Kunkel, C.; Schober, C.; Margraf, J. T.; Reuter, K.; Oberhofer, H. *Chem. Mater.* **2019**, *31*, 969–978.

[169] Kuo, M.-Y.; Chen, H.-Y.; Chao, I. *Chemistry–A European Journal* **2007**, *13*, 4750–4758.

[170] Atahan-Evrenk, S.; Atalay, F. B. *J. Phys. Chem. A* **2019**, *123*, 7855–7863.

[171] Pyzer-Knapp, E. O.; Li, K.; Aspuru-Guzik, A. *Advanced Functional Materials* **2015**, *25*, 6495–6502.

[172] Gryn'Ova, G.; Lin, K. H.; Corminboeuf, C. *J. Am. Chem. Soc.* **2018**, *140*, 16370–16386.

[173] Suh, C.; Fare, C.; Warren, J. A.; Pyzer-Knapp, E. O. *Annual Review of Materials Research* **2020**, *50*, 1–25.

[174] Tkatchenko, A. *Nature Communications* **2020**, *11*, 1–4.

[175] Narayanan, R.; Neuberger, H. *Physical Review Letters* **1993**, *71*, 3251.

[176] Jiao, L.; Howard, S.; Ran, S.; Wang, Z.; Rodriguez, J. O.; Sigrist, M.; Wang, Z.; Butch, N. P.; Madhavan, V. *Nature* **2020**, *579*, 523–527.

[177] Rodger, A.; Nordén, B. Circular dichroism and linear dichroism. 1997.

[178] Kiran, V.; Mathew, S. P.; Cohen, S. R.; Hernández Delgado, I.; Lacour, J.; Naaman, R. *Advanced Materials* **2016**, *28*, 1957–1962.

[179] Yang, Y.; Da Costa, R. C.; Fuchter, M. J.; Campbell, A. J. *Nature Photonics* **2013**, *7*, 634–638.

[180] Canary, J. W. *Chemical Society Reviews* **2009**, *38*, 747–756.

[181] Schweinfurth, D.; Zalibera, M.; Kathan, M.; Shen, C.; Mazzolini, M.; Trapp, N.; Crassous, J.; Gescheidt, G.; Diederich, F. *Journal of the American Chemical Society* **2014**, *136*, 13045–13052.

[182] Matell, M. *Ark. Kemi* **1953**, *5*, 17–37.

[183] Linol, J.; Morelli, T.; Petit, M.-N.; Coquerel, G. *Crystal growth & design* **2007**, *7*, 1608–1611.

[184] Flack, H. D. *Helvetica Chimica Acta* **2003**, *86*, 905–921.

[185] Nespolo, M.; Aroyo, M. I.; Souvignier, B. *Journal of Applied Crystallography* **2018**, *51*, 1481–1491.

[186]

[187] Tanaka, H.; Inoue, Y.; Mori, T. *ChemPhotoChem* **2018**, *2*, 386–402.

[188] Heffern, M. C.; Matosziuk, L. M.; Meade, T. J. *Chemical reviews* **2014**, *114*, 4496–4539.

[189] Esposito, M.; Tasco, V.; Todisco, F.; Cuscunà, M.; Benedetti, A.; Sanvitto, D.; Passaseo, A. *Nature communications* **2015**, *6*, 1–7.

[190] Martin, R. H. *Angewandte Chemie International Edition in English* **1974**, *13*, 649–660.

[191] Gingras, M. *Chemical Society Reviews* **2013**, *42*, 968–1006.

[192] Isla, H.; Crassous, J. *Comptes Rendus Chimie* **2016**, *19*, 39–49.

[193] Newman, M. S.; Lutz, W. B.; Lednicer, D. *Journal of the American Chemical Society* **1955**, *77*, 3420–3421.

[194] Newman, M. S.; Lednicer, D. *Journal of the American Chemical Society* **1956**, *78*, 4765–4770.

[195] Weitzenböck, R.; Klingler, A. *Monatshefte für Chemie und verwandte Teile anderer Wissenschaften* **1918**, *39*, 315–323.

[196] Hoffmann, R.; Hopf, H. *Angewandte Chemie International Edition* **2008**, *47*, 4474–4481.

[197] Rice, B.; LeBlanc, L. M.; Otero-De-La-Roza, A.; Fuchter, M. J.; Johnson, E. R.; Nelsona, J.; Jelfs, K. E. *Nanoscale* **2018**, *10*, 9410.

[198] Shen, Y.; Chen, C.-F. *Chemical reviews* **2012**, *112*, 1463–1535.

[199] Irfan, A.; Assiri, M.; Al-Sehemi, A. G. *Organic Electronics* **2018**, *57*, 211–220.

[200] Zhang, X.; Clennan, E. L.; Petek, T.; Weber, J. *Tetrahedron* **2017**, *73*, 508–518.

[201] Yamamoto, K.; Shimizu, T.; Igawa, K.; Tomooka, K.; Hirai, G.; Suemune, H.; Usui, K. *Scientific reports* **2016**, *6*, 1–8.

[202] Staab, H. A.; Saupe, T. *Angewandte Chemie International Edition in English* **1988**, *27*, 865–879.

[203] Zirnstein, M. A.; Staab, H. A. *Angewandte Chemie International Edition in English* **1987**, *26*, 460–461.

[204] Staab, H. A.; Zirnstein, M. A.; Krieger, C. *Angewandte Chemie International Edition in English* **1989**, *28*, 86–88.

[205] Furche, F.; Ahlrichs, R.; Wachsmann, C.; Weber, E.; Sobanski, A.; Vögtle, F.; Grimme, S. *Journal of the American Chemical Society* **2000**, *122*, 1717–1724.

[206] Nakai, Y.; Mori, T.; Inoue, Y. *The Journal of Physical Chemistry A* **2012**, *116*, 7372–7385.

[207] Brandt, J. R.; Salerno, F.; Fuchter, M. J. *Nature Reviews Chemistry* **2017**, *1*, 0045.

[208] Jiang, Q.; Han, Y.; Zou, Y.; Phan, H.; Yuan, L.; Herng, T. S.; Ding, J.; Chi, C. *Chemistry–A European Journal* **2020**,

[209] Seibel, J.; Allemann, O.; Siegel, J. S.; Ernst, K.-H. *Journal of the American Chemical Society* **2013**, *135*, 7434–7437.

[210] Amemiya, R.; Yamaguchi, M. *Organic & biomolecular chemistry* **2008**, *6*, 26–35.

[211] Anger, E.; Iida, H.; Yamaguchi, T.; Hayashi, K.; Kumano, D.; Crassous, J.; Vanthuyne, N.; Roussel, C.; Yashima, E. *Polymer Chemistry* **2014**, *5*, 4909–4914.

[212] An, Z.; Yamaguchi, M. *Chemical Communications* **2012**, *48*, 7383–7385.

[213] Chen, J.; Captain, B.; Takenaka, N. *Organic Letters* **2011**, *13*, 1654–1657.

[214] Ben Hassine, B.; Gorsane, M.; Pecher, J.; Martin, R. *Bulletin des Sociétés Chimiques Belges* **1985**, *94*, 597–603.

[215] Mishra, A.; Fischer, M. K.; Bäuerle, P. *Angewandte Chemie International Edition* **2009**, *48*, 2474–2499.

[216] Ooyama, Y.; Ito, G.; Fukuoka, H.; Nagano, T.; Kagawa, Y.; Imae, I.; Komaguchi, K.; Harima, Y. *Tetrahedron* **2010**, *66*, 7268–7271.

[217] Ooyama, Y.; Inoue, S.; Asada, R.; Ito, G.; Kushimoto, K.; Komaguchi, K.; Imae, I.; Harima, Y. *European Journal of Organic Chemistry* **2010**, *2010*, 92–100.

[218] Laleu, B.; Mobian, P.; Herse, C.; Laursen, B. W.; Hopfgartner, G.; Bernardinelli, G.; Lacour, J. *Angewandte Chemie* **2005**, *117*, 1913–1917.

[219] Gilbert, A. M.; Katz, T. J.; Geiger, W. E.; Robben, M. P.; Rheingold, A. L. *Journal of the American Chemical Society* **1993**, *115*, 3199–3211.

[220] Sudhakar, A.; Katz, T. J. *Journal of the American Chemical Society* **1986**, *108*, 179–181.

[221] Katz, T. J.; Pesti, J. *Journal of the American Chemical Society* **1982**, *104*, 346–347.

[222] Katz, T. J.; Sudhakar, A.; Teasley, M. F.; Gilbert, A. M.; Geiger, W. E.; Robben, M. P.; Wuensch, M.; Ward, M. D. *Journal of the American Chemical Society* **1993**, *115*, 3182–3198.

[223] Katz, T. J.; Slusarek, W. *Journal of the American Chemical Society* **1979**, *101*, 4259–4267.

[224] Takaishi, K.; Yamamoto, T.; Hinoide, S.; Ema, T. *Chemistry–A European Journal* **2017**, *23*, 9249–9252.

[225] Brandt, J. R.; Wang, X.; Yang, Y.; Campbell, A. J.; Fuchter, M. J. *Journal of the American Chemical Society* **2016**, *138*, 9743–9746.

[226] Singh, R.; Narayanan Unni, K.; Solanki, A.; Deepak, *Optical Materials* **2012**, *34*, 716–723.

[227] Wan, L.; Wade, J.; Salerno, F.; Arteaga, O.; Laidlaw, B.; Wang, X.; Penfold, T.; Fuchter, M. J.; Campbell, A. J. *ACS nano* **2019**, *13*, 8099–8105.

[228] Lovinger, A. J.; Nuckolls, C.; Katz, T. J. *Journal of the American Chemical Society* **1998**, *120*, 264–268.

[229] Feng, F.; Miyashita, T.; Okubo, H.; Yamaguchi, M. *Journal of the American Chemical Society* **1998**, *120*, 10166–10170.

[230] Verbiest, T.; Van Elshocht, S.; Kauranen, M.; Hellemans, L.; Snauwaert, J.; Nuckolls, C.; Katz, T. J.; Persoons, A. *Science* **1998**, *282*, 913–915.

[231] Ernst, K.-H.; Neuber, M.; Grunze, M.; Ellerbeck, U. *Journal of the American Chemical Society* **2001**, *123*, 493–495.

[232] Zhou, Y.; Lei, T.; Wang, L.; Pei, J.; Cao, Y.; Wang, J. *Advanced Materials* **2010**, *22*, 1484–1487.

[233] Kim, C.; Marks, T. J.; Facchetti, A.; Schiavo, M.; Bossi, A.; Maiorana, S.; Licandro, E.; Todescato, F.; Toffanin, S.; Muccini, M.; Graiff, C.; Tiripicchio, A. *Organic Electronics* **2009**, *10*, 1511–1520.

[234] Bushby, R. J.; Lozman, O. R. *Current Opinion in Colloid & Interface Science* **2002**, *7*, 343–354.

[235] Mahmood, M. S.; Said, S. M.; Chatterjee, A.; Sabri, M. F. M.; Mainal, A.; Daud, M. N.; Sairi, N. A. *Materials Research Express* **2018**, *5*, 126306.

[236] Nuckolls, C.; Katz, T. J. *Journal of the American Chemical Society* **1998**, *120*, 9541–9544.

[237] Shcherbina, M. A.; Zeng, X.-b.; Tadjiev, T.; Ungar, G.; Eichhorn, S. e.; Phillips, K. E.; Katz, T. J. *Angewandte Chemie International Edition* **2009**, *48*, 7837–7840.

[238] Phillips, K. E.; Katz, T. J.; Jockusch, S.; Lovinger, A. J.; Turro, N. J. *Journal of the American Chemical Society* **2001**, *123*, 11899–11907.

[239] Hatakeyama, T.; Hashimoto, S.; Oba, T.; Nakamura, M. *Journal of the American Chemical Society* **2012**, *134*, 19600–19603.

[240] Nakano, K.; Oyama, H.; Nishimura, Y.; Nakasako, S.; Nozaki, K. *Angewandte Chemie* **2012**, *124*, 719–723.

[241] Karamertzanis, P. G.; Pantelides, C. C. *Journal of Computational Chemistry* **2005**, *26*, 304–324.

[242] Rice, E. Computational modelling of electronic states, charge transfer and charge transport in organic semiconductors. Ph.D. thesis, 2017.

[243] Jensen, F. *Introduction to computational chemistry*; John wiley & sons, 2017.

[244] Monticelli, L.; Tieleman, D. P. *Biomolecular simulations* **2013**, 197–213.

[245] Wang, X.; Ramírez-Hinestrosa, S.; Dobnikar, J.; Frenkel, D. *Physical Chemistry Chemical Physics* **2020**, *22*, 10624–10633.

[246] Williams, D. *Journal of molecular structure* **1999**, *485*, 321–347.

[247] Pyzer-Knapp, E. O.; Thompson, H. P.; Day, G. M. *Acta Crystallographica Section B: Structural Science, Crystal Engineering and Materials* **2016**, *72*, 477–487.

[248] Koch, W.; Holthausen, M. A Chemist's Guide to Density Functional Theory, Wiley. 2001.

[249] Attila Szabo, N. S. O., Ed. *Modern Quantum Chemistry, Introduction to Advanced Electronic Structure Theory*, first edit ed.; Dover Publications Inc.: Mineola, New York, 1996.

[250] Kohn, W.; Becke, A. D.; Parr, R. G. *The Journal of Physical Chemistry* **1996**, *100*, 12974–12980.

[251] Becke, A. D. *The Journal of chemical physics* **2014**, *140*, 18A301.

[252] Perdew, J. P.; Burke, K.; Ernzerhof, M. *Physical review letters* **1996**, *77*, 3865.

[253] Becke, A. D.; Johnson, E. R. *The Journal of chemical physics* **2007**, *127*, 154108.

[254] Otero-De-La-Roza, A.; Johnson, E. R. *The Journal of chemical physics* **2012**, *136*, 174109.

[255] Christian, M. S.; Otero-de-la Roza, A.; Johnson, E. R. *Journal of chemical theory and computation* **2016**, *12*, 3305–3315.

[256] Becke, A. *The Journal of chemical physics* **1986**, *85*, 7184–7187.

[257] Becke, A. D. *The Journal of Chemical Physics* **1993**, *98*, 5648–5652.

[258] Grimme, S. *Wiley Interdisciplinary Reviews: Computational Molecular Science* **2011**, *1*, 211–228.

[259] Kristyán, S.; Pulay, P. *Chemical physics letters* **1994**, *229*, 175–180.

[260] Hermann, J.; DiStasio Jr, R. A.; Tkatchenko, A. *Chemical Reviews* **2017**, *117*, 4714–4758.

[261] Berland, K.; Cooper, V. R.; Lee, K.; Schröder, E.; Thonhauser, T.; Hyldgaard, P.; Lundqvist, B. I. *Reports on Progress in Physics* **2015**, *78*, 066501.

[262] Beran, G. J. *Chemical reviews* **2016**, *116*, 5567–5613.

[263] Tkatchenko, A.; DiStasio Jr, R. A.; Car, R.; Scheffler, M. *Physical review letters* **2012**, *108*, 236402.

[264] Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. *J. Chem. Phys.* **2010**, *132*, 154104.

[265] Grimme, S.; Hansen, A.; Brandenburg, J. G.; Bannwarth, C. *Chemical reviews* **2016**, *116*, 5105–5154.

[266] Johnson, E. R.; Otero-de-la Roza, A. *Journal of chemical theory and computation* **2012**, *8*, 5124–5131.

[267] Otero-De-La-Roza, A.; Johnson, E. R.; Contreras-García, J. *Physical Chemistry Chemical Physics* **2012**, *14*, 12165–12172.

[268] LeBlanc, L. M.; Otero-de-la Roza, A.; Johnson, E. R. *Journal of chemical theory and computation* **2018**, *14*, 2265–2276.

[269] LeBlanc, L. M.; Johnson, E. R. *CrystEngComm* **2019**, *21*, 5995–6009.

[270] Bannwarth, C.; Ehlert, S.; Grimme, S. *J. Chem. Theory Comput.* **2019**, *15*, 1652–1671.

[271] Vydrov, O. A.; Van Voorhis, T. *The Journal of chemical physics* **2010**, *133*, 244103.

[272] Grimme, S.; Bannwarth, C.; Shushkov, P. *Journal of chemical theory and computation* **2017**, *13*, 1989–2009.

[273] Caldeweyher, E.; Ehlert, S.; Hansen, A.; Neugebauer, H.; Spicher, S.; Bannwarth, C.; Grimme, S. *The Journal of chemical physics* **2019**, *150*, 154122.

[274] Minenkov, Y.; Chermak, E.; Cavallo, L. *Journal of chemical theory and computation* **2015**, *11*, 4664–4676.

[275] Guo, Y.; Riplinger, C.; Becker, U.; Liakos, D. G.; Minenkov, Y.; Cavallo, L.; Neese, F. *The Journal of chemical physics* **2018**, *148*, 011101.

[276] Edmiston, C.; Krauss, M. *The Journal of Chemical Physics* **1966**, *45*, 1833–1839.

[277] Karamertzanis, P.; Pantelides, C. *Molecular Physics* **2007**, *105*, 273–291.

[278] Frisch, M. J. et al. Gaussian16 Revision C.01. 2016.

[279] Stone, A. J. *Journal of Chemical Theory and Computation* **2005**, *1*, 1128–1132.

[280] Coombes, D. S.; Price, S. L.; Willock, D. J.; Leslie, M. *The Journal of Physical Chemistry* **1996**, *100*, 7352–7360.

[281] Feurer, M.; Klein, A.; Eggensperger, K.; Springenberg, J. T.; Blum, M.; Hutter, F. *Automated Machine Learning*; Springer, Cham, 2019; pp 113–134.

[282] others,, et al. Tensorflow: A system for large-scale machine learning. 12th {USENIX} symposium on operating systems design and implementation ({OSDI} 16). 2016; pp 265–283.

[283] Alpaydin, E. *Introduction to Machine Learning*; MIT Press, 2014.

[284] MacKay, D. J. *NATO ASI series F computer and systems sciences* **1998**, *168*, 133–166.

[285] Strehl, A.; Ghosh, J. *Journal of machine learning research* **2002**, *3*, 583–617.

[286] Topchy, A.; Jain, A. K.; Punch, W. *IEEE transactions on pattern analysis and machine intelligence* **2005**, *27*, 1866–1881.

[287] Newman, M. E.; Girvan, M. *Physical review E* **2004**, *69*, 026113.

[288] Blondel, V. D.; Guillaume, J.-L.; Lambiotte, R.; Lefebvre, E. *Journal of Statistical Mechanics: Theory and Experiment* **2008**, *2008*, P10008.

[289] Landrum, G. *Read. Writ.* **2011**, *1*, 1–79.

[290] Willett, P.; Barnard, J. M.; Downs, G. M. *Journal of chemical information and computer sciences* **1998**, *38*, 983–996.

[291] Bajusz, D.; Rácz, A.; Héberger, K. *J. Cheminform.* **2015**, *7*, 20.

[292] Hofmann, T.; Schölkopf, B.; Smola, A. J. *The annals of statistics* **2008**, 1171–1220.

[293] Pronobis, W.; Müller, K.-R. *Machine Learning Meets Quantum Physics*; Springer, 2020; pp 25–36.

[294] Scholkopf, B. *Advances in neural information processing systems* **2001**, 301–307.

[295] Shervashidze, N.; Schweitzer, P.; Van Leeuwen, E. J.; Mehlhorn, K.; Borgwardt, K. M. *Journal of Machine Learning Research* **2011**, *12*.

[296] Weisfeiler, B.; Leman, A. *NTI, Series* **1968**, *2*, 12–16.

[297] Nikolentzos, G.; Meladianos, P.; Vazirgiannis, M. Matching node embeddings for graph similarity. Proceedings of the AAAI Conference on Artificial Intelligence. 2017.

[298] Grauman, K.; Darrell, T. *Journal of Machine Learning Research* **2007**, *8*.

[299] Lazebnik, S.; Schmid, C.; Ponce, J. Beyond bags of features: Spatial pyramid matching for recognizing natural scene categories. 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'06). 2006; pp 2169–2178.

[300] Grauman, K.; Darrell, T. The pyramid match kernel: Discriminative classification with sets of image features. Tenth IEEE International Conference on Computer Vision (ICCV'05) Volume 1. 2005; pp 1458–1465.

[301] Kirkpatrick, J. Calculating Intermolecular Charge Transport Parameters in Conjugated Materials. Ph.D. thesis, Faculty of Natural Sciences, Imperial College London, 2007.

[302] Marcus, R. A. *Annual Review of Physical Chemistry* **1964**, *15*, 155–196.

[303] Kirkpatrick, J. *Int. J. Quantum Chem.* **2008**, *108*, 51–56.

[304] Wu, Q.; Van Voorhis, T. *The Journal of chemical physics* **2006**, *125*, 164105.

[305] Cave, R. J.; Newton, M. D. *Chemical physics letters* **1996**, *249*, 15–19.

[306] Fonari, A.; Sutton, C.; Brédas, J.-L.; Coropceanu, V. *Physical Review B* **2014**, *90*, 165205.

[307] Xie, X.; Santana-Bonilla, A.; Troisi, A. *Journal of chemical theory and computation* **2018**, *14*, 3752–3762.

[308] Yang, Y.; Rice, B.; Shi, X.; Brandt, J. R.; Correa Da Costa, R.; Hedley, G. J.; Smilgies, D. M.; Frost, J. M.; Samuel, I. D.; Otero-De-La-Roza, A.; Johnson, E. R.; Jelfs, K. E.; Nelson, J.; Campbell, A. J.; Fuchter, M. J. *ACS Nano* **2017**, *11*, 8329–8338.

[309] Lluch, J. M. *Theor. Chem. Acc.* **2000**, *103*, 231–233.

[310] Sakanoue, K.; Motoda, M.; Sugimoto, M.; Sakaki, S. *The Journal of Physical Chemistry A* **1999**, *103*, 5551–5556.

[311] Andrienko, D.; Kirkpatrick, J.; Marcon, V.; Nelson, J.; Kremer, K. *physica status solidi (b)* **2008**, *245*, 830–834.

[312] Salerno, F.; Rice, B.; Schmidt, J. A.; Fuchter, M. J.; Nelson, J.; Jelfs, K. E. *Phys. Chem. Chem. Phys.* **2019**, *21*, 5059–5067.

[313] Brédas, J.-L.; Beljonne, D.; Coropceanu, V.; Cornil, J. *Chemical reviews* **2004**, *104*, 4971–5004.

[314] da Silva Filho, D. A.; Kim, E.-G.; Brédas, J.-L. *Advanced Materials* **2005**, *17*, 1072–1076.

[315] Lin, B. C.; Cheng, C. P.; Lao, Z. P. M. *The Journal of Physical Chemistry A* **2003**, *107*, 5241–5251.

[316] Ingel, R. P.; III, D. L. *Journal of the American Ceramic Society* **1988**, *71*, 265–271.

[317] Xu, M.; Kitazawa, H.; Takano, Y.; Ye, J.; Nishida, K.; Abe, H.; Matsushita, A.; Tsujii, N.; Kido, G. *Applied Physics Letters* **2001**, *79*, 2779–2781.

[318] Hoofman, R. J.; Siebbeles, L. D.; de Haas, M. P.; Hummel, A.; Bloor, D. *The Journal of chemical physics* **1998**, *109*, 1885–1893.

[319] Reese, C.; Bao, Z. *Advanced Materials* **2007**, *19*, 4535–4538.

[320] Wade, J.; Steiner, F.; Niedzialek, D.; James, D. T.; Jung, Y.; Yun, D.-J.; Bradley, D. D.; Nelson, J.; Kim, J.-S. *Journal of Materials Chemistry C* **2014**, *2*, 10110–10115.

[321] Dong, H.; Wang, C.; Hu, W. *Chem. Commun.* **2010**, *46*, 5211–5222.

[322] Wang, C.; Liu, Y.; Ji, Z.; Wang, E.; Li, R.; Jiang, H.; Tang, Q.; Li, H.; Hu, W. *Chemistry of Materials* **2009**, *21*, 2840–2845.

[323] Wachsmann, C.; Weber, E.; Czugler, M.; Seichter, W. *European Journal of Organic Chemistry* **2003**, *2003*, 2863–2876.

[324] Schaack, C.; Sidler, E.; Trapp, N.; Diederich, F. *Chemistry – A European Journal* **2017**, *23*, 14153–14157.

[325] D'Avino, G.; Muccioli, L.; Olivier, Y.; Beljonne, D. *The Journal of Physical Chemistry Letters* **2016**, *7*, 536–540, PMID: 26785294.

[326] Griffith, O. L.; Anthony, J. E.; Jones, A. G.; Lichtenberger, D. L. *Journal of the American Chemical Society* **2010**, *132*, 580–586.

[327] Zeidell, A. M.; Jennings, L.; Frederickson, C. K.; Ai, Q.; Dressler, J. J.; Zakharov, L. N.; Risko, C.; Haley, M. M.; Jurchescu, O. D. *Chemistry of Materials* **2019**, *31*, 6962–6970.

[328] Harrelson, T. F. et al. *Mater. Horiz.* **2019**, *6*, 182–191.

[329] Steed, K. M.; Steed, J. W. *Chemical reviews* **2015**, *115*, 2895–2933.

[330] Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 1372–1377.

[331] Breneman, C. M.; Wiberg, K. B. *Journal of Computational Chemistry* **1990**, *11*, 361–373.

[332] Sugden, I.; Adjiman, C. S.; Pantelides, C. C. *Acta Crystallographica Section B: Structural Science, Crystal Engineering and Materials* **2016**, *72*, 864–874.

[333] Giannozzi, P. et al. *Journal of Physics. Condensed Matter* **2017**, *29*.

[334] Blöchl, P. E. *Physical review B* **1994**, *50*, 17953.

[335] Dovesi, R.; Erba, A.; Orlando, R.; Zicovich-Wilson, C. M.; Civalleri, B.; Maschio, L.; Rérat, M.; Casassa, S.; Baima, J.; Salustro, S.; Kirtman, B. *WIREs Computational Molecular Science* **2018**, *8*, e1360.

[336] Cutini, M.; Civalleri, B.; Corno, M.; Orlando, R.; Brandenburg, J. G.; Maschio, L.; Ugliengo, P. *Journal of chemical theory and computation* **2016**, *12*, 3340–3352.

[337] Monkhorst, H. J.; Pack, J. D. *Physical review B* **1976**, *13*, 5188.

[338] Ramdas, S.; Thomas, J. M.; Jordan, M.; Eckhardt, C. J. *The Journal of Physical Chemistry* **1981**, *85*, 2421–2425.

[339] Salentin, S.; Schreiber, S.; Haupt, V. J.; Adasme, M. F.; Schroeder, M. *Nucleic acids research* **2015**, *43*, W443–W447.

[340] Yang, J.; Roy, A.; Zhang, Y. *Nucleic acids research* **2012**, *41*, D1096–D1103.

[341] Ong, S. P.; Richards, W. D.; Jain, A.; Hautier, G.; Kocher, M.; Cholia, S.; Gunter, D.; Chevrier, V. L.; Persson, K. A.; Ceder, G. *Computational Materials Science* **2013**, *68*, 314–319.

[342] Shawe-Taylor, J.; Cristianini, N. *Kernel Methods for Pattern Analysis*; Cambridge University Press, 2004.

[343] Peach, R. L.; Arnaudon, A.; Schmidt, J. A.; Palasciano, H. A.; Bernier, N. R.; Jelfs, K. E.; Yaliraki, S. N.; Barahona, M. *Patterns* **2021**, 100227.

[344] Girvan, M.; Newman, M. E. J. **2002**,

[345] Brandes, U. *Journal of mathematical sociology* **2001**, *25*, 163–177.

[346] Brandes, U. *Social Networks* **2008**, *30*, 136–145.

[347] Lu, L.; Zhang, M. In *Encyclopedia of Systems Biology*; Dubitzky, W., Wolkenhauer, O., Cho, K.-H., Yokota, H., Eds.; Springer New York: New York, NY, 2013; pp 647–648.

[348] Siglidis, G.; Nikolentzos, G.; Limnios, S.; Giatsidis, C.; Skianis, K.; Vazirgiannis, M. *Journal of Machine Learning Research* **2020**, *21*, 1–5.

[349] Milligan, G. W. *psychometrika* **1980**, *45*, 325–342.

[350] Ben-Hur, A.; Guyon, I. *Functional genomics*; Springer, 2003; pp 159–182.

[351] Macrae, C. F.; Sovago, I.; Cottrell, S. J.; Galek, P. T.; McCabe, P.; Pidcock, E.; Platings, M.; Shields, G. P.; Stevens, J. S.; Towler, M.; Wood, P. A. *Journal of Applied Crystallography* **2020**, *53*, 226–235.

[352] Nyman, J.; Day, G. M. *CrystEngComm* **2015**, *17*, 5154–5165.

[353] Newman, M. E. *Proceedings of the national academy of sciences* **2006**, *103*, 8577–8582.

[354] Lancichinetti, A.; Fortunato, S. *Physical review E* **2009**, *80*, 056117.

[355] Liao, Y.-L.; Lin, C.-Y.; Liu, Y.-H.; Wong, K.-T.; Hung, W.-Y.; Chen, W.-J. *Chemical communications* **2007**, 1831–1833.

[356] Jhulki, S.; Mishra, A. K.; Chow, T. J.; Moorthy, J. N. *Chem. Eur. J.* **2016**, *22*, 9375–9386.

[357] Carpenter, J. E.; Grünwald, M. *Journal of the American Chemical Society* **2020**, *142*, 10755–10768.

[358] Schmidt, J. A.; Weatherby, J. A.; Sugden, I. J.; Santana-Bonilla, A.; Salerno, F.; Fuchter, M. J.; Johnson, E. R.; Nelson, J.; Jelfs, K. E. *Crystal Growth & Design* **2021**, *21*, 5036–5049.

[359] Lan, S.; Yan, Y.; Yang, H.; Zhang, G.; Ye, Y.; Li, F.; Chen, H.; Guo, T. *Journal of Materials Chemistry C* **2019**, *7*, 4543–4550.

[360] Bruno, I. J.; Cole, J. C.; Edgington, P. R.; Kessler, M.; Macrae, C. F.; McCabe, P.; Pearson, J.; Taylor, R. *Acta Crystallographica Section B* **2002**, *58*, 389–397.

[361] Rappé, A. K.; Casewit, C. J.; Colwell, K. S.; Goddard, W. A.; Skiff, W. M. *J. Am. Chem. Soc.* **1992**, *114*, 10024–10035.

[362] Roos, K.; Wu, C.; Damm, W.; Reboul, M.; Stevenson, J. M.; Lu, C.; Dahlgren, M. K.; Mondal, S.; Chen, W.; Wang, L.; Abel, R.; Friesner, R. A.; Harder, E. D. *J. Chem. Theory Comput.* **2019**, *15*, 1863–1874.

[363] Turcani, L.; Berardo, E.; Jelfs, K. E. *J. Comput. Chem.* **2018**, *39*, 1931–1942.

[364] Kudin, K. N.; Scuseria, G. E.; Cances, E. *The Journal of chemical physics* **2002**, *116*, 8255–8261.

[365] Neese, F. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2018**, *8*, e1327.

[366] Neese, F.; Hansen, A.; Liakos, D. G. *The Journal of chemical physics* **2009**, *131*, 064103.

[367] Riplinger, C.; Neese, F. *The Journal of chemical physics* **2013**, *138*, 034106.

[368] Chen, T.; Guestrin, C. XGBoost: A scalable tree boosting system. Proc. ACM SIGKDD Int. Conf. Knowl. Discov. Data Min. 2016; pp 785–794.

[369] Robert, C. Machine Learning, a Probabilistic Perspective. 2014.

[370] Rogers, D.; Hahn, M. *J. Chem. Inf. Model.* **2010**, *50*, 742–754.

[371] Price, S. L.; Leslie, M.; Welch, G. W. A.; Habgood, M.; Price, L. S.; Karamertzanis, P. G.; Day, G. M. *Phys. Chem. Chem. Phys.* **2010**, *12*, 8478–8490.

[372] Hu, H.; Lu, Z.; Yang, W. *Journal of chemical theory and computation* **2007**, *3*, 1004–1013.

[373] Otero-de-la Roza, A.; Cao, B. H.; Price, I. K.; Hein, J. E.; Johnson, E. R. *Angewandte Chemie International Edition* **2014**, *53*, 7879–7882.

[374] Leblanc, L. M.; Weatherby, J. A.; Otero-De-La-Roza, A.; Johnson, E. R. *J. Chem. Theory Comput.* **2018**, *14*, 5715–5724.

[375] Perdew, J. P.; Burke, K.; Ernzerhof, M. *Phys. Rev. Lett.* **1996**, *77*, 3865–3868.

[376] Becke, A. D. *J. Chem. Phys.* **1986**, *85*, 7184–7187.

[377] Troullier, N.; Martins, J. L. *Phys. Rev. B* **1991**, *43*, 8861–8869.

[378] Hamann, D. R.; Schlüter, M.; Chiang, C. *Phys. Rev. Lett.* **1979**, *43*, 1494.

[379] Bachelet, G. B.; Hamann, D. R.; Schlüter, M. *Phys. Rev. B* **1982**, *26*, 4199–4228.

[380] Carter, D. J.; Rohl, A. L. *J. Chem. Theory Comput.* **2014**, *10*, 3423–3437.

[381] Pack, J. D.; Monkhorst, H. J. *Phys. Rev. B* **1977**, *16*, 1748–1749.

[382] Nielsen, C. B.; White, A. J.; McCulloch, I. *The Journal of organic chemistry* **2015**, *80*, 5045–5048.

[383] Nematiaram, T.; Troisi, A. *Mater. Horiz.* **2020**, *7*, 2922–2928.

[384] Illig, S.; Eggeman, A. S.; Troisi, A.; Jiang, L.; Warwick, C.; Nikolka, M.; Schweicher, G.; Yeates, S. G.; Geerts, Y. H.; Anthony, J. E.; Sirringhaus, H. *Nature communications* **2016**, *7*, 1–10.

[385] Zhang, B.; Kan, Y.-H.; Geng, Y.; Duan, Y.-A.; Li, H.-B.; Hua, J.; Su, Z.-M. *Organic Electronics* **2013**, *14*, 1359 – 1369.

[386] Tang, M. L.; Oh, J. H.; Reichardt, A. D.; Bao, Z. *Journal of the American Chemical Society* **2009**, *131*, 3733–3740.

[387] Li, M.-J.; Long, M.-Q.; Chen, K.-Q.; Xu, H. *Solid state communications* **2013**, *157*, 62–67.

[388] Jones, B. A.; Facchetti, A.; Wasielewski, M. R.; Marks, T. J. *J. Am. Chem. Soc.* **2007**, *129*, 15259–15278.

[389] Di Donato, E.; Fornari, R. P.; Di Motta, S.; Li, Y.; Wang, Z.; Negri, F. *J. Phys. Chem. B* **2010**, *114*, 5327–5334.

[390] Maiti, B.; Schubert, A.; Sarkar, S.; Bhandari, S.; Wang, K.; Li, Z.; Geva, E.; Twieg, R. J.; Dunietz, B. D. *Chem. Sci.* **2017**, *8*, 6947–6953.

[391] Koch, N.; Vollmer, A.; Duhm, S.; Sakamoto, Y.; Suzuki, T. *Adv. Mater.* **2007**, *19*, 112–116.

[392] Kleemann, H.; Schuenemann, C.; Zakhidov, A. A.; Riede, M.; Lüssem, B.; Leo, K. *Org. Electron.* **2012**, *13*, 58–65.

[393] Tang, M. L.; Reichardt, A. D.; Wei, P.; Bao, Z. *J. Am. Chem. Soc.* **2009**, *131*, 5264–5273.

[394] Mottishaw, J. D.; Sun, H. *J. Phys. Chem. A* **2013**, *117*, 7970–7979.

[395] Trotter, J. *Acta Crystallographica* **1963**, *16*, 605–608.

[396] Hendon, C. H.; Tiana, D.; Murray, A. T.; Carbery, D. R.; Walsh, A. *Chemical Science* **2013**, *4*, 4278–4284.

[397] Bazterra, V. E.; Ferraro, M. B.; Facelli, J. C. *The Journal of Chemical Physics* **2002**, *116*, 5984–5991.

[398] Schneider, E.; Vogt, L.; Tuckerman, M. E. *Acta Crystallographica Section B: Structural Science, Crystal Engineering and Materials* **2016**, *72*, 542–550.

[399] Della Valle, R. G.; Venuti, E.; Brillante, A. *Chemical physics* **1995**, *198*, 79–89.

[400] Day, G. M. *Crystallography Reviews* **2011**, *17*, 3–52.

[401] McIntosh, A.; Robertson, J. M.; Vand, V. *Journal of the Chemical Society (Resumed)* **1954**, 1661–1668.

[402] Bédard, A.-C.; Vlassova, A.; Hernandez-Perez, A. C.; Bessette, A.; Hanan, G. S.; Heuft, M. A.; Collins, S. K. *Chemistry–A European Journal* **2013**, *19*, 16295–16302.

[403] Kuroda, R. *J. Chem. Soc., Perkin Trans. 2* **1982**, 789–794.

[404] Kaehler, T.; John, A.; Jin, T.; Bolte, M.; Lerner, H.-W.; Wagner, M. *European Journal of Organic Chemistry* **2020**, *2020*, 5847–5851.

[405] Hahn, T. *Acta Crystallographica* **1958**, *11*, 825–825.

[406] Dračínskỳ, M.; Storch, J.; Církva, V.; Císařová, I.; Sỳkora, J. *Physical Chemistry Chemical Physics* **2017**, *19*, 2900–2907.

[407] Fuchter, M. J.; Weimar, M.; Yang, X.; Judge, D. K.; White, A. J. *Tetrahedron Letters* **2012**, *53*, 1108–1111.

[408] Fujino, S.; Yamaji, M.; Okamoto, H.; Mutai, T.; Yoshikawa, I.; Houjou, H.; Tani, F. *Photochemical & Photobiological Sciences* **2017**, *16*, 925–934.

[409] Mori, K.; Murase, T.; Fujita, M. *Angewandte Chemie* **2015**, *127*, 6951–6955.

[410] Stuart, J. G.; Quast, M. J.; Martin, G. E.; Lynch, V. M.; Simonsen, S. H.; Lee, M. L.; Castle, R. N.; Dallas, J. L.; John, B. K.; Johnson, L. F. *Journal of heterocyclic chemistry* **1986**, *23*, 1215–1234.

[411] Alvarez, S. *Comptes Rendus Chimie* **2019**, *22*, 437–444.

[412] Bernstein, J. *Crystal Growth & Design* **2011**, *11*, 632–650.

[413] Ravat, P. *Chemistry (Weinheim an der Bergstrasse, Germany)* **2021**, *27*, 3957.

[414] Stone, M. T.; Heemstra, J. M.; Moore, J. S. *Accounts of chemical research* **2006**, *39*, 11–20.

[415] Hinkle, D. E.; Wiersma, W.; Jurs, S. G. *Applied statistics for the behavioral sciences*; Houghton Mifflin College Division, 2003; Vol. 663.

[416] Sinnokrot, M. O.; Sherrill, C. D. *The Journal of Physical Chemistry A* **2006**, *110*, 10656–10668.

[417] Ma, Z.; Geng, H.; Wang, D.; Shuai, Z. *Journal of Materials Chemistry C* **2016**, *4*, 4546–4555.

[418] Onwubiko, A. et al. *Nature communications* **2018**, *9*, 1–9.

[419] Fu, C.; Beldon, P. J.; Perepichka, D. F. *Chemistry of Materials* **2017**, *29*, 2979–2987.

[420] Cardia, R.; Malloci, G.; Mattoni, A.; Cappellini, G. *The Journal of Physical Chemistry A* **2014**, *118*, 5170–5177.

[421] Francese, T.; Ribas-Arino, J.; Novoa, J. J.; Havenith, R. W.; Broer, R.; de Graaf, C.; Deumal, M. *Physical Chemistry Chemical Physics* **2018**, *20*, 20406–20416.

[422] Schenning, A. P.; Meijer, E. *Chemical Communications* **2005**, 3245–3258.

[423] Hoeben, F. J.; Jonkheijm, P.; Meijer, E.; Schenning, A. P. *Chemical reviews* **2005**, *105*, 1491–1546.

[424] Lindner, S. M.; Kaufmann, N.; Thelakkat, M. *Organic electronics* **2007**, *8*, 69–75.

[425] Hong, J.; Xiao, X.; Liu, H.; Fu, L.; Wang, X.-C.; Zhou, L.; Wang, X.-Y.; Qiu, Z.; Cao, X.-Y.; Narita, A.; Müllen, K.; Hu, Y. *Chemical Communications* **2021**,

Appendix A

Appendix to Chapter 3: Structure-property Relationships of 2,2'-Diethynyl- and 2,2'-Dicyano-[6]helicenes: Do Chiral Columns Promote Charge Transport?

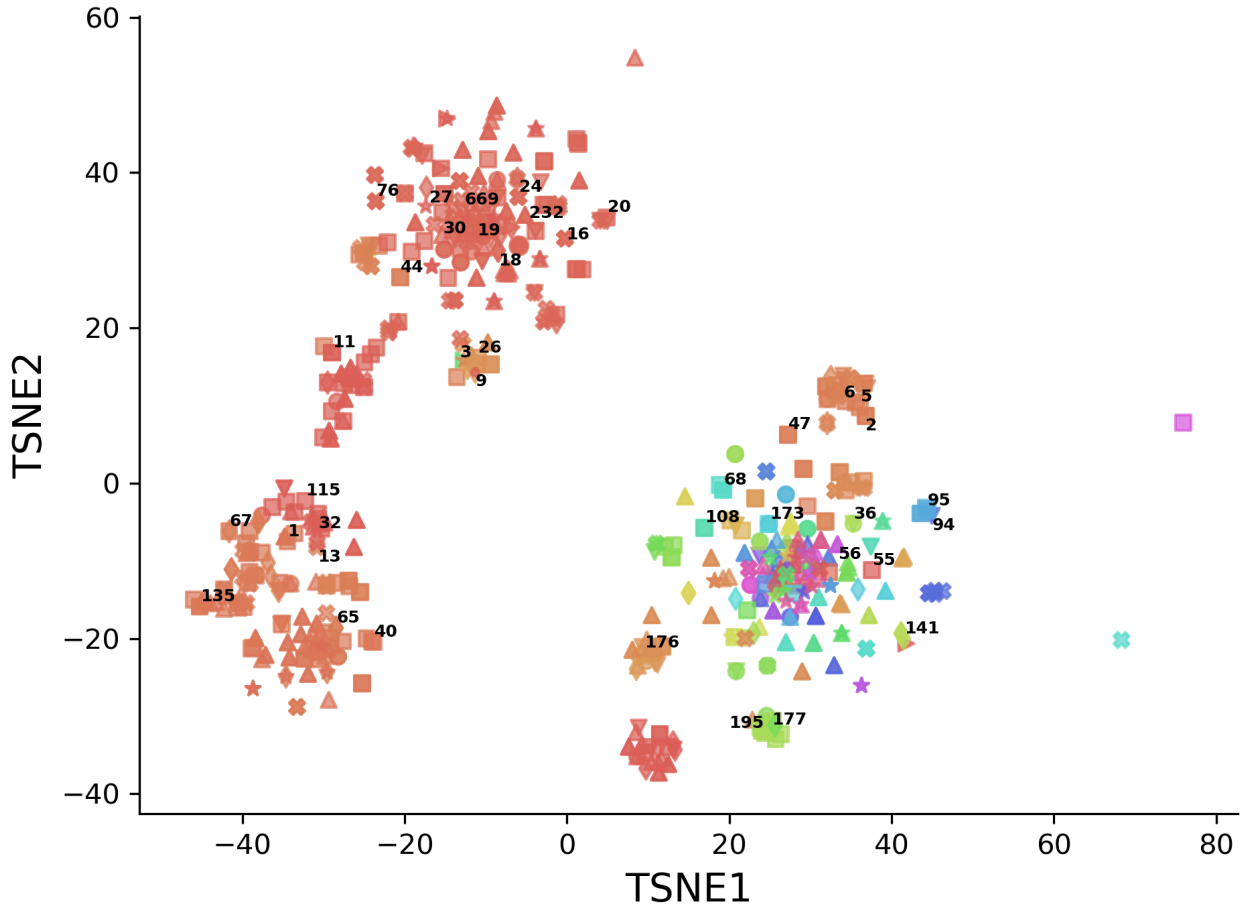


Figure A.1: Louvain clustering on the PM (4,6) graph similarity kernel of 2,2'-dicyano[6]helicene identified 100 unique clusters. Colors illustrate cluster affiliation, markers indicate which of the three common dimer motifs were found in the five closest unique dimers of the crystal neighbouring shell.

Appendix B

Appendix to Chapter 4: Computational Screening of Chiral Organic Semiconductors: Exploring Side-group Functionalisation and Assembly to Optimise Charge Transport

Table B.1: Results of a dimer-based motif search in substituted [6]helicene CSD crystals. The table highlights whether the crystal structure contains the translational motif (or any of the previously classified motifs in Figure 1.5) (1) or not (0). The motifs were identified by visual inspection. Reprinted with permission from 358. Copyright 2021 American Chemical Society.

REFCODE	**Translational**	**Herringbone**	**Homochiral back-to-back**	**Heterochiral back-to-back**	**Homochiral interlocked**	**Heterochiral interlocked**
BORGOG	0	0	0	0	0	0
DATLOB	0	0	0	1	0	0
DMHXIC	0	1	0	0	0	0
DOMTUW	0	0	0	0	0	0
EGOLOD	0	1	0	0	0	0
FHXHLB	0	0	0	0	0	0

Table B.1 – continued from previous page

REFCODE	Trans-lational	Herring-bone	Homochiral back-to-back	Heterochiral back-to-back	Homochiral interlocked	Heterochiral interlocked
FIJMUI	0	0	0	0	0	0
GEBTIT	0	1	0	0	0	0
GEBTOZ	0	0	0	0	0	0
GEFJUY	0	1	0	0	0	0
GOJDAK	0	0	0	1	0	0
LIYQAL	0	0	0	0	0	0
NEMKOI	0	0	0	0	0	0
NOQTIW	0	0	0	0	0	0
ODAHAF	0	0	0	0	0	0
OTUROM	0	0	0	1	0	1
OWUSOQ	0	0	0	0	0	0
PAWQAI	0	1	0	0	0	0
PAWQEM	0	0	0	0	0	0
PIFSUV	0	0	0	1	0	1
POZFIT	0	0	0	0	0	0
POZFIT01	0	0	0	0	0	0
QEPFEZ	0	0	0	0	0	0
QEYWAS	0	0	0	0	0	0
QONHUW	0	0	0	0	0	0
QONJAE	0	0	0	0	0	0
QONJEI	0	0	0	0	0	0
SEQHUR	0	0	0	0	0	0
TASHED	0	0	0	0	0	0
TASHON	0	0	0	0	0	0
TEFTIJ	0	0	0	0	0	0
TEFTOP	0	0	0	0	0	0

Table B.1 – continued from previous page

REFCODE	Trans-lational	Herring-bone	Homochiral back-to-back	Heterochiral back-to-back	Homochiral interlocked	Heterochiral interlocked
TEFVEH	0	0	0	0	0	0
TIFLUS	0	0	0	0	0	0
TODWAM	0	1	0	0	0	0
TODWEQ	0	0	0	0	0	0
UDILUR	0	0	0	0	0	0
UJOCUR	0	0	0	0	0	0
UJODEC	0	1	0	1	0	0
UKISAJ	0	0	0	0	0	0
UKISAJ01	0	0	0	1	0	0
UKISAJ02	0	0	0	0	0	0
UKISAJ03	0	0	0	1	0	0
UKISAJ04	0	0	0	1	0	0
VASWOE	0	1	0	0	0	0
WEXVIF	0	0	0	0	0	0
WURPIK	0	0	0	0	0	0
WURTAG	0	0	0	0	0	0
XEMNOT	0	0	0	0	0	0
XIFHOM	0	0	0	0	0	0
XITZIJ	0	0	0	1	0	1
XOCXAQ	0	1	0	0	0	0
YALSOV	0	0	0	0	0	0
YALSUB	0	1	1	0	0	0
YIGHED	0	0	0	0	0	0
YIGHUT	0	0	0	0	0	1
YIGJAB	0	0	0	0	0	0
YIHXAR	0	0	0	0	0	0

Table B.1 – continued from previous page

REFCODE	Trans-lational	Herring-bone	Homochiral back-to-back	Heterochiral back-to-back	Homochiral interlocked	Heterochiral interlocked
ALUQUT	1	1	1	0	0	0
BHXHEL20	1	0	1	0	1	0
BOZXUK	1	0	0	0	0	0
CIYSIM	1	0	0	0	0	0
DONBAL	1	0	0	0	0	0
FHXHLA	1	1	0	1	0	0
FILNUL	1	0	1	0	1	0
FILPOH	1	0	1	0	1	0
FILPUN	1	0	1	0	1	0
LIYSUJ	1	1	0	0	0	0
MEHXHE	1	0	0	1	0	1
MHXHEL10	1	0	1	0	1	0
NEMJOH	1	1	0	0	0	0
NEMJUN	1	0	0	1	0	1
NEMKUO	1	1	1	0	0	0
OWUSAC	1	1	0	0	0	0
OWUSEG	1	1	0	0	0	0
PICFEO01	1	1	1	0	0	0
QEPFID	1	1	1	0	0	0
QUENZER	1	0	0	0	0	0
TEFVAD	1	0	1	0	0	0
UGOTAM	1	0	0	1	0	1
UJOCOL	1	1	1	0	0	0
UJODAY	1	0	0	1	0	1
VASWUK	1	0	0	1	0	1
XIFHUS	1	0	0	0	0	0

Table B.1 – continued from previous page

REFCODE	Trans-lational	Herring-bone	Homochiral back-to-back	Heterochiral back-to-back	Homochiral interlocked	Heterochiral interlocked
YIGHIH	1	0	0	0	0	0
ZERQAR	1	0	0	0	0	0
ZERQEV	1	0	1	0	0	0

Table B.2: A list of translational dimer intermolecular distances found for hypothetical crystal structures obtained from the 6-triacetylene-[6]helicene CSP. The crystal ID refers to the original crystal number from the global search. The motifs were identified by visual inspection. The lowest-energy crystals containing the translational motif and whose mobilities are discussed in the main paper are thus crystal IDs: 52, 948, 175 and ranked by energy from the global minimum: crystal # 2,5,10 respectively. The energy (ΔE) is given in kJ mol^{-1} measured relative to the lowest energy structure (global minimum). The density (ρ) is given in absolute values and units of g cm^{-3}. The translational dimer intermolecular distance (d_{trans}) is given in Å. In case where the translational dimer was not 100% purely translational, due to the triacetylene group being attached at opposite ends of the backbone, and the exact intermolecular distance thus varied slightly with each atom, an approximated intermolecular distance is provided using ~ approximation symbol. Reprinted with permission from 358. Copyright 2021 American Chemical Society.

Crystal ID	ΔE	ρ	d_{trans}
52	0.28	1.22	5.55
948	0.87	1.22	5.55
175	3.47	1.23	5.35
2	4.50	1.22	~5.5
6	4.52	1.24	~5.52
83	4.58	1.19	6.31
10	4.59	1.22	5.35
3	4.66	1.23	~5.5
517	6.27	1.16	5.44

Continued on next page

Table B.2 – continued from previous page

Crystal ID	ΔE	ρ	d_{trans}
153	6.76	1.20	5.51
193	6.78	1.18	5.46
298	7.71	1.18	5.4
15	7.74	1.25	5.47
125	7.87	1.21	5.47
68	7.90	1.23	5.49
23	7.99	1.20	5.96
318	8.07	1.22	5.39
98	8.13	1.23	5.39
648	9.05	1.17	5.40
195	9.12	1.21	5.36
32	9.34	1.22	5.72
309	10.87	1.18	5.53
48	11.21	1.20	6.21
440	12.03	1.21	5.60
174	12.67	1.20	5.43
130	12.92	1.20	7.27
386	13.32	1.19	5.77
204	13.33	1.22	5.59
703	13.98	1.17	6.39
302	14.53	1.21	5.43
74	14.94	1.19	8.58
1265	15.43	1.18	5.61
287	16.24	1.15	6.38
340	16.38	1.18	6.38
736	16.96	1.18	~5.0

Continued on next page

Table B.2 – continued from previous page

Crystal ID	ΔE	ρ	d_{trans}
63	17.12	1.19	~6.32
788	18.02	1.20	6.21
122	19.31	1.18	8.24
147	20.76	1.18	8.72
308	21.80	1.18	8.69
508	24.19	1.17	9.07

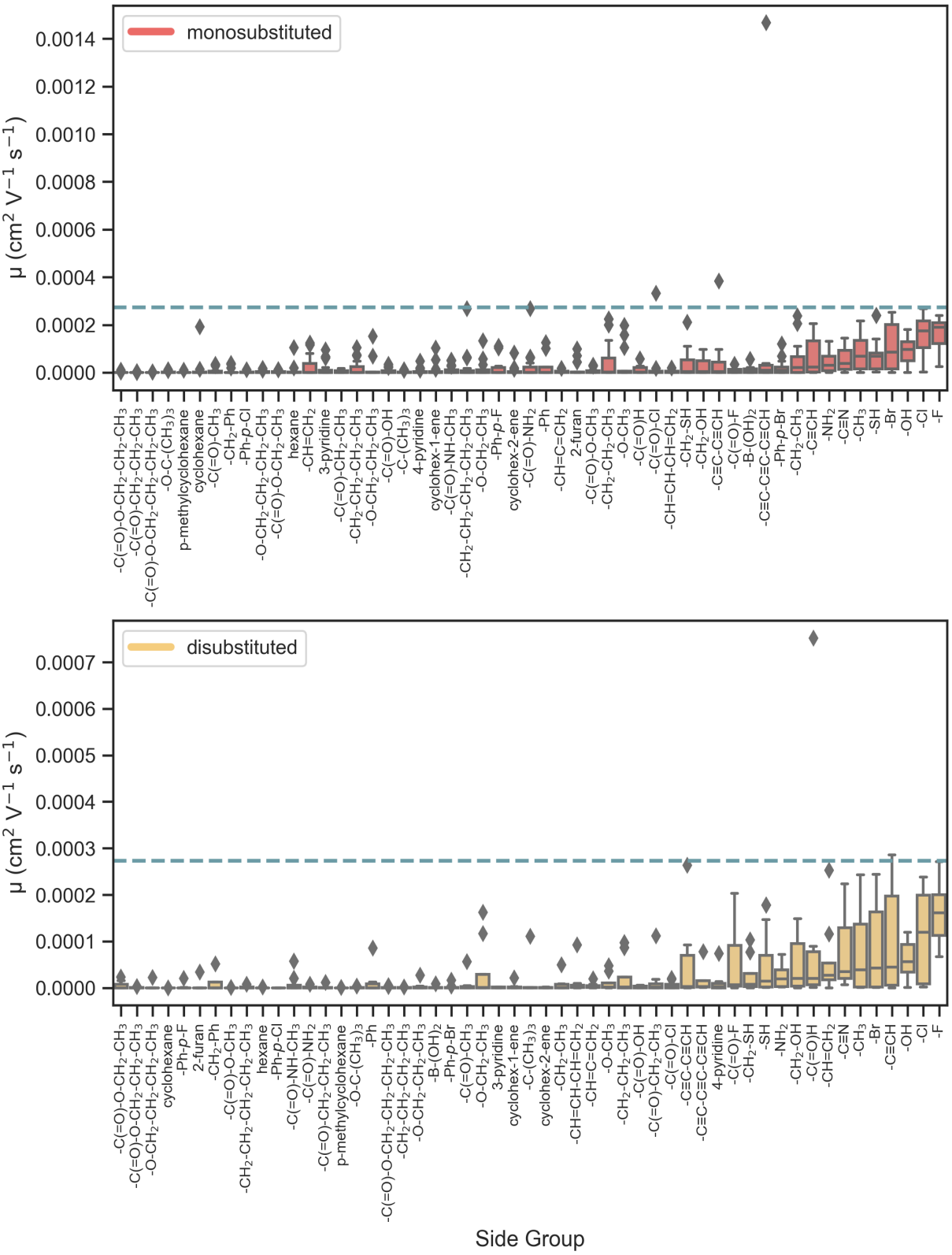


Figure B.1: Low field mobilities per functional group attached to the [6]helicene backbone, computed across the translational dimer model and sorted by ascending median. Reprinted with permission from 358. Copyright 2021 American Chemical Society.

Appendix C

Appendix to Chapter 5: On the Effect of [n]helicene Length on Crystal Packing Interactions and Polymorphism

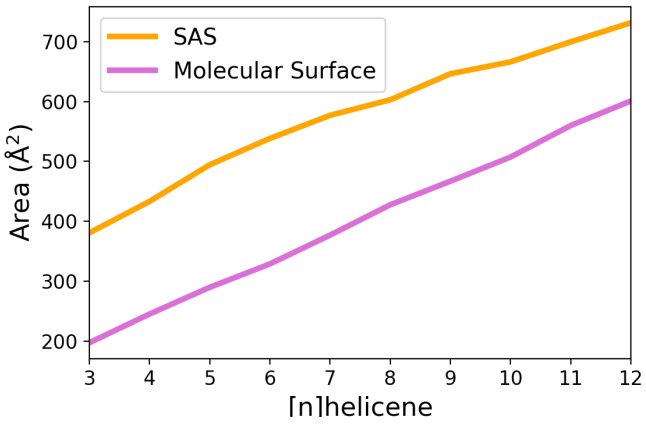


Figure C.1: Molecular surface area per [n]helicene molecule. The molecular surface area is given in pink and the solvent accessible surface area (SASA) in orange. The areas were computed based on the DFT-optimised single molecule geometries and using PyMOL (PyMOL Molecular Graphics System, Version 2.3.5, Schrödinger, LLC).

C.1 CRYSTACK Technical Details

The package consists of two main classes; a dimer class was created to store molecule A and molecule B and their respective identification numbers in the crystal lattice. Similarly, an inter-

actions class was created to obtain and store all intermolecular interactions information, such as π-π crystal stacking interactions, short intermolecular contacts and hydrophobic interactions across molecules.

An aromatic ring is considered planar if the maximum deviation from planarity (*i.e.* 0° or 180°) of the ring is within 5°.[339] Furthermore, for the purpose of the helicenes, the smallest permitted aromatic ring size was set to 4 (4-membered rings) and the largest permitted ring size was set to 6 (6-membered rings). For each detected ring, the ring centre was computed as the mean position across all ring atoms (assuming carbon-based rings only). Each detected (6-membered) ring was then analysed for whether it is a terminal or a bridged aromatic ring. To identify whether a ring was terminal or bridged (fused), the number of common ring atoms was compared with other rings detected for each molecule. If the aromatic ring consisted of four unique ring atoms, not identified as making part of other rings, the aromatic ring was considered terminal. Conversely, if only two unique ring atoms were identified and all remaining ring atoms were shared with other aromatic rings, the aromatic ring was considered bridged.

For the identification of π-π stacking interactions, maximum ring-to-ring distance cutoff for parallel and shifted parallel π-π stacking interactions was set to 5.5 Å (McGaughey, 1998).[339] The maximum angular deviation from a perfectly parallel (*i.e.* 0° or 180°) or perpendicular orientation (90°) was set to 30°.[339] The maximum offset of π-π stacking between two rings, *i.e.* if the ring centres are not perfectly on top of each other (shifted parallel vs. perfectly parallel π-π stacking), parallel π-π stacking interactions can have a maximum offset of 2.0 Å.[339] Hence, a parallel π-π stacking interaction was found if the angle between the two aromatic rings was between 0°±30° at a maximum separation ≤5.5 Å and a maximum offset between the ring centres of 2.0 Å. Conversely, a C-H$\cdots\pi$ stacking interaction was found if the angle between the two aromatic rings was near 90°, *i.e.* between 60°-120° (90°±30°) at a maximum separation ≤ 5.5 Å and a maximum offset between the ring centres of 2.0 Å.

Hydrophobic intermolecular contacts, *i.e.* close contacts between aromatic (hydrophobic) carbon centres (carbon-carbon or hydrogen-carbon) across two molecules, are considered if they are observed within a maximum distance cutoff of 4.0 Å.[339]

Methods to identify all short contacts was created inside the Interactions class. Inside the short

contacts method all atoms on molecule A and all atoms on molecule B are screened for potential short contacts by using the euclidean distance between each pair of atoms and the van der Waals radii of the respective atoms. The atomic van der Waals radii were obtained from the van der Waals radii table using a static method inside the Interactions class. If the euclidean distance is shorter than the sum of the van der Waals radii, a short contact was found. If more short contacts, *i.e.* shortest connections across molecules, are desirable the sum of the van der Waals radii can be increased by some float value (default 0.0 Å).

CCC Marketplace™

This is a License Agreement between Julia A. Schmidt ("User") and Copyright Clearance Center, Inc. ("CCC") on behalf of the Rightsholder identified in the order details below. The license consists of the order details, the CCC Terms and Conditions below, and any Rightsholder Terms and Conditions which are included below.

All payments must be made in full to CCC in accordance with the CCC Terms and Conditions below.

Order Date	21-Jul-2021	**Type of Use**	Republish in a thesis/dissertation
Order License ID	1134921-1		
ISSN	1466-8033	**Publisher**	ROYAL SOCIETY OF CHEMISTRY
		Portion	Chart/graph/table/figure

LICENSED CONTENT

Publication Title	CrystEngComm	**Publication Type**	e-Journal
Article Title	Polymorphism in p-aminobenzoic acid	**Start Page**	2034
		End Page	2042
Author/Editor	Royal Society of Chemistry (Great Britain)	**Issue**	13
		Volume	21
Date	01/01/1999	**URL**	http://www.rsc.org/Publishing/Journals/ce/index.asp
Language	English		
Country	United Kingdom of Great Britain and Northern Ireland		
Rightsholder	Royal Society of Chemistry		

REQUEST DETAILS

Portion Type	Chart/graph/table/figure	**Distribution**	Worldwide
Number of charts / graphs / tables / figures requested	1	**Translation**	Original language of publication
		Copies for the disabled?	No
Format (select all that apply)	Print, Electronic	**Minor editing privileges?**	Yes
Who will republish the content?	Academic institution	**Incidental promotional use?**	No
Duration of Use	Life of current edition	**Currency**	GBP
Lifetime Unit Quantity	Up to 499		
Rights Requested	Main product		

NEW WORK DETAILS

Title	Predicting structure-property relationships to accelerate chiral organic electronic materials discovery	**Institution name**	Imperial College London
		Expected presentation date	2022-07-01
Instructor name	Dr Kim Jelfs		

ADDITIONAL DETAILS

Order reference number	N/A	**The requesting person / organization to appear on the license**	Julia A. Schmidt

REUSE CONTENT DETAILS

Title, description or numeric reference of the portion(s)	Polymorphism in p-aminobenzoic acid	**Title of the article/chapter the portion is from**	Polymorphism in p-aminobenzoic acid
Editor of portion(s)	Ward, Martin R.; Sugden, Isaac J.; Oswald, Iain D. H.; Davey, Roger J.; Cruz-Cabeza, Aurora J.	**Author of portion(s)**	Ward, Martin R.; Sugden, Isaac J.; Oswald, Iain D. H.; Davey, Roger J.; Cruz-Cabeza, Aurora J.
Volume of serial or monograph	21	**Issue, if republishing an article from a serial**	13
Page or page range of portion	2034-2042	**Publication date of portion**	2019-01-01

CCC Terms and Conditions

1. Description of Service; Defined Terms. This Republication License enables the User to obtain licenses for republication of one or more copyrighted works as described in detail on the relevant Order Confirmation (the "Work(s)"). Copyright Clearance Center, Inc. ("CCC") grants licenses through the Service on behalf of the rightsholder identified on the Order Confirmation (the "Rightsholder"). "Republication", as used herein, generally means the inclusion of a Work, in whole or in part, in a new work or works, also as described on the Order Confirmation. "User", as used herein, means the person or entity making such republication.

2. The terms set forth in the relevant Order Confirmation, and any terms set by the Rightsholder with respect to a particular Work, govern the terms of use of Works in connection with the Service. By using the Service, the person transacting for a republication license on behalf of the User represents and warrants that he/she/it (a) has been duly authorized by the User to accept, and hereby does accept, all such terms and conditions on behalf of User, and (b) shall inform User of all such terms and conditions. In the event such person is a "freelancer" or other third party independent of User and CCC, such party shall be deemed jointly a "User" for purposes of these terms and conditions. In any event, User shall be deemed to have accepted and agreed to all such terms and conditions if User republishes the Work in any fashion.

3. Scope of License; Limitations and Obligations.

    3.1. All Works and all rights therein, including copyright rights, remain the sole and exclusive property of the Rightsholder. The license created by the exchange of an Order Confirmation (and/or any invoice) and payment by User of the full amount set forth on that document includes only those rights expressly set forth in the Order Confirmation and in these terms and conditions, and conveys no other rights in the Work(s) to User. All rights not expressly granted are hereby reserved.

    3.2. General Payment Terms: You may pay by credit card or through an account with us payable at the end of the month. If you and we agree that you may establish a standing account with CCC, then the following terms apply: Remit Payment to: Copyright Clearance Center, 29118 Network Place, Chicago, IL 60673-1291. Payments Due: Invoices are payable upon their delivery to you (or upon our notice to you that they are available to you for downloading). After 30 days, outstanding amounts will be subject to a service charge of 1-1/2% per month or, if less, the maximum rate allowed by applicable law. Unless otherwise specifically set forth in the Order Confirmation or in a separate written agreement signed by CCC, invoices are due and payable on "net 30" terms. While User may exercise the rights licensed immediately upon issuance of

the Order Confirmation, the license is automatically revoked and is null and void, as if it had never been issued, if complete payment for the license is not received on a timely basis either from User directly or through a payment agent, such as a credit card company.

3.3. Unless otherwise provided in the Order Confirmation, any grant of rights to User (i) is "one-time" (including the editions and product family specified in the license), (ii) is non-exclusive and non-transferable and (iii) is subject to any and all limitations and restrictions (such as, but not limited to, limitations on duration of use or circulation) included in the Order Confirmation or invoice and/or in these terms and conditions. Upon completion of the licensed use, User shall either secure a new permission for further use of the Work(s) or immediately cease any new use of the Work(s) and shall render inaccessible (such as by deleting or by removing or severing links or other locators) any further copies of the Work (except for copies printed on paper in accordance with this license and still in User's stock at the end of such period).

3.4. In the event that the material for which a republication license is sought includes third party materials (such as photographs, illustrations, graphs, inserts and similar materials) which are identified in such material as having been used by permission, User is responsible for identifying, and seeking separate licenses (under this Service or otherwise) for, any of such third party materials; without a separate license, such third party materials may not be used.

3.5. Use of proper copyright notice for a Work is required as a condition of any license granted under the Service. Unless otherwise provided in the Order Confirmation, a proper copyright notice will read substantially as follows: "Republished with permission of [Rightsholder's name], from [Work's title, author, volume, edition number and year of copyright]; permission conveyed through Copyright Clearance Center, Inc. " Such notice must be provided in a reasonably legible font size and must be placed either immediately adjacent to the Work as used (for example, as part of a by-line or footnote but not as a separate electronic link) or in the place where substantially all other credits or notices for the new work containing the republished Work are located. Failure to include the required notice results in loss to the Rightsholder and CCC, and the User shall be liable to pay liquidated damages for each such failure equal to twice the use fee specified in the Order Confirmation, in addition to the use fee itself and any other fees and charges specified.

3.6. User may only make alterations to the Work if and as expressly set forth in the Order Confirmation. No Work may be used in any way that is defamatory, violates the rights of third parties (including such third parties' rights of copyright, privacy, publicity, or other tangible or intangible property), or is otherwise illegal, sexually explicit or obscene. In addition, User may not conjoin a Work with any other material that may result in damage to the reputation of the Rightsholder. User agrees to inform CCC if it becomes aware of any infringement of any rights in a Work and to cooperate with any reasonable request of CCC or the Rightsholder in connection therewith.

4. Indemnity. User hereby indemnifies and agrees to defend the Rightsholder and CCC, and their respective employees and directors, against all claims, liability, damages, costs and expenses, including legal fees and expenses, arising out of any use of a Work beyond the scope of the rights granted herein, or any use of a Work which has been altered in any unauthorized way by User, including claims of defamation or infringement of rights of copyright, publicity, privacy or other tangible or intangible property.

5. Limitation of Liability. UNDER NO CIRCUMSTANCES WILL CCC OR THE RIGHTSHOLDER BE LIABLE FOR ANY DIRECT, INDIRECT, CONSEQUENTIAL OR INCIDENTAL DAMAGES (INCLUDING WITHOUT LIMITATION DAMAGES FOR LOSS OF BUSINESS PROFITS OR INFORMATION, OR FOR BUSINESS INTERRUPTION) ARISING OUT OF THE USE OR INABILITY TO USE A WORK, EVEN IF ONE OF THEM HAS BEEN ADVISED OF THE POSSIBILITY OF SUCH DAMAGES. In any event, the total liability of the Rightsholder and CCC (including their respective employees and directors) shall not exceed the total amount actually paid by User for this license. User assumes full liability for the actions and omissions of its principals, employees, agents, affiliates, successors and assigns.

6. Limited Warranties. THE WORK(S) AND RIGHT(S) ARE PROVIDED "AS IS". CCC HAS THE RIGHT TO GRANT TO USER THE RIGHTS GRANTED IN THE ORDER CONFIRMATION DOCUMENT. CCC AND THE RIGHTSHOLDER DISCLAIM ALL

OTHER WARRANTIES RELATING TO THE WORK(S) AND RIGHT(S), EITHER EXPRESS OR IMPLIED, INCLUDING WITHOUT LIMITATION IMPLIED WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE. ADDITIONAL RIGHTS MAY BE REQUIRED TO USE ILLUSTRATIONS, GRAPHS, PHOTOGRAPHS, ABSTRACTS, INSERTS OR OTHER PORTIONS OF THE WORK (AS OPPOSED TO THE ENTIRE WORK) IN A MANNER CONTEMPLATED BY USER; USER UNDERSTANDS AND AGREES THAT NEITHER CCC NOR THE RIGHTSHOLDER MAY HAVE SUCH ADDITIONAL RIGHTS TO GRANT.

7. Effect of Breach. Any failure by User to pay any amount when due, or any use by User of a Work beyond the scope of the license set forth in the Order Confirmation and/or these terms and conditions, shall be a material breach of the license created by the Order Confirmation and these terms and conditions. Any breach not cured within 30 days of written notice thereof shall result in immediate termination of such license without further notice. Any unauthorized (but licensable) use of a Work that is terminated immediately upon notice thereof may be liquidated by payment of the Rightsholder's ordinary license price therefor; any unauthorized (and unlicensable) use that is not terminated immediately for any reason (including, for example, because materials containing the Work cannot reasonably be recalled) will be subject to all remedies available at law or in equity, but in no event to a payment of less than three times the Rightsholder's ordinary license price for the most closely analogous licensable use plus Rightsholder's and/or CCC's costs and expenses incurred in collecting such payment.

8. Miscellaneous.

    8.1. User acknowledges that CCC may, from time to time, make changes or additions to the Service or to these terms and conditions, and CCC reserves the right to send notice to the User by electronic mail or otherwise for the purposes of notifying User of such changes or additions; provided that any such changes or additions shall not apply to permissions already secured and paid for.

    8.2. Use of User-related information collected through the Service is governed by CCC's privacy policy, available online here:https://marketplace.copyright.com/rs-ui-web/mp/privacy-policy

    8.3. The licensing transaction described in the Order Confirmation is personal to User. Therefore, User may not assign or transfer to any other person (whether a natural person or an organization of any kind) the license created by the Order Confirmation and these terms and conditions or any rights granted hereunder; provided, however, that User may assign such license in its entirety on written notice to CCC in the event of a transfer of all or substantially all of User's rights in the new material which includes the Work(s) licensed under this Service.

    8.4. No amendment or waiver of any terms is binding unless set forth in writing and signed by the parties. The Rightsholder and CCC hereby object to any terms contained in any writing prepared by the User or its principals, employees, agents or affiliates and purporting to govern or otherwise relate to the licensing transaction described in the Order Confirmation, which terms are in any way inconsistent with any terms set forth in the Order Confirmation and/or in these terms and conditions or CCC's standard operating procedures, whether such writing is prepared prior to, simultaneously with or subsequent to the Order Confirmation, and whether such writing appears on a copy of the Order Confirmation or in a separate instrument.

    8.5. The licensing transaction described in the Order Confirmation document shall be governed by and construed under the law of the State of New York, USA, without regard to the principles thereof of conflicts of law. Any case, controversy, suit, action, or proceeding arising out of, in connection with, or related to such licensing transaction shall be brought, at CCC's sole discretion, in any federal or state court located in the County of New York, State of New York, USA, or in any federal or state court whose geographical jurisdiction covers the location of the Rightsholder set forth in the Order Confirmation. The parties expressly submit to the personal jurisdiction and venue of each such federal or state court.If you have any comments or questions about the Service or Copyright Clearance Center, please contact us at 978-750-8400 or send an e-mail to support@copyright.com.

v 1.1

RightsLink®

Home | Help | Email Support | Julia Schmidt

Computational Screening of Chiral Organic Semiconductors: Exploring Side-Group Functionalization and Assembly to Optimize Charge Transport

Author: Julia A. Schmidt, Joseph A. Weatherby, Isaac J. Sugden, et al

Publication: Crystal Growth and Design

Publisher: American Chemical Society

Date: Aug 1, 2021

Copyright © 2021, American Chemical Society

PERMISSION/LICENSE IS GRANTED FOR YOUR ORDER AT NO CHARGE

This type of permission/license, instead of the standard Terms and Conditions, is sent to you because no fee is being charged for your order. Please note the following:

- Permission is granted for your request in both print and electronic formats, and translations.
- If figures and/or tables were requested, they may be adapted or used in part.
- Please print this page for your records and send a copy of it to your publisher/graduate school.
- Appropriate credit for the requested material should be given as follows: "Reprinted (adapted) with permission from {COMPLETE REFERENCE CITATION}. Copyright {YEAR} American Chemical Society." Insert appropriate information in place of the capitalized words.
- One-time permission is granted only for the use specified in your RightsLink request. No additional uses are granted (such as derivative works or other editions). For any uses, please submit a new request.

If credit is given to another source for the material you requested from RightsLink, permission must be obtained from that source.

BACK CLOSE WINDOW

© 2021 Copyright - All Rights Reserved | Copyright Clearance Center, Inc. | Privacy statement | Terms and Conditions
Comments? We would like to hear from you. E-mail us at customercare@copyright.com