

Universität Stuttgart

Computational Analysis of Periodic Systems for Covalent Organic Frameworks and Molecules in High Electric Fields

Vom SC SimTech und der Fakultät für Chemie der Universität Stuttgart zur Erlangung der
Würde eines Doktors der Naturwissenschaften (Dr. rer. nat.) genehmigte Abhandlung

Vorgelegt von

Robin Schuldt

aus Sindelfingen

Hauptberichter:

Prof. Dr. Johannes Kästner

Mitberichter:

Prof. Dr. Thomas D. Kühne

Prüfungsvorsitzender:

Prof. Dr. Blazej T. Grabowski

Tag der mündlichen Prüfung:

17. Juni 2022

Institut für Theoretische Chemie

Universität Stuttgart

2022

“ “ *The task is not to see what has never been seen before,
but to think what has never been thought before about
what you see everyday.*

” ”

Erwin Schrödinger

Danksagung

Ohne die Unterstützung zahlreicher Menschen die mir durch ihre Beiträge und Unterhaltungen vielfältige Eindrücke und Inspirationen lieferten, wäre das Schreiben dieser Arbeit nicht möglich gewesen.

An erster Stelle möchte ich mich bei Johannes Kästner bedanken, der es mir ermöglichte meine Arbeit in einem so abwechslungsreichen Themenbereich zu verfassen. Ihm verdanke ich nicht nur eine ausgezeichnete Betreuung während der gesamten Promotion, sondern auch die Freiheit meine Arbeit nach eigenen Vorstellungen zu gestalten. In meiner Zeit in seiner Arbeitsgruppe konnte ich viel über den professionellen Umgang mit Mitarbeitern, Kooperationspartnern und etwaigen Komplikationen lernen.

Besonders möchte ich mich bei Thomas Kühne für die Übernahme des Zweitgutachtens sowie bei Blazej Grabowski für die Übernahme des Prüfungsvorsitzes bedanken.

Weiter danke ich der Deutschen Forschungsgemeinschaft für die Finanzierung meiner Arbeit im Rahmen des Sonderforschungsbereichs “Molekulare heterogene Katalyse in definierten, dirigierenden Geometrien” (SFB 1333), sowie der europäischen Union, die mich über das European Union’s Horizon 2020 research and innovation programme (grant agreement No. 646717, TUNNELCHEM) unterstützte.

Während meiner Promotion führten Gespräche mit Kollegen und Kooperationspartner stets zu neuen Einblicken und Erkenntnissen, auch in Bereichen außerhalb meiner Hauptforschungsfeldern und beflügelten meine Forschungsmotivation.

Ein besonderer Dank gilt an dieser Stelle Stefan Naumann der mir erste Einblicke in die organische Chemie N-Heterozyklischer Olefine ermöglichte und mich beim Schreiben meiner ersten Veröffentlichung unterstützte. Sebastian Emmerling möchte ich für die stets bereichernden Diskussionen über COFs sowie deren strukturellen Eigenschaften danken. Christopher Kessler möchte ich für die abwechslungsreichen Themen unserer Debatten, sowie die problemlose Kommunikation während unserer Projekte zur Adsorption in COFs danken. Außerdem danke ich Helena Solodenko für interessante Gespräche zur Atomsondentomographie und die gute Zusammenarbeit.

Eine Doktorarbeit schreibt sich nicht über Nacht, viele Kollegen begleiteten mich während meiner gesamten Promotionszeit, ihnen allen danke ich für ihre vielfältige Unterstützung.

Vier Jahre in einem Büro... An dieser Stelle danke ich Daniel Born, wir kamen, wir sahen und wir lachten. Durch viele schwierige Zeiten sind wir gegangen und viel haben wir von einander gelernt, was wäre die Promotion ohne Balkoni gewesen?

Außerdem danke ich Sina Klostermann für etwaige Korrekturlesungen und so manche abwechslungsreiche Kaffeepause sowohl mit als auch ohne Kekse. Katrin Gugeler, Juliane Heitkämper und Alexander Waigum danke ich für ihre stets offenen Ohren und etwaiges Korrekturlesen. An dieser Stelle danke ich auch Ai' Lan Dietrich, wer weiß wie viele Termine ich ohne sie verpasst hätte. Des Weiteren danke ich Germán Molpeceres für seine Ratschläge sowie Viktor Zaverkin für inspirierende Gespräche.

Ich danke auch Manoj Kumar Kesharwani, der mich vor allem zu Beginn meiner Promotion durch seine Erfahrung unterstützte. Hier möchte ich auch meinem ersten Mentor Jörg Main danken, der mich auf den Anfängen meiner akademischen Laufbahn begleitete.

Des Weiteren danke ich meinen Kollegen aus der technischen Chemie sowie der theoretischen Physik für abwechslungsreiche Gespräche.

An dieser Stelle danke ich auch all meinen Freunden außerhalb der Universität für ihre Zeit und die Abwechslung die Sie mir von der Promotion boten. Jeder hat auf die eine oder andere Art und Weise zum Gelingen dieser Arbeit beigetragen.

Zu guter Letzt gilt mein Dank meiner Familie, ohne die ich niemals in die Position gekommen wäre auch nur an eine Doktorarbeit zu denken. Seit fast 27 Jahren konnte ich mich auf die finanzielle sowie moralische Unterstützung meiner Eltern verlassen. Viele meiner Launen habt ihr ausgehalten und mich unterstützt wo ihr nur konntet, ich schaue zurück im Wissen immer in eurer Schuld zu stehen.

Es bleibt mein liebevoller Dank für meine Schwester Tamara, die mir stets ein "Tacitus" im Leben war und mich so gut sie konnte unterstützte. Du weisst: "Wer eine Schwester hat, der hat immer Rückendeckung."

Abstract

Technological progress relies on the discovery of previously unknown effects, however, in modern days their understanding requires also advanced model descriptions. This thesis aims at providing a theoretical framework for modeling structures and processes in the emerging scientific fields of atom probe tomography (APT) and covalent organic frameworks (COFs). While both fields are frequent targets of experiments, the contrary holds for theoretical treatments.

Atom Probe Tomography

In APT, sharp needles are subjected to strong electric fields, which lead to the desorption of single atoms or small clusters. Their trajectory can be reconstructed to provide information on the 3D geometry of the sample with chemical information. Developed initially for the analysis of metals, recent technical advances opened up APT's applicability to the analysis of covalently bound systems. To aid interpretation of experimentally obtained fragmentation patterns, the desorption of self-assembled monolayers (SAMs) from a gold substrate is modeled using density functional theory. Previous investigations of SAMs on metal surfaces were performed for gas-phase models of isolated molecules whereas this work uses a bulk approach. SAMs of decanethiolate (C10-SAM) and its modifications amino-decanethiolate (NH₂-SAM) and per-fluoro-decanethiolate (F-SAM) were investigated. Preliminary studies of SAM-substrate interfaces revealed common, and modification specific, properties for the three materials systems. Simulations revealed molecular tilting, stabilization due to periodic images and staggered segments of adjacent molecules for all species. Furthermore, NH₂-SAM showed hydrogen bonds between adjacent molecular head-groups whereas F-SAM showed increased intermolecular repulsion. Strategies for increasing the flexibility of density description were developed to overcome self consistent field (SCF)-convergence issues associated with the systems' polarization and the periodic electric field. Analysis of electronic properties during the modeled evaporation process revealed stepwise charge accumulation in the molecules' head groups yielding increased polarization in field direction. Stabilization due to interaction with periodic images and intermolecular interactions for surface models resulted in smaller

fragments and higher evaporation fields than gas-phase simulations. The computed evaporation field strengths agreed well with the experimental ones. The insights gained will facilitate APT analysis of covalently bound structures and shed light onto necessary fundamentals for simulating field evaporation with periodic models.

Covalent Organic Frameworks

Structural models for theoretical analysis of two-dimensional COFs usually assume perfect stacking schemes, avoiding the consideration of structural impurities due to interlayer slipping and mixed stacking. Largely unconsidered in previous studies, tuning of slipping effects and stacking is of great significance for the design of large-pore COFs, where pore occlusion and collapse commonly prevail. Rationalizing the design of possible species, therefore, requires understanding of the impact of structural features on the effective pore size. A suitable description for COFs based on density functional theory is established for literature-known examples, targeting structural properties affecting stacking order and pore size. The study revealed the relevance of non-covalent interlayer interactions for preference of specific stacking schemes. The observations made for component associated interlayer interactions inspired the modeling of a hypothetical system, which is predicted to maximize pore sizes by favored eclipsed stacking. Furthermore, the possible stabilizing impact of methoxy groups is exemplified by comparison of a literature-known non-planar example and its methoxylated version. The obtained methodology is further applied to explain the correlation of methoxylation for an isorecticular series of large-pore imine COFs with its structural stabilization by systematic analysis of components' interactions. Reduction of interlayer slipping and increase in order of component orientations for eclipsed configurations by addition of methoxy groups acting as molecular "anchors". Secondly, the established methodology of methoxylation was investigated in combination with fluorinated phenyl linkers. Studies of possible linker arrangements predicting T-shape stacking to be energetically favored and stabilizing for eclipsed configurations. The obtained framework for COF modeling was shown to successfully explain experimental observations of increased pore size and stability corroborating its predictive potential. The conducted studies illuminate possible pathways for future design of large-pore COFs based on theoretical predictions.

Zusammenfassung

Technologischer Fortschritt basiert stets auf dem Entdecken bisher unbekannter Effekte, jedoch werden heutzutage zur Beschreibung experimenteller Beobachtungen aufgrund steigender Komplexität auch fortschrittlichere Modelle benötigt. Im Rahmen dieser Arbeit wurden Methoden zur theoretischen Beschreibung von Strukturmodellen und Prozessen in den neuerdings rasch an Bedeutung gewinnenden Feldern der kovalenten organischen Gerüstverbindungen (eng. covalent organic frameworks (COFs)) und Atomsondentomographie (eng. atom probe tomography (APT)) entwickelt. Die Ausschöpfungsmöglichkeiten beider Forschungsfelder werden durch zahlreiche experimentelle Arbeiten ständig erweitert, dennoch liegt bis zum heutigen Tage lediglich eine sehr überschaubare Zahl theoretischer Untersuchungen vor.

Atomsondentomographie

In der Atomsondentomographie werden Substrate auf Metallspitzen einem starken elektrischen Feld ausgesetzt, was zur Verdampfung von ionisierten Fragmenten führt. Das Nachverfolgen der resultierenden Ionenflugbahnen erlaubt die dreidimensionale Rekonstruktion der zu untersuchenden Oberfläche. Neuerdings ermöglichen technische Verbesserungen, der anfangs nur für Metalle konzipierten Atomsonden, die Erweiterung ihres Anwendungsbereiches auf Flüssigkeiten und organische Verbindungen. Zur Unterstützung der Interpretation von experimentell erhaltenen Massenspektren, wird die Feldverdampfung selbstorganisierender Einzelschichten (eng. self-assembled monolayers SAMs) von einer Goldoberfläche, mithilfe der Dichtefunktionaltheorie untersucht. Bisherige theoretische Studien zu SAMs beschränkten sich auf einzelne SAM-Moleküle und modellierten die zugehörige Goldoberfläche mit einzelnen Atomen bzw. kleinen Verbunden, wohingegen diese Arbeit ein periodisches Oberflächenmodell verwendet. Drei Typen von SAM-Oberflächen, bestehend aus Decanthiolat (C10-SAM) und seinen modifizierten Versionen amino undecanthiolat (NH₂-SAM) sowie Perfluordecanthiolat (F-SAM), wurden eingehend untersucht. Der Vergleich der charakterisierten Oberflächenmodelle zeigte spezifische Eigenschaften, welche aus der periodischen Beschreibung sowie den Modifikationen der SAM Moleküle resultierten. Für alle drei Molekültypen wurden Verkippungen, eine stabilisierende Wirkung durch periodische Bilder, sowie die al-

ternierende Stapelung der Segmente festgestellt. Für Systeme des NH_2 -SAM wurden des Weiteren Wasserstoffbrücken zwischen den Kopfgruppen der Moleküle beobachtet. Im Falle der F-SAM-Systeme erzeugten F Atome eine erhöhte intermolekulare Abstoßung. Analysen der elektronischen Eigenschaften während des Feldverdampfungsprozesses zeigten eine schrittweise Erhöhung der netto-Partialladungen in der Kopfgruppe der SAM-Moleküle, resultierend in einer Systempolarisierung in Feldrichtung. Auftretende Konvergenzprobleme der Methode des selbstkonsistenten Feldes (eng. self-consistent field) sowie der periodischen Feldbeschreibung, konnten durch Entwicklung von Strategien zur Flexibilitätssteigerung der Elektronendichte überwunden werden. Die im Vergleich zu Gasphasenmodellen erhöhten Verdampfungsfeldstärken wurden auf eine Stabilisierung der Oberflächenmoleküle durch periodische Bilder sowie intermolekulare Wechselwirkungen zurückgeführt. Die theoretisch berechneten Feldstärken entsprachen in guter Näherung den experimentell gemessenen. Die durchgeführte Simulation des APT-Prozesses für organische Moleküle mithilfe einer Festkörperbeschreibung wird zukünftige Arbeiten in diesem Feld, durch die erhaltenen Einblicke in die fundamentalen Eigenschaften des periodischen Ansatzes, erleichtern.

Kovalente Organische Gerüstverbindungen

Strukturmodelle zur Beschreibung zweidimensionaler COFs nehmen in der Regel eine perfekte Stapelung der Schichten an, um die Behandlung von Strukturverunreinigungen, Verschiebungen zwischen zwei Schichten sowie gemischt auftretender Stapelmotive zu vermeiden. Für die Konzipierung von großporigen COFs, sind die bisher weitestgehend ununtersuchten Zwischenschichtwechselwirkungen zur gezielten Manipulation von Stapelmotiven, von signifikanter Bedeutung. Das Blockieren von Poren und deren Zusammenbrüche nach Befüllen mit Adsorbatmolekülen sind häufig festzustellende Probleme. Die systematische Konzipierung möglicher großporiger Spezies wird durch ein tiefgreifendes Verständnis des Zusammenhangs zwischen Strukturkomponenten und der erhaltenen effektiven Porengröße signifikant erleichtert. Hierzu wird eine geeignete Beschreibung mittels DFT anhand von literaturbekannten COFs, für die Zielgrößen der Stapelordnung und Porengröße, erstellt. Es zeigte sich der bedeutende Einfluss von nicht kovalenten Wechselwirkungen auf die Bevorzugung von Stapelmotiven. Die Beobachtungen für literaturbekannte COFs, wurden als Grundlage zur

Konzeption eines COFs mit maximierter Porengröße, durch Bevorzugung des ekliptischen Stapelmotivs, verwendet. Darüber hinaus konnte der stabilisierende Einfluss von Methoxygruppen anhand eines literaturbekannten nicht planaren COFs gezeigt werden. Mit dem entwickelten Vorgehen zur Strukturanalyse von COFs konnte der experimentell gefundene Zusammenhang zwischen Vergrößerung der Poren und hinzugefügten Methoxygruppen der isoretikularen PPB-TAB-COF Serie erläutert werden. Hierzu wurden die Wechselwirkungen zwischen einzelnen Komponenten sowie kleineren Verbänden in einer hierarchischen Studie analysiert, welche die Ankerwirkung der Methoxygruppen durch reduzierte Zwischenschichtverschiebungen und gesteigerter Ordnung der Komponentenorientierung verdeutlichte. In einer zweiten Studie wurde das entwickelte Vorgehen außerdem zur Untersuchung von fluorinierten Phenylgruppen verwendet. Die Betrachtung möglicher Orientierungen identifizierte die T-Stapelung als bevorzugtes Stapelmotiv, welches weiterführend die der ekliptische Anordnung stabilisiert. Die erfolgreiche Erklärung der experimentell gefundenen Ergebnisse, zu Porengröße und Stabilität, untermauern die Aussagekraft der theoretischen Methoden für mögliche Vorhersagen weiterer günstiger Wechselwirkungen für großporige COFs. Die Gesamtheit der durchgeführten Analysen, wirft ein Licht auf den potenziellen Weg zur gezielten Konzipierung großporiger COFs auf Basis theoretischer Untersuchungen.

Peer-Reviewed Publications

Articles included in this Thesis

- Carolin A. Dietrich, **Robin Schuldt**, Daniel Born, Helena Solodenko, Guido Schmitz and Johannes Kästner, “Evaporation and fragmentation of organic molecules in strong electric fields simulated with DFT”, *J. Phys. Chem. A*, 124(41), 8633-8642 (2020)
- Sebastian T. Emmerling, **Robin Schuldt**, Sebastian Bette, Liang Yao, Robert E. Dinnebier, Johannes Kästner and Bettina V. Lotsch “Interlayer Interactions as Design Tool for Large-Pore COFs”, *J. Am. Chem. Soc.*, 143(38), 15711-15722 (2021)
- Christoher Kessler, **Robin Schuldt**, Sebastian T. Emmerling, Bettina V. Lotsch, Johannes Kästner, Joachim Gross, and Niels Hansen, “Influence of layer slipping on adsorption of light gases in covalent organic frameworks: A combined experimental and computational study”, *Microporous Mesoporous Mater.*, 336, 111796 (2022)

Articles published during the PhD Studies not included in this Thesis

- **Robin Schuldt**, Johannes Kästner and Stefan Naumann, “Proton affinities of N-heterocyclic olefins and their implications for organocatalyst design”, *J. Org. Chem.*, 84(4), 2209-2218 (2019)
- Johannes Mögerle, **Robin Schuldt**, Johannes Reiff, Jörg Main and Rigoberto Hernandez, “Transition state dynamics of a driven magnetic free layer”, *Commun. Nonlinear Sci. Numer. Simul.*, 105, 106054 (2022)
- Viktor Zaverkin, David Holzmüller, **Robin Schuldt**, and Johannes Kästner, “Predicting properties of periodic systems from cluster data: A case study of liquid water”, *J. Chem. Phys.*, 156(11), 114103 (2022)

Contents

| | |
|---|-----------|
| Danksagung | V |
| Abstract | VII |
| Zusammenfassung | IX |
| List of Abbreviations | XXIII |
| 1 Introduction | 1 |
| 1.1 Simulation of Field Evaporation for Atom Probe Tomography | 1 |
| 1.2 Theoretical Description of Covalent Organic Frameworks | 3 |
| 1.3 Structure of this Thesis | 5 |
| I Theory | 9 |
| 2 Density Functional Theory | 11 |
| 2.1 The Challenge of Quantum Chemistry | 12 |
| 2.2 The General Idea of Hartree-Fock Methods | 15 |
| 2.3 Density Functional Theory | 18 |
| 2.3.1 Properties of the Density | 18 |
| 2.3.2 Hohenberg-Kohn Theorems | 19 |
| 2.4 Kohn-Sham Density Functional Theory | 20 |
| 2.5 The Exchange and Correlation Functional | 23 |
| 2.5.1 Local Density Approximation (LDA) | 23 |
| 2.5.2 Generalized Gradient Approximation (GGA) | 25 |
| 2.5.3 Meta-GGA Functionals | 27 |

Contents

| | | |
|----------|--|-----------|
| 2.5.4 | Hybrid Functionals | 28 |
| 2.6 | The LCAO Approach | 30 |
| 2.6.1 | Slater-Type Functions | 31 |
| 2.6.2 | Gaussian-Type Functions | 32 |
| 2.6.3 | Self Consistent Field Approach (SCF) | 32 |
| 2.7 | Density Functional Theory for Periodic Systems | 34 |
| 2.8 | Description of the Solid State | 37 |
| 2.8.1 | Periodic Atom Centered Gaussian Basis Functions | 38 |
| 2.8.2 | Plane-Wave Basis Sets | 39 |
| 2.9 | Energy Evaluation for Periodic Systems (\mathbf{k} -Point Sampling) | 40 |
| 2.10 | Goedecker-Theter-Hutter-Pseudopotentials | 42 |
| 2.11 | Gaussian Plane Wave DFT | 43 |
| 2.11.1 | The Basic Idea of the Gaussian and Plane Waves Approach | 44 |
| 2.11.2 | GPW Preparation Scheme | 45 |
| 2.11.3 | Coulomb Interaction Treatment in CP2K | 47 |
| 2.11.4 | Exchange-Correlation Term | 50 |
| 2.12 | Non-Local Corrections to Density Functional Theory | 51 |
| 2.12.1 | Pairwise Dispersion Correction | 52 |
| 2.13 | Non-Local Exchange Correlation Functionals | 53 |
| 3 | <i>Polarization and Electric Fields in</i> | |
| | <i>Periodic Systems</i> | 55 |
| 3.1 | Interaction with External Electric Fields | 55 |
| 3.1.1 | The Polarization Problem | 57 |
| 3.2 | Modern Theory of Polarization | 58 |
| 3.2.1 | Berry-Phase-Approach and King-Smith-Vanderbilt Formula | 60 |
| 3.2.2 | The Position Operator in Extended Systems | 62 |
| 3.3 | Wannier-Functions | 66 |
| 3.4 | Polarization Quantum | 68 |
| 3.5 | Fields in DFT | 69 |
| 3.6 | Periodic Electric Fields | 71 |

| | | |
|-----------|---|------------|
| 3.7 | Field Effects Illustrated | 74 |
| 3.7.1 | Orbital Shifts | 75 |
| 3.7.2 | Charge Shifts | 76 |
| 3.7.3 | Bond Profiles | 76 |
| 4 | <i>Tight Binding Methods</i> | 79 |
| 4.1 | The Concepts of Tight Binding Models in the Solid State | 80 |
| 4.1.1 | Ideal one Atomic Crystal | 82 |
| 4.2 | Density Functional Tight Binding | 83 |
| 4.2.1 | DFTB-Energy Expression | 84 |
| 4.2.2 | Basic DFTB | 85 |
| 4.2.3 | DFTB2-Extension | 87 |
| 4.2.4 | DFTB3-Extension | 89 |
| 4.3 | GFN-xTB | 90 |
| II | <i>Analysis Tools</i> | 95 |
| 5 | <i>Fourier Random Displacement Generator</i> | 97 |
| 5.1 | The Initial Guess | 100 |
| 5.2 | The Difference Function | 101 |
| 5.2.1 | The Difference Function in 2D | 101 |
| 5.3 | The Optimization Scheme | 102 |
| 5.4 | Results and Applications | 104 |
| 6 | <i>Charge Symmetrizer</i> | 105 |
| 6.1 | Working Scheme | 105 |
| 6.2 | On the Description of Atomistic Environments | 106 |
| 6.2.1 | Angular and Radial Distribution Functions | 107 |
| 6.2.2 | Comparison of Atomistic Environments | 111 |
| 6.3 | Exemplary Results and Applications | 115 |

Contents

| | | |
|------------|---|------------|
| 6.4 | General Applicability | 119 |
| 6.4.1 | Multilayer Treatment | 119 |
| 7 | <i>Pore Size Determinator</i> | 123 |
| 7.1 | The Definition of the Pore Size | 123 |
| 7.2 | Workflow | 128 |
| 7.3 | Exemplary Results and Applications of the PSD-Analysis | 130 |
| III | <i>Applications</i> | 133 |
| 8 | <i>Investigation of Surface Models for Self-Assembled Monolayers on Metal Substrates</i> | 135 |
| 8.1 | Principles of Thiolate Self-Assembled Monolayers | 136 |
| 8.2 | Investigated Thioliates in the Gas Phase | 138 |
| 8.3 | Construction of Models for Metal Substrates | 140 |
| 8.3.1 | Exemplary Workflow for Construction of a Gold (111) Surface | 141 |
| 8.3.2 | Bulk Lattices using GFN-xTB | 144 |
| 8.3.3 | Applicability of the Theoretical Models | 145 |
| 8.4 | SAM-Molecules on Gold-Surfaces | 146 |
| 8.4.1 | Interaction Energies | 147 |
| 8.5 | Decanethiol | 149 |
| 8.5.1 | Considerations for the Smallest System | 149 |
| 8.5.2 | One C10-SAM Molecule | 152 |
| 8.5.3 | Explicit Intermolecular Effects (2 Molecules) | 154 |
| 8.5.4 | Surface Arrangement with three Molecules per Unit-Cell | 154 |
| 8.6 | Perfluoro-Decanethiol | 158 |
| 8.6.1 | Explicit Intermolecular Effects (2 Molecules) | 161 |
| 8.7 | Amino-Undecanethiol | 162 |
| 8.7.1 | Explicit Intermolecular Effects (2 Molecules) | 163 |
| 8.8 | Comparison of the investigated Species | 166 |
| 8.8.1 | Discussion of the Setup Preparation | 168 |

| | | |
|-----------|---|------------|
| 9 | <i>Field Evaporation of Self Assembled Monolayers</i> | 171 |
| 9.1 | Consequences of Applied Fields within PBCs | 172 |
| 9.1.1 | Bond Profiles in the Gas Phase | 172 |
| 9.1.2 | Bond Profiles within PBCs | 174 |
| 9.1.3 | Charge Shifts and Equilibrium Distance in Electric Fields | 177 |
| 9.1.4 | Polarization in Electric Fields | 178 |
| 9.2 | Basics of Atom Probe Tomography | 179 |
| 9.3 | Simulations of Field Evaporation using GPW-DFT | 181 |
| 9.3.1 | Technical Details for Performing Simulations of Evaporation Sequences using GPW-DFT | 182 |
| 9.3.2 | Simulation of Evaporation for a pure Gold Surface | 186 |
| 9.3.3 | Evaporation of Decanethiole from Gold | 187 |
| 9.3.4 | Evaporation of Amino-undecanethiol from Gold | 193 |
| 9.3.5 | Evaporation of Perfluoro-Decanethiol from Gold | 198 |
| 9.3.6 | Discussion of the Observed Effects using GPW-DFT | 204 |
| 9.4 | Simulations of Field Evaporation using GFN-xTB | 207 |
| 9.4.1 | Evaporation of Gold | 207 |
| 9.4.2 | Evaporation of Decanethiol from Gold | 208 |
| 9.4.3 | Evaporation of Amino-undecanethiol from Gold | 212 |
| 9.4.4 | Evaporation of Fluorinated Decanethiol from Gold | 215 |
| 9.5 | Discussion of the Results | 218 |
| 9.5.1 | Limitations of the utilized Models and Outlook | 220 |
| 10 | <i>Theoretical Analysis for Covalent Organic Frameworks</i> | 223 |
| 10.1 | What are COFs? | 224 |
| 10.2 | The Choice of a Suitable Theoretical Description | 225 |
| 10.3 | Electrostatic Interactions | 226 |
| 10.4 | Construction of the GPW-DFT-Setup | 229 |
| 10.4.1 | Functional Choice | 230 |
| 10.4.2 | Basis Set Choice | 233 |
| 10.4.3 | PW-Cutoff and Rel-Cutoff | 234 |

Contents

| | | |
|-----------|---|------------|
| 10.4.4 | The impact of Dispersion Corrections | 234 |
| 10.4.5 | Conclusion for the Standard Setup | 240 |
| 10.5 | Investigations using GFN-xTB | 240 |
| 10.5.1 | Interpolation of the Potential Energy Surface | 242 |
| 10.6 | Overview of the Investigated Species | 246 |
| 10.6.1 | Planar Species | 247 |
| 10.6.2 | Non-planar Species | 264 |
| 10.7 | Insights Gained for the description of COF-systems | 270 |
| 11 | <i>Theoretical Investigations of Interlayer Interactions for Large-Pore COFs</i> | 273 |
| 11.1 | Categorization of the PP-TAB COFs | 274 |
| 11.2 | Component Analysis by means of DFT | 277 |
| 11.2.1 | Phenylphenanthridine-Combinations | 279 |
| 11.2.2 | Imine-linker-Combinations | 281 |
| 11.2.3 | Block Analysis with GPW-DFT | 282 |
| 11.2.4 | Wall Analysis with GPW-DFT | 287 |
| 11.2.5 | Influences of the Methoxy Groups | 289 |
| 11.2.6 | Analysis of Methoxy Group Orientations | 291 |
| 11.3 | Stability Analysis of the PP-TAB-COF Series | 294 |
| 11.3.1 | Construction of the Eclipsed Structure | 295 |
| 11.3.2 | Pore Size Analysis using PSD | 298 |
| 11.4 | Flourinated-Linker Design | 303 |
| 11.4.1 | T-Shaped Analysis for the F-PP-COF System | 305 |
| 11.5 | Stability Analysis of the F-PP-TAB COF Series | 308 |
| 11.5.1 | Theoretical Investigation of Displacements for F-PP-TAB-COFs | 308 |
| 11.6 | Lessons Learned for the Design of Large-Pore COFs | 312 |
| 12 | <i>Summary</i> | 315 |
| 12.1 | Investigations of Surface Models for Self-Assembled Monolayers on Metal Substrates | 316 |
| 12.2 | Field Evaporation of Self-Assembled-Monolayers | 317 |

| | | |
|----------|---|------------|
| 12.3 | Theoretical Analysis for Covalent Organic Frameworks | 319 |
| 12.4 | Theoretical Investigations of Interlayer Interactions for Large-Pore COFs . . . | 321 |
| A | <i>Theory</i> | 325 |
| A.1 | Separation of the Time Dependence | 325 |
| A.2 | Born-Oppenheimer Approximation | 326 |
| A.3 | Variational Principle | 328 |
| A.4 | Theorems for Gaussian-Functions | 330 |
| A.5 | Hohenberg-Kohn Theorem I | 332 |
| A.6 | Parsevals Theorem | 333 |
| A.7 | Plancherets Theorem | 333 |
| A.8 | Mean Value Theorem for Lattice Periodic Functions | 334 |
| B | <i>Modern Theory of Polarization</i> | 335 |
| B.1 | Position Operator and Polarization | 335 |
| B.1.1 | Example Systems | 336 |
| B.2 | Derivation of the Berry Phase for Adiabatic Change of State | 337 |
| C | <i>Tight Binding Methods</i> | 341 |
| C.1 | Reformulation of the DFT Energy Expression for DFTB Approximation . . . | 341 |
| C.2 | Derivation of the DFTB-Energy Expression | 342 |
| C.3 | Compressed Atomic Basis Functions | 344 |
| C.4 | Cyclic Cluster Model | 345 |
| D | <i>SAM-Basics</i> | 347 |
| D.1 | SAM Densities for Au (111) | 347 |
| D.2 | SAM Densities for Cu and Pt (111) | 348 |
| D.3 | Molecular Length and Tilting Angle | 350 |
| D.4 | Interaction Energies for SAMs on Gold (111) Surfaces | 351 |
| D.5 | Comparison of Interaction Energies | 352 |
| E | <i>Field Evaporation</i> | 353 |
| E.1 | Fragmentation Patterns - GFN-xTB | 353 |

Contents

| | | |
|----------|---|------------|
| E.2 | Charge Redistribution for the H ₂ -Example | 355 |
| F | COF-Basics | 357 |
| F.1 | GPW-DFT Benchmarks | 357 |
| F.2 | PW-Fit Coordinates | 362 |
| G | Analysis of Large Pore COFs (PP-TAB series) | 363 |
| G.1 | T-shape analysis | 363 |
| G.2 | Band Analysis Data (GPW-DFT) | 364 |
| G.3 | Data for the full COF-analysis | 366 |
| G.4 | Data Points for the Stability Analysis | 366 |
| | List of Figures | 371 |
| | List of Tables | 389 |
| | Bibliography | 397 |
| | Declaration of Authorship | 415 |

List of Abbreviations

| | |
|----------------|---|
| AO | Atomic orbitals |
| APT | Atom probe tomography |
| BLYP | Becke Lee Yung Parr |
| BvK | Born van Karman |
| BZ | Brillouin zone |
| C10-SAM | Decanethiolate |
| CCM | Cyclic cluster model |
| CI | Configuration interaction |
| CM | Clausius Mosotti |
| COF | Covalent organic framework |
| CRC | Collaborative research center |
| CS | Charge symmetrizer |
| DFT | Density functional theory |
| DFTB | Density functional tight binding |
| F-SAM | Per-fluoro decanethiolate |
| FPW | Field-polarized-Wannier |
| FRDG | Fourier random displacement generator |
| GFN-xTB | Geometries Frequencies non-covalent Interactions extended tight binding |
| GGA | Generalized gradient approximation |
| GPW | Gaussian plane waves |
| GTH | Goedecker-Theter-Hutter |

List of Abbreviations

| | |
|---------------------------|---------------------------------------|
| HF | Hartree–Fock |
| LCAO | Linear combination of atomic orbitals |
| LDA | Local density approximation |
| LR | Localization region |
| LSD | Local spin density |
| LYP | Lee Yung Parr |
| MEP | Molecular electrostatic potential |
| MGC | Mauri Galli Car |
| MTP | Moder theory of polarization |
| NH₂-SAM | Amino-undecanethiolate |
| NL | Non-local |
| OBC | Open boundary condition |
| PBC | Periodic boundary conditions |
| PBE | Perdew Burke Ernzerhof |
| PP | Pseudo potential |
| PSD | Pore size determinant |
| PW | Plane wave |
| QM | Quantum mechanical |
| QS | Quick step |
| SAM | Self-assambled monolayer |
| SCC | Self-consistent charge |
| SCF | Self-consistent field |
| SD | Slater determinant |

| | |
|-------------|--------------------------------|
| TB | Tight binding |
| TOF | Time of flight |
| TPSS | Tao–Perdew–Staroverov–Scuseria |
| XC | Exchange and correlation |

Chapter 1

Introduction

Theoretical models, capable of describing real-world applications accurately, allow for an in-depth analysis of effects which are not experimentally observable, and the prediction of properties' changes due to system modifications. Within the framework of the collaborative research center (CRC) 1333 “Molecular Heterogeneous Catalysis in Confined Geometries”, porous structures providing confined environments, such as covalent organic frameworks (COFs) are designed specifically for applications in catalysis. Furthermore, the structural analysis of catalytic hybrid materials by extending the applicability of atom probe tomography (APT) to organic materials marks the next step in its development. Both are emerging fields of research with the possibilities of theoretical investigation being, to present day, largely unexploited.

1.1 Simulation of Field Evaporation for Atom Probe

Tomography

Atom probe tomography (APT) is a unique destructive analysis technique that allows chemical material characterization on near-atomic scale in three dimensions, by consecutive removal of surface atoms. Substrates, placed on needle shaped metal tips with curvature radii < 50 nm, are subjected to intense electric fields (around 10 V/nm), resulting in the ionization of the surface atoms¹. Subsequent, ions are repelled from the surface and accelerated in field direction

1 Introduction

towards a position sensitive detector². The ionized species are identified by time of flight (TOF) mass spectrometry within a standing field, where voltage or laser pulses are used to trigger single evaporation events. Triggering of the laser pulse marks the starting point for TOF measurements, terminated upon arrival at the detector. Only the topmost atoms are desorbed stepwise allowing for three-dimensional reconstruction of the investigated sample³⁻⁵. Consequently, APT provides enough chemical sensitivity for spatial reconstruction on the subnanometer scale and is commonly used for metals, insulators and semiconductors^{6,7}. The generalization of APT for analysis of organics and soft matter requires, besides further development of the experimental methodology, the necessity of additional insights' into the underlying desorption mechanisms⁸⁻¹⁰. Recent advances in cryo-freezing techniques for liquids opened up a new field of analysis for APT and have been exemplary conducted for frozen water as well as glucose samples^{11,12}. Here, the interpretation of peaks in the mass-spectra was supported by theoretical analysis of the stability of ionic-species with corresponding mass to charge ratio. Calibration of the APT methodology for possible application to organic systems requires the characterization of prototypical examples. Well-ordered organic systems are provided by self-assembled monolayers (SAMs), which consist of chainlike molecules and organize themselves in organic thin films. By tuning the head-group they bind to a range of different substrates with their layer-thickness depending on molecular length and tilt angle, while surface functionality can be adjusted by tuning the end group^{13,14}. During APT measurements, the electric field penetrates the SAM layer and breaks the chainlike molecules into various hydrocarbon fragments leading to mass spectra displaying combinations of ionized $C_xH_y^{n+}$ -fragments instead of single-atom ions¹⁵⁻¹⁷. The identification of the obtained species is ambiguous in comparison to metal systems and requires fundamental understanding of field evaporation mechanisms of organic species.

Previous theoretical works targeting the process of field evaporation for SAMs on metal surfaces utilized single atoms and small clusters to model the metal substrate^{18,19}. Contrarily, this work provides direct analysis of the evaporation process for surface models constructed for SAM-metal interfaces as schematically illustrated in Figure 1.1. During the interpretation of the obtained fragmentation patterns, emphasis is put on the crucial differences stemming from the solid state description. In parallel, the nature of the modeled evaporation process is

investigated by targeting the evolution of the system's polarization due to redistribution of charges and the corresponding structural reorganization.

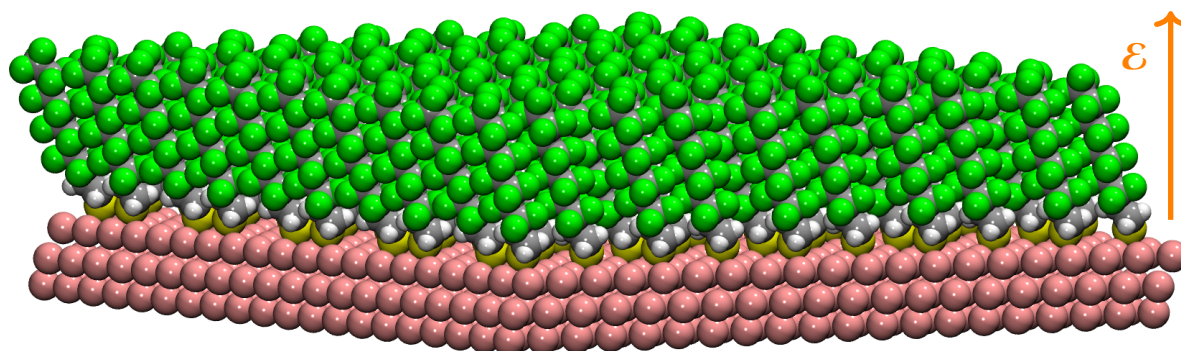


Figure 1.1: Schematic representation of an extended surface model as utilized for the simulation of field evaporation of SAMs. The direction of the applied electric field ϵ is indicated in orange.

1.2 Theoretical Description of Covalent Organic Frameworks

Covalent organic frameworks (COFs) are a class of highly stable and crystalline polymers forming porous networks as illustrated in Figure 1.2, assembled from covalently bound organic components consisting of light elements (e.g. C, N, H, O, F)²⁰. The ease of manipulation of COFs due to the variability in the design of specific components allows the creation of frameworks with a wide range of structural and chemical diversity^{21,22}. Consequently, COFs have been attracting great interest for various applications such as catalysis^{23–28}, gas separation and storage^{29–32}, water harvesting^{33,34}, energy storage^{35,36}, chemical sensing³⁷, optoelectronic materials³⁸ and medical purposes^{39,40}. To design and optimize COFs for applications, materials with high crystallinity and accessible surface areas, as well as chemical and thermal stability, are targeted⁴¹. The design of COFs for a specific task is fine tuned by changing linkers or adding functional groups^{42–45}. The vast amount of possible materials and framework designs, requires their categorization as established in COF-data bases, providing crystal structures and framework properties^{46–48}. Besides “chemical intuition”, the design of COFs can be further rationalized by analysis of suitable theoretical models elucidating the expected properties of a hypothetical species. Due to the large unit-cell size of the underlying crystal structures, theoretical models on the atomic length scale require efficient treatment of the quantum

1 Introduction

mechanical problem. Identification of potentially well-defined structures with accessible docking sites and large pore sizes become increasingly important especially when considering their application for catalysis.

While micro- and small mesoporous COFs are widespread today, large-pore COFs with pore diameters exceeding 5 nm and high crystallinity are still rare^{49–54}. The class of two-dimensional COFs, forms pores by direct stacking of layers organized by non-covalent interactions, whereas three-dimensional species are connected by covalent bonds also in stacking direction. Already conducted approaches often suffer from pore blocking and collapse caused by guest molecules as a result of slipped stacking schemes and low crystallinity^{55–57}. For two-dimensional COFs tuning interlayer interactions⁵⁸ was already employed and is a promising approach for systematically reducing the lateral slipping as observed for stacked two-dimensional layers⁵⁹ by favoring eclipsed stacking. Consequently, the obtained pore size and surface area are increased while avoidance of pore collapse by rigid linkers keeps the porous structure intact⁶⁰.

To this end, an in-depth understanding of the factors affecting pore structure and stability is necessary to avoid the loss of crystallinity and porosity. For two-dimensional COF-systems as utilized within the CRC-1333 the target properties, namely pore size and stacking order, are determined by non-covalent interlayer interactions.

This thesis provides computational tools as well as an analysis of the theoretical framework available for treating interlayer interactions and their impact on the preference of specific stacking schemes. Structural models are investigated by means of periodic density functional theory (DFT) and the semi empirical geometries frequencies non-covalent interactions extended tight-binding (GFN-xTB)⁶¹ method. The efficiency of periodic treatment of porous systems by standard plane wave codes suffers from large vacuum spaces, requiring large numbers of basis functions to fill the whole unit cell. This is accounted for by utilization of the Gaussian and plane waves (GPW) approach, as implemented in the quick step module within the CP2K program package⁶², for simulation of structural models for COFs. The results obtained from the study of planar and non-planar COFs provide insights into possibilities for rationalizing the design of large-pore COFs based on theoretically predicted properties.

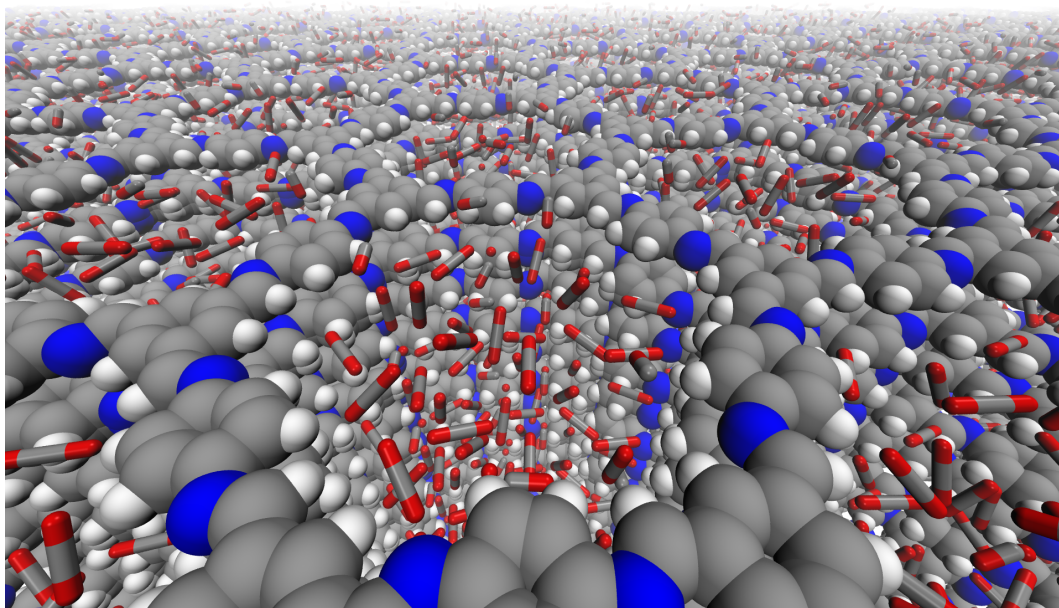


Figure 1.2: Exemplary illustration of an eclipsed COF-LZU1-pore network filled with CO₂.

1.3 Structure of this Thesis

This thesis consists of three distinct parts, Part I provides a detailed review on the utilized theoretical methods while Part II explains the computational tools developed for the analysis of COFs. Part III summarizes the main results of the conducted studies for simulations of APT and the theoretical analysis of COFs.

- Part I - Theory:
 - Chapter 2 reviews the fundamentals of the DFT formalism and its generalization to periodic systems. Due to its uniqueness special emphasis is put on the explanation of the gaussian and plane waves (GPW) approach as implemented in the CP2K program package.
 - Chapter 3 is dedicated to an in-depth description of polarization as a bulk property following the modern theory of polarization utilized for the analysis of field evaporation. Furthermore, it explains basic effects of periodic field implementations within the DFT framework.
 - Chapter 4 explains the basic ideas of the hierarchically improved density functional tight binding (DFTB) models and their parametrization. The chapter aims at

1 Introduction

illustrating the commonly stated connection of GFN-xTB to the DFTB3-family of methods by a detailed review of its derivation.

- Part II - Analysis Tools:
 - Chapter 5 explains the working scheme of the Fourier random displacement generator (FRDG) utilized for the creation of periodic random displacement paths.
 - Chapter 6 reviews the mathematical background for the description of atomistic environments by means of environmental descriptors utilized within the developed charge symmetrizer (CS). The chapter is kept as general as possible to highlight the transferability of the algorithm to other properties than atomic partial charges.
 - Chapter 7 provides the basic considerations necessary for geometric pore analysis as utilized in the developed pore size determinant (PSD). Based on the definition of the geometric pore size \mathfrak{P} , the general concepts used for its determination are explained in detail with emphasis put on the application of the PSD with FRDG-generated random paths.

- Part III - Applications:
 - Chapter 8 provides an overview of the investigated SAM-species, the creation of the utilized surface models as well as the specifics obtained from their analysis. Understanding of the fundamentals of the utilized models facilitates later interpretation and highlights the strengths of the used surface model.
 - Chapter 9 highlights the differences arising from gas-phase and solid state treatment of external electric fields within DFT at the example of H₂. The effects observed in the initial study reveal the descriptive features of the periodic treatment, the knowledge of which, simplifies further interpretation. The studies conducted revealed difficulties which were overcome by strategies outlined for follow up investigators. Subsequently, the main results of the simulated fragmentation processes for the three SAM species using GPW-DFT and GFN-xTB are reported. A detailed analysis of the evaporation process by means of GPW-DFT is conducted for the smallest reference systems illustrating the nature of the evaporation process.

Comparison of both methods for their applicability, limitations as well as the underlying reasons close up the studies.

- Chapter 10 treats the basics for the theoretical analysis of COFs with emphasis on interlayer interactions arising due to aromatic components and the creation of a suitable GPW-DFT setup. General structural features of COFs and related effects on the obtained pore size are exemplified at literature known COFs namely COF-LZU1, N3-COF and TpPa-COF. The significance of the analyzed interlayer interactions on structural arrangements is tested for a hypothetical framework. In analogy to the previous study, the impact of methoxylation is discussed for the example of the PP-DAAQ-COF system. Finally, based on the insights gained, a workflow for characterization of COFs by construction of suitable structural models is obtained and presented.
- Chapter 11 applies the methodology of the previous investigations for an in-depth study of the isorecticular imine based PP-TAB-COF series. The experimentally observed stabilization due to methoxy groups attached to the COFs imine linkers is explained by a hierarchical analysis of components. Further stability analysis of the obtained structural model is conducted leading to identification of methoxy groups as molecular “anchors”, stabilizing the eclipsed configuration. Subsequent studies on layer buckling and interlayer shifts are conducted to find possible connections between both phenomena. With the applicability of the method shown, it is further used for analysis of T-shaped stacking for fluorinated phenyl linkers, predicting it as a second type of suitable interlayer interactions. In summary, the insights gained are utilized to suggest the main features necessary for rationalization of COF design by theoretical predictions. Previous studies on modifications of two of the three main components suggest modification of the third as the next step, here aryl interactions as observed for the N3-COF offer a promising third type of modification.

Part I

Theory

Chapter 2

Density Functional Theory

In the following chapter the basic concepts of density functional theory (DFT) are presented based on the fundamental problem of quantum chemistry. Starting from the basic approach of solving the Schrödinger equation by explaining the corresponding Hamiltonian and its components, the most crucial steps for simplification of the many-body problem namely the separation of the time dependence and the Born-Oppenheimer (BO) approximation are explained, splitting off the kinetic energy of the nuclei from the general Hamiltonian. The resulting description allows the determination of the electronic ground state for a given molecular geometry which is subsequently used to solve the nucleic problem. These basic considerations are firstly explained at the example of the fundamental Hartree-Fock (HF) approach, illustrating the connection to ground state DFT as utilized within this work. Starting from the Hohenberg-Kohn theorems, the Kohn-Sham equations and the fundamental aspects mark the beginning of this discussion.

As the crucial aspect for possible descriptions utilizing DFT, the exchange-correlation problem is described along the functionals mostly used within this thesis.

Followed by the algebraic formulation of the fundamental problem, the linear combination of atomic orbitals (LCAO) approach is described with consecutive introduction of Slater and Gaussian-type basis functions for explanation of the self-consistent field (SCF) concept. The description of the so far molecular treatment, is further generalized to the solid state by introduction of periodic-systems. The necessary adaptation of the previous theoretical framework is shown by introducing the concepts of real and reciprocal grids. The resulting

2 Density Functional Theory

consequences for calculation of bulk properties are explained followed by the necessary generalization of basis sets to the periodic case. Finally, the benefits of a plane-wave (PW) basis over Gaussians are discussed for periodic treatments.

Based on the explained framework, the idea of pseudopotentials (PPs) is illustrated at the example of the specialized Goedecker-Theter-Hutter (GTH) type, thereby introducing the last concept necessary for understanding the Gaussian and plane waves (GPW) approach.

Within this thesis, the CP2K-program package⁶² was used for all calculations of periodic systems. Compared to standard Gaussian or full PW-treatment, the GPW-workflow is rather special, therefore, the main features of its implementation in the quickstep (QS) module within CP2K, are reviewed. A special emphasis is taken on the treatment of the electrostatic contributions as it inherits the main advantages by reducing the computational effort. The chapters review on the DFT basics is completed by introducing the concepts of pair-wise dispersion corrections and non-local exchange correlation functionals.

2.1 The Challenge of Quantum Chemistry

In quantum chemistry, the basic problem consists in finding a solution to the molecular Schrödinger equation, which is fundamentally given as

$$i\hbar \frac{\partial}{\partial t} \Psi(t, \mathbf{r}) = \hat{\mathbf{H}} \Psi(t, \mathbf{r}) \quad . \quad (2.1)$$

Here, i is the imaginary unit, \hbar Planck's constant, $\hat{\mathbf{H}}$ the described system's Hamilton operator and $\Psi(t, \mathbf{r})$ the system's wave function depending on time t and space \mathbf{r} respectively. For most systems, only the energies of stationary states are of interest. In the absence of external perturbations, the Hamiltonian is independent of time allowing for the separation of the time-dependent part of the Schrödinger equation by a product ansatz as shown in Section A.1.

With the time dependence separated, the remaining time independent Schrödinger equation marks the starting point for further investigations. The differences in the description of the system at hand, therefore, rely on the form of $\hat{\mathbf{H}}$ and the method for solving Eq (A.3b). For a molecule consisting of N electrons and M nuclei, the system's many-body Hamiltonian in

2.1 The Challenge of Quantum Chemistry

atomic units ($e = \hbar = \frac{1}{4\pi\epsilon_0} = m_e = 1$) in the absence of external fields, is given as:

$$\begin{aligned}
 \hat{\mathbf{H}} &= \hat{\mathbf{T}}_e + \hat{\mathbf{T}}_n + \hat{\mathbf{V}}_{ee} + \hat{\mathbf{V}}_{nn} + \hat{\mathbf{V}}_{en} \\
 &= \frac{1}{2} \sum_i^N \hat{\mathbf{p}}_i^2 + \frac{1}{2} \sum_A^M \frac{\hat{\mathbf{P}}_A^2}{m_A} + \sum_i^N \sum_{j \neq i}^N \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} \\
 &\quad + \frac{1}{2} \sum_A^M \sum_{B \neq A}^M \frac{Z_A Z_B}{|\mathbf{R}_A - \mathbf{R}_B|} - \frac{1}{2} \sum_i^N \sum_A^M \frac{Z_A}{|\mathbf{r}_i - \mathbf{R}_A|}
 \end{aligned} \tag{2.2}$$

Here, $\hat{\mathbf{p}}_i$ is the momentum operator and \mathbf{r}_i the position of electron i , $\hat{\mathbf{P}}_A$ the momentum, \mathbf{R}_A the position of the nuclei, m_A the mass, and Z_A the core charge of atom A . The general many-body Hamiltonian can be separated into two main parts, one depending on the electrons and one on the nuclei. In Eq (2.2), the movement of nuclei and electrons is described simultaneously. The problem is significantly simplified by splitting off the kinetic energy of the cores, using the Born-Oppenheimer approximation. Here, the separation of electronic and nucleic movement is justified by consideration of the ratio of masses between both particles $\frac{M_n}{m_e} \approx 2000$ leading to the assumption of quasi independent movement of electrons. To be more vivid, the system of electrons adjusts itself on a timescale three orders of magnitude faster than the nucleic system. Therefore, the positions of the nuclei are acting as quasi static environmental variables for the electronic movements. Considering changes of the electron system, this can be interpreted in analogy to an adiabatic change of state, as known from thermodynamics. For this reason, the BO-approximation is also referred to as adiabatic approximation. The description is considered accurate for effects slower than the mentioned timescale and the approximation is applicable without significant loss of information. Splitting off the core movement, leads to two separate differential equations for electrons and nuclei respectively. The nucleic coordinates act only as an environmental parameter for the electrons, resulting in the electronic Schrödinger equation, given as

$$\hat{\mathbf{H}}_e \psi(\mathbf{r}, \mathbf{R}) = E_e(\mathbf{R}) \psi(\mathbf{r}, \mathbf{R}) \quad , \tag{2.3}$$

where $\hat{\mathbf{H}}_e = \hat{\mathbf{T}}_e + \hat{\mathbf{V}}_{ee} + \hat{\mathbf{V}}_{ne} + \hat{\mathbf{V}}_{nn}$ marks the electronic Hamilton operator. The Schrödinger

2 Density Functional Theory

equation describing the cores contains the remaining term of Eq (2.2) resulting in

$$(\hat{\mathbf{T}}_n + E_e(\mathbf{R}))\eta(\mathbf{R}) = E_{\text{tot}}\eta(\mathbf{R}) \quad , \quad (2.4)$$

where η is the core wave-function. The occurrence of E_e in Eq (2.4) shows the interdependence of the total energy eigenvalue E_{tot} from the electronic one. Consequently, the solution of the nucleic problem is calculated after solving the electronic Schrödinger equation Eq (2.3) for E_e . The Hamiltonian yields a description of the system's energy levels, evaluated for a given wave function. Principally the coupling of nuclei and electrons is not fully resolved in the new equations and the decoupled equations resulting from the BO-approximation, mark a simplification that is explained in more detail in Section A.2. The remaining task consists in finding a suitable way for evaluation of expectation values for the system at hand. In practical applications, this corresponds to the determination of a suitable wave function. In the usual case, the focus of interest lies within the electronic energies. The main task consists therefore in finding an electronic wave function first and a nucleic one afterwards. The core problem will not be treated in the following, as the main task of this work consists in evaluation of the electronic structure.

Electronic movements are in general highly correlated, hence, numerical methods have to be used in order to solve Eq (2.3). Various approaches are available for solving the electronic Schrödinger equation, which can be generally divided in two groups, namely wave function and density functional based methods.

Wave-function based methods directly calculate the many-body wave function by solving a system of N coupled differential equations with N variables. Consequently, wave-function based methods scale heavily with the number of electrons N .

An alternative way is offered by calculating the system's electron density ρ . Using this method allows a reformulation of the problem in terms of a system of N independent equations, significantly reducing the computational effort. The most fundamental approach for wave-function based methods is based on Hartree-Fock (HF) theory, which is described in the following.

2.2 The General Idea of Hartree-Fock Methods

To solve Eq (2.3), inter-particle interactions have to be considered in a suitable manner. A simple approach developed by Hartree and Fock⁶³ treats the electronic interactions not as pairwise, but rather sums up all interactions acting on one specific electron, resulting in a so called mean field. The mean field acting on each of the system's electrons is calculated for an initial state, which is further used to solve Eq (2.3) yielding a new field and electronic state. This process is repeated until a self-reproducing solution is obtained (meaning there is no further change achieved by the process itself). The obtained solution is considered self consistent, giving this method its name 'self-consistent field' (SCF). The underlying methodology is widely used in quantum chemistry for analogous problems. In the following the general approach to solve the HF problem is described as it marks the foundation for most subsequent methods. In the following, the well known Dirac notation is used, for better readability. Solving Eq (2.3) requires the determination of an expression for the N -electron wave function $|\Psi_0\rangle$ corresponding to the ground state energy E_0 . An approximate wave function can be found by a variational ansatz, which aims at minimizing the energy of a trial wave function $|\Psi_{\text{trial}}\rangle$ (note that in the following the "index" trial is suppressed). With this ansatz the overall minimization problem is given as

$$E_0 = \min E[\Psi] = \min \langle \Psi | \hat{T} + \hat{V}_{\text{ne}} + \hat{V}_{\text{ee}} | \Psi \rangle \quad . \quad (2.5)$$

While the approximate ground-state wave function is constructed by an optimization procedure, an initial guess of the desired wave function is required as starting point. In usual HF-theory, $|\Psi\rangle$ is constructed under the constraints of the Pauli principle by using a Slater determinant Φ_{SD} of N one electron wave functions χ_i . Each electron state consists of a superposition of spacial and spin orbitals, resulting in the expression

$$\chi_i(\mathbf{x}) = \phi_i(\mathbf{r})\sigma(s) \quad , \quad (2.6)$$

for the spin orbitals. Here, \mathbf{x} considers spacial \mathbf{r} and spin s coordinates respectively.

2 Density Functional Theory

A Slater determinant is constructed from N one electron states is then given as

$$\Phi_{\text{SD}} = \frac{1}{\sqrt{N!}} \begin{vmatrix} \chi_1(\mathbf{x}_1) & \chi_2(\mathbf{x}_1) & \cdots & \chi_N(\mathbf{x}_1) \\ \chi_1(\mathbf{x}_2) & \chi_2(\mathbf{x}_2) & \cdots & \chi_N(\mathbf{x}_2) \\ \vdots & \vdots & \ddots & \vdots \\ \chi_1(\mathbf{x}_N) & \chi_2(\mathbf{x}_N) & \cdots & \chi_N(\mathbf{x}_N) \end{vmatrix} . \quad (2.7)$$

In a more general case, the wave function is constructed as a linear combination of configuration functions which describe different possible configurations of spin states (which in turn are constructed from linear combination of multiple determinants), the resulting methodology is referred to as configuration interaction (CI). However, the system is described only with maximum accuracy when all possible configurations (excitations of states) are considered (full-CI), leaving the problem computationally highly demanding even for the smallest systems.

In HFtheory, particles are assumed to be interacting indirectly by being placed in the external field provided by V_{HF} describing the mean-interaction with all other particles. The fundamental equations are then provided as

$$\hat{\mathbf{H}}_{\text{HF}} |\Psi\rangle = \sum_i^N \hat{\mathbf{f}}_i \Phi_{\text{SD},i} = \sum_i^N \epsilon_i \Phi_{\text{SD},i} = E_{\text{HF}}^0 |\Psi\rangle . \quad (2.8)$$

The functions χ_i are referred to as spin orbitals, since they are composed of a spatial $\phi_i(\mathbf{r})$ and spin component $\sigma(s)$ respectively. The variation of $|\Psi\rangle$ is thus executed over the N spin orbitals χ_i , under the orthonormality constraint, to minimize the resulting energy E_{HF} . Reformulation of the problem, using the approximate description of one electron in the field provided by the $N - 1$ others, results in N decoupled HF equations of form

$$\hat{\mathbf{f}}_i |\chi_i\rangle = \epsilon_i |\chi_i\rangle . \quad (2.9)$$

Here, ϵ_i is the energy of orbital i and the effective one electron Fock-operators $\hat{\mathbf{f}}_i$ are given as

$$\hat{\mathbf{f}}_i = -\frac{1}{2} \nabla_i^2 - \sum_A^M \frac{Z_A}{r_{iA}} + V_{\text{HF}}^i . \quad (2.10)$$

It is necessary to note, that the interaction of electron i with all other electrons is combined

into the external HF potential V_{HF}^i , which results from the collective repulsion of the remaining $N - 1$ electrons. The Hartree-Fock potential consists of the sum of two pairwise electronic contributions for all N electrons, and can be written for electron i as

$$V_{\text{HF}}^i(\mathbf{x}_i) = \sum_j^N \left(\hat{\mathbf{J}}_j(\mathbf{x}_i) - \hat{\mathbf{K}}_j(\mathbf{x}_i) \right) \quad . \quad (2.11)$$

With $\hat{\mathbf{J}}_j$ in Eq (2.11) being called the Coulomb-operator, describing the Coulomb interaction potential acting on electron i , arising from the averaged charge distribution of electron j in orbital χ_j . For a pair of electrons denoted as 1, 2, with $r_{12} = |\mathbf{x}_2 - \mathbf{x}_1|$ it can be written as

$$\hat{\mathbf{J}}_j = \int |\chi_j(\mathbf{x}_2)|^2 \frac{1}{r_{12}} d\mathbf{x}_2 \quad . \quad (2.12)$$

The second term in Eq (2.11), $\hat{\mathbf{K}}_j$ is called the exchange operator. The operator $\hat{\mathbf{K}}_j$ describes the energetic contribution arising from a hypothetical exchange of spin orbitals, explicitly this leads to the formulation

$$\hat{\mathbf{K}}_j \chi_i(\mathbf{x}_1) = \int \chi_j(\mathbf{x}_2)^* \frac{1}{r_{12}} \chi_j(\mathbf{x}_2) d\mathbf{x}_2 \chi_j(\mathbf{x}_1) \quad . \quad (2.13)$$

When calculating the expectation values for Eq (2.11), the physically wrong treatment of self interaction for $j = i$ arises. However, in HF theory this is irrelevant since for identical orbitals, $\hat{\mathbf{J}}$ and $\hat{\mathbf{K}}$ exactly cancel each other, eliminating this side effect.

HF-Exchange-Contribution

In the framework of HF theory, the exchange contribution to the total energy is calculated by integrating Eq (2.13) and given as

$$E_x^{\text{HF}} = \sum_{i,j} \int \int \frac{\chi_i(\mathbf{r}_1) \chi_i(\mathbf{r}_2) \chi_j(\mathbf{r}_2) \chi_j(\mathbf{r}_1)}{|\mathbf{r}_{12}|} d\mathbf{r}_1 d\mathbf{r}_2 \quad . \quad (2.14)$$

Generally it is not highly intuitive to interpret the physical meaning of Eq (2.14) in a conceptual sense. The exchange contribution arising between two electrons, is based on their indistinguishable nature, and acts in addition to the classical interaction of separated charge

2 Density Functional Theory

distributions given in Eq (2.12). Conceptually, it is not known which particle is located at a specific position resulting in an indeterminacy of the charge distributions associated with possible switches of the electrons' localizations. Naively speaking it is not known which part of the wave function is associated with a specific electron and thus the additional interactions arise if the wave functions of the individual electrons overlap, otherwise the overlap integrals in Eq (2.14) vanish.

2.3 Density Functional Theory

The main idea of density functional theory is to find a way of solving Eq (2.2) without explicitly calculating the system's wave function. The descriptive difference to HF-methods is founded in its crucial variable, which in the case of DFT is given by the electronic density ρ .

2.3.1 Properties of the Density

The electronic density of a quantum system with N particles and wave functions $\Psi_{i,s}(\mathbf{r}) = \phi_i(\mathbf{r})\chi_s$, is defined as

$$\rho(\mathbf{r}) = \sum_{i=1}^N \sum_s |\Psi_{i,s}|^2 \quad . \quad (2.15)$$

The sum over i runs over all orbitals with s accounting for spin up and down states respectively. It follows from its definition, that integration of the correct density over the whole space is equal to the total number of electrons N . This can be summarized as:

$$N = \int \rho(\mathbf{r}) d\mathbf{r} \quad (2.16)$$

For the suitable description of molecular quantum chemical systems, the density has to secure the locality of the electronic system yielding the condition $\rho(\mathbf{r} \rightarrow \infty) = 0$. In other words, the chance of finding a particle outside of the scope is zero and the total number of particles must therefore be found inside the described spacial vicinity. The electronic density, of a molecular system, yields a maximum probability of finite value at the position of nuclei due to

the attractive force between nuclear charge and electrons. The behavior is thus related to the nuclear charge as it stems from the divergence of the term $\frac{Z_A}{|r_{i,A}|}$ in the Hamiltonian. This term diverges for $|r_{i,A}| \rightarrow 0$, leading to a cusp of the electronic density and therefore a discontinuity in its gradient. Therefore, the core charge Z_A can be reconstructed according to the cusp's height.

2.3.2 Hohenberg-Kohn Theorems

The main advantage of DFT becomes visible when reformulating the quantum mechanical (QM) many-body problem in terms of the electronic density. The substantiation for this reformulation is provided by the fundamental theorems of Hohenberg and Kohn⁶⁴, providing the essential proofs for the validity of the DFT description. They can be summarized as:

- **Theorem I**

The system's ground state is described by a single density $\rho(\mathbf{r})$, determining the system's Hamiltonian uniquely and, therefore, all of its properties. This is the result of the density uniquely fixing the external potential to within a constant, consequently the many-particle ground state is a unique functional of $\rho(\mathbf{r})$. A short proof is provided in Section A.5.

- **Theorem II**

The ground state energy of a system, can be determined by the application of the variational theorem. Any density constructed by a non complete basis set expansion, results in an energy bigger than the one of the real ground state as shown in Section A.3.

The two theorems provide the necessary conditions to describe the system's Hamiltonian completely via the electron density by re-expression of Eq (2.2) in terms of functionals of ρ . Simultaneously, a framework for an optimization problem as further elaborated in Section 2.6 is provided. A functional provides a value depending on a utilized function, whereas a function itself provides a value depending on a variable. As an example, a closer look at the electron-nuclei interaction potential shows how a functional behaves while giving insight into the reformulation of the many-body problem. The electron-nuclei interaction in terms of the

2 Density Functional Theory

density is given as

$$V_{\text{ne}}[\rho] = \sum_K \int \rho(\mathbf{r}) \frac{Z_K}{|\mathbf{r} - \mathbf{r}_K|} d\mathbf{r}. \quad (2.17)$$

Here, square brackets indicate that a functional is used, K corresponds to the K^{th} nuclei, Z_K is its corresponding charge and \mathbf{r}_K its position. The value of V_{ne} depends on the used function for the density ρ .

2.4 Kohn-Sham Density Functional Theory

The Hohenberg-Kohn theorems provide the mathematical foundation for the possible description of the system solely in terms of its density, requiring reformulation of its Hamiltonian. The nuclear-electronic interaction can be expressed as Eq (2.17), which leaves the kinetic energy of the electrons T_e and V_{ee} to be re-expressed in terms of $\rho(\mathbf{r})$. These terms are independent of the external potential provided by the nuclei and can, for simplification, be merged into one universal functional, which is usually referred to as $F[\rho] = T[\rho] + V_{\text{ee}}[\rho]$. The remaining problem of minimization can be summarized as

$$E_0 = \min(F[\rho] + \int V_{\text{ne}}\rho(\mathbf{r})d\mathbf{r}) \quad . \quad (2.18)$$

The solution of Eq (2.18) for the ground state can be found by applying the variational theorem. The last step in the optimization workflow, consists now in finding a suitable expression for T_e and V_{ee} , as their exact form is unknown. Multiple approaches are discussed in detail in Section 2.5.

Multiple early approaches have been developed to directly solve the minimization problem of Eq (2.18) by means of so called orbital-free DFT^{65,66}. Even though these were successful for specific problems closely related to their initial motivations, they generally proved unsuited for molecular systems due to their inherent inability to describe chemical bonds. For the most part in modern day problems, approaches following the initial work of Kohn and Sham (KS) are applied, using an orbital based ansatz for construction of ρ ⁶⁷. The main idea exploited in the KS ansatz, is the description of the interacting many-body system by electrons moving

in the effective potential of a non-interacting reference system. Electrons are described by decomposing the system's density into one particle wave functions Φ_i analogous to HF. The Φ_i are then used to construct the systems total wave function Ψ , by combining them in a single Slater determinant (SD), given as

$$\Psi = \frac{1}{\sqrt{N}} \det(\Phi_1 \Phi_2 \dots \Phi_N) \quad . \quad (2.19)$$

The so called 'orbitals' Φ_i used for construction of the density, have no direct physical meaning as they are not eigenfunctions of the full interacting many-body Hamiltonian. With the wave function Φ constructed as in Eq (2.19), the density is given as

$$\rho = \sum_{i=1}^N |\Phi_i|^2 \quad . \quad (2.20)$$

In the orbital picture, $T[\rho]$ can be expressed by the newly constructed wave function of the non-interacting system. It is then given as

$$T_s[\rho] = -\frac{1}{2} \sum_{i=1}^N f_i \langle \Phi_i | \Delta | \Phi_i \rangle \quad , \quad (2.21)$$

where f_i denotes the occupation number of the respective orbital Φ_i . However, this expression does not directly correspond to the kinetic energy of the full interacting system $T[\rho]$, resulting in an unknown deviation between the two expressions. Furthermore, the Coulomb interaction of the electrons contained in V_{ee} is split into its classical analogue $J[\rho]$ and an unknown fully quantum mechanical part $V_{qm}[\rho]$. For the classical Coulomb interaction, the so called self-interaction term of a spread charge density is considered and given as

$$J[\rho] = -\frac{1}{2} \int \int \frac{\rho(\mathbf{r}_1)\rho(\mathbf{r}_2)}{|\mathbf{r}_1 - \mathbf{r}_2|} d\mathbf{r}_1 d\mathbf{r}_2 \quad , \quad (2.22)$$

where the indices 1,2 denote two different parts of the density respectively. Consequently, the integration can be seen as corresponding to interactions of a finite part of ρ with the rest of the charge density, summed over all parts of ρ . The unknown parts of the energy contributions can then be combined into one newly introduced additional functional. The resulting exchange-

2 Density Functional Theory

correlation functional $E_{xc}[\rho]$ describes the neglected terms within the reformulation procedure. These can be separated in two main contributions, firstly the error made by describing the kinetic energy within the non-interacting reference system and secondly, the error made by using the classical analogue of the Coulomb interaction denoted as V_{qm} . The total expression for E_{xc} is then given as

$$E_{xc}[\rho] = T[\rho] - T_s[\rho] + \underbrace{V_{ee}[\rho] - J[\rho]}_{V_{qm}[\rho]} \quad . \quad (2.23)$$

Following Kohn and Sham, the system of equations can be decoupled, by introduction of an external potential v_{eff} . The Kohn-Sham effective potential describes the potential provided by the environment, acting on a single orbital Ψ_j . This new property is constructed as a functional derivative of the density and is given as

$$v_{\text{eff}}(\mathbf{r}) = v_{\text{ext}}(\mathbf{r}) + \frac{\partial J[\rho]}{\partial \rho} + \frac{\partial E_{xc}}{\partial \rho} \quad , \quad (2.24)$$

consisting of an external potential part $v_{\text{ext}}(\mathbf{r})$, as well as, the derivatives of $J[\rho]$ and E_{xc} with respect to the density. Combining these terms together leads to N decoupled equations for the orbitals given as

$$\left(-\frac{1}{2}\Delta + v_{\text{eff}}(\mathbf{r}) \right) \Phi_i = \epsilon_i \Phi_i \quad , \quad (2.25)$$

describing the original many-body problem in principle exactly. The total energy is obtained by multiplying Eq (2.25) with Ψ^* and subsequent integration over \mathbf{r} . Therefore, the functional derivatives cancel out, leading to the expression of the total energy given as

$$E[\rho] = T_s[\rho] + J[\rho] + \int v_{\text{ext}}(\mathbf{r})\rho(\mathbf{r})d\mathbf{r} + E_{xc}[\rho] \quad . \quad (2.26)$$

All terms besides E_{xc} are known, and can therefore be calculated directly. In order to solve the system, an expression for E_{xc} has to be found. Various DFT methods have been developed, differing in their definition of E_{xc} , which will be discussed in the following. Common approaches include physically motivated models like the local density approximation (LDA)

as well as fitted functionals. However, up to present day, there exists no purely analytical ansatz for the construction of the correct functional, leaving its exact shape unknown.

2.5 The Exchange and Correlation Functional

To solve the KS problem, the construction of an explicit expression for E_{xc} is necessary. The fundamental considerations entering the chosen expressions of exchange and correlation contributions define the utilized framework, leading to subdivision of the DFT methods into various categories.

2.5.1 Local Density Approximation (LDA)

Historically, the first and simplest approach to define an approximation for E_{xc} is based on the so called local-density-approximation (LDA), where the systems E_{xc} is assumed to behave locally as the one of the uniform electron gas. Here, the exchange and correlation ϵ_{xc} of a uniform electron gas of the same density ρ are used, resulting in the expression of form

$$E_{xc}[\rho] = \int \rho \epsilon_{xc}(\rho) d\mathbf{r} \quad . \quad (2.27)$$

Typically, approximations are split into exchange ϵ_x and correlation ϵ_c parts respectively as

$$\epsilon_{xc} = \epsilon_x + \epsilon_c \quad . \quad (2.28)$$

For the uniform electron gas the exchange-interaction can be expressed analytically^{64,67} and is given as

$$\epsilon_x = \frac{3}{4} \left(\frac{3}{\pi} \right)^{1/3} \int \rho(\mathbf{r})^{4/3} d\mathbf{r} \quad . \quad (2.29)$$

The correlation part ϵ_c has to be approximated, as no exact formulation is known between the high and low density limits. Typical approaches involve quantum Monte Carlo simulations leading to a parameterized expression^{68–70}. For usual applications the influence of the correlation energy is an order of magnitude smaller than the one of the exchange one⁷¹. The major

2 Density Functional Theory

drawback of utilizing the LDA, is the assumption that the reference density for ϵ_{xc} is uniformly distributed. This approximation is especially unsuited for molecular systems which include strongly varying density parts. For chemically interesting systems, this usually leads to a poor description of fundamental properties such as bond-energies and the LDA, therefore, has limited applicability in theoretical chemistry. However, systems with slowly varying density and highly ordered structures, as for example in metals, allow its useful application.

Spin-Polarized Electrons and Local Spin Density Approximation (LSD)

When a system's magnetic effects are of interest or that consist of an odd number of electrons, as they are found for example in radicals and high spin states, it is necessary to split the orbitals into spin up and down contributions. Within the DFT framework, this is formally accomplished by introduction of an additional magnetization density $\mathbf{m}(\mathbf{r})$ ⁷¹ considering the possible orientations of spins with respect to each other. For most systems all spins are oriented either parallel or anti-parallel. Therefore, for most systems the explicit introduction of \mathbf{m} can be avoided by splitting the density in two spin channels. The correlations between states of different spin are considered as a modification of the exchange correlation functional. The total density is then comprised of the two channels' contributions as

$$\rho(\mathbf{r}) = \rho_{\uparrow}(\mathbf{r}) + \rho_{\downarrow}(\mathbf{r}) \quad , \quad (2.30)$$

with the spin channels constructed as

$$\rho_{\uparrow}(\mathbf{r}) = \sum_i \Phi_{i,\uparrow}^*(\mathbf{r})\Phi_{i,\uparrow}(\mathbf{r}) \quad \text{and} \quad \rho_{\downarrow}(\mathbf{r}) = \sum_j \Phi_{j,\downarrow}^*(\mathbf{r})\Phi_{j,\downarrow}(\mathbf{r}) \quad . \quad (2.31)$$

When the total number of spin up and spin down states is not equal, the expression for E_{xc} is modified to account for the non-vanishing total spin. To this end, the relative spin polarization ζ is introduced, given as

$$\zeta = \frac{\rho_{\uparrow} - \rho_{\downarrow}}{\rho_{\uparrow} + \rho_{\downarrow}} \quad . \quad (2.32)$$

2.5 The Exchange and Correlation Functional

For the exchange contribution of the net spin polarization, the spin scaling function $F(\zeta)$ is introduced, securing the high density limits as in this case the simple form of ζ is a non-sufficient. A simple form of $F(\zeta)$ has been derived for the LDA by Wang and Perdew in 1990⁷² carrying the name of local spin density (LSD) approximation and is given as

$$F(\zeta) = \frac{1}{2}[(1 + \zeta)^{4/3} + (1 - \zeta)^{4/3}] \quad . \quad (2.33)$$

With these additional considerations, the resulting exchange correlation term is given as

$$E_{xc}^{\text{LSD}}[\rho_{\uparrow}, \rho_{\downarrow}] = \int \rho(\mathbf{r})[\epsilon_x(\rho)F(\zeta) + \epsilon_c(r_s(\mathbf{r}), \zeta)]d^3r \quad . \quad (2.34)$$

This simple modification marks a good approximation for small-amplitude density fluctuations around the uniform density. Similar expressions are used for consideration of spin polarization in more advanced methods as described in the following.

2.5.2 Generalized Gradient Approximation (GGA)

As an improvement to the descriptions provided by the LDA functional, it is suitable to include the density gradient dependence in E_{xc} to account for possible inhomogeneities in the density distribution. Using this additional dependence, leads to the so called generalized gradient approximation (GGA). Unlike the LDA, the form of a GGA functional is not uniquely defined and up to the present day a multitude of functionals have been published. Compared to the LDA these approaches lead to much better descriptions of molecular systems, as they consider density fluctuations. Their general form is structured as:

$$E_{xc} = \int \rho(\mathbf{r})\epsilon_{xc}^{\text{GGA}}(\rho, \nabla\rho)d\mathbf{r} \quad (2.35)$$

In simple terms, one can describe the GGA-Functionals as a modification to the LDA, with their explicit expressions differing vastly by the utilized functional. The existing functionals are differentiated into classes, depending on the number of used parameters, fulfilled physical boundaries and empiricism.

2 Density Functional Theory

PBE-GGA

A standard example for a functional constructed on the basis of physical boundaries, is the PBE functional designed by Perdew, Burke and Ernzerhof in 1996⁷³. For the construction of this functional, physical boundaries that are determined to be energetically significant were used for improvement of the LDA expressions in regimes where they are not sufficient. The PBE exchange part ϵ_x is obtained by modifying the LDA exchange with the scaling function $\mathcal{F}(s)$, where s marks the system's dimensionless gradient term.

$$\epsilon_x^{\text{PBE}} = \epsilon_x^{\text{LDA}} \mathcal{F}_x(s) \quad . \quad (2.36)$$

The scaling function is given as

$$\mathcal{F}_x(s) = 1 + \kappa - \frac{\kappa}{1 - \mu s^2 / \kappa} \quad , \quad (2.37)$$

with the scaling parameters $\kappa = 0.804$ and $\mu \approx 0.21951$ and $s = \frac{|\nabla\rho|}{2\rho(3\pi^2\rho)^{1/3}}$ a dimensionless density gradient. For the correlation part, the basic LSD correlation is corrected by a gradient dependent term as

$$E_c^{\text{PBE}} = E_c^{\text{LSD}}(r_s, \zeta) + H(r_s, \zeta, s) \quad . \quad (2.38)$$

Herein r_s marks the systems Seitz radius, with the function H adjusted to fulfill physical boundary conditions. The second order gradient expansion has to be reproduced in the slow varying limit $s \rightarrow 0$, leading to a vanishing correlation for $s \rightarrow \infty$, if the system varies rapidly. Another boundary is the vanishing singularity for uniform scaling to the high density limit $\rho(\mathbf{r}) \rightarrow \lambda\rho(\lambda\mathbf{r})$ and $\lambda \rightarrow \infty$, leading to a constant correlation energy⁷⁴. These conditions are fulfilled by the ansatz

$$H = \gamma F(\zeta)^3 \ln \left(1 + \frac{1}{\chi s^2 / F(\zeta)^2 + (\chi s^2 / F(\zeta)^2)^2} \right) \quad , \quad (2.39)$$

where $\gamma \approx 0.031091$ and $\chi = 0.72161$ are parameters determined from references to physical boundaries⁷³. Combined with the previously described expressions for the exchange part, the

two obtained expressions lead to the full $E_{xc}^{\text{PBE}}(\zeta, \nabla\rho)$ -functional given as

$$E_{xc}^{\text{PBE}} = \epsilon_x^{\text{LSD}} \mathcal{F}_x(s) + \epsilon_c^{\text{LSD}} - H(r_s, \zeta, s) \quad , \quad (2.40)$$

showing the modifications to the corresponding two parts of the LSD approximation. This thesis mostly utilizes the standard version of the PBE functional as described in this section, alongside its modified versions revPBE⁷⁵ and PBEsol⁷⁶.

BLYP-GGA

Another commonly used GGA-functional is the BLYP functional, constructed by combining the B88 exchange functional (B)⁷⁷, and the Lee-Yang-Parr correlation functional (LYP)⁷⁸ that reformulates the HF correlation based formula by Colle and Salvetti⁷⁹. Both parts of the functional are constructed semi empirically with fitted parameters. The combined BLYP functional is then given as

$$E_{xc}^{\text{BLYP}} = E_x^{\text{LSD}} + E_x^{\text{B88}} + E_c^{\text{LYP}} \quad , \quad (2.41)$$

considering the B88 correction to the LSD exchange and the LYP correlation with equal parts.

2.5.3 Meta-GGA Functionals

A second step for improving the accuracy of the exchange and correlation treatment is provided by including the second derivative namely the Laplace operator of the density. However, in most cases the so called kinetic energy density τ_{kin} defined as

$$\tau_{\text{kin}} = \frac{1}{2} \sum_i^{\text{occ}} |\nabla \Psi_i|^2 \quad , \quad (2.42)$$

is used instead, due to higher numerical stability. The general form of dependence for meta-GGA functionals is given as

$$E_{xc} = \int \rho(\mathbf{r}) \epsilon_{xc}^{\text{Meta}}(\rho, \nabla\rho, \tau_{\text{kin}}) d\mathbf{r} \quad . \quad (2.43)$$

2 Density Functional Theory

Generally, meta-GGA functionals lead to improved accuracy compared to GGA and LSD calculations, however, they still belong to the group of semi-local functionals, meaning that the exchange correlation density at a point does solely depend on the properties at this specific point. Prominent functionals include for example the TPSS functional by Tao, Perdew, Staroverov and Scuseria⁸⁰.

2.5.4 Hybrid Functionals

Another way of enhancing the description of E_{xc} consists in adding a fraction of exact Hartree-Fock reference exchange interaction as shown in Eq (2.14), calculated with the Kohn-Sham orbitals. This approach is justified by the adiabatic connection formula, smoothly increasing the amount of electron-electron interaction from the purely non-interacting Kohn-Sham reference system. The adiabatic connection⁸¹ is given as

$$E_{xc} = \int_0^1 \left(\left\langle \sum_{i>j} \frac{1}{r_{ij}} \middle| \sum_{i>j} \frac{1}{r_{ij}} \right\rangle_{\lambda} - J[\rho^{\lambda}] \right) d\lambda = \int_0^1 U_{xc}^{\lambda} d\lambda \quad , \quad (2.44)$$

with λ being the parameter smoothly increasing the contribution of electron-electron interaction. From Eq (2.44), the Kohn-Sham case is obtained for $\lambda = 0$ whereas $\lambda = 1$ leads to a description of the full interacting system. Hence, U_{xc}^0 leads to the HF-exchange term calculated with the Kohn-Sham orbitals as

$$U_{xc}^0 = -\frac{1}{2} \sum_{ij} \int \frac{\Phi_i(\mathbf{r}_1)\Phi_i(\mathbf{r}_2)\Phi_j(\mathbf{r}_2)\Phi_j(\mathbf{r}_1)}{|\mathbf{r}_{12}|} d\mathbf{r}_1 d\mathbf{r}_2 \quad , \quad (2.45)$$

with the sum running over all pairs of occupied Kohn-Sham orbitals. The hybrid functionals are defined according to the amount of exact exchange added upon the calculation of Eq (2.45). This contribution is controlled by the parameter a_x leading to the expression

$$E_{xc}[\rho] = a_x U_{xc}^0 + (1 - a_x) E_{xc}^{\text{semilocal}}[\rho] \quad . \quad (2.46)$$

Here, $E_{xc}^{\text{semilocal}}$ marks the exchange contribution of the corresponding standard approximation used for E_{xc} e.g. a GGA functional. For the PBE functional the PBE0 add-on was designed

2.5 The Exchange and Correlation Functional

as the natural extension on the basis of theoretical justifications by using a fraction of 25% HF-exchange⁸², resulting in the form

$$E_{xc}^{\text{PBE0}} = b_0 E_x^{\text{HF}} + (1 - b_0) E_x^{\text{PBE}} + E_c^{\text{PBE}} \quad , \quad (2.47)$$

where $b_0 = 0.25$ considers the mentioned HF-exchange contribution. Using this approach usually increases greatly the calculation time as the advantages of the DFT-simplifications for the calculations of integrals are lost due to exact exchange. The resulting formulation of the exchange correlation functional is no longer semi-local, as the exact HF-exchange is considered. This can be formally understood due to the necessary double integration in space in Eq (2.13). The by far most popular functional in quantum chemistry is the B3LYP hybrid functional by Becke-Lee-Yang-Parr which is build using three parameters on the basis of the previously described empirically constructed BLYP functional. The general form of the B3LYP-exchange correlation contribution is given as

$$E_{xc}^{\text{B3LYP}} = (1 - a_0) E_x^{\text{LSDA}} + a_0 E_x^{\text{HF}} + a_x E_x^{\text{B88}} + a_c E_c^{\text{LYP}} + (1 - a_c) E_c^{\text{VWN}} \quad , \quad (2.48)$$

with the contributions weighted by the parameters as $a_0 = 0.2$, $a_x = 0.72$ and $a_c = 0.81$. Here, the correlation contribution E_c^{VWN} is given by the Vosko-Wilk-Nusair correction to the LDA⁸³.

2.6 The LCAO Approach

For the description of real systems, an explicit expression for the Kohn-Sham orbitals Φ_i has to be found. A suitable approach consists in expansion of the orbitals to avoid the search for a completely unknown function and the highly non-linearity of the optimization problem. Historically the choice of functions for the expansion in atomic orbitals is influenced by those known in explicit form for QM problems of similar nature, namely the solutions of the hydrogen-atom. Those are separable into radial $R_{n,l}$ and angular part $Y_{l,m}$ respectively.

$$\Phi_{n,l,m} = R_{n,l}Y_{l,m} \quad (2.49)$$

Herein the angular dependence is given in terms of the spherical harmonic functions $Y_{l,m}$ associated with the magnetic quantum number m and the angular quantum number l as

$$Y_{l,m} = Ne^{im\phi} P_l^m(\theta, \phi) \quad , \quad (2.50)$$

with the normalization constant N and the associated Legendre polynomials P_l^m obtained by solution of the generalized Legendre-differential equations. After substitution of $r = \frac{2Zr}{na_0}$, the radial part is given as

$$R_{l,n}(r) = \sqrt{\left(\frac{2Z}{na_0}\right)^3 \frac{(n-l-1)!}{2n((n+l)!)^3}} e^{-r/2} r^l \mathcal{L}_{n-l-1}^{2l+1}(r) \quad , \quad (2.51)$$

with n the principal quantum number, l being the angular momentum quantum number, a_0 the hydrogen atoms Bohr-radius and Z the charge of the nuclei. The $\mathcal{L}_{n,l}$ are the generalized Laguerre polynomials that are constructed following Rodriguez' formula.

The orbital wave-functions of electrons for elements other than hydrogen can not be calculated analytically, even in the atomic case. A possible approach is given by modeling the wave function in a linear combination of orbitals resembling the hydrogen wave function. Consequently, the obtained description is based on expansion of the orbitals, entering the total wave-function, in a sum of so called atomic orbitals (AO). This ansatz is usually referred to as linear combination of atomic orbitals (LCAO). An exact expression for AOs is only known

for the H-atom, for other elements they have to be approximated by a set of basis functions usually chosen to resemble the hydrogen type. Consequently, a basis set expansion of the molecular orbital Φ_i in atomic orbitals ϕ_α is given as

$$\Phi_i = \sum_{\alpha=1}^R c_{\alpha,i} \phi_\alpha \quad . \quad (2.52)$$

The quality of the expansion depends on the number and shape of the utilized basis functions for the system at hand. Multiple ways of finding possible functions for this purpose are popular within the community and will be briefly discussed in terms of their efficiency as well as suitability for specific types of problems.

2.6.1 Slater-Type Functions

The natural choice for expanding the AOs consists in a function that mimics the properties of the hydrogen-orbitals. In a first approach a set of modified functions reproducing the main features of those of the hydrogen type was used by Slater⁸⁴ which have the general form:

$$\phi^{\text{STO}} = \mathcal{N} r^{n-1} e^{-\zeta r} Y_{l,m}(\theta, \phi) \quad (2.53)$$

Slater-type functions are designed to reproduce the main characteristics of the hydrogen wave functions. These consist of a cusp for $r \rightarrow 0$ and the asymptotic decrease $\propto e^{-\zeta r}$. With these functions, usually only few basis functions are necessary to gain good approximations for the orbitals entering the many-body wave function. These properties come at the cost of an unfavorable mathematical shape, requiring numerical integration for many terms contributing to Eq (2.26). These are extremely costly when calculating the multi-center integrals of Eq (2.26). Due to this immense drawback these functions have limited application in modern quantum chemistry. As the precursors of later applications especially when considering Gaussian type basis functions, basis sets are often characterized in reference to Slater type orbitals e.g. (STO-3G) where three Gaussian orbitals are used to approximate one Slater Orbital.

2.6.2 Gaussian-Type Functions

Even though Slater-type orbitals offer the correct behaviors in reference to the hydrogen wave functions, it is straight forward to use themselves as reference for a computationally cheaper approach. The prominent choice for an alternative, is given by Gaussian functions, offering great mathematical simplifications, summarized in Section A.4, making their use extremely cost efficient. A Gaussian-type orbital $\eta^{\text{GTO}}(\mathbf{r})$ corresponding to quantum numbers n, l, m , is generally constructed by the superposition of a spherical harmonic function $Y_{l,m}$ and a Gaussian function with exponent ζ as

$$\eta^{\text{GTO}}(\mathbf{r}) = \mathcal{N} r^l Y_{l,m}(\mathbf{r}) e^{-\zeta(\mathbf{r}-\mathbf{R}_A)^2} \quad . \quad (2.54)$$

Here, \mathcal{N} is a normalization constant and the term r^l accounts for higher angular momentum. As the Slater-Type orbitals offer the most suitable description, therefore, a sum of multiple Gaussians is utilized to approximate one Slater-Type function. The resulting combined basis function ϕ^{CGTO} , is called contracted Gaussian basis, and can be expressed as

$$\phi^{\text{CGTO}} = \sum_k d_k \eta_k \quad . \quad (2.55)$$

The contraction coefficients d_k weigh the portion of the corresponding primitive Gaussian η_k , which only differ in terms of their exponents ζ values. For example in a standard triple- ζ basis set, for each orbital three different ζ values are used for the radial part of the Gaussians. The trade-offs when using Gaussians are the incompleteness and non orthogonality of the basis set. Furthermore, there is no systematic way of increasing the basis sets accuracy in an asymptotically convergent manner.

2.6.3 Self Consistent Field Approach (SCF)

Solving the Kohn-Sham problem by introduction of molecular orbitals expanded in atomic ones as described in Section 2.4, leads to

$$\hat{\mathbf{K}}\Phi_i = \epsilon_i \Phi_i \quad , \quad (2.56)$$

where the Kohn-Sham operator $\hat{\mathbf{K}}$ defined as

$$\hat{\mathbf{K}} = \left(-\frac{1}{2}\Delta + v_{\text{eff}}(\mathbf{r}) \right) , \quad (2.57)$$

has been introduced. The effective Kohn-Sham potential v_{eff} is directly depending on the molecular reference density as can be concluded from its definition Eq (2.24). Using a basis set expansion for the orbitals Φ_i , leads to

$$\hat{\mathbf{K}} \sum_{\alpha=1}^R c_{\alpha,i} \phi_{\alpha} = \epsilon_i \sum_{\alpha=1}^R c_{\alpha,i} \phi_{\alpha} . \quad (2.58)$$

Where, as the basis functions ϕ_{α} are fixed, the problem is only depending on the expansion coefficients $c_{\alpha,i}$. Multiplying Eq (2.58) with an arbitrary basis function ϕ_{λ} and integration over \mathbf{r} yields the algebraic basic equation for the description of the DFT problem given as

$$\sum_{\alpha=1}^R c_{\alpha,i} \int \phi_{\lambda}^* \hat{\mathbf{K}} \phi_{\alpha} d\mathbf{r} = \epsilon_i \sum_{\alpha=1}^R c_{\alpha,i} \int \phi_{\lambda}^* \phi_{\alpha} d\mathbf{r} . \quad (2.59)$$

This results in R equations that can be compacted by introducing $R \times R$ matrices, consisting of the energy matrix ϵ the Kohn-Sham matrix \mathbf{K} and the overlap matrix \mathbf{S} . Their elements are given as

$$S_{\lambda,\alpha} = \int \phi_{\lambda}^* \phi_{\alpha} d\mathbf{r}, \quad K_{\lambda,\alpha} = \int \phi_{\lambda}^* \hat{\mathbf{K}} \phi_{\alpha} d\mathbf{r} \text{ and } \epsilon_{\lambda,\alpha} = \epsilon_{\alpha} \delta_{\lambda,\alpha} . \quad (2.60)$$

Inserting the newly defined matrices in Eq (2.59), the resulting eigenvalue problem can be summarized as

$$\mathbf{KC} = \mathbf{SC}\epsilon , \quad (2.61)$$

with \mathbf{C} being the coefficient matrix. Solving for the whole systems' energy, the density can be expressed in terms of the expanded orbitals and takes the form

$$\rho(\mathbf{r}) = \sum_i^N \sum_{\lambda}^R \sum_{\alpha}^R f_i c_{i,\lambda}^* \phi_{\lambda}^* c_{i,\alpha} \phi_{\alpha} = \sum_{\lambda,\alpha}^R P_{\lambda,\alpha} \phi_{\lambda}^* \phi_{\alpha} . \quad (2.62)$$

2 Density Functional Theory

Here, the newly introduced matrix \mathbf{P} , is called density matrix and contains all relevant information about the density, consisting of the respective coefficients and occupation numbers of the corresponding orbitals Φ_i . Thus the whole DFT-problem can be expressed in terms of the basis functions and is now variational in the coefficients c , which can be summarized as

$$E_0 = \min_c [E_{\text{kin}}(c) + E_{\text{ext}}(c) + E_{\text{H}}(c) + E_{\text{xc}}(c)] \quad . \quad (2.63)$$

The ground state energy contributions correspond to the E_{kin} kinetic energy of electrons, E_{ext} external potential contributions, E_{H} Hartree or Coulomb interactions and E_{xc} to the exchange correlation energy respectively. During each SCF step, these coefficients are varied in order to minimize the energy functional in equation (Eq (2.63)). Each iteration changes the molecular orbital coefficients c and consequently also the reference density ρ (compare Eq (2.62)) entering the definition of v_{eff} . It is therefore not sufficient to simply solve Eq (2.63) once but rather to loop over it until a self reproducing solution is obtained. The initial choice of coefficients c is further optimized until the relative change in v_{eff} between two consecutive SCF-steps is no longer significant.

2.7 Density Functional Theory for Periodic Systems

For systems with crystal symmetry, an extended bulk can be described by expansion of a unit cell via periodic boundary conditions (PBCs). As the system is periodic this is also the case for the underlying potential and thus additional features depending on the shape of the underlying periodicity are introduced in the potential term of $\hat{\mathbf{H}}$. A periodic system in space, is constructed by repetition of unit cells. The lattice can be described in the three dimensional space via its unit vectors \mathbf{a}_i with $i \in \{1, 2, 3\}$. The total system is then built up by a basis consisting of an atom, molecule or ionic group placed in the unit-cell, that is periodically repeated in the desired dimensions. In Figure 2.1 an exemplary two-dimensional lattice with cell vectors \mathbf{a}_i is shown. The following basic discussion of systems with PBCs is presented^{85,86}.

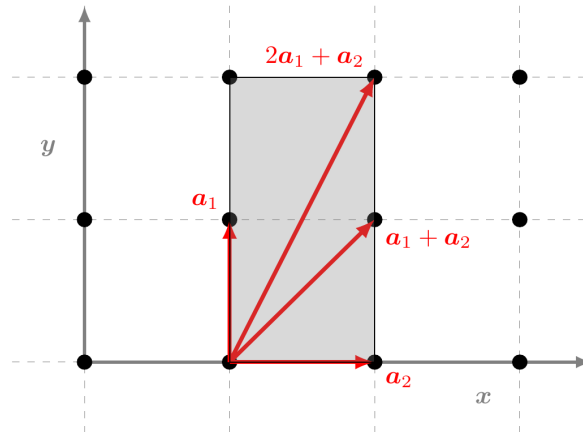


Figure 2.1: Schematic representation of the lattice description in real space in two dimensions with cell vectors \mathbf{a}_1 and \mathbf{a}_2 .

Direct Lattice

Any lattice in real space, can be defined by a unit cell matrix \mathbf{A} consisting of the cell vectors \mathbf{a}_i in the form

$$\mathbf{A} = \begin{pmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \mathbf{a}_3 \end{pmatrix} . \quad (2.64)$$

The volume Ω of the corresponding unit cell box can then be calculated as

$$\Omega = \det(\mathbf{A}) . \quad (2.65)$$

Any point on the lattice can be referred to in terms of a lattice vector built up of multiples of the cell vectors and can be written in three dimensions as

$$\mathbf{R}_{n_1, n_2, n_3} = n_1 \mathbf{a}_1 + n_2 \mathbf{a}_2 + n_3 \mathbf{a}_3 . \quad (2.66)$$

Where the integer valued multipliers $n_1, n_2, n_3 \in \mathbb{Z}$ identify the specific position with respect to the elementary unit cell. The potential of a periodic system in real space has to be invariant

2 Density Functional Theory

under translation with an arbitrary lattice vector \mathbf{R} and fulfills the condition

$$V(\mathbf{r} + \mathbf{R}) = V(\mathbf{r}) \quad . \quad (2.67)$$

Considering the potentials periodicity, the wave functions describing the eigenstates of the Hamiltonian, with a periodic potential of form Eq (2.67), have to be modified accordingly.

Reciprocal Lattice

For every real space lattice a corresponding one exists in reciprocal space. Both representations are connected via a Fourier-transform. The reciprocal cell vectors \mathbf{b}_i can be constructed from the cell vectors of the real space system as

$$\mathbf{b}_i = 2\pi \frac{\mathbf{a}_j \times \mathbf{a}_k}{\mathbf{a}_i \cdot [\mathbf{a}_j \times \mathbf{a}_k]} = \frac{2\pi}{\Omega} (\mathbf{a}_j \times \mathbf{a}_k) \quad . \quad (2.68)$$

The connection to the real space lattice vectors can be summarized as

$$\mathbf{a}_i \cdot \mathbf{b}_j = 2\pi \delta_{i,j} \quad \text{with} \quad i, j \in \{1, 2, 3\} \quad (2.69)$$

Any vector on the reciprocal lattice can thus be constructed by

$$\mathbf{G}_{h,k,l} = h\mathbf{b}_1 + k\mathbf{b}_2 + l\mathbf{b}_3 \quad , \quad (2.70)$$

with $h, k, l \in \mathbb{Z}$ which are also known as Miller-indices. Consequently, the Hamiltonian describing the periodic system has to be modified in comparison to the isolated system. These modifications can be compacted as a sum over additional interaction terms with adjacent unit cells. Resulting from the periodic description, the basis functions have to be extended periodically to fully describe the extended system, otherwise they wouldn't yield a periodic density.

2.8 Description of the Solid State

Solid state systems can be generally divided in crystalline and amorphous materials. The prior ones are in the usual case, constructed with inherent symmetries which can be exploited by describing them via periodic expansion of a unit cell. In amorphous materials, this is not possible and utilized models require super-cell approaches with thousands of atoms per unit cell. In order to describe crystalline systems, a lattice description as described in Section 2.7 can be used along with PBCs.

The introduction of PBCs adds additional conditions for the corresponding wave-functions, therefore, the description of the solid state differs formally from the molecular case.

The main differences in the description, arise from the typically $N \approx 10^{24}$ particles of a usual solid bulk that is de facto modeled. Therefore, even though only a unit cell with a limited number of atoms is considered, additional effects stemming from the total system have to be considered. In summary the additional features consist of:

- The crystal potential is described as cell periodic.
- The wave-function has to fulfill periodic boundary conditions.
- The additional treatment requires inclusion of “all” possible electronic states within the given boundary conditions which results in the formation of energy bands.

These additional requirements are accounted for by modification of the underlying description, starting from the periodic crystal potential entering the Hamiltonian. New eigenfunctions can be constructed according to Blochs theorem⁸⁶. Due to the translational symmetry of the lattice, the wave function can only change by a phase factor under translation by an arbitrary lattice vector \mathbf{R} as defined by Eq (2.66), resulting in

$$\Psi(\mathbf{r} + \mathbf{R}) = e^{i\mathbf{k}\mathbf{R}}\Psi(\mathbf{r}) \quad . \quad (2.71)$$

The eigenstates of a periodic Hamiltonian should lead to the same value of observables at equivalent positions in different lattices sites. This is accounted by the lattice periodic part (also considered Bloch factor) for the wave function fulfilling $u_{\mathbf{k}}(\mathbf{r} + \mathbf{R}) = u_{\mathbf{k}}(\mathbf{r})$. Here, it has to be noted that the Pauli principle has to be fulfilled and two electrons can't be in the

2 Density Functional Theory

exact same quantum state. To account for the multitude of possible \mathbf{k} -values, and by means of periodicity, the function can be multiplied by a plane wave $e^{i\mathbf{k}\mathbf{r}}$ accounting for all possible contributions for different \mathbf{k} values. In summary this can be expressed as

$$\Psi_{\mathbf{k}} = e^{i\mathbf{k}\mathbf{r}} u_{\mathbf{k}}(\mathbf{r}) \quad , \quad (2.72)$$

where compared to the vacuum case, the wave functions additionally depend on the crystal momentum \mathbf{k} . This momentum acts as an additional quantum number that differentiates the electrons of different cells originally located at equivalent orbitals in the gas phase. The combination of a number of electrons as modeled for a solid, results in a multitude of energies for various \mathbf{k} -values and therefore introduces an additional dependency $E_n(\mathbf{k})$. Therefore, in the solid state energies are not only quantized in analogy to the corresponding molecular orbitals but are further split into continuous energy-bands depending on \mathbf{k} and, in the regular case, energetically separated for different principal quantum numbers n .

As a result of the additional \mathbf{k} -dependence, the eigenfunctions of the periodic Hamiltonian are no longer localized, meaning they are expanded in the full periodic grid rather than limited on the description of a specific unit cell. In general, Bloch-functions leave up a free phase factor as well as a gauge indeterminacy, consequently, the application of any unitary transformation onto these orbitals yields the same physical description⁸⁷. Further details are expanded in Section 3.3, where the solid state analogue to molecular orbitals, the so called Wannier-functions and their general interpretation are discussed.

2.8.1 Periodic Atom Centered Gaussian Basis Functions

In order to describe a periodic system with localized Gaussian basis functions, the previously introduced formulation of Section 2.6.2, has to be adjusted to suite the new description. The localized basis functions are extended by summarizing over adjacent cells, differing by multiples of the lattice vectors. The resulting functions are then used as Bloch-factors $u_{\mathbf{k}}(\mathbf{r})$ in Eq (2.72). A simple approach to construct a wave function from localized orbitals, is to

expand it into a basis set as

$$\Psi_n^k = \sum_A c_{A,n}^k \Phi_A^k \quad , \quad (2.73)$$

where n describes the number of the corresponding electrons as given in Eq (2.61). Here, the basis functions Φ_A^k have to be Bloch functions following Eq (2.72). Using localized functions as Bloch-factor for the plane wave dependence, results in a superposition of the local basis functions given as

$$\Phi_A^k = \sum_L e^{ikL} \phi_A^L \quad . \quad (2.74)$$

Here, the summation is carried out over real space grid vectors L . In all applications used in this thesis and also in most popular codes, contracted Gaussian functions are used for the $\phi_A^L(\mathbf{r})$ and can be constructed as

$$\phi_A^L(\mathbf{r}) = \sum_i d_{A,i} \eta_{A,i}^L(\mathbf{r}) \quad , \quad (2.75)$$

where the index i runs over all primitive Gaussian functions $\eta_{A,i}^L$ ⁸⁸.

2.8.2 Plane-Wave Basis Sets

Another possibility to expand the orbitals Ψ_i , especially suitable for periodic systems, is given by plane waves. These functions are often used in solid state calculations with PBCs due to their inherent periodicity with respect to a unit cell. Let this unit cell for instance be spanned by the cell vectors $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3$ with the volume $\Omega = \mathbf{a}_1(\mathbf{a}_2 \times \mathbf{a}_3)$. For a periodic function, the Fourier-theorem states that an expansion in terms of plane waves (PWs) is always possible. Such an expansion for the orbitals takes the general shape

$$\Phi_i^{\text{PW}}(\mathbf{r}) = \frac{1}{\sqrt{\Omega}} \sum_G c_i(\mathbf{G}) e^{i\mathbf{G}\mathbf{r}} \quad , \quad (2.76)$$

where \mathbf{G} denotes the reciprocal lattice vector. The inherent periodicity makes PWs the straight forward basis set for crystallographic problems, but also marks their main disadvantage in case

2 Density Functional Theory

of non-periodic systems. PWs form in principle a complete basis set for $G \rightarrow \infty$, meaning that increasing the maximum value of possible values of G , allows for a systematically more accurate description. In real applications a cutoff value has to be defined E_{cut} , which is usually done in analogy to the kinetic energy of the plane waves. The corresponding relation is given as

$$E_{\text{kin}} = -\frac{1}{2} \langle \Phi | \Delta | \Phi \rangle = -\frac{1}{2\Omega} G^2 \leq E_{\text{cut}} \quad , \quad (2.77)$$

with the basis set size N_{PW} determined as

$$N_{\text{PW}} \approx \frac{1}{2\pi^2} \Omega E_{\text{cut}}^{3/2} \quad . \quad (2.78)$$

Due to their non-locality, plane waves are independent of the atomic positions and span the whole space of the box volume Ω . A result of this property is the usually high number of basis functions even in regions where no electrons are located. As a consequence the number of basis functions scales directly with the box size as does the computational effort.

2.9 Energy Evaluation for Periodic Systems (*k*-Point Sampling)

In contrast to the unbounded case in a periodic system the total energy is defined in reference to the unit cell, or to be more precise, the Brillouin-zone⁸⁹. This is a direct consequence of the wave function satisfying Bloch's theorem⁸⁶, which requires the consideration of all possible boundary conditions of type Eq (2.71), with \mathbf{k} -vectors lying in the unit-cell of the reciprocal lattice, also known as the first Brillouin zone (BZ). The total energy of a system within PBCs can be determined by an integral over the Brillouin-zone given as

$$E_{\text{tot}} = \frac{1}{\Omega} \int_{\text{BZ}} E(\mathbf{k}) d\mathbf{k} \quad , \quad (2.79)$$

where Ω denotes the corresponding Brillouin-zone volume. This integral can be evaluated by using the mean value theorem (compare Section A.8), for any lattice periodic function by

2.9 Energy Evaluation for Periodic Systems (\mathbf{k} -Point Sampling)

expansion in a Fourier-series. An example is provided by

$$E(\mathbf{k}) = \sum_j^{j_{\max}} E_j e^{i\mathbf{k}\mathbf{R}_j} \quad , \quad (2.80)$$

which can be used to evaluate Eq (2.79) by determining the first coefficient E_0 of the series. To use this assumption, the function $E(\mathbf{k})$ has to be evaluated at certain points in space, as the coefficients in Eq (2.80) are determined by

$$E_j = \frac{1}{\omega_j} \int_{\text{BZ}} e^{-i\mathbf{k}\mathbf{R}_j} E(\mathbf{k}) d\mathbf{k} \quad . \quad (2.81)$$

In order to determine the coefficients E_j exactly, the knowledge of the function $E(\mathbf{k})$ in the full \mathbf{k} -space is required. This would not lead to any computational benefit and in order to avoid this, the integral in Eq (2.81) is approximated by a weighted sum over special points in \mathbf{k} space. In a very simplistic approach this leads to

$$E_0 \approx \sum_i^{N_{\text{special}}} w_i E(\mathbf{k}_i) \quad . \quad (2.82)$$

The points in \mathbf{k} -space are chosen in order to represent the main features of Eq (2.80), with the upper boundary j_{\max} usually determined by means of computational power. For the choice of \mathbf{k} points for evaluation, multiple schemes are well-known in the literature, for example the method of Monkhorst and Pack⁹⁰. Most schemes utilize symmetry assumptions for justification of the sampling points⁹¹. However, the reciprocal lattice volume is inversely connected to the real-space lattice, which results in small reciprocal lattices for large real-space unit cells. For large unit-cells it is therefore sufficient to evaluate the Brillouin zone only at the Γ -point, which marks the origin of the reciprocal lattice. In summary, it is suitable to approximate Eq (2.79) by

$$E_{\text{tot}} = E_0 \approx E(\mathbf{k}_{\Gamma}) \quad . \quad (2.83)$$

This approximation is referred to as Γ -point sampling and was used for all calculations for periodic systems throughout this thesis.

2.10 Goedecker-Theter-Hutter-Pseudopotentials

Pseudopotentials (PP) offer an elegant way for reduction of computational effort by avoiding the explicit treatment of core electrons⁹². This is especially desirable when utilizing PW basis sets, as the wave functions of core electrons are usually very peaked, requiring large numbers of PWs for their description (see Section 2.8.2). Additionally, scalar relativistic effects can be included leading to potential improvements of the description. In order to avoid direct calculation of core electrons, they are approximated by a potential mimicking their influence on other electrons. For electronic systems that are mainly determined by valence properties such as covalent bonds, the PP treatment of the core electrons has usually no crucial effect on the accuracy and an explicit treatment of the valence is sufficient. Due to the reduction of the size of the explicitly calculated system during the SCF, the effective system size is reduced anyway, leading to reduction in calculation times. The use of PPs is therefore predestined for application with a PW basis, or when atoms with numerous electrons such as transition metals e.g. gold, silver or platinum are treated.

PPs are constructed by explicitly calculating the all-electron core wave function in the atomic case and approximating the influence of the core electrons on the valence, by introduction of an artificial potential. This simplification, removes the difficulties of peaked electronic core wave-functions close to $r = 0$. Therefore, not only the total number of basis functions needed to approximate the system's density is reduced, but also the wave-function cusp at the core is eliminated. Exemplary ways to construct PPs can be found in works of Blöchl⁹³, Stoll⁹⁴, Hamann⁹⁵, Kresse⁹⁶, Mi⁹⁷ and many others. The construction of PPs is still a topic of current investigation, as can be exemplarily seen in⁹⁸. The potentials created by Goedecker Theter and Hutter (GTH), that are used throughout this work, are of the non-local norm conserving dual-space type and specifically constructed for the application with Gaussian basis functions. The shape of the Potential functions is chosen to resemble a Gaussian in order to simplify follow up calculations^{99–101}. GTH-PPs are separable into a sum of a non-local and local contribution with the total potential $V_{\text{tot}}^{\text{PP}}$ given as

$$V_{\text{tot}}^{\text{PP}} = V_{\text{loc}}^{\text{PP}} + V_{\text{nl}}^{\text{PP}} \quad . \quad (2.84)$$

The local contribution is calculated as

$$V_{\text{loc}}^{\text{PP}}(r) = -\frac{Z_{\text{ion}}}{r} \text{erf}(\alpha^{\text{PP}} r) \sum_{i=1}^4 C_i^{\text{PP}} (\sqrt{2}\alpha^{\text{PP}} r)^{2i-2} e^{-(\alpha^{\text{PP}} r)^2}, \quad (2.85)$$

with $\alpha_{\text{PP}} = \frac{1}{\sqrt{2}r_{\text{loc}}^{\text{PP}}}$, Z_{ion} the charge of the corresponding reference ion, C_i^{PP} the expansion coefficients and r_{loc} the cutoff radius for the local part. The non-local contribution is given as

$$V_{\text{nl}}^{\text{PP}}(\mathbf{r}, \mathbf{r}') = \sum_{l,m} \sum_{i,j} \langle \mathbf{r} | p_i^{lm} \rangle h_l^{ij} \langle p_j^{lm} | \mathbf{r}' \rangle. \quad (2.86)$$

Here, the Gaussian projectors on the real space are given as

$$\langle p_j^{lm} | \mathbf{r} \rangle = N_i^l Y_{l,m}(\hat{r}) r^{l+2i-2} e^{-\frac{1}{2}\left(\frac{r}{r_l}\right)^2}. \quad (2.87)$$

The resulting total PPs can be fitted by using a small set of parameters for each element¹⁰¹. The potentials used in this work were constructed using relativistic all electron DFT reference calculations for specific functionals. The nomenclature will be explained at the example of lithium and the PBE functional, where the corresponding PP is denoted as GTH-PBE-q1, with PBE corresponding to the used functional and q1 indicating the number of electrons that are explicitly calculated. The use of GTH-PPs limits the applicability of general all electron basis sets and those constructed for the use with other types of PPs. For the applicability with their GTH-PPs, a special type of Gaussian basis sets denoted with GTH, was constructed uniquely for the use with their GTH-PP counterparts. Details on the construction of the associated basis sets are described in⁸⁸.

2.11 Gaussian Plane Wave DFT

Previously in this chapter the fundamental benefits and limitations of using either a Gaussian or plane-wave basis sets were discussed for isolated molecular systems and the solid state. However, in surface chemistry, situations where both methods reach the limits of applicability such as molecule-surface interfaces, are common. Here, plane waves are suitable for the surface's description, whereas the attached molecule is better treated by Gaussian functions.

2 Density Functional Theory

The fundamental idea of the Gaussian and plane wave (GPW) approach, is to combine the two descriptions into one work flow, exploiting the benefits of both basis sets. This approach is rather unique and up to the present day only few quantum chemistry codes use the GPW method. The general work flow of the method as it is implemented in the program package CP2K¹⁰², which is mainly used for the applications in this thesis, is presented in the following. The description follows the way it is presented in works by Hutter, Krack, Iannuzzi and Vandevondele^{62,103–105}.

2.11.1 The Basic Idea of the Gaussian and Plane Waves Approach

During the process of solving the SCF equations with Gaussian basis functions, the contributions to the total energy functional differ in terms of their necessary computational effort. The GPW approach aims at eliminating the drawbacks of Gaussian functions by switching to an auxiliary basis set given in plane waves. Finding a solution for Eq (2.63), requires the calculation of five respective integrals. The kinetic energy integral I_{kin} , given as

$$I_{\text{kin}} = \int \Phi_i^*(\mathbf{r}) \Delta \Phi_i(\mathbf{r}) d\mathbf{r} \quad . \quad (2.88)$$

The integral considering the external potential's contribution, given as

$$I_{\text{ext}} = \int V_{\text{ext}} \rho(\mathbf{r}) d\mathbf{r} \quad . \quad (2.89)$$

The Coulomb four center integrals I_{Coulomb} of form

$$I_{\text{Coulomb}} = \int \int \rho(\mathbf{s}) \frac{1}{|\mathbf{r} - \mathbf{s}|} \rho(\mathbf{r}) d\mathbf{s} d\mathbf{r} \quad . \quad (2.90)$$

The exchange and correlation integral I_{xc} , given as

$$I_{\text{xc}} = \int F[\rho] \rho(\mathbf{r}) d\mathbf{r} \quad , \quad (2.91)$$

and last but not least the overlap integrals I_{overlap} , given as

$$I_{\text{overlap}} = \int \phi_i^*(\mathbf{r})\phi_j(\mathbf{r})d\mathbf{r} \quad . \quad (2.92)$$

The properties of Gaussian functions allow to solve the kinetic energy integral Eq (2.88), the overlap integrals Eq (2.92) as well as the external potential integral Eq (2.89) analytically. However, the bottleneck, in terms of computational effort, is usually the calculation of the Coulomb four-center integrals. The number of four center integrals, scales heavily with the system size and therefore limits the applicability of Gaussian basis functions for large systems. Multiple approaches try to minimize the effort, for example the resolution of identity method (RI) which reduces the four center to three center integrals. The scaling with system and basis-set size is too high for the efficient application of Gaussian basis functions for large systems. Furthermore, the exchange and correlation contribution requires numerical integration in any case, making it computationally expensive as well. It is known, that utilizing a PW basis set allows for drastic simplifications of the Coulomb term treatment. Consequently, the highest efficiency for calculating Eq (2.63) is achieved by combining the two basis sets in one approach for simplified treatments. Here Gaussians are used for Eq (2.88), Eq (2.92),Eq (2.89) whereas PWs are used for Eq (2.90) and Eq (2.91). Another crucial step in terms of efficiency, is given when calculating all types of Coulomb interactions in a single procedure as it will be shown in Section 2.11.3.

2.11.2 GPW Preparation Scheme

The most important steps of the GPW-workflow as implemented in the QUICKSTEP module will be explained in the following¹⁰⁶. Initially, the electron density ρ is constructed using a linear combination of atom centered contracted Gaussian basis functions ϕ_α for the orbitals Φ_i as

$$\Phi_i = \sum_{\alpha} c_{\alpha,i}\phi_{\alpha} \quad . \quad (2.93)$$

2 Density Functional Theory

The resulting density matrix \mathbf{P} entering the SCF-calculation as shown in Eq (2.61) is then of shape

$$P_{\lambda\alpha} = \sum_i f_i c_{\lambda,i}^* c_{\alpha,i} \quad , \quad (2.94)$$

with f_i being the occupation number. With this matrix the total density of the system is expressed as

$$\rho = \sum_{\lambda,\alpha} P_{\lambda,\alpha} \phi_\lambda \phi_\alpha \quad . \quad (2.95)$$

The density after construction with Gaussian basis functions can then be expanded in plane waves by using a Fourier transform. In order to expand the density it is suitable to reformulate equation Eq (2.95) by exploiting the properties of Gaussian functions, presented in Section A.4. The product of two Gaussians is again a Gaussian with shifted center following the treatment in Eq (A.22). This connection, leads to product functions $\phi_{\lambda,\alpha} = \phi_\lambda \phi_\alpha$ of shape

$$\phi_{\lambda,\alpha}(\mathbf{r}) = N_p e^{-\alpha_p(\mathbf{r}-\mathbf{R}_p)^2} \quad , \quad (2.96)$$

where N_p is the normalization constant, α_p the width parameter and \mathbf{R}_p the center of the shifted Gaussian as shown in Section A.4. With the shape of the product function being Gaussian, the plane wave expansion can be easily expressed as the Fourier transform of a Gaussian remains a Gaussian in reciprocal space, as can be seen in Eq (A.24). The shifted Gaussians are expanded in plane waves like

$$\phi_{\lambda,\alpha}(\mathbf{r}) = \sum_{\mathbf{G}} \phi_{\lambda,\alpha}(\mathbf{G}) e^{i\mathbf{G}\mathbf{r}} \quad , \quad (2.97)$$

where the sum over all lattice vectors $\mathbf{G} = \mathbf{G}_{h,k,l}$ is truncated at a maximum value N in each respective dimension of the reciprocal space. The auxiliary density $\tilde{\rho}$ is then given as

$$\tilde{\rho} = \sum_{\lambda,\alpha} P_{\lambda,\alpha} \sum_{\mathbf{G}} \phi_{\lambda,\alpha}(\mathbf{G}) e^{i\mathbf{G}\mathbf{r}} \quad .$$

This process is mathematically similar to the methods used for density fitting, as explained for example in¹⁰⁷. The newly defined auxiliary density $\tilde{\rho}$ is a function of the $c_{\lambda,\alpha}$ and the PW functions. This new representation can now be readily applied for simplifying the integrals whose calculation would be challenging utilizing Gaussian basis functions. In summary, the total energy functional in GPW formalism is given as

$$E^{\text{GPW}} = E_{\text{kin}}(\rho) + E_{\text{pseudo}}(\rho) + E_{\text{xc}}(\tilde{\rho}) + E_{\text{H}}(\tilde{\rho}) + E_{\text{overlap}} - E_{\text{self}} \quad . \quad (2.98)$$

Here, the auxiliary basis $\tilde{\rho}$ is used for calculation of exchange-correlation and Hartree terms, whereas the remaining terms are calculated using the Gaussian representation ρ . The resulting total energy expression E^{GPW} remains variational in the coefficients c of the Gaussians. Therefore, the energy expression can be minimized within an SCF-approach as explained in Section 2.6.3.

2.11.3 Coulomb Interaction Treatment in CP2K

The main goal of utilizing PWs is the avoidance of the explicit calculation of the electronic repulsion integrals. The utilization of GTH-PPs gives rise to an additional electrostatic term which, together with the external potential provided by electron-nuclei and the electron-electron interaction term, has to be calculated separately. In order to avoid unnecessary numerical integrations caused by this particularization of the Coulomb term, an approach to combine all electrostatic interactions into one term seems promising. In summary, the electrostatic contributions to the total energy are given as

$$E_{\text{C}} = \underbrace{\frac{1}{2} \int \int \frac{\rho(\mathbf{s})\rho(\mathbf{r})}{|\mathbf{r} - \mathbf{s}|} d\mathbf{s}d\mathbf{r}}_{E_{\text{ee}}} - \underbrace{\sum_{\text{A}} Z_{\text{A}} \int \frac{\rho(\mathbf{r})}{|\mathbf{r} - \mathbf{R}_{\text{A}}|} d\mathbf{r}}_{E_{\text{en}}} + \underbrace{\sum_{\text{A} \neq \text{B}} \frac{Z_{\text{A}}Z_{\text{B}}}{|\mathbf{R}_{\text{A}} - \mathbf{R}_{\text{B}}|}}_{E_{\text{nn}}} \quad , \quad (2.99)$$

where Z_{A} and Z_{B} are the charges and R_{A} and R_{B} the positions of ion A and B. A mathematical simplification is not possible without reformulation of the problem. To this end, a new type of density is defined, describing both, the electronic and the nuclear charge distribution. The

2 Density Functional Theory

total density is then given as the sum of both contributions

$$\rho_{\text{tot}}(\mathbf{r}) = \rho_{\text{el}}(\mathbf{r}) + \sum_A \rho_A(\mathbf{r}) \quad , \quad (2.100)$$

where ρ_{el} is the system's total electronic, whereas ρ_A describes only the nuclear contribution of atom A , to the total density. The main difference arising from this formulation, is given by the finite extension of the nuclear charge around the atoms position. This is realized by placing a Gaussian shaped charge distribution around R_A carrying the total core charge Z_A of the corresponding Ion A . The core charge distribution ρ_A then takes the form

$$\rho_A(\mathbf{r}) = Z_A \left(\frac{\alpha}{\pi} \right)^{3/2} e^{-\alpha(\mathbf{r}-\mathbf{R}_A)^2} \quad . \quad (2.101)$$

In the next step the total electrostatic energy is rewritten in terms of the total density following Eq (2.100), resulting in the newly defined term structure for the total electrostatics of form

$$E_{\text{C,tot}} = \frac{1}{2} \int \int \frac{\rho_{\text{tot}}(\mathbf{s})\rho_{\text{tot}}(\mathbf{r})}{|\mathbf{r} - \mathbf{s}|} d\mathbf{s}d\mathbf{r} \quad , \quad (2.102)$$

with α being the width parameter of the Gaussian. Considering the total density and substituting it in Eq (2.102), the total electrostatic contribution can be expanded as

$$E_{\text{C}} = -\frac{1}{2} \int \int \frac{\rho_{\text{el}}(\mathbf{s})\rho_{\text{el}}(\mathbf{r}) + \rho_{\text{el}}(\mathbf{s})\rho_A(\mathbf{r}) + \rho_A(\mathbf{s})\rho_{\text{el}}(\mathbf{r}) + \rho_A(\mathbf{s})\rho_A(\mathbf{r})}{|\mathbf{r} - \mathbf{s}|} d\mathbf{s}d\mathbf{r} \quad , \quad (2.103)$$

where terms corresponding to the ones in Eq (2.99) are denoted with the same colors. However, these terms do not exactly represent the same situation as previously nuclear point charges were treated, thus, additional correction terms have to be considered resulting in the total electrostatic contribution given as

$$E_{\text{C}} = E_{\text{C,tot}} + E_{\text{C,err}} \quad , \quad (2.104)$$

where the correction $E_{C, \text{err}}$ is the sum of three separate terms and given as

$$E_{C, \text{err}} = \frac{1}{2} \sum_{A \neq B} \frac{Z_A Z_B}{|\mathbf{R}_A - \mathbf{R}_B|} \text{erfc}(\sqrt{2}\alpha|\mathbf{R}_A - \mathbf{R}_B|) - \sum_A Z_A \frac{\text{erf}(\alpha(\mathbf{r} - \mathbf{R}_A))}{|\mathbf{r} - \mathbf{R}_A|} \rho_{el}(\mathbf{r}) - \sum_A \frac{\alpha}{\sqrt{2\pi}} Z_A^2 . \quad (2.105)$$

The first term corresponds to the overlap of two nuclear Gaussians, the second term to the electron core overlap and the last term to the interaction of the extended core-charge distributions with themselves. The second part is included within the specialized GTH-PPs in equation (Eq (2.85)) leaving only terms independent of \mathbf{r} in the energy correction $E_{C, \text{err}}$. Consequently, the correction is trivially given and adds minimal computational effort during the calculation of the energy. With this treatment, only a single term is left for integration

$$E_{C, \text{tot}} = \frac{1}{2} \int \int \frac{\rho_{\text{tot}}(\mathbf{s})\rho_{\text{tot}}(\mathbf{r})}{|\mathbf{r} - \mathbf{s}|} d\mathbf{s} d\mathbf{r} . \quad (2.106)$$

The remaining task for solving the DFT problem in this formalism, is now the solution of the Coulomb-integral for the total charge density ρ_{tot} . For this purpose, it is helpful to consider that the electrostatic potential $\mathcal{V}(\mathbf{r})$ is defined as

$$\mathcal{V}(\mathbf{r}) = \int \frac{\rho_{\text{tot}}(\mathbf{s})}{|\mathbf{s} - \mathbf{r}|} d\mathbf{s} , \quad (2.107)$$

which can be identified in Eq (2.106) for further simplification. It is known, that $\mathcal{V}(\mathbf{r})$ is the solution of Poisson's equation for electrostatics, which is of shape

$$\Delta_{\mathbf{r}} \mathcal{V}(\mathbf{r}) = -4\pi\rho_{\text{tot}}(\mathbf{r}) . \quad (2.108)$$

Furthermore, if the total density is given in plane waves as

$$\rho_{\text{tot}}(\mathbf{r}) = \sum_{\mathbf{G}} \rho_{\text{tot}}(\mathbf{G}) e^{i\mathbf{G}\mathbf{r}} , \quad (2.109)$$

2 Density Functional Theory

the solution of Poisson's equation is trivially given. Using Eq (2.109) in the expression of Eq (2.107) and substituting it in Eq (2.108) leads to

$$\Delta_{\mathbf{r}} \mathcal{V}(\mathbf{r}) = 4\pi \sum_{\mathbf{G}} \rho_{\text{tot}}(\mathbf{G}) e^{i\mathbf{G}\mathbf{r}} \quad , \quad (2.110)$$

which can be further simplified¹⁰⁸⁻¹¹⁰ if the potential itself is expected to be representable in plane waves. A right hand approach for the electrostatic potential results in the simple equation

$$\mathcal{V}(\mathbf{r}) = \frac{4\pi \rho_{\text{tot}}(\mathbf{r})}{\mathbf{G}^2} \quad . \quad (2.111)$$

As a result, the potential is obtained by simply dividing the coefficients in the sum of the PW expansion by the corresponding value of \mathbf{G}^2 . Therefore, Eq (2.106) can be re-expressed using Parseval's theorem, for integrals over functions expandable in plane waves, described in Section A.6, leading to

$$E_{\text{C,tot}} = \frac{1}{2} \int \mathcal{V}(\mathbf{r}) \rho_{\text{tot}}(\mathbf{r}) d\mathbf{r} = \frac{1}{2} \sum_{\mathbf{G}} \mathcal{V}^*(\mathbf{G}) \rho_{\text{tot}}(\mathbf{G}) \quad , \quad (2.112)$$

where * marks the complex conjugated. Utilizing the fact that $\mathcal{V}(\mathbf{G})$ is directly connected to the density, the Coulomb term can be further simplified to

$$E_{\text{C,tot}} = -\frac{2\pi}{\Omega} \sum_{\mathbf{G}} \frac{\rho_{\text{tot}}^*(\mathbf{G}) \rho_{\text{tot}}(\mathbf{G})}{\mathbf{G}^2} \quad . \quad (2.113)$$

It is obvious that the used description makes the solution of the Coulomb contribution trivial as it is only determined by a sum. The computationally limiting step is, therefore, the fast Fourier transformation (FFT) resulting in a scaling behavior $\propto N \log(N)$ with the system size.

2.11.4 Exchange-Correlation Term

The last remaining problem term in the GPW energy expression is the exchange correlation term. The approach to simplify I_{XC} uses similar methods as previously introduced for the Coulomb term, as it underlies PBCs, a simplification in analogy to Eq (2.106) by using

Parseval's theorem is possible. Therefore, it can be rewritten as

$$E_{xc} = \int d\mathbf{r} \epsilon_{xc}(\mathbf{r}) \rho_{tot}(\mathbf{r}) d\mathbf{r} \stackrel{\text{Parseval}}{=} \Omega \sum_G \epsilon_{xc}(\mathbf{G}) \rho_{tot}^*(\mathbf{G}) \quad . \quad (2.114)$$

However, ϵ_{xc} is not local in \mathbf{G} -space, therefore, an alternative to Eq (2.114) has to be found. A possible solution employed in CP2K is the redefinition of the functional in \mathbf{G} space, which is achieved using a plane wave expansion of the exchange correlation functional and results in the expression

$$E_{xc} = \frac{\Omega}{N_x N_y N_z} \sum_R \tilde{\epsilon}_{xc}(\mathbf{R}) \rho_{tot}(\mathbf{R}) = \Omega \sum_G \tilde{\epsilon}_{xc}(\mathbf{G}) \rho_{tot}(\mathbf{G}) \quad . \quad (2.115)$$

As a consequence of the redefinition the so called ripple effect arises, which was observed at multiple occasions throughout this thesis. Due to the exchange energy no longer being translation invariant, translations by a multiple of the integration grids spacing leave the total energy invariant. Quote: J. Hutter "It's like optimizing a molecule sitting on an orange peel"⁶².

2.12 Non-Local Corrections to Density Functional Theory

KS-DFT as a theoretical framework, is a localized theory meaning that the electron density does only depend on the immediate surrounding of the investigated point in space. This treatment omits three effects of interest for quantum chemical investigations, namely dispersion effects as observed in halogen bonding, π - π interactions and hydrogen bridges. All of these interactions are badly described by GGA functionals alone, with most of them overestimating repulsive interactions¹¹¹. In this thesis all of these effects play a crucial role for the investigation of interlayer interactions in covalent organic frameworks and will be discussed in Section 10.4.4. For a better description of the underlying phenomena, it is suitable to correct the energy obtained by an additional term, explicitly considering dispersion corrections to the used functional. When assuming that other non-local phenomena, e.g. quadrupole moments, are already considered in the base functional, the total energy E_{tot} can be determined as

$$E_{tot} = E_{base} + E_{disp} \quad , \quad (2.116)$$

2 Density Functional Theory

where the energy correction E_{disp} is simply added upon the energy E_{base} , obtained by the functional alone.

2.12.1 Pairwise Dispersion Correction

The most prominent way of including dispersion effects is the consideration of atom pair wise interactions with the scheme

$$E_{\text{disp}} = -\frac{C_6^{\text{AB}}}{R_{\text{AB}}^6}, \quad (2.117)$$

where $R_{\text{AB}} = |\mathbf{R}_A - \mathbf{R}_B|$ and C_6 is the so called dispersion-coefficient. In the simplest case it can be determined via London's formula

$$C_6^{\text{AB}} = \frac{3}{2} \cdot \frac{\alpha_A I_A \alpha_B I_B}{I_A + I_B}, \quad (2.118)$$

where α is the polarizability of the particular atom and I its ionization potential¹¹². These coefficients depend directly on the polarizability, which itself depends on the size of the atoms as well as the number of valence electrons¹¹³. Due to its limited practical applicability, more sophisticated methods can be used as for example from second order perturbation theory. Here, excited states and respective dipole transitions are explicitly treated and the dispersion coefficients are determined as

$$C_6^{\text{B}} = \frac{2}{3} \sum_{n_A n_B} \frac{\langle 0 | \boldsymbol{\mu}_A | n_A \rangle \langle n_A | \boldsymbol{\mu}_A | 0 \rangle \langle 0 | \boldsymbol{\mu}_B | n_B \rangle \langle n_B | \boldsymbol{\mu}_B | 0 \rangle}{\mathcal{E}_{n_A} + \mathcal{E}_{n_B}}. \quad (2.119)$$

The sums for $n_{A,B}$ run over the excited states of atoms A and B respectively, with the associated excitation energy $\mathcal{E}_{n_A} = E_{n_A} - E_0$ for state n_A . Density fluctuations in an investigated atom in the time independent case, can only be explained by a change of electronic states. Therefore, these effects have to be associated with the probabilities of state switching, as they change the density distribution and, thus, potentially influence a reference pair atom. Generally, the dispersion correction has to account for more than just dipole interactions and can be split in

2.13 Non-Local Exchange Correlation Functionals

multi-pole parts including their respective coefficients in a series

$$E_{\text{disp}} = -\frac{C_6^{\text{AB}}}{R_{\text{AB}}^6} - \frac{C_8^{\text{AB}}}{R_{\text{AB}}^8} - \frac{C_{10}^{\text{AB}}}{R_{\text{AB}}^{10}} - \dots \quad (2.120)$$

With the coefficients C_8^{AB} describing the quadrupole and C_{10}^{AB} the octopole term respectively. The coefficients are obtained by the Casimir-Polder formula following

$$C_{2n}^{\text{AB}} = \sum_{l=1}^{n-2} \frac{(2n-2)!}{2\pi(2l)!(2n-2l-2)!} \int_0^\infty \alpha_l^{\text{A}}(i\omega) \alpha_{n-l-1}^{\text{B}}(i\omega) d\omega \quad , \quad (2.121)$$

considering the frequency dependence of the polarizabilities $\alpha(i\omega)$ ¹¹¹. It has to be noted, that the frequency dependence is not directly accessible in ground state DFT, but can be calculated explicitly by other methods¹¹⁴. General contributions of the van-der-Waals corrections to the field of DFT can be found in the literature, e.g.¹¹⁵.

2.13 Non-Local Exchange Correlation Functionals

Another approach to account for the lack of non-local contributions in DFT-functionals is given by directly correcting the base functional. Using this approach, the correction is no longer added upon the total energy of the base system as shown in Eq (2.116), but rather modifies the correlation energy of a functional in the form

$$E_c = E_c^{\text{LDA/GGA}} + E_c^{\text{nl}} \quad . \quad (2.122)$$

The introduction of non-local (NL) correlations E_c^{nl} was pioneered by Dion et. al.¹¹⁶ and is given as

$$E_c^{\text{nl}} = \frac{\hbar}{2} \int d^3\mathbf{r} \int d^3\mathbf{r}' \rho(\mathbf{r}) K(\mathbf{r}, \mathbf{r}') \rho(\mathbf{r}') \quad , \quad (2.123)$$

where the kernel $K(\mathbf{r}, \mathbf{r}')$ is chosen to be symmetric for \mathbf{r} and \mathbf{r}' and defined as

$$\lim_{R \rightarrow \infty} : K = -\frac{3e^4}{2m^2 \omega_0(\mathbf{r}) \omega_0(\mathbf{r}')} [\omega_0(\mathbf{r}) + \omega_0(\mathbf{r}')] |\mathbf{r} - \mathbf{r}'|^6 \quad . \quad (2.124)$$

2 Density Functional Theory

Here, the limes is introduced for the variable $R = |\mathbf{r} - \mathbf{r}'|$. From the original idea of Dion, multiple modifications were considered differing in the utilized kernel $K(\mathbf{r}, \mathbf{r}')$. The most important functional for this thesis is the rVV10 modified specifically for the application with periodic boundary conditions¹¹⁷. Drawbacks of the NL functionals are based on the double integration in real space and the interplay with the different base functionals. For NL functionals, the long range limits are known however the overlap region is not, leading to potential double counting which reduces their predictive power. E_{xc}^{rVV10} is given as

$$E_{xc}^{rVV10} = E_x^{rPW86} + E_c^{LDA} + E_{c-nl}^{rVV10} \quad , \quad (2.125)$$

where E_x^{rPW86} corresponds to the refitted Perdew-Wang exchange functional. The energy contribution E_{c-nl}^{rVV10} of the rVV10 functional follows Eq (2.123) with the kernel

$$K^{rVV10}(\mathbf{r}, \mathbf{r}') = -\frac{3e^4}{2m^2} \cdot \frac{1}{(qR^2 + 1)(q'R^2 + 1)(qR^2 + q'R^2 + 2)} \quad . \quad (2.126)$$

Here, the auxiliary function $q(\mathbf{r}) = \omega_0(\rho(\mathbf{r}), |\nabla\rho(\mathbf{r})|)/k(\rho(\mathbf{r}))$ is introduced in analogous manner for \mathbf{r}' . The factor k is defined in analogy to the VV10 kernel as $k(\mathbf{r}) = 3b(\omega_p/k_s^2) = 3\pi b \left(\frac{\rho(\mathbf{r})}{9\pi}\right)^{\frac{1}{6}}$ where k_s is the Thomas-Fermi screening wave vector, b is an empirically determined parameter and $\omega_p = \frac{4\pi\rho(\mathbf{r})e^2}{m}$ is the plasma frequency¹¹⁸.

Chapter 3

Polarization and Electric Fields in Periodic Systems

Molecular interactions with external electric fields are determined by the system's polarization and the field itself. In this chapter, the basic concepts of both properties are briefly revisited at the example of gas-phase systems. Afterwards the problems arising from the generalization to the solid state are shown at a one dimensional example system. In the following, the main ideas leading to the modern theory of polarization (MTP) are explained, with the goal of illustrating the meaning of the expressions entering the polarization's definition for extended systems. Subsequently, the treatment of homogeneous electric fields in periodic systems are discussed putting emphasis on the most common problems and how they can be circumvented. With the provided overview of the theoretical framework, the application of MTP onto DFT-problems will be explained following the review provided by R. Resta in 2020¹¹⁹. The chapter closes with an overview of field related effects on molecular systems, as they are later observed in simulations of field evaporation.

3.1 Interaction with External Electric Fields

For consideration of a molecular system's interaction with external electric fields, the molecular Hamiltonian in the absence of fields $\hat{\mathbf{H}}_0$, does no longer provide a sufficient description. To account for the additional interaction with an external field, a term $\hat{\mathbf{H}}_{\text{ext}}$ is introduced. The

resulting total Hamiltonian $\hat{\mathbf{H}}_{\text{tot}}$, is given as the sum of both contributions

$$\hat{\mathbf{H}}_{\text{tot}} = \hat{\mathbf{H}}_0 + \hat{\mathbf{H}}_{\text{ext}} \quad . \quad (3.1)$$

The field interaction term in Eq (3.1), depends on the atomic charge distribution and the applied electric field. In the macroscopic case a description is provided by the Clausius Mossotti (CM) picture^{120,121}, which describes the interaction as the superposition of localized electric dipole contributions. In general, the dipole moment of N point charges with value q_i , is defined as their sum, weighted by their relative positions \mathbf{r}_i with respect to the center of charges

$$\mathbf{d} = \sum_{i=1}^N q_i \mathbf{r}_i \quad . \quad (3.2)$$

Suppose a DFT-problem, consisting of N paired electrons and M ionic nuclei with charges Z_A at positions \mathbf{R}_A . The total dipole moment can then be expressed as the sum of nucleic \mathbf{d}_{nuc} and electronic \mathbf{d}_{el} contributions

$$\mathbf{d}_{\text{tot}} = \mathbf{d}_n + \mathbf{d}_{\text{el}} = \sum_{A=1}^M Z_A \mathbf{R}_A - 2e \sum_{i=1}^{N/2} \langle \Phi_i | \hat{\mathbf{r}}_i | \Phi_i \rangle \quad , \quad (3.3)$$

where $\hat{\mathbf{r}}_i$ denotes the electronic position operator, e the elementary charge and Φ_i the corresponding KS-orbitals for the electron i . The sum in Eq (3.3) runs over i to $N/2$, as every orbital Φ_i is doubly occupied. Using open boundary conditions (OBC), the wave functions in Eq (3.3) vanish at infinity and the dipole moment of the investigated sample, can be obtained¹¹⁹. When an electric field is applied, the basic interaction of dipoles, is a reorientation that is described by the coupling of a system's total polarization \mathbf{P} with the applied field \mathcal{E} (note that in the following, the term field refers to the electric field vector). In the Hamiltonian Eq (3.1), the additional external coupling to the field $\hat{\mathbf{H}}_{\text{ext}}$ is provided in the molecular case as

$$\hat{\mathbf{H}}_{\text{ext}} = \mathbf{P} \mathcal{E} = \mathbf{d}_{\text{tot}} \mathcal{E} = (\mathbf{d}_n - 2e \sum_{i=1}^{N/2} \hat{\mathbf{r}}_i) \mathcal{E}. \quad (3.4)$$

As a result, the eigenfunctions of the original Hamiltonian $\hat{\mathbf{H}}_0$ are no longer eigenfunctions of $\hat{\mathbf{H}}_{\text{tot}}$, and thus new ones have to be calculated to find the new ground state. Consequently,

the charge distribution is shifted, which in turn corresponds to a change in the distribution of molecular orbitals for determination of the new ground state energy as well as a possible structural rearrangement of the new minimum energy structure. In the molecular case this treatment leads to good results and is applied to determine molecular response to electric fields, as well as net polarization changes in various systems.

However, in the case of bulk materials, the unique determination of \mathbf{P} becomes a lot more difficult and the introduction of additional concepts is necessary as will be illustrated in Section 3.2. Bulk description is usually based on a unit cell treatment with Born van Karman (BvK) boundary conditions as described later in Eq (3.9). A naive approach for this case, provided by the CM model, defines the polarization \mathbf{P} of a bulk sample as the combined dipole moment \mathbf{d} per unit cell volume V . For a continuous charge distribution $\rho(\mathbf{r})$, this leads to

$$\mathbf{P} = \frac{\mathbf{d}}{V} = \frac{1}{V} \int_V \mathbf{r} \rho(\mathbf{r}) d\mathbf{r} \quad . \quad (3.5)$$

Such a definition for the polarization is not applicable for the periodic case as cell boundaries are not uniquely defined, and the determination of the net polarization, depends on the choice of reference^{122–124}. This leads to the formulation only being applicable to systems where the crystal structure is providing the charge locations uniquely, which is only the case for continuous distributed crystals, an example can be found in Section B.1.1. Furthermore, for a uniformly distributed charge density, the average of the bulk charge is zero. This leaves the polarization to be a surface property, however, it is intrinsically a bulk property.

3.1.1 The Polarization Problem

In order to illustrate the previously stated problem of non-uniqueness, the fundamental flaws in the assumptions entering Eq (3.5) will be shown at the example of a one dimensional insulator-chain. Consider a system consisting of charges with $q_+ = e$ and $q_- = -e$ on a one dimensional lattice with unit cell length L . The one dimensional polarization following

Eq (3.5), is then given as

$$P_x = \frac{1}{L} \int_0^L x\rho(x)dx \quad . \quad (3.6)$$

The lattice may consist of a two body unit cell as shown in Figure 3.1 and as the system is neutrally charged, the net polarization should vanish. A shift of the unit-cell boundaries, leads to different positions of the charges and thus influences the result of Eq (3.6) as explained by Spaldin¹²⁵. When calculating P for two different charge distributions on the lattice, with their distance within the unit cell given as d , the resulting expressions are given as

$$\rho_a(\mathbf{r}) = e\delta\left(\frac{L}{2} + \frac{d}{2} - \mathbf{r}\right) - e\delta\left(\frac{L}{2} - \frac{d}{2} - \mathbf{r}\right) \quad , \quad (3.7)$$

as illustrated in Figure 3.1a and

$$\rho_b(\mathbf{r}) = e\delta\left(L - \frac{d}{2} - \mathbf{r}\right) - e\delta\left(0 + \frac{d}{2} - \mathbf{r}\right) \quad , \quad (3.8)$$

as illustrated in Figure 3.1b respectively, with δ being the Dirac delta distribution. The obtained values for the polarization P if $L = 1$ and $d = \frac{1}{4}$ is chosen, are $P_a = -\frac{1}{4}e$ and $P_b = -\frac{3}{4}e$ which differ, although only the position of cell reference is changed. This indeterminacy can not be circumvented, unless the cell-length is chosen to be $L = 2d$ which corresponds to the uniform crystal. (For a detailed calculation see Section B.1.1)

3.2 Modern Theory of Polarization

As shown in the previous section, a treatment of the bulk polarization in the CM picture is not suitable due to non-uniqueness. The main reason for this, is that the position operator is not translational invariant¹²⁶, which makes it unsuited for treatment with periodic boundary conditions. In the following, a structural approach to solve the previously discovered problems in the definition of the polarization for bulk materials as it was historically discovered in the 90's by Resta, Vanderbilt, King-Smith and others will be presented. The review of the early developments, will result in the framework of what is known in the literature as modern

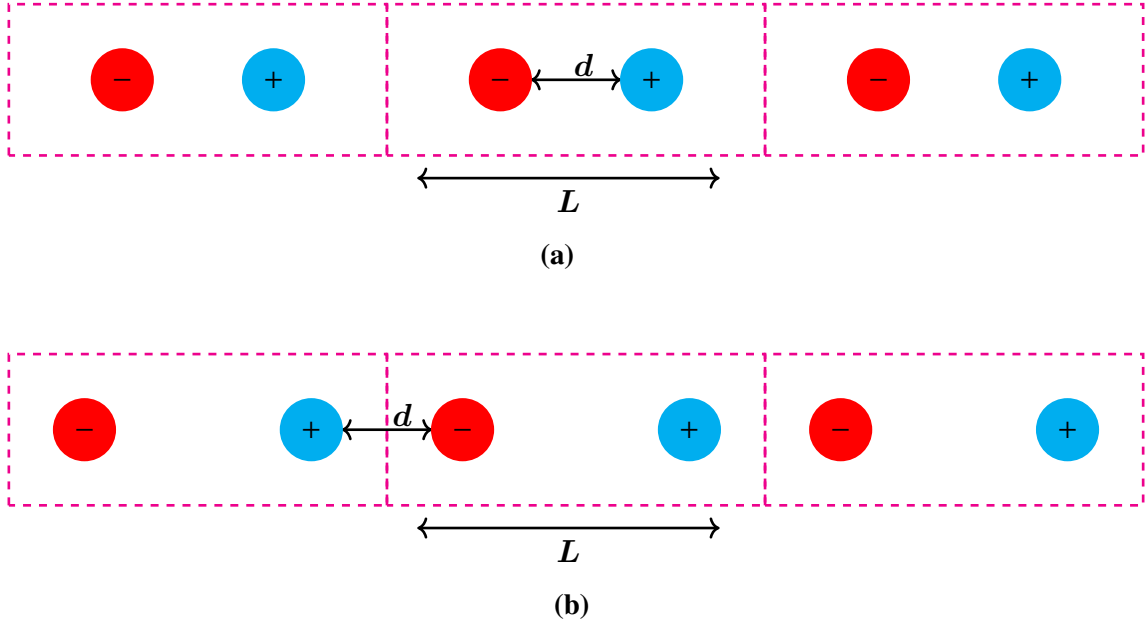


Figure 3.1: Schematic representation of two lattice arrangements for a neutral chain with charges placed at distance d from each other in a unit cell with cell-vector L . a) The dipole placed in the center of the unit cell. b) The dipole placed over the cell boundaries. Using Eq (3.6), leads to two different values for the polarization per unit cell volume.

theory of polarization (MTP). In the bulk description, the wave function has to fulfill the BvK-boundary conditions for every coordinate given as

$$\Psi = \Psi(\mathbf{r}_1, \dots, \mathbf{r}_i, \dots, \mathbf{r}_N) = \Psi(\mathbf{r}_1, \dots, \mathbf{r}_i + \mathbf{a}_j \cdot \mathbf{m}_j, \dots, \mathbf{r}_N) \quad , \quad (3.9)$$

where the \mathbf{a}_j are the primitive lattice vectors of the systems Bravais-lattice and m_j integers. As already stated, the position operator has to be re-expressed in order to suite the cyclic topology of a periodic system. A mathematical argument to illustrate the problem in one dimension, is given by analyzing the behavior of the position operator $\hat{\mathbf{X}} = \sum_{i=1}^N \hat{\mathbf{x}}_i$ for an N -electron system in one dimension, with a unit cell of length L . For a short explanation of the problems underlying the description and its results, see Section B.1.

3.2.1 Berry-Phase-Approach and King-Smith-Vanderbilt Formula

In order to compensate for the problems of defining the polarization in a unique way for periodic systems, fundamental formulations were derived in the 1990's, with the main idea of investigating polarization changes under the assumption that these are performed adiabatically¹²⁷. This can intuitively be understood, as the change of polarization $\Delta\mathbf{P}$ is also the property that is measurable in experiments¹²². The electronic contribution to the change of polarization can be expressed as the electronic current integrated over the process-time, leading to

$$\Delta\mathbf{P} = \int_0^{\Delta t} \frac{d\mathbf{P}}{dt} dt = \mathbf{P}(\Delta t) - \mathbf{P}(0) \quad . \quad (3.10)$$

Here, the adiabatic process is assumed over a time interval $[0, \Delta t]$, which leaves the total change of polarization in form of a two point formula. It follows, that the quantum mechanical current for electrons \mathbf{j}_{el} , given as

$$\mathbf{j}_{\text{el}} = \frac{d\mathbf{P}}{dt} = \frac{e}{2\pi} \text{Im}(\langle \Psi_i | \hat{\mathbf{H}} | \Psi_i \rangle + \langle \Psi_i | \hat{\mathbf{H}} | \Psi_i \rangle) \quad , \quad (3.11)$$

has to be calculated in order to obtain information about \mathbf{P} . For systems within PBCs, a possible Hamiltonian following the periodic gauge, as first proposed by Kohn¹²⁸ is given as

$$\hat{\mathbf{H}}_{\mathbf{k}} = \frac{(\hat{\mathbf{p}} + \hbar\mathbf{k})^2}{2m} + V(\mathbf{r}) \quad . \quad (3.12)$$

Here, the newly introduced modification to the momentum depends on \mathbf{k} , which was originally denoted as flux or twist¹¹⁹. This add on corresponds to a gauge transformation and leaves the physics of the system untouched, derived in the original paper by Kohn. An expression for the total electron current can be calculated by means of adiabatic perturbation theory¹²⁹. Suppose an adiabatic change of state acting on the system of interest also leads to a change in polarization (In analogy to Eq (3.10)), given as

$$\Delta\mathbf{P} = \int_0^1 \frac{d\mathbf{P}}{d\lambda} d\lambda \quad . \quad (3.13)$$

Then all quantities in Eq (3.12) depend implicitly on a parameter λ that is varied during the adiabatic process. If the change of λ is small in time, the wave function acquires a first order perturbation given as

$$|\delta\Psi_n, \mathbf{k}\rangle = -i\hbar\dot{\lambda} \frac{\langle\Psi_{m,\mathbf{k}}|\partial_\lambda\Psi_{n,\mathbf{k}}\rangle}{E_{n,\mathbf{k}} - E_{m,\mathbf{k}}} |\Psi_{m,\mathbf{k}}\rangle \quad , \quad (3.14)$$

wherein $\dot{\lambda}$ corresponds to $\frac{d\lambda}{dt}$ and ∂_λ is the partial derivative with respect to the parameter λ . From here, it follows in the first order perturbative treatment when averaging over the BZ, that the expression for the current arising from orbital n is given as

$$\mathbf{j}_n = \frac{d\mathbf{P}}{dt} = \frac{i\hbar e\dot{\lambda}}{(2\pi)^3 m} \sum_{m \neq n} \int_{\text{BZ}} \left(\frac{\langle\Psi_{n,\mathbf{k}}|\hat{\mathbf{p}}|\Psi_{m,\mathbf{k}}\rangle \langle\Psi_{n,\mathbf{k}}|\partial_\lambda\Psi_{m,\mathbf{k}}\rangle}{E_{n,\mathbf{k}} - E_{m,\mathbf{k}}} + \text{c.c.} \right) d\mathbf{k} \quad , \quad (3.15)$$

where c.c. in the sum, marks the complex conjugated of the previous expression. Here, the time derivative can be eliminated by expanding the fraction on the left hand side of Eq (3.15) using $\frac{d\mathbf{P}}{d\lambda} \frac{d\lambda}{dt}$. After utilizing ordinary perturbation theory in \mathbf{k} for the expression Eq (3.12), one obtains after some manipulations

$$\frac{d\mathbf{P}_n}{d\lambda} = \frac{ie}{(2\pi)^3} \int \langle\nabla\mathbf{k}u_{n,\mathbf{k}}|\partial_\lambda u_{n,\mathbf{k}}\rangle + \text{c.c.} d\mathbf{k} \quad , \quad (3.16)$$

as the contribution of band n to the total change of the electronic contribution to the systems polarization. Here, $u_{n,\mathbf{k}}$ is the Bloch factor corresponding to the periodic wave function Ψ_n . In applications, the formula in Eq (3.16) has to be evaluated by integration and results in a two point formula given as

$$\Delta\mathbf{P} = \mathbf{P}(1) - \mathbf{P}(0) \quad . \quad (3.17)$$

Reformulation of the expression and summation over all bands leads to a formally very elegant expression in terms of a Berry-phase¹²⁷, named after their developers the King-Smith-Vanderbilt formula, given as

$$\mathbf{P}_{\text{el}}(\lambda) = -i \frac{2e}{(2\pi)^3} \sum_i \int_{\text{BZ}} \langle u_{i,\mathbf{k}}|\nabla_{\mathbf{k}}|u_{i,\mathbf{k}}\rangle d\mathbf{k} \quad . \quad (3.18)$$

Here, the integrand can be identified as the sum over Berry connections of type Eq (B.20) and as a result, the total polarization can be identified in terms of a Berry-phase. This, leads to the commonly used name of Berry-phase approach in literature when referring to the MTP. A derivation of the explicit formula for the Berry-phase, obtained by adiabatic change of state, following the original approach of Berry¹³⁰ can be found in Section B.2. Generally this formulation can be interpreted as the change of movement direction along a closed path, which in this case corresponds to the total phase of the wave function, gathered by crossing the unit cell. Here, it has to be noted that in general, the integral over the complete Brillouin zone in Eq (3.18), has to be evaluated following a k-point sampling scheme as described in Section 2.9. For applications within this thesis, Γ -point sampling was used. In the following the expression for the single k -point case will be further investigated to put emphasis on the features of the polarization's new definition and the way it is implemented in quantum chemistry codes¹²⁵.

3.2.2 The Position Operator in Extended Systems

In 1998 a new formulation for the expectation value of the position operator in extended systems was introduced by Resta¹²⁶, in order to find an alternative way of facilitating the derivation of what was previously known as the Berry-phase formulation Section 3.2.1. This new formulation of the position operator makes it applicable to periodic systems and marks the first step of the definition of the polarization in bulk systems, by directly extending Eq (3.5). It has to be noted, that the formulation provided in this section, leads to the same equations as originally thought of by King-Smith and Vanderbilt Eq (3.18), however, the derivation is a lot more intuitive¹²⁷. The following treatment is explained exemplary for one dimensional systems on a lattice with length L . Based on the generalization of the position operator for many-body systems, it is straight forward to generalize \hat{x} for N particles as

$$\hat{X} = \sum_{i=1}^N \hat{x}_i \quad . \quad (3.19)$$

Using this expression to calculate the position expectation value for the ground state density results in

$$\langle \hat{X} \rangle = \langle \Psi_0 | \hat{X} | \Psi_0 \rangle = \int x \rho(x) dx \quad , \quad (3.20)$$

where $|\Psi_0\rangle$ is the ground state wave function, given as a Slater determinant of the Bloch orbitals within the KS-DFT approach. However, within PBCs this formulation is not applicable due to non-uniqueness and an alternative formulation for the position operator has to be found. A suitable expression was found by Resta¹²⁶. It compensates for the lack of uniqueness by defining the expectation value Eq (3.20) modulo the unit cell. The main observation here is the fact, that due to the cyclic topology of the problem considering PBCs, the position can be expressed via angles $\theta_i = \frac{2\pi x_i}{L}$. This can be intuitively understood as moving along a circle, where the initial and final position correspond to the cell boundaries. The expectation value of the position operator for periodic systems can therefore be written as

$$\langle \hat{X} \rangle = \text{Im} \left\{ \ln \left(\langle \Psi_0 | e^{i \sum_{i=1}^N \hat{\theta}_i} | \Psi_0 \rangle \right) \right\} \quad . \quad (3.21)$$

Eq (3.21) only leads a contribution, if the angle hits the position of the corresponding electrons. This can be easily understood when using localized Wannier-functions as basis for the expansion of Ψ_0 . With this formulation, the expression for the polarization can be obtained by substituting Eq (3.21) in the expression for the total dipole moment Eq (3.2), leading to the systems electronic contribution to the polarization given as

$$\mathbf{P}_{\text{el}} = -2 \frac{eL}{\pi} \text{Im} \left[\ln \left(\langle \Psi_0 | e^{i \sum_{i=1}^N \hat{\theta}_i} | \Psi_0 \rangle \right) \right] = 2 \frac{eL}{\pi} \gamma_{\text{el}} \quad . \quad (3.22)$$

Here, the factor of two accounts for doubly occupied orbitals, again leading to charge neutrality in the full system. With this formulation, the polarization is no longer dependent on the choice of the cell boundary. The total polarization for a system of N paired electrons, is then constructed as the sum of the nucleic and electronic contributions respectively as

$$\mathbf{P}_{\text{tot}} = \mathbf{P}_{\text{nuc}} + \mathbf{P}_{\text{el}} = \mathbf{P}_{\text{nuc}} - 2 \frac{e}{L} \text{Im} \left[\ln \left(\langle \Psi_0 | e^{i \sum_{i=1}^{N/2} \hat{\theta}_i} | \Psi_0 \rangle \right) \right] \quad . \quad (3.23)$$

The expectation value of the position operator, can thus be interpreted as a Berry-phase in disguise¹¹⁹, resulting in the previously introduced King-Smith-Vanderbilt formulation when generalizing to three dimensions. The same procedure can be applied for the nucleic part, where it is obvious, that the electronic wave function will commute with the atomic position operators $\hat{\mathbf{X}}$. The resulting nucleic part, can therefore be written as

$$\mathbf{P}_{\text{nuc}} = \frac{eL}{2\pi} \text{Im}\left\{\ln\left(e^{i\frac{2\pi}{L}\sum_A Z_A X_A}\right)\right\} = \frac{eL}{2\pi}\gamma_{\text{nuc}} \quad , \quad (3.24)$$

as a result both parts can be combined into a single Berry-phase, given as

$$\frac{e}{2\pi}\gamma_{\text{tot}} = \frac{e}{2\pi}(\gamma_{\text{nuc}} + 2\gamma_{\text{el}}) = \frac{e}{2\pi} \text{Im}\left\{\ln\left(\langle\Psi_0| e^{i\frac{2\pi}{L}(\sum_A Z_A \hat{X}_A - 2\sum_i \hat{x}_i)} |\Psi_0\rangle\right)\right\} \quad . \quad (3.25)$$

This new defined total phase, is invariant under translation of the coordinate origin, whereas both independent parts are not. Another feature of the total description is that it conserves charge neutrality within a unit cell. In common implementations, the expectation value is further simplified by identifying

$$|\tilde{\Psi}_0\rangle = e^{i\frac{2\pi x}{L}} |\Psi_0\rangle \quad , \quad (3.26)$$

as a second Slater determinant. The determinant $|\Psi_0\rangle$ is constructed according to

$$|\Psi_0\rangle = \frac{1}{\sqrt{M^N}} \mathbf{A} \prod_{m=1}^{N/M} \prod_{s=0}^{M-1} \Psi_{m,k_s} \quad , \quad (3.27)$$

where \mathbf{A} is the antisymmetrization operator. The basis functions are chosen according to M equally spaced Bloch vectors $k_s = \frac{2\pi}{Ma}s$, which are normalized to a crystal length a , whereas in comparison, the total system length is given as $L = Ma$. Here, it has to be noted that the full determinant is normalized to the whole system's size giving rise to the prefactor $\frac{1}{\sqrt{M^N}}$. As the overlap of two determinants is equal to the determinant of the overlap matrix among the orbitals. Introducing the overlap matrix S and utilizing Eq (3.26) leads to a significant

simplification of Eq (3.23) resulting in

$$\gamma_{\text{el}} = -\text{Im} \left(\ln \left(\prod_{s=0}^{M-1} \det(S(k_s, k_{s+1})) \right) \right) . \quad (3.28)$$

Here, the fact that the determinants are constructed by M basis functions build from orthogonal Bloch orbitals leads to vanishing elements except when the condition $k_{s'} = k_s + \frac{2\pi}{L}$ is fulfilled, which corresponds to 'adjacent' functions with $s' = s + 1$. For a specific band, the contribution to γ_{el} is further simplified using

$$S(k_s, k_{s+1}) = \langle u_{k_s} | u_{k_{s+1}} \rangle , \quad (3.29)$$

where the u_k are the Bloch factors of the corresponding orbitals. Using the laws for reformulation of the logarithm yields

$$\gamma_{\text{el}} = - \sum_{s=0}^{M-1} \text{Im} \left(\ln \langle u_{k_s} | u_{k_{s+1}} \rangle \right) . \quad (3.30)$$

Finally, the connection to the original formula Eq (3.18), can be established by utilizing

$$- \text{Im} \left(\ln \langle u_k | U_{k+\Delta k} \rangle \right) \cong i \left\langle u_k \left| \frac{du}{dk} \right. \right\rangle \Delta k , \quad (3.31)$$

which is applicable, if u is a differentiable function of k and a discrete spacing (as was introduced in the beginning when constructing the determinants), is present. Generalizing the obtained formula by assuming $M \rightarrow \infty$ and rewriting the sum as an integral over the unit cell, yields

$$\gamma_{\text{el}} = i \int_0^{2\pi/a} \left\langle u_k \left| \frac{du}{dk} \right. \right\rangle dk . \quad (3.32)$$

This integral is a one dimensional Berry-phase as first investigated by Zak¹³¹. The whole electronic contribution is then obtained by summing over all bands and, in three dimensions, yields the previously mentioned King-Smith and Vanderbilt formula Eq (3.18).

3.3 Wannier-Functions

In order to circumvent the delocalization problem while calculating polarization differences, an alternative to the basic description in terms of Bloch-functions can be utilized in order to illustrate the basic concepts in a different way. Consider eigenfunctions to a Hamiltonian with lattice periodicity \mathbf{R} and Bloch functions of the form

$$\Psi_{n,\mathbf{k}}(\mathbf{r}) = u_{n,\mathbf{k}}(\mathbf{r})e^{i\mathbf{k}\mathbf{r}} \quad , \quad (3.33)$$

where the corresponding function $u_{n,\mathbf{k}}(\mathbf{r})$ is the Bloch-factor. However, for every \mathbf{k} a different function is constructed and due to the orthogonality of the wave functions it is possible to construct wave functions by a superposition of Bloch functions⁸⁷. One special way of constructing new wave functions is to superimpose all Bloch functions in the Brillouin zone via

$$w_0(\mathbf{r}) = A \int_{\text{BZ}} \Psi_{n,\mathbf{k}}(\mathbf{r})d\mathbf{k} \quad , \quad (3.34)$$

with a normalization constant A, resulting in a Wannier function. As a result of the periodicity in reciprocal space, the Wannier function is the same for every Brillouin-Zone. The natural choice is to normalize the integral over the Brillouin zone via

$$A = \frac{\Omega}{(2\pi)^3} \quad , \quad (3.35)$$

where Ω is the volume of the real space unit-cell. A phase factor can be introduced in the Wannier function in order to construct a new set corresponding to a Fourier-transform of the Bloch functions

$$w_n(\mathbf{r}) = \frac{\Omega}{(2\pi)^3} \int_{\text{BZ}} \Psi_{n,\mathbf{k}}(\mathbf{r})e^{-i\mathbf{k}\mathbf{R}}d\mathbf{k} = |\mathbf{R}n\rangle \quad . \quad (3.36)$$

Here, the index \mathbf{R} marks the reference cell where the Wannier function is located. As a result of the introduced phase $e^{i\mathbf{k}\mathbf{R}}$, the Bloch functions can be recalculated from the Wannier functions by an inverse Fourier transform. The functions of form Eq (3.36) are connected to

Bloch functions of a single energy band $E_n(\mathbf{k})$. In a more general approach, multiple bands are considered, resulting in a combined expression for the underlying Bloch functions as

$$|\tilde{\Psi}_{n,\mathbf{k}}(\mathbf{r})\rangle = \sum_{m=1}^J U_{mn}^{\mathbf{k}} |\Psi_{m,\mathbf{k}}\rangle \quad , \quad (3.37)$$

where $U_{mn}^{\mathbf{k}}$ is any unitary matrix of dimension J which is periodic in \mathbf{k} . With this considerations, a more general set of possible Wannier functions can be constructed as

$$|\mathbf{R}n\rangle = \frac{\Omega}{(2\pi)^3} \int_{\text{BZ}} d\mathbf{k} e^{-i\mathbf{k}\mathbf{R}} \sum_{m=1}^J U_{mn}^{\mathbf{k}} |\Psi_{m,\mathbf{k}}\rangle \quad . \quad (3.38)$$

If bands overlap at some points in the \mathbf{k} -space, the resulting Wannier functions are no longer localized⁸⁷. In order to find the maximally localized subset of possible Wannier functions, the Bloch orbitals are chosen to minimize a localization function

$$D = \sum_n \left(\langle \mathbf{0}n | \mathbf{r}^2 | \mathbf{0}n \rangle - \langle \mathbf{0}n | \mathbf{r} | \mathbf{0}n \rangle^2 \right) \quad , \quad (3.39)$$

by optimizing the chosen $U_{mn}^{\mathbf{k}}$ entering Eq (3.37). The function D measures the spread of the total set of J Wannier functions in the original unit cell ($\mathbf{R} = 0$) around their respective centers. The minimization can be performed using any optimization algorithm whereas in an explicit example by Vanderbilt, a steepest-decent method was used¹³². The resulting maximally localized Wannier functions can then be used in order to determine analogs to the molecular orbitals in gas phase calculations. Another insightful application for the MTP is found, when using localized Wannier functions for the calculation of the electronic contributions to Eq (3.23).

Wannier-Functions for Calculations of the Polarization

Due to the delocalized character, the polarization as defined in Eq (3.23) is not gauche invariant and thus depends on the choice of functions when Bloch orbitals are used. However, when using maximally localized Wannier functions, the orbitals of a respective unit cell can be reconstructed and allow for the representation of charge distributions. These can be obtained

by calculating the gauche invariant Wannier centers, defined as

$$\langle \bar{r}_n \rangle = \langle \mathbf{R}n | \hat{\mathbf{r}} | \mathbf{R}n \rangle \quad . \quad (3.40)$$

This expression, can be evaluated without using the Wannier functions explicitly when switching into \mathbf{k} -space representation. The corresponding representation as derived by Blount¹³³, is given as follows

$$\langle \bar{r}_n \rangle = i \frac{\Omega}{(2\pi)^3} \int_{\text{BZ}} \langle u_{n,\mathbf{k}} | \nabla_{\mathbf{k}} | u_{n,\mathbf{k}} \rangle d\mathbf{k} \quad , \quad (3.41)$$

which can be identified with the Berry-phase formulation of the electronic contribution in Eq (3.18). It is therefore possible, to use the Wannier centers in order to arrive at a formulation that connects the former Eq (3.5) in the solid state. As a result, the total electronic contribution to the polarization in terms of maximally localized Wannier functions corresponding to one unit cell (here $\mathbf{R} = 0$) can be expressed as

$$P_{el} = -\frac{e}{\Omega} \sum_{n=1}^{n_b} \langle \bar{r}_{n,0} \rangle \quad . \quad (3.42)$$

Where the sum over n runs over all n_b bands, contributing to the polarization within the first unit cell. Interpretation of the Wannier functions as localized electronic orbitals in the reference cell, leads to an equivalent picture as the original CM formula suggests. The obtained spatial distributions can then be used to calculate the Wannier centers which can be interpreted as the center of electronic charge.

3.4 Polarization Quantum

It follows from Eq (3.23), that the overall polarization is related to a phase angle. As a result, the two point formulas for polarization differences ΔP given in Eq (3.10) are only defined modulo a polarization quantum¹²². This polarization quantum of the electronic contribution, as stemming from the King-Smith and Vanderbilt formula Eq (3.18), is given for the 3d case

as

$$p_n = 2 \frac{eR}{\Omega} \quad , \quad (3.43)$$

where again Ω is the unit cell volume and the factor two stems from double occupancy. A similar factor is the consequence of the nuclear contribution, resulting in an additional integer times $\frac{eR}{\Omega}$. In practical calculations this is rarely a problem, as the change of polarization is for most common perturbations much smaller than the quantum. If the change is bigger, the overall treatment can be modified by subdividing the calculation of the path in Eq (3.10) running from 0 to Δt in multiple steps. However, information is lost when the polarization changes by more than p_n , as a consequence of the modulo indeterminacy as the “observed” change in polarization $\Delta\tilde{P}$ is given as

$$\Delta\tilde{P} = \text{mod}(\Delta P, p_n) \quad . \quad (3.44)$$

3.5 Fields in DFT

The first question when considering electric fields, is how to determine its influence on the investigated system. It can either be considered by perturbation treatments or by directly adding the electric part to the systems Hamiltonian. When using perturbation treatments, the systems eigenfunctions can be used in order to obtain the perturbed ones. However, this approach is only valid, as long as the applied field is small and thus the resulting energy shift is also. Perturbative treatments have been successfully used throughout the literature^{134–137}. When dynamical properties or strong electric fields are the focus of interest, the direct consideration of finite electric fields comes into focus. Within gas phase calculations, the potential providing an electric field, can be directly added to the system’s potential and leads to a superposition of the potentials. An example scheme is depicted in Figure 3.2. As a result, the total energy functional \mathfrak{F} obtained is the sum of the ground state energy and the field interaction term given as

$$\mathfrak{F}[\Psi] = E^0[\Psi] - \mathcal{E}P[\Psi] \quad . \quad (3.45)$$

Where E^0 denotes the system's ground-state energy in the absence of an external field and Ψ is the electronic wave function. Minimization of the expression of Eq (3.45) with respect to the wave functions, leads to the systems ground state energy in the presence of the electric field. The field itself is provided by an electrostatic potential via

$$\mathcal{E} = -\nabla V_{\mathcal{E}}(\mathbf{r}) \quad . \quad (3.46)$$

The simplest finite field, is provided in the homogeneous case. It is provided as a result of a linear potential, which, in the one dimensional case, can exemplarily be given as

$$V_{\mathcal{E}}(x) = -\mathcal{E}x + c \quad , \quad (3.47)$$

which causes a resulting homogeneous field of strength \mathcal{E} in x -direction. As a direct consequence, the total potential acting onto the electrons is now given by a superposition of $V_{\mathcal{E}}$ and the cores Coulomb potential V_c . A schematic functional form of a total potential V_{tot} is given in Figure 3.2.

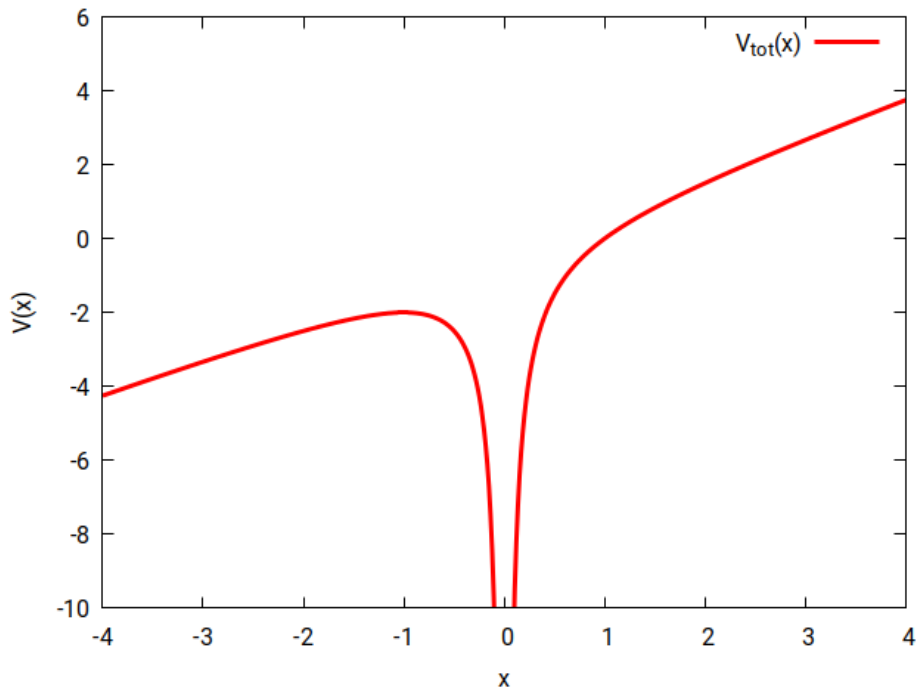


Figure 3.2: Schematic representation of a cross section along the x -direction showing the resulting total potential $V_{\text{tot}} = V_c + V_{\mathcal{E}}(x)$.

3.6 Periodic Electric Fields

When fields are to be considered within PBCs, multiple aspects have to be accounted for and the simple minimization of Eq (3.45) is no longer sufficient. The main issues, are:

- a) The potential providing the electric field, has to be lattice periodic.
- b) The definition of polarization has to be adjusted, in order to suit the cyclic topology.
- c) The Bloch states arising for $\hat{\mathbf{H}}_0$ are no longer eigenfunctions to the field Hamiltonian, and newly created functions do not necessarily carry the Bloch symmetry¹³⁸.

Even in the most simplistic case of a homogeneous electric field, the superposition of $V_{\mathcal{E}}$ with the lattice potential leads to a variation of the total potential acting on the electrons within the super-cell. The superposition can no longer fulfill the BvK, as is shown for an example situation in one dimension in Figure 3.3. In order to provide a lattice periodic potential $V_{\mathcal{E}}$, the field's potential has to be periodic.

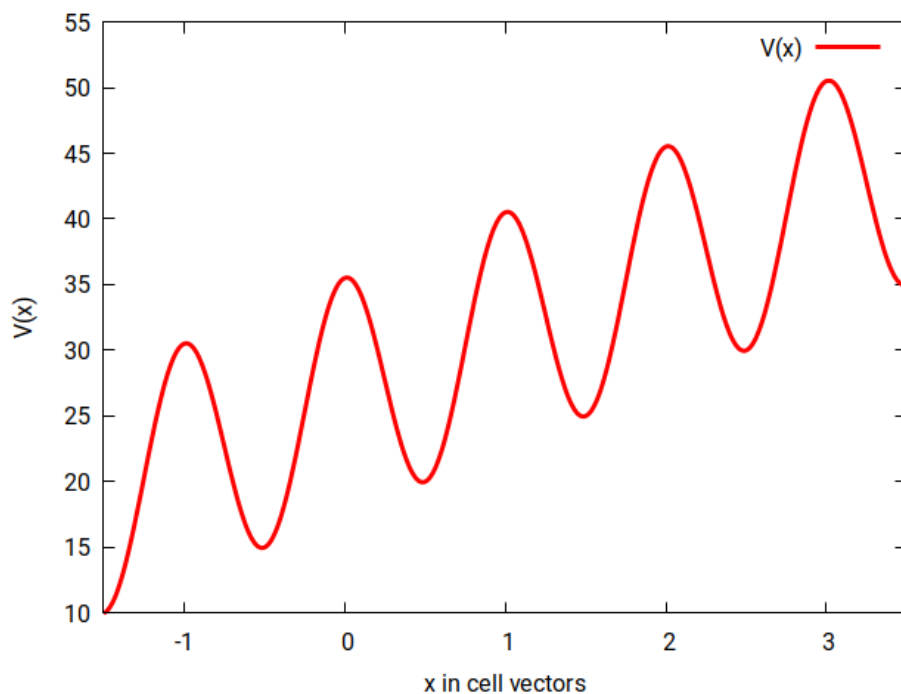


Figure 3.3: Schematic representation of a periodic potential superimposed with an external field in x direction. Depicted are multiple unit cells showing that the potential in adjacent cells gets displaced by a constant offset.

A simple approach to deal with these effects by utilizing a super-cell treatment was investigated by Kunc and Resta¹³⁹. In an approximation, the potential providing an external field, can be reset in the middle of the cell, resulting in the saw-tooth shape as depicted for the one dimensional example in Figure 3.4. Two possible molecular placements can be thought of with the most intuitive being shown in Figure 3.4a. Alternatively, the molecular geometry can be placed over the boundaries to save computational time, as the density empty regions are then smaller as shown in Figure 3.4b. However, such approaches come usually with additional side effects due to the field's divergence at the potential's step position or large computational times needed for treatment of sufficiently large slabs in field direction¹⁴⁰.

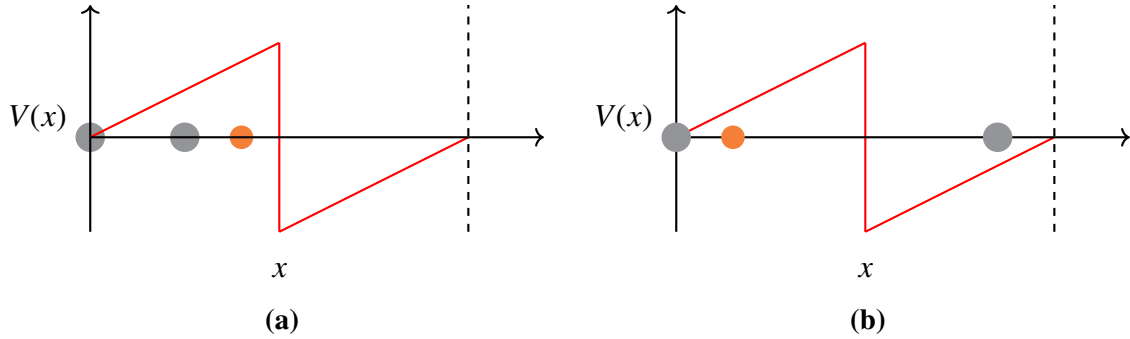


Figure 3.4: Schematic representation for molecular arrangements when utilizing a periodic saw-tooth potential in x direction. Grey circles represent the atomic positions with lowest/highest x -coordinate, the red circle indicates the minimum distance necessary to avoid effects of the divergence created by the potential step. a) Possible molecular arrangement considering the cell up to the potential step. b) Possible arrangement utilizing the periodic boundary conditions.

A fundamental way of avoiding the implementation of a saw-tooth potential and the use of a super-cell approach, is provided by a real space localized representation based on field-polarized-Wannier (FPW) functions¹⁴¹. Here, the definition of P as provided by the MTP is used in combination with the FPW functions. By introducing a set of FPW-functions $w_{0,n}^{\mathcal{E}}$ that are localized in R -space, the possibility of a localized treatment for \mathcal{E} is given. Combining all introduced concepts of the previous sections, namely redefinition of the polarization Section 3.2, rewriting the orbitals of the original unit cell in terms of localized Wannier functions Section 3.3 and combination of these into the field functional Eq (3.45), leads to

$$\mathfrak{F}(W_0^{\mathcal{E}}) = E(W_0^{\mathcal{E}}) - \mathcal{E}e \sum_{n=1}^{n_b} \langle w_{0,n}^{\mathcal{E}} | \hat{\mathbf{r}} | w_{0,n}^{\mathcal{E}} \rangle \quad . \quad (3.48)$$

$W_0^{\mathcal{E}}$ is the total electronic ground state wave function in terms of the FPWs in the presence of an external electric field, and the summation runs over all orbitals n_b contributing within the first unit cell. This treatment is especially promising, as it allows for a calculation of the ground state without considering the “whole” bulk. The localization of the constructed wave functions to a specific localization region (LR), leaves the properties to vary freely inside the LR and zero outside. Therefore, the overall consideration, can as well be limited to a finite amount of LR (in the following only the fundamental real space unit cell with $\mathbf{R} = 0$ is considered). These localized basis functions, can be created by a unitary transformation of the Bloch eigenfunctions, and are then optimized for their orthogonality in order to span the maximum amount of the occupied subspace of the periodic Hamiltonian as proposed by Mauri, Galli and Car (MGC)¹⁴². The natural choice for basis functions to represent the density is given by the localized Wannier functions $w_{0,n}$. These are per definition localized within a corresponding unit cell Section 3.3, but are not necessarily orthogonal. This can be compensated by considering the orthogonality in the total energy functional, using idempotency as additional constraint. The initial trial density $\tilde{\rho} = \sum_j^M |w_{0,n}\rangle \langle w_{0,n}|$ can be constructed via a unitary transformation from the system’s Bloch functions. Idempotency is enforced via consideration of

$$\tilde{\rho} = \rho \sum_{b=0}^B (2 \cdot \mathbf{1} - \rho)^b \quad , \quad (3.49)$$

where B controls the accuracy of the orthogonality constraint and is chosen to be odd numbered. Optimization of localized orbitals, is performed with the $w_{0,n}$ being varied within their respective LR. For periodic systems the boundary conditions can be included, via translation of the Wannier functions of the fundamental unit cell, where the full bulk density is reproduced by

$$\rho = \sum_{\mathbf{R},n} |w_{\mathbf{R},n}\rangle \langle w_{\mathbf{R},n}| \quad , \quad (3.50)$$

where the summation runs over cell vectors(indices) \mathbf{R} and band indices n respectively. Such an approach and its generalization to the electric field treatment, was first proposed and tested by Nunes¹⁴¹. Following this approach and consideration of the corresponding energy

expression in terms of Eq (3.50) given as

$$E_0[\rho] = 2 \sum_{n=1}^M \langle w_{0,n} | \hat{\mathbf{H}} | w_{0,n} \rangle - \sum_{n,m=1}^M \sum_{\mathbf{R}} \langle w_{0,n} | w_{l,n} \rangle \langle w_{l,n} | \hat{\mathbf{H}} | w_{0,n} \rangle \quad , \quad (3.51)$$

with $\hat{\mathbf{H}} = \hat{\mathbf{H}}_0 + e\mathcal{E}\hat{\mathbf{r}}$ as well as the polarization in terms of Wannier functions, given as

$$\mathbf{P}_e[\rho] = \frac{e}{\Omega} \left[2 \sum_{n=1}^M \langle w_{0,n} | \hat{\mathbf{r}} | w_{0,n} \rangle - \sum_{n,m=1}^M \sum_{\mathbf{R}} \langle w_{0,n} | w_{\mathbf{R},n} \rangle \langle w_{\mathbf{R},n} | \hat{\mathbf{r}} | w_{0,n} \rangle \right] \quad , \quad (3.52)$$

the total energy functional in the presence of an electric field, can be expressed as

$$\tilde{\mathcal{F}}[\rho^{\mathcal{E}}] = E_0[\rho^{\mathcal{E}}] - e\mathcal{E}\mathbf{P}_e[\rho^{\mathcal{E}}] \quad . \quad (3.53)$$

Here, the density $\rho^{\mathcal{E}}$ is constructed in terms of the FPW, which no longer describe the bands of the corresponding system at zero field. However, they are expected to be reasonably similar for small fields. A minimization of the functional Eq (3.53) at fixed \mathcal{E} can directly be performed using any optimization algorithm¹⁴¹. An alternative way was thought of by Souza¹⁴³ and uses a \mathbf{k} -space representation of polarized Bloch orbitals for the calculation.

3.7 Field Effects Illustrated

The previous sections describe the differences in the mathematical formalism in the presence of electric fields. The application of electric fields leads to physical effects on the molecular structures, that are also observable within the quantum chemical description. The induced polarization causes charge shifts, and leads to a shift in the electronic structures of the observed molecules. As a result, various properties change in the presence of \mathcal{E} and can be obtained by means of quantum chemical investigations. In the following, some crucial effects are explained schematically. In the later presented investigations in Section 9.1, these effects are illustrated at the example of the H₂ molecule, with investigations performed by means of GPW-DFT.

3.7.1 Orbital Shifts

The application of external fields, leads not only to charge shifts but also to shifts in the structures of the molecular orbitals. The most important effects can be observed when analyzing the energy gap of the highest occupied molecular orbital (HOMO) and lowest unoccupied molecular orbital (LUMO). The HOMO-LUMO gap is calculated as

$$E_{\text{GAP}} = E_{\text{LUMO}} - E_{\text{HOMO}} \quad , \quad (3.54)$$

which is positive for a stable molecular configuration, but becomes negative if an unstable state is reached. As an external electric field leads to further polarization within the molecules, it also increases the respective interactions of individual atoms with the field, leading to energy shifts of the atomic orbitals and consequently for the molecular orbitals. With increasing polarization the molecule becomes less stable due to stronger interactions with the electric field, resulting in a general reduction of the HOMO-LUMO-gap. For a spin unrestricted molecule with distinct spin up and down states, the configurational shift is illustrated in Figure 3.5, with the stable state shown in Figure 3.5a and the unstable state depicted in Figure 3.5b. As can be seen, the highest occupied orbital in the spin down frame $E_{\downarrow}^{\text{H}}$, lies above the lowest unoccupied in the spin up frame E_{\uparrow}^{L} . As a consequence, the resulting electronic configuration is unstable and a separation of the atoms included in the previous bonding case, would be suitable in order to allow for relaxation in the now more energetically favorable non-bonding configuration.

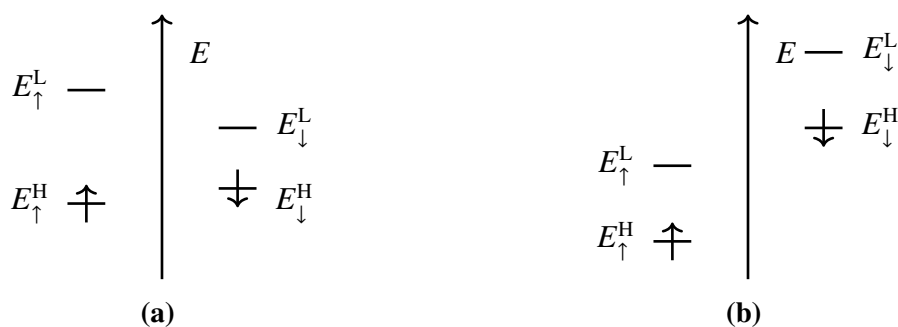


Figure 3.5: The configurations of the HOMO orbital E_i^{H} and LUMO orbital E_i^{L} for both spin states in a) the stable case and b) the unstable case after the field has reached the bond-breaking threshold.



Figure 3.6: Schematics of possible charge shifts in a homo-nuclear dimer. The total charge shift between the two atoms is indicated with δ and can be observed in the simulation via partial charge calculations. The length of the equilibrium geometry before and after field application is indicated by r_0 for the field free case and $r_{\mathcal{E}}$ in the presence of fields.

3.7.2 Charge Shifts

As already mentioned in the introduction Section 3.7, the induced polarization on a molecular geometry, leads to charge shifts. This separation occurs even in molecules that are neutrally charged in the absence of a field. These charge shifts are directly related to the energetic shifts of orbitals as depicted in Figure 3.5. A schematic representation is shown in Figure 3.6, where the situation is explained for a simple homo nuclear dimer. To analyze charge shifts in the presence of electric fields, partial charges associated with individual atoms can be calculated. In a DFT framework, this can be understood as a shift of electron density. In the presented example this is connected to the “flow” of electron density from the bonding state towards atom two, thus resulting in a shift of the total charges. In the presented alignment, this results in the formation of a positive partial charge at atom one and a corresponding negative charge at atom two. In the case of large electric fields, the bond character changes from covalent to ionic. The forces acting on them via an electric field, also cause a new equilibrium distance for the molecule, which is discussed in the following.

3.7.3 Bond Profiles

As a consequence of the previously described shift in orbital energies in the presence of external fields, the molecular equilibrium geometries are influenced. This can be understood by considering the already described charge shifts. Potentially, for increased field strengths,

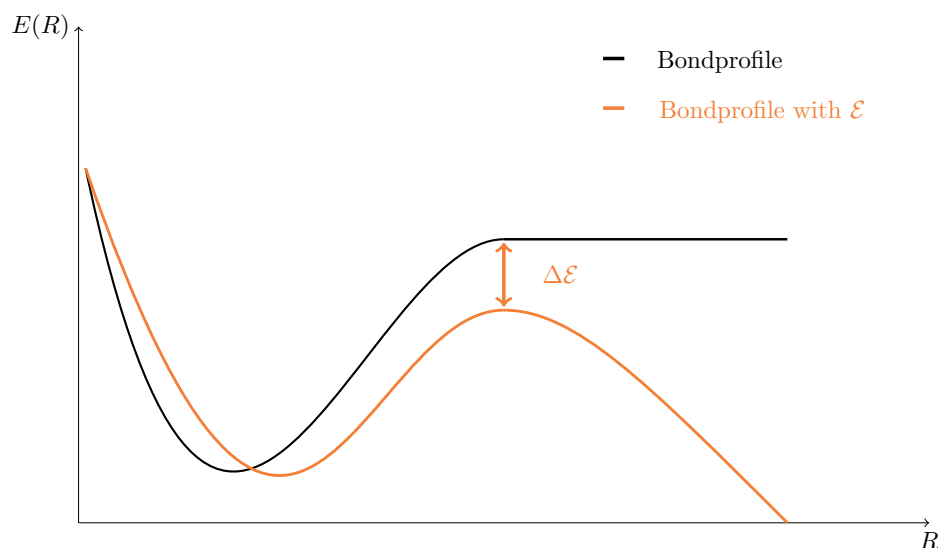


Figure 3.7: Schematic representation of a hypothetical bond profile, for a homo-nuclear dimer, with the bond profile without electric field depicted in black, the modified profile with finite \mathcal{E} is presented in orange.

this leads to a partial ionization of the atoms involved in molecular bonds. In the case of a single bond, consequently, a performed geometry optimization results in a rearrangement of the nuclei's equilibrium distance. This can be illustrated by comparing bond profiles and the resulting energetic minima. A second observable consequence, is the reduction of the dissociation barrier connected to the bond energy. This can be understood, as the bonding interactions holding the molecule together are superposed by the field interaction with the polarized system. On the mathematical level, the energies of the gas phase systems for various nucleic arrangements $E(\mathbf{R})$ are superimposed with the electric field contribution Eq (3.47). The resulting forces lead to repulsion between the charges and partially compensate the bonding interactions. In summary, the molecular rearrangements and shrinking dissociation barriers depend on the charge separation as well as the applied field strength. A schematic representation is given in Figure 3.7, showing the shift of minimum distance and the shrinking of the dissociation barrier shrinking by ΔE for an arbitrary field smaller than the bond breaking threshold. The barrier vanishes as soon as the critical field strength $\mathcal{E}_{\text{crit}}$ is reached, leading to the dissociated structure being the energetically most favorable.

Chapter 4

Tight Binding Methods

The applicability of DFT and other ab-initio methods is generally limited due to bad scaling behavior with respect to the system size, as a result of numerous explicit integrals. For extended super cell systems, simplifications to reduce the computational effort are necessary. Tight binding (TB) methods offer a way of keeping parts of the quantum mechanical description by approximating integrals with parametrized analytic functions. Within this thesis semi empirical density functional tight binding (DFTB) based methods and corresponding extensions are used excessively.

This chapter aims at providing an overview of the fundamental principles for the used TB methods. The idea of integral approximations by analytic functions and other computation time saving principles, are demonstrated for the idealized one dimensional crystal. Subsequently, the hierarchically extended description of DFTB methods, describing the density of a molecular system build from atomic references, is explained in detail. A special emphasis is put on the treatment of density fluctuations and approximations of the exchange and correlation contribution for the self-consistent charge (SCC) formalism. The respective methodologies of DFTB2 and its third order extension DFTB3 are investigated for their modifications. With these basic concepts in mind, the working mechanisms of the DFTB3-based geometries frequencies non-covalent interactions extended tight binding (GFN-xTB) model, are reviewed, focusing on its connection to the DFTB3 description.

4.1 The Concepts of Tight Binding Models in the Solid State

Understanding the DFTB formalism is simplified when considering the basic ideas of the TB approximation. In the following these are presented for the one dimensional crystal. The working scheme as presented in the following is mainly inspired by^{144,145}. As described in Section 2.8, the band structure of a solid is directly connected to the energy levels of the isolated atoms within its unit cell. The simplification within the TB approximation consists of separating the systems total Hamiltonian $\hat{\mathbf{H}}_{\text{tot}}$ in two main contributions

$$\hat{\mathbf{H}}_{\text{tot}} = \hat{\mathbf{H}}_{\text{iso}} + \hat{\mathbf{U}} \quad , \quad (4.1)$$

where $\hat{\mathbf{H}}_{\text{iso}}$ is the Hamiltonian of the isolated atoms and $\hat{\mathbf{U}}$ is a correction arising due to inter-atomic interactions within a molecule or crystal. It is further assumed that $\hat{\mathbf{U}} \rightarrow 0$ in the center of each respective atom where $\mathbf{r} \rightarrow \mathbf{R}$. The number of energy bands in a periodic description is determined by the electronic orbitals of the atoms comprising the elementary unit-cell as described in Section 2.8. The band wave functions for orbital n are constructed using Bloch's theorem following Eq (2.71) as superposition of Bloch functions $\Phi_{i,\mathbf{k}}$ given as

$$\Phi_{i,\mathbf{k}} = \frac{1}{\sqrt{N}} \sum_{\mathbf{R}} e^{i\mathbf{k}\mathbf{R}} \phi_i(\mathbf{r} - \mathbf{R}) \quad . \quad (4.2)$$

Where \mathbf{R} is a real space lattice vector as defined in Section 2.7, ϕ_i are atomic basis functions, $\frac{1}{\sqrt{N}}$ is a normalization constant and \mathbf{k} is the corresponding crystal momentum for BZ treatment as described in Section 2.9. The atomic orbitals $\phi_i(\mathbf{r})$ are eigenfunctions to the isolated atoms Hamiltonian $\hat{\mathbf{H}}_{\text{iso}}$ and therefore fulfill

$$\hat{\mathbf{H}}_{\text{iso}}\phi_i = \epsilon_i\phi_i \quad . \quad (4.3)$$

For a system consisting of an N_b -atomic basis with one relevant orbital per atom contributing to the total energy, the band wave function can be constructed via linear combination of N_b Bloch-orbitals $\Phi_{i,\mathbf{k}}$.

4.1 The Concepts of Tight Binding Models in the Solid State

This leads to the expression

$$\Psi_{n,\mathbf{k}(\mathbf{r})} = \sum_{i=1}^{N_b} c_{i,\mathbf{k}} \Phi_{i,\mathbf{k}}(\mathbf{r}) \quad . \quad (4.4)$$

The ground state energy is then found by variation of the coefficients $c_{i,\mathbf{k}}$ in analogy to Section A.3. The eigenvalue equation for band n is therefore given as

$$\hat{\mathbf{H}}_{\text{tot}} \Psi_{n,\mathbf{k}} = E_{n,\mathbf{k}} \Psi_{n,\mathbf{k}} \quad . \quad (4.5)$$

This equation can be solved as explained in Section 2.6.3 by rewriting it in matrix form resulting in

$$\sum_i (H_{ij} - \delta_{ij} E(\mathbf{k})) c_{j\mathbf{k}} = 0 \quad , \quad (4.6)$$

where H_{ij} corresponds to the matrix elements of the Hamiltonian given as $\langle \Phi_{i,\mathbf{k}} | \hat{\mathbf{H}}_{\text{tot}} | \Phi_{j,\mathbf{k}} \rangle$. The resulting problem can be summarized as

$$\langle \Psi_{n,\mathbf{k}} | \hat{\mathbf{H}}_{\text{iso}} + \hat{\mathbf{U}} | \Psi_{n,\mathbf{k}} \rangle = \langle \Psi_{n,\mathbf{k}} | \hat{\mathbf{H}}_{\text{iso}} | \Psi_{n,\mathbf{k}} \rangle + \langle \Psi_{n,\mathbf{k}} | \hat{\mathbf{U}} | \Psi_{n,\mathbf{k}} \rangle \quad . \quad (4.7)$$

Assuming zero overlap between different lattice sites alongside orthonormal basis functions, leads to a vanishing overlap integral S_{ij} , given as

$$S_{ij} = \int \phi_i(\mathbf{r}) \phi_j(\mathbf{r} + \mathbf{R}) d\mathbf{r} \quad , \quad (4.8)$$

for $\mathbf{R} \neq 0$ and $i \neq j$. The resulting formalism is referred to as orthogonal tight binding formalism and marks the most simplistic form of description.

4.1.1 Ideal one Atomic Crystal

The fundamental principles are explained in a simple manner for the single atom one dimensional crystal. For only one contributing orbital ϕ to the crystal states, rewriting the integration of Eq (4.7), results in the total energy expression:

$$E(\mathbf{k}) = \frac{1}{N} \sum_{\mathbf{R}} \sum_{\mathbf{R}'} e^{i\mathbf{k}(\mathbf{R}'-\mathbf{R})} \int \phi^*(\mathbf{r} - \mathbf{R})H\phi(\mathbf{r} - \mathbf{R}')\mathbf{d}\mathbf{r} \quad (4.9)$$

Here \mathbf{R} and \mathbf{R}' are translation vectors on the one dimensional lattice. They are given as $\mathbf{R} = na_0\mathbf{e}_x$, where n is an integer a_0 the lattice constant and \mathbf{e}_x the unit vector in x direction. The closest neighbors translation vectors are then given as $\tau = \pm a_0\mathbf{e}_x$. Interpretation is simplified by reformulating the expression in three steps:

- Substituting \mathbf{r} with $\mathbf{x} = \mathbf{r} - \mathbf{R}$.
- Identifying $\mathbf{R}' - \mathbf{R}$ with a new lattice vector \mathbf{R}'' and executing the summation over \mathbf{R} leads to a factor of N .
- Separating the $\mathbf{R}'' = 0$ contribution as it leads to the atomic energy term and only including nearest neighbors contributions τ .

The resulting expression for the dispersion relation is then given as

$$E(\mathbf{k}) = \epsilon + \sum_{\tau} e^{i\mathbf{k}\tau} \int \phi^*(\mathbf{x})H\phi(\mathbf{x} - \tau)\mathbf{d}\mathbf{x} \quad . \quad (4.10)$$

The energy of the isolated atom ϵ is corrected by an interaction term corresponding to the nearest neighbors in the crystal chain. In empirical tight binding methods, these expressions are replaced with a parameter following

$$\gamma(|\tau|) = \int \phi^*(\mathbf{x})H\phi(\mathbf{x} - \tau)\mathbf{d}\mathbf{x} \quad . \quad (4.11)$$

The resulting expression for the energy is given as

$$E(\mathbf{k}) = \epsilon + \sum_{\tau} e^{i\mathbf{k}\tau} \gamma(|\tau|) \quad , \quad (4.12)$$

where for a one dimensional chain, the wave vectors are directed along the lattice in x -direction as $\mathbf{k} = ke_x$. The resulting dispersion relation can therefore be explicitly written as

$$E(k) = \epsilon + \gamma(a_0)(e^{ika_0} + e^{-ika_0}) = \epsilon + 2\gamma(a_0) \cos(ka_0) \quad . \quad (4.13)$$

The dispersion relation gives the energy corresponding to a given crystal momentum, where the energetic location is mainly determined by the energy of the isolated atom ϵ and the energetic bandwidth can be identified to be 4γ . The dispersion relation is periodic with respect to the reciprocal cell and therefore only k values modulo $2\pi/a_0$ contribute different values of energy. Adjusting the parameter γ directly determines the bandwidth and, thus, the allowed energetic maximum and minimum of the associated band. For practical applications, the parameter γ is therefore adjusted to fit the experiment. The parameter γ depends in general on the atomic distances and therefore varies with changing atomic positions, directly influencing the obtained band structure. In typical tight binding models, $\gamma(\tau)$ is modeled by a function, specifically parameterized for the corresponding atom and orbital pair, if more than one type is used. The parameterization is based either on experimental values or more sophisticated theoretical descriptions as discussed in the following section.

4.2 Density Functional Tight Binding

The Density functional tight binding method (DFTB), is based on the concepts of KS-DFT, utilizing the ideas of the semi-empirically tight binding models^{146,147}. It is based on a second order expansion of the DFT energy expression around a reference density as shown in Section 4.2.2. Hamilton-matrix elements are calculated within a minimal LCAO. The remaining terms are approximated such that the diagonalization of the pre-computed Hamilton-matrix dominates the efforts necessary^{148,149}. All approximations are centered around the calculated reference ρ_0 and its deviation $\delta\rho$ with respect to the DFT ground state density ρ ^{150,151}. The simplest version of DFTB is in essence a stationary approximation to DFT-functionals, allowing for the description of small charge transfers within the description^{152,153}. However, it fails when treating intermediate charge shifts, where the self consistent charge (SCC) extension comes into account¹⁵⁴. The basic assumptions used within DFTB are explained in

4 Tight Binding Methods

the following, where the discussion is limited to the energy corresponding to a single \mathbf{k} -point. The generalization over the whole BZ is performed in analogy to the DFT case, as shown in Section 2.9. An overview of the utilized descriptions from DFTB to GFN-xTB is provided in Figure 4.1, showing the hierarchical improvements in the underlying description.

4.2.1 DFTB-Energy Expression

The derivation of the basic DFTB energy expression is illustrated as presented by Seifert¹⁵⁵. The KS-energy expression introduced in Section 2.4 can be reformulated in terms of the non interacting orbitals and pairwise correction terms (resulting from the intermolecular density rearrangements). The resulting formulation marks the basis for the used DFTB approximations and is given as follows

$$E_{\text{DFT}}[\rho(\mathbf{r})] = \sum_i n_i \epsilon_i - \frac{1}{2} \int V_{\text{ee}}[\rho(\mathbf{r})] \rho(\mathbf{r}) \mathbf{d}\mathbf{r} - \int V_{\text{xc}}[\rho(\mathbf{r})] \rho(\mathbf{r}) \mathbf{d}\mathbf{r} + E_{\text{xc}} + E_{\text{nn}} \quad . \quad (4.14)$$

A detailed derivation is presented in Section C.1. The main idea of DFTB is the description of the total molecular electronic density ρ in terms of the KS-orbitals Ψ_i is constructed as the sum of an atomic reference density ρ_0 and small fluctuations $\Delta\rho$. The resulting connection is given as

$$\Delta\rho = \rho - \rho_0 \quad , \quad (4.15)$$

describing the fluctuations of the density as the difference of the molecular from the atomic case. The obtained density fluctuations have to fulfill the charge conservation

$$\int \Delta\rho \mathbf{d}\mathbf{r} = 0 \quad . \quad (4.16)$$

Substituting Eq (4.15) in Eq (4.14) and sorting the terms according to their dependence on charge fluctuations and reference density, leads to the general DFTB energy expression given

as

$$E_{\text{DFTB}}[\rho(\mathbf{r})] = \sum_i n_i \langle \Psi_i | \hat{\mathbf{H}}_0(\rho_0) | \Psi_i \rangle - \frac{1}{2} \int \int \frac{\rho(\mathbf{r})\rho(\mathbf{r}') + \delta\rho(\mathbf{r})\delta\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} \mathbf{d}\mathbf{r}\mathbf{d}\mathbf{r}' \quad (4.17)$$

$$- \int V_{\text{xc}}[\rho_0]\rho(\mathbf{r})\mathbf{d}\mathbf{r} + E_{\text{xc}}(\rho_0 + \delta\rho) + E_{\text{nn}} \quad (4.18)$$

A detailed derivation of the reformulated equation is provided in Section C.2. Subsequently, the exchange correlation term E_{xc} is expanded in a Taylor series around the reference density ρ_0 . The expanded exchange correlation energy is given as Eq (C.11), leading to the different levels of flexibility within the developed DFTB model, depending on its truncation. In DFTB3, E_{xc} is expanded up to third order, leading after insertion in Eq (4.18) to the full DFTB3-energy expression given as

$$\begin{aligned} E_{\text{DFTB3}}[\rho_0 + \delta\rho] &= \sum_i n_i \langle \Psi_i | \hat{\mathbf{H}}_0(\rho_0) | \Psi_i \rangle - \int V_{\text{xc}}[\rho_0]\rho(\mathbf{r})\mathbf{d}\mathbf{r} \\ &\quad - \frac{1}{2} \int \int \frac{\rho_0(\mathbf{r})\rho_0(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} \mathbf{d}\mathbf{r}\mathbf{d}\mathbf{r}' + E_{\text{xc}}[\rho_0] + \underbrace{\frac{1}{2} \sum_{a,b} \frac{Z_a Z_b}{R_{ab}}}_{E_{\text{nn}}} \\ &\quad + \underbrace{\frac{1}{2} \int \int \left(\frac{1}{|\mathbf{r} - \mathbf{r}'|} + \frac{\delta^2}{\delta\rho(\mathbf{r})\delta\rho(\mathbf{r}')} \Big|_{\rho_0} \right) \delta\rho(\mathbf{r})\delta\rho(\mathbf{r}') \mathbf{d}\mathbf{r}\mathbf{d}\mathbf{r}'}_{E_2[\rho_0, (\delta\rho)^2]} \\ &\quad + \underbrace{\frac{1}{6} \int \int \int \frac{\delta^3 E_{\text{xc}}[\rho]}{\delta\rho(\mathbf{r})\delta\rho(\mathbf{r}')\delta\rho(\mathbf{r}'')} \Big|_{\rho_0} \mathbf{d}\mathbf{r}\mathbf{d}\mathbf{r}'\mathbf{d}\mathbf{r}''}_{E_3[\rho_0, (\delta\rho)^3]} \quad (4.19) \end{aligned}$$

The basic concepts leading to the most recent DFTB3 formalism are explained by investigating the hierarchical contributions leading from DFTB to DFTB2 and then DFTB3.

4.2.2 Basic DFTB

Simple TB models neglect second order and higher terms in the density fluctuations¹⁵⁶, leading to the basic concept of describing the molecular case as isolated atoms and position depending pairwise correction terms as illustrated in Section 4.1. In the most fundamental

4 Tight Binding Methods

DFTB formalism, a systems energy is given as

$$E_{\text{DFTB}}[\rho_0] = \sum_i n_i \epsilon_i - \underbrace{\int V_{\text{xc}}[\rho_0] \rho(\mathbf{r}) \mathbf{d}\mathbf{r} + E_{\text{nn}} - \frac{1}{2} \int \int \frac{\rho_0(\mathbf{r}) \rho_0(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} \mathbf{d}\mathbf{r} \mathbf{d}\mathbf{r}' + E_{\text{xc}}[\rho_0]}_{\approx E_{\text{REP}}} \quad , \quad (4.20)$$

consisting only of the non interacting orbitals, the classical Coulomb self interaction of the reference density ρ_0 , nuclear-nuclear repulsion E_{nn} and exchange correlation contributions V_{xc} and E_{xc} ^{150,157}. The latter ones are approximated as a repulsive potential E_{REP} , being comprised of a nuclear repulsion term V_a^{rep} as well as pairwise interaction potentials V_{ab}^{rep} and can be written as

$$E_{\text{REP}}[\rho_0] = \sum_a V_a^{\text{rep}}[\rho_0^a] + \frac{1}{2} \sum_{ab} V_{ab}^{\text{rep}}[\rho_0^a, \rho_0^b, r_{ab}] \quad . \quad (4.21)$$

Here, the sum runs over all atoms a and b with their reference densities ρ_0^a and ρ_0^b respectively. The atomic contributions sum up to a constant energy shift and are neglected in DFTB as properties based on energy differences are the scope of interest. The repulsive potential therefore consists only of pair wise contributions and the total DFTB energy expression is given as

$$E_{\text{DFTB}}[\rho_0] = \sum_i n_i \epsilon_i + \frac{1}{2} \sum_{ab} V_{ab}^{\text{rep}}[\rho_0^a, \rho_0^b, r_{ab}] \quad , \quad (4.22)$$

which shows the same structure as the initially illustrated example of the ideal crystal Eq (4.10). The terms in Eq (4.22) are density independent and the systems energy can therefore be obtained by constructing the orbitals Ψ_i from an LCAO ansatz as $\Psi_i = \sum_{\mu} c_{i,\mu} \phi_{\mu}$. The $c_{i,\mu}$ are obtained by solving atomic KS-equations in an additional confinement potential for each orbital Ψ_i respectively as described in Section C.3. For known $c_{i,\mu}$ the energy can be directly calculated following Eq (4.22). However, this treatment does not include any flexibility for reorganization of charges as it is solely determined by the atomic references and the parametrized atomic interactions. A schematic representation of the constructed density ρ_0 from atomic reference entering the energy calculation, is provided in Figure 4.1a.

4.2.3 DFTB2-Extension

To allow for charge reorganization, the second order terms in Eq (4.19) become important, leading to charge redistribution for a given molecular arrangement. The second order terms $E^{2\text{nd}}$ are given as

$$E^{2\text{nd}}[\rho_0 + \Delta\rho] = \frac{1}{2} \int \int \left(\frac{1}{|\mathbf{r} - \mathbf{r}'|} + \left[\frac{\delta^2 E_{\text{xc}}}{\delta\rho\delta\rho'} \right]_{\rho_0, \rho'_0} \right) \Delta\rho(\mathbf{r})\Delta\rho(\mathbf{r}')\mathbf{d}\mathbf{r}\mathbf{d}\mathbf{r}' \quad . \quad (4.23)$$

In order to avoid explicit integration, these terms are approximated by analytic functions depending on the charge fluctuation $\Delta\rho$ given as the sum of atomic contributions ρ_α as

$$\Delta\rho = \sum_{\alpha} \Delta\rho_{\alpha} \quad . \quad (4.24)$$

Each charge distribution can be expanded in terms of spherical harmonics. In DFTB2 the series is truncated at the monopole term for maximal computational efficiency. The fluctuations can therefore be written as changes in a spherical charge distributions as

$$\Delta\rho_{\alpha} \approx \Delta q_{\alpha} F_{\alpha}^{00} Y^{00} \quad , \quad (4.25)$$

Where F_{α}^{00} is the shape and Y^{00} the zeroth order spherical harmonic function. The charge shift Δq_{α} is defined as the difference of the Mulliken charge q_{α} within the molecule and its isolated reference q_{α}^0 , given as

$$\Delta q_{\alpha} = q_{\alpha} - q_{\alpha}^0 \quad . \quad (4.26)$$

The Mulliken charge is calculated as

$$q_a = \sum_i n_i \sum_{\mu \in a} \sum_b \sum_{\nu \in b} c_{\mu i} c_{\nu i} S_{\mu\nu} \quad , \quad (4.27)$$

4 Tight Binding Methods

with the expansion coefficients $c_{\mu i}$ of basis function ϕ_μ of orbital Ψ_i , with the overlap matrix elements $S_{\mu\nu}$ given as

$$S_{\mu\nu} = \langle \phi_\mu | \phi_\nu \rangle \quad . \quad (4.28)$$

Within this approximation, density transfer is limited to charge fluctuations Δq_a on a respective atom a . Due to the monopole approximation, the deformation of charge densities is also neglected. Furthermore, exponentially decaying charge densities are assumed, leading for the shape function to

$$F_a^{00} = \frac{\tau_a^3}{4\sqrt{\pi}} e^{-\tau_a |r - \mathbf{R}_a|} \quad , \quad (4.29)$$

with the atom specific coefficients τ_a . Inserting Eq (4.27) and Eq (4.24) under consideration of Eq (4.29) into Eq (4.23), allows for the definition of an analytic function γ for approximating the second order integrals¹⁴⁶. The resulting second order charge fluctuation term is then given as

$$E^{2\text{nd}} \approx E_\gamma = \frac{1}{2} \sum_{a,b} \gamma_{ab} \Delta q_a q_b \quad . \quad (4.30)$$

The function γ_{ab} describes the interaction of the fluctuations between atom a and b . The functional form of γ is chosen to resemble the Coulomb interaction $\propto \frac{1}{r_{ab}}$ for large inter atomic distances r_{ab} . For $a = b$, the function represents the on site self repulsion resulting from the fluctuation within the same atom. In this case, the functional value is predefined as

$$\gamma_{aa} = U_a = \frac{\partial^2 E_{\text{at}}^a}{\partial q_{\text{at}}^2} \quad , \quad (4.31)$$

where the Hubbard parameter U_a is introduced, given as the second derivative of the total atomic energy with respect to the occupation number of the highest occupied atomic orbital. Furthermore, U_a is used to define the parameter $\tau_a = \frac{16}{5} U_a$ determining the decay rate of the charge fluctuation Δq_a correcting it to the covalent radius R_a^{cov} by inverse proportionality as

$U_a \propto \frac{1}{R_a^{\text{cov}}}$. With these considerations, the energy expression of DFTB2 is given as

$$E_{\text{DFTB2}}[\rho_0 + \Delta\rho] = \sum_i n_i \epsilon_i - \frac{1}{2} \sum_{ab} \Delta q_b (q_a + q_a^0) \gamma_{ab} + E_{\text{REP}} \quad . \quad (4.32)$$

Within the previously described extension, charge transfer between atoms is possible, however it is still lacking the possibility of adjusting an atoms "size" accordingly with changing charge state. The reason lies within the determination of the Hubbard parameter from neutral atomic reference calculations which becomes unhandy when treating molecules with large charge shifts. A schematic representation is provided in Figure 4.1b, showing the initially determined ρ_0 and charge equilibration provided by the charge fluctuations $\Delta\rho$ between the systems atoms.

4.2.4 DFTB3-Extension

In order to compensate for the remaining drawbacks in the "static" parameterization of the γ function due to fixed predefined U , the inclusion of third order contributions is straight forward. The further possibilities of adjustment as provided by the DFTB3 model, lead to charge adjusted U parameters as well as the inclusion of changes within the charge shifts¹⁵⁴. Based on the third order term of the Taylor expansion in Eq (C.11), given as

$$E^{\text{3ird}}[\rho_0 + \Delta\rho] = \frac{1}{6} \int'' \int' \int \frac{\delta}{\delta\rho(\mathbf{r}'')} \left[\frac{\delta^2 E_{\text{xc}}[\rho]}{\delta\rho(\mathbf{r}')\delta\rho(\mathbf{r})} \right]_{\rho_0, \rho'_0, \rho''_0} \Delta\rho(\mathbf{r}) \Delta\rho(\mathbf{r}') \Delta\rho(\mathbf{r}'') \mathbf{d}\mathbf{r}'' \mathbf{d}\mathbf{r}' \mathbf{d}\mathbf{r} \quad . \quad (4.33)$$

Using the same approximation procedure as in Section 4.2.3 namely monopole expansion of atomic fluctuations $\Delta\rho_a$ as well as the introduction of γ_{ab} provides the third order contribution

$$E^{\text{3ird}} \approx E^\Gamma = \frac{1}{6} \sum_{abc} \Delta q_a \Delta q_b \Delta q_c \left. \frac{d\gamma_{ab}}{dq_c} \right|_{q_c^0} \quad , \quad (4.34)$$

where the derivative of the γ function with respect to charge occurs¹⁵⁷. The contributions can be split into diagonal and off diagonal terms leading to the formulation

$$E^\Gamma = \frac{1}{6} \sum_a \Delta q_a^3 \left. \frac{\partial \gamma_{aa}}{\partial q_a} \right|_{q_a^0} + \frac{1}{6} \sum_{ab} \Delta q_a \Delta q_b \left(\Delta q_a \left. \frac{\partial \gamma_{ab}}{\partial q_a} \right|_{q_a^0} + \Delta q_b \left. \frac{\partial \gamma_{ab}}{\partial q_b} \right|_{q_b^0} \right) \quad , \quad (4.35)$$

4 Tight Binding Methods

where the resulting derivatives of the γ function can be summarized as

$$\Gamma_{aa} = \left. \frac{\partial \gamma_{aa}}{\partial q_a} \right|_{q_a^0} \quad \Gamma_{ab} = \left. \frac{\partial \gamma_{ab}}{\partial q_a} \right|_{q_a^0} \quad \Gamma_{ba} = \left. \frac{\partial \gamma_{ab}}{\partial q_b} \right|_{q_b^0} \quad (4.36)$$

leading to the compact expression

$$E^\Gamma = \frac{1}{6} \sum_{ab} \Delta q_a \Delta q_b (\Delta q_a \Gamma_{ab} + \Delta q_b \Gamma_{ba}) = \frac{1}{3} \sum_{ab} \Delta q_a^2 \Delta q_b \Gamma_{ab} \quad . \quad (4.37)$$

Inclusion of the third order correction therefore leads via the connection

$$\left. \frac{\partial \gamma_{aa}}{\partial q_a} \right|_{q_a^0} = \left. \frac{\partial \gamma_{aa}}{\partial U_a} \frac{\partial U_a}{\partial q_a} \right|_{q_a^0} \quad , \quad (4.38)$$

to a change in the Hubbard parameter U and therefore in the effective size of the treated atom if charge fluctuations occur^{145,150,151}. The total energy expression in DFTB3 can be summarized as

$$E_{\text{DFTB3}} = E_{H_0} + E_\gamma + E_\Gamma + E_{\text{REP}} \quad . \quad (4.39)$$

A schematic representation of the constructed density ρ is provided in Figure 4.1c, indicating the initially constructed ρ_0 , charge fluctuations $\Delta\rho$ and related changes in the atomic size due to changes in the Hubbard parameter ΔU .

4.3 GFN-xTB

GFN-xTB formalism marks a tight binding approach that is based upon the previously described DFTB3 formalism. Its main purpose however is the description of geometries frequencies and non-covalent interactions⁶¹ by extension of the atomic basis set as well as the used Hamiltonian. The energy in GFN-xTB formalism is given as

$$E_{\text{xTB}} = E_{\text{el}} + E_{\text{REP}} + E_{\text{disp}} + E_{\text{XB}} \quad . \quad (4.40)$$

The contributions can be described along their respective purposes as follows:

- The electronic contribution E_{el} accounts for charge readjustments within the molecule and is the focus of the following work.
- The repulsive pair potential E_{REP} in analogy to the previously described one in Eq (4.21) in the DFTB formalism.
- An empirical dispersion correction E_{disp} is added, based on the D3 method and the Becke Johnson damping scheme¹⁵⁸, without considering three body terms.
- A halogen bonding correction E_{XB} accounting for deficiencies arising from the charge monopole description.

The non-electronic terms are pairwise contributions and not optimized during the SCC cycles. The electronic contributions⁶¹ using the DFTB nomenclature, to show analogies to the previous descriptions, are given as

$$E_{\text{el}} = \sum_i n_i \langle \Psi_i | \hat{\mathbf{H}}_0 | \Psi_i \rangle + \frac{1}{2} \sum_{ab} \sum_{l \in a} \sum_{l' \in b} \Delta p_a^l \Delta p_b^{l'} \gamma_{ab, ll'} + \frac{1}{3} \sum_a q_a^3 \Gamma_a - \underbrace{T_{\text{el}} S_{\text{el}}}_{\text{Entropic-Term}} \quad (4.41)$$

Here $\hat{\mathbf{H}}_0$ is the zero order Hamiltonian in analogy to the one utilized within Eq (4.18), however, its form differs from the DFTB one and avoids element pair specific parameterization⁶¹. The function $\gamma_{ab, ll'}$ describes the distance dependent Coulomb interaction and fulfills the same purpose as γ_{ab} in Section 4.2.3. However, its functional form is different from the one used in DFTB3. Finally Γ_a is given as the derivative of the Hubbard parameter with respect to charge in analogy to the third order diagonal terms described in Section 4.2.4. A major difference consists in the choice of analytic function $\gamma_{ab, ll'}$ as well as the treatment of partial charges q_a . Similar to DFTB, these are calculated as Mulliken charges, however, for each atomic shell l respectively. The shell charges Δp_a^l entering in the second order term are given as

$$\Delta p_a^l = p_{a,0}^l - \sum_v \sum_{\mu \in a, \mu \in l}^{N_{\text{AO}}} S_{\mu\nu} P_{\mu\nu} \quad (4.42)$$

For each shell l of atom a in the molecule, the charge is calculated as deviation from the atomic reference $p_{a,0}^l$. The sum over all occupied shells results in the total Mulliken charge of the

4 Tight Binding Methods

atom, given as

$$q_a = \sum_{l \in a} p_a^l \quad . \quad (4.43)$$

In contrast to the treatment described in Eq (4.23), the fluctuation between distinct shells rather than full atoms is considered. Furthermore, fractional occupations arising from a Fermi smearing in combination with the entropic term in Eq (4.41) are realized. The occupation number n_i is therefore provided for a given temperature T_{el} by a Fermi distribution for orbital Ψ_i , leading to

$$n_i(T_{\text{el}}) = \frac{1}{1 + e^{\frac{(\epsilon_i - \epsilon_F)}{k_B T_{\text{el}}}}} \quad , \quad (4.44)$$

with the orbitals energy ϵ_i and the Fermi energy level ϵ_F . The orbitals Ψ_i are constructed based on an LCAO approach, where Slater type atomic basis functions are approximated by multiple Gaussian orbitals as described in Section 2.6.2. The resulting expansion is given as

$$\Psi_i = \sum_{\mu=1}^{N_{\text{AO}}} c_{\mu,i} \phi_{\mu} \quad , \quad (4.45)$$

where the sum runs over all used basis functions μ upon the total number of atomic orbitals N_{AO} . Typically four to six Gaussian functions are used for the approximation following Stewarts approach¹⁵⁹. Using the expanded orbitals for the calculation of the energies in analogy to the SCC-DFTB procedure leads to the appropriate KS-equations that are then solved self consistently. The procedure of solving the KS-equations in the GFN-xTB formalism is described in detail in¹⁶⁰. An in-depth discussion of the non-electronic pair potentials and the parameterization process of the xTB Hamiltonian, can be found in the corresponding literature^{61,160,161}. An illustration of the density construction is provided in Figure 4.1d. The charge shifts Δp between respective shells p^l leading to the newly constructed density ρ after adjusting the initially determined reference ρ_0 . The associated changes in the atomic size are indicated by changes within the Hubbard parameter ΔU . The used description within the GFN-xTB framework yields accurate geometries even for large systems consisting of more than 1000 atoms. Finally it has to be noted, that GFN-xTB was mainly parametrized for molecular

systems⁶¹. The described GFN-xTB formalism can be used for periodic systems via the cyclic cluster implementation. The cyclic cluster model (CCM) working scheme is described in some detail in Section C.4¹⁴⁴. Within this thesis, periodic GFN-xTB as implemented in CP2K⁶² is excessively used for the description of crystalline polymers.

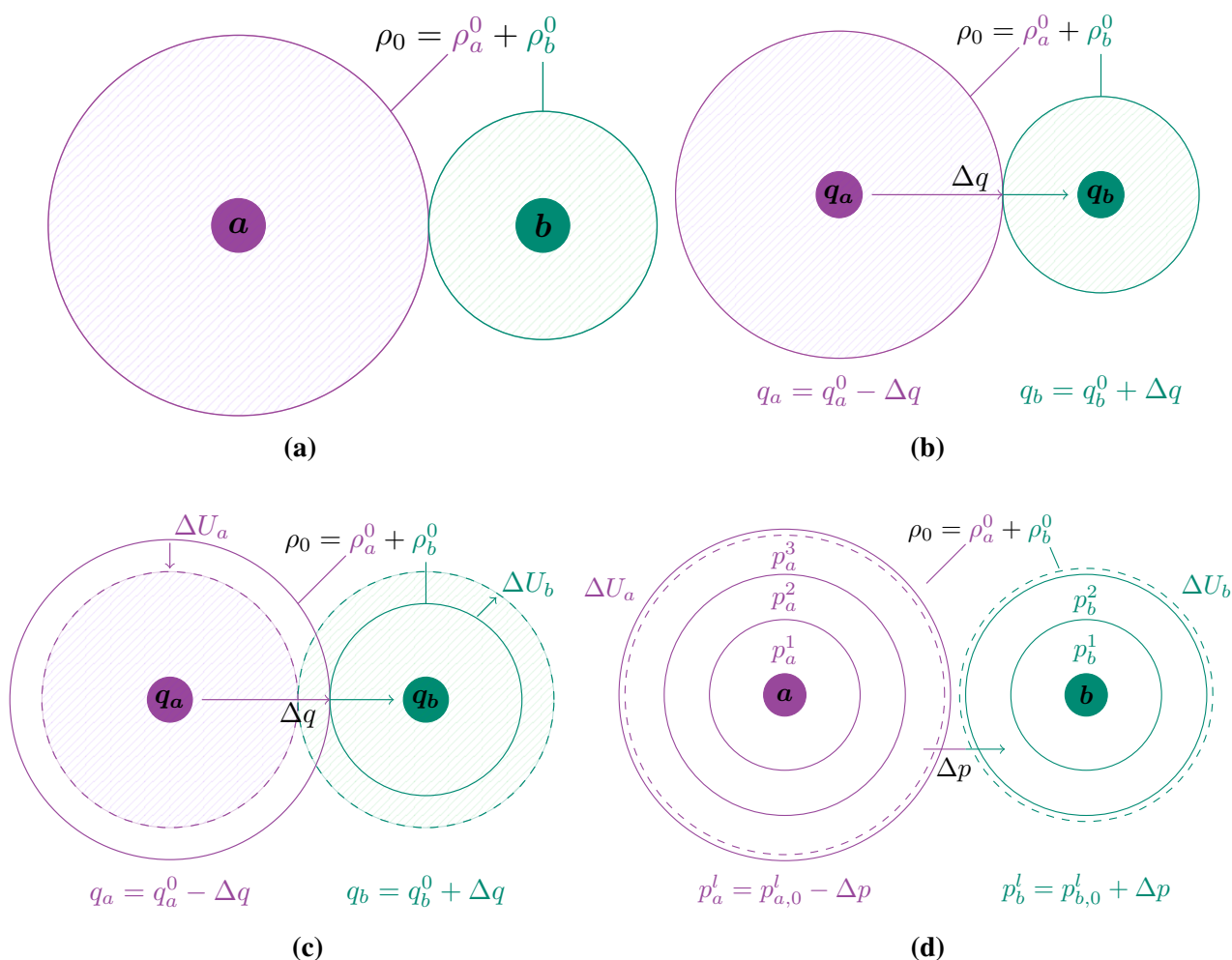


Figure 4.1: Schematic representation of the charge treatment for the hierarchically improved DFTB-family of models. a) DFTB with energies solely determined from the density ρ_0 constructed from atomic references. b) DFTB2 based on reference density ρ_0 and additional charge fluctuation Δq between atoms. c) DFTB3 with charge fluctuations Δq additionally leading to changes in the chemical hardness ΔU , influencing the interatomic repulsion (effective atomic size). d) GFN-xTB, treating the atomic shells separately, leading to charge fluctuations between different shells p which in turn lead to changes in the atomic size U .

Part II

Analysis Tools

Chapter 5

Fourier Random Displacement Generator

Comparability of structural models with experimental data in surfaces or crystalline structures, as often used in applications of solid state physics, is generally limited due to the absence of defects and natural randomness in theoretical models. The analysis of perfectly ordered systems often acts as idealized reference and makes direct comparison of statistical properties difficult. An example from the field of COFs was investigated by Pütz et. al.⁵⁹, where random interlayer displacements were observed for two-dimensional covalent organic frameworks. These real-stacking effects are insignificant for an individual layer's properties. An example are structural features such as inter-layer interactions, component adjustment and preferable stacking types. Absence of real layer effects becomes a major challenge when imperfect alignments are necessary, as such models lack arbitrariness. For compensation of the gap between theory and experiment, the theoretical model has to be modified. The simulation of adsorption isotherms in a multi-layer system as investigated within this thesis, showed the suitability of properly generated random stacking¹⁶². Isotherms simulated with randomized stacking patterns, generated by the Fourier random displacement generator (FRDG), showed improved performance of the obtained structural models when compared to the experiment.

In the context of random displacements, a naive approach is given by simple displacement of adjacent layers with constant shifts in random directions. Models for theoretical description of unordered structures in the solid state employ the super-cell approach. Doing so guarantees some sort of entropic behavior, however constructed systems still have to fulfill PBCs. This is also the case when structural shifts are considered and can be accounted for, by use of an

5 Fourier Random Displacement Generator

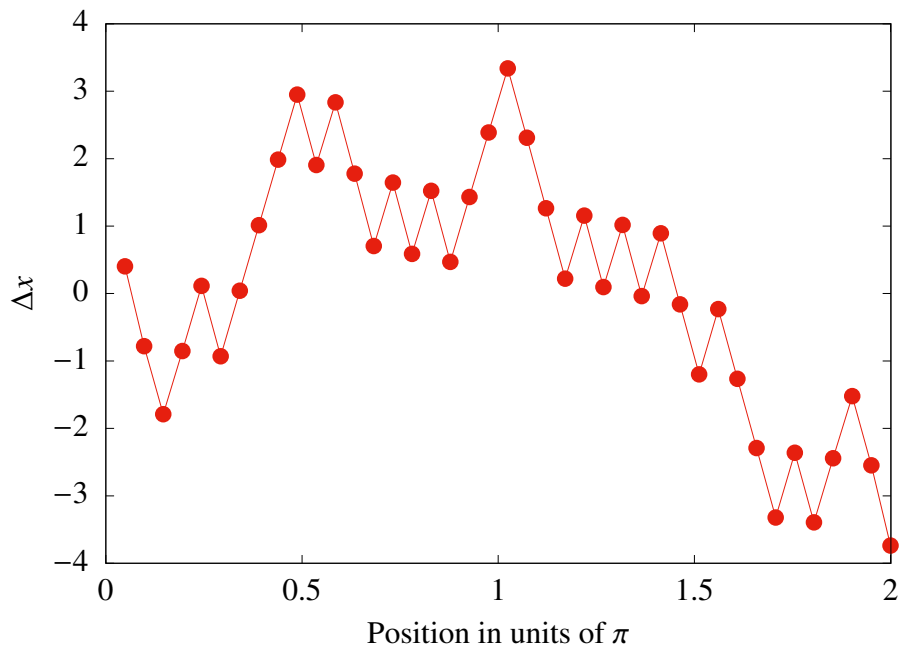
inherent symmetry (statistical) in the choice of directions. For clarification it has to be noted that large displacements over the boundary do not prohibit a periodic treatment. However, such systems might be no longer suited, due to unwanted large displacements over the boundary. An illustrative example, is shown in Figure 5.1a, for a one dimensional displacement path. A program for generation of random periodic displacement paths should therefore satisfy the following set of conditions.

- Conditions:
 1. Automatized random displacement generation.
 2. Continuity of displacements. (Regulated within a certain boundary)
 3. Avoidance of statistically equilibrated choices of displacement direction, without manual introduction of artificial weights.
 4. Continuity of displacements over the PBCs.

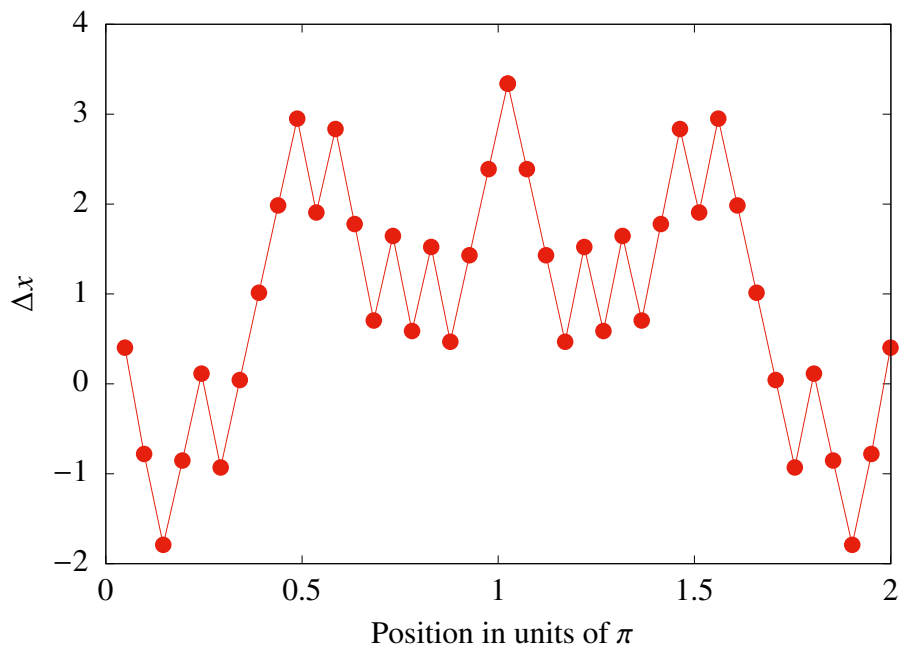
Naive approaches include for example:

- Random choice of displacement-direction. → on average equally distributed directions.
- Random displacement until the middle of the cell and mirroring the pattern to fulfill the boundary continuity. → Introduction of an inherent symmetry.
- Restricting directions already chosen, in order to avoid preferred displacement in one direction, thus avoiding large steps over the boundary.

None of the mentioned approaches offers a possibility of creating vastly differing displacement patterns for periodic systems with continuity at the boundary. Another way of securing these conditions while offering the maximum amount of automation and random characteristics possible is of interest for the “systematically” working investigator.



(a)



(b)

Figure 5.1: Examples for possible approaches for random path generation. Ordered displacement a) without continuity condition at the boundary, b) with continuity at the boundary due to inherent symmetry.

5.1 The Initial Guess

When treating periodic systems, the Fourier expansion offers a method of describing discrete periodic functions via a set of complex coefficients. Any 2π -periodic function with N complex data points a_j , is described by its complex Fourier coefficients via

$$\hat{a}_k = \sum_{j=0}^{N-1} e^{-2\pi i \frac{kj}{N}} a_j \quad . \quad (5.1)$$

Written in matrix-vector form this leads to the transformation rule for an arbitrary complex vector $\hat{\mathbf{a}}$:

$$\hat{\mathbf{a}} = \mathbf{W} \mathbf{a}, \quad (5.2)$$

where \mathbf{W} is the matrix of the Fourier transform, with elements $W_{k,j} = e^{-2\pi i \frac{jk}{N}}$. Reversing the transformation, thus, leads to the construction of a complex periodic function from its coefficients. The back-transformation is given as

$$a_k = \frac{1}{N} \sum_{j=0}^{N-1} e^{2\pi i \frac{kj}{N}} \hat{a}_j \quad (5.3)$$

A special case is given when describing a real valued function where the coefficients fulfill:

- $a_0 \in \mathbb{R}$
- $\hat{a}_i = \hat{a}_{N-i}^*$ for $i > 0$

where $*$ signifies the complex conjugate. The resulting real valuedness of the generated function, results from the coefficients exactly canceling out their complex arguments in Eq (5.3). Therefore, choosing a set of N random values for the components of the complex valued coefficients \hat{a}_k , yields N discrete points of a real valued periodic function distributed on the interval $[0, 2\pi)$. Combination of these manipulations offers a way of creating a periodic function of random shape, fulfilling the desired conditions 1. and 3. respectively. However, when choosing a small set of discrete points of the resulting function, the obtained path is not necessarily continuous. The created path, thus, has to be optimized afterwards to suit the

application at hand, securing the remaining conditions 2. and 4. respectively.

5.2 The Difference Function

For quantifying the continuity of the obtained path, a function describing the deviation from the desired case is introduced. The most important property for random displacements is given by the distance between adjacent points (the length of the displacement vector in multi-dimensional cases). Therefore, it is necessary to optimize the inter-point distances along the path, including the periodic connection. In one dimension, a way of defining a property measuring the discrepancy from the desired reference, is provided by a harmonic "potential". In analogy to a spring like model, a force pulling the points into the desired arrangement is obtained. The potential used in the FRDG is of form

$$D(\mathbf{x}) = b(\sqrt{(x_1 - x_N)^2} - c)^2 + b \sum_i^{N-1} (\sqrt{(x_i - x_{i+1})^2} - c)^2 \quad , \quad (5.4)$$

where the extra term outside the sum assures the continuity over the PBCs. The parameter c signifies the desired distance between adjacent points on the displacement path and b the steepness of the harmonic potentials. Minimizing the function $D(\mathbf{x})$ with respect to the coordinate vector \mathbf{x} consisting of all positions x_i , results in an arbitrary periodic function with desired distances. Avoiding inherent directional adjustment allows for displacement paths with preferred directions, providing a more sophisticated methodology. An example for an optimized is shown in Figure 5.2.

5.2.1 The Difference Function in 2D

Generalization of the difference function Eq (5.4) to two-dimensional in-plane displacements, is achieved by simultaneous treatment of x_i and y_i coordinates of point i with a position vector \mathbf{r}_i . As a result, an N dimensional tuple of coordinates is summarized in \mathbf{r} . The generalization

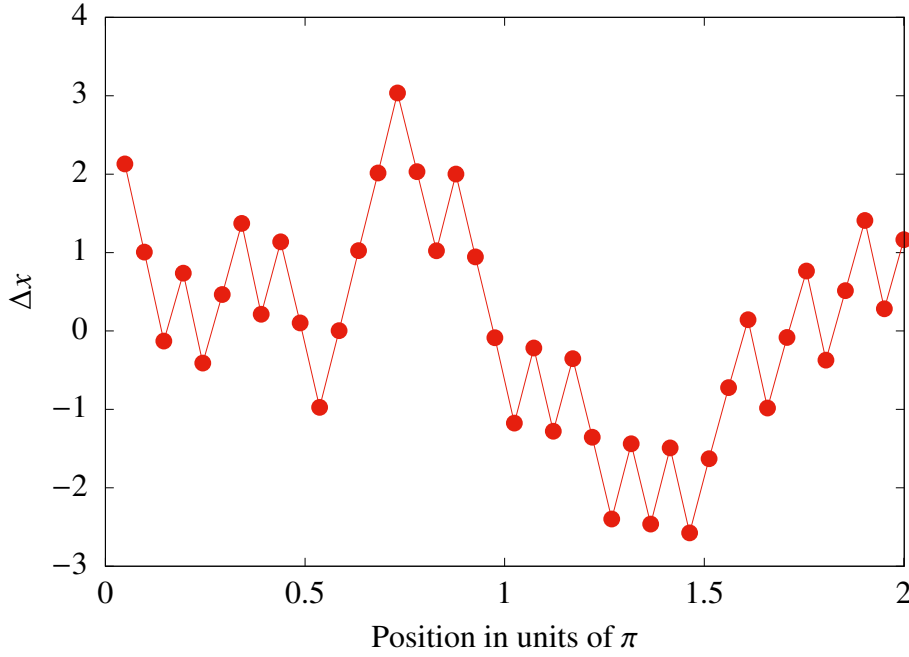


Figure 5.2: Periodic displacement with continuity at the boundary due to the utilized optimization via the displacement function Eq (5.4), for $c = 1.2$.

of Eq (5.4) to two dimensions is given as

$$D(\mathbf{r}) = b(\sqrt{(r_1 - r_N)^2} - c)^2 + b \sum_i^{N-1} (\sqrt{(r_i - r_{i+1})^2} - c)^2 \quad . \quad (5.5)$$

Here, b is the strength of the harmonic potential while the target distance is given by the parameter c . In the two-dimensional case, the number of parameters is given as $2N$, as each contributing displacement point i contributes x and y coordinates.

5.3 The Optimization Scheme

The initial values for the periodic displacement function are obtained by evaluation of an initial random guess of the complex Fourier coefficients \hat{a}_i by use of the inverse Fourier transform. Afterwards, the initial value of the difference function $D(x)$ is calculated following Eq (5.4) for the a priori obtained positions r_i , followed by its gradients and Hessians, calculated by finite differences. An iterative procedure consisting of a three-loop based optimization is used to find $\min(D(\mathbf{x}))$. Discretization of the randomly generated function might lead to

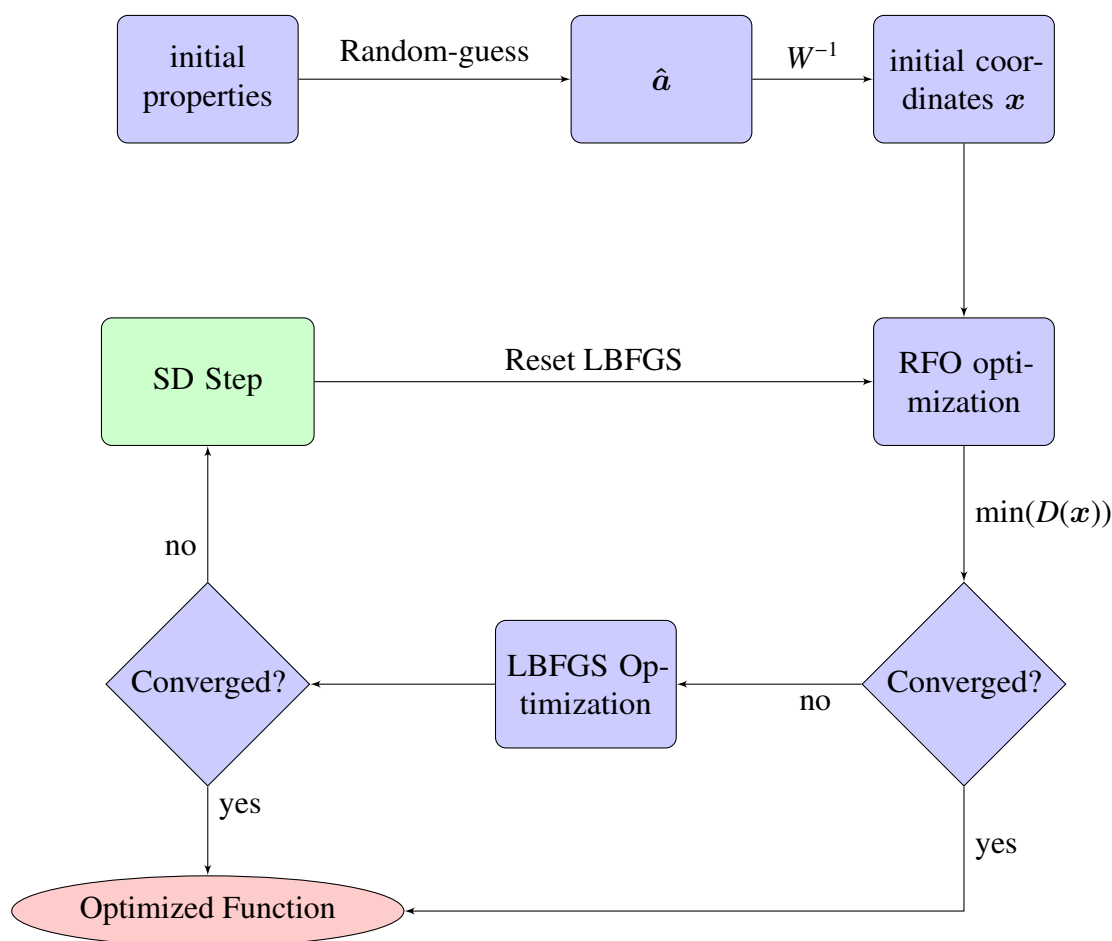


Figure 5.3: Schematic workflow of the Optimization algorithm used for creation of a periodic displacement path by minimizing the difference function $D(x)$.

values far from the desired minimum, consequently, a newton-type optimizer is used for a few initial optimization steps. The optimization is based on the RFO module as implemented in DL-Find¹⁶³. To compensate for possibly large calculation times and non-convergence, an outer loop using the LBFGS algorithm is utilized. It was observed that for a good initial guess, the LBFGS algorithm leads generally to a much faster optimization. A steepest decent step of fixed length is used in combination with a reset of the LBFGS-history, as a convergence securing mechanism to provide a systematic reset. The overall work-flow of the optimization is shown in Figure 5.3. The combination of three optimizers might seem over engineered, however, it proved necessary in one-dimensional test cases with few data points. Furthermore, the provided method becomes more suitable the more points and dimensions are used in the randomly generated path.

5.4 Results and Applications

Two-dimensional displacement functions for stacked COFs are obtained by randomly chosen Fourier coefficients for both in-plane dimensions. The initial guess for the optimization was obtained via inverse Fourier transform. These initial coordinates are then used for calculation of the difference function as given in Eq (5.5). The obtained difference function is then minimized following the scheme presented in Figure 5.3. The FRDG was successfully used to construct periodic displacements of arbitrary shape with constraints on the inter-point distances (displacement vectors). The current implementation is suited for random displacements between stacked layers as necessary for 2D-COF structures and has no known convergence issues. Exemplary in-plane displacements for a system consisting of 15 layers for COF-LZU1 as utilized for analysis of interlayer slippings on adsorption isotherms¹⁶² are shown in Figure 5.4a and Figure 5.4b.

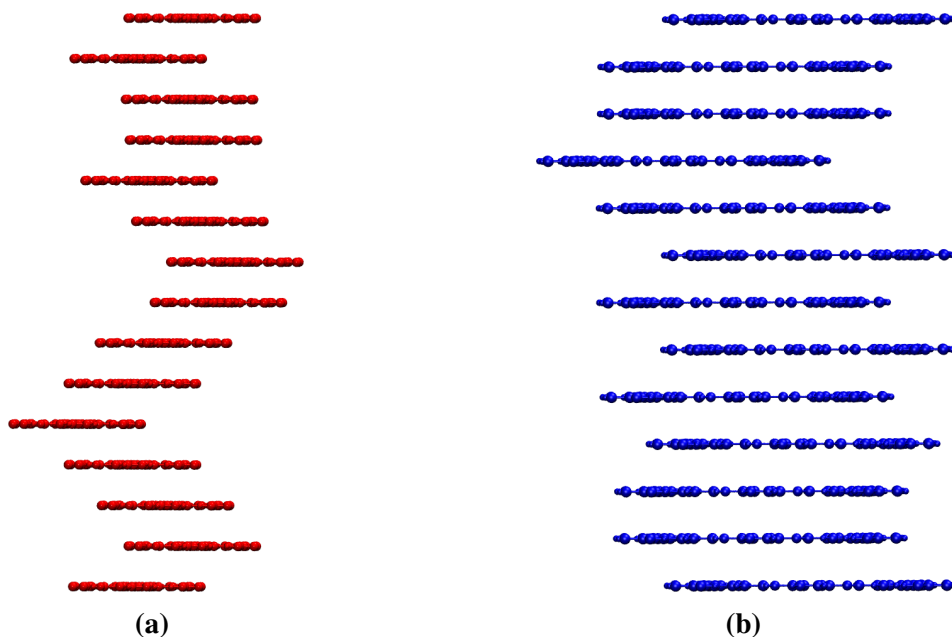


Figure 5.4: Visualization of two examples of FRDG-generated displacement paths for 15 layers of COF-LZU1. The components of the two-dimensional displacement path are illustrated with the view aligned with a) the molecular axis, b) perpendicular to the molecular axis.

Chapter 6

Charge Symmetrizer

The treatment of large crystalline super-cell structures with inherent symmetry, brings methods for averaging over atomic groups with equivalent environments, into the scope of interest. Identification of these atomic groups, requires a suitable description of the atomistic environment and appropriate conditions for their evaluation. For symmetrization of atomic partial charges, the concept of environmental descriptors obtained from spacial distribution functions is applicable. During investigations of symmetric structures by means of DFT, discrepancies arise due to imperfections in the model itself. To this end, a tool for automatic symmetrization of atomistic properties based on the identification of equivalent environments has been developed. The procedure was explicitly used for symmetrization of partial charges, however, it is also applicable for ensemble averages of equivalent atoms. The used description for atomistic environments, symmetrization work flow and possible applications of the created charge symmetrizer (CS), are presented in the following.

6.1 Working Scheme

For description of atomistic environments it is suitable to introduce a quantification for systems' similarities, based on the investigation of an atoms nearest neighbors, bond lengths and angles for each connection. An elegant way of including all this information at once is provided by so called environmental descriptors, which are denoted as \mathbf{d} in the following. A typical descriptor characterizes the environment by storing the essential information in form of a vector, built

up of several blocks of components corresponding to different environmental features. For illustration of the corresponding method of environmental description, the basic concept for construction of the specifically utilized ACU-descriptor, is explained in the following.

6.2 On the Description of Atomistic Environments

In computational chemistry, all crucial atomistic properties are determined by an atoms' environment. An easy and illustrative example is the combination of two isolated H atoms into an H₂ molecule, altering the atomic properties with respect to the isolated case, with the only difference being the presence of another H atom. Furthermore, the main environmental differences within a molecular structure can be quantified by bond profiles and nearest neighbor arrangement. The task of describing an atomistic environment is essential for comparison of similarities between systems for direct access on the deviation of properties. The simplest approach is given by internal coordinates consisting of a set of inter-atomic distances and bond lengths, bond angles and dihedrals. For the mathematical description of an environment, the chosen method has to fulfill a set of invariances, in order to be physically meaningful. Those include invariance under translation and rotation of the whole system as well as exchange invariance with respect to atoms of identical type. For methods describing individual atoms rather than whole systems, the corresponding parameters for each atom a , can be summarized into a descriptor vector \mathfrak{d}_a . Possible shape and interpretation of \mathfrak{d}_a depend on the used method of description. When treating large systems, considering the information necessary for describing the environment becomes important for the efficiency of the used method. A prominent approach is the limitation of the described vicinity, by introducing a spherical cutoff r_c around the atoms position. The remaining task, is the environmental description itself. Popular choices for construction of descriptors include for example Symmetry functions^{164,165}, Gaussian Moments¹⁶⁶, and expansion of distribution functions in a complete basis of Chebychev polynomials¹⁶⁷. In the following, the method of description via distribution functions and the corresponding expansion as proposed by Arthrit et. al.¹⁶⁷ is described, and the main features as implemented in Aenet¹⁶⁸ as well as their relevance for later application in this work are illustrated.

6.2.1 Angular and Radial Distribution Functions

Distribution functions offer a tool for mathematical description of atomistic environments. The vicinity of each atom is described by a probability distribution, containing the information of finding an atom at a specific location. To simplify the problem, a separation into an angular and radial part is straight forward. Atomic positions \mathbf{R}_j are measured in respect to the reference atoms position \mathbf{R}_i in comparison to the probability of finding an atom at the same position in an ideal gas. If the position of an atom is exactly known, the probability of finding it is given by a delta-function located at it's respective position in the distribution. This results in a radial distribution function (RDF) associated with atom i of the form

$$\text{RDF}_i(r) = \sum_j \delta(r - |\mathbf{R}_j - \mathbf{R}_i|) \quad . \quad (6.1)$$

Where r denotes the distance from atom i , and the sum is executed over all atoms j . To account for the finite extension of the atomic vicinity σ , the environment is described only within the cutoff radius R_c . Therefore, a cutoff function $f_c(r)$ is introduced, that falls off to 0 until R_c is reached. This introduces a distance related weighting with respect to the reference atom i . The chosen cutoff function is

$$f_c(r) = \frac{1}{2} (\cos(r \cdot \pi/R_c) + 1) \quad , \quad (6.2)$$

following the works of Miksch and Artrith^{169,170}. Up until now, the distribution function does not include any element specific information.

To account for different elements in the vicinity σ , an additional element specific weighting factor w_j can be introduced. Combination of all components results in the following radial distribution function for atom i

$$\text{RDF}_i(r) = \sum_{\mathbf{R}_j \in \sigma} \delta(r - |\mathbf{R}_j - \mathbf{R}_i|) f_c(r) w_j \quad . \quad (6.3)$$

For the total description the treatment of the radial part for the distribution is not sufficient, therefore an angular distribution function (ADF) is used in addition. It's definition is chosen in analogy to the radial part. The defining variable of the ADF is given as the bond angle

6 Charge Symmetrizer

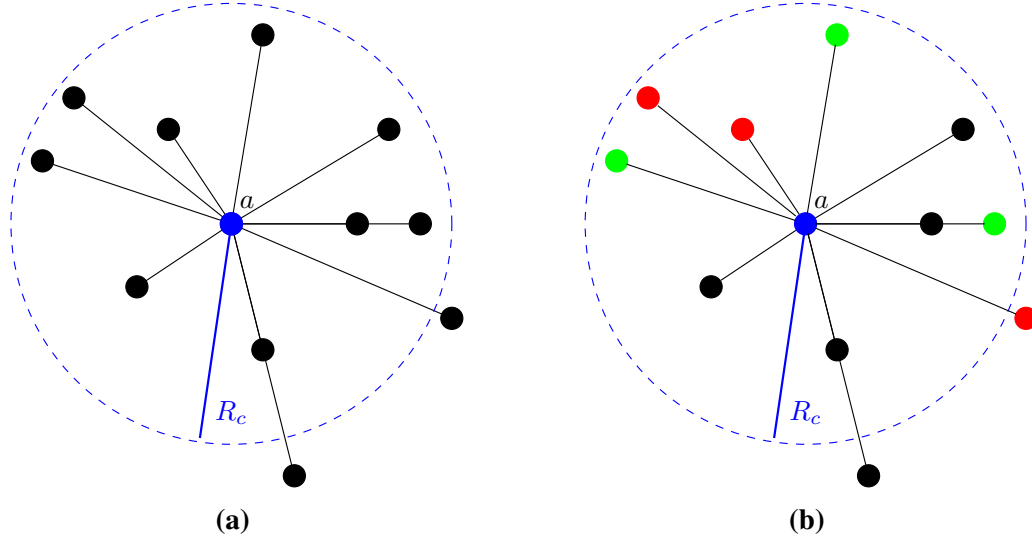


Figure 6.1: Schematic representation of the situation utilized for the atomistic environmental description by RDF. a) Only considering the atomic position of atoms within the cutoff radius R_c around the reference atom a indicated in blue. b) Considering 3 different atom types indicated by their respective colors. An analogous procedure is used for the ADF¹⁷⁰.

between three atoms of reference $\theta_{ijk} = \angle(\mathbf{R}_i - \mathbf{R}_j, \mathbf{R}_j - \mathbf{R}_k)$ and thus the resulting function is given as

$$\text{ADF}_i(\theta) = \sum_{\mathbf{R}_j, \mathbf{R}_k \in \sigma} \delta(\theta - \theta_{ijk}) f_c(R_{ij}) f_c(R_{ik}) w_j w_k \quad . \quad (6.4)$$

In the following the environment is described by expanding the distribution functions in a complete basis. To allow for a mathematically stable expansion, the distribution build up from delta peaks is convoluted with Gaussian functions. The resulting resolution is depending on the width of the Gaussians used for the convolution as can be seen in Figure 6.2. Due to the resulting distribution function being smoothed it is now suitable for a numerically stable basis set expansion.

The expansion of the RDF for atom i in terms of basis functions $\phi_\alpha(r)$, can be written as

$$\text{RDF}_i(r) = \sum_{\alpha}^N c_{\alpha} \phi_{\alpha}(r) \quad , \quad (6.5)$$

with the maximum number of basis functions N . The expansion coefficients c_{α} can be

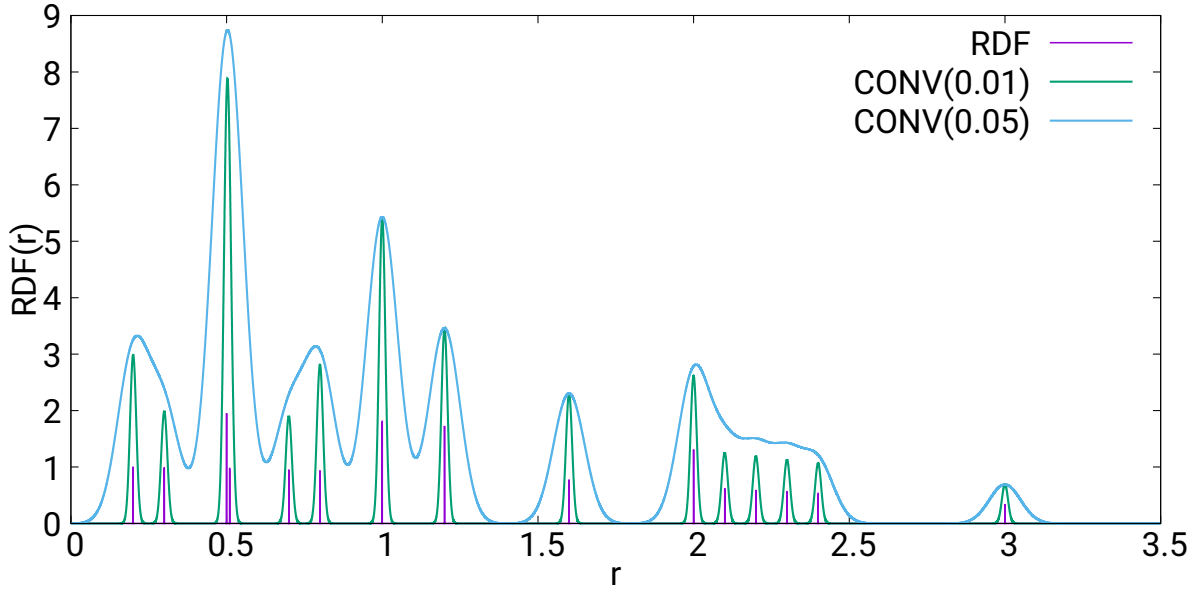


Figure 6.2: Exemplary arbitrarily chosen radial distribution function considering atoms at the delta-functions distances indicated in purple. The convolutions with a Gaussian function of width 0.01 and 0.05 are indicated in green and blue respectively.

determined by

$$c_{\alpha}^i = \int_0^{R_c} \bar{\phi}_{\alpha}(r) \text{RDF}_i(r) dr \quad , \quad (6.6)$$

where $\bar{\phi}_{\alpha}$ is an element of the corresponding orthogonal dual-basis to the basis functions ϕ_{α} . These fulfill the relation:

$$\int \bar{\phi}_{\beta}(r) \phi_{\alpha}(r) dr = \delta_{\alpha,\beta} = \begin{cases} 1 & \alpha = \beta \\ 0 & \text{otherwise.} \end{cases} \quad (6.7)$$

The RDF with cutoff function as written in Eq (6.3) only differs from zero in the interval 0 to R_c , therefore the integrals boundaries in Eq (6.6) can be expanded to the whole space. With the expression of the RDF for atom i , the coefficients c_{α}^i are given as

$$c_{\alpha}^{\text{RDF},i} = \sum_{\mathbf{R}_j \in \sigma_i} \int \bar{\phi}_{\alpha}(r) \delta(r - |\mathbf{R}_j - \mathbf{R}_i|) f_c(r) w_j dr \quad . \quad (6.8)$$

Executing the integration over the delta functions and rewriting $\mathbf{R}_{ij} = |\mathbf{R}_i - \mathbf{R}_j|$ results in the

6 Charge Symmetrizer

simple expression for the coefficients of form

$$c_{\alpha}^{\text{RDF},i} = \sum_{\mathbf{R}_j \in \sigma_i} \bar{\phi}_{\alpha}(\mathbf{R}_{ij}) f_c(\mathbf{R}_{ij}) \mathbf{w}_j \quad . \quad (6.9)$$

The form of the ADF can be derived in analogous manner and is obtained as

$$c_{\alpha}^{\text{ADF},i} = \sum_{\mathbf{R}_j \mathbf{R}_k \in \sigma_i} f_c(\mathbf{R}_{ij}) \bar{\phi}_{\alpha}(\theta_{ijk}) f_c(\mathbf{R}_{ik}) \mathbf{w}_j \mathbf{w}_k \quad . \quad (6.10)$$

As a possible choice for the basis functions, Chebychev polynomials are used in this work. They are defined by the recursion relation

$$\begin{aligned} T_0(x) &= 1 \\ T_1(x) &= x \\ T_{n+1}(x) &= 2xT_n(x) - T_{n-1}(x) \quad . \end{aligned} \quad (6.11)$$

The polynomials fulfill the orthogonality condition on the interval $[-1,1]$. The orthogonality condition is given by the relation

$$\int_{-1}^1 T_{\alpha}(x) T_{\beta}(x) \frac{dx}{\sqrt{1-x^2}} \quad , \quad (6.12)$$

which can be adjusted to the problem at hand, by modifying its applicability onto the interval $[0, R_c]$. This is accomplished by using radial basis functions ϕ_{α} of form

$$\phi_{\alpha}(r) = \frac{k}{2\pi \sqrt{\frac{r}{R_c} - \frac{r^2}{R_c^2}}} T_{\alpha}\left(\frac{2r}{R_c} - 1\right) \quad (6.13)$$

with $k = 1$ for $\alpha = 0$ and $k = 2$ otherwise. The corresponding dual basis is given as

$$\bar{\phi}_{\alpha}(r) = T_{\alpha}\left(\frac{2r}{R_c} - 1\right) \quad . \quad (6.14)$$

A detailed discussion about the Chebychev polynomials and the corresponding series was provided by Mason¹⁷¹. An expression for angular basis functions can be derived in analogous

6.2 On the Description of Atomistic Environments

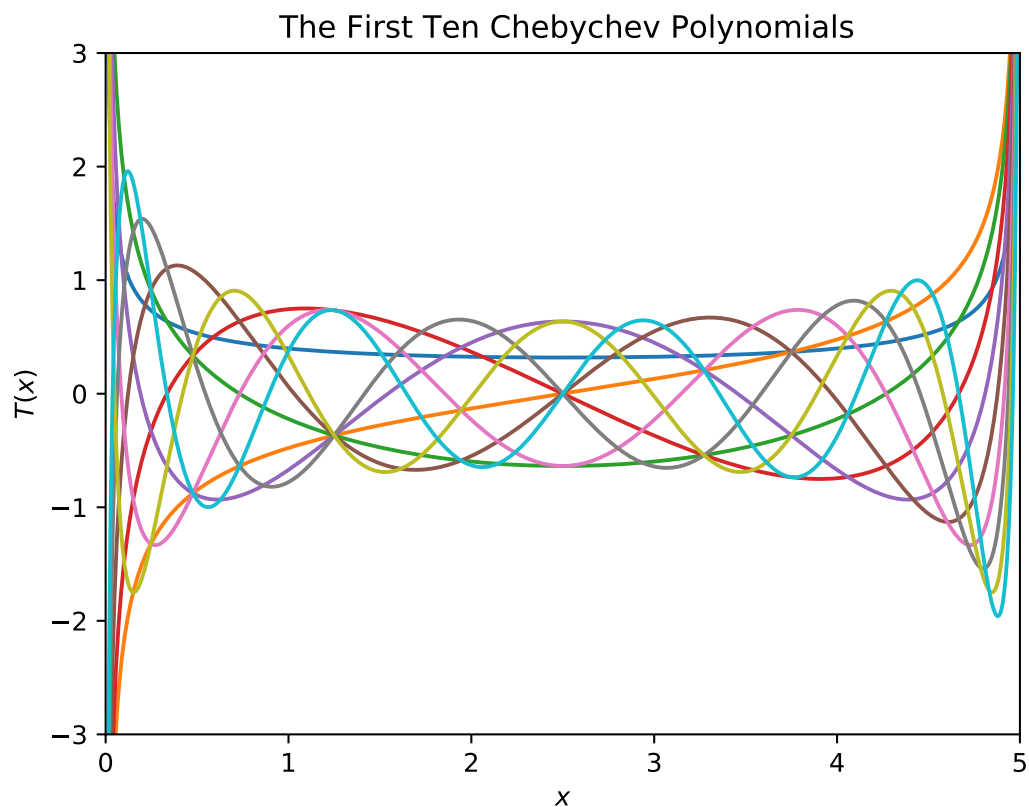
manner with the variable being θ_{ijk} and the corresponding interval being $[0,\pi]$, meaning that R_c in Eq (6.13) and Eq (6.14) has to be replaced by π . An example of the expanded function in Chebychev polynomials, can be seen in Figure 6.3b. If for every individual atom the distribution function is expanded with polynomials up to order N , each atom is also associated with N expansion coefficients. These can be summarized in a descriptor vector \mathbf{d} which therefore holds all necessary information of the environment, and if the model is adjusted properly, is sufficient to reconstruct the atomistic environment. As the expansion coefficients only depend on the shape of RDF and ADF respectively, the vector \mathbf{d} uniquely identifies an individual atoms environment. For each environment σ_i , the distribution function is unique and thus the expansion coefficients differ. Therefore they are suitable to measure similarities in atomic environments. In order to improve resolution, two sets of coefficients are calculated and used to build the corresponding \mathbf{d}_i . One part is only taking care of positions, meaning that for all elements $w = 1$ is chosen, whereas the second part also includes the element specific weighting $w \neq 1$. The resulting ACU descriptors (named after its developers Arthrit, Ceder and Urban¹⁷⁰) contains all information necessary, making the comparison of environments possible for further applications.

6.2.2 Comparison of Atomistic Environments

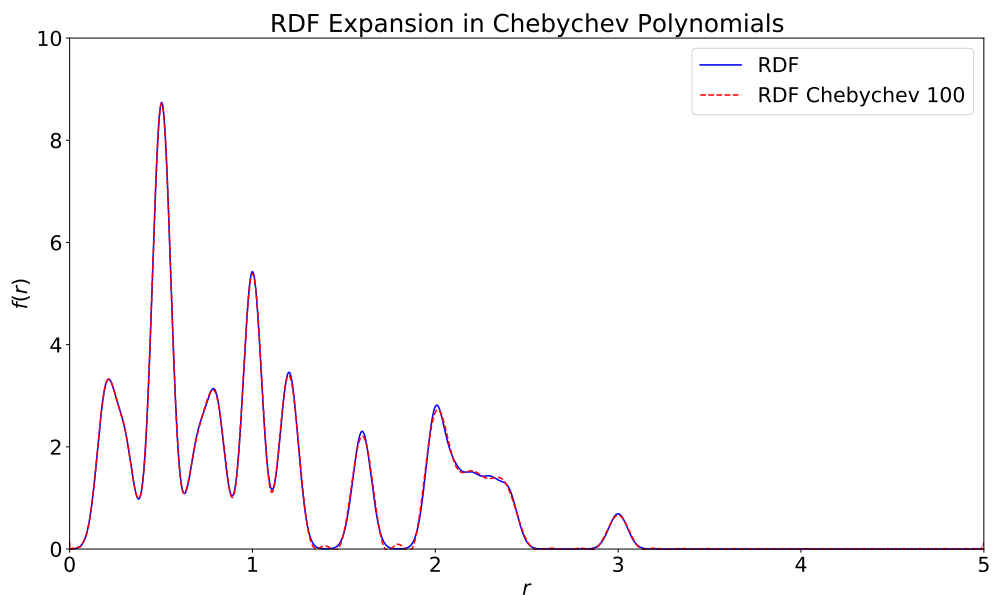
For quantification of similarities in atomistic environments, the atom specific descriptors have to be compared. Suppose two atoms a and b with their respective environmental descriptors \mathbf{d}_a and \mathbf{d}_b are studied. Equivalent environments should lead to the same expansion coefficients in Eq (6.5) and lead to similar descriptors. For a perfectly symmetric system with atoms in environments $a \in \mathcal{A}$ and $b \in \mathcal{B}$ connected via the symmetry operation \mathcal{L} as $a = \mathcal{L}b$ and a complete basis set expansion $N \rightarrow \infty$, both atoms should have exactly the same descriptors and therefore satisfy the condition

$$\mathbf{d}_a - \mathbf{d}_b = 0 \quad . \quad (6.15)$$

However this is not realizable due to truncation of the expansion Eq (6.5) at finite N and imperfections in numerically optimized symmetric structures. It is therefore expectable, that



(a)



(b)

Figure 6.3: Overview of the methods for expansion of the convoluted RDF a) The first ten Chebychev polynomials renormalized onto the interval $[0, 5]$, b) Expansion of the convoluted radial distribution function with 100 polynomials.

in real applications Eq (6.15) will never be fulfilled. The difference of two descriptors will always be nonzero. This circumstance can be used for finding a weaker criterion for similarity that is actually more suitable for real applications. For comparison, the length of the difference vector \mathbf{r} is compared to a reference value r_{ref} determining the tolerance for deviation from the ideal case Eq (6.15). This results in the similarity condition

$$\mathfrak{R}(a, b) : r_{\text{ref}} \geq r_{a,b} = |\mathbf{d}_a - \mathbf{d}_b| \quad . \quad (6.16)$$

If the size N of the basis set as well as the threshold r_{ref} for the similarity condition for the descriptors \mathbf{d} are carefully adjusted, the desired correlation between vicinities can be tuned at will. It has to be noted, that the relation connecting the imperfect case as given in Eq (6.16), does not necessarily fulfill all three conditions of an equivalence class. Suppose that an equivalence is given as \mathcal{S} over the relation \mathfrak{R} .

- Equivalence Properties

1. Reflexivity: $\mathfrak{R}(a, a) \in \mathcal{S}$
2. Symmetry: $\mathfrak{R}(a, b) \in \mathcal{S} \Leftrightarrow \mathfrak{R}(b, a) \in \mathcal{S}$
3. Transitivity: $\mathfrak{R}(a, b) \in \mathcal{S} \text{ and } \mathfrak{R}(b, c) \in \mathcal{S} \rightarrow \mathfrak{R}(a, c) \in \mathcal{S}$

In the following the terminology equivalence class is used as a synonym for classes of similar atoms and, mathematically, only refers to a tolerance class. The underlying reason is that Eq (6.16) only satisfies conditions (1) and (2) but not necessarily (3). In most investigated systems for the obtained set of atoms, the differences $r_{a,b}$ for atoms differ largely if they don't belong to the same class. For the discrete set of \mathbf{d} , the equivalence conditions are therefore quasi fulfilled if the threshold r_{ref} is tuned accordingly. However, formally this can not be guaranteed in general. The terminology is used anyway as it indicates the desired purpose of application. In the perfect case, the necessary condition could be replaced by Eq (6.15) and atoms would be equivalent in the mathematical sense. In the following the main focus is put on identifying classes of equivalent atoms in a given super cell structure. The used methodology has to fulfill the following set of main aspects

- Identification of similar bonding profiles.
- Special weighting of the environmental atom types.

6 Charge Symmetrizer

This means that two atoms with similar bonding profile, further consisting of the same atomic types, are considered equivalent in the sense of Eq (6.16). In this way, each descriptor is compared to all others. If two atoms a and b are considered equivalent they form an equivalence class \mathfrak{F} consisting of $\{a, b \mid \mathfrak{R}(a, b)\}$ that is extended by all atoms that are equivalent (in the sense of Eq (6.16)) to the already assigned ones. The assigned atoms are then excluded from further comparison and the process is continued until all atoms have been uniquely assigned to their respective classes. This leads for the connection $\cup_i \mathfrak{F}_i = \mathfrak{T}$ for the total set of atoms \mathfrak{T} . Here it has to be noted, that transitivity of Eq (6.16) is not necessarily fulfilled, however, it can be adjusted accordingly. As a result of the division in equivalence classes, the whole system of interest can be split up to the investigators desire. In the following, the explicit application for identifying atoms according to symmetry, will be illustrated at the example of atomic partial charges. The tool at hand highlights atoms of each equivalence class and allows for charge equilibration by averaging over the obtained \mathfrak{F}_i and repartitioning onto all associated atoms. The total charge for an equivalence class is given by

$$Q(\mathfrak{F}_i) = Q_i = \sum_j q_j \quad \forall j \in \mathfrak{F}_i \quad . \quad (6.17)$$

Suppose that, the total number of elements forming one equivalence class \mathfrak{F}_i is given as n_i , then the averaged charge for each atom j of the respective class is given as

$$\bar{q}_j = \frac{Q_i}{n_i} \quad , \quad (6.18)$$

forming the equilibrated charges for \mathfrak{F}_i . Applying this procedure to all classes i yields the charge equilibrated system. The overall working scheme of the symmetrizer is illustrated in Figure 6.4.

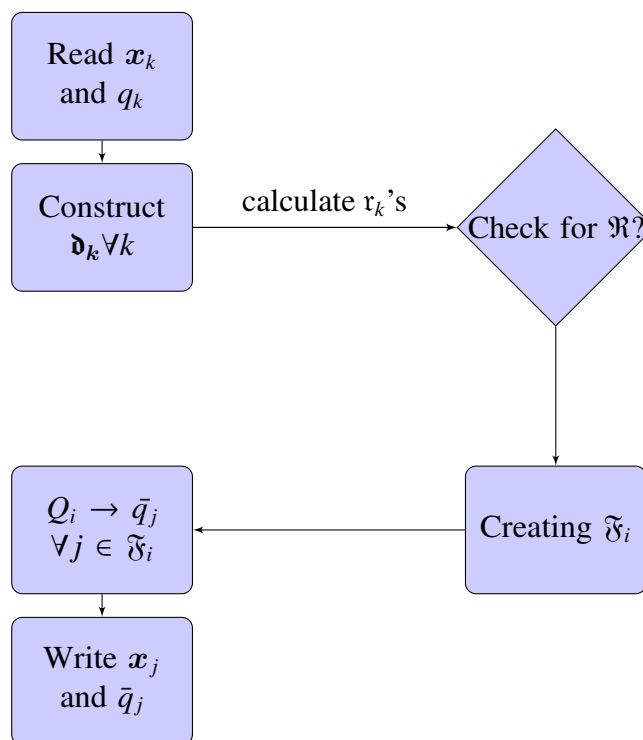


Figure 6.4: Schematic workflow of the steps in the symmetrization algorithm for determination of the charge-equilibrated system.

6.3 Exemplary Results and Applications

The successful application of the developed algorithm relies on the construction of descriptors with appropriate environmental resolution. For the utilized Chebyshev descriptors two main criteria are of significant:

- The cutoff radius R_c of f_c used for angular and radial distribution functions respectively.
- The number of basis-functions N assigned for the expansion of the radial and angular distribution functions.

Distinguishability between different elements is maximized by using different numbers of basis functions for each atomic type. Technically, this is realized by zeroing the missing elements of the smaller descriptor vectors \mathbf{d} , maximizing r in Eq (6.16). The resolution of the modeled environments, is tunable by the number of basis functions N . Therefore N has to be chosen with respect to the reference value r_{ref} , in order to determine the desired partitioning. In most observed cases, higher accuracy lead to more detailed partitioning where also smaller

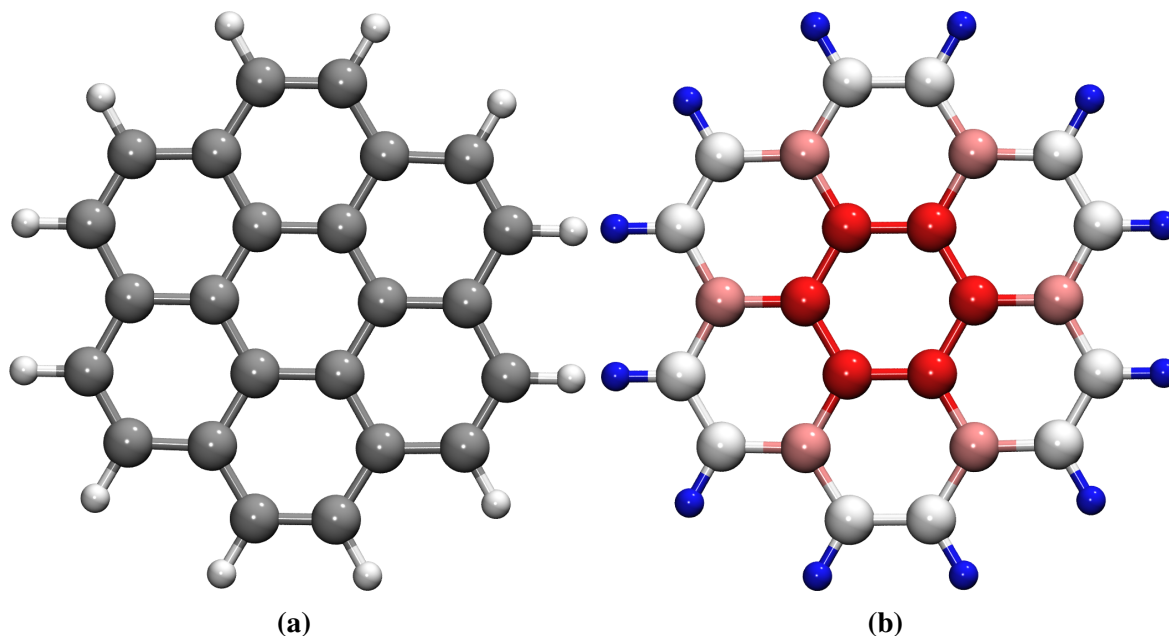


Figure 6.5: Visualization of the symmetry arrangement for the coronene molecule. a) Basic structure of coronene, carbon atoms in gray, hydrogen atoms in white. b) The four obtained equivalence classes for the coronene molecule differentiated by color.

differences are highlighted leading to increased lengths of the difference vectors r . For most applications, r_{ref} has to be adjusted accordingly, mostly yielding rather loose equivalence criteria in the sense of Eq (6.16). In the following, exemplary results are shown for molecular as well as solid state systems.

All presented molecular structures are geometry-optimized using dl-find interfaced via CHEMSHELL using TURBOMOLE for energy calculations. The later presented periodic structures were optimized using a direct cell optimization as implemented in CP2K. An example for the applicability of the described method is shown for the coronene molecule in Figure 6.5. Analysis of coronene's structural arrangement depicted in Figure 6.5a by the CS was conducted using $R_{\text{cut}} = 5.0 \text{ \AA}$ for radial and angular distribution functions as well as $r_{\text{ref}} = 1.0$. The resulting distribution of equivalent atoms shown in Figure 6.5b shows four equivalence classes consisting of six atoms each, nicely illustrating the molecules D_{6h} symmetry. Coronene acting as a simplistic example, illustrating the adaptability of the used description method. A single nitrogen atom is introduced in the inner ring (in the following, this molecule is referred to as N-coronene), breaking the molecules D_{6h} symmetry as illustrated in Figure 6.6, with the resulting N-coronene structure shown in Figure 6.6a. The CS analysis

6.3 Exemplary Results and Applications

with $R_{\text{cut}} = 4.0 \text{ \AA}$ leads to a total of 12 equivalence classes symmetric to the mirror plane “cutting” the molecule in half as shown in Figure 6.6b. However, not all atomistic environments include the N atom prohibiting the additional particularization of the whole molecule. Increase of the cutoff value to $R_{\text{cut}} = 10.0 \text{ \AA}$ shows its influence, leading to a complete particularization as the N-atom is now included within all atomistic environments σ_i . The resulting color pattern indicates an in-plane σ symmetrical arrangement. Atoms on the mirror plane form their own equivalence class, whereas all others are part of the mirror classes consisting of two elements each. The resulting groups are shown in Figure 6.6c showing the additional partitioning.

6 Charge Symmetrizer

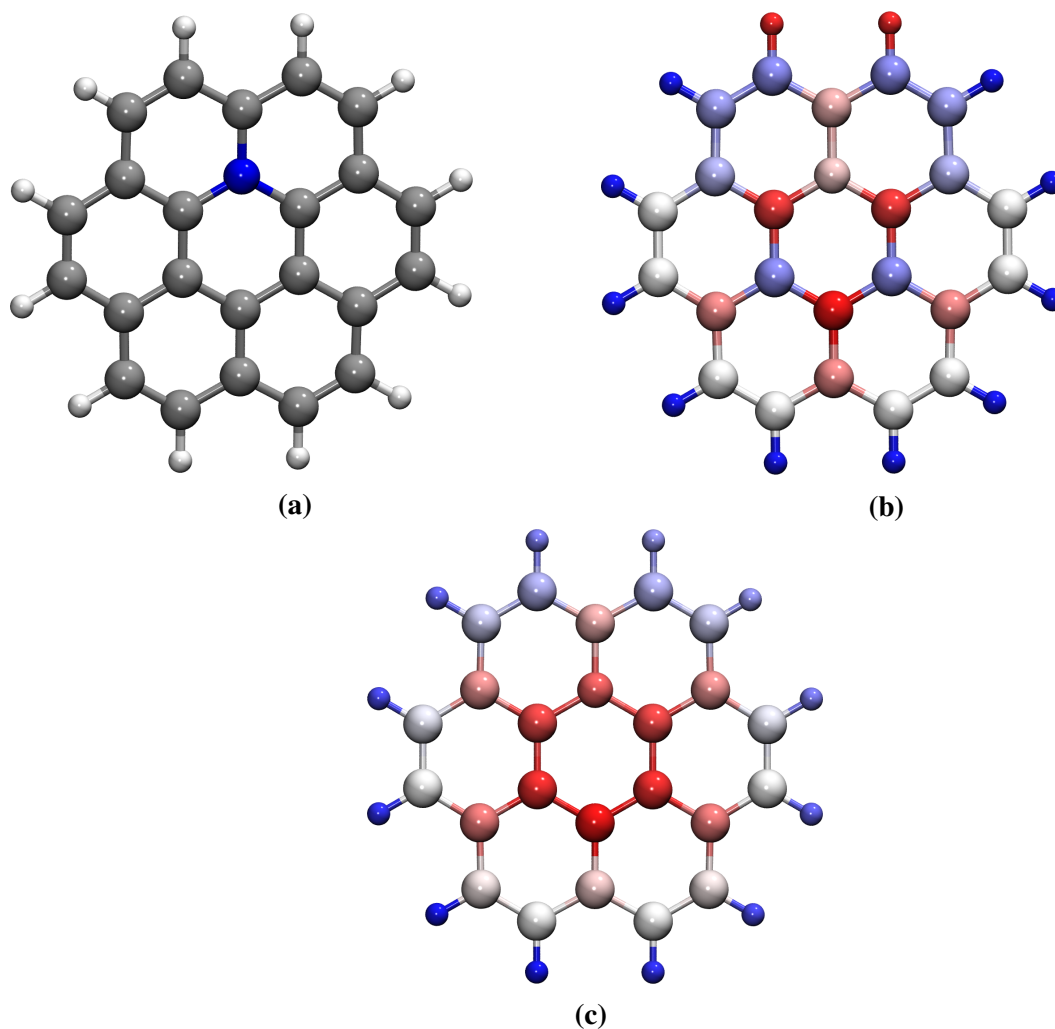


Figure 6.6: Illustration of the CS treatment for the N-coronene molecule. a) Basic structure of N-coronene, carbon atoms in gray, hydrogen atoms in white and nitrogen atoms in blue b) Equivalence classes for the N-coronene molecule differentiated by color. Using a value of $R_{\text{cut}} = 4.0 \text{ \AA}$ resulted in a total of 12 equivalence classes. c) Increase to $R_{\text{cut}} = 10 \text{ \AA}$, resulting in a total of 20 equivalence classes.

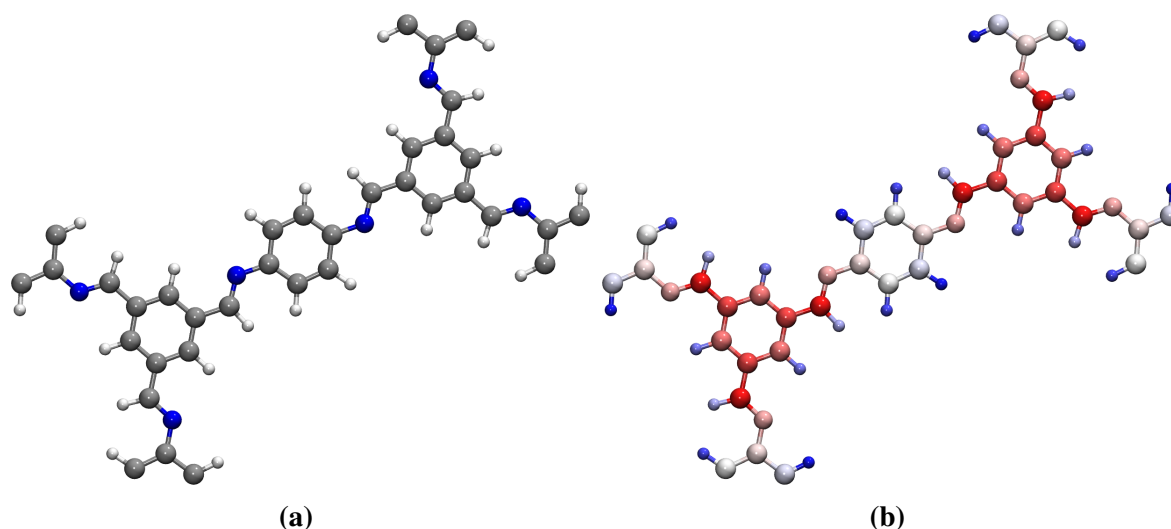


Figure 6.7: Visualization of the symmetry arrangement for the COF-LZU1 unit-cell. a) Basic structure of the COF-LZU1 unit-cell, carbon atoms are indicated in gray, hydrogen atoms in white and nitrogen atoms in blue. b) The eleven equivalence classes for the COF-LZU1 unit cell differentiated by color.

6.4 General Applicability

The previously described treatment, as demonstrated for isolated molecules, can also be applied to periodic systems. In order to obtain suitable descriptors, the underlying distribution functions are determined via extended unit cells. Consequently boundary atoms are identified analogous to atoms closer to the unit cell's center. An illustrative example is given in Figure 6.7 where the COF-LZU1 unit cell is illustrated with respect to possible CS application. As expected, the analysis of the COF-LZU1 structure Figure 6.7a, lead to a partitioning in eleven equivalence classes consisting of six elements each shown in Figure 6.7b.

6.4.1 Multilayer Treatment

Applications requiring the treatment of multiple layers give rise to additional assumptions when identifying equivalent atoms. In the simplest case when the structures themselves are not breaking the initial symmetry, this does not lead to additional effects as for example in the case of the so called perfectly eclipsed stacking of layers shown in Figure 6.8. Here the descriptors for each atom are influenced by the adjacent layers (If the cutoff is chosen to be larger than the interlayer displacement) but the number of equivalent environments as defined in Eq (6.16) is

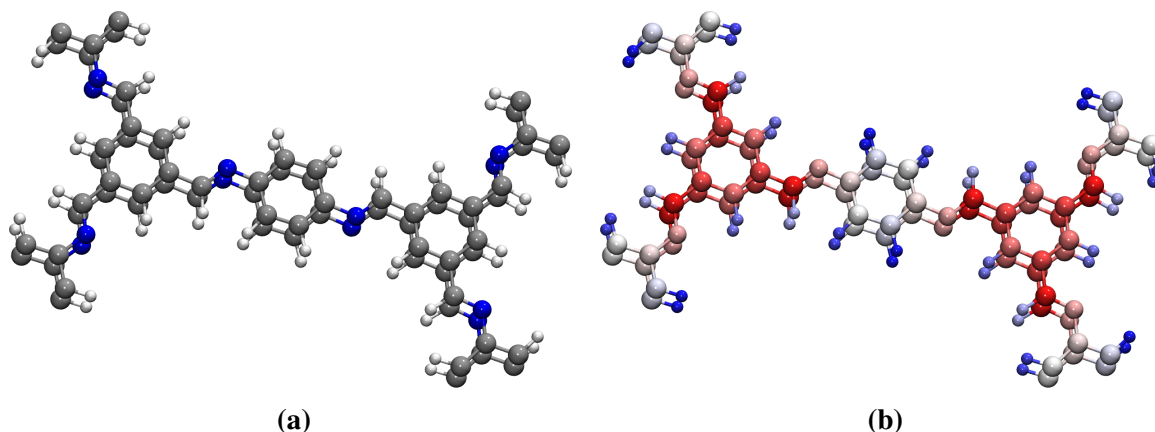


Figure 6.8: Example of the CS application on a two layered COF-LZU1 system. a) The structural arrangement of layers within the treated unit cell. b) The obtained eleven equivalence classes for the eclipsed system, differentiated by color.

not influenced. As a result, the structural arrangement shown in Figure 6.8a resembles exactly the periodic extension of one layer in stacking direction, leaving the number of equivalence classes invariant. As a result of the CS analysis, again eleven equivalence classes consisting of twelve atoms as shown in Figure 6.8b are obtained.

Realistic COF systems are not expected to be perfectly eclipsed, but rather slight displacements between the layers are expected. Considering a two layered system as shown in Figure 6.9, where the respective layers are identical but slightly shifted with respect to each other. The displaced system consists of additional environments due to interlayer positioning, and direct CS analysis results in the identification of more equivalence classes if R_{cut} is chosen large enough to include the adjacent layer. An example can be seen in Figure 6.9b where the stacking gives rise to 64 equivalence classes with only two atoms in each of them. The full partitioning is correct as the layers are identical and therefore symmetric under rotation along the axis cutting the unit cell in half. It has been shown in Section 6.4, that for each of the individual layer, only twelve different atom groups are included. In cases where a full particularization as shown in Figure 6.9b is undesired, the layers should be investigated independently. For avoidance of the inclusion of atoms from adjacent layers in the description, a naive approach would be to decrease the cutoff for the descriptors functions below the inter layer displacement value. However this comes with a loss of accuracy in the determination of equivalence classes and is ultimately only possible for systems where the correct intra-layer particularization does not require larger values of R_{cut} . An alternative way is to treat the layers

in separation when constructing the descriptors and recombining them afterwards. Within the CS-algorithm this is done for two layer systems by identifying each layer and separating them by 20.0 Å. Constructing the descriptors after separation and relocating the layers afterwards in their original position, leads to identification of equivalent groups in both layers according to Eq (6.16). The option to treat layers separately can be tuned in the programs setup script. The result of the separated multilayer treatment for the shifted COF-LZU1 unit cell are shown in Figure 6.9c where eleven equivalence classes consisting of twelve atoms each are identified.

Exemplary Application of the CS-Tool:

Charges were symmetrized for ideal initial conditions of COFs used for investigation of layer slipping on simulations of adsorption isotherms¹⁶². Furthermore, the CS was successfully used for charge symmetrization in COFs as obtained for various stages during the adsorption process of light Gases. By averaging the relative charge shifts for differently loaded frameworks for equivalent atoms, the inflicted polarization by light gas guest atoms could be investigated by a fully automatized work-flow as explained in more detail in Section 10.6.1.

6 Charge Symmetrizer

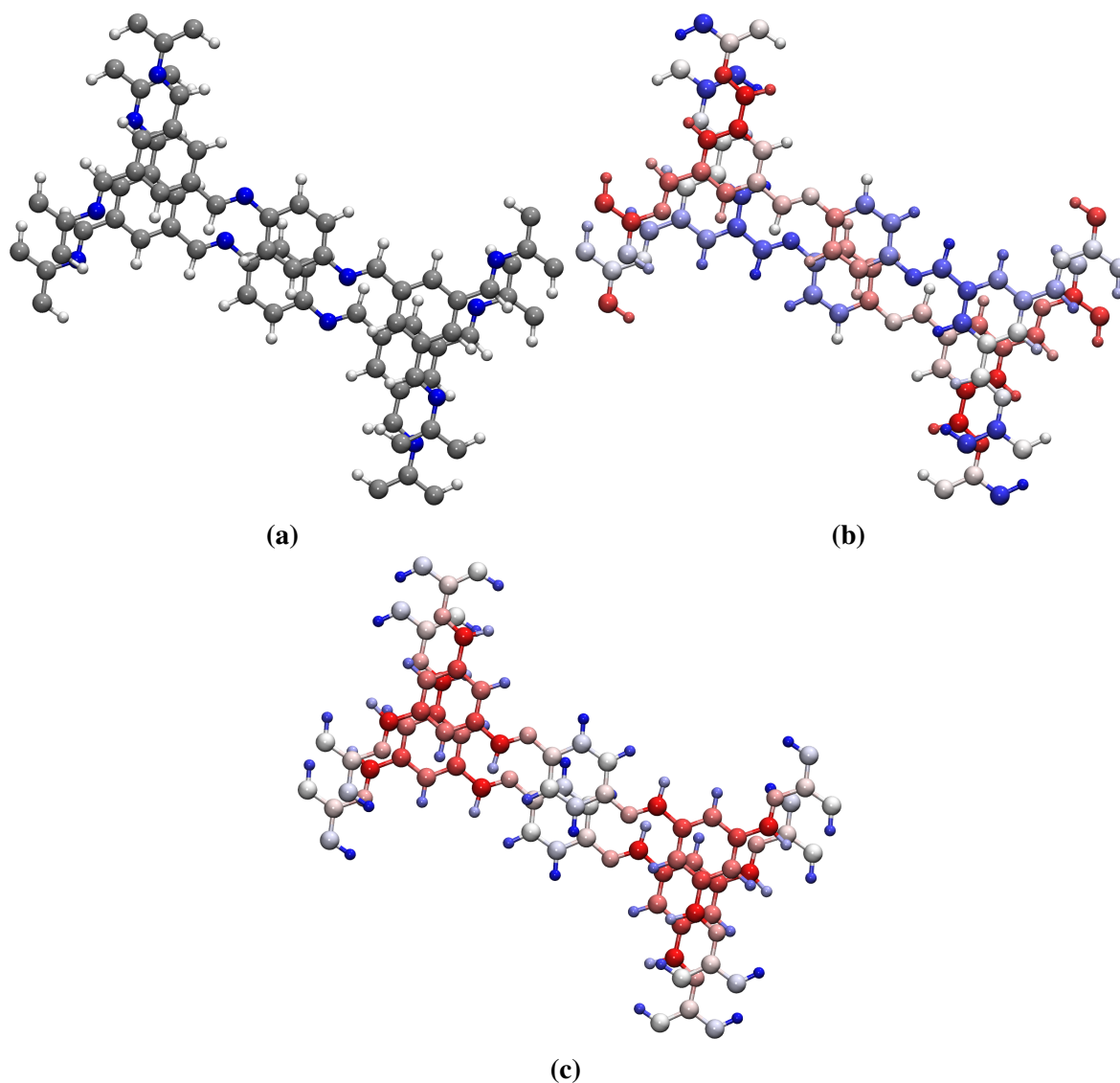


Figure 6.9: Overview of the multilayer treatment for the shifted COF-LZU1 unit cell. a) The structural alignment of the treated unit cell. b) Visualization of the symmetry arrangement for the shifted COF-LZU1 with $R_{\text{cut}} = 4.0 \text{ \AA}$, thus including the adjacent layer environments, and resulting in a total of 64 equivalence classes. c) Visualization of the equivalence classes for the shifted layers of the COF-LZU1 unit-cell after separated treatment, resulting again in the 11 equivalence classes.

Chapter 7

Pore Size Determinator

Experimentally, pore sizes are determined indirectly, as for example in standard BET gas adsorption measurements¹⁷². However, in theoretical studies they can be directly determined from the calculated pore structures. A possibility of defining pore sizes is given by the smallest distance between atoms within the pore walls. This procedure offers a way of “exact” determination, as the coordinates are directly accessible and results in a precise estimate, provided the treated structures are in accordance with the real systems of interest. To this end, an algorithm for automatic determination of pore sizes was developed. The definition of the pore size as well as the methodology used in the pore size determinator’s (PSD) work flow is described in the following. The chapter closes with an example treatment as utilized for the theoretical analysis of covalent organic frameworks (COFs), within this thesis.

7.1 The Definition of the Pore Size

For the determination of pore sizes, firstly a suitable definition of the property itself has to be found. Specifically for the treatment of porous media, there are multiple possible definitions available, depending on the complexity and accuracy of the problem at hand. For applications in chemistry, the size of a pore is related to the largest molecule that is able to pass through it. Suppose one layer of a pore as exemplary shown in Figure 7.1. A straight forward geometrical definition of the pore size is the diameter d of the largest circle fitting into given structure. For an ideally constructed system, the pore center can be placed in the origin, where the distance

7 Pore Size Determinator

of each atom to the center is calculated and the smallest determines the pore size. The pore "surface" A can be estimated by the radius r_{\max} determined from the minimum distance from the pore center. The corresponding accessible surface area is given by

$$A = \pi r_{\max}^2 \quad . \quad (7.1)$$

Subsequently this estimation can be modified by comparison with experiments, for instance by chemically motivated criteria e.g. inter-atomic repulsion considering the finite extension of particles as well as their charges. Inter-atomic repulsion limits the possibilities of atoms to approach closer than the vdW radii r_{vdW} and therefore reduces the effective pore size. Alternatively the pore size could be estimated by precisely sampling the edge of the pore walls.

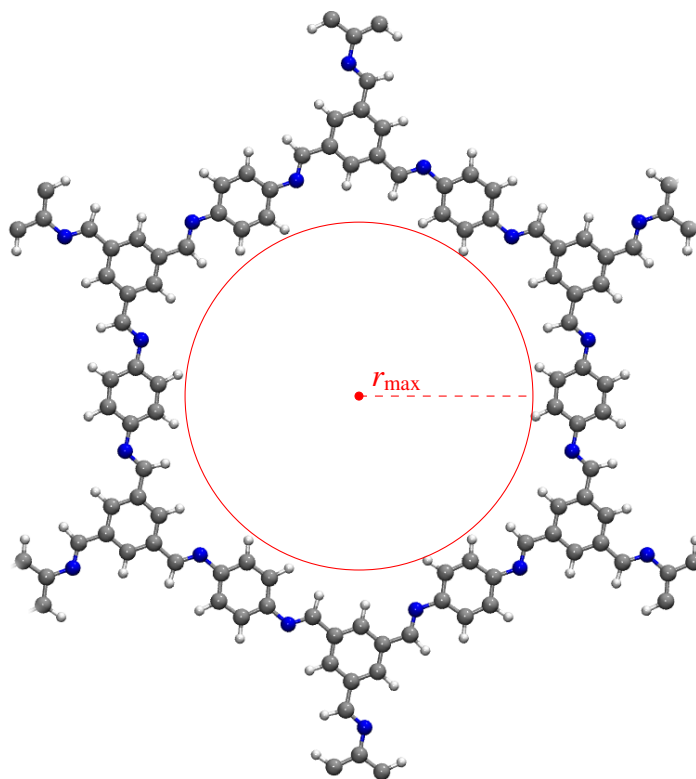


Figure 7.1: Schematic representation of the described pore size definition, the red circle indicates the pore surface resulting from the circle with largest radius r_{\max} fitting inside the pore.

This definition of the pore size is directly applicable for two-dimensional systems, however its suitability for models of real pores in three dimensions is limited. The definition has to be modified in accordance with the desired application.

In Figure 7.2 a molecule passing through a two-dimensional layer is schematically depicted, however real pores are three-dimensional systems and the simple definition is suitable only for perfectly eclipsed stacking types. An example is the treatment of multiple shifted layers as shown in the work of Emmerling et. al.⁶⁰. In three dimensions, a molecule does not only have to fit through the in-plane pore spanned by the respective layer, but can also be blocked by atoms with smaller distances from to adjacent ones. The situation is schematically presented in Figure 7.3 and motivates the modification of the current definition. The first step for improving the treatment is the inclusion of atoms from adjacent layers up to a relevant distance depending on the investigated system.

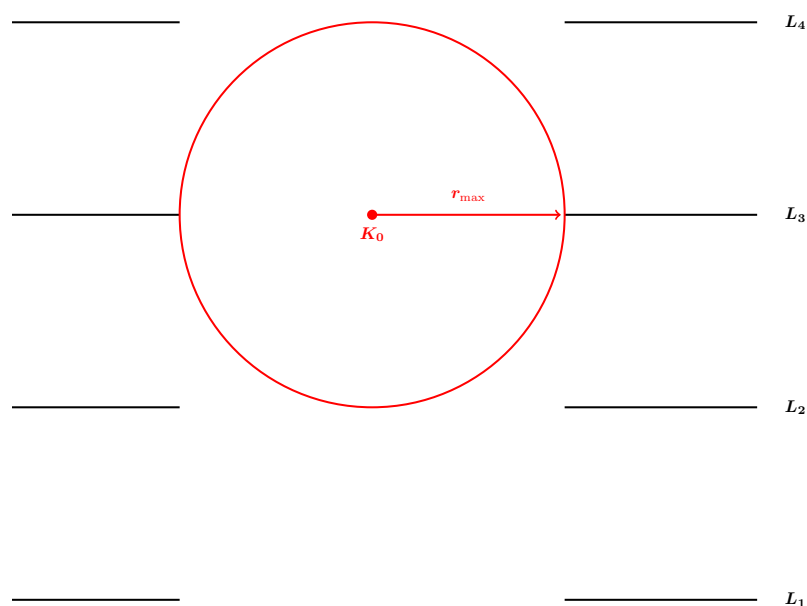


Figure 7.2: Scheme of an eclipsed stacking type, where the simplest definition of the pore size indicated by the red circle with radius r_{\max} around the pores center is still applicable.

Most possible guest molecules will not be planar in shape and also rotate in dynamic processes. To account for possible rotations and out of plane extension the molecule can be interpreted as a sphere, with the diameter d of the biggest inter-atomic distance inside. This sphere accounts for all possible rotations for the investigated molecule that are expected to occur as an ensemble average. The maximum pore size of a shifted system is, thus, given for the biggest sphere with radius $r_{\max} = \frac{d}{2}$ that is still able to propagate through the

7 Pore Size Determinator

pore. The cross-section of the sphere can be taken as the resulting in-plane pore size and is again calculated according to the simple (Eq (7.1)). For an explicit calculation following the described scheme, a point of reference has to be defined in the initial step. In the case of two-dimensional layers as used in the case of the investigated COFs, an obvious choice is the center of a respective layers pore. Every molecule passes through the in-plane alignment, therefore marking the system's narrow point. For ideally constructed systems, the pore center can be determined by symmetry assumptions during the layer creation process. The center can then be adjusted for each layer according to its relative shift to the layer below. The maximum pore size can therefore be calculated by determination of the minimum distance of all nearest atoms, from the in-plane pore center as illustrated in Figure 7.3. This definition is insufficient when the layer shifts are larger and an in-plane relocation of the sphere becomes necessary as shown in Figure 7.4. This can be accounted for by sampling the in-plane positions and search for the characteristic maximum size of a layers pore. From the experiment it is known, that real COF systems, show arbitrary random displacements⁵⁹. It is obvious that the pore size can not be determined by a simple treatment of an exemplary arrangement (small unit cell) but rather has to be calculated for every layer respectively. The pore size can then be estimated either from the average or the minimum value of the calculated spheres radii.

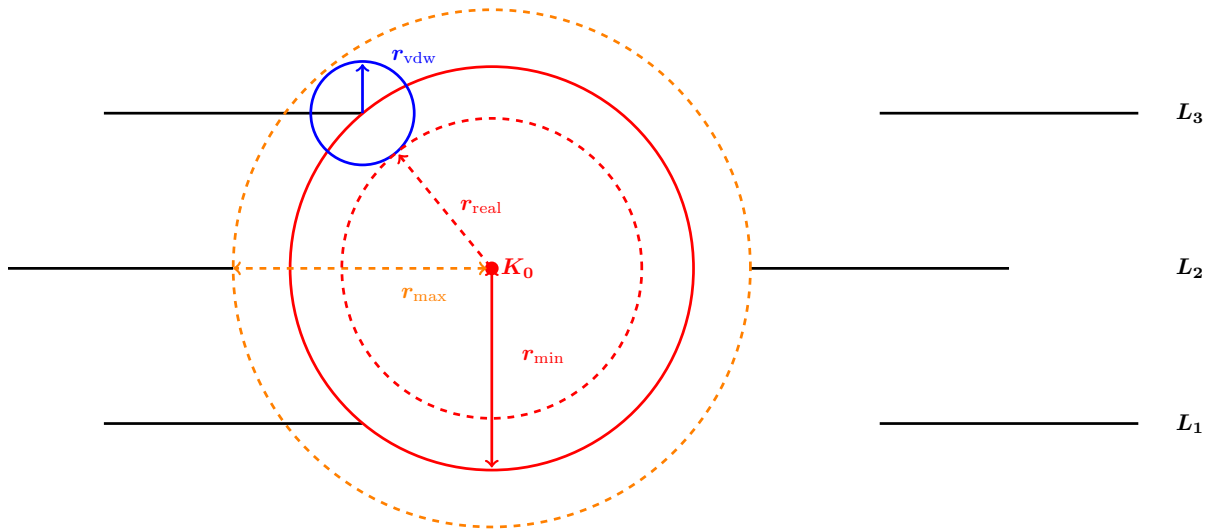


Figure 7.3: Schematic representation of the situation for shifted layers. The original pore size as resulting from eclipsed stacking is indicated by a dashed orange circle. The geometric pore size for shifted layers, considering point like atoms, is indicated by a red circle with radius r_{\min} . The even smaller effective pore considering the atoms repulsion of radius r_{real} , is drawn in dashed red with the van-der-Waals radius r_{vdw} around the edge atoms indicated in blue.

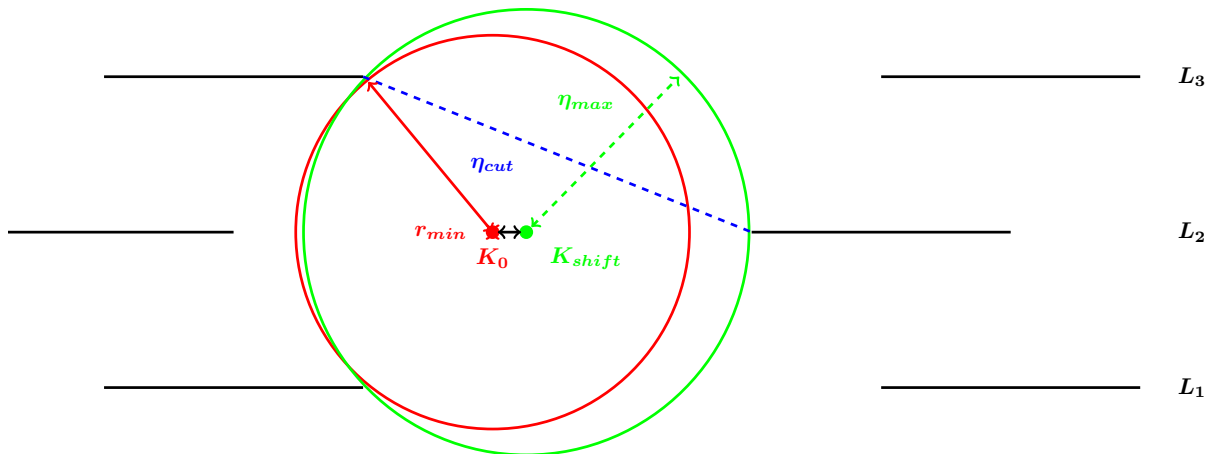


Figure 7.4: The reduced pore size resulting from shifted layers, when calculated from the pores center K_0 for L_2 is indicated in red. The corrected pore size after relocating the center of the constructed sphere to K_{shift} is represented in green. The determined maximum radius determined by consideration of the pores' center is indicated by r_{\max} . However, the maximum diameter of a sphere truly able to propagate through the whole layer arrangement, is indicated by d_{cross} .

7.2 Workflow

To determine the maximum pore size, the most accurate method presented in the previous section was chosen. To do so, the system initially gets divided into separate layers with the displacements between the layers saved into a file. Later the pore centers of the layers are reconstructed according to the displacement information. With the pore centers known, the overall distances of atoms to the center for each respective layer are calculated within the range of relevant adjacent layers defined by the user in accordance with the system at hand and the minimum distance is determined. As mentioned in the description of the possible definitions, this pore size estimation underestimates the pore size for most systems, because the center of the layer might not be the most suitable reference anymore if layers are shifted arbitrarily. Thus, a shifting of the reference point defining the sphere's center leads to a better description. To modify the choice of the reference, the vicinity around the pore center is sampled on a grid to find the most suitable arrangement. The maximum sphere is then calculated for every point on the grid, and the one with the biggest sphere radius is determined. This method is applied to all layers L_i and the resulting pore size is estimated by the sphere with the smallest radius r_{\min} . In principle the maximum radius r_{\max} resulting from the sampling process, for each layer, has to be compared to possible cross diameters d_{cross} , calculated across multiple layers as shown in Figure 7.4. However, the COF systems investigated within this thesis, are calculated without consideration of this effect, due to their large in-plane pore sizes. As mentioned beforehand, an alternative is to calculate the average pore diameter, which considers all layers. Both values are compared for an exemplary system in table Table 7.1.

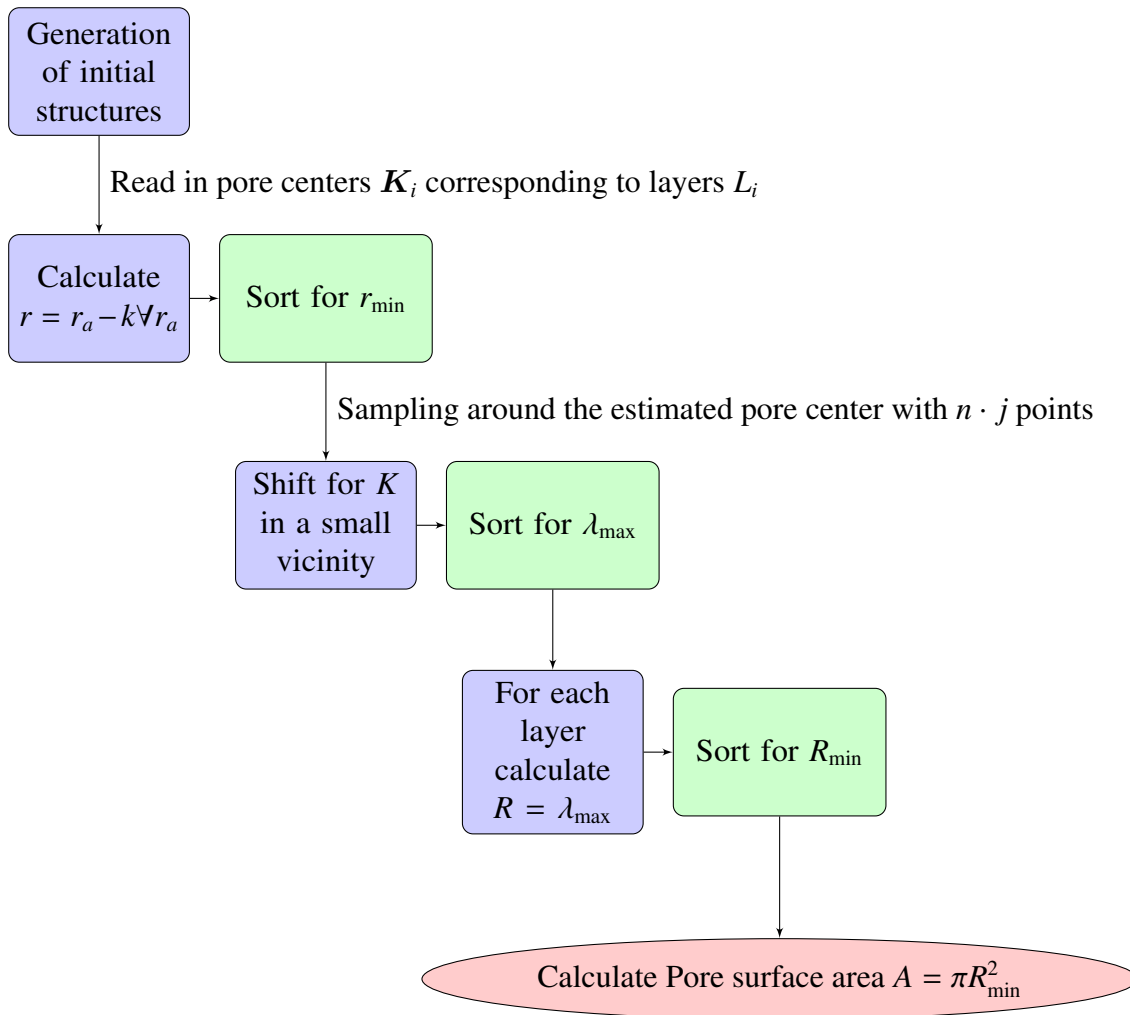


Figure 7.5: Schematic work flow of the PSD algorithm.

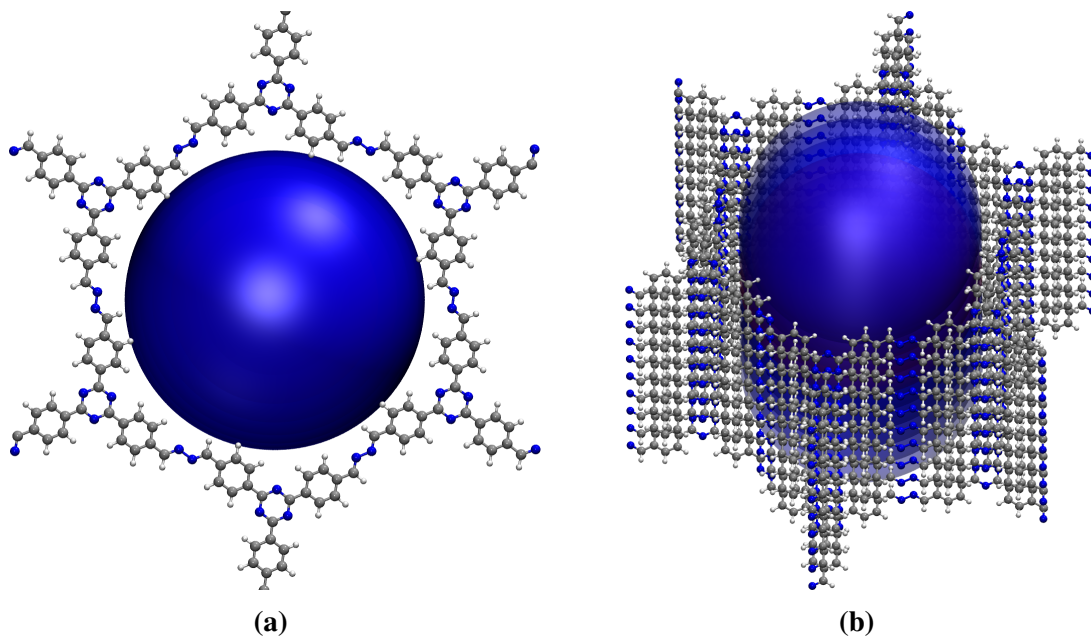


Figure 7.6: Exemplary situations of the N-COF unit cell in perfectly eclipsed stacking. a) Top view onto the unit cell, showing the sphere of a one layered pore. b) Extended system with the pore size determined for every individual layer, indicated in blue.

7.3 Exemplary Results and Applications of the PSD-Analysis

The algorithm was tested for a system of ten layers of the N3-COF, stacked without displacement. The top view in Figure 7.6a shows the exactly identical positions and acts as a representation of the in-plane circle showing the estimated maximum accessible pore surface. The arrangement can be further seen by the side view in Figure 7.6b, where the calculated spheres are shown in blue. The corresponding radii could be determined as $r_{\max} \approx 12.0 \text{ \AA}$. The maximum pore surface could be determined as $A \approx 452.28 \text{ \AA}^2$, here minimum and averaged size of the pores are identical due to the eclipsed arrangement and the identical spheres.

In Figure 7.7 the results for a pore analysis for a randomly shifted system, with the random displacements being determined by the choice of $n \in \mathbb{N}$. The displacements are given as

$$\mathbf{v}(n) = 1.5 \text{ \AA} \begin{pmatrix} \cos\left(\frac{\pi}{3}n\right) \\ \sin\left(\frac{\pi}{3}n\right) \end{pmatrix}. \quad (7.2)$$

Here \mathbf{v} is the respective in-plane displacement vector. Its value is only determined by the choice of the equally distributed random numbers $n \in 1, 2, 3, 4, 5, 6$. The total pore size according

7.3 Exemplary Results and Applications of the PSD-Analysis

to the smallest sphere with radius $r_{\min} \approx 9.61 \text{ \AA}$, for this example, could be determined to be $A_{\min} \approx 290.17 \text{ \AA}^2$. While considering the average radii of the spheres, it was determined to be $\bar{A} \approx 337.46 \text{ \AA}^2$. All values of the respective spheres' centers, radii and corresponding surface-areas are summarized in Table 7.1.

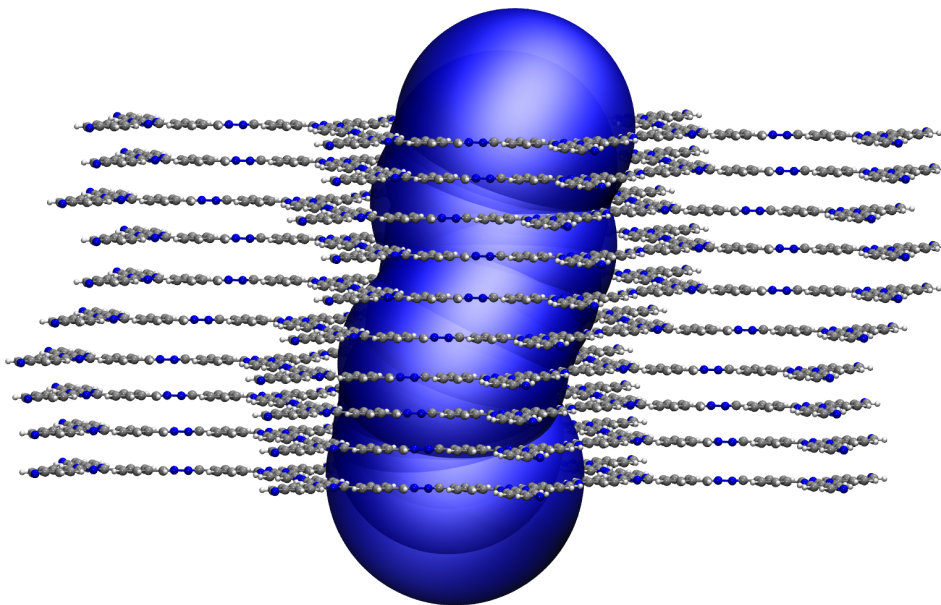


Figure 7.7: Relevant pore volume, indicated by the overlapping spheres in blue, for a system of 10 stacked and randomly displaced layers of the N3-COF.

Table 7.1: Components of the determined sphere position \mathbf{K} with respect to the pore center and maximal radii r_{\max} for each layer L_i shown in Figure 7.7 from top to bottom respectively.

| L_i | $K_x / \text{\AA}$ | $K_y / \text{\AA}$ | $K_z / \text{\AA}$ | $r_{\max} / \text{\AA}$ |
|-------|--------------------|--------------------|--------------------|-------------------------|
| 1 | 0.00 | -0.50 | 0.00 | 11.51 |
| 2 | -1.0 | -1.00 | 3.47 | 10.71 |
| 3 | -2.10 | -0.77 | 6.93 | 10.23 |
| 4 | -3.20 | -1.05 | 10.40 | 9.61 |
| 5 | -4.80 | 1.27 | 13.87 | 10.55 |
| 6 | -5.40 | 1.50 | 17.34 | 10.31 |
| 7 | -4.80 | 2.77 | 20.80 | 9.61 |
| 8 | -2.20 | 3.54 | 24.27 | 9.91 |
| 9 | -1.50 | 6.04 | 27.74 | 10.44 |
| 10 | -1.10 | 6.81 | 31.20 | 10.62 |

Applications of the PSD within this Thesis:

The designed PSD-tool was used for determination of the average pore size of FRDG generated random structures corresponding to experimentally estimated interlayer slippings of COFs. The pore sizes \mathfrak{P} were determined by averaging over all values of the respective layers and showed good agreement with the experimental estimate. The proclivity of systems towards larger layer slipping indicating staggered arrangements, showed the validity of the constructed structural model for the studied isorecticular PP-TAB and PP-TAPB-COF series⁶⁰.

Furthermore, the PSD tool is used for determination of the validity of structural arrangements determined by sampling of interlayer displacements for planar COF systems as studied in Section 10.6.1. For comparison with the experiment, various FRDG-generated random displacement paths were created for values of d . Determination of the connection for $\mathfrak{P}(d)$, allows the identification of the most probable interlayer displacement d .

Part III

Applications

Chapter 8

Investigation of Surface Models for Self-Assembled Monolayers on Metal Substrates

The following chapter is dedicated to the aspects of modeling surfaces built from thiolates on metal substrates and aims at providing information to simplify follow up interpretations within the framework of field evaporation. After giving the reader an initial overview of thiolate Self-Assembled Monolayers (SAMs) based on recent literature, the three types of SAM-molecules are investigated by means of gas phase DFT-calculations. Subsequently, the procedure for the creation of models for the metal substrates is illustrated. Starting from a reference bulk constructed from literature lattice constants, description based differences due to GPW-DFT and the periodic implementation of GFN-xTB⁶¹ within CP2K⁶², are investigated. Discrepancies between model and literature and consequences for the applicability and interpretation are shown at the example of metals used in the experiment. All investigated metals crystallize in an fcc structure. The procedure used for modeling ideal as well as reconstructed surfaces are shown for the case of gold atoms.

Based on the initial treatments of surfaces and SAM molecules, the molecule-surface interfaces is constructed. Firstly, the basic description is introduced for the smallest system, where only one SAM-molecule is explicitly treated. Due to the periodic treatment and the

GPW-DFT framework, differences to cluster simulations utilizing TURBOMOLE¹⁷³ are discussed for all investigated SAM species. Furthermore impacts of the SAM density are studied, utilizing Au, Pt and Cu surfaces, each providing different surface areas due to their respective in plane lattice vectors. In order to obtain further information about intermolecular contributions of the SAM species, systems consisting of two and more molecules per unit cell are investigated. Here, a special emphasis is taken on the differences arising from the modifications with respect to the C10 reference for F-SAM and NH₂-SAM respectively. The obtained results are compared to previously performed studies of SAMs on metal surfaces and possibilities for further improvements are discussed.

8.1 Principles of Thiolate Self-Assembled Monolayers

Self assembly describes a process of system construction without external guidance. The formation of SAMs from organic molecules offers a general way of surface construction with easily tunable properties such as thickness, functionality and density. SAMs are created by the spontaneous adsorption of molecules from the liquid or gas phase onto a substrate. During this process, the resulting surface layers are organized in surface domains in the $x - y$ plane¹⁷⁴. Until today, SAMs proved to be suitable for a large variety of applications including the design of nanoelectric devices, sensor arrays, super-capacitors and catalysts¹⁷⁵. A possible blueprint for carbon chain based design, is depicted in Figure 8.1, where each indicated group is responsible for a distinct SAM-surface property. The investigated SAM type consists of a head group, a carbon chain and a functional group. The head-group crucially determines the interaction with the surface and its choice is motivated by the used type of substrate. In the case of metal substrates, charged head groups lead to increased electrostatic interactions with the surface. Commonly used head groups for applications on metals include thioles, silanes and phosphates¹³. The length of the utilized carbon chains, determines the thickness of the obtained SAM and the internal molecule-molecule interactions. Applicability of the surfaces is determined by the type of the attached functional groups. The SAM-formation process by adsorption can be characterized by three steps according to the density of SAM molecules per surface area¹³:

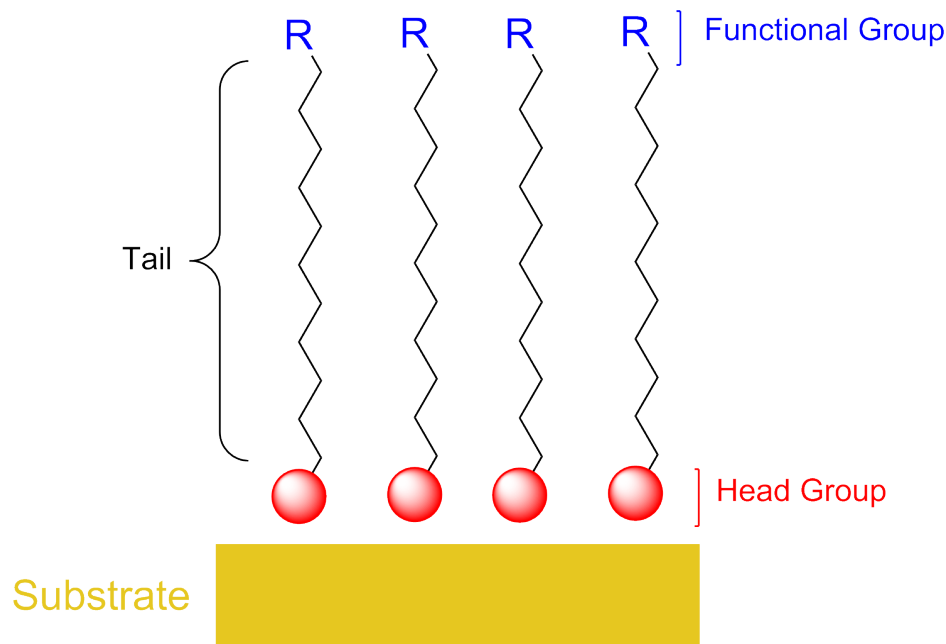


Figure 8.1: Scheme of a SAM with the substrates-surface, consisting of a headgroup, a carbon chain and if required a functional group.

1. Low density: Random dispersion/distribution of molecules on the surface.
2. Medium density: Conformational disordered and 'flat' lying molecules on the surface.
3. High density: Closed-packed order, molecules standing upright on the surface.

The occurrence of a specific phase depends on the time spent for layer organization and the properties of the created surface. For low densities, the interaction with the surface is dominated by van-der-Waals (vdW) interactions of highly tilted molecules. As soon as a medium density of molecules per surface is reached, the head group determines the main molecule-surface interaction. Therefore, the head groups affinity for chemisorption/physisorption on the substrate, greatly influences the time necessary for crystallization in the SAM formation process, as well as the obtained structural order. The choice of substrates is influencing the formation process and the resulting SAM structure. In this work, three types of organic-SAMs on metallic surfaces are investigated. In the following, the basic properties of three thiolates are discussed by theoretical investigations. Followed by an explanation of the used model for SAM on metal surfaces and related observations.

8.2 Investigated Thiolates in the Gas Phase

Three different species of SAM molecules are investigated in the following by means of DFT. For reference purposes, the obtained structures are then compared to the ones created using the CCM implementation of GFN-xTB. Initially a simple unmodified reference molecule, decanethiol (C10-SAM) is investigated by means of DFT. From this basic molecule, the modified versions perfluoro decanethiol (F-SAM) and amino-undecanethiole (NH₂-SAM) are studied. Afterwards a comparison of the effects arising from the respective modifications compared to the simple C10 case is conducted.

In the initial gas phase study, all energy calculations were performed using TURBO-MOLE¹⁷⁶. Geometry optimizations were conducted with dl-find¹⁶³ interfaced via CHEMSHELL¹⁷⁷. To gain an introductory overview, a geometry optimization at the def2-TZVPP/B3LYP-D3 level was performed for all three SAMs respectively, using their protonated as well as radical versions respectively. The resulting geometries are summarized in Figure 8.2. As can be seen from the structures itself, the carbon chain is arranged linearly. The molecular lengths of the

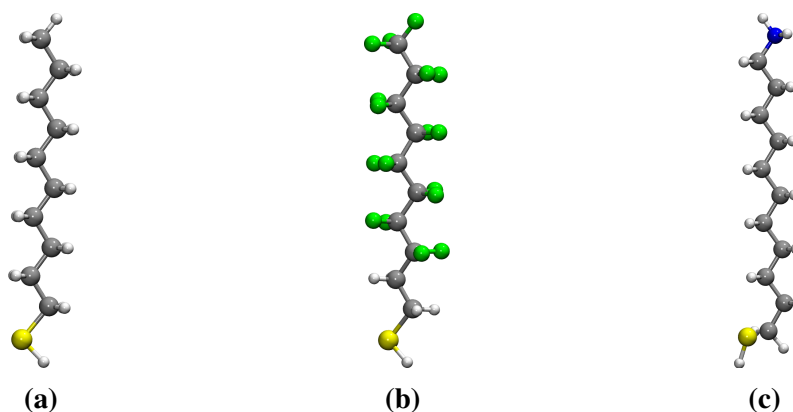


Figure 8.2: Gas phase optimized geometries on the def2-TZVPP/B3LYP-D3 level for saturated a) decanethiol C10-SAM, b) perfluoro-decanethiol F-SAM and c) amino-undecanethiol NH₂-SAM.

used SAM molecules, provides a structural reorganization, which directly effects the thickness of the created SAM-surface. Molecule lengths, could be determined by measuring the distances between the last carbon atom and the sulfur head group for F-SAM and C10-SAM as

Table 8.1: The molecular length d_{sd} and the carbon-sulfur bond distance d_{CS} of the three SAM-molecules calculated from their molecular geometry optimized on the B3LYP-D3/def2-TZVPP and PBE-D3/def2-TZVPP level respectively.

| Molecule | B3LYP | | PBE | |
|----------------------|-----------------------|-----------------------|-----------------------|-----------------------|
| | $d_{sd} / \text{\AA}$ | $d_{CS} / \text{\AA}$ | $d_{sd} / \text{\AA}$ | $d_{CS} / \text{\AA}$ |
| C10-SAM | 12.97 | 1.84 | 12.97 | 1.84 |
| F-SAM | 13.16 | 1.84 | 13.19 | 1.83 |
| NH ₂ -SAM | 14.43 | 1.84 | 14.39 | 1.84 |

well as the N-S distance for the NH₂-SAM. The molecule length d_{sd} can be calculated using

$$d_{sd} = |\mathbf{R}_{N/C_{10}} - \mathbf{R}_S| \quad , \quad (8.1)$$

where $\mathbf{R}_{N/C_{10}}$ and \mathbf{R}_S are the positions of the last carbon atom/nitrogen and sulfur respectively. The molecular lengths d_{sd} as well as the average C-C bond distance, were extracted from the geometry optimized structures. For comparison with the later performed surface studies presented in Section 8.4, calculations were also performed using the PBE functional.

It is not entirely clear from the experimental observation, whether SAM-molecules attach directly to the surface or they are deprotonized beforehand. However it is expected, that at some point during the chemisorption, the molecules interact with the surfaces as ions^{174,178–180}. For comparison, the case of radicals is investigated subsequently. Calculations were performed in analogy to the protonated case. However as the SAM-radicals inherit an unpaired spin orbital, open shell calculations as described in Section 2.5.1, have to be performed.

When comparing the SAM alignments in Figure 8.3 and Figure 8.2, it becomes clear that structural differences between radicals and saturated molecules are limited. When comparing the lengths of the SAM molecules, it becomes obvious that the radical structures are shorter, due to the reduced C-S bond length in the radical case. Generally it can be assumed that thiolates interact with metals as ions. From literature it is not exactly known, which portion of molecules is present in radical and protonated form respectively¹⁷⁴. For the conducted surface studies described in Section 8.4, the head group hydrogens are not considered and a neutrally charged system is assumed. In principle this corresponds to SAM-radicals not ions, this might

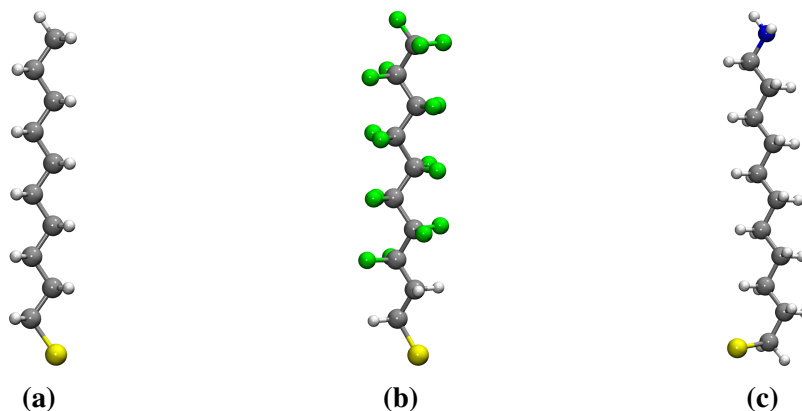


Figure 8.3: Gas phase optimized geometries on the def2-TZVPP/B3LYP-D3 level for saturated a) decanethiol C10-SAM, b) perfluoro-decanethiol F-SAM and c) amino-undecanethiol H-SAM.

Table 8.2: The spatial distance d_{sd} and the carbon-sulfur bond distance d_{CS} for the radical version of the three SAM-molecules, calculated from their molecular geometry optimized on B3LYP-D3/def2-TZVPP and PBE-D3/def2-TZVPP, respectively using the UKS formalism for unpaired spin orbitals.

| Molecule | B3LYP | | PBE | |
|----------------------|---------------------|---------------------|---------------------|---------------------|
| | $d_{sd} / \text{Å}$ | $d_{CS} / \text{Å}$ | $d_{sd} / \text{Å}$ | $d_{CS} / \text{Å}$ |
| C10-SAM | 12.97 | 1.81 | 12.96 | 1.80 |
| F-SAM | 13.25 | 1.80 | 13.13 | 1.79 |
| NH ₂ -SAM | 14.49 | 1.81 | 14.45 | 1.78 |

seem counterintuitive given the assumption of charged molecules but can be justified by the ability of the metals to redistribute the charge surplus.

8.3 Construction of Models for Metal Substrates

Surface models can either be constructed using experimentally obtained crystal structures from literature or directly from calculated bulk references. For crystalline systems, the applicability of literature values within a given theory framework can be estimated by comparison of the literature and calculated bulk crystal structures. In this section, the general procedure used for reference calculation by comparison of lattice constants and inter atomic distances is provided. Deviations arise due to limitations of the used quantum chemical description as

well as the uncertainty of weighting for the present interactions. When using the GPW-DFT framework for calculation of periodic crystal structures, localized basis sets are used which are not the natural choice for describing idealized metals as electron density is heavily distributed. However by including specialized diffuse functions of the MOLOPT-type⁸⁸ as well as the auxiliary plane wave basis as described in Section 2.11.1, these shortcomings in the description are overcome.

A generally applicable procedure for creating surfaces of crystalline materials, is given by optimization of a unit cell in the minimal three dimensional periodic bulk¹⁸¹. The obtained equilibrium lattice vectors within the given theoretical description can be directly compared to the literature values. Extending the optimized unit cell to a supercell by periodic repetition and removing periodicity in one direction, creates an extended surface model. In the following, the work-flow is illustrated and discussed step by step at the example of a gold surface, which was used for the following simulations of field evaporation in Chapter 9.

8.3.1 Exemplary Workflow for Construction of a Gold (111) Surface

In order to compare the GPW-DFT model with literature values, a bulk description based on a hexagonal unit cell of $3 \times 3 \times 3$ gold atoms was used. This shape of unit cell was motivated by the predominantly observed (111) surface which is the main scope of all follow up calculations. This surface type can be constructed from the fcc-bulk by cutting along the $hkl = 111$ lattice plane (compare to Section 2.8). Surface arrangements for the (111) are created by use of literature distances via a FORTRAN interface. The initial distribution of atoms leading to the arrangement depicted in Figure 8.4a was created using hexagonal symmetry with in plane lattice vectors $a_{\text{Au}} = 2.89 \text{ \AA}$. A cell optimization was performed using GPW-DFT starting from the literature unit cell. For the description of gold atoms, the specialized DZVP-MOLOPT-SR-GTH basis set was used, as it is constructed using more diffuse functions. For the electrons of the inner shells, the GTH-PBE-q19 PPs were used, leaving only 11 electrons for explicit treatment. Calculations were performed using the PBE exchange and correlation functional with Grimmes D3 dispersion correction.

The in plane lattice vector $a_{\text{opt,Au}} = 2.87 \text{ \AA}$ after cell optimization deviates only slightly from the literature value. Therefore, it is justified to directly use literature values within the utilized

description. The literature bulk arrangement is directly used in the simulations without further optimizations. Using the ideal (111) fcc bulk unit cell and removing the periodicity in one spatial direction, leads to the creation of a surface model. This procedure was used for most applications within this thesis and is a standard procedure used e.g. in ¹⁸².

Alternatively, the optimized three dimensional bulk unit cell within PBCs, can be used while re-optimizing a certain number of layers after increasing the lattice constant in z -direction. Increase of the cell vectors in direction of the surface normal, leads to a vacuum spacing, keeping the utility of the periodic treatment, while excluding interactions between the surface structures. The environmental differences of the surface layers lead to surface reconstruction effects.

From this stage on, follow up simulations can either be performed for the ideal structures or considered surface re-constructions. This option, allows for the inclusion of crystal imperfections as well as to account for the surface termination. The used procedure consists of geometry optimization of the uppermost layers while keeping the cell as well as the lower layers in the ideal arrangement. In the following studies, due to computational affordability, up to three layers are considered in the later model.

Surface effects are expected to range further than one layer and in principle the treatment becomes more sophisticated when considering reconstruction effects for additional layers. The later described applications for field evaporation are computationally very demanding. Therefore, not more than three layers were considered for the conducted surface studies and reconstruction is limited to the uppermost layer. The resulting reorganization for the exemplary gold (111) surface is shown in Figure 8.4c, where a geometry optimization utilizing the DFT-GPW scheme was performed for the uppermost layer of gold while constraining the bottom two layers to the ideal (111) configuration. For the calculations the TZVP-MOLOPT-SR-GTH basis set, PBE-D3 functional, a plane wave cutoff of 300 Ry were used alongside GTH-PBE-q11 PPs.

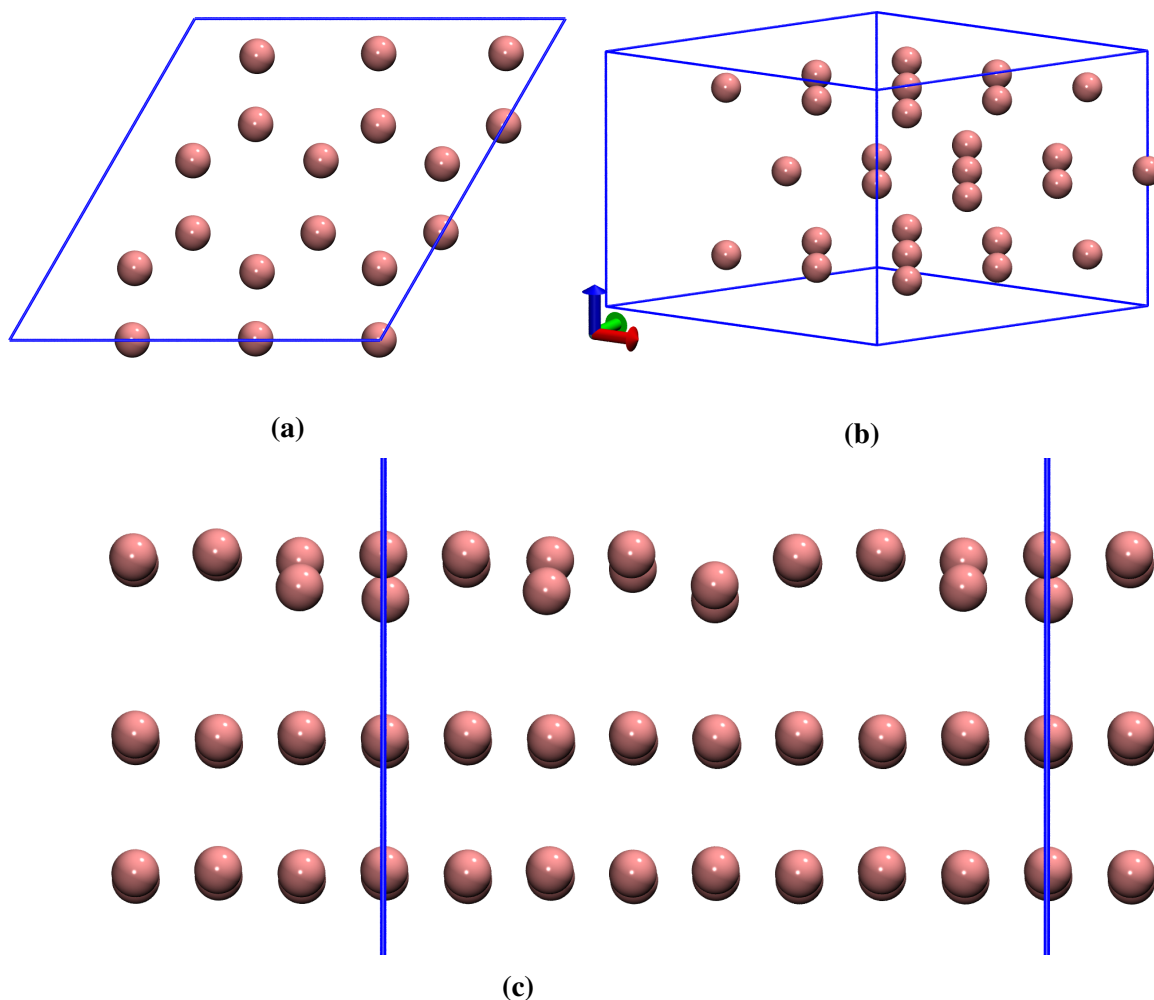


Figure 8.4: Illustration of the utilized systems for surface construction with a) top view onto the optimized fcc-bulk gold structure in $x - y$ plane, b) view onto the crystal structure of fcc gold after cell optimization, c) reconstructed surface utilizing a three layered 4×4 initial structure.

Bulk Reference Calculations for other Metals

The procedure described at the example of gold, was also used for all metals of the conducted surface study namely Pt, Cu, Ag, Pd and Al, which all crystallize in the fcc structure. The obtained lattice parameters deviate only slightly from the literature as shown in Table 8.3. Therefore, the use of ideal surfaces from literature bulk configurations is justified within the DFT-GPW description. Note: Optimizations were performed utilizing the corresponding DZVP-MOLOPOT-SR-GTH basis sets for each element with PPs utilizing GTH-PBE-q18 for Pt, GTH-PBE-q11 for Cu and GTH-PBE-q3 for Al respectively, following the nomenclature

Table 8.3: Overview of the deviations of calculated in plane lattice vectors lengths from literature values. The cell vectors were extracted from GPW-DFT optimized fcc bulk structures with constrained hexagonal symmetry.

| Element | a_{lit} in Å | a_{opt} in Å | Deviation |
|---------|-----------------------|-----------------------|-----------|
| Au | 2.88 | 2.89 | 0.34 % |
| Ag | 2.88 | 2.86 | 0.69 % |
| Pt | 2.78 | 2.86 | 2.87 % |
| Pd | 2.76 | 2.83 | 3.18 % |
| Cu | 2.55 | 2.50 | 2.00 % |

of PPs described in Section 2.10.

8.3.2 Bulk Lattices using GFN-xTB

For later applications, a comparison between GPW-DFT and calculations with GFN-xTB is provided. Being the computationally cheaper method, GFN-xTB offers the possibility of calculating large systems. The possibility to treat large surfaces in super-cell calculations offers a possibly more accurate representation of the real world situation as a larger variety of arrangements is possible. However, the description deviates in some cases heavily from the DFT reference. Possible reasons lie in the highly parametrized interactions and the minimal basis set approach as described in Section 4.3. When constructing extended systems for large scale simulations, deviations arising from the ideal case become more important and cell optimized structures have to be used.

In analogy to the previously described procedure for the DFT case, cell optimizations for the bulk case were performed utilizing GFN-xTB with periodic boundary conditions as implemented in CP2K. The obtained lattice constants in comparison to the literature values are summarized in Table 8.4 showing deviations below 10%, but are large enough to be considered when trying to conduct follow up calculations, especially for super-cell systems. As an example for an exceptional case, calculations were also conducted for aluminum which has a deviation of more than 30% which disqualifies it for further treatment. All follow up calculations using GFN-xTB, are utilizing surfaces constructed from optimized fcc bulks using a (111) surface description via a hexagonal unit cell in analogy to the previously described work-flow. The usage of optimized unit cells is necessary, as deviations are higher than in

Table 8.4: Overview of the lattice parameters deviations from literature in the GFN-xTB optimized fcc bulk structures.

| Element | a_{lit} in Å | a_{opt} in Å | Deviation |
|---------|-----------------------|-----------------------|-----------|
| Au | 2.88 | 2.91 | 1.04 % |
| Ag | 2.88 | 2.92 | 1.39 % |
| Pt | 2.78 | 3.01 | 8.27 % |
| Pd | 2.78 | 2.92 | 6.60 % |
| Cu | 2.55 | 2.60 | 1.96 % |
| Al | 2.86 | 3.86 | 34.96 % |

the DFT case. This leads to non convergence of the SCC procedure, when treating extended supercell structures which are the main focus of GFN-xTB application within this thesis.

8.3.3 Applicability of the Theoretical Models

When treating ideal surfaces, the non-periodic direction leads to difficulties in the description. One reason is the non vanishing surface dipole firstly investigated by Kohn in 1953^{183–185}. This effect was taken care of in many works that introduced a dipole correction based on Bengtssons' approach¹⁸⁶ and corresponding modifications¹⁸⁷. However, when treating non ideal and highly assymmetric surfaces as is the case for the studied SAM surfaces, the direct consideration of the surface dipole correction can be avoided. For all systems used within this thesis, the surface dipole correction was not considered as the systems show high anisotropy. Furthermore, the deviations of the optimized bulk lattice constants from the literature values in the DFT case are small, therefore the use of idealized literature values for the lattice constants is justified. The gold (111) surface created from bulk references, showed deviations smaller then 0.5% with respect to the literature values. Deviations in the case of GFN-xTB optimized bulks are generally higher and range up to $\approx 8\%$ in the case of platinum. The GFN-xTB description is assumed to be less accurate as DFT due to high level of integral parametrization with respect to gas phase references as described in Section 4.3. However its main advantage, lies in the possibility to treat explicitly extended supercell systems. Possible deviations arising within the description are extended to the full system as part of the supercell creation by repetition of the minimal unit cell. Consequently, deviations observed for the minimal bulk lattice have more impact on follow up calculations. In order to create a coherent setup for simulations

based on the created surfaces, these drawbacks are circumvented by using optimized bulk structures. The created systems therefore deviate from literature/experimental reference but secure consistency within the model itself.

8.4 SAM-Molecules on Gold-Surfaces

The main target of this chapter, is the explanation of building the SAM-substrate interface and the investigation of the modeled SAM interactions.

From literature it is known, that thiolate gold interactions lead to a surface re-construction of the uppermost layer¹⁸⁸. However in this study, the surface reconstruction was not considered, as the complexity of the system and thus the computational effort are increased significantly since more metal layers would be required as . The setup used in this thesis was constructed based on literature of previous investigations and the requirements of the following simulations of field evaporation processes^{15,17,18}.

The SAM-substrate interface, was constructed from the previously created idealized metal surfaces based on literature parameters as described in Section 8.3.1. Each effect will be shown at the example of an ideal gold (111) surface, due to gold being used as the metal substrate of choice in the later simulated experiments of field evaporation¹⁹.

The strength of the molecule-surface interaction is mainly determined by the interactions arising within the interface region. In the thiolate case, it is expected to be dominated by the sulfur metal bond. Theoretical studies of small to medium size carbon thiolates have been performed in the past including for example^{174,178,189–191}.

To model surface effects, calculations utilizing periodic boundary conditions were performed, using 20 Angstrom of vacuum spacing in direction of the surface normal, eliminating unphysical interactions with the periodic images (as explained in Section 8.3.1). As stated already, molecules are expected to minimize their surface energy by maximizing interactions with the metal substrate. Investigations of SAM molecules in the literature, show that SAM-density is directly influencing the tilting of molecules^{10,192}. The molecule tilting is directly related to the thickness of the created SAM-surface as shown in Figure 8.1. Furthermore, the layer thickness is effected by the overall flexibility of SAM molecules to arrange themselves

e.g. tilting. Most effects are directly related to the SAM density ρ_{SAM} , that can be defined as

$$\rho_{\text{SAM}} = \frac{N_{\text{SAM}}}{A_{\text{cell}}} \quad , \quad (8.2)$$

with the number of SAM molecules per unit-cell N_{SAM} and the surface area A_{cell} , calculated according to the underlying cell symmetry. In the hexagonal case, it is given as

$$A_{\text{cell}} = \frac{2}{\sqrt{3}}ab \quad , \quad (8.3)$$

where a, b are the lengths of the in plane cell constants. In the following, investigations are performed for a gold (111)-surface with various numbers of layers and surface areas. The SAM-density per unit cell is calculated using the lattice plane distance of gold given as $a_{\text{Au}} = 2.88 \text{ \AA}$. A short overview of possible densities associated with respective surface sizes is given in Section D. In the following, systems corresponding to densities as presented in Table 8.5 are created and analyzed using idealized gold (111) surface constructed as described in Section 8.3.1. For better readability, the inverse SAM density ρ_{SAM}^{-1} is used in the following throughout this thesis.

Table 8.5: The inverse densities ρ_{SAM}^{-1} utilized in the investigated models depending on the number of SAM molecules N_{SAM} , and the integer multiples n_a, n_b of the cell vectors $a = b = 2.88 \text{ \AA}$.

| N_{SAM} | n_a | n_b | ρ_{SAM}^{-1} in $1/\text{\AA}^2$ |
|------------------|-------|-------|--|
| 1 | 2 | 2 | 28.73 |
| 1 | 3 | 3 | 64.64 |
| 2 | 3 | 2 | 21.54 |
| 3 | 3 | 2 | 14.36 |
| 4 | 4 | 4 | 28.73 |
| 5 | 4 | 4 | 22.98 |

8.4.1 Interaction Energies

Structural properties are mainly determined by the surface coverage, however another important property is the molecule surface interaction itself as it plays a crucial role for the

8 Investigation of Surface Models for Self-Assembled Monolayers on Metal Substrates

systems stability. In order to investigate the adsorption properties of the simulated surface models, an investigation of the underlying interaction energies should give further insight. The contributions for the total interaction energy E_{int} can be split in three major parts:

1. Sulfur-metal bond.
2. Intermolecular SAM interactions.
3. Interactions with periodic images.

The sulfur-metal bond is of major interest as it determines the adsorption strength of the SAM molecules on the substrate. However, obtaining the bond-profile in full periodic calculations is highly computationally demanding and in the case of dense packing not uniquely possible. An estimation for the sulfur-metal bond can be calculated by identifying its contribution to the total interaction energy E_{int} . In comparison to typical gas phase problems, the additional contributions have to be determined first, in order to identify the corresponding energetic portion.

In analogy to a gas phase calculation, the total interaction energy of the SAM surface interface, can be calculated as

$$E_{\text{int}} = E_{\text{SAM+Metal}}^{\text{per}} - (E_{\text{SAM}}^{\text{per}} + E_{\text{Metal}}^{\text{per}}) \quad , \quad (8.4)$$

where the superscript per indicates periodic treatment, $E_{\text{SAM}}^{\text{per}}$ is the energy of the SAM molecules, $E_{\text{Metal}}^{\text{per}}$ the energy of the metal substrate and $E_{\text{SAM+Metal}}^{\text{per}}$ the energy of the combined system¹⁹³. The energy of the combined surface interface, contains contributions due to the S-metal bond and interaction with periodic images. A separation of both energetic contributions allows for quantification of surface properties. Difficulties arise, when the identification of the periodic SAM contributions has to be performed. The contribution arising due to the images of additional SAM molecules, is estimated by calculating the single point energies for the surface optimized SAMs without metal substrate. The reference contribution due to periodic images of SAM molecules $E_{\text{per,ref}}^{\text{SAM}}$ can then be calculated following

$$E_{\text{per,ref}}^{\text{SAM}} = E_{\text{SAM}}^{\text{per}} - E_{\text{SAM}}^{\text{vac}} \quad , \quad (8.5)$$

given as the difference between periodic $E_{\text{SAM}}^{\text{per}}$ and gas phase $E_{\text{SAM}}^{\text{vac}}$ energies. Furthermore, energetic contributions arising due to intermolecular interactions within the unit cell $E_{\text{int}}^{\text{SAM-Cell}}$ can be estimated by calculating the SAMs energy without PBCs and comparison to isolated molecules. The energetic contribution is then given as

$$E_{\text{int}}^{\text{SAM-Cell}} = E_{\text{SAM}}^{\text{vac}} - N E_{\text{SAM}}^{\text{iso}} \quad . \quad (8.6)$$

Here N is the number of SAM molecules per unit cell and $E_{\text{SAM}}^{\text{iso}}$ the energy of a single isolated molecule.

8.5 Decanethiol

For comparison of additional interactions for modified decanethiol versions, the investigation of the unmodified reference molecule is necessary. For the used model, SAM radicals are placed on a gold (111) surface, with the chemisorption modeled by the sulfur gold bond. The nature of the chemisorption process as well as the organization of the SAM in radical or saturated form with and without surface reconstruction is a debated topic in the literature^{178,188,194–197}. Besides the fact that it is unclear whether protonated or radical molecules adsorb onto the substrate, the self organization procedure is unknown as well. However within this thesis, the target properties are structural arrangement of the already chemisorbed SAM surface. For the performed investigations, the uncertainties are not of major relevance as the investigation focuses on tilting and molecular length. In the following, the model for SAM surfaces and its suitability are illustrated and discussed.

8.5.1 Considerations for the Smallest System

To investigate the properties of the utilized model, the examination of the smallest physical meaningful system was chosen as starting point. Increasing the surface size as well as the number of SAM molecules leads to further improvements of the model by considering additional SAM-SAM as well as SAM-surface interactions. The smallest system that can be constructed by considerations of the vdW radii as well as PBCs is one molecule placed on a

8 Investigation of Surface Models for Self-Assembled Monolayers on Metal Substrates

2×2 unit cell which is illustrated for the case of decanethiol in Figure 8.6. Hypothetically, a system created by a 2×1 unit cell leaves enough space for the molecule to arrange, however even when ideal conditions are assumed and the molecule is placed in the cell center, the resulting configuration can be considered nonphysical as hydrogen atoms arrange with too little distance between each other. The situation is schematically represented in Figure 8.5b, where the proportions are chosen to suit the inter atomic distances within a unit cell.

This can easily be shown by measuring the equilibrium distances of a simplified model. Suppose a C-C bond in the molecule with two fluorine atoms attached to each of them, resulting for a parallel arrangement to the surface normal in a rectangular arrangement if seen from the top as shown in Figure 8.5b. Here, the atoms belonging to one C_2H_4 segment of the molecule chain are indicated, where hydrogen atoms are represented by gray and carbon atoms by black circles. For a relaxed molecule aligned in z direction with a tilting of $\phi = 0$, the observed distance between two C-C atoms of the simulated carbon chain in the $x - y$ plane is given on average by $r_{CC} = 1.2 \text{ \AA}$. Here, the spatially most favorable arrangement is assumed when the C-C bond is positioned at the cell's center along the cell bisection indicated in orange. The ideal position is assumed to be the one with the largest distances to the cell boundaries. The length of the bisection is exactly the same as for the cell vector \mathbf{b} . Therefore, the maximum distance d_{\max} can be calculated following

$$d_{\max} = |\mathbf{b}| - \frac{r_{CC}}{2} \quad . \quad (8.7)$$

The positions of the carbon atoms within the unit cell can thus be determined following

$$\mathbf{r}_{C_{1,2}} = \frac{\mathbf{a} + \mathbf{b} \pm r_{CC}\mathbf{e}_b}{2} \quad . \quad (8.8)$$

Within the reference molecule, the relaxed hydrogen atoms are positioned in a plane with the carbon atoms they are attached to and tilted by an angle $\theta_{HCH} \approx 110^\circ$ from each other while the bond distance is given as $r_{CH} \approx 1.10 \text{ \AA}$. The position of the hydrogen atoms $\mathbf{r}_{H_{1,2}}$ in the $x - y$ plane within the unit cell, can thus be determined following

$$\mathbf{r}_{H_{1,2}} = \mathbf{r}_{C_1} - r_{CH} \cos\left(\frac{\theta_{HCH}}{2} + \gamma\right) \quad , \quad (8.9)$$

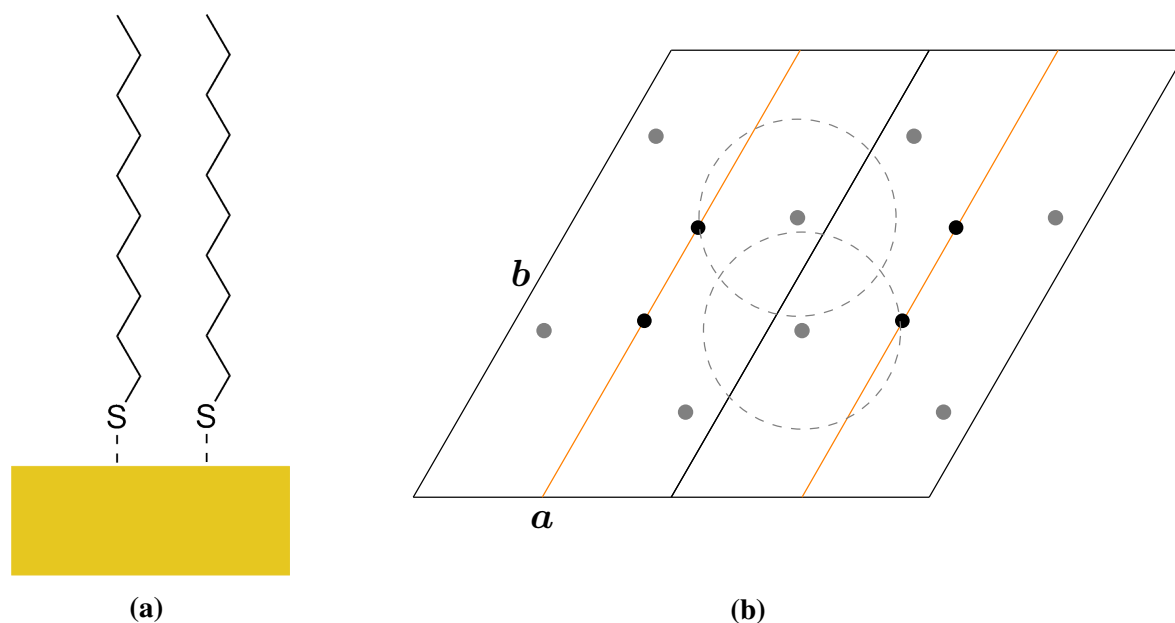


Figure 8.5: Schematic considerations of the created models. a) Chemical structure of the C10-SAM molecules attached on the substrates surface. b) Schematic unit cell arrangement illustrating the situation for a 2×1 unit cell assuming a perfectly aligned C10 molecule. Carbon atoms are represented in black, hydrogen atoms in gray. The cell vectors a and b are drawn in red, whereas the cell diagonal cutting the system in half is shown in orange. The vdW radii of the hydrogen atoms closest to the cell boundaries are illustrated by dashed lines in gray.

with γ being the cell angle between the vectors a and b . It is further assumed through the field of chemistry, that the minimum distance measure for closed packed atoms of the same type is given by the van-der-Waals radii as threshold parameters. To indicate the minimum H-H distance of $r_{\text{vdW}} = 1.1 \text{ \AA}$, the van-der-Waals radii are represented in Figure 8.5b via dashed gray circles. Even in this idealized case, the distances to adjacent hydrogen atoms within neighboring cells are too small to provide a suitable setup. This can be further illustrated considering the huge idealizations considered within this simple geometrical proof. The suppression of molecular tilting, head group arrangement as well as surface attachment effects. Considering these difficulties, the smallest system for might be calculable in the C10 case, however for its modified versions it is not. This is illustrated in Section 8.6 for the fluorinated case, eliminating the possibility of comparison and therefore disqualifying it as reference. The smallest used model system, is therefore chosen based on a 2×2 unit cell in all simulated cases.

8.5.2 One C10-SAM Molecule

The system consisting of one C10 molecule on a two layerd 2×2 gold surface, is relaxed by geometry optimization of the SAM molecule while constraining the gold atoms. Energy calculations are performed by means of GPW-DFT using the PBE-D3 functional and TZV2P-GTH basis sets with corresponding PPs for the non metal atoms. Gold atoms are described by DZVP-MOLOPT-SR-GTH basis sets and PBE-GTH-q11 PPs. The top view onto the optimized unit cell shown in Figure 8.6a, reveals the tilted molecule ranging over the whole unit cell. The side view onto the relaxed system is depicted in Figure 8.6b. The total molecular length could be determined to be $d_{C10,sd}^{2 \times 2} \approx 11.57 \text{ \AA}$ whereas the tilting angle with respect to the surface normal is given as $\phi_{2 \times 2}^{C10} \approx 26.5^\circ$. The smallest system of one SAM molecule per 2×2 unit cell following Eq (8.2) corresponds to a density of $\rho_{SAM}^{-1} \approx 28.7 \text{ \AA}^2$. Analysis of the interaction energy following Eq (11.1) leads to an estimate of $E_{int}^{PER} \approx -79.2 \text{ kcal/mol}$, meaning that the constructed surface is stabilizing the molecule. From literature it is known that the sulfur metal bond calculated with the PBE functional is quantified to contribute a portion of $E_{SM} \approx 55 \text{ kcal/mol}$ ¹⁹⁸. Assuming this estimation leads to the conclusion that besides the bond stabilization also the periodic images interact favorably. However further analysis of the interaction energy arising due to the images following Eq (8.5) leads to the contribution of $E_{SAM}^{PER} \approx -49.3 \text{ kcal/mol}$. From the both contributions it can be concluded that in summary the sulfur metal bond can be quantified to contribute around 40 kcal/mol. This bond is of great significance for simulations within electric fields, as it marks the maximal barrier for full desorption of molecules from the surface.

To model SAM-layers with lower densities, the surface area per unit cell can be increased while keeping the number of SAM-molecules constant. This is exemplary done for decanethiol on a 3×2 gold surface as shown in Figure 8.7. After optimization, the molecule is oriented towards the surface, reducing its effective length to $d_{sd,3 \times 2}^{C10} \approx 10.9 \text{ \AA}$ and the tilting angle increases to $\Phi_{3 \times 2}^{C10} \approx 33.9^\circ$. This effect can be attributed to the maximization of the vdW-interaction between SAM and surface, possible due to the lowered molecular density. The treatment of one molecule per unit cell allows for the investigation of surface related properties, that differ from the cluster calculations. Therefore, one molecule per unit cell is suitable in order to determine effects stemming from the solid state treatment, describing a SAM with

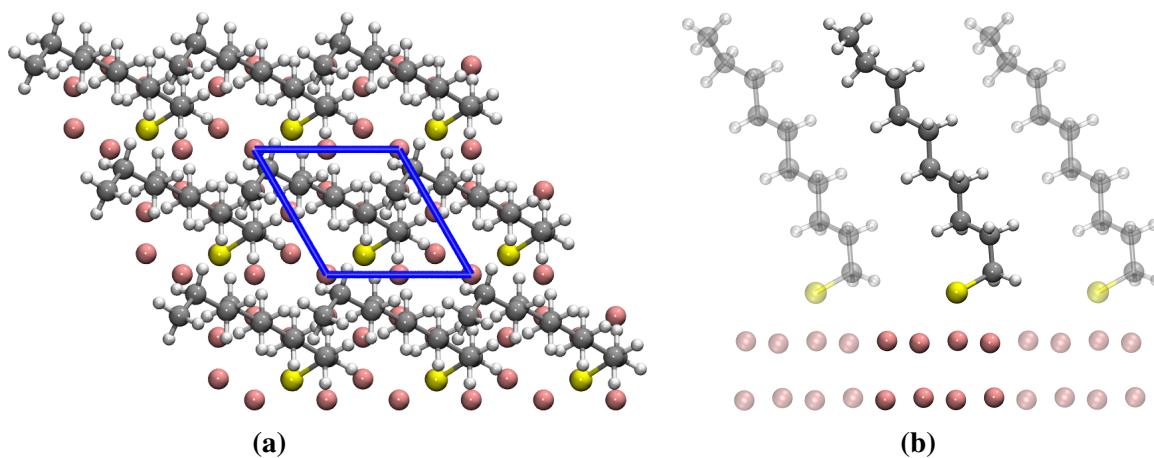


Figure 8.6: The unit cell arrangement after optimization for one molecule for the C10-SAM on a 2×2 gold surface. a) Top view along the surface normal with the cell boundaries indicated in blue. b) side view of the unit-cell, with the periodic images in transparent colors.

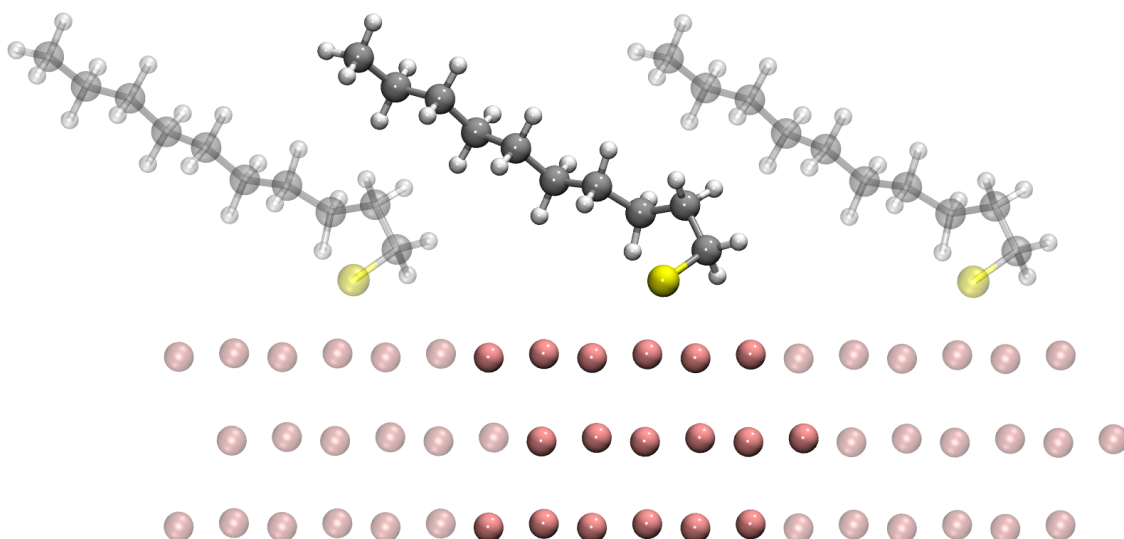


Figure 8.7: Side view onto the unit cell for one decanethiol molecule on a three layered 3×2 gold surface, revealing the increased tilting of the SAM molecules towards the surface.

complete restriction and perfect order in its molecules. A significant increase in description quality is achieved when treating more than one molecule per unit cell.

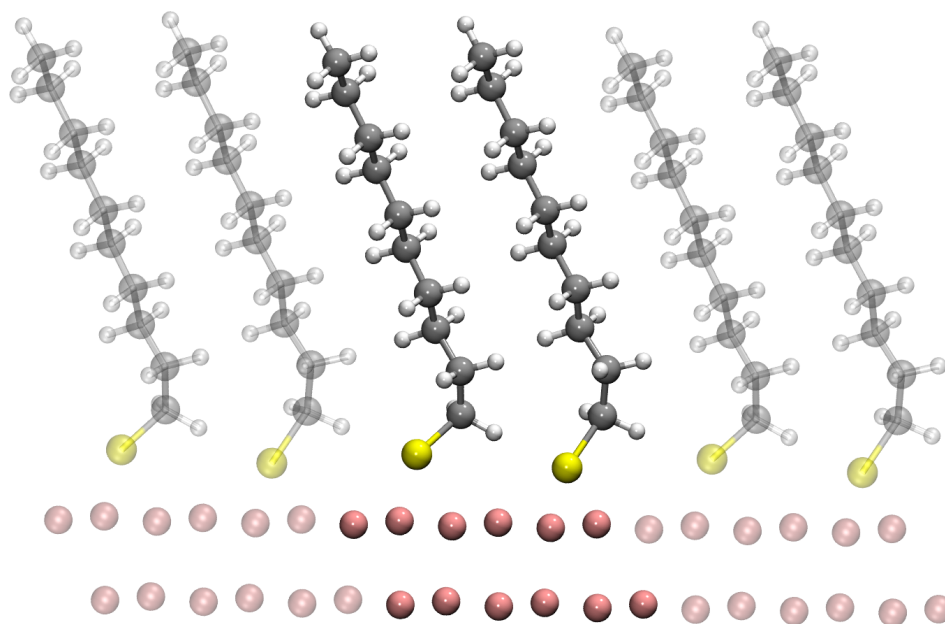
8.5.3 *Explicit Intermolecular Effects (2 Molecules)*

The treatment of intermolecular interactions with increased flexibility for structural rearrangement, is possible when using two and more molecules per unit cell. A significant increase in the description quality is already achieved when treating two molecules. For two C10-SAM molecules on a 3×2 gold (111) surface, the relaxed system is shown in Figure 8.8. The side view on the molecules shown in Figure 8.8a reveals reorientations between the carbon chains with increased tilting angles. The observed behavior for carbon thioles is consistent with previous investigations in the literature^{174,178,192}. Two molecules on a 3×2 surface correspond to a SAM density of $\rho_{\text{SAM},3 \times 2}^{-1} \approx 21.55 \text{ \AA}^2$. The average molecular tilting could be determined to be $\bar{\phi}_{\text{C10},3 \times 2} \approx 13.3^\circ$ and the average molecular length as $\bar{d}_{\text{sd},3 \times 2}^{\text{C10}} \approx 12.36 \text{ \AA}$. Analysis of the surface stability by calculating the interaction energies following Eq (11.1) leads to $E_{\text{int}}^{\text{PER}} \approx -29.8 \text{ kcal/mol}$.

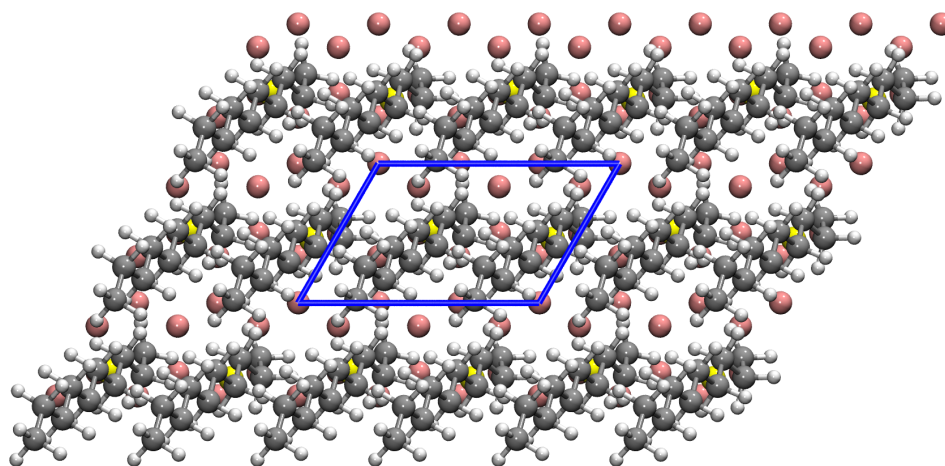
From the obtained rearrangements depicted in the top view onto the extended unit cell depicted in Figure 8.8b, it becomes apparent, that molecules tend to tilt in order to fill the whole unit cell. Furthermore, molecules arrange to minimize repulsion between adjacent chains by staggered arrangements of the hydrogen atoms.

8.5.4 *Surface Arrangement with three Molecules per Unit-Cell*

In order to improve the description by additional flexibility within the unit cell, more SAM molecules can be considered explicitly. Doing so allows for more rearrangements, due to intermolecular interactions which are not considered with mirror images. The resulting arrangements can then be analyzed for their respective efficiency considering model improvement and additional computational effort. An example for three SAM molecules on a 3×3 surface is illustrated in Figure 8.9. From the side view shown in Figure 8.9a, it can be concluded that all molecules arrange very similarly. The top view illustrated in Figure 8.9b reveals high molecular tilt. Molecules align themselves throughout the whole unit cell even more than in the case of two molecules depicted in Figure 8.8b. Interestingly three molecules on a 3×3



(a)



(b)

Figure 8.8: The unit cell arrangement after optimization for one decanethiol molecule on a two layered 3×2 gold surface. Top view onto the unit cell of decanethiol after optimization, with the cell boundaries in blue. Side view of the decanethiol unit cell, with the periodic images in transparent colors, the cell boundaries are indicated in blue.

8 Investigation of Surface Models for Self-Assembled Monolayers on Metal Substrates

surface correspond to a density of $\rho_{\text{SAM},3\times3}^{-1} \approx 21.55 \text{ \AA}^2$ which was also modeled in the case of two molecules on a 3×2 surface. The obtained average tilting of $\bar{\phi}_{\text{C10},3\times3} \approx 22.13^\circ$ and a further reduction of the average molecular length to $\bar{d}_{\text{sd},3\times3}^{\text{C10}} \approx 11.83 \text{ \AA}$.

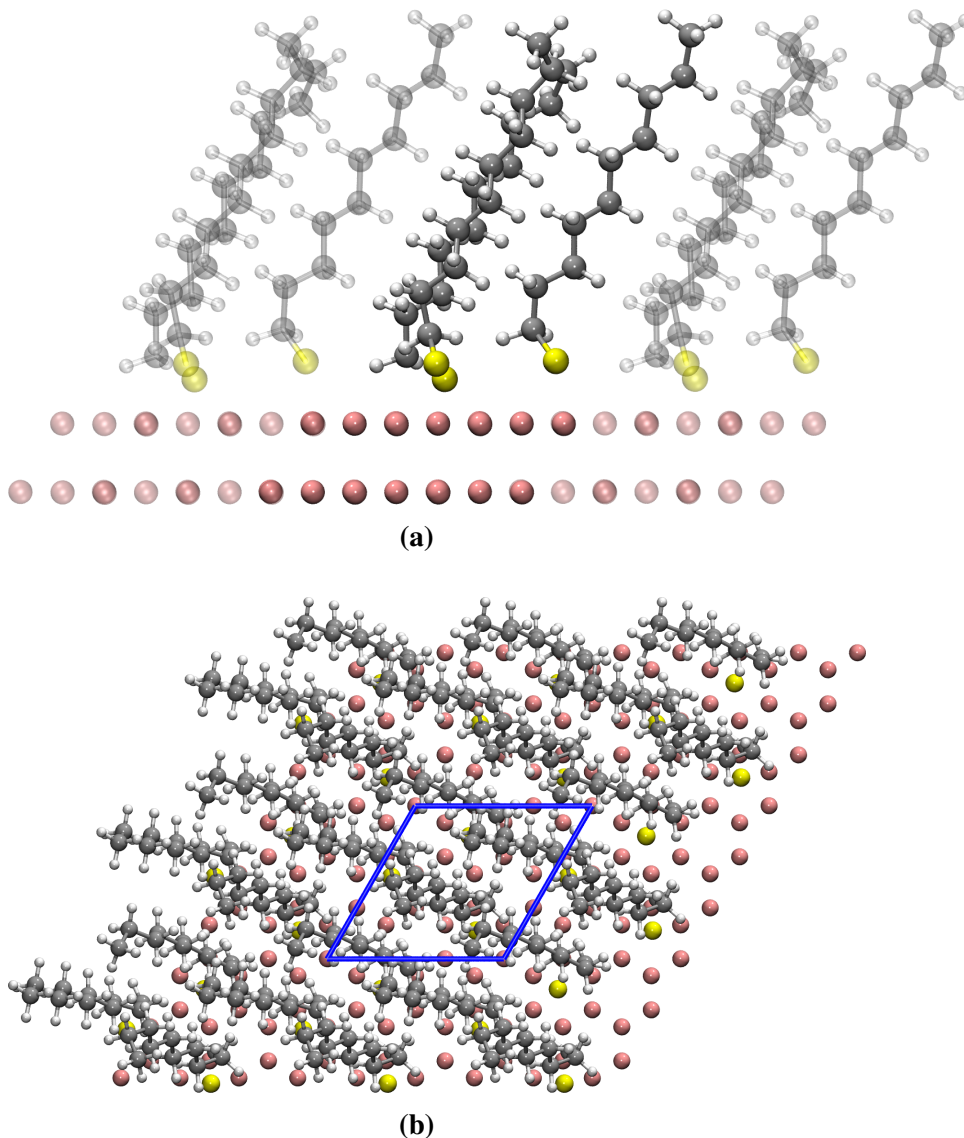


Figure 8.9: The unit cell arrangement after optimization for three C10 molecules a two layered 3×3 gold surface. Top view onto the unit cell of three C10-SAM molecules per unit cell after optimization, with the cell boundaries in blue. b) Side view onto the optimized unit cell including three SAM-molecules.

It can be seen from the combined perspectives, that molecules are rearranging themselves in order to fill the free surface area, leading to a tilting of the molecules. Interestingly the obtained rearrangements are larger than in the case of the previously described case of two molecules. This leads to the conclusion that larger surface area as well as more molecules

allow more rearrangements.

The previously investigated system of one SAM molecule for lower surface densities indicated that density and tilting are correlated. However combined with the results of the model for three molecules, the molecular tilting can be attributed to the size of the modeled unit cell, as it marks the major difference between the models. In summary the observations allow to conclude that adding additional degrees of freedom to the system influences the observed model behavior.

Discussion of the Surface Models Suitability

Compared to the simulation utilizing a cluster model, the surface model provides additional effects due to intermolecular interactions within the explicitly treated unit cell as well as periodic images. For lower densities, the molecular arrangement becomes similar to the cluster case as cell boundaries are further away from the treated molecules. The description is improved due to interactions with mirror images, however in the 2×2 case these do not add additional flexibility for rearrangements. The surface model has no model boundaries, therefore unphysical edge effects, leading to adsorption at the surface termination in the $x - y$ plane are suppressed. Avoidance of these effects as well as the possibility of simulating extended layers rather than isolated ones, make the surface model in principle the natural choice for follow up SAM investigations. The most crucial improvement is achieved when two molecules per unit cell are considered, allowing for intermolecular adjustments as well as reorientation on the surface.

Comparing to typically used SAM densities in the experiment $\rho_{\text{exp}}^{-1} = 21.5 \text{ \AA}^2$, using two molecules seems the most suitable compromise for expensive follow up simulations. Adding further molecules leads to better adjustability coming at the cost of more expensive SCF calculations. For the example of the C10-SAM, the increase of the system size showed small differences, however the improvements seem to be limited. Therefore, the system size is only increased when it becomes qualitatively necessary.

8.6 Perfluoro-Decanethiol

Substituting H with F for the uppermost eight carbon atoms of the C10-SAM molecule, leads to the perfluoro-decanethiol molecule (F-SAM). The additional fluorine atoms ($r_{\text{vdW}}^{\text{F}} = 1.47 \text{ \AA}$) require more space when compared to hydrogen atoms ($r_{\text{vdW}}^{\text{H}} = 1.10 \text{ \AA}$) and are strongly repulsive with respect to each other due to their high electron density in the valence. In combination, this leads to higher intermolecular distances and higher “effective” surface coverage by a smaller number of molecules. The spatial requirements of fluorine atoms are illustrated in Figure 8.10. For F-SAM systems, it is expected that higher SAM densities are less favorable and molecular tilting is lowered. This effects can be easily illustrated when considering the hypothetical arrangements for a 2×1 unit cell in analogy to Section 8.5.1 for the C10-SAM. Conducting comparable considerations for the in plane arrangement as in the fluorinated case, results in the model shown in Figure 8.10b. The additional spatial requirements of the fluor attachments become apparent when comparing with the schematics for the C10-SAM in Figure 8.5b. Fluorine atoms are represented as green, carbon atoms as gray circles with the vdW-radii of the boundary atoms indicated by dashed green circles. The distances of the atoms are too small to allow for a physically interpretable model arrangement, even in this idealized case. Here, the previously described procedure is used, considering the fluorine carbon bond distance $r_{\text{CF}} = 1.36 \text{ \AA}$ as well as the tilting angle $\theta_{\text{FCF}} \approx 109^\circ$ along side the increased vdW radii for fluor atoms $r_{\text{vdW}}^{\text{F}} = 1.47 \text{ \AA}$. Starting again from the simplistic surface model containing one SAM molecule per 2×2 unit cell, the effects for explicit intermolecular treatment (e.g. more than one molecule per unit cell) as well as SAM-density related effects are investigated analogously to the previous study for the C10-SAM. The unit cell of the smallest investigated system after geometry optimization, is shown in Figure 8.11. As can be concluded from the top view onto the unit cell in Figure 8.11a, the F-SAM molecules require more space and are therefore less tilted than in the analogous case for C10. The tilting angle of perfluoro-decanethiol with respect to the surface normal for the smallest system, could be determined to be $\phi_{\text{F},2 \times 2} = 22.1^\circ$. The segments of the molecule are aligned straight without rotation along the molecular axis as illustrated in Figure 8.11b, leading to a molecular length of $d_{\text{sd},2 \times 2}^{\text{F}} \approx 13.19 \text{ \AA}$. Calculation of the interaction energy following Eq (11.1) shows a value of $E_{\text{int}} \approx 71.25 \text{ kcal/mol}$ indicating that the surface arrangement is in principle less

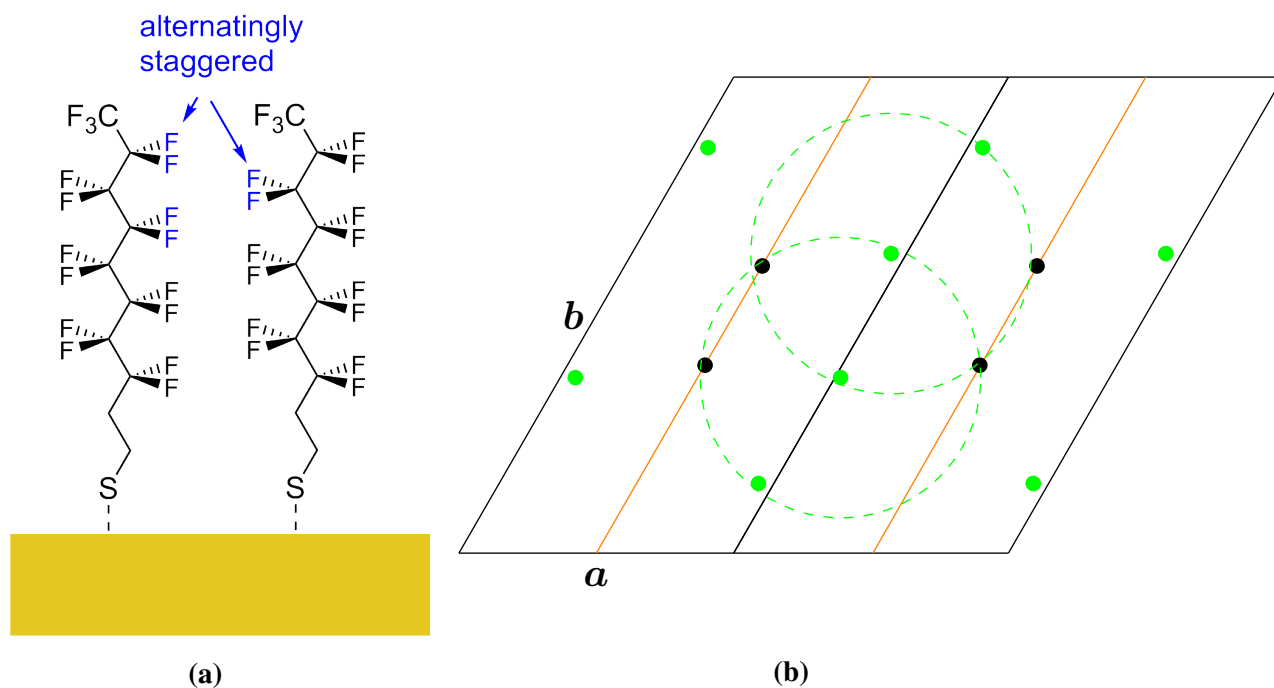


Figure 8.10: Overview of the structural properties of the F-SAM arrangements. a) Schematic representation of the fluorinated decanethiol on a metal surface illustrating the observed staggered arrangement. b) Illustration of the placement problem for a 2×1 unit cell. The distances were chosen in analogy to the relaxed molecule, and an idealized alignment of the molecule along its axis was assumed. The van-der-Waals radii of the fluorine atoms are indicated by dashed lines.

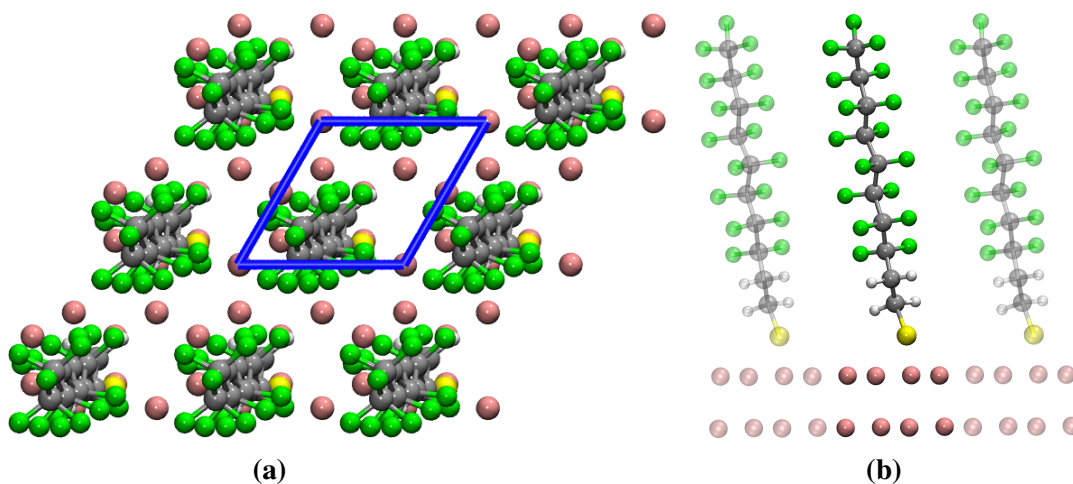


Figure 8.11: Unit cell for one F-SAM molecule on a two layered 2×2 gold surface. a) Top view onto the unit cell after optimization, with the cell boundaries in blue. b) Side view of the unit cell, with the periodic images in transparent colors.

favorable than the isolated molecule case. The reason for this lies in the repulsive forces between adjacent fluorine atoms of the molecule itself and its mirror images, compared to the case of decanethiol, the fluorine atoms interact stronger while being closer to each other. The limited amount of possible reorganization during the optimization on a 2×2 surface makes the arrangement overall unfavorable. Analyzing the contribution of the SAM images following Eq (8.5) reveals the energetic portion to be $E_{\text{int}}^{\text{PER}} \approx -22.75$ kcal/mol. When investigating lower densities, this behavior changes as shown Figure 8.12b where one relaxed SAM-F molecule is placed on a 3×3 gold (111) surface. As expected from the C10 case, the tilting of the molecule also increases for the F-SAM and is given as $\phi_{\text{F},3 \times 3} = 65.0^\circ$. The tilting increase leads to a clear reorientation towards the surface, maximizing vdW interactions between fluorine and gold atoms. However in comparison to the C10 case, the fluorine atoms are more repulsive with respect to the surface, leaving them as far away as possible while increasing the hydrogen surface interaction.

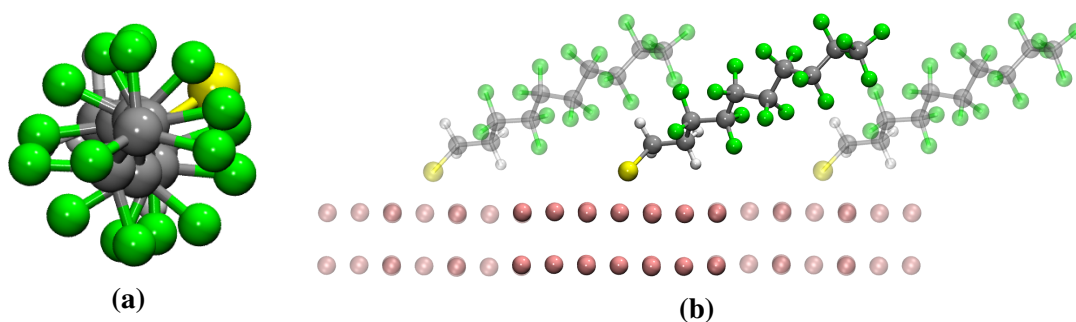


Figure 8.12: Overview of the F-SAM arrangement in the low density regime. a) View aligned with the molecules axis in the case of one per-fluoro decanethiol molecule on a 3×3 surface. Surface atoms are excluded for better visibility. b) Side view onto the unit cell for one SAM-molecule on a 3×3 gold surface, with periodic extension of the surface atoms. The PBC-SAM-pictures were excluded due to visibility reasons.

Another feature is the internal tilt of this SAM molecule, which was first observed in gas phase calculations^{10,18}. This effect can be seen in the low density surface model, when aligning the view with the molecule axis. This indicates that lowering the density of SAM molecules, recovers somehow the gas phase features. This can be easily understood by considering that the main difference to the single molecule gas phase calculation lies in the interaction with the periodic images.

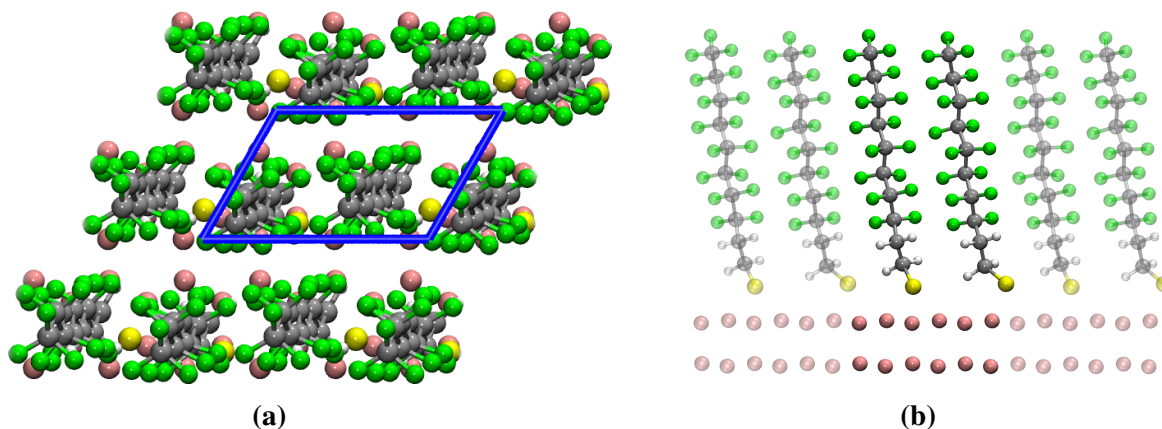


Figure 8.13: Unit cell arrangement for two SAM molecules per unit cell on a two layered 3×2 gold surface. a) Top view onto the unit cell of the perfluoro-decanethiol unit cell after optimization, with the cell boundaries in blue. b) Side view of the perfluoro-decanethiol unit cell, with the periodic images in transparent colors.

When comparing with typical experimental conditions utilizing fluorinated decanethiol¹⁵, densities are estimated to be around $\rho_{\text{SAM}}^{-1} = 21.5 \text{ \AA}^2$. Therefore, it can be assumed that the general tilting of molecules in the experiment is rather small, thus, systems with low density are not the focus of the following studies.

8.6.1 Explicit Intermolecular Effects (2 Molecules)

Analogous to the C10 case, the most crucial modification to the minimal system is provided when treating two molecules per unit cell. The explicit intermolecular interactions lead to observable differences of the fluor substituents during the optimization process. For comparison with the previous C10-SAM model, again two SAM molecules were placed on a 3×2 gold surface considering three layers in stacking direction. The resulting unit cell arrangements after geometry optimization on the PBE-D3/TZV2P-GTH level are shown in Figure 8.13. As can be seen in Figure 8.13a, the relaxed system is organized with fluorine atoms in alternating orientations, leading to a slight displacement between adjacent molecules as seen from the top view. Another observation can be made when considering the side view shown in Figure 8.13b, showing a total molecular length of $d_{\text{sd},3 \times 2}^{2\text{F}} \approx 13.12 \text{ \AA}$ as well as average molecular tilting of $\bar{\phi}_{3 \times 2}^{2\text{F}} \approx 20.7^\circ$. Here, it becomes obvious that fluorine atoms are not only arranged in alternating manner in the $x - y$ plane, but also in the direction of the surface normal.

The reorientations between the two carbon chains due to the repulsive forces between the F

atoms cause the molecules to straighten, this can be seen when comparing the tilting angle and SAM length to the C10 case where $\bar{\phi}_{3 \times 2}^{\text{C10}} \approx 25.2^\circ$ as well as $\bar{d}_{\text{sd},3 \times 2}^{\text{C10}} \approx 11.03 \text{ \AA}$ was obtained. Within the literature, the importance of interactions arising from the molecule backbone, could be determined as crucial for SAM arrangements^{192,199}. However comparing the interaction energy of $E_{\text{int}} \approx 6.58 \text{ kcal/mol}$ following Eq (11.1) to the case of one SAM-F molecule on a 2×2 surface shows a strong stabilization. Possible reasons for this observation are found in the staggered arrangement of segments as well as the molecules themselves as becomes apparent when considering Figure 8.11. Further indication for this effect is found when considering the contribution of the periodic SAM images given as $E_{\text{int}}^{\text{PER}} \approx -11.71 \text{ kcal/mol}$. In summary, as initially expected, F-SAM molecules need more space compared to the C10 ones, leading to smaller tilting and increased molecular length. The intermolecular interactions for identical SAM-densities are much stronger than in the C10 case. As a result, the investigated arrangements allow only limited flexibility between the molecular chains and thus the fluorinated SAM is expected to be more rigid. Positive interaction energies reveal the repulsive nature of this species' arrangements when compared to the decanethiol case.

Consideration of the calculated interaction energies show that the F-SAM molecules in the simulated densities lead to unfavorable arrangements, which are minimized when treating smaller densities, leaving more space for molecules to arrange. An indication can be found in the contribution of periodic images to the total interaction energy as shown in Figure 8.18, where a systematic reduction of the destabilizing contributions is observed.

8.7 Amino-Undecanethiol

Adding an amino group NH_2 and an additional carbon segment CH_2 to the C10-SAM, leads to the formation of the amino-undecanethiol molecule (NH_2 -SAM). This species is expected to have largely the same intermolecular interactions stemming from the backbone as in the case of C10. However the amino-headgroup adds additional features. The modified SAM molecules are schematically shown in Figure 8.14.

The surface models for the NH_2 -SAM were constructed analogous to the previous cases, starting from a two layered 2×2 gold (111) surface with one molecule per unit cell. The

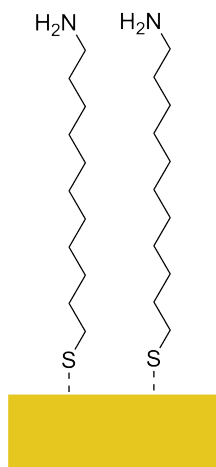


Figure 8.14: Schematics of the chemical structure for the NH_2 -SAM molecules attached on a metal substrate.

resulting unit cell organization is presented in Figure 8.15. It is directly visible, that the molecule arranges in a similar way as the unmodified C10-SAM molecule, with a tilting angle of $\phi_{2 \times 2}^{\text{NH}_2} = 26.76^\circ$ corresponding to an additional straightening in the alignment, leading to a total molecule length of $d_{\text{sd},2 \times 2}^{\text{NH}_2} \approx 14.48 \text{ \AA}$. Considering the interaction energy following Eq (11.1) leads to $E_{\text{int}} \approx -78.42 \text{ kcal/mol}$ which is of similar magnitude as in the C10 case. Contribution of the SAM periodic images following Eq (8.5) leads to $E_{\text{int}}^{\text{PER}} \approx -57.4 \text{ kcal/mol}$ also showing the similarities to the C10 molecule.

8.7.1 Explicit Intermolecular Effects (2 Molecules)

In the case of one molecule per unit cell, the head groups interactions with the periodic images allow no additional flexibility and the situation is largely the same as for the C10-SAM. Investigations for the inter molecular interactions were performed on the PBE-D3/TZV2P-GTH level by placing two molecules on a two layered 3×2 gold surface. The optimized unit cell is illustrated in Figure 8.16, where the top view in Figure 8.16a reveals a rotation of the head group along the molecular axis, whereas the side view illustrated in Figure 8.16b reveals minor tilting of the molecules with respect to the surface normal. The average molecular length could be determined as $\bar{d}_{\text{sd},3}^{2\text{H}} \approx 15.4 \text{ \AA}$ as well as the average tilt of $\bar{\phi}_{\text{H},3 \times 2} \approx 29.12^\circ$. Interestingly, the total interaction energy following Eq (11.1) leads to increased stabilization when compared to the case of the corresponding C10 molecules.

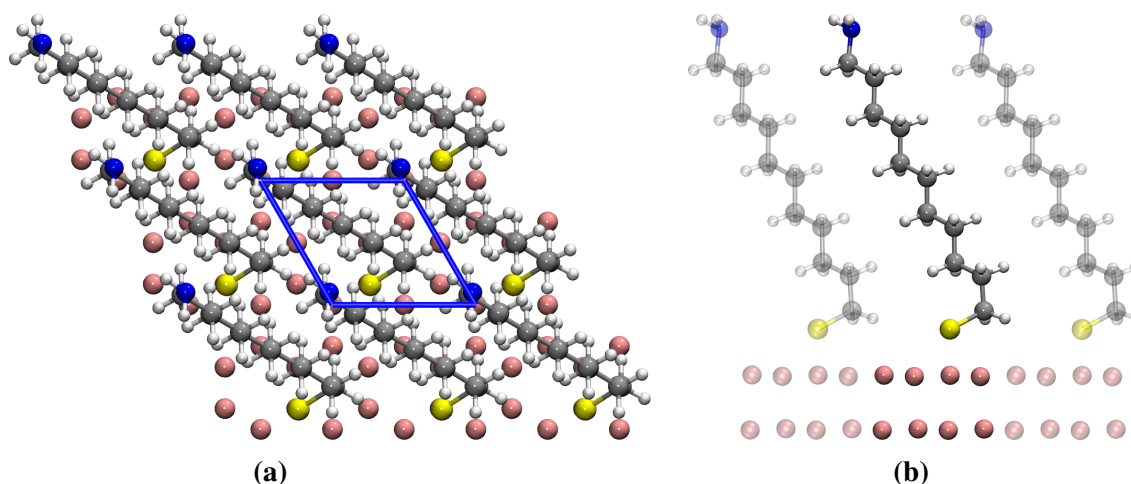


Figure 8.15: Unit cell arrangement for one NH_2 -SAM molecule on a 2×2 gold surface. a) Top view onto the unit cell of the amino-undecanethiol unit cell after optimization. The cell boundaries in the $x - y$ plane, are depicted with blue lines. b) Side view of the amino-undecanethiol unit cell, with the periodic images in transparent colors.

The side view reveals the already known alternating arrangement of H atoms in the carbon chains. However, a closer look shows the formation of hydrogen bonds between the amino groups which are contributing as a stabilizer. Consequently, the functional groups of the molecules are approaching each other by tilting the molecular “heads” of the adjacent chains. Considering the interaction energy of the molecules themselves when compared to an isolated gas phase situation following Eq (8.6) leads to $E_{\text{NH}_2, \text{int}}^{\text{MOL}} \approx 4.0$ kcal/mol which is more stable than in the C10 case where a value of $E_{\text{C10, int}}^{\text{MOL}} \approx 10.0$ kcal/mol is obtained. As shown in Figure 8.17, the distance between the hydrogen and the accepting nitrogen atom is determined to be $R_{\text{N-HN}} = 2.01 \text{ \AA}$, which corresponds to the usual definition of a hydrogen bond. The resulting effect is an additional stabilization of the system as can be seen from the interaction energies $E_{\text{int}}^{\text{PER}} = -97.2$ kcal/mol. Structurally, the formation of molecule pairs and tilting in the head groups orientation is observed. The behavior of the NH_2 -SAM is in large parts similar to the C10 case as can be seen from the minimal system. However the interactions within the unit cell differ due to the possibility of hydrogen bond formation, increasing the layer stability when compared to the raw case.

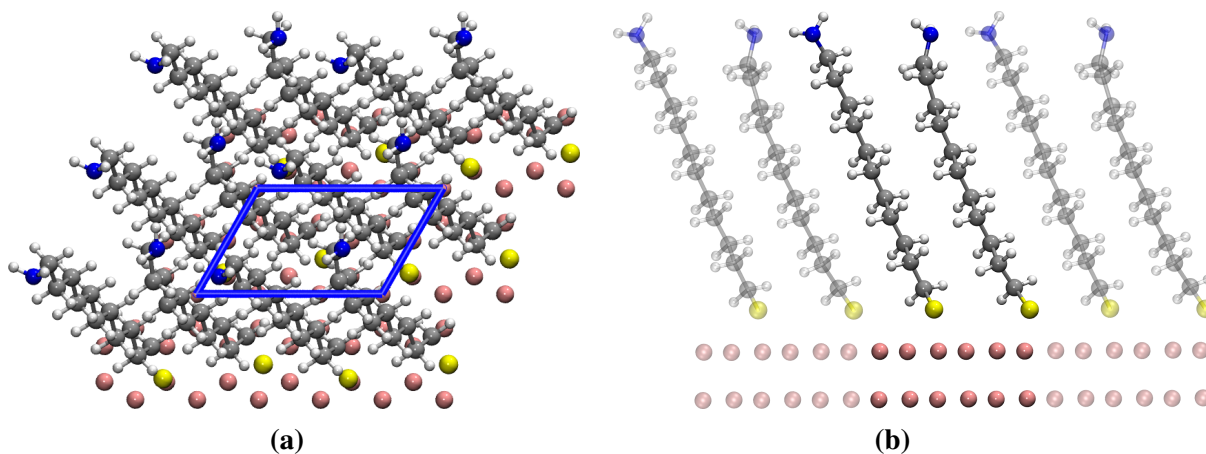


Figure 8.16: Unit cell arrangement for one NH_2 -SAM molecule on a two layered 3×2 gold surface. a) Top view onto the unit cell of the NH_2 -SAM after optimization. The cell boundaries in the $x - y$ plane, are depicted with blue lines. b) Side view of the unit cell, with the periodic images in transparent colors.

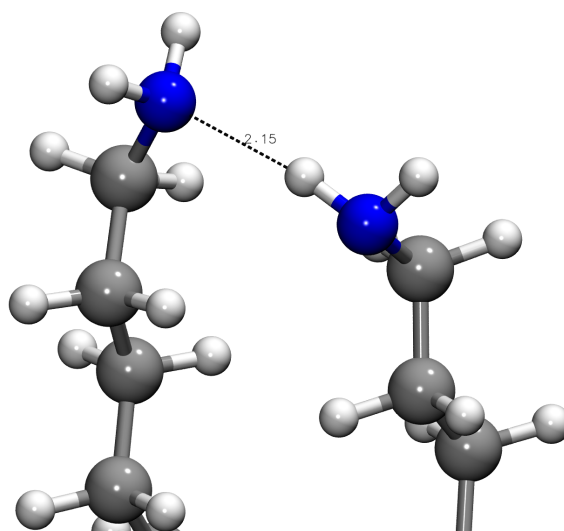


Figure 8.17: Zoom onto the hydrogen bond arrangement in the functional groups for two NH_2 -SAM molecules on a 3×2 gold unit cell, illustrating the formation of molecule pairs. The distance between the donor nitrogen and acceptor hydrogen in angstrom is indicated by a dashed black line.

8.8 Comparison of the investigated Species

The three investigated molecular species differ with respect to their intermolecular interactions. The constructed surface interfaces show different behavior of the molecules in terms of directly treated intermolecular interaction as well as interactions with periodic images. The main differences between the two contributions arise from the possibility to readjust the atoms within the unit cell, which is not possible in the case of the periodic images. Considering the total interaction energy when investigating the surface interfaces following Eq (11.1) shows the distinct differences between the molecules. In the case of the C10 molecule the smallest system shows stabilizing interaction given as $E_{C10,int} \approx -79.20$ kcal/mol which is almost identical to the case of the NH₂ SAM where the total interaction is given as $E_{NH_2,int} \approx -78.5$ kcal/mol. Contrarily, the total interaction for the F-SAM interface is given as $E_{F,int} \approx 71.25$ kcal/mol which means the arrangement is disfavoured. For the more densely packed systems the stabilizing interaction is reduced to $E_{2C10,int} \approx -29.7$ kcal/mol showing less stabilization through the interface. However the contrary is observed for the NH₂-SAM where an increase to $E_{NH_2,int} \approx -134.45$ kcal/mol was observed. Possible reasons lie within the formation of hydrogen bonds as well as the consequently smaller tilting angles. In the case of the F-SAM molecule, the arrangement becomes more favorable as well, as can be concluded from the reduction of interaction contribution to $E_{2F,int} \approx 6.58$ kcal/mol.

Intermolecular Interactions

Intermolecular interactions are compared when explicitly treating two molecules in the unit cell. Energies are obtained for surface optimized geometries when compared to isolated molecules. Single point energy calculations were performed using GPW-DFT with the PBE-D3 functional and GTH-TZV2P basis sets along side GTH-PPs. For the basic C10-SAM, the slightly repulsive nature of the considered molecules shows in the obtained interaction energy of $E_{C10,int}^{MOL} \approx 9.81$ kcal/mol. It is expected that total repulsion scales with the chain length of the investigated molecules. Calculation of the NH₂-SAM with an additional amino group (CH₂NH₂) shows less repulsive interaction between the molecules from the corresponding interaction energy $E_{NH_2,int}^{MOL} \approx 3.75$ kcal/mol. This behavior can be explained by the observed hydrogen bond between the molecules head groups. The repulsive effect is increased for the

8.8 Comparison of the investigated Species

fluorinated F-SAM, where a total interaction energy of $E_{F,int}^{MOL} \approx 34.51$ kcal/mol was obtained, showing the major difference when compared to the other systems.

Interaction with Periodic Images

The already observed trends for the intermolecular interactions between SAM molecules are further extended when investigating interactions with periodic images. Comparing the effects on one SAM molecule in the smallest system 2×2 as well as two molecules in the 3×2 unit cell, show continuation of the expected trends. The corresponding densities given as $\rho_{2 \times 2}^{-1} \approx 28.7 \text{ \AA}^2$ and $\rho_{3 \times 2}^{-1} \approx 21.5 \text{ \AA}^2$, indicate the reduction of the modelled density. For the F-SAM, additional interactions following Eq (8.5) are calculated to be $E_{F,int}^{PER} \approx -22.8$ kcal/mol in the case of the smallest system. However when treating a system of two molecules on a 3×2 surface the obtained interaction results in $E_{2F,int}^{PER} \approx -11.7$ kcal/mol, leading to less stabilization due to periodic images. However the differences are small when considering a density increase of $\approx 33\%$. This effect can be explained by partial compensation of the repulsive forces between molecules and periodic images due to staggered configuration. For the C10 SAM, the contrary effect is observed and the system is significantly more stabilized with the periodic contribution of $E_{2C10,int}^{PER} \approx -75.9$ kcal/mol, compared to the one in the smaller system $E_{C10,int}^{PER} \approx -49.3$ kcal/mol. A similar effect is observed for the NH₂-SAM where interactions are expected to be similar, leading to a stabilization of $E_{NH_2,int}^{PER} \approx -57.4$ kcal/mol to $E_{2NH_2,int}^{PER} \approx -97.2$ kcal/mol. The observed trends are furthermore extended by utilizing molecular configurations for unit cells modeling other densities. For their construction, the optimized metal surfaces for Pt and Cu were used as reference. The results of the obtained interaction energies per considered SAM molecule for various densities are shown in Figure 8.18 further highlighting the differences between the molecular species. It can be concluded, that for the C10 and NH₂-SAM systems of similar densities considering only one molecule, are generally more stabilized due to forced interaction with the periodic images and little flexibility. This can be seen from the large jump in the interaction energies around $\rho^{-1} \approx 22 \text{ \AA}^2$ where two systems consisting of one and two molecules respectively are treated. In the case of F-SAM systems it can be seen, that for smaller densities, the interactions arising due to periodic images are destabilizing which can be associated with the limited surface area

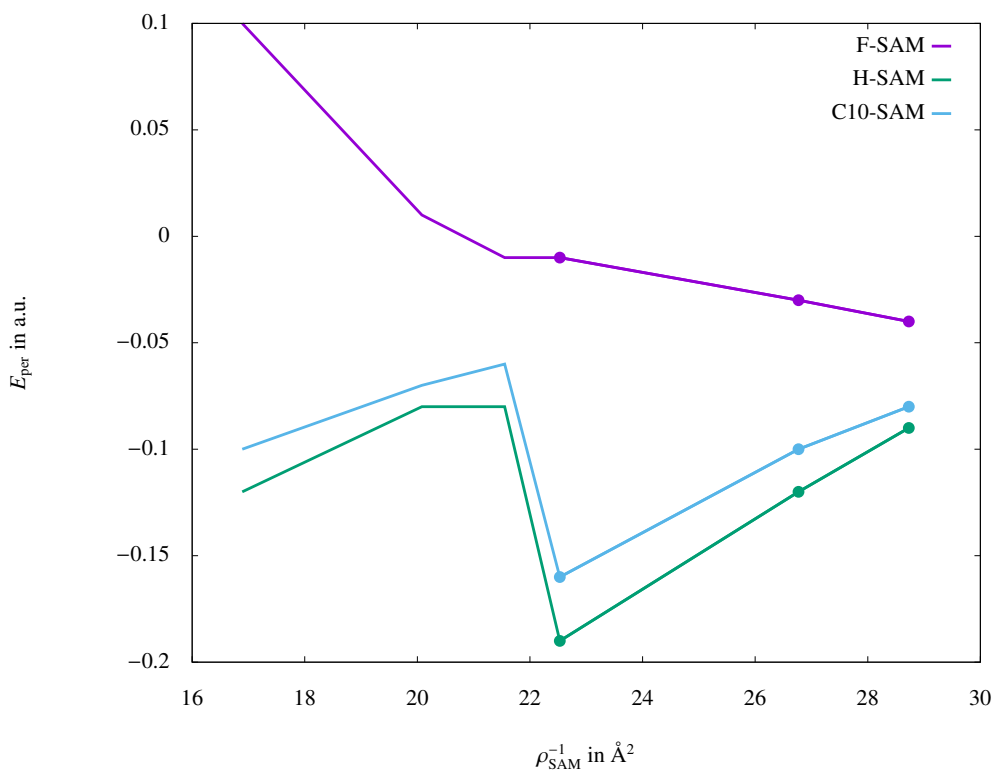


Figure 8.18: Periodic contribution to the interaction energies in a.u. following Eq (8.5) per SAM molecule, for various SAM-densities. Energies for systems with one molecule per unit cell are indicated by dots.

available for rearrangements. The effects for the F-SAM are generally smaller as possibilities for rearrangement are small.

8.8.1 Discussion of the Setup Preparation

The described methods for the creation of surface interfaces were used as initial setup for follow up simulations of field evaporation. The investigation of the initials simplifies further interpretation of the follow up field desorption model. For all three investigated types of SAMs, differences between gas phase and solid state calculations were considered. The contributions due to periodic images as shown in Figure 8.18 reveals the stabilizing effects of the hydrogen bridges in the case of amino-undecanethiol when compared to the simple decanethiol molecules. Also the sudden stabilization upon introducing a second chain can be associated at least partially to the possibility of alternately arranged molecules. In principle the position of sulfur on the metal surfaces is not uniquely defined, but due to expensive costs

8.8 Comparison of the investigated Species

and limited relevance for the follow up calculations, a separate investigation of attachment positions was not conducted. In analogy to literature known investigations, it was assumed that the positioning of sulfur in the center of four gold atoms is the most stable^{194,195,200}. The periodic contribution to the interaction energies calculated for various SAM densities, showed crucial differences. In the case of F-SAM molecules, the repulsive contribution is lowered correlated with lower SAM density. Contrarily, in the case of C10 and NH₂ molecules, the lowering of the density reduces the stabilizing nature of the surface configurations. In summary the presented study shows the used methodology for the interface creation and shows specific intermolecular interactions between the investigated species. The observations simplify follow up interpretations in the performed simulations of field evaporation Section 9. It could be observed throughout the investigated systems, that models using only one SAM molecule per unit cell miss important intermolecular interactions. However they act as a good starting point to set up the used model procedure. Unit cells including two molecules per unit cell, cover the main interactions between SAM-molecules. Systems of two molecules on a 3×2 surface of metal atoms are chosen as the “workhorse” for follow up investigations.

Chapter 9

Field Evaporation of Self Assembled Monolayers

This chapter contains information published in¹⁹. Firstly, the main considerations when treating electric fields in the periodic case, within the DFT-GPW framework as implemented in CP2K, are illustrated at the example of H₂. An additional aim is to provide the reader with technical details of the performed simulations for follow up work within the field. Furthermore, observed difficulties during the simulation process as well as collected information on how to resolve SCF convergence issues during the simulations are provided. The work described here was carried out within the collaborative research center CRC 1333 to support the experimental development of atom probe tomography to be applicable to organic molecules.

After an initial overview on external electric fields within PBCs, the basic principles of atom probe tomography (APT) are reviewed to allow for interpretation of the results of the performed simulations. Starting by analyzing the possible effects of the electric field on the utilized metal surfaces, the desorption field strength for metal atoms are calculated. Afterwards, the desorption process of the previously analyzed SAMs introduced in Chapter 8, is simulated by means of GPW-DFT as well as CCM implementation of GFN-xTB in the CP2K program package. The results of both methods are compared with each other highlighting the differences in the description. Furthermore, differences between surface simulations and the experimental data as well as simulations of cluster models are discussed. The chapter

closes with a discussion of possibilities for general improvement of the utilized models, their applicability as well as difficulties in the experimental interpretation.

9.1 Consequences of Applied Fields within PBCs

Simulations of field evaporation require calculations in the presence of electric fields. In order to evaporate segments of an investigated system, the electric field strength is incrementally increased until the bond breaking threshold is reached. Using this approach, leads to a static suppression of the dissociation barrier. The effects presented in Section 3.7 are further illustrated by example calculations for the H₂ dimer. All reference calculations performed aim at illustrating electric field as well as specific PBC related effects. Both are of great significance, as they impact all calculations utilizing the PERIODIC_EFIELD routine within CP2K. In the following, the electric field strength is denoted in a.u. = $E_h/(a_0e)$ with the Bohr radius a_0 , the elementary charge e and the Hartree energy $E_h \approx 27$ eV. The connection to SI-units is given by 0.01 a.u. ≈ 5.14 V/nm.

9.1.1 Bond Profiles in the Gas Phase

In the presence of external electric fields, molecular bond profiles are influenced by the field's potential. For H₂, the situation in the presence of a homogeneous electric field is exemplary shown in Figure 9.1. The bond profile was obtained by varying the distance r between the two H atoms and performing a single point energy calculation for every displacement. The obtained energies were calculated using the TURBOMOLE program package and an external field of $\mathcal{E} = 0.04$ a.u. and the B3LYP-D3 functional with Ahlrichs def2-TZVPP basis set, for interatomic displacements from 1.3 \AA to 12.5 \AA . The obtained curvature shows a linear section for larger distances at which the field's influence is dominating.

A comparison of bond profiles calculated on the same level for varying field strengths is illustrated in Figure 9.2. Increasing the strength of the applied electric field, leads to suppression of the dissociation barrier of the molecule. This effect can be directly understood from Eq (3.4) as described earlier. At the critical field strength, the atoms dissociate due to the separation of both atoms is more favorable than the bonding configuration. This can be

directly associated with the vanishing barrier in the bond profile for field strengths larger than $\mathcal{E}_{crit} = 0.15$ a.u..

As can be seen from Figure 9.2, the bond distance is increased while ramping up the field strength. For incrementally ramped field strengths, geometry optimizations are performed with constrained position for one H atom to avoid both atoms from drifting. Searching for the minimum energy for a given configuration starting on the left side of the barrier in Figure 9.1, leads to successively increased bonding distance. However, when the critical field strength \mathcal{E}_{crit} is reached, the H atom dissociates. This leads to a non convergent geometry optimization, resulting from the continuous decrease of energies in field direction, preventing convergence during geometry optimization for vanishing barriers. This is the direct result of the electric potential dominating this regime, which results in a linear decay in energy in the case of homogeneous fields.

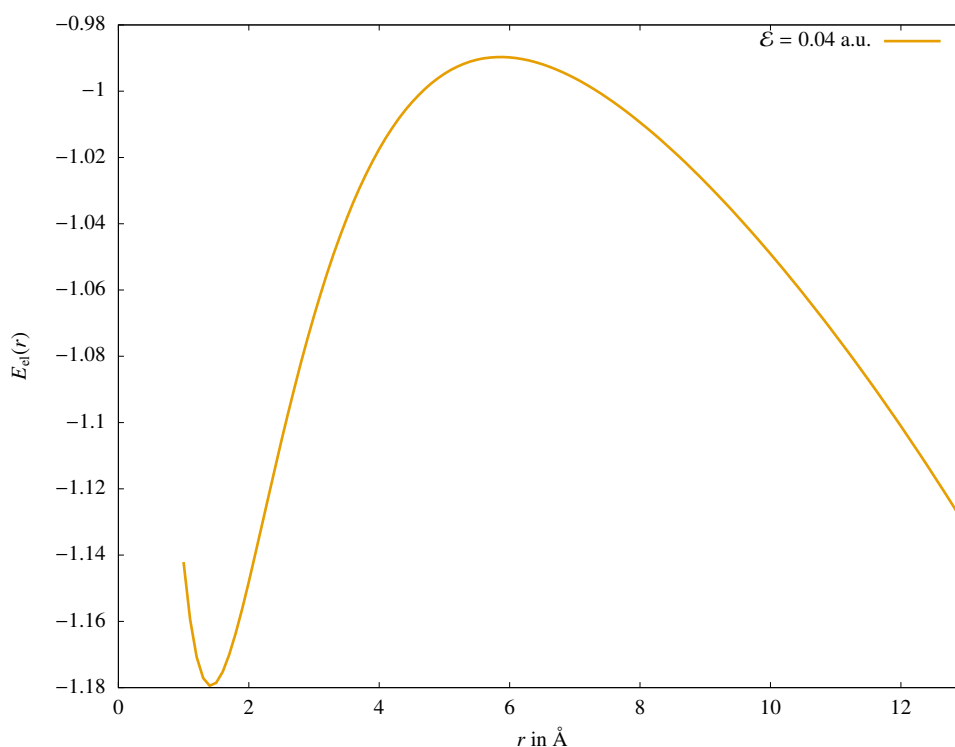


Figure 9.1: Exemplary bond profile considering the electronic energy E_{el} , calculated for H_2 in the gas phase. Single point energy calculations were performed on the B3LYP-D3/def2-TZVPP level, for an applied field of $\mathcal{E} = 0.04$ a.u. in displacement direction, every 0.01 Å.

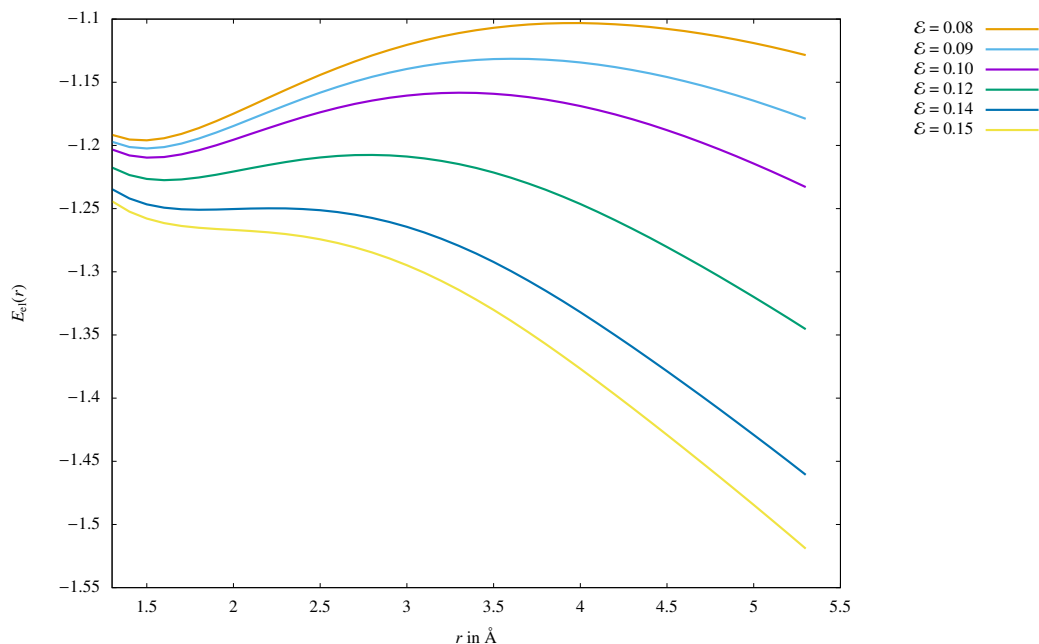


Figure 9.2: Bond profiles of H_2 in vacuum calculations, performed with the TURBOMOLE program package for various field strengths in displacement direction. The single point energy calculations were performed on the B3LYP-D3/def2-TZVPP level every 0.02 Å.

9.1.2 Bond Profiles within PBCs

Initial investigations of the effects on molecular properties within PBCs show crucial differences when comparing to gas-phase calculations. In the example of H_2 , fundamental effects are illustrated. For systematic comparison, it is suitable to firstly investigate the effects arising purely as a consequence of the treatment within PBCs. Due to interactions with periodic images, the bond profiles for gas-phase and periodic bulk calculations are only equivalent for large cell sizes. In this case, only the elementary cell contributes to the total energy, describing the same molecular arrangements. However, when comparing results for smaller cell sizes, bond profiles show differences. By means of a purely didactic example, the H_2 bond profile was calculated for a cubic unit cell of 10 Å. Calculations were performed using GPW-DFT formalism with a TZV2P-PBE-GTH basis sets and the PBE functional. Bond profiles were obtained by varying the inter-atomic distance r between both H atoms within the unit cell along the molecular axis. For comparison, calculations were performed for the closed shell and open-shell system as well as open shell within a periodic electric field $\mathcal{E} = 0.01$ a.u..

Figure 9.3 shows the influences of PBCs as well as the electric field on the obtained bond profiles.

Consequences of the periodic description can be identified by comparing bond profiles with the performed gas phase equivalents. The calculated profiles within solid state description are symmetric with respect to the unit cell center, due to the periodic extension of the system. The example system was constructed such that these interactions are negligible for small distances, however upon passing the middle of the unit cell, they lead to the reversed bond profile. After crossing the middle of the cell, the displaced reference atom approaches the periodic image of the second atom placed in the origin. As a result, the reversed bond profile leading to the symmetric course of the curve is obtained. In the open shell case, the separation of the molecular orbital in two separated atomic ones is favored approximately at distances larger than $r \approx 2.45 \text{ \AA}$ which is the point where the curve differs from the closed shell reference. Correspondingly it is again energetically disadvantageous when $r \approx 7.55 \text{ \AA}$ is surpassed, as the bonding molecular orbital with the periodic image is preferred.

Consequences of the electric field can be determined by comparing the energetic distribution for both cases. For small distances between the H atoms, the field leads to a slight shift in energy when comparing with the field free case. However, the minima of the obtained curves do not overlap, indicating a new equilibrium bond distance within the applied field, as was also observed in the gas phase case. The bond profiles are qualitatively similar for small distances. However, when inter-atomic distances are increased, the field contribution to the total energy becomes dominant and, thus, the field lowers the overall energy drastically as can be understood from Eq (3.53). The field therefore favors further increase of the inter-atomic distance as expected from the gas phase references. As can be seen from the barrier heights, the applied field of $\mathcal{E} = 0.04 \text{ a.u.}$ leads to a drastic reduction of the barrier. In the following, further increase of the applied field strength leads to a completely vanishing barrier, upon reaching the critical field strength of $\mathcal{E}_{\text{crit}} \approx 0.145 \text{ a.u.}$. Fragmentations related to this effect enable simulations of field evaporation as investigated within this thesis.

Effects of the Field Implementation in PBCs

Comparison of the field effects obtained by simulations with PBCs as shown in Figure 9.3 with the gas phase Figure 9.1, reveal quite striking differences. A main issue that has to be addressed, is the divergent behavior in the cell's center in direction of the applied field. This effect corresponds to the electrostatic potential drop in the cell's center as a consequence of the PBCs as presented in Section 3.6. This effect can be partially avoided, by increasing the system size in field direction. Doing so, leads to vanishing electron density in the vicinity around the potential drop, avoiding divergent behavior. However, the differences arising due to the utilized formalism as highlighted in Chapter 3, lead to additional differences in the observed energetic behavior. While the overall slope of the bonding profiles are almost equivalent for small distances, higher distances lead to different slopes of the profile. This indicates that field influences within PBCs differ fundamentally from the gas phase analog as a consequence of the application of the modern theory of polarization (MTP). From the fundamental equations explained in Section 3.2, it is directly visible that the interaction of the electric field with the bulk polarization is described fundamentally different from the gas phase case.

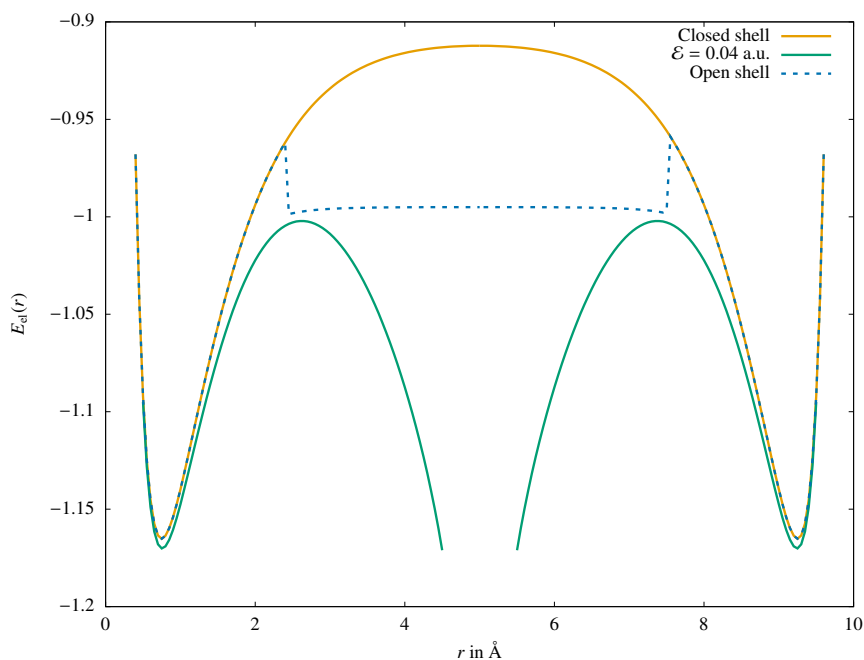


Figure 9.3: Bond profiles, obtained by singlepoint energy calculations for displacements r in field direction oriented in parallel to the molecular axis. Results for closed shell calculations are indicated in yellow, open-shell results in dashed turquoise and results considering an applied field of $\mathcal{E} = 0.04$ a.u. in green.

9.1.3 Charge Shifts and Equilibrium Distance in Electric Fields

As an illustrative example, the charge shifts upon application of an external field can be investigated by calculating the partial charges of both atoms after performing a geometry optimization with incrementally increased field strength. Optimizations using a cubic unit cell with lattice vectors of $a = b = c = 20\text{\AA}$ were used. One H atom was placed at the origin and constrained during the optimization to avoid drifting throughout the unit cell. Calculations were performed on the PBE-D3/TZV2P-GTH-level with corresponding GTH-PPs and a plane wave cutoff of 300 Ry. The obtained equilibrium distances over the applied field strengths until reaching the critical field are shown in Figure 9.4. The H_2 molecule was observed to dissipate in two supposedly charged H fragments after reaching $\mathcal{E}_{\text{crit}} \approx 0.145$ a.u.. An analysis of the individual atom's partial charges, is shown in Table E.4. It can be concluded, that the charge splitting and therefore the polarization is directly related to the applied field strength. For smaller field values, the distance remains almost static and therefore the charge separation is mainly determined by \mathcal{E} .

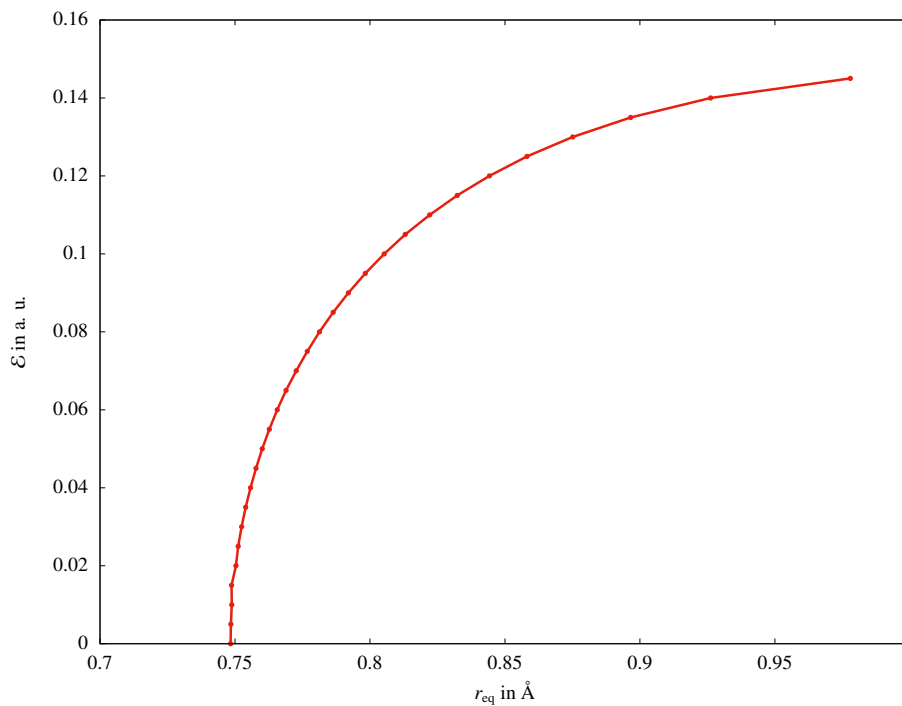


Figure 9.4: Equilibrium distance r_{eq} of H_2 , calculated by geometry optimization for incrementally increased field strengths.

9.1.4 Polarization in Electric Fields

A significant property of principle interest is the polarization resulting from the previously investigated charge shifts. For the applied field strengths and optimized geometries of the previous investigation, the polarization was calculated following Eq (3.18) and its dependence on the applied field strength is shown in Figure 9.5. Increasing the strength of the applied field also increases the systems total polarization as the charge shifts within H_2 are coupled to the new equilibrium distances. The initially linear behavior upon $\mathcal{E} \approx 0.07$ a.u. can be attributed mostly to a pure field induced polarization as the equilibrium distance shown Figure 9.4 is only slightly adjusted. Larger fields lead to a stronger polarization increase due to the increase of the dipole contribution associated with the new equilibrium distances as well as the charge separation. Following the visible trends, it can be assumed that upon dissociation of the H_2 molecule, the systems polarization would diverge.

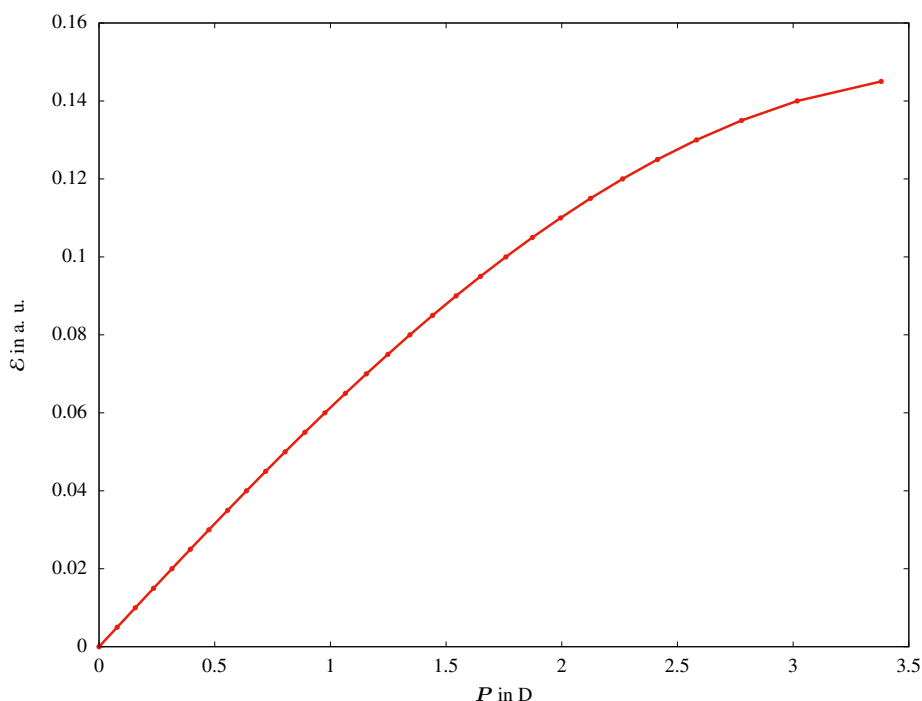


Figure 9.5: Total polarization P calculated at the equilibrium distance for H_2 for incrementally increased field strength.

9.2 Basics of Atom Probe Tomography

Microscopy marks a crucial aspect of modern day technological advancements, as it offers a scope onto properties that are classically out of view for human beings²⁰¹. Since the early approaches of the late 80s, a newly developed methodology of microscopy has come to life, namely atom probe tomography (APT). APT can be characterized as a destructive bulk analysis technique that allows for a materials characterization on an atomic scale. The possible atomic resolution as well as the time of flight (TOF) mass spectra analysis, offer insights of a materials chemical composition at the nano scale²⁰². In typical experimental applications, it has been successfully used for characterizing metals²⁰³, semiconductors²⁰⁴, ceramics²⁰⁵ and recently also gained influence in geoscience²⁰⁶. However, the analysis of organic molecules such as the investigated SAMs are a development of the recent years with some, but few, studies already conducted^{9,15,16,207}. In general multiple ways of performing atom probe tomography experiments are distinguished within the community, depending on the specific field of application. An exemplary overview is provided by Amouyal⁷ as well as a more recent one with applications to the geosciences²⁰⁶.

In order to analyze a material with APT, the surface of interest, is placed on needle shaped metal tip and subjected to a voltage in the kV regime, leads in combination with the low apex diameter to high electric field strengths ($\mathcal{E} \approx 10\text{V/nm}$). Due to the application of electric fields, surface atoms are ionized, repelled from the surface and subsequently accelerated in the electric field to be detected by a position sensitive detector. To trigger a controlled evaporation, the field strength resulting from the applied DC voltage is increased until a value remaining below the critical field strength $\mathcal{E}_{\text{crit}}$ for evaporation. The remaining gap in the field strength, is then overcome by applying a pulse in form of an electric voltage or a laser, triggering the evaporation event. The start of the pulse application, is associated with the initiation of the corresponding time of flight measurement, which is terminated as the ionized fragment hits the position sensitive detector. A crucial task remains then in the reconstruction of the particles trajectory while flying, in order to determine its initial position²⁰². The ion trajectory follows the electric field lines of the standing DC voltage between sample and a position sensitive detector, acting as a counter electrode. It has been elaborated that only the topmost atoms are desorbed, allowing for a layer to layer reconstruction of the 3D sample^{3,4}. The

9 Field Evaporation of Self Assembled Monolayers

exact algorithms for reconstruction are still a topic of research and constantly improved⁵. A typical setup scheme for the vacuum chamber is depicted in Figure 9.6, illustrating the most important features during the measurement. When investigating SAMs as the sample of

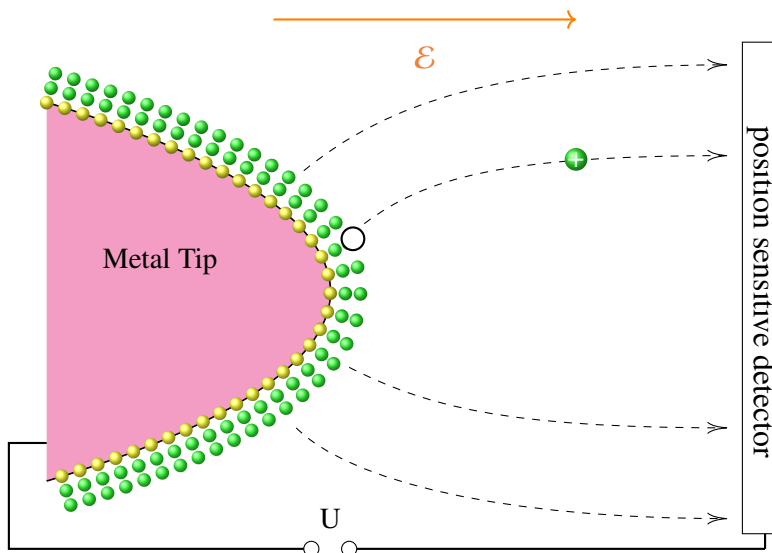


Figure 9.6: APT setup when performing experiments with the metal tip in pink with the attached sample molecules depicted in green/yellow and the flight trajectory to the detector upon evaporation in dashed lines. The applied DC-Voltage is illustrated by a capacitor connected to a voltage U , resulting in the electric field \mathcal{E} , determining the flight trajectory of ions in the chamber.

interest, the obtained mass spectrum shows the evaporation of a multitude of segments from the carbon chains instead of single ionized atoms^{15,16}. The electric field penetrates the layer and breaks the carbon chains into various fragments resulting in complicated combinations of $C_xH_y^{n+}$ ions, which are not uniquely identifiable. Possible reasons for that are to be found in the determination of mass to charge ratios when detecting fragments as well as a lack of understanding in the physical mechanism of field evaporation for organic species.

Simulations of mass spectra without external electric fields are available in the literature. They employ an ever-evolving multitude of methods, including, for example, models based on well-established statistical theories such as quasi-equilibrium theory (QET) and Rice-Ramsperger-Kassel-Marcus (RRKM) theory. The fragments charge distribution is determined by evaluating the ionization potential. Furthermore, the abundance of data on mass spectra also allows for the increasingly popular use of neural networks in simulation. In contrast to these examples, in this thesis, a direct approach is taken which does not require statistics

or sampling methods for finding possible stable fragments or reproducing the distribution of peaks in the mass spectrum. Instead, it focuses explicitly on investigating the process of field-assisted fragmentation itself. Investigation of the effects arising from the external field prior to the evaporation event, including charge shifts induced polarization as well as structural reorganization for the investigated surface models.

9.3 Simulations of Field Evaporation using GPW-DFT

The simulation of the evaporation process in APT, was conducted specifically for surface models. Therefore, calculations using the GPW-DFT approach with external fields were performed. The main target was the direct analysis of the evaporation mechanisms in the solid state. Just as the experiments on the field evaporation of organic molecules are a rather new development, so are the theoretical examinations of those molecules. Until present day, only limited preliminary studies have been published with none besides¹⁹ targeting organic molecules in the solid state. A general workflow for such simulations utilizing a surface description will be presented in the following for metals as well as for SAMs on gold surfaces. A continuum picture of the electrostatics throughout the SAM-surface interface¹⁹. A small fraction of the tip of the SAM-covered surface can be approximated planar. The electric field is applied in parallel to the surface normal. Within the gold metal bulk, the electrostatic field vanishes and therefore, the electrostatic potential is constant. A positive surface charge develops at the interface between gold and the dielectric SAM, which exactly screens the field within the conducting metal. The SAM is polarized upon the application of the electric field, which leads to another layer of surface charges. The field within the SAM is shielded by the dielectric and therefore it is smaller than in the vacuum space above it. This is indicated by the smaller slope of the potential with respect to the z coordinate. Electrons are pulled into the bulk until enough electron density is withdrawn to cause a bond to break. The resulting cation is then accelerated away from the surface within the vacuum region by the electric field. The aim of using the SAM surface interface created as described in Section 8.4, is the modeling of the experimentally used metal tips. A typically used system consists of a capped gold tip with an apex diameter of about 25 nm with a SAM placed on its surface in a strong electric

field. The computational restrictions require the choice of smaller structural models, surface fragments with only one to two SAM molecules in the model are at the edge of the available computational resources. In such small surface areas, the surface can be approximated to be planar. The electric field is applied perpendicular to the surface. Even though it is known that most of the spherical gold tips is amorphous or in a high-index surface direction, a (111) surface was chosen as the model of choice to demonstrate the possibility of conducting such simulations. The electric field of the APT apparatus was modeled as a homogeneous external electric field, which is included as a linear external electrostatic potential in the Hamiltonian and the modified Kohn-Sham equations are solved in a self-consistent manner as described in detail in Section 3.6.

9.3.1 Technical Details for Performing Simulations of Evaporation Sequences using GPW-DFT

As the subject of APT simulations is a rather unique and new topic, the pure setup “construction” was a main issue encountered within this thesis. The following section aims at providing the reader with difficulties the author encountered during his extensive study of the used method. Furthermore, possible explanations for the underlying problems are discussed in order to provide an idea of the challenges that arose during the setup creation. A fundamental desire of the author, lies in providing enough information to simplify the life of following researchers using the same or similar sets of methods. It has to be noted, that possible transferability of the explained concepts, highly depends on the used DFT formalism as well as the field implementation.

During the authors attempts at simulating the evaporation of molecules from surfaces using the GPW-DFT framework and PERIODIC_EFIELD routine within CP2K, multiple convergence difficulties concerning SCF and Geometry optimization were observed. To simplify following works, the used solutions based on the experience gained during the authors studies, are briefly explained. The arising main issues can be categorized in three cases:

- Non-convergent SCF due to bad initial guess of the wave function for higher field strengths:

- The case within a field step.
- The case for too high initial fields.
- Non-convergent SCF due to lack of vacuum spacing for higher field strengths.
- Non-convergent geometries due to badly described overlap matrix elements for non sufficiently large EPS_DEFAULT cutoff values.

The crucial aspects leading to the encountered difficulties and possible theoretical explanations are discussed in the following.

The Importance of the Initial Wave Function

Using the atomic guess for construction of the initial Slater determinant used for entering the SCF-calculation as described in Section 2.6.3, is suitable when the molecular density is expected to be closely related to the atomic one. In the presence of strong external fields, this is not the case due high induced polarization and as a result a convergent SCF-cycle is usually not achieved. This behavior can be accounted for by calculating energies for wave functions corresponding to successively increasing fields. Furthermore, the calculation of low accuracy wave functions followed by an increase in accuracy is required in difficult cases, leading to a double loop optimization for the initial wave function. This effect is a direct consequence of successive charge reorganization that occurs when calculating densities in external fields. As described in Section 9.1, distortions of the electron density from the field free case, are potentially large. To account for these in an energy optimization, flexibility in the basis set is required. Naturally, a larger number of basis functions including polarization functions as well as more diffuse ones are suitable. However, more degrees of freedom in the optimization lead to the limits of computational affordability. Additionally, more flexible basis sets potentially require a more accurate description of the overlap matrix and a careful tuning is required in many cases. Using more flexible basis sets allows for the successful description of potentially larger density changes during the SCF energy calculation. When fields are applied, the initially calculated density with more flexibility, was observed to be far off the equilibrium in many cases, which creates another potential failure within the optimizations. To avoid this fact, an initial guess from a density calculated for a smaller basis can be used.

The Importance of the Geometry

Using geometries in electric fields that are far off the realistic one, lead to similar effects as described in Section 9.3.1, however the underlying reason is reversed. When using geometries far off the equilibrium in electric fields, the change of geometry and charge reorganization during the optimization cause difficulties between the steps of a geometry-optimization. This is a direct result of a bad initial guess for subsequent geometries, that possibly lead to large changes in the charge distribution. Atoms are rearranging in a minimum that only exists for the given geometry due to strong distortion of the underlying potential energy surface. As a result the evaporation of rather unlikely species was observed, e.g. (F^+) ions when applying a field to a F-SAM molecule lying on the surface rather than being relaxed in a lower field where it aligns with the field axis. Subsequent optimization of the geometries is therefore the natural choice for the performed simulations. The necessary workflow is summarized in Figure 9.7. However, as previously described, the initial guess of the wave function plays a crucial role too. Therefore, a careful scheme of subsequent optimizations for both initials, geometry and wave function, is required. Combining both described considerations, leads the required triple loop optimization for really difficult systems. However, for most systems a geometry optimizations for every field increment is sufficient.

The Vacuum Spacing and Periodic Potential Drop

Due to the periodic potential drop described in Section 3.6, the vacuum spacing in direction of the surface normal, is crucial for convergent SCF-cycles. In order to avoid parts of the density being influenced by the divergence in the electric field, it has to be 0 in the surrounding vicinity. The cutoff values for possible overlap elements within the density matrix entering Eq (3.53) has to be chosen large enough to avoid the consideration of long tails in the density. However, from the authors experience, the basis set expansion used for the construction of polarized Wannier functions as described in Section 3.3, requires a high minimum level of accuracy when calculating the overlaps for construction of the density in applied fields (converging SCF). The possibility to tune the accuracy when calculating overlaps by the EPS_DEFAULT value is therefore limited. Other ways to account for the potential drop rather than increasing the vacuum spacing are unsuited in most cases. Here, a balance between additional computational

costs arising from the increased unit cell (compare Section 2.11.1) and the required density fall-off, has to be found. The vacuum spacing was chosen ad-hoc from experience, rather than convergence tests, for the simulations described in the following. Additionally, they are chosen larger than necessary for calculating the energy for the initial geometry during the optimization process. This directly accounts for the fact, that evaporated fragments are simulated up to a cutoff value of 5 Å from the remaining molecules.

Accounting for the Difficulties

Acknowledging the previously described difficulties, leads to the mentioned triple loop scheme as illustrated in Figure 9.7, which could also be used in the most difficult cases and is explained in the following. The initial optimization of the density has two main starting schemes indicated for the initial wave function ρ_{init} . The left-hand approach in Figure 9.7 directly uses the optimized wave function at zero field for the next field step, whereas the right-hand approach uses an iterative application of the external field with ρ_{opt} of the previous step as new initial. The prior approach leads to a quick initial guess, which is usually applicable for low fields. The latter approach constructs an initial guess based on a field \mathcal{E}_{aux} close to the simulated one \mathcal{E} , which simplifies restarts for geometries optimized in high fields e.g. after fragmentation. Depending on the quality of the initial geometry \mathbf{R} , the free energy functional $\mathfrak{F}(\mathbf{R})$ following Eq (3.53), is already in a geometry minimum which allows for the direct increase of the field. In case of non convergence, the geometry is further optimized using recalculation of ρ_{opt} for the new geometry \mathbf{R}_{new} until $\mathfrak{F}(\mathbf{R}_{\text{new}})$ is in a minimum. For high electric fields, the evaporation of fragments is possible, leading to increased distance of the fragment's atoms from the remaining SAM molecule, which is deleted after a minimum distance of $\bar{R}_{\text{fragment}} > 3 \text{ \AA}$ is surpassed. The density of the remaining system is then successively re-optimized using the right-hand optimization scheme to reenter the optimization sequence. This process can be repeated until either full fragmentation is reached, or the reference field strength \mathcal{E}_{max} for nonphysical events is surpassed.

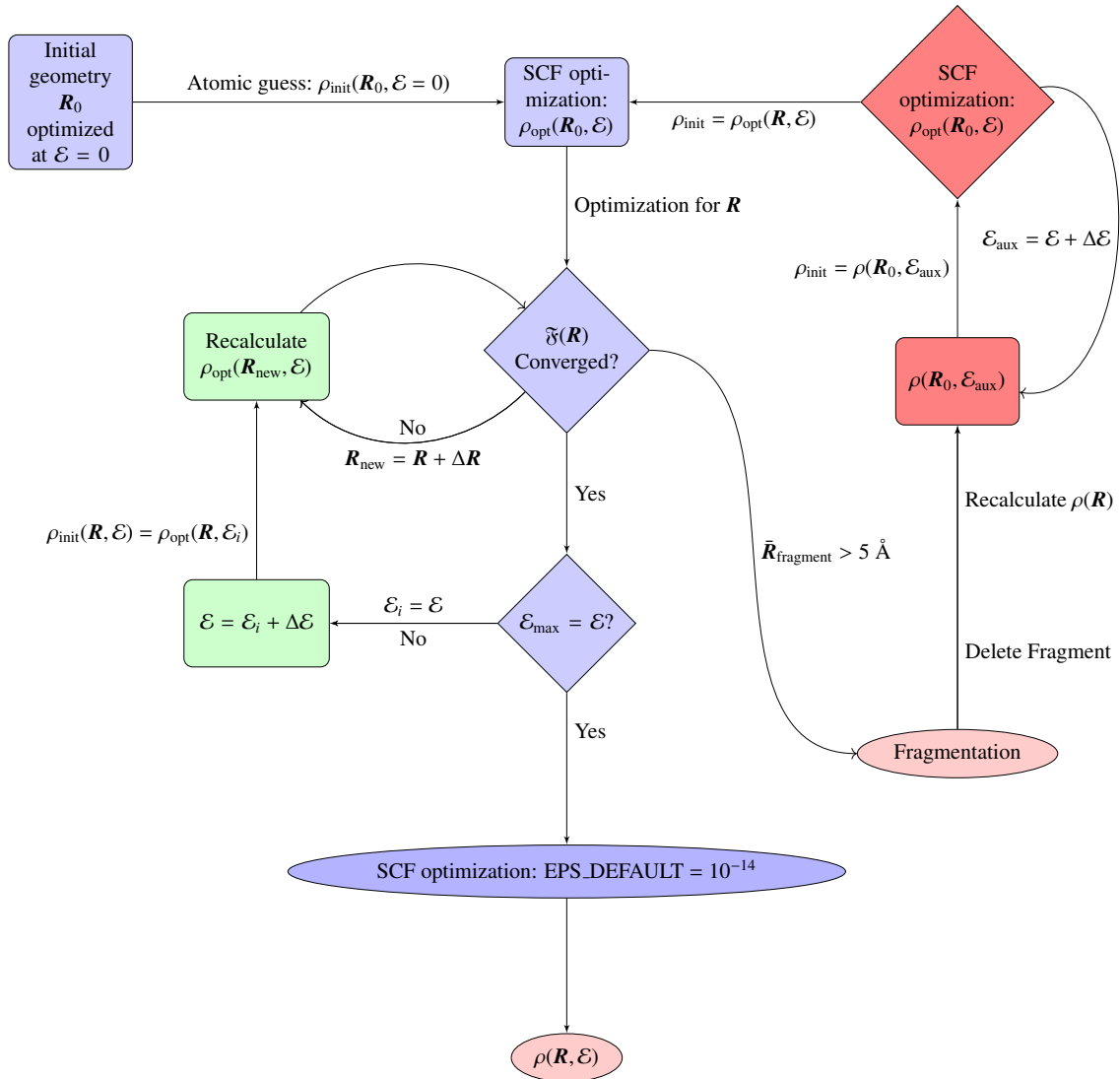


Figure 9.7: Schematic workflow of the loops necessary for the most complicated calculations during the evaporation process.

9.3.2 Simulation of Evaporation for a pure Gold Surface

In order to investigate the models applicability, the desorption of single gold atoms from a (111) surface was studied. Simulations were performed for a reconstructed three-layered 3×2 super-cell. The uppermost layer was free to relax during the geometry optimization, while the lower ones remained fixed. A profile of the surface reconstruction for three layers can be seen in Figure 9.8a with the reconstruction performed as previously described in Section 8.3.1. The obtained surface was subjected to an external field whose strength was incrementally increased with geometry optimization of the uppermost layer at every step, until the first

gold atom was ionized and evaporated during the field step $\Delta\mathcal{E} = 0.05 - 0.06$ a.u.. The rearrangement during the desorption process can be seen in Figure 9.8b. Here, the periodic images of the desorbed atoms have been omitted for better visibility. The desorption field strength for gold $\mathcal{E}_{\text{thresh}} = 0.06$ a.u., can be considered an estimate for a maximum threshold for the desorption of molecules from the surface and is further referred to as such. Even though the effective field strength at the gold surface may be weakened by the polarization of the attached molecules, higher fields are expected to desorb atoms of the gold surface and therefore the whole molecules. Within the model framework, if molecules do not desorb upon applying the desorption field of gold, it can be concluded that the system's description includes too many stabilizing effects. This is a direct consequence of keeping the gold atoms positions constrained during the optimizations.

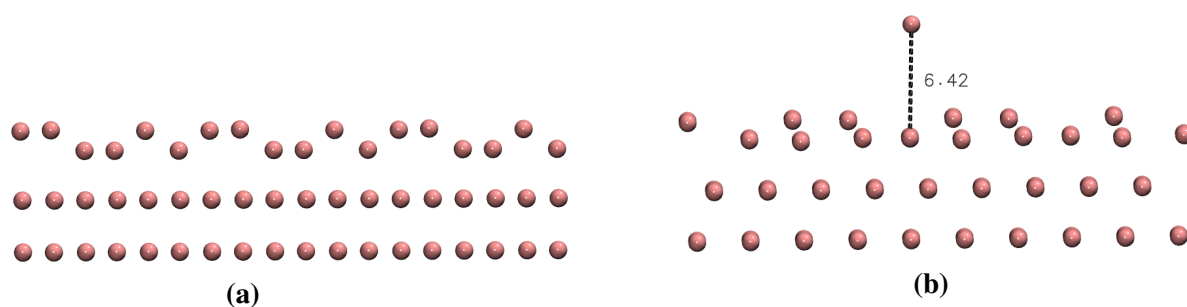


Figure 9.8: Setup for the evaporation of gold, a) The reorganization of the surface after relaxation of the uppermost layer at $\mathcal{E} = 0$ a.u., b) Desorption process during the field step $\mathcal{E} = 0.05 - 0.06$ a.u..

9.3.3 Evaporation of Decanethiole from Gold

Following the previously introduced nomenclature, decanethiole is further referred to as C10-SAM (compare Figure 8.3). As the initial system of reference, simulations of field evaporation for surfaces of the unmodified C10-SAM are simulated in order to gain insight on its fragmentation behavior. Systems were constructed for a two-layered 2×2 and 3×2 gold surface respectively. The simplest model consisting of one SAM molecule acts as initial reference for comparison with gas phase single molecule calculations. The larger surface model consisting of two molecules on a 3×2 surface, leads to additional intermolecular interactions and a possibly better description of the surface. From the surface study it follows that the systems with two molecules lead to higher total interaction energies, however per

9 Field Evaporation of Self Assembled Monolayers

molecule they are lower. Therefore, it can be assumed, that the molecules in the 2×2 case are initially more stable on the surface as discussed in Section 8.8. For densities corresponding to the experimental observation, the unmodified C10-SAM molecules interact only slightly with the periodic images when compared to the other investigated species. This allows for a clean reference point for follow up simulations, as "pure" field effects can be investigated. In the following, Hirshfeld charges are always given in a.u. after using the SCC-algorithm with the shape of the Gaussians used for their calculation, chosen according to the density envelope.

One C10-SAM Molecule per Unit-Cell

The simulation of field desorption for the C10-SAM was conducted using the initially relaxed surface interfaces as presented in Section 8.5.2. The external field was incrementally increased in steps of 0.01 a.u. with a geometry optimization of the SAM molecule after every step, if not stated otherwise. As discussed in Section 8.4, these smaller 2×2 models are not adequate representations of the experimental situation due to lower density and poor intermolecular rearrangements. However, they act as nice reference systems to show the observed behaviors during the calculations.

The initial geometry after optimization of the SAM molecule is shown in Figure 9.9a. Upon application of the external field, the molecular tilting is only slightly impacted as shown in Figure 9.9b before the evaporation of the first H^+ -ion. Afterwards, the remaining molecule aligns its segments with the field direction. This can be seen from the slight reduction of molecular tilt from initially $\phi_{2 \times 2}^{C10}(0.00 \text{ a.u.}) = 26.49^\circ$ to $\phi_{2 \times 2}^{C10}(0.03 \text{ a.u.}) = 23.01^\circ$ and finally to $\phi_{2 \times 2}^{C10}(0.04 \text{ a.u.}) = 4.9^\circ$ during the evaporation of the H^+ ion as illustrated in Figure 9.9c. The structural data of the C10-SAM up to first fragmentation are summarized in Table 9.3. Furthermore, this reorientation is accompanied by a charge redistribution leading to the system's polarization. Negative charge is step wise transferred from the SAM molecule into the substrate as can be seen from calculating the net charges for both components respectively. Charge neutrality is conserved throughout the whole simulation process.

Initially, the molecular headgroup carries the net charge of $q_{CH_3} = -0.016$ and can therefore be considered neutral. After increase of the field, the head groups fragment gets more and more positively charged until it fragments with a net Hirshfeld charge of $q_{CH_3} \approx +0.5$, indicating the

ionization of the fragment. During increase of the field, the observed charge transfer leads to the metal bulk carrying a surplus of partially compensated by increased charge in the sulfur atom. Initially, a charge of $q_{\text{Au}} = -0.942$ is observed on the metal substrate and $q_{\text{S}} = +0.156$ on the sulfur atom respectively. Upon fragmentation, the charge distribution is shifted towards a higher surplus in the metal substrate leading to $q_{\text{Au}} = -1.306$ as well as larger ionization of the SAM molecule accompanied by accumulated positive charge at the sulfur atom $q_{\text{S}} = +0.23$ increasing the ionic character of the sulfur-metal bond, while the bond distance remains almost unchanged at $d_{\text{S-Metal}}(0.0 \text{ a.u.}) \approx 2.0 \text{ \AA}$. As a result, the sulfur metal attachment increases its bond strength. The remaining structure after deletion of the H^+ fragment is then re-optimized in a field of $\mathcal{E} = 0.03 \text{ a.u.}$ showing molecular alignment in parallel to the field axis. Further increase of the field strength leads to evaporation of the ionized molecular backbone $\text{C}_{10}\text{H}_{20}^+$ during the critical field step $\mathcal{E}_{\text{crit}} = 0.035 - 0.04 \text{ a.u.}$, while the sulfur atom stays attached.

During the increase of the field, a charge separation within the molecule and the surface is observed indicating the system's polarization. The polarization of the fragments further indicates that during the evaporation process, an electron is successively pulled into the substrate, breaking the bond to the now positively charged ion as described by Dietrich et. al.¹⁹. An overview of the atomic partial charges contributing to the charge shifts is summarized in Table 9.1. Furthermore, the calculation of the system's bulk polarization P during incremental field steps, reveals an increase of the absolute value as shown for the data summarized in Table 9.2 resulting in a value of $P(\mathcal{E} = 0.03 \text{ a.u.}) = D$. Specifically the polarization's z -component is gradually increased with every field increment.

Table 9.1: Distribution of the Hirschfeld partial charges for the C10-SAM systems during incremental field increase.

| \mathcal{E} in a.u. | Au - (2×2) | S | SAM | CH_3 - Head group |
|-----------------------|-----------------------|-------|-------|----------------------------|
| 0.00 | -0.942 | 0.156 | 0.945 | -0.016 |
| 0.01 | -1.057 | 0.196 | 1.072 | 0.020 |
| 0.02 | -1.161 | 0.228 | 1.177 | 0.051 |
| 0.03 | -1.306 | 0.236 | 1.325 | 0.098 |
| 0.04 | -1.157 | 0.205 | 1.254 | 0.383 |

Table 9.2: Development of the total bulk Polarization P_{tot} and its components P_i following Eq (3.23), during increase of the external electric field for one C10-SAM molecule.

| \mathcal{E} in a.u. | P_x in D | P_y in D | P_z in D | P_{tot} in D |
|-----------------------|------------|------------|------------|-----------------------|
| 0.00 | -6.946 | 9.033 | 1.356 | 11.476 |
| 0.01 | -1.634 | -5.873 | 7.115 | 9.369 |
| 0.02 | -2.574 | -7.620 | 13.102 | 15.373 |
| 0.03 | -16.854 | -8.735 | 20.505 | 27.943 |

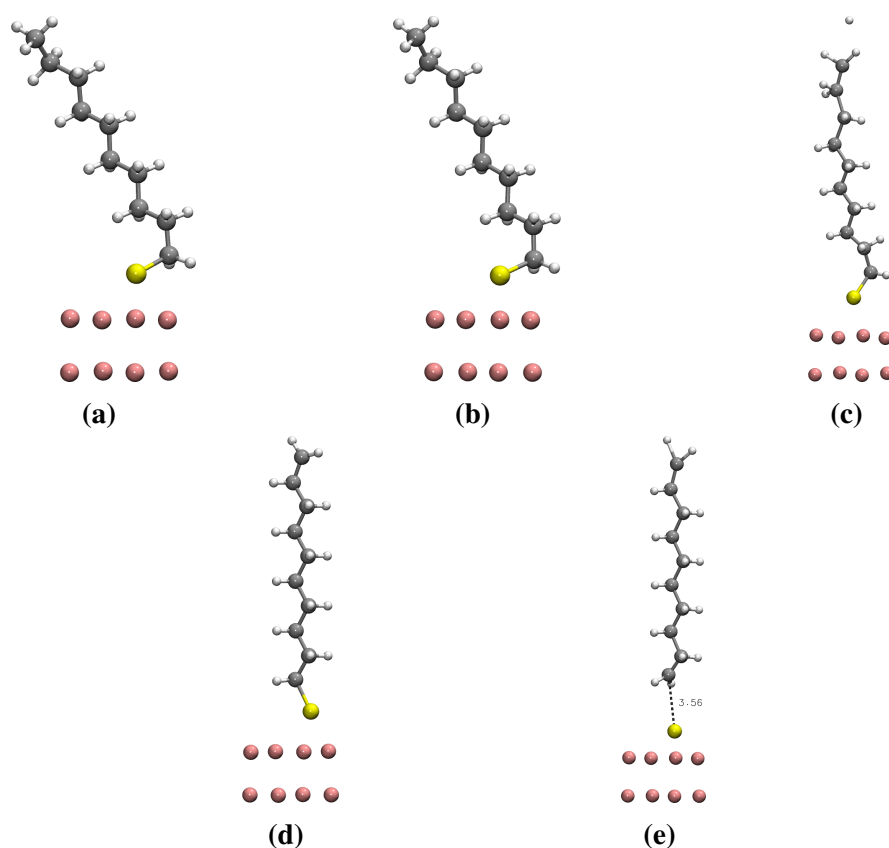
**Figure 9.9:** Stages of the desorption for C10 molecules on a 2×2 gold surfaces. a) The initial structure in the absence of an electric field. b) The structure of the C10 molecule in an external field of $\mathcal{E} = 0.03$ a.u., showing only slight reorientation. c) The structure during the desorption of a single H^+ ion from the SAM-chain during the field step $\Delta\mathcal{E} = 0.03 - 0.04$ a.u. showing alignment of the remaining chain with the field axis. d) The relaxed C10-molecule without the evaporated hydrogen in an external field of $\mathcal{E} = 0.03$ a.u., e) The situation during the critical field step $\mathcal{E} = 0.035 - 0.04$ a.u. showing the evaporation of the molecular backbone.

Table 9.3: Molecular tilting and length for the relaxed geometries for incrementally increased external field for the investigated 2×2 C10-SAM system.

| Basis Set | \mathcal{E} in a.u. | d in Å | ϕ in ° |
|--|-----------------------|--------|-------------|
| TZV2P-GTH | 0.00 | 11.57 | 26.49 |
| | 0.01 | 11.58 | 26.31 |
| | 0.02 | 11.59 | 26.16 |
| | 0.03 | 11.56 | 23.01 |
| - C ₁₀ H ₂₁ ⁺ | 0.04 | — | 4.9 |

Two C10-SAM Molecules per Unit-Cell

For two molecules on a 3×2 surface, the relaxed geometries discussed in Section 8.5.3 are used as initial structures for the simulation. The optimized structure is shown in Figure 9.10a. The external field was incrementally increased in steps of $\mathcal{E} = 0.01$ a.u. until fragmentation of ions is observed. During the initial increase of the field strength, the molecular rearrangement changes only slightly, which can be seen from the optimized structure at $\mathcal{E} = 0.03$ as illustrated in Figure 9.10b. Upon the field step $\Delta\mathcal{E} = 0.03 - 0.04$ a.u. a CH₃⁺ fragment is evaporated from chain II. During the evaporation process, the split of the CH₃⁺ fragment in a CH₂⁺ and H⁺ is observed before the threshold of 3 angstrom is reached, indicating a simultaneous evaporation event as can be seen in Figure 9.10c. The optimized remaining structure after deletion of CH₃⁺, shows a stronger molecular alignment with the field axis, which is indicated by the reduced average tilting of both molecules from $\bar{\phi}_{3 \times 2}^{\text{C10}}(0.03 \text{ a.u.}) = 17.82^\circ$ before and $\bar{\phi}_{3 \times 2}^{\text{C10}}(0.03 \text{ a.u.}) = 10.11^\circ$ after evaporation in a field of $\mathcal{E} = 0.03$ a.u.. During the following field step $\Delta\mathcal{E} = 0.03 - 0.035$ a.u. leads to the evaporation of a single H⁺ ion from chain I. After deletion of the evaporated fragment, the simulation is continued in the field of $\mathcal{E} = 0.04$ a.u., leading to the desorption of the remaining chain I as a C₁₀H₁₂S⁺ fragment. After re-optimization of the remaining chain II in a field of $\mathcal{E} = 0.03$ a.u., the second molecule is evaporated as a C₉H₁₈S⁺ fragment. The field steps during evaporation as well as the observed fragments are summarized in Table 9.4.

Table 9.4: Summary of the obtained fragmentation patterns for the investigated C10-SAM systems.

| molecules per unit cell | fragment | \mathcal{E}_{sim} in a.u. |
|-------------------------|-------------------------------------|------------------------------------|
| 1 | H^+ | 0.03-0.04 |
| | $\text{C}_{10}\text{H}_{23}^+$ | 0.03-0.035 |
| 2 | CH_3^+ (II) | 0.03-0.04 |
| | H^+ (I) | 0.03-0.04 |
| | $\text{SC}_9\text{H}_{18}^+$ (II) | 0.03-0.035 |
| | $\text{SC}_{10}\text{H}_{20}^+$ (I) | 0.03-0.035 |

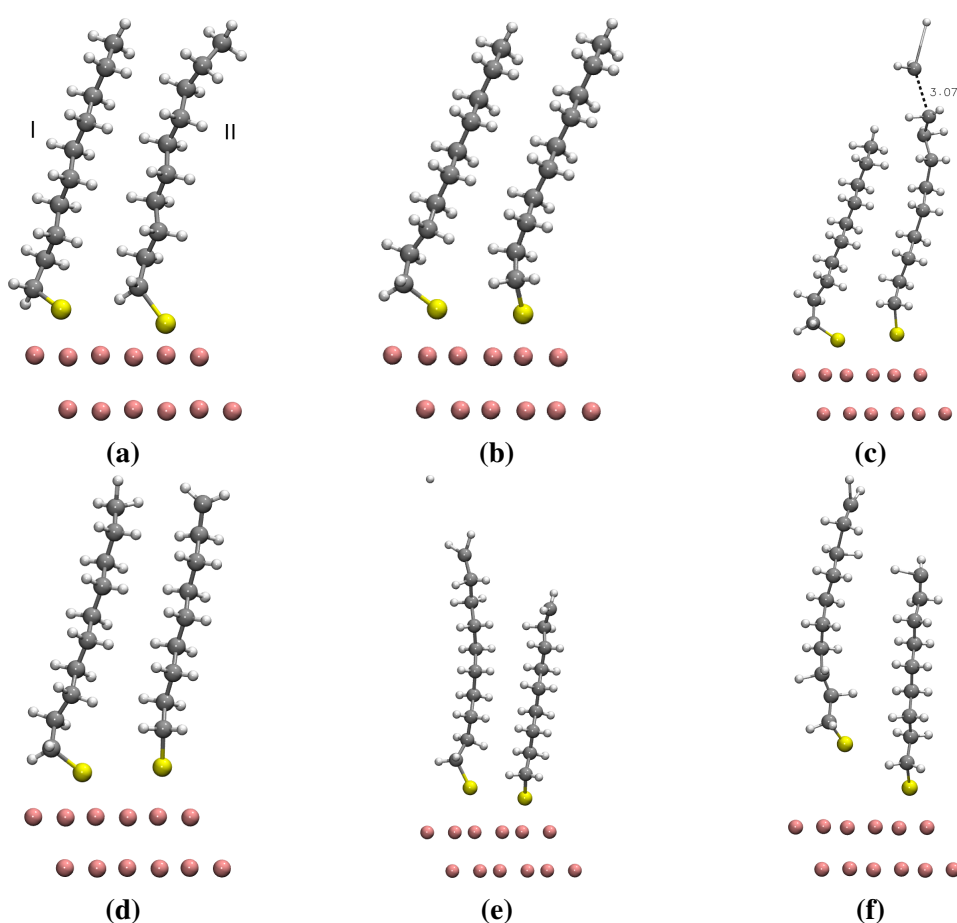


Figure 9.10: Simulated stages of the desorption process for two C10-SAM molecules on a two-layered 3×2 gold surface. a) The initial geometry after optimization at $\mathcal{E} = 0.00$ a.u., b) the slightly straightened molecules after relaxation in a field of $\mathcal{E} = 0.03$ a.u., c) Evaporation of a CH_3^+ -ion from chain II during the field step $\Delta\mathcal{E} = 0.03 - 0.04$ a.u., d) The re-optimized SAM geometry in a field of $\mathcal{E} = 0.03$ a.u. e) evaporation of a single H^+ ion from chain I during re-increasing the field to $\Delta\mathcal{E} = 0.03 - 0.04$ a.u., f) continuation of the evaporation process in a field of $\mathcal{E} = 0.04$ a.u. leading to desorption of chain I as $\text{C}_{10}\text{H}_{20}\text{S}^+$ ion.

9.3.4 Evaporation of Amino-undecanethiol from Gold

Following the previously introduced nomenclature, amino-undecanethiol is further referred to as NH₂-SAM (compare Figure 8.3). In analogy to the previously studied C10-SAM, investigations were performed for amino-undecanethiol. Systems with one and two molecules on a two-layered 2 × 2 and 3 × 2 gold (111) surface were studied respectively. The smaller system is used to highlight the general effects of the applied field onto the SAMs. The 3 × 2 system, however, corresponds nicely to the experimentally observed SAM-density, while including explicit intermolecular rearrangements. All calculations were performed utilizing TZV2P-GTH basis sets for all organic atoms along with the PBE functional and D3 dispersion correction. GTH-PBE optimized PPs were used for all organic compounds. The gold atoms were again treated with specialized DZVP-SR-MOLOPT-GTH basis sets containing more diffuse Gaussian functions and GTH-q19 PPs. The initially relaxed structures were placed in an electric field acting perpendicular to the surface in *z*-direction which was incrementally increased with performed geometry optimization after every field step. During all simulations, the lattice parameters as well as the gold atoms were fixed and not rearranged within the optimization process.

One NH₂-SAM Molecule

The simulations of one SAM molecule on the surface were conducted in analogy to Section 9.3.3. After initial application of a field with strength $\mathcal{E} = 0.01$ a.u., a slight straightening of the molecule due to alignment of its segments with the applied field direction can be observed. This can be seen in the reduced tilting angle, which was initially observed as $\phi(0 \text{ a.u.}) = 63.24^\circ$ without external field and reduced upon field application to $\phi(0.01 \text{ a.u.}) = 61.41^\circ$. The geometry without external field is shown in Figure 9.11a, whereas the straightened alignment upon field application is illustrated Figure 9.11b. Application of the field with strength 0.01 a.u. also leads to the first indications of a charge redistribution. Negative charge is shifted towards the substrate as can be seen from the increase of net charge by $\Delta q_{\text{Au}} \approx -0.09$.

Further increase lead to minor rearrangements of the chains segments until a total alignment with the applied field at a field strength of $\mathcal{E} = 0.04$ a.u. with an observed tilting angle of

9 Field Evaporation of Self Assembled Monolayers

$\phi(0.04 \text{ a.u.}) = 2.24^\circ$. The initially tilted head group was reoriented as well with the rest of the chain, indicating a change in the charge distribution. The Hirschfeld partial charges revealed a transfer from the molecule into the substrate by the increased net charge by $\Delta q_{\text{Au}} = -0.227$ and the accumulation of positive charge of $q_{\text{CH}_2\text{NH}_2} = +0.385$ in the head group, while the whole SAM compensates for the transfer to the gold substrate by $\Delta q_{\text{SAM}} = +0.18$. Upon further increase of the field strength above $\mathcal{E} = 0.04 \text{ a.u.}$ which was determined as the critical field step, the whole carbon chain of the NH_2 -SAM molecule is evaporated. However, the sulfur atom responsible for the bonding to the surface, stays attached and is placed in the middle of 3 gold atoms. Data of the molecular properties are summarized in Table 9.7 as well as the charge distribution summarized in Table 9.5. Calculation of the system's bulk polarization P for the optimized geometries for every field step, reveals an increase of the absolute value as shown for the data summarized in Table 9.6. As observed for the C10-SAM system, increasing the electric field results in a gradual increase of the polarization's z -component, however for the NH_2 -SAM it this behavior is more pronounced with a maximum value of $P(\mathcal{E} = 0.04 \text{ a.u.}) \approx 68.9 \text{ D}$.

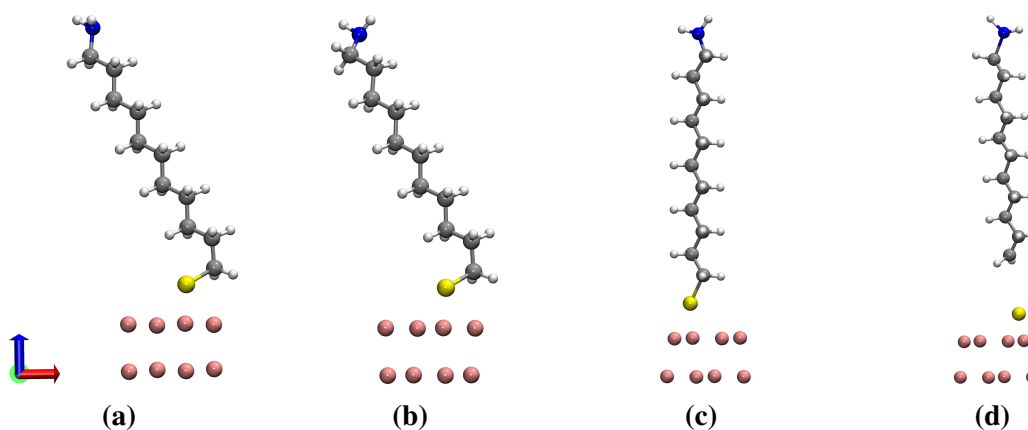


Figure 9.11: Simulated stages of the desorption process for one SAM molecule on a two-layered 2×2 gold surface. a) The initial geometry after optimization at $\mathcal{E} = 0 \text{ a.u.}$, b) the slightly straightened geometry after application of a field $\mathcal{E} = 0.01 \text{ a.u.}$, c) the completely straightened molecule at $\mathcal{E} = 0.04 \text{ a.u.}$, d) the situation during the field step $\Delta\mathcal{E} = 0.04 - 0.045 \text{ a.u.}$ resulting in desorption of the carbon chain without sulfur.

Table 9.5: Distribution of the Hirschfeld partial charges for the NH₂-SAM systems during incremental field increase.

| \mathcal{E} in a.u. | Au - (2×2) | S | SAM | NH ₂ CH ₂ - Head group |
|-----------------------|-----------------------|-------|-------|---|
| 0 | -0.945 | 0.155 | 1.06 | 0.012 |
| 0.01 | -1.038 | 0.193 | 1.06 | -0.009 |
| 0.015 | -1.095 | 0.2 | 1.098 | 0.026 |
| 0.02 | -1.180 | 0.191 | 1.182 | 0.103 |
| 0.03 | -1.218 | 0.175 | 1.235 | 0.352 |
| 0.04 | -0.944 | 0.146 | 0.962 | 0.385 |

Table 9.6: Development of the total bulk Polarization P_{tot} and its components P_i following Eq (3.23), during increase of the external electric field for one NH₂-SAM molecule.

| \mathcal{E} in a.u. | P_x in D | P_y in D | P_z in D | P_{tot} in D |
|-----------------------|------------|------------|------------|-----------------------|
| 0.00 | -1.458 | -3.655 | 1.179 | 4.109 |
| 0.01 | -15.857 | -6.064 | 7.982 | 18.760 |
| 0.015 | -2.452 | -6.927 | 11.810 | 13.909 |
| 0.02 | -10.343 | 4.494 | 16.782 | 20.219 |
| 0.03 | -0.636 | -9.816 | 46.796 | 47.819 |
| 0.04 | -11.198 | -4.000 | 67.878 | 68.912 |

Table 9.7: Molecular tilting and length for the relaxed geometries for incrementally increased external field for the investigated 2×2 NH₂-SAM system.

| Basis Set | \mathcal{E} in a.u. | d in Å | ϕ in ° |
|-----------------|-----------------------|--------|-------------|
| TZV2P-GTH | 0.00 | 14.47 | 26.76 |
| | 0.01 | 14.32 | 28.60 |
| | 0.015 | 14.32 | 28.56 |
| | 0.02 | 14.40 | 26.68 |
| | 0.03 | 14.90 | 13.40 |
| | 0.04 | 16.47 | 2.24 |
| DZVP-MOLOPT-GTH | 0.04 | 16.56 | 2.86 |
| | 0.0425 | 16.81 | 2.72 |

Two NH₂-SAM Molecules

The relaxed surface arrangement for two SAM molecules on a two-layered 3×2 gold surface, gives rise to a hydrogen bond as observed in Section 8.7.1, resulting in the formation of molecule pairs. The extended unit-cell shown in Figure 9.12a with its periodic images in transparent colors showing the hydrogen bonded head groups. Upon application of $\mathcal{E} = 0.01$ a.u., this additional feature leads to the stabilization of the molecule pair and rather than alignment in field direction as for the previous species, a collective reorientation could be observed as shown in Figure 9.12b. This is quantified by the slight reduction of the tilting angle for molecules I and II from initially $\phi_{\text{I}}(0.00 \text{ a.u.}) = 60.27^\circ$ and $\phi_{\text{II}}(0.00 \text{ a.u.}) = 61.48^\circ$ to $\phi_{\text{I}}(0.02 \text{ a.u.}) = 73.64^\circ$ and $\phi_{\text{II}}(0.02 \text{ a.u.}) = 67.86^\circ$ respectively. The further increase of the field from $\mathcal{E} = 0.02$ a.u. to $\mathcal{E} = 0.03$ a.u. leads to the breaking of the head groups hydrogen bond as illustrated in Figure 9.12d. Due to the absence of the head groups interaction, both molecules aligned with the field and the tilting angles are further reduced to $\phi_{\text{I}}(0.03 \text{ a.u.}) \approx 7.20^\circ$ and $\phi_{\text{II}}(0.03 \text{ a.u.}) \approx 14.90^\circ$. After further increasing the field to $\mathcal{E} = 0.04$ a.u. , the evaporation of the higher positioned molecule I was observed as shown in Figure 9.12e. The second molecule is subsequently evaporated after restarting the calculation without chain I. Increasing the basis set for the gold atoms from DZVP-MOLOPT-SR-GTH to TZV2P-MOLOPT-SR-GTH lead to no change in the desorption behavior. The same behavior was also observed when utilizing TZV2P-MOLOPT-GTH basis sets for the organic atoms as well. The fragmentation data are summarized in Table 9.8.

Table 9.8: Summary of the obtained fragmentation patterns for the investigated NH₂-SAM systems.

| molecules per unit cell | fragment | \mathcal{E}_{sim} in a.u. |
|-------------------------|---|------------------------------------|
| 1 | C ₁₁ NH ₂₄ ⁺ | 0.04 - 0.045 |
| 2 | C ₁₁ NH ₂₄ ⁺ (I) | 0.04 - 0.045 |

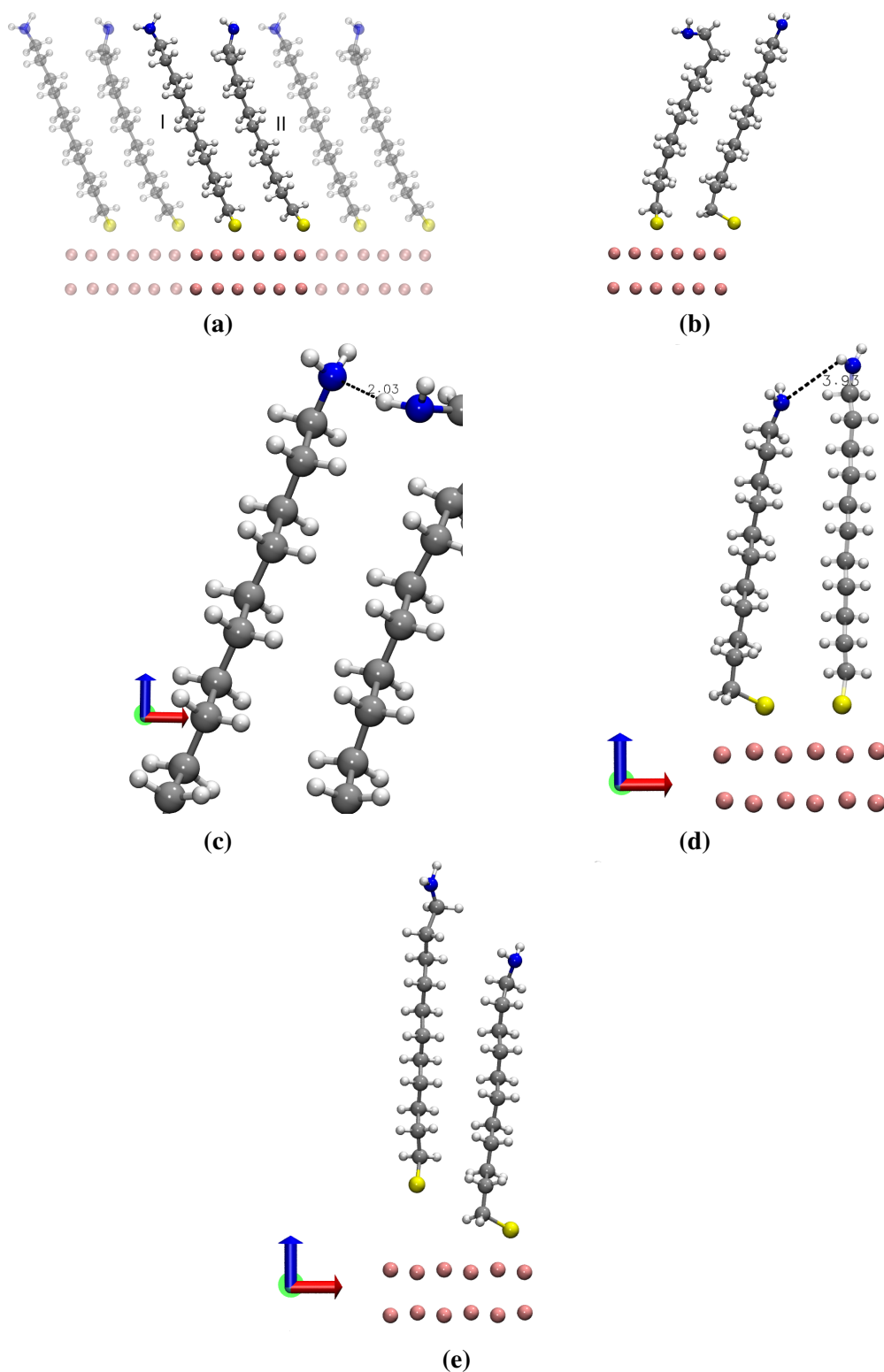


Figure 9.12: Simulated stages of the desorption process for two SAM molecule on a two-layered 3×2 gold surface. a) The extended initial geometry after optimization at $\mathcal{E} = 0.00$ a.u. with the periodic images in x -direction in transparent colors, b) The unit cell after optimization in a field of $\mathcal{E} = 0.01$ a.u., c) The head groups of chain I and II, remaining hydrogen bonded after large reorganizations during increasing the field to $\mathcal{E} = 0.01$ a.u., d) the straightened molecules after further increasing the field to $\mathcal{E} = 0.03$ a.u., e) the situation after increasing the field to $\mathcal{E} = 0.04$ a.u. showing the desorption of molecule I from the surface.

9.3.5 Evaporation of Perfluoro-Decanethiol from Gold

Following the previously introduced nomenclature, perfluoro-decanethiol is further referred to as F-SAM (compare Figure 8.3). In experimental studies related to atom probe tomography of F-SAM molecules¹⁵, a surface area of $A = 21.5 \text{ \AA}^2$ per molecule was estimated for each F-SAM molecule. From the previously investigated surface models, a suitable set of applicable surface coverage can be obtained. The models of choice include one SAM molecule on a 2×2 as well as two molecules on a 3×2 gold surface. The prior one acts as a possible reference system due to its simplicity whereas the latter one nicely corresponds to the experimental density, while including explicit intermolecular interactions Section 8.6.1.

One F-SAM Molecule per Unit-Cell

Unit cells constructed with one F-SAM molecule on a two-layered 2×2 gold surface consist of in-plane cell vectors of $a = b = 5.76 \text{ \AA}$, resulting in a surface area of $A = 28.7 \text{ \AA}^2$ per molecule. The lower SAM coverage leads to a tilting of the SAM molecules to optimize the intermolecular interactions as investigated in the previously conducted surface study. In the absence of an electric field, the molecular tilting angle is given as $\phi_{2 \times 2}^F(0.0 \text{ a.u.}) = 22.0^\circ$. After application of the electric field, the molecule straightens from $d_{2 \times 2}^F(0.0 \text{ a.u.}) \approx 12.26 \text{ \AA}$ to $d_{2 \times 2}^F(0.02 \text{ a.u.}) \approx 12.8 \text{ \AA}$ and the tilting angle is increase to $\phi_{2 \times 2}^F(0.02 \text{ a.u.}) = 34.75^\circ$ and almost vanishes when reaching the evaporation field strength of the first fragment. The first fragment, CF_3^+ , evaporates while increasing the field strength from 0.04 to 0.045 a.u.. This behavior was observed for both utilized basis sets (DZVP-GTH and TZV2P-GTH). The molecule is re-optimized after deleting the CF_3^+ ion from the system, the remaining chain aligned almost perpendicular to the surface with the remaining tilting angle of $\phi(0.03 \text{ a.u.}) \approx 5^\circ$, after relaxation in a field of 0.045 a.u.. This alignment maximizes the distance between the new head group and its images, which carry a Hirschfeld partial charge of about +0.5 each. Within the used model, all CF_3 head groups of the whole modeled layer desorbed synchronously due to the periodic boundary conditions. The behavior of the F-SAM fragment at higher fields depends on the choice of the basis sets for the organic atoms. With the smaller DZVP-GTH basis set, the molecule remained stable, without additional fragments evaporating up to a field strength of $\mathcal{E} = 0.065 \text{ a.u.}$, which is higher than the determined evaporation field of gold

$\mathcal{E}_{\text{tresh}} = 0.06$ a.u.. Higher field strengths are inconsistent within the model framework and no further field steps are investigated. However, for the larger TZV2P-GTH basis set, a second fragment, CF_2^+ , desorbed during the field step $\mathcal{E} = 0.055$ - 0.06 a.u.. The remaining molecule was then again relaxed and remained stable until the field strength of $\mathcal{E} = 0.065$ a.u. was surpassed. From these initial calculations it follows that the TZV2P-GTH basis set leads to better description and it is therefore chosen for all follow up simulations. Calculation of the systems polarization, reveals a much less systematic increase as in the previous cases as can be seen from the data presented in Table 9.10. From its components, a continuous increase in z -direction is observed with similar magnitude as in the case of the C10-SAM system. However, the shifts in the off field direction components are

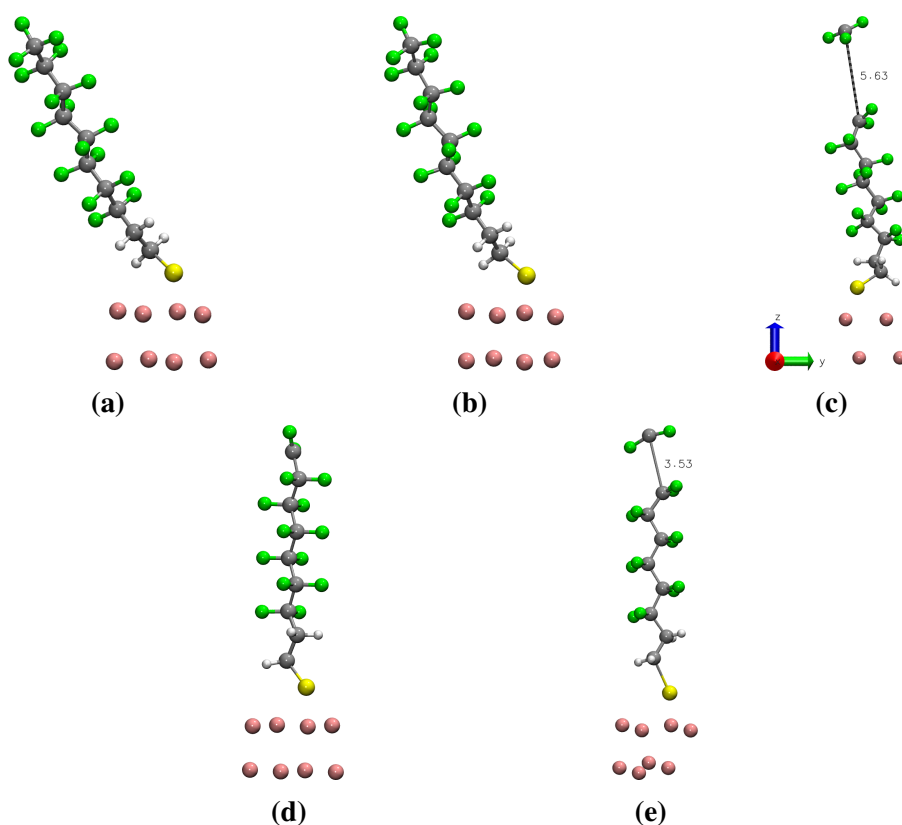


Figure 9.13: Simulated stages of the desorption process for one SAM molecules on a two-layered 2×2 gold surface utilizing the TZV2P-GTH basis. a) The molecule after initial geometry optimization. b) The slightly straightened molecule after application of the electric field $\mathcal{E} = 0.02$ a.u.. c) The situation during the desorption of the CF_3^+ -head group in a field of $\mathcal{E} = 0.045$ a.u., d) Reorganization after geometry optimization without the head group in an electric field of $\mathcal{E} = 0.045$ a.u. e) Desorption process of the CF_2^+ ion at the field strength of $\mathcal{E} = 0.065$ a.u..

Table 9.9: Distribution of the Hirschfeld partial charges for the F-SAM systems during incremental field increase.

| \mathcal{E} in a.u. | Au - (2×2) | S | SAM | CF ₃ - Head group |
|-----------------------|-----------------------|-------|-------|------------------------------|
| 0 | -0.846 | 0.213 | 0.868 | 0.008 |
| 0.02 | -1.189 | 0.354 | 1.194 | 0.073 |
| 0.03 | -1.424 | 0.318 | 1.423 | 0.092 |
| 0.035 | -1.476 | 0.295 | 1.475 | 0.133 |
| 0.04 | -1.515 | 0.261 | 1.510 | 0.199 |
| 0.045 | -1.470 | 0.258 | 1.473 | 0.439 |

Table 9.10: Development of the total bulk Polarization P_{tot} and its components P_i following Eq (3.23), during increase of the external electric field for one NH₂-SAM molecule.

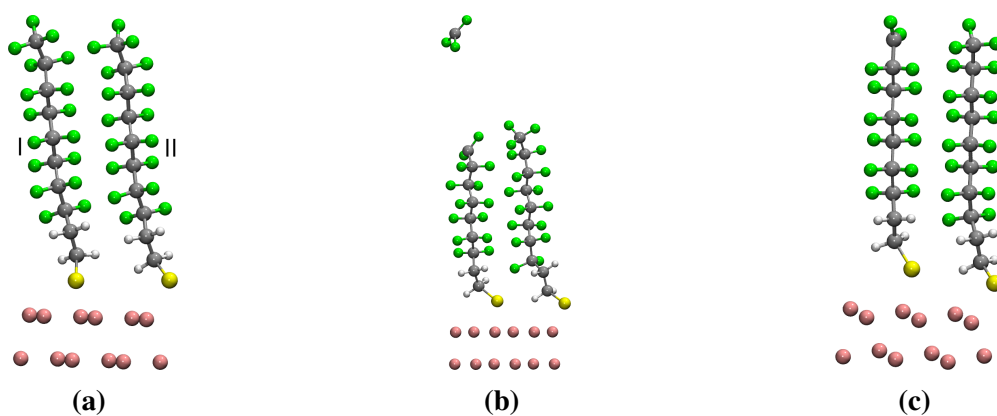
| \mathcal{E} in a.u. | P_x in D | P_y in D | P_z in D | P_{tot} in D |
|-----------------------|------------|------------|------------|-----------------------|
| 0.00 | -18.580 | -9.013 | -0.473 | 20.656 |
| 0.02 | 0.961 | 11.258 | 14.594 | 18.457 |
| 0.03 | 9.907 | -1.138 | 18.697 | 21.190 |
| 0.035 | -8.294 | -1.400 | 24.258 | 25.675 |
| 0.04 | -5.093 | 9.541 | 31.477 | 33.284 |

Two F-SAM Molecules per Unit-Cell

A better description of possible surface corrugation and intermolecular interactions is obtained by utilizing two-layered 3×2 super cells of gold with two F-SAM molecules. The modeled SAM density of $\rho_{\text{sim}}^{-1} \approx 21.56 \text{ \AA}^2$ perfectly matches the experimental target value of $\rho_{\text{exp}}^{-1} \approx 21.5 \text{ \AA}^2$ ¹⁵. Both molecules are further referred to as I and II respectively as shown in Figure 9.14a. Both molecules align their segments in an alternately staggered. Investigation of the resulting organization reveals atoms of molecule I to be positioned slightly further away from the surface, than those of II. In the absence of an external field, both molecules are slightly tilted by $\phi_{\text{I}}^{\text{F}}(0 \text{ a.u.}) = 10.3^\circ$ and $\phi_{\text{II}}^{\text{F}}(0 \text{ a.u.}) = 13.7^\circ$, which is founded on the fluorines' strong repulsion as discussed in Section 8.8. During the increase of the field, the SAM-geometry changes only slightly as the molecules are already almost in perpendicular alignment to the surface. While increasing the applied field from 0.04 a.u., the first CF₃⁺ fragment was observed to evaporate from molecule I as shown in Figure 9.14b. Subsequently, due to staggered

Table 9.11: Molecular tilting and length for the relaxed geometries for incrementally increased external field for the investigated 2×2 F-SAM system.

| Basis Set | \mathcal{E} in a.u. | d in Å | ϕ in ° | |
|--------------------------------|--------------------------------|--------|-------------|-------|
| DZVP-GTH | 0.00 | 13.29 | 9.61 | |
| | 0.02 | 13.19 | 10.30 | |
| | 0.03 | 13.17 | 10.27 | |
| | 0.04 | 13.17 | 10.27 | |
| - CF ₃ ⁺ | 0.045 | 12.44 | 11.01 | |
| TZV2P-GTH | 0.00 | 12.26 | 21.96 | |
| | 0.02 | 12.79 | 34.75 | |
| | 0.03 | 12.24 | 32.07 | |
| | 0.035 | 12.28 | 31.14 | |
| | 0.04 | 12.32 | 28.7 | |
| | - CF ₃ ⁺ | 0.045 | 11.32 | 25.45 |
| | | 0.05 | 12.65 | 4.99 |
| | - CF ₂ ⁺ | 0.055 | 12.98 | 3.25 |
| | | 0.06 | 11.46 | 3.79 |
| | - CF ₂ ⁺ | 0.07 | 10.91 | 7.15 |

**Figure 9.14:** The initial stages of the simulated desorption process for two F-SAM molecules on a two-layered 3×2 gold surface. a) The initial geometry at $\mathcal{E} = 0.00$ a.u., b) The evaporation of chain I's CF₃⁺ head group during the step $\mathcal{E} = 0.03 - 0.04$ a.u. . c) Relaxation of chain I without head group at $\mathcal{E} = 0.03$ a.u. .

9 Field Evaporation of Self Assembled Monolayers

segments, an alternating evaporation of fragments, from the molecules' end groups, is favored. At 0.04 a.u., the two chains rearrange, as shown in Figure 9.14c. The simulation is continued with the relaxed geometries of the first evaporation series shown in Figure 9.14, leading to the second fragmentation series illustrated in Figure 9.15. An increase of the field to 0.045 a.u. leads to simultaneous evaporation of two fragments: the CF_3^+ head group of chain II, followed by CF_2^+ from chain I shown in Figure 9.15a and Figure 9.15a. Subsequently, the field was decreased to 0.04 a.u. to allow relaxation of the molecules as shown in Figure 9.15c. Otherwise, the whole remaining SAM was found to desorb. Increasing the field to 0.045 a.u. yields desorption of CF_2^+ from chain II as depicted in Figure 9.15d. Subsequently, even after a relaxation at 0.03 a.u as illustrated in Figure 9.15e, increasing the field leads to desorption of the whole remaining layer at 0.045 a.u. shown in Figure 9.15f. The data are also summarized in Table 9.12.

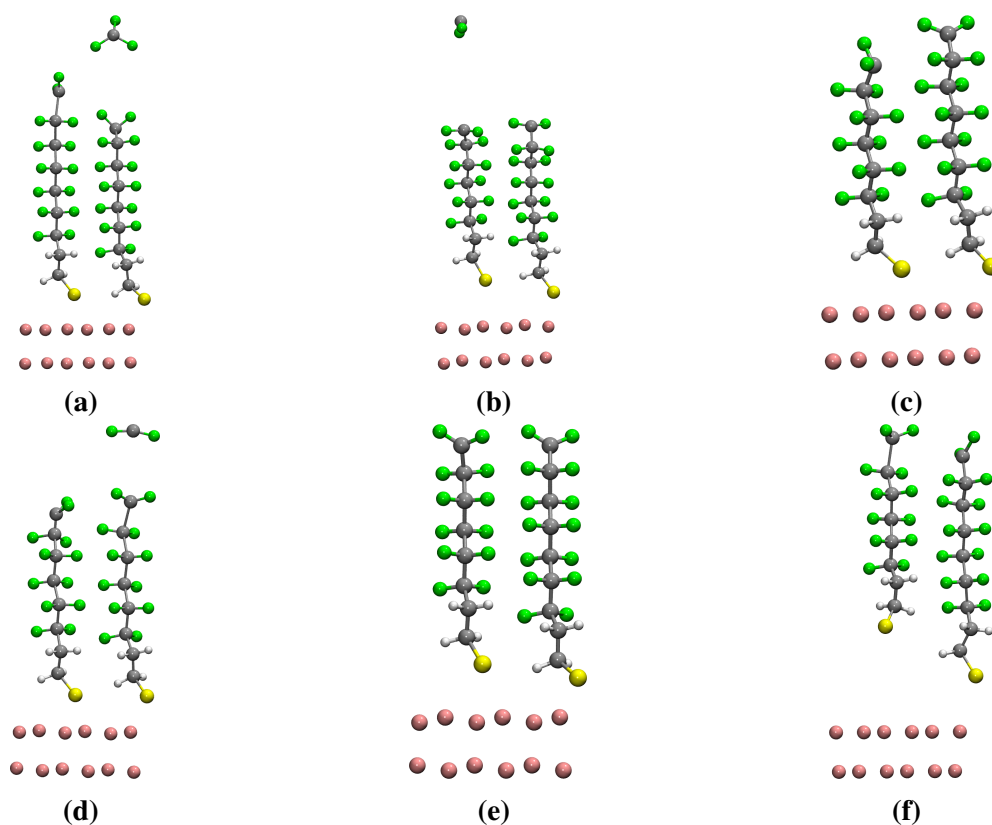


Figure 9.15: Stages of the continued simulation process. a-b) Simultaneous evaporation of CF_3^+ from chain II followed by CF_2^+ from chain I during $\mathcal{E} = 0.04 - 0.045$ a.u. . c) Relaxed structure after removing the fragments at $\mathcal{E} = 0.03$ a.u. . d) The evaporation of CF_2^+ from chain II during the field step $\mathcal{E} = 0.04 - 0.045$ a.u. . e) Relaxed structure in a field of $\mathcal{E} = 0.045$ a.u., f) Full desorption of the remaining molecules during the field step $\mathcal{E} = 0.045 - 0.05$ a.u. .

Table 9.12: Summary of the obtained fragmentation patterns for the investigated F-SAM systems.

| molecules per unit cell | fragment | \mathcal{E}_{sim} in a.u. |
|-------------------------|--|------------------------------------|
| 1 | CF_3^+ | 0.04 - 0.045 |
| | CF_2^+ | 0.05 - 0.055 |
| | CF_2^+ | 0.065 - 0.07 |
| 2 | CF_3^+ (I) | 0.03 - 0.04 |
| | CF_3^+ (II) | 0.04 - 0.045 |
| | CF_2^+ (I) | 0.04 - 0.045 |
| | CF_2^+ (II) | 0.04 - 0.045 |
| | $\text{SC}_8\text{F}_{12}\text{H}_4^+$ | 0.045 - 0.05 |

9.3.6 Discussion of the Observed Effects using GPW-DFT

The simulation of evaporation for the pure gold surface, showed the possible applicability for metals in general. In analogy, simulations were also performed for platinum and copper surfaces constructed following Section 8.3.1. The obtained evaporation field strengths show the magnitude of maximally applicable fields within the used framework. All observed evaporation features for systems of two molecules per unit cell, appeared for fields below the estimated threshold. For the idealized 2×2 reference systems however, the fragmentation patterns were in some cases not completed before reaching $\mathcal{E}_{\text{thresh}} = 0.06$ a.u. e.g. for the F-SAM.

In summary, the simulation of field desorption of the investigated SAMs from a gold surface was successfully conducted, utilizing GPW-DFT with external fields, for all three investigated species. The comparison with the simulations using a cluster model approach shows direct consequences of the “dense” surface packing as a result of the continuous solid state description. During the simulations, besides the fragmentation patterns, general effects of the field applications were observed. These include charge separation, molecular straightening and alignment in field direction. These could be shown to occur in the 2×2 example systems for all three SAM species. The observed effects as well as possible interpretations are summarized in the following.

Charge Separation and Polarization

For all three investigated species, an increase in the systems polarization depending on the applied electric field could be observed. Comparison of the bulk polarization for the optimized geometries at a field of $\mathcal{E} = 0.04$ a.u. with the unmodified decanethiol, showed crucial differences for the two modifications. Overall, the polarization in z direction was increased with the applied field for all systems, showing a direct dependence as already described for H_2 . For the F-SAM the total polarization was barely increased, due to its high initial values in the components perpendicular to the field axis. For the NH_2 -SAM, the obtained total polarization $P(\mathcal{E} = 0.04 \text{ a.u.}) \approx 69.0 \text{ D}$ is largest within the conducted study. The polarization is drastically increased upon the field step $\Delta\mathcal{E} = 0.02 - 0.03$ a.u. which is accompanied by a molecular straightening and the re-orientation of the amino head-group in field direction. A possible

reason for the larger increase of P is found in the higher molecular length as well as the amino head-groups affinity towards charge separation. Furthermore, the high field induced polarization in the case of NH_2 -SAM, is proportional to the length of the molecule. During the reorientation, the amino headgroups net partial charge increases drastically, showing the direct transfer of electrons from the headgroup into the substrate. In contrast to the other observed species, the F-SAM system is already polarized with a total value of $P_{\text{tot}} = 20.66$ D due to the high initial tilt in the system and the fluorine atoms carrying higher electron density. The reorientation in field direction is limited due to the larger spacings required for the fluorinated backbone and shows only a slight increase of P_z for low field strengths when compared to the other two molecules. In summary, an extraction of electrons from the SAM to the substrate and following redistribution is indicated by the change in the total net Hirschfeld partial charges for substrate and SAM respectively. This effect is observed to continue until the ionization of the molecular fragment that is next to evaporate, is large enough to suppress the bond-barrier. The effect was also observed during vacuum calculations of NH_2 -SAM and F-SAM¹⁹, showing consistency between the models.

Molecular Straightening and Realignment

Due to the charge reorganization, the investigated systems are polarized, which can be seen from the increasing positive charge in the molecular headgroups. The resulting reorientation along the surface normal can be seen as direct consequence of the molecule acting as an electric dipole. This effect is observed for all three investigated SAM species in the 2×2 case and accompanied by a straightening of the molecules. In the case of one C10 molecule, the initial molecule is highly tilted and only aligns itself with the field direction after losing its first fragment. A similar effect is observed in the case of two C10 molecules on a 3×2 surface, where notable reorientation of the molecules is only observed after the first evaporation event. For the F-SAM, the initial molecule in the 2×2 case is highly tilted as well and a re-orientation in field direction is observed only after the first fragmentation. The reorganization as well as re-orientation for two F-SAM molecules, is not as pronounced as the initial molecules are not as tilted as in the C10 case. For the NH_2 -SAM, the complete alignment in field direction is observed before the first evaporation which is also influencing the observed polarization

9 Field Evaporation of Self Assembled Monolayers

as mentioned earlier. As this behavior is not observed for the other two molecules, it can be directly associated with the amino headgroup. In the case of two NH₂-SAM molecules, the re-orientation of the molecules is accompanied by the breaking of the hydrogen bond, which increases the effective layer thickness by molecular straightening.

Evaporation Sequences

For the simple C10-SAM as well as NH₂-SAM, the fragmentation of molecules was not observed. However, the complete molecules evaporated during the increase of the applied field. Possible reasons are found when comparing the results of various descriptions. In the case of the NH₂-SAM, the breaking of the molecules hydrogen bond in the case of two molecules could be observed at a field of $\mathcal{E} = 0.03$ a.u. followed by the whole molecules at $\mathcal{E} = 0.045$ a.u.. The same behavior was also observed in the C10 case, however, without the intermediate stage of the absent hydrogen bond. The whole molecules are evaporated at field strengths of $\mathcal{E} = 0.04$ a.u.. For the simulated F-SAM, fragments of CF₃⁺ and CF₂⁺ were obtained for both systems, followed by the remaining C₈F₁₆S⁺ fragment. Interestingly, only for SAM-F, a multi fragment desorption pattern could be observed. A possible explanation lies in the difference of interaction due to the much more repulsive fluorine atoms not allowing for the evaporation of larger fragments until significant rearrangement from the staggered configuration was achieved.

9.4 Simulations of Field Evaporation using GFN-xTB

The DFT description of field evaporation as previously presented in Section 9.3, offers a high amount of possible improvements given within the method itself by adjusting basis sets, possible functional implementations and charge treatment, however at the price of extremely high computational costs. A promising alternative is provided by periodic implementations of semi-empirical tight binding methods. The following analysis uses the CCM implementation of GFN-xTB⁶¹ within the CP2K program package⁶², following the described workflow presented in Section 4.3. This section presents the results of the performed simulations within the GFN-xTB framework for the previously treated SAM-substrate interfaces. The desorption process is modeled in analogy to the already described GPW-DFT simulations. The utilized gold surfaces were created using the bulk optimized lattice constants as described in Section 8.3.2. The initial structures of the SAM-Metal interfaces were then constructed as presented in Section 8.4. Calculations utilizing GFN-xTB require significantly less computational effort and therefore allow, in principle, for the treatment of super-cells. However, for comparability, the same systems as in the GPW-DFT case are investigated namely one molecule on a two-layered 2×2 and two molecules on a 3×2 (111) surface.

9.4.1 Evaporation of Gold

In an initial investigation, the threshold for evaporation from a reconstructed three-layered 3×2 gold surface was determined. The gold surface was constructed as explained in Section 8.3.1 utilizing the bulk optimized in plane lattice vector of $a_{\text{Au},xTB} = 2.91 \text{ \AA}$. The initial geometry was then created by geometry optimization of the uppermost layer of gold keeping the bottom two layers fixed. The reorganized structure is shown in Figure 9.16a.

Subsequently, the external field was incrementally increased in steps of 0.01 a.u. with geometry optimization, of the uppermost layer, performed at every step. At the field step of $\mathcal{E} = 0.05 - 0.06$ a.u., two gold atoms desorbed simultaneously from the surface. The corresponding situation is shown in Figure 9.16b. In analogy to the calculations performed for the GPW-DFT simulations, the obtained threshold for desorption acts as an estimate for the highest applicable field strength for follow up simulations. However, the field acting on

the gold surface below the SAM layers as investigated in the following simulations, might be lowered due to the dielectric shielding provided by the SAM coverage.

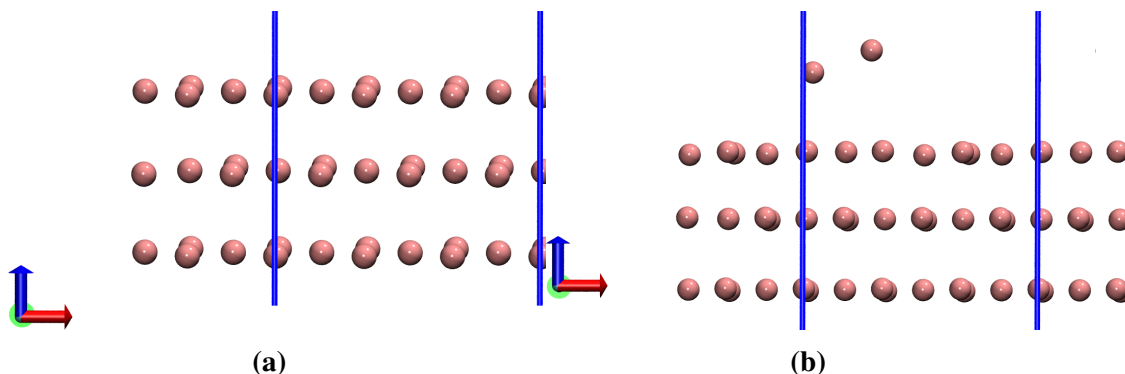


Figure 9.16: Exemplary situations for the evaporation of gold in a field applied perpendicular to the surface. a) Shows the minor reorganization of the top layer in a field of strength $\mathcal{E} = 0.01$ a.u., b) The situation during the field step $\mathcal{E} = 0.055 - 0.06$ a.u. where two gold atoms evaporate simultaneously.

9.4.2 Evaporation of Decanethiol from Gold

As an initial example, the evaporation of unmodified decanethiole was investigated. The external field was incrementally increase in steps of 0.01 a.u. with geometry optimization of the SAM molecule after every step, if not stated otherwise. For the required Ewald summation necessary for a treatment with the CCM implementation of GFN-xTB a mesh of 80 points in every spacial direction is used. Simulations were performed using the previously introduced described in plane lattice vectors of $a_{\text{Au},xTB} = 2.91 \text{ \AA}$. The SAM-metal interfaces were created using the methodology described in Section 8.4. The modeled SAM densities are slightly lower than in the DFT case due to the reduced surface coverage and are given as $\rho_{\text{SAM},2 \times 2}^{-1} \approx 29.3 \text{ \AA}^2$ and $\rho_{\text{SAM},3 \times 2}^{-1} \approx 22.0 \text{ \AA}^2$. After each field step, geometries are optimized while keeping the metal atoms fixed. In analogy to the DFT case, a 2×2 system is used as initial reference for each species respectively. The importance of explicit intermolecular interactions is investigated using the two-layered 3×2 gold surfaces with two C10-SAM molecules. The cell parameter in z direction was chosen ad-hoc to be 60 \AA in order to account for the electrostatic potential drop, as described in Section 3.6. The obtained fragmentation patterns are summarized in Table E.1.

One C10-SAM Molecule per Unit-Cell

The stages of the fragmentation process, simulated for the smallest system using the C10-SAM molecule, are presented in Figure 9.18. The initially highly tilted molecule shown in Figure 9.17 is straightened after field application as illustrated for the case of $\mathcal{E} = 0.04$ a.u. in Figure 9.18a. This effect continues until full alignment with the field axis is achieved at $\mathcal{E} = 0.06$ a.u. as shown in Figure 9.18b. Further increase of the field leads to evaporation of the whole C10-SAM molecule including the sulfur atom, during the step $\mathcal{E} = 0.07 - 0.08$ a.u. as shown in Figure 9.18c. In a second simulation, the sulfur atom was held fixed during the whole desorption process. During the field step from $\mathcal{E} = 0.09 - 0.10$ a.u. a single H^+ ion is evaporated from the molecule. The situation is shown in Figure 9.18d, where the remaining molecule is completely aligned with the field axis. After deletion of the H^+ fragment from the system, the molecule is reoptimized in a field of $\mathcal{E} = 0.09$ a.u. with the optimized geometry shown in Figure 9.18e. Further increase of the field, leads to evaporation of the full $\text{C}_{10}\text{H}_{20}^+$ fragment during the field step $\mathcal{E} = 0.09 - 0.10$ a.u. as shown in Figure 9.18d.

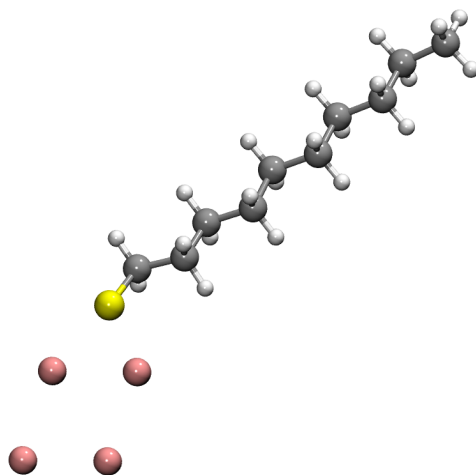


Figure 9.17: The initial geometry of the simulated 2×2 C10-SAM system at $\mathcal{E} = 0$.

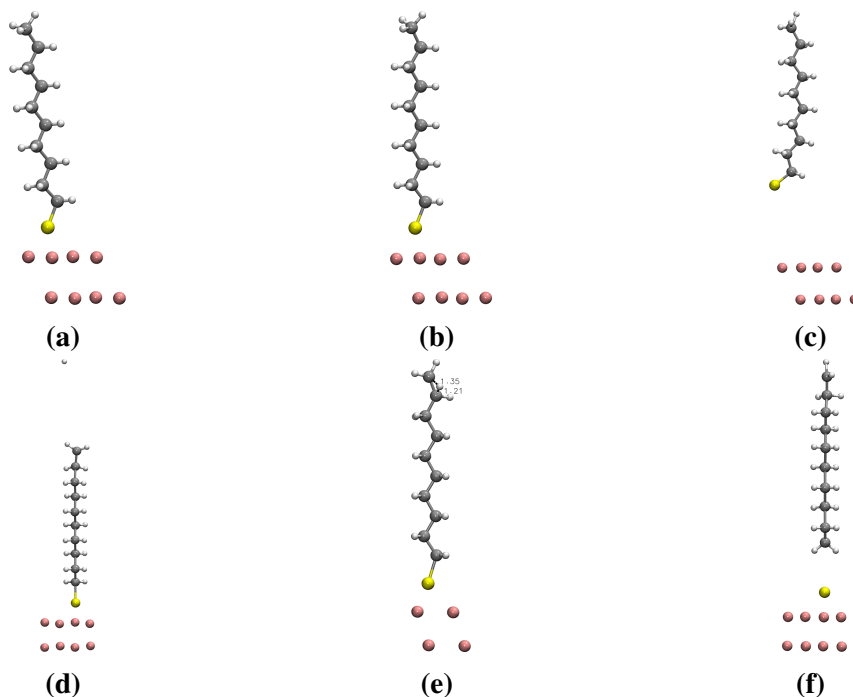


Figure 9.18: Overview of the simulated desorption for one C10-SAM molecule on a two-layered 2×2 gold surface. a) The molecule almost aligned with the field axis at $\mathcal{E} = 0.04$ a.u.. b) The molecule oriented in parallel to the field axis at $\mathcal{E} = 0.06$ a.u. c) Desorption of the whole C10-molecule during the field step $\mathcal{E} = 0.07 - 0.08$ a.u. d) Evaporation of a single H^+ ion, during the field step $\mathcal{E} = 0.09 - 0.10$ a.u. for optimization with fixed sulfur atom. e) Reoptimized C10-molecule at $\mathcal{E} = 0.09$ a.u. showing the headgroups reorganization, for fixed sulfur atom. f) Desorption of the remaining $\text{C}_{10}\text{H}_{20}^+$ -fragment during the field step $\mathcal{E} = 0.09 - 0.10$ a.u. for fixed sulfur atom.

Two C10-SAM Molecules per Unit-Cell

For two molecules on a 3×2 surface, the relaxed geometries are constructed based on the GFN-xTB optimized initial structures in analogy to Section 8.4. The stages of the fragmentation process, simulated for two C10-SAM molecules on a 3×2 surface, are presented in Figure 9.19. The relaxed initial structures are shown in Figure 9.19a, showing reduced tilting when compared to the case of the smallest system.

Upon application of an external field of $\mathcal{E} = 0.03$ a.u. the molecules are completely straightened and aligned with the field axis. After optimization, one chain is positioned higher with respect to the surface plane, the situation after optimization is shown in Figure 9.19b. During the field step $\mathcal{E} = 0.035 - 0.04$ a.u., the higher positioned chain is ionized and

evaporated as shown in Figure 9.19c. The remaining molecule is then optimized in a field of $\mathcal{E} = 0.04$ a.u. with the obtained geometry shown in Figure 9.19d. Upon increase of the field from $\mathcal{E} = 0.05 - 0.06$ a.u. the remaining molecule is ionized and desorbed from the surface. The situation is shown in Figure 9.19e, completing the fragmentation sequence.

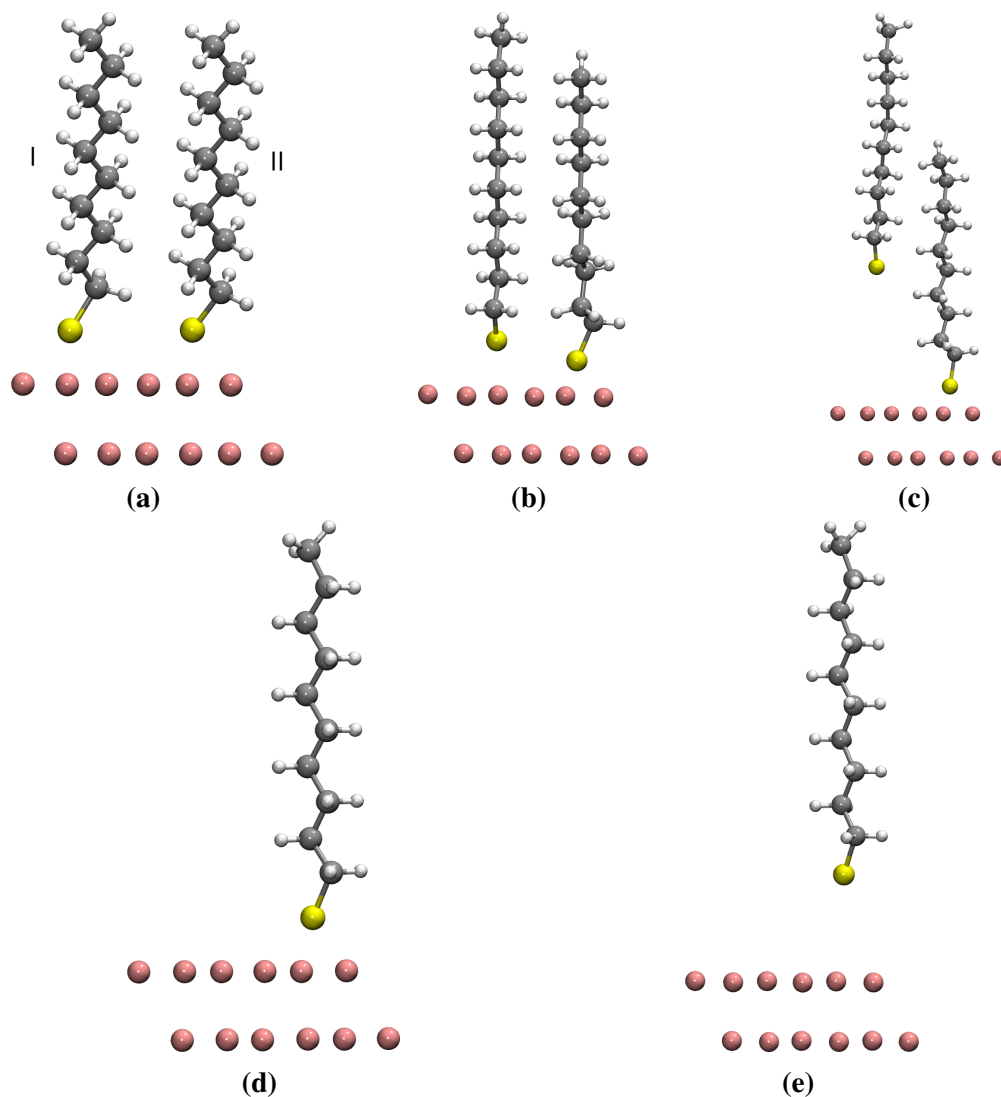


Figure 9.19: Overview of the distinct stages while simulating the desorption of two C10-SAM molecules on a two-layered 3×2 gold surface. a) Initial geometry for $\mathcal{E} = 0$. b) Molecular rearrangement with one chain lying higher than the other at $\mathcal{E} = 0.04$ a.u. c) Desorption of the higher positioned C10-molecule during the field step $\mathcal{E} = 0.03 - 0.04$ a.u. d) Optimized structure of the remaining C10-molecule at a field strength of $\mathcal{E} = 0.04$ a.u. e) Evaporation of the second molecule during the field step $\mathcal{E} = 0.05 - 0.06$ a.u.

9.4.3 Evaporation of Amino-undecanethiol from Gold

Simulations of amino-undecanethiol on gold surfaces have been performed analogous to the previously described ones for the unmodified decanethiol. Starting from the smaller two-layered 2×2 gold (111) surface with one NH_2 -SAM molecule, the basic behavior of this SAM species is investigated. It is followed by the slightly bigger system consisting of two SAM molecules on a two-layered 3×2 gold (111) surface treating intermolecular interactions explicitly. As previously described for the C10-SAM case, the bulk optimized in plane lattice vectors were used for the surface construction. Calculations were performed upon initial relaxation of the SAM-system by incrementally increasing the field in steps of 0.005 a.u. with geometry optimizations performed after each increase of the field. The geometries resulting from every optimization were used as new initials for the following step. The obtained fragmentation patterns are summarized in Table E.2.

One NH_2 -SAM Molecule per Unit-Cell

The initially optimized unit cell shown in Figure 9.20 shows that the molecule in zero field is highly tilted with an angle of $\phi_{2 \times 2}^{\text{NH}_2} \approx 54.6^\circ$. The stages of the fragmentation process, simulated for the smallest system using the NH_2 -SAM molecule, are presented in Figure 9.21. For the following simulations under the applied electric field, the SAM molecule was geometry optimized while the metal atoms were held fixed. Upon application of an electric field, the NH_2 -molecule is aligning itself with the field axis as shown for the case of $\mathcal{E} = 0.04$ a.u. in Figure 9.21a. Further increase of the field, leads to evaporation of the whole NH_2 -SAM

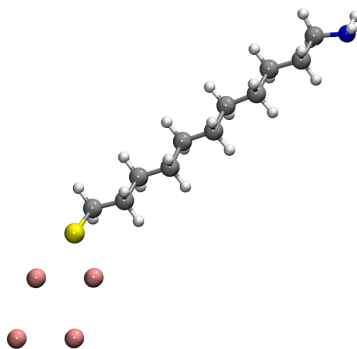


Figure 9.20: Optimized initial geometry of the smallest NH_2 -SAM system after geometry optimization at $\mathcal{E} = 0$ a.u. with GFN-xTB.

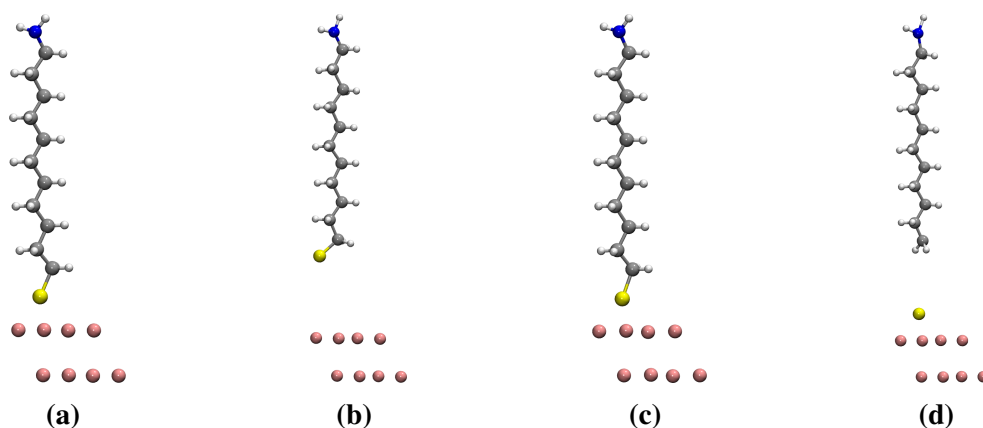


Figure 9.21: Simulated stages of the desorption process for one NH_2 -SAM molecule on a two-layered 2×2 gold surface. a) The straightened geometry after application of a field $\mathcal{E} = 0.04$ a.u., b) Complete desorption of the NH_2 -SAM molecule during the field step $\mathcal{E} = 0.08$ a.u., c) The optimized geometry with fixed sulfur atom at a field strength of $\mathcal{E} = 0.06$ a.u., d) Desorption of the $\text{NH}_2\text{C}_{11}\text{H}_{22}^+$ fragment during the field step $\mathcal{E} = 0.07 - 0.08$ a.u.

molecule during the field step $\mathcal{E} = 0.075 - 0.08$ a.u. as illustrated in Figure 9.21b. In a second evaporation series, the sulfur atom was fixed during the optimization. The remaining atoms of the flexible chain, were observed to align with the field axis as expected from the previous simulations. The situation is shown for an external field of $\mathcal{E} = 0.06$ a.u. in Figure 9.21c. During the field step $\mathcal{E} = 0.075 - 0.08$ a.u. a $\text{NH}_2\text{C}_{11}\text{H}_{22}^+$ fragment is evaporated as shown in Figure 9.21d, evaporating all flexible atoms.

Two NH_2 -SAM Molecules per Unit-Cell

The stages of the fragmentation process, simulated for two NH_2 -SAM molecules on a 3×2 surface, are presented in Figure 9.22. In this case, the initial tilt of the molecules is much less pronounced leading to an average tilting of $\bar{\phi}_{3 \times 2}^{\text{NH}_2} \approx 39.9^\circ$. Upon application of the external field, the molecules are rearranged with chain I positioned higher than chain II. Furthermore, the formation of a hydrogen bond could be observed as shown in Figure 9.22b for an external field of $\mathcal{E} = 0.03$ a.u.. The higher positioned chain is then observed to evaporate during the field step $\mathcal{E} = 0.035 - 0.04$ a.u. as shown in Figure 9.22c. Reoptimization of the remaining NH_2 molecule in a field of $\mathcal{E} = 0.035$ a.u. leads to alignment of the remaining molecule with the field axis. Upon increasing the field to $\mathcal{E} = 0.04$ a.u. Figure 9.22e, the remaining NH_2 -SAM molecule is evaporated.

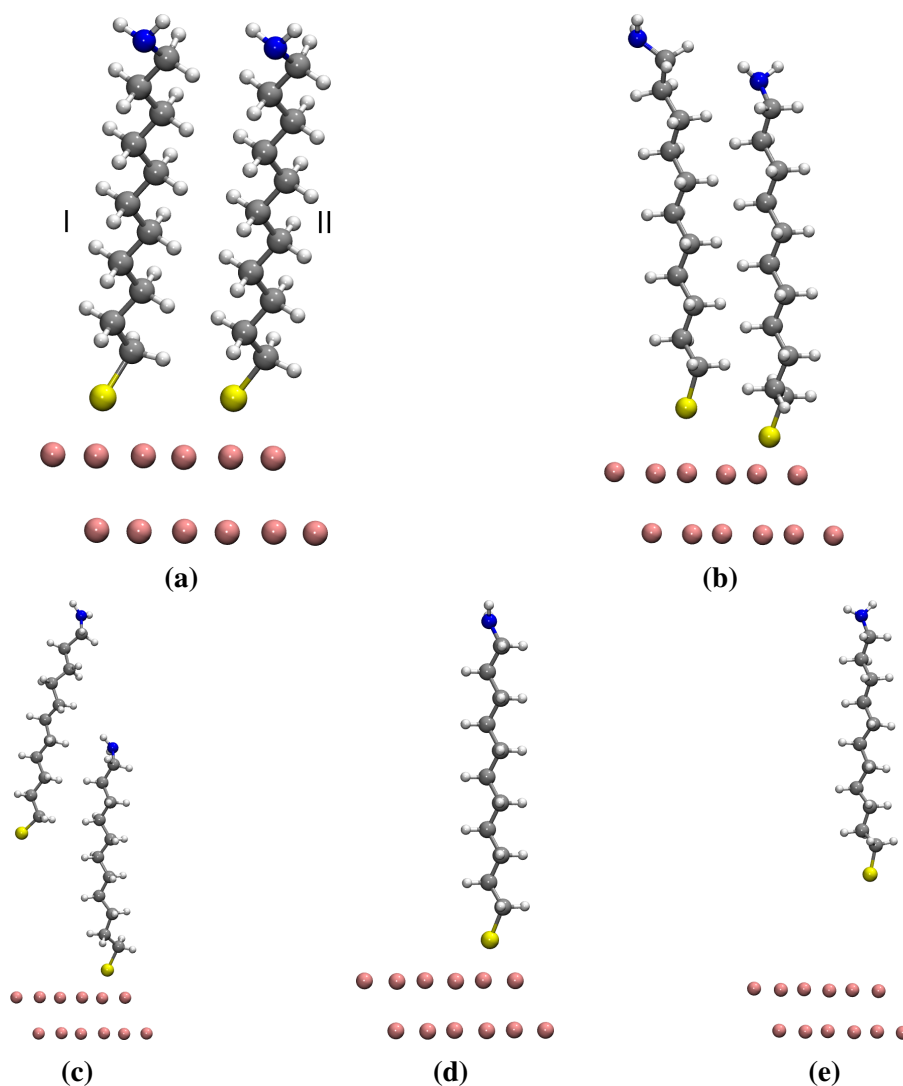


Figure 9.22: Simulated stages of the desorption process for two SAM-NH₂ molecules on a two-layered 3 × 2 gold surface. a) The initial geometry after optimization at $\mathcal{E} = 0$ a.u., b) The head groups showing a hydrogen bonded after increasing the field to $\mathcal{E} = 0.03$ a.u., c) Desorption of the higher positioned NH₂-molecule during the field step $\mathcal{E} = 0.03 - 0.04$ a.u., d) Reoptimized geometry for the remaining NH₂-molecule in a field of $\mathcal{E} = 0.04$ a.u., e) Desorption of the second NH₂-molecule during the field step $\mathcal{E} = 0.06 - 0.07$ a.u.

9.4.4 Evaporation of Fluorinated Decanethiol from Gold

Simulations of fluorinated decanethiol on gold surfaces have been performed in analogy to the previously described ones for the C10-SAM, for two-layered 2×2 and 3×2 gold (111) surfaces with in plane lattice vectors obtained from the cell optimized bulk calculations. Calculations were performed upon initial relaxation of the SAM-system by incrementally increasing the field in steps of 0.005 a.u. with geometry optimization for every field step. The geometries resulting from every optimization were used as new initial ones for the following calculations alongside the wave function calculated in the last optimization step. The obtained fragmentation patterns are summarized in Table E.3.

One F-SAM Molecule per Unit-Cell

The stages of the fragmentation process, simulated for the smallest system using the F-SAM molecule, are presented in Figure 9.23. The initial structure of the F-SAM molecule, shows lower total tilt with respect to the surface normal as in the NH_2 and C10 case. However, it is the only of the three investigated species with internal bending observed at $\mathcal{E} = 0$ as shown in Figure 9.23a. Upon increase of the electric field, the F-SAM molecule aligns itself with the field axis as exemplary shown for $\mathcal{E} = 0.04$ a.u. in Figure 9.23b. The head-group is evaporated as a CF_3^+ ion during the field step $\mathcal{E} = 0.055 - 0.06$ a.u. as shown in Figure 9.23c. During reoptimization of the remaining molecule in a field of $\mathcal{E} = 0.06$ a.u., the new headgroup is reoriented and slightly rotates itself with respect to the other segments. The remaining chain is evaporated during the field step $\mathcal{E} = 0.075 - 0.08$ a.u. leading to a $\text{C}_9\text{F}_{14}\text{H}_4\text{S}^+$ fragment.

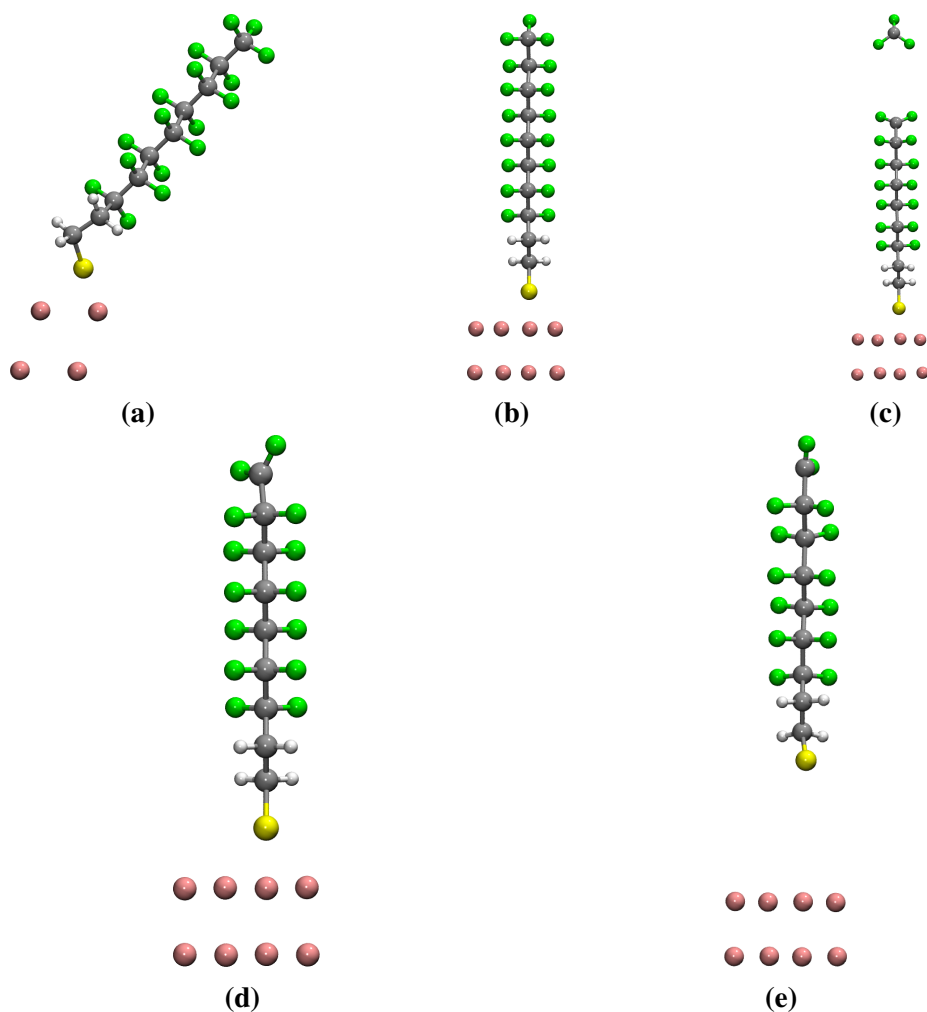


Figure 9.23: Simulated stages of the desorption process for one SAM molecule on a two-layered 2×2 gold surface utilizing GFN-xTB. a) The molecule after initial geometry optimization. b) Reoriented molecule after optimization in a field of $\mathcal{E} = 0.04$ a.u. c) Desorption of a CF_3^+ -ion during the field step $\mathcal{E} = 0.055 - 0.06$ a.u.. d) Relaxed structure after deletion of the CF_3^+ fragment at a field strength of $\mathcal{E} = 0.06$ a.u.. e) Desorption of the remaining molecule at the field step $\mathcal{E} = 0.075 - 0.08$ a.u. .

Two F-SAM Molecules per Unit-Cell

The stages of the fragmentation process, simulated for two F-SAM molecules on a 3×2 surface, are presented in Figure 9.24. Comparing the F-SAM molecules to the previously studied SAM systems, reveals them to be less tilted. The relaxed molecules are shown in Figure 9.24a, where the staggered arrangement of fragments as described in Section 8.6 is observed in order to minimize repulsive inter-molecular forces. Upon application of an

external field, one chain is elevated which is indicated by an increase in the sulfurs distance from the gold surface. For chain I the distance of sulfur to the plane of the uppermost gold layer is given as $d_S(0.00 \text{ a.u.}) = 1.71 \text{ \AA}$, whereas it is given as $d_S(0.04 \text{ a.u.}) = 2.41 \text{ \AA}$ for chain II. After increasing the field, a CF_3^+ fragment is evaporated from chain II, during the field step $\mathcal{E} = 0.055 - 0.06 \text{ a.u.}$. Reoptimization of the remaining molecules after deletion of the CF_3^+ ion in a field of $\mathcal{E} = 0.05 \text{ a.u.}$, leads to the situation shown in Figure 9.24d. The whole system is evaporated during the field step $\mathcal{E} = 0.065 - 0.07 \text{ a.u.}$ leading to evaporation of a $\text{C}_9\text{F}_{14}\text{H}_4\text{S}^+$ as well as a $\text{C}_{10}\text{F}_{17}\text{H}_4\text{S}^+$ ion.

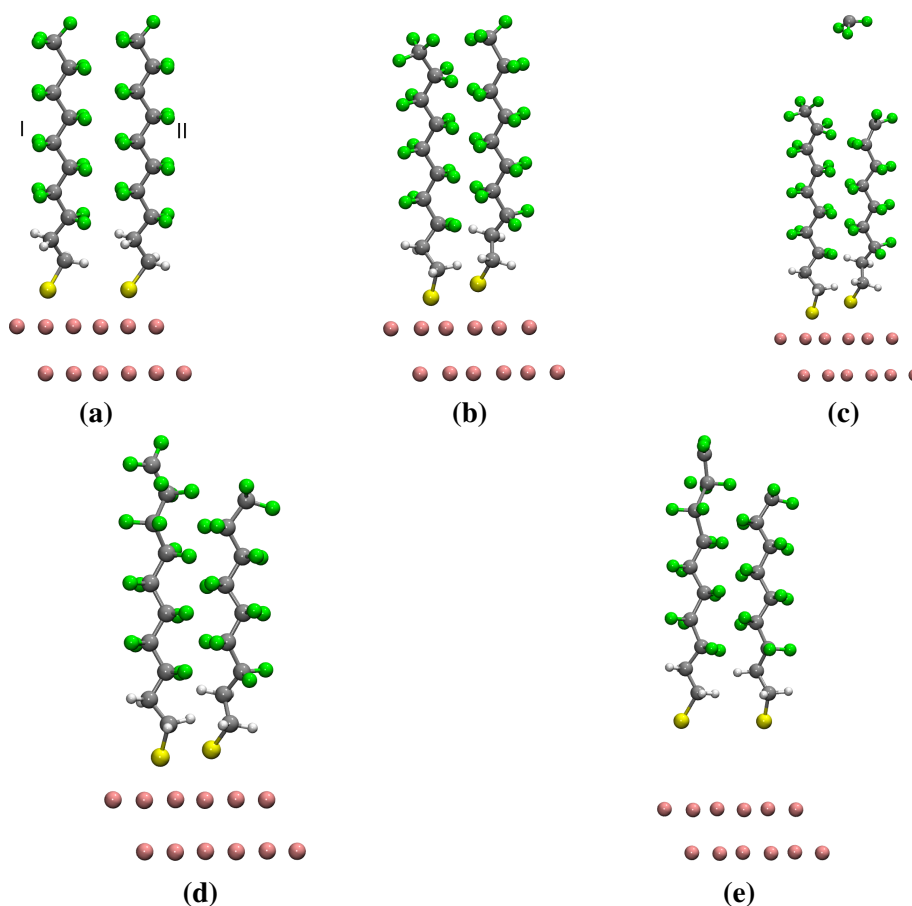


Figure 9.24: Simulated stages of the desorption process for two F-SAM molecules on a two-layered 3×2 gold surface utilizing GFN-xTB. a) The molecules after initial geometry optimization. b) Reorganized molecules after optimization in a field of $\mathcal{E} = 0.04 \text{ a.u.}$ c) Desorption of a CF_3^+ -ion during the field step $\mathcal{E} = 0.055 - 0.06 \text{ a.u.}$ d) Relaxed structure after deletion of the CF_3^+ fragment at a field strength of $\mathcal{E} = 0.05 \text{ a.u.}$ e) Desorption of the remaining molecules during the field step $\mathcal{E} = 0.065 - 0.07 \text{ a.u.}$.

9.5 Discussion of the Results

The conducted study showed the successful application of both utilized descriptions for all presented systems. For both methods, the initially conducted studies from the reconstructed gold surface, showed an evaporation field strength of $\mathcal{E}_{\text{crit}}^{\text{DFT}} = 0.065$ a.u. as explained in Section 9.3.2 and $\mathcal{E}_{\text{crit}}^{\text{GFN-xTB}} = 0.065$ a.u. in the case of GFN-xTB Section 9.4.1. From experiments, an evaporation field of around $\mathcal{E}_{\text{crit}}^{\text{exp}} = 0.055 - 0.06$ a.u. is determined for pure gold tips²⁰⁸. The evaporation fields determined for the reconstructed surfaces, are therefore in good agreement with the experimental observation for both methods. In simulations of SAM surfaces, also higher fields might be observed without evaporation of substrate atoms due to the dielectric shielding of the surface caused by charge accumulation at the interfaces between the layers¹⁹.

Investigations of the decanethiole reference system showed the fundamental charge shifts arising within the thioles backbone during the application of an electric field. The Hirshfeld charges obtained from the GPW-DFT calculations of the 2×2 system, indicated a charge reorganization in the SAM molecules, leading to the systems polarization. Furthermore, the transfer of an electron from the molecule to the metal surface is indicated by the charge surplus within the SAM, which is observed to be largest right before the evaporation. This can be exemplary seen for the completely ionized C10-molecule as documented in Table 9.1. The same was observed for the NH_2 -SAM Table 9.5, but was most pronounced in the F-SAM case Table 9.9. Investigation of the systems bulk polarization as defined in the MTP following Eq (3.23), showed a systematic increase in its field parallel component P_z for all three SAM species. Here, the behavior of C10-SAM and F-SAM was similar, differing only due to the initially more polarized fluorine system as can be seen from the in plane components P_x and P_y . For the NH_2 -SAM, the total polarization increases twice as strong with the fields increase and is mainly found in the P_z component alone. A possible reason for this, can be found due to the molecules symmetry of segments canceling out the resulting in plane components for the total system. Furthermore, the P_z component is impacted by the straightening of the molecule during field increase accompanied with the strong charge transfer from the head group to the surface.

During the increase of the external field, the titling of the molecules with respect to the field

axis was reduced and vanished in all three cases upon ionization of the first fragment in the case of GPW-DFT. In simulations utilizing GFN-xTB this effect was already observed before ionization of the first fragment as observed in Section 9.4.2. Crucial differences between the simplest model of one molecule on a 2×2 surface and two molecules on a 3×2 surface were observed. Using GPW-DFT, evaporation fields for the whole SAM in the 2×2 case were slightly higher than in the larger system when treating the C10-SAM and NH₂-SAM. In simulations using GFN-xTB, these differences are even more pronounced with the evaporation fields being $\mathcal{E} = 0.03 - 0.04$ a.u. and $\mathcal{E} = 0.09 - 0.10$ a.u. for the ionized C10 chain. Similar effects were observed for the NH₂-SAM where in the smaller system the molecule is desorbed at the field step $\mathcal{E} = 0.07 - 0.08$ a.u.. However, in this case the GFN-xTB calculations lead to two separate field strengths $\mathcal{E} = 0.03 - 0.04$ a.u. and $\mathcal{E} = 0.06 - 0.07$ a.u. for the respective chains in the 3×2 case. A possible reason lies in the reduced space available for reorganization in the smaller system as well as more "rigid" modelling due to intermolecular interactions only being present in form of the periodic images. This assumption is further supported by the results of the F-SAM investigations where these effects are observed most clearly. As worked out in Chapter 8, the F-SAM molecules, require more space due to the fluorine atoms at the molecular backbone, which is especially consequential for the smallest system (compare Section 8.6). In the DFT case, the F-SAM molecule in the 2×2 system is fragmented, however the remainders stay attached to the surface even for fields larger than the critical strength for gold. In the larger system, both molecules are evaporated at field strengths between $\mathcal{E} = 0.04 - 0.045$ a.u., with CF₃⁺ and CF₂⁺ fragments observed beforehand. In the GFN-xTB case, a similar effect is observed however with smaller differences in the evaporation field strengths e.g. $\mathcal{E}_{2 \times 2} = 0.075 - 0.08$ a.u. and $\mathcal{E}_{3 \times 2} = 0.065 - 0.07$ a.u. in the larger case. Possible reasons can be attributed to the slightly larger surface areas in the GFN-xTB calculations, arising from the usage of bulk optimized in plane lattice constants. In summary, the evaporation of the SAMs from the surface was observed for both methods and all species.

The results of both methods used, are similar, but differ when considering the obtained evaporation field strengths that are significantly higher in the GFN-xTB calculations. This can be exemplarily seen for the first fragment of the F-SAM series CF₃⁺ that evaporates at

9 Field Evaporation of Self Assembled Monolayers

$\mathcal{E}^{\text{DFT}} = 0.03 - 0.04$ a.u. in the DFT case and $\mathcal{E}^{\text{GFN-xTB}} = 0.055 - 0.06$ a.u. using GFN-xTB. Interestingly, only in the case of the F-SAM molecule real fragmentation patterns could be observed. This leads to the conclusion, that intramolecular C-C bonds are described as stronger than the sulfur metal bond for the unfluorinated SAMs. As can be concluded from the observed charge shifts, the CF_3 headgroup gets ionized much stronger than the others, showing more flexibility of the electronic redistribution. As discussed in¹⁹, the flexibility of electron polarization is strongly impacted by the diffuse character of the describing basis functions. This leads to the impression, that the electronic distribution is much less distorted in the case of F-SAM molecules than in the unfluorinated molecules. Therefore, diffuse functions are much more crucial to describe the polarization of the C10 and NH_2 -SAM. However, in the utilized framework of GPW-DFT, they are not affordable due to high condition numbers as well as computational costs¹⁰².

When comparing the fragmentation patterns for F-SAM with previously conducted approaches of isolated molecules attached on single metal atoms¹⁸ as well as cluster simulations¹⁹, fundamental differences occur. These consist mainly in different field strengths for the first fragment CF_3^+ which is observed to evaporate at $\mathcal{E}^{\text{iso}} \approx 0.015 - 0.02$ a.u. in the isolated molecule case and $\mathcal{E}^{\text{surf}} = 0.035 - 0.04$ a.u. in the surface model. Furthermore, larger fragment types such as C_2F_4^+ are observed in the non periodic models, a possible reason lies in the dense packing magnifying the influence of the staggered arrangement of SAM molecules. This assumption is indicated by the alternating evaporation of identical fragments from chain I and II.

9.5.1 Limitations of the utilized Models and Outlook

The presented work marks only a first step towards the description of APT experiments and the field desorption process for solids. Experiments for the evaporation of molecules in APT, use a static field to align the SAM and allow the propagation of evaporated ions towards a detector. However, a pulsed laser is used to trigger the evaporation event. As direct consequence, static calculations as described in this chapter, are not accounting for the dynamic character of the laser induced evaporation. A model using molecular dynamics simulations within a standing field and a time dependent Gaussian pulse would serve this purpose. Direct usage of an electric

pulse would eliminate the necessity of assuming a thermal triggering of the fragmentation process. Significant improvement of the surface description would be achieved when treating more layers of a reconstructed metal surface as well as flexible lattice constants during the field application. A straight forward improvement would consist therefore in increasing the number of gold layers as well as the surface size e.g. to four molecules on a four-layered 3×3 reconstructed gold surface. Furthermore, the description of the complete electronic structure by avoiding the use of pseudopotentials as well as using more diffuse functions for the DFT description, are expected to improve the description. Unfortunately, the impact of different exchange and correlation functionals is unknown, especially considering hybrid and non-local contributions. However, this is unlikely to change in the near future as the computational costs for their application are not affordable for the systems at hand. Improvements of the obtained results with GFN-xTB by using supercell calculations, is not expected. This can be attributed to the method itself, as it relies heavily on integral parameterization in the field free case, which is not aiming at suitable description of polarization effects and external electric fields (compare Section 4.3). Furthermore GFN-xTB uses a minimal number of compressed atomic basis functions (compare Section C.3), limiting the electronic reorganization when compared with more diffuse functions as expected to be necessary for describing the C10 and NH₂-SAMs fragmentation. Additionally, comparison of fragments with the experimentally obtained mass spectra, is difficult due to possible post ionization events as well as laser induced effects.

Chapter 10

Theoretical Analysis for Covalent Organic Frameworks

The theoretical methods used for the analysis of covalent organic frameworks (COFs) and their general properties are presented in this chapter. Subsequently, the method framework is applied for determination of COF-characteristics and the influence of interlayer interactions on the obtained pore size. Initially, an overview of COFs as a class of crystalline polymers, the different types and general properties available in the literature, are briefly discussed. Afterwards, the setup and basic properties of the utilized GPW-DFT and CCM implemented GFN-xTB⁶¹ methods within CP2K⁶² are explained for the investigated example COF systems. Starting from the GPW-DFT setup adjustment, the most significant contributions for interlayer interactions are shown at the example of two-layered unit cells for the planar COF-LZU1 and N3-COF systems. The setup is adjusted by analysis of the impacts arising from XC-functional, basis set, PW-cutoff and especially vdW-corrections on the target properties. Based on efficiency considerations, the used standard setup is chosen, balancing the required accuracy and computational effort. Additionally, the GFN-xTB and GPW-DFT descriptions are compared and discussed in Section 4.3, in order to justify the applicability of the by far cheaper GFN-xTB description for the calculation of interlayer displacement maps as explained in Section 10.5.1 utilizing a PW-interpolation. With the presented set of methods, the investigated COF systems are characterized by creation of the most suitable setup and

compared to the experimental literature. For this purpose, GPW-DFT and GFN-xTB as well as the whole spectrum of constructed tools presented in Part II are applied. The suitability of the respective methods is then interpreted by comparison with the experimental data. The chapter closes with a “manual” for describing COFs of the investigated types by theoretical means, to simplify future work on COF-systems.

10.1 What are COFs?

Covalent organic frameworks (COFs) are a class of crystalline polymers building porous materials by strong covalent bonds of organic secondary building units (SBUs) composed of light elements (e.g. C, N, H, O, F)²⁰. Generally they are classified in two and three dimensional species. In the two dimensional case, layers are stacked and connected non-covalently by e.g. vdW- and π - π interactions. Whereas in the three dimensional case, the stacked layers are connected by covalent bonds. The most significant characteristics of COFs are high thermal stability, specific surface areas and porosity²⁰⁹. In contrast to the class of metal organic frameworks (MOFs) that, besides the organic compounds, also contain metal atoms, the molecular weight and therefore the density of COFs is comparably low^{210,211}. These characteristics allow the use of COFs for applications such as gas storage, separation, drug delivery, sensing, catalysis, optoelectronic materials and medical purposes^{21,22,27–29,32,38–40,45}. Due to the variety of available SBUs and the tremendous number of their possible combinations for COF synthesis, large numbers of frameworks were either synthesized or hypothetically conceptualized²¹². Choosing the most suitable framework for the task at hand, therefore marks a crucial step for their actual applicability. For a specific task, frameworks are generally fine tuned by changing linkers or incorporating functional groups⁴². The vast amount of possible materials and framework designs, requires their categorization as established in COF-data bases, providing crystal structures and framework properties^{46–48}. However, performing experiments for such a large number of possible frameworks, seems highly non economic and the screening of COFs for the desired properties is necessary. Theoretical investigations in combination with automated workflows for characterization are therefore of main importance. Such methodology is applied e.g. in the field of gas capture and separation^{213–215}, requiring

theoretical treatment of the crystal structure²¹⁶. The reliability of the used work flow is determined by the choice of theoretical description and its accuracy for the target properties, most fundamentally determined by the crystal structure.

10.2 The Choice of a Suitable Theoretical Description

Up to present day, only a small number of theoretical studies have been performed for COFs including e.g.^{217–219}. A main issue for possible investigations arises due to typically large unit cells with ≈ 80 atoms per unit cell per layer for small systems. The regular size of even the smallest COF systems, excludes high scaling methods such as wave function based approaches or hybrid DFT, from application. Remaining, are methods based on classical force fields (FFs), density functional tight-binding (DFTB) and density functional theory (DFT). Furthermore, as COFs are crystalline polymers their successful description requires the use of a solid state description by PBCs. Therefore, the use of PW-DFT methods seems straight forward, however as COFs are highly porous, large fractions (typically more than $\approx 60\%$) of the unit-cell volume is empty. Consequently the use of PWs leads to unnecessary large computational effort, as PWs scale with the systems unit-cell size as described in Section 2.8.2. Using GPW-DFT, allows for much more efficient treatment by avoiding these drawbacks as described in Section 2.11.1. Another possibility used in this thesis is given by the density functional based CCM implemented GFN-xTB description, allowing for an efficient treatment of even the largest molecular systems analyzed within this thesis (compare Section 4.3).

One major obstacle when comparing theoretically obtained structures with the experiment, is the non-uniqueness of the obtained X-ray diffraction patterns and the inherent inaccuracies in real world measurements. The exact structural arrangement, even of a crystalline structures such as COFs, is not directly reproducible as the investigated sample will always consist of a multitude of stacking schemes.

To minimize the size of an investigated system, it is suitable to reduce the unit cell to only the necessary features, e.g. treatment of only two individual layers in order to determine favorable stacking schemes. Here, the fundamental structural properties can be determined as representative for the computationally much more expensive super-cell systems. The

corresponding target properties are intra-molecular arrangements (in-plane cell vectors a, b), interlayer distance (stacking vector c), charge distribution (ρ_{el}), interlayer displacements (d) and the related energetic displacement landscape $E(d)$ influencing the resulting pore size ($\Phi(d)$). Crucial for the determination of these properties is a suitable description of the underlying interactions (e.g. electrostatics) between adjacent layers.

10.3 Electrostatic Interactions

The overall electrostatic interactions in planar two dimensional COFs of the investigated types, are predominantly determined by delocalized electrons in aromatic rings. Non covalent interactions between aromatic rings are referred to as $\pi - \pi$ interactions, however, the nomenclature is not unique²²⁰. The underlying phenomenon is well studied in the literature and a complete overview of the related work is impossible. Main focus points important for this thesis are related to theoretical work treating unmodified aromatic rings^{221–224} as well as influences of substituent effects^{225–228}. Interestingly, suitable tuning of the mentioned interactions allows the creation of completely non-covalent frameworks based on π -stacking²²⁹. In aromatic systems, the delocalization of electrons in the ring leads to the formation of π orbitals. The accumulation of negative charge in the ring's center, leads to a separation of charges as illustrated in Figure 10.1.

The illustrated surfaces are created by calculating a property called molecular electrostatic potential (MEP) at positions r_{iso} corresponding to equal values of the electron density $\rho_{\text{el}}(r_{\text{iso}})$.

The MEP is defined as the force acting on a positive probe charge and calculated for points r_{iso} on the isosurface calculated by means of DFT on the B3LYP/TZVP-level using the GAUSSIAN08²³⁰ program package. Visualization was performed using GAUSSVIEW²³¹ the graphical interface to GAUSSIAN.

In the case of benzene rings, delocalized electrons in the center of the rings, lead to the accumulation of negative charge acting perpendicular to the molecular plane and positive charge at the molecular edges as can be seen in Figure 10.1a. The density surface in between the ring atoms is colored in orange, indicating the delocalized π orbitals, leading to electron surplus interacting attractive with a positive probe charge. H-atoms positioned on the outside

of the ring carry a positive partial charge and consequently, the resulting charge distribution shown in Figure 10.1b, acts like an electronic quadrupole. The cross-section of the electrostatic potential lines further illustrates this interpretation. As a result, the interaction of two benzene rings leads to three distinct stable categories, illustrated in Figure 10.2. Here the underlying types of arrangements, referred to as sandwich Figure 10.2a, parallel displaced Figure 10.2b and T-shaped Figure 10.2c are indicated by chemical formula.

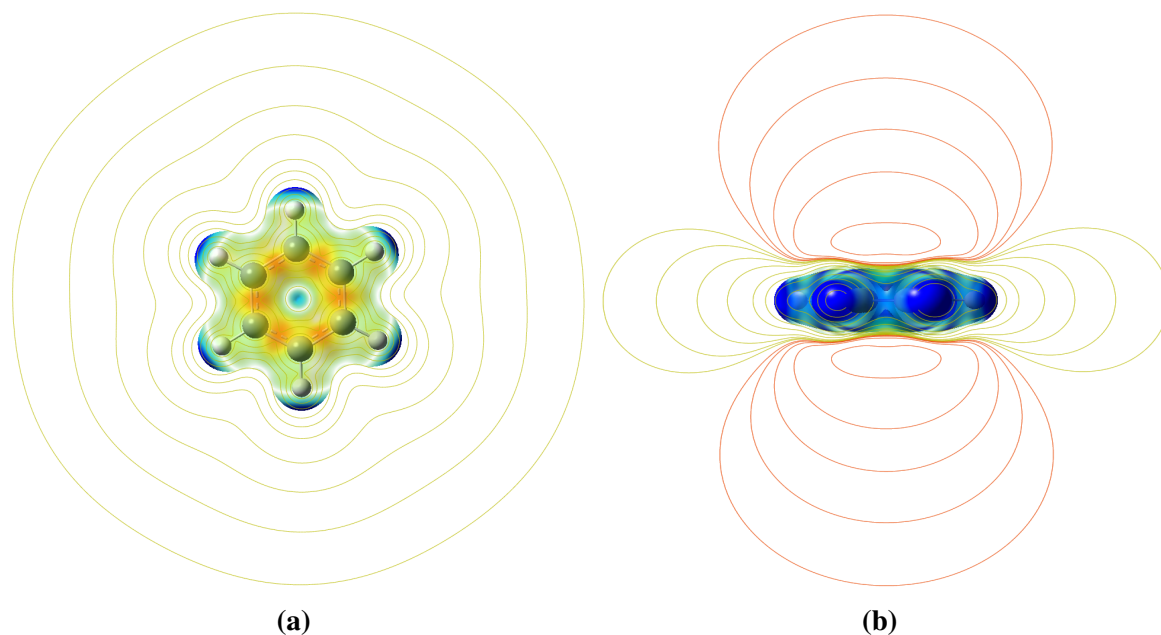


Figure 10.1: Overview of the electrostatics for a benzene ring as calculated by means of DFT on the B3LYP/TZVP level using the GAUSSIAN program package. The shown surface was created by mapping the MEP onto an iso surface of the electron density ρ_{el} , created for a density isovalue of $\rho = 0.02$. Red corresponds to a MEP value of -0.02 a.u. and blue with 0.2 a.u. . a) Top view onto the molecule with the in-plane potential lines indicating the hexagonal symmetry. b) Side view onto the molecule with the electrostatic potential arising perpendicular to the molecular axis.

In polymeric crystals that are dominated by aromatic compounds, the resulting stacking schemes are determined by the additive interactions of various contributions. The most prominent example for $\pi - \pi$ stacking is given by graphite, where graphene sheets are stacked with an approximate in-plane shift of $r = 1.55 \text{ \AA}$. This leads to the overlap of vicinities of different charge interacting attractively and therefore stabilizes the geometry in this configuration. Similar effects arise in the investigated two-dimensional COFs, where shifts between the layers are observed to be determined by overlapping interactions of aromatic rings between adjacent

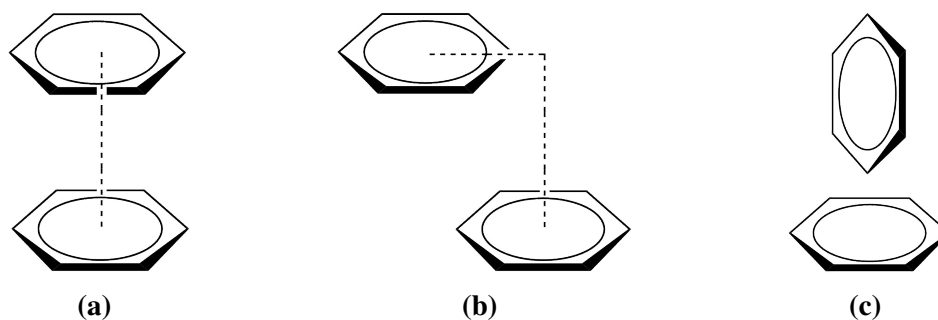


Figure 10.2: The three main types of interactions arising between two benzene molecules depending on their stacking scheme. a) The sandwich stacking scheme. b) The parallel displaced stacking scheme. c) The T-shape stacking scheme.

layers. The sandwich and parallel displaced structures are observed in the investigated study of planar COF systems.

When analyzing planar COFs, the main interactions between adjacent layers are attributed to the electrostatics. In turn, these are given by superimposed interactions stemming from aromatic rings similarly to isolated benzene dimers as shown in Figure 10.3 at the example of a shifted two-layered COF-LZU1 system. The systems density ρ_{el} and electrostatic potential V_C was calculated using CP2K for the cell optimized system for slightly displaced initial structures. The obtained distribution of electrostatics shows the similarities with the isolated benzene case. Here the color scheme is given in reverse as V_C is mapped onto an iso-surface of ρ_{el} , leading to negative potentials being displayed in blue whereas attractive is illustrated in red. For eclipsed structures, the interactions of the aromatic rings correspond largely to the sandwich type where same sign charges are positioned directly above each other. This arrangement of course leads to maximized repulsive interaction between the COF-sheets, which can be partially compensated by slightly displacing the layers with respect to each other. Displacing the systems from eclipsed configuration (AA) leads to a symmetry break, which allows readjustment of the layers during optimization by minimizing the repulsive overlap by displacement of aromatic rings. Within COF systems, the π - π interactions between layers are accompanied by others such as repulsion between imine bridge interactions of subsequent layers.

T-shaped stacking:

Another typical interaction of benzene rings is the so called T-shape arrangement schematically

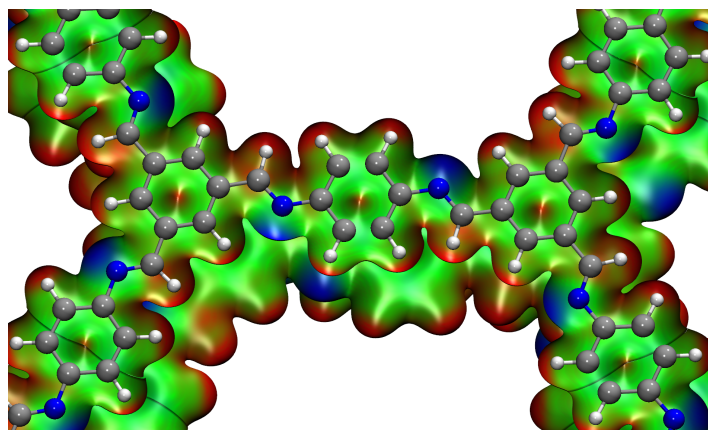


Figure 10.3: Top view onto V_C mapped onto a density iso surface of ρ_{el} for the shifted COF-LZU1 system. Structures were obtained by cell-optimization on the PBE-D3/TZV2P-GTH level with corresponding GTH-PPs for all elements.

shown in Figure 10.2c, where two rings are oriented perpendicular with respect to each other. This arrangement is stabilized due to favorable electric quadrupole interactions within the dimer. When occurring in stacked sheets within a crystal, in contrast to the previously described stackings, the T-shape illustrated in Figure 10.4, requires out of plane rotation. Consequently, it is only relevant for the non-planar COF systems which are introduced at a simple example of the adj-COF in Section 10.6.2. However, T-shape stacking will be of major interest for systematically increasing the stability of large-pore COFs in Chapter 11. Here additional attachments are placed at the pores walls, allowing for systematic linkage by non-covalent interactions between adjacent layers.

10.4 Construction of the GPW-DFT-Setup

The investigation of COFs by means of DFT, requires a suitable setup providing the best accuracy available while still being computational affordable. A crucial simplification is the description of extended layers via their unit cell with PBCs. COFs form porous systems, leading to big vacuum regions within a unit cell. The GPW-DFT approach as implemented in CP2K is suited for the treatment of periodic system due to its usage of a PW-basis set for the electrostatic and exchange contributions (compare Section 2.11.3). Furthermore, the GPW approach capitalizes on the fact that the electron density vanishes for large parts of the unit cell, by using Gaussian basis functions (compare Section 2.8.1). In the following, the main

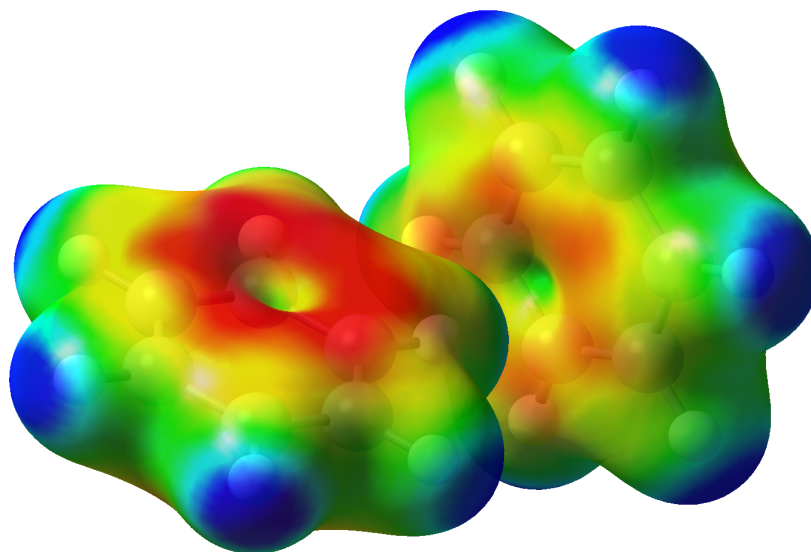


Figure 10.4: Molecular arrangement of two benzene rings forming a T-shape arrangement with the MEP mapped onto an iso value of ρ_{el} . The arrangement shows opposite sign charges orienting towards each other after optimization by means of DFT on the B3LYP/TZVP level.

considerations for a suitable applicability of the method are explained, resulting in the used standard setup. Emphasis is taken on the choice of XC-functional, Gaussian basis as well as the PW-cutoff for the corresponding Fourier expansion. Further technical details such as cutoff values for the overlap matrix elements, SCF-accuracy and integration grids are briefly discussed. The initial adjustments are benchmarked at the example of COF-LZU1 and the N3-COF systems. Furthermore, the impact of dispersive effects is investigated by considering the DFT-D series as well as non-local (NL) XC-functionals.

10.4.1 Functional Choice

Using the PBE functional for the exchange and correlation in GPW-DFT calculations within CP2K is the natural choice, due to optimized PBE-GTH-PPs and a wide variety of basis functions being available (compare Section 2.10). However, functional specific optimized PPes also exist for BLYP, BP86 and the PADE implementation of the LDA. A limited number of optimized basis sets is available for BLYP. The three GGA functionals mark the possible standard choices for DFT calculations, with BLYP being mostly used for molecular systems as mentioned in Section 2.5.2. In turn, the PBE functional is the standard choice for solid

state calculations. The BP86 functional can furthermore be used as a reference, however, it was excluded as a possible choice, due to unavailability of optimized basis sets and a lack of available parameters for the DFT-D series. For highly homogeneous electron densities, as arising when treating one atomic crystal structures with PPs, the LDA usually leads to accurate descriptions. This is not expected in the case of COFs, as these are molecular systems with highly fluctuating density regimes (compare Section 2.5), disqualifying the LDA as a possible choice. The PBE functional marks not only the literature standard for solid state calculations but CP2K provides PBE optimized basis sets, PPs and DFT-D parameters. Consequently PBE is used as standard functional for all calculations within this chapter. It is known from literature, that GGA functionals generally describe π -stacked systems as to repulsive. This effect is partially compensated by dispersion corrections such as the used DFT-D type. However, specifically for the PBE functional, the obtained equilibrium distances for π - π interactions are generally described as to repulsive^{232,233}. Furthermore, it is known that energetic differences from PBE-D3 calculations are modeled to small for π -stacked systems²³⁴. In turn, the PBE-D3 description has been shown to yield productive results for aromatic components in periodic calculations²³⁵. For COF systems especially, the BLYP functional therefore marks an alternative description, as it is known to describe energetics for non-covalent interactions more accurately for a multitude of test sets²³⁶. However, functionals with the LYP correlation are known for their inherent tendency to over stabilize vdW-bonds²³⁷. Which of the tendencies is more important and the associated functional choice depends highly on the specifics of the investigated system.

The impact of the XC-functional, was tested for the N3-COF by cell optimization using a TZV2P-GTH basis set with functional optimized GTH-PPs. The initial structures were created by displacing the top layer from the perfectly eclipsed case by 1.0 Å in x and y -direction respectively. The obtained unit-cell arrangements for both functionals are shown in Figure 10.5 and the obtained lattice constants are summarized in Table 10.1. The deviations of the obtained in-plane lattice constants are marginal with a deviation of $\Delta a \approx 0.07$ Å. This corresponds to a negligible deviation of $\approx 0.23\%$ between the functionals, showing that bonds within a planar COF system are described almost identical in both cases. Interestingly, the deviation of $\Delta c \approx 0.38$ Å in stacking direction corresponds to a deviation of $\approx 5\%$ purely due to

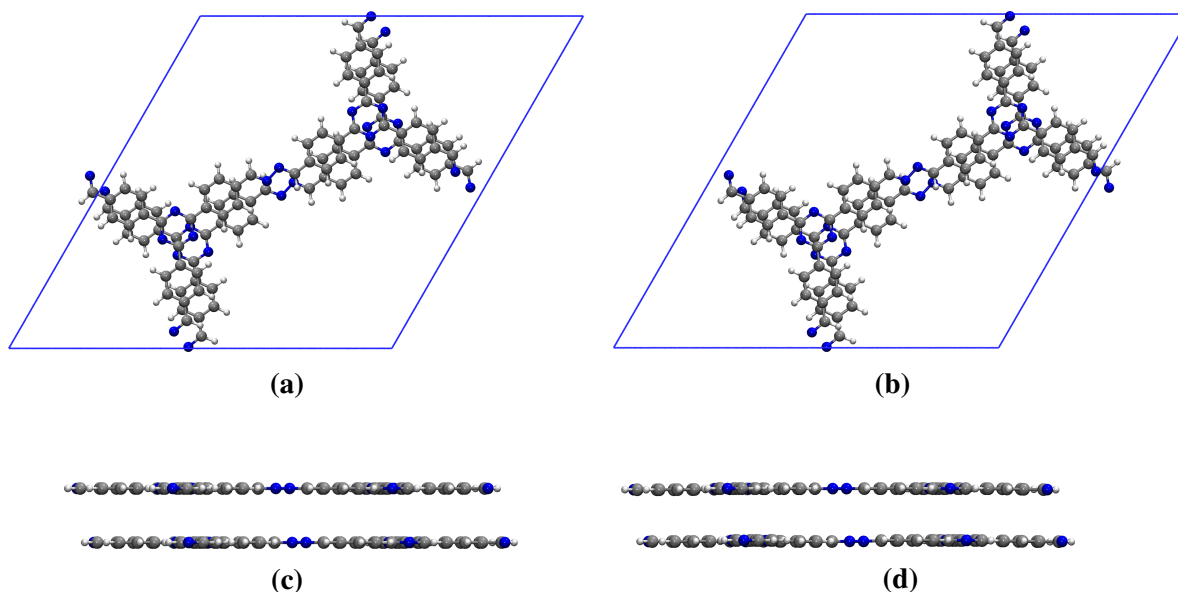


Figure 10.5: Overview of the relaxed unit cells for the N3-COF for both GGA functionals. a) Top view using the PBE functional, b) Top view using the BLYP-functional, c) Side view using the PBE functional, d) Side view using the BLYP functional.

XC-functional choice, directly affecting the modeled stacking behavior. For planar systems, the interlayer distance r_{int} can be extracted from the obtained c lattice constant, after dividing by the number of layers. An interlayer distance in the range of 3.2-3.8 Å is expected for planar π -stacked systems. In the investigated cases they were determined as $r_{\text{int}}^{\text{PBE}} \approx 3.65$ Å and $r_{\text{int}}^{\text{BLYP}} \approx 3.46$ Å. Experimentally, a value of $r_{\text{int}}^{\text{exp}} \approx 3.44$ Å was obtained, showing the BLYP case in perfect agreement for the investigated system. The interlayer slipping for adjacent layers was estimated using the in-plane distances of the central imine bridge nitrogens, following

$$d = \sqrt{(x_{\text{N1}} - x_{\text{N2}})^2 + (y_{\text{N1}} - y_{\text{N2}})^2} \quad , \quad (10.1)$$

with the coordinates x_{Ni} and y_{Ni} of the identical nitrogen atoms in layer one and two respectively. The obtained estimates for interlayer slipping were calculated to be $d^{\text{PBE}} = 1.58$ Å and $d^{\text{BLYP}} = 1.67$ Å. Both obtained geometries show identical alignment as can be seen in Figure 10.5a and Figure 10.5b. Furthermore both structures are perfectly flat as can be seen from Figure 10.5c and Figure 10.5d. When comparing the lattice constants for COF-LZU1 as presented later on, the PBE functional leads to better agreement with the experimental value as described in Section 10.6.1.

Table 10.1: The lattice constants for the optimized AA' unit cells for the two GGA functionals.

| Functional | $a = b$ in Å | c in Å | d in Å |
|------------|--------------|----------|----------|
| PBE | 29.47 | 7.30 | 1.58 |
| BLYP | 29.54 | 6.93 | 1.67 |

As there is no significant difference in the obtained in-plane vectors, the interlayer distance acts as the only criteria for the functional choice.

Meta GGA and Hybrid functionals (compare to Section 2.5) are not applicable in the case of COF unit-cell systems as the computational effort is too high. Due to readily available basis sets as well as GTH-PBE-PPs, the PBE functional as the literature standard for solid state calculations was utilized for all calculations in this chapter if not stated otherwise.

10.4.2 Basis Set Choice

The choice of a suitable basis set is motivated by finding the most efficient description with respect to accuracy and computational effort. Considering the properties of the investigated systems such as π - π and vdW-interactions as well as density smearing in aromatic rings, the density is in large parts not compressed. The addition of polarization functions is essential as dominant interactions arise due to hybridization in aromatic systems (π -stacks). Furthermore, long range interactions and density smearing indicate the use of diffuse functions in the basis set. However, their efficient use for systems as large as COF unit cells, treated by means of GPW-DFT, is impossible. Here, additional computational effort arises due to integration of the long Gaussian tails themselves, as well as the needed PW-expansion.

At the example of the N3-COF, a convergence study of the total molecular energy for polarized Gaussian basis sets was performed, using the default setup. ($E_{\text{cut}} = 300$ Ry, $N_{\text{GRIDS}} = 4$, $N_{\text{GRID,cut}} = 60$ Ry, PBE-D3, PBE-GTH-PPs) The obtained results show only a slight decrease in energy for increasing the basis set size further than TZV2P-GTH. Therefore, this basis set is chosen as default for the follow up investigations. An interesting addition is provided when using the slightly more diffuse basis sets of the MOLOPT type⁸⁸ that provide the lowest energy for the given system, indicating the potentially best description. These basis sets add extensive computational effort and are therefore not an option for large production runs, however, they

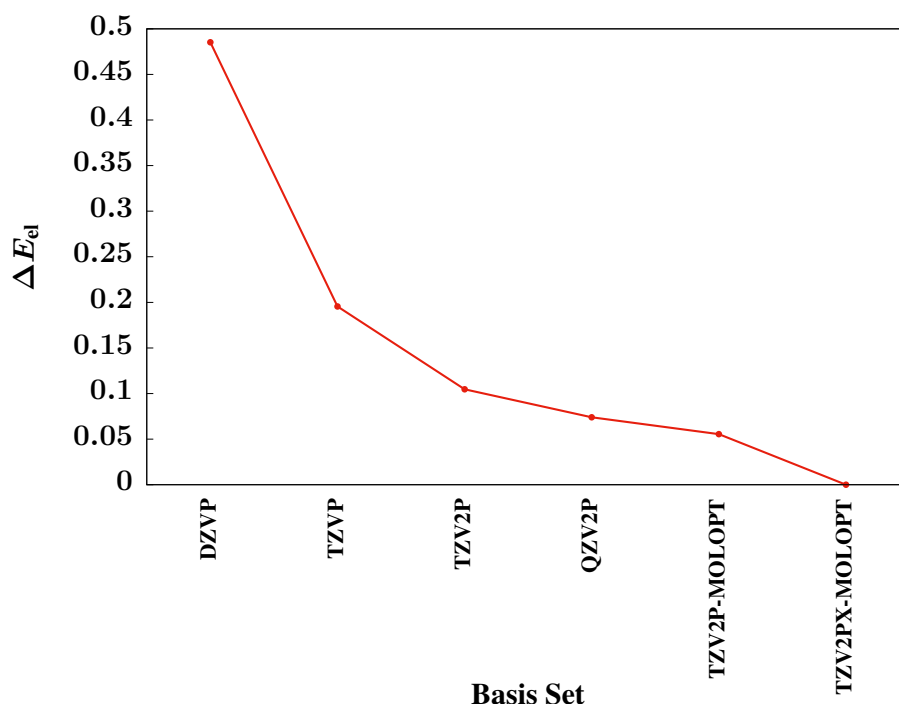


Figure 10.6: Deviation ΔE_{el} from the obtained minimum value $E_{el}^{TZV2PX-MOLOPT}$ for each basis set respectively.

are used for high accuracy single point energy calculations.

10.4.3 PW-Cutoff and Rel-Cutoff

A suitable cutoff value for the maximally allowed energy within the utilized PW-expansion (compare Section 2.11.1), is estimated by calculation of total energies for an optimized N3-COF unit cell in steps of 50 Ry for values $E_{cut} = 50 - 500$ Ry. The calculations indicate cutoffs around $E_{cut} = 300$ Ry to be accurate enough even though larger values yield even more accuracy. Sampling of the relative cutoff for the finest integration grid E_{rel} revealed no major impact and a value of $E_{rel} = 60$ Ry is chosen accordingly. For calculations where a higher cutoff is affordable and necessary, the value of $E_{cut} = 500$ Ry is chosen, showing the highest accuracy within the conducted study (compare Section 10.6.1).

10.4.4 The impact of Dispersion Corrections

Interlayer interactions between adjacent layers, mark the crucial property for determination of structural parameters in 2D-COF systems, which are dominated by non-covalent dispersion

forces. Different ways of accounting for dispersion are possible e.g. the highly parametrized Grimme DFT-D series¹⁵⁸ as well as non-local functionals. The impact of dispersion on the obtained molecular arrangements, is shown by cell optimizing a two-layered perfectly stacked unit cell with dispersion corrected and regular DFT. Calculations were performed using cell optimizations for the N3-COF on the PBE/TZV2P-GTH level with corresponding GTH-PPs, with and without DFT-D3 dispersion correction. The dispersion correction is most influential for non-covalent contributions and therefore directly influences the stacking behavior of COFs. The interlayer distances are mainly determined by van-der-Waals (vdW) interactions and r_{inter} as schematically shown in Figure 10.7. Comparison of the obtained interlayer distance for the optimized unit cell with uncorrected DFT and the D3 corrected case, shows the interlayer distance to decrease significantly when dispersion correction is applied. The resulting deviation in obtained lattice constants $\Delta c = 1.0 \text{ \AA}$ in stacking direction is a direct consequence of the pure repulsive description of $\pi - \pi$ interactions, provided by the used GGA-functionals^{158,238,239}. For a two layered completely planar system, this translates into a relative difference in the modeled interlayer distance of $\Delta r_{\text{inter}} = 0.5 \text{ \AA}$. Comparing the interlayer distances of $r_{\text{inter}}^{\text{PBE-D3}} = 3.6 \text{ \AA}$ with and $r_{\text{inter}}^{\text{PBE}} = 4.1 \text{ \AA}$ without dispersion correction, shows the dispersion corrected case to be significantly closer to the experimentally obtained value of $r_{\text{inter}}^{\text{exp}} = 3.55 \text{ \AA}$ ⁴⁰. Consequently, dispersion corrections are not only indispensable when investigating planar COF systems, but their exact nature seems to influence the obtained target parameters. In the planar case, the DFT-D3 correction is sufficient for describing most systems with the required accuracy. However, this is not necessarily the case especially when non-planar systems are investigated as will be discussed in Section 10.6.2. Consequences of using non-local functionals or other corrections of the Grimme type are discussed in the following.

A study of COF-LZU1 as well as the N3-COF, act as representative for planar COFs, showing the general impact of vdW-corrections of the Grimme type as well as NL-functionals. A study for the other planar COFs is not conducted separately, as the dominant interactions are of the same nature for all investigated systems. This can be easily seen when comparing the electrostatics of these systems illustrated by the MEPs for optimized AA' stackings as exemplary shown in Figure 10.3. Predominantly, these interactions are determined by stacked

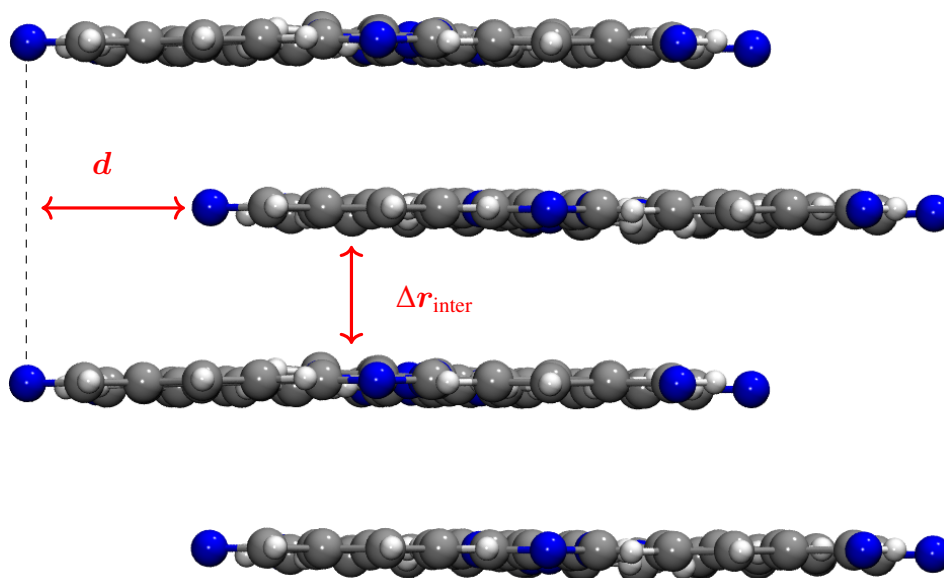


Figure 10.7: Side view onto the exemplary AA'-Stacking type investigated for COF-LZU1, with d being the in-plane displacement and Δr_{inter} the interlayer distance.

aromatic rings with their properties previously discussed in Section 10.3. The impact of the used correction on the resulting structural arrangement, is further investigated in the following, relating the obtained accuracy of target properties with the computational costs necessary.

DFT-D Corrections

The most commonly used correction for dispersion effects is given by Grimmes DFT-D¹⁵⁸ series, whose concepts are described in Section 2.12.1. The main advantage of the DFT-D corrections, is given by their low additional computational cost, due to a force-field like description. Within CP2K, three versions of the correction are available, namely D3(BJ), D3 and D2. Each of these has a different weighting of long and short-range contributions and therefore yields a specific description. Optimizations of the COF-LZU1's AA' configuration using the PBE/TZV2P-GTH description, showed possible differences arising due to the utilized dispersion corrections. The optimized structure using the D3(BJ) correction remains planar, whereas the D2 correction leads to 3 rotated phenyl groups as shown in Figure 10.8, indicating a strongly differing energetic landscape.

The resulting cell-parameters after optimization of the AA' unit cells for COF-LZU1 and N3-COF, summarized in Table 10.2 and Table 10.3, show the trends associated with the used correction types. The deviations arising for the in-plane lattice constants $a = b$ are below 1%

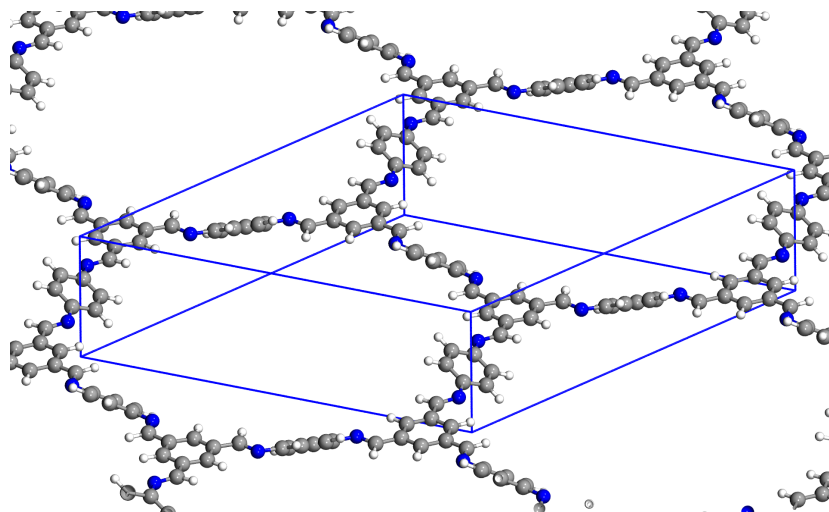


Figure 10.8: Cell-optimized structure showing the extended layers utilizing PBE/TZV2P-GTH with DFT-D2 correction.

Table 10.2: The lattice constants for the optimized COF-LZU1 AA' unit cells using the PBE functional corrected by the DFT-D series.

| Correction | $a = b$ in Å | c in Å |
|------------|--------------|----------|
| D2 | 22.35 | 7.56 |
| D3 | 22.47 | 7.38 |
| D3(BJ) | 22.45 | 7.25 |

and therefore insignificant, for both systems respectively.

In the case of COF-LZU1, the rotated phenyls of the D2 case, are associated with an increase in the c -lattice constant when compared to the previously used D3 standard. Contrarily, using the damped D3(BJ) description leaves the system planar while reducing the c -constant. Planarity of all optimized structures for the N3-COF reveals the direct influence of the chosen method onto the c -lattice constant. It can be concluded, that the D2 correction describes the $\pi - \pi$ interactions as less repulsive. This effect further indicates that rotated phenyl groups, as observed for COF-LZU1, potentially lead to larger c unit-cell vectors. Overall, the choice of dispersion correction influences the obtained interlayer distance up to 4.5 %.

Non-Local XC-Functionals

Alternatively, non-local functionals as conceptually described in Section 2.12 can be used to account for dispersion effects. In the CP2K program package the functionals DRSSL¹¹⁶,

Table 10.3: The lattice constants for the optimized N3-COF AA' unit cells using the PBE functional corrected by the DFT-D series.

| Correction | $a = b$ in Å | c in Å |
|------------|--------------|----------|
| D2 | 29.54 | 7.12 |
| D3 | 29.53 | 7.33 |
| D3(BJ) | 29.50 | 7.21 |

RVV10¹¹⁷ and LMKLL²⁴⁰ are available and the structural differences obtained for the COF-LZU1 are shown in Figure 10.9. Cell-optimizations are performed on the PBE/TZV2P-GTH level with corresponding PPs using standard SCF-parameters. The obtained structural parameters show mainly differences in the cell-vector c in stacking direction which is mostly depending on non-covalent interactions. Structures for COF-LZU1 showed rotated phenyl groups after optimization for vdW-DF1(DRSLL) and RVV10. In these cases, the lattice vector c varies from the planar case of the vdW-DF2(LMKLL) functional by being larger. For three rotated phenyl groups, this leads to a deviation of $\Delta c \approx 0.2\text{Å}$ whereas one rotated phenyl lead to a deviation of $\Delta c \approx 0.15\text{Å}$. Structures of the N3-COF remain planar during all optimizations, with the obtained in-plane cell vectors a and b slightly varying between the used functionals. In summary it can be seen from the obtained cell parameters, that non-local functionals have as much of an impact on the c -lattice constant as the choice of a different DFT-D correction or XC-functional. All three functionals used, describe the systems energy landscape differently and therefore lead to vastly differing descriptions. In the case of COF-LZU1, the out of plane rotation of the phenyl groups linking the edge is observed to various extend, with one group rotated for DRSLL Figure 10.9a, none in the case of LMKLL as shown by the side view in Figure 10.9b and three in the case of RVV10 as illustrated in Figure 10.9c. The resulting c -lattice parameters are larger the more rotated phenyls are present as can be seen from Table 10.4.

In the N3-COF case, planarity of all structures after optimization reveals the impact of dispersion correction onto the obtained c lattice constants. This can be seen from the data summarized in Table 10.5, revealing the significant of the used description type, onto this target property. The influence of the chosen NL functional on the obtained c -lattice constant, is shown to be below 6.5%

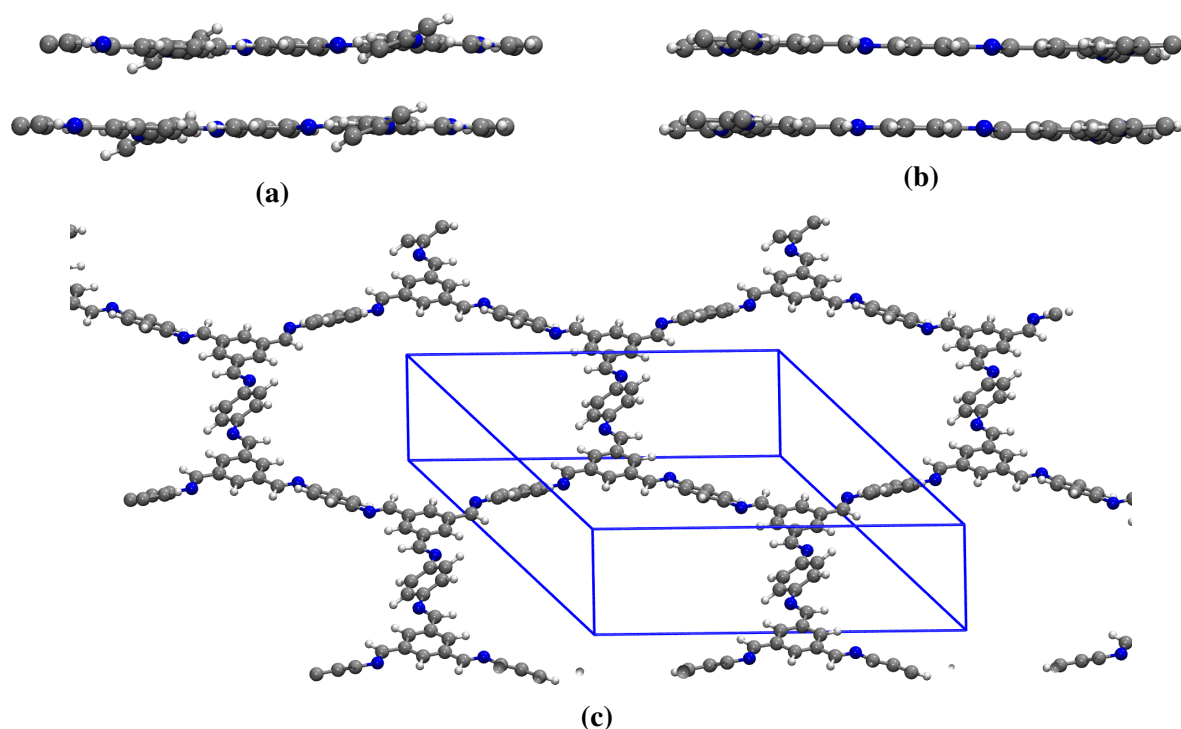


Figure 10.9: Structures after cell-optimizations using GTH-PPs and a TZV2P-GTH basis set. a) The side view after optimization with the vdW-DF1 (DRSLL) functional. b) The side view after optimization with the vdW-DF2 (LMKLL) functional, c) One optimized layer of the extended system using the RVV10 non-local functional, with the unit-cell boundaries indicated in blue. The second layer is excluded for reasons of visibility.

Table 10.4: The lattice constants for the optimized COF-LZU1 AA' unit cells using the PBE functional combined with non-local-vdW functionals.

| Non-Local Kernel | $a = b$ in Å | c in Å |
|------------------|--------------|----------|
| vdW-DF1 (DRSLL) | 22.52 | 7.52 |
| vdW-DF2 (LMKLL) | 22.57 | 7.36 |
| RVV10 | 22.41 | 7.59 |

Table 10.5: The lattice constants for the optimized N3-COF AA' unit cells using the PBE functional combined with non-local-vdW functionals.

| Non-Local Kernel | $a = b$ in Å | c in Å |
|------------------|--------------|----------|
| vdW-DF1 (DRSLL) | 29.57 | 7.52 |
| vdW-DF2 (LMKLL) | 29.62 | 7.30 |
| RVV10 | 29.51 | 7.04 |

10.4.5 Conclusion for the Standard Setup

From the previously conducted investigations it could be concluded, that a sufficiently accurate setup for planar COF-systems is provided by using the PBE functional alongside a TZV2P-GTH basis set with GTH-PBE-PPs, a PW-cutoff of $E_{\text{cut}} = 300$ Ry with $E_{\text{rel}} = 60$ Ry as well as Grimmes D3-correction. The choice of this setup is motivated solely by the necessity of reducing the computational effort as much as possible. Secondly, the same setup with the BLYP functional lead to better description for the N3-COF and is therefore used as the second choice for backup. A general trend for choosing the right functional is given by the overall impact of the repulsive interactions arising from aromatic rings as observed in the following study of planar systems.

For most systems, the TZV2PX-MOLOPT-GTH basis alongside a higher PW-cutoff will lead to a higher level of system description. However, the higher level of flexibility in the optimization of the orbitals does not systematically increase the reproduce ability of structural target parameters for comparison with the experiment. This is also the case for NL-functionals which are the superior choice for the description of non-covalent interactions. Consequently, a higher level of accuracy is only used when the standard description does not reproduce the structural target properties with sufficient accuracy. An exception is the case of high-accuracy calculations for obtaining the electron density ρ_{el} for the calculation of atomistic properties, such as partial charges q .

10.5 Investigations using GFN-xTB

The CCM implementation of GFN-xTB within the CP2K program package offers a promising alternative for two-dimensional COF systems as its parameterization is optimized for non-covalent interactions within organic molecules⁶¹(compare Section 4.3). Furthermore, avoidance of explicit integrations due to the high level of parameterization allows the treatment of large systems. Comparison of the GFN-xTB description with GPW-DFT is conducted using the cell-optimized structures for COF-LZU1 and the N3-COF system. Optimizations are performed for the AA and AA' configurations, with the AA' initial structure created by displacing the upper layer by 1.0 Å in x and y direction from the AA configuration.

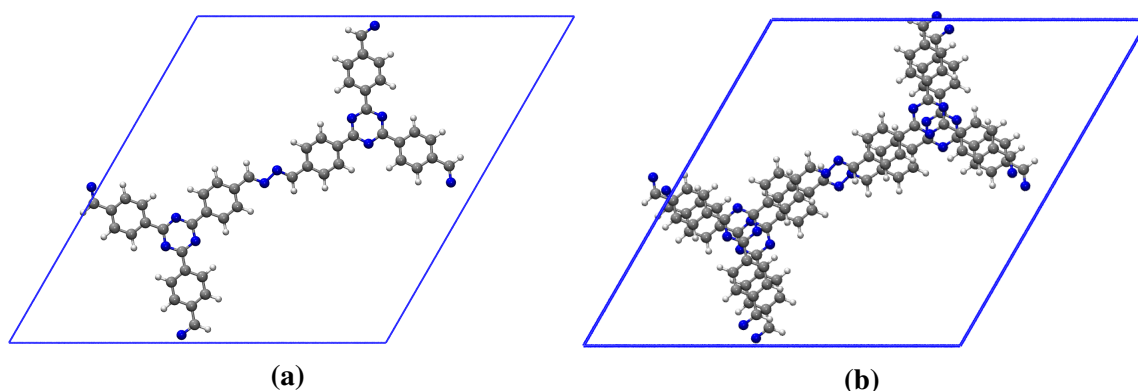


Figure 10.10: Unit cells of the N3-COF optimized using GFN-xTB. a) The case of AA stacking leading to $a = b \approx 29.11 \text{ \AA}$ and $c \approx 6.97 \text{ \AA}$, b) The case of AA' stacking $a = b \approx 29.13 \text{ \AA}$ and $c \approx 6.46 \text{ \AA}$.

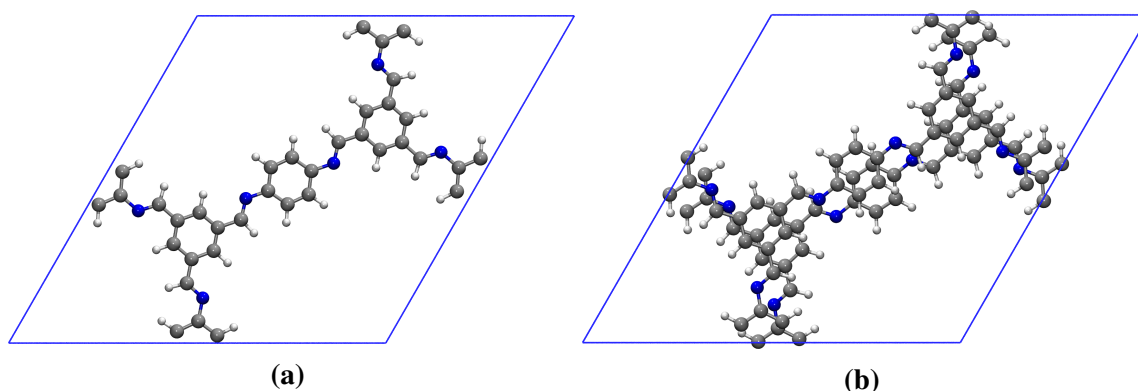


Figure 10.11: Side view onto the cells of the COF-LZU1 optimized using GFN-xTB. a) The case of AA stacking leading to $a = b \approx 22.23 \text{ \AA}$ and $c \approx 7.06 \text{ \AA}$, b) The case of AA' stacking $a = b \approx 22.23 \text{ \AA}$ and $c \approx 6.51 \text{ \AA}$.

Comparison of the obtained arrangements after cell optimization indicates significantly more repulsion between the layers in the AA-structures for both systems. For the N3-COF the AA' arrangement is energetically more favorable by $\Delta E \approx 11.17 \text{ kcal/mol}$ compared to the AA case. Similarly for COF-LZU1, the AA' arrangement is energetically more favorable by $\Delta E \approx 9.06 \text{ kcal/mol}$. Comparison with the calculated GPW-DFT data on the standard level (compare Section 10.4.5) summarized in Table 10.6 and Table 10.9, shows smaller energetic differences between the arrangements. The obtained cell-parameters for both structural arrangements show large deviations between the methods for both COF systems. The deviations in the cell-parameters for the AA configurations are determined as $\Delta a = b \approx 0.25 \text{ \AA}$ and $\Delta c \approx 1.05 \text{ \AA}$, for the COF-LZU1 system. In the case of the N3-COF, the deviations between the methods for the

10 Theoretical Analysis for Covalent Organic Frameworks

obtained lattice parameters are determined as $\Delta a = b \approx 0.34 \text{ \AA}$ and $\Delta c \approx 1.1 \text{ \AA}$. For the AA' N3-COF system, deviations in the obtained lattice constants of $\Delta a = b \approx 0.4 \text{ \AA}$ and $\Delta c \approx 0.7 \text{ \AA}$ arise. Similarly, for the COF-LZU1 system, they are determined as $\Delta a = b \approx 0.23 \text{ \AA}$ and $\Delta c \approx 0.77 \text{ \AA}$. For all systems and configurations, lattice parameters obtained for GFN-xTB optimized unit cells are smaller than in the corresponding GPW-DFT case. Even though the obtained cell parameters deviate strongly, the obtained energetic differences show qualitatively similar trends. Overall it can be concluded, that GFN-xTB describes the interlayer interactions much less repulsive than the used GPW-DFT standard. The differences might partially result from GGA-functionals describing π - π interactions generally as too repulsive^{158,232,233,238}. However, both methods used apply the D3-dispersion correction and consequently the obtained energy profiles are qualitatively similar, as differences predominantly arise due to non-covalent interactions. Technically the observations made can be attributed to the electrostatics with GFN-xTB approximating the DFT treatment by a monopole expansion as described in Section 4.2.3 and identical dispersion treatment Section 4.3. The observed trends are further investigated in the following by sampling the displacements within the unit cell for both methods.

10.5.1 Interpolation of the Potential Energy Surface

Insight into the thermal stability of COFs is gained by calculating the energies of two-layered systems, modeling various displacements d . These are created by shifting the center of mass of the upper layer by a displacement d within the x - y -plane of the unit cell. For COF systems with hexagonal symmetry (D_6) sampling of the lower triangle spanning only a quarter of the full unit cell is sufficient. For each point on the grid, the corresponding energy $E(x, y)$ was obtained by simultaneous geometry and cell-optimization using constraint hexagonal symmetry for the lattice vectors. Energies are calculated using CCM implemented GFN-xTB within the CP2K program package. Optimizations are performed using the conjugated gradients (CG) method and considered converged when the total forces acting on the molecules are smaller than $4.5 \cdot 10^{-4}$ Hartree/Bohr with the relative change in geometry being smaller than $3.0 \cdot 10^{-3}$ Bohr and the pressure tolerance with respect to the environment below 100bar was reached.

Initially, both layers are displaced such that the x - y -coordinates of their center of masses (COMs) are located at the origin of the elementary real space unit cell. The upper layer is

then displaced by multiples of the cell vectors in x and y direction with the sampling points chosen as listed in Table F.8. The generated grid is generated under the assumption of D_{6h} -symmetric COF molecules which allows the use of the corresponding in-plane D_6 symmetry (full hexagonal), significantly reducing the necessary sampling area. Consequently, the lower triangle of the unit cell as shown in Figure 10.12 is sufficient to estimate the full PES. The obtained energy values E for the optimized unit cells associated with their initial in-plane displacements x and y , are expanded in a PW basis-set as

$$E(x, y) = \sum_n c_n \phi_n(x, y) \quad , \quad (10.2)$$

finding the optimal coefficients c_n in a least square manner using the singular value decomposition. The used PW-basis functions are chosen to fulfill hexagonal symmetry and given as

$$\phi_n(x, y) = \cos(n\alpha \cdot x) + \cos\left(\frac{n\alpha}{2}(x + \sqrt{3}y)\right) + \cos\left(\frac{n\alpha}{2}(-x + \sqrt{3}y)\right) \quad , \quad (10.3)$$

with $\alpha = \frac{4\pi}{\sqrt{3}}$, n the index of the used basis function and x, y the in-plane spacial position. Alternatively, the used basis set can be written in terms of polar coordinates with the two-dimensional position vector \mathbf{r} and the angle φ between the components of the wave vector $\mathbf{k}_n = n_1 \mathbf{b}_1 + n_2 \mathbf{b}_2$ with $n_1 + n_2 \leq n$. The basis functions $\phi_n(\varphi)$ can then be expressed by using the scalar product of \mathbf{k}_n and \mathbf{r} , leading to

$$\phi_n(\varphi) = \sum_{i=0}^6 \cos(\mathbf{k}_{n,i} \cdot \mathbf{r}) \quad , \quad (10.4)$$

with the $\mathbf{k}_{n,i}$ given as

$$\mathbf{k}_{n,i} = k_n \cdot \left(\begin{array}{c} \cos\left(\alpha \cdot \left(\varphi + \frac{\pi}{3}(i-1)\right)\right) \\ \sin\left(\alpha \cdot \left(\varphi + \frac{\pi}{3}(i-1)\right)\right) \end{array} \right) \quad , \quad (10.5)$$

fulfilling the assumed hexagonal symmetry. The PW interpolated PES is constructed by calculating energies for the expanded $E(x, y)$ on a grid of 250×250 data points in the x - y plane, spanning the area of 2×2 unit cells. The origin (0,0) of the elementary unit cell given in

terms of the basis vectors as $x = \frac{1}{3}a$ and $y = \frac{1}{2}b$, is placed in the middle of the sampling area.

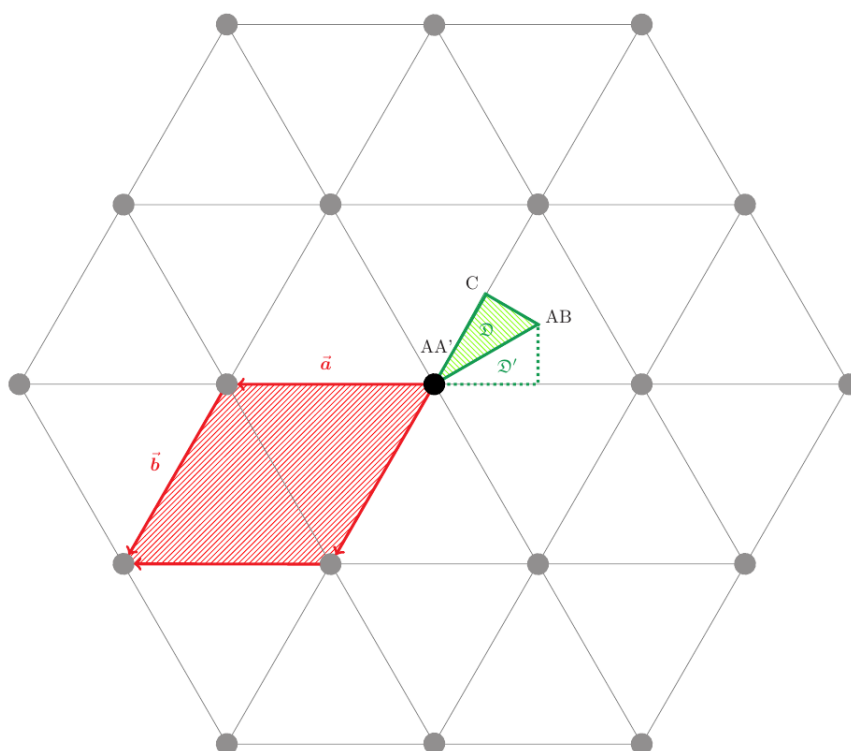


Figure 10.12: Schematic representation of the assumptions entering in the PW fit of the energy landscape. The treatment of the hexagonal unit cell shown in red with cell vectors \mathbf{a} and \mathbf{b} was further simplified by using the properties of the D_6 symmetry, which leaves only the green triangle \mathfrak{D} , spanning only 1/12 of the unit cell, for calculation. The remaining part of the unit cell is constructed by mirroring the interpolated triangle \mathfrak{D} as indicated by \mathfrak{D}' and rotation by multiples of $\frac{\pi}{3}$.

An example investigation was conducted for COF-LZU1, where two grids of energy values are calculated comparing the GFN-xTB with the DFT obtained energy landscape as shown in Figure 10.13. The obtained PES entering the interpolation obtained from GPW-DFT on the standard level are shown in Figure 10.13a, whereas the results obtained via GFN-xTB are shown in Figure 10.13b. Comparison of both energy profiles shows differences in the magnitude of energy shifts but are qualitatively similar. The determined initial displacements resulting in the energetic minima and maxima are identical, however, comparison of the absolute range of energetic differences shows both methods to differ quite significantly. This can be attributed to the parametrization of GFN-xTB, which was not designed specifically to

reproduce energetic differences. Consequently in Figure 10.13a, an additional barrier indicated in purple can be observed around the AA' stacked position. This indicates clearly a local minimum when compared to the almost flat area in Figure 10.13b. Both methods describe the main trends similarly, qualifying GFN-xTB usage for the unit cell sampling. However, for possible interpretations, the flattening of the obtained PES in the GFN-xTB case has to be kept in mind. As the by far cheaper method, GFN-xTB provides the possibility of sampling the PES for shifts within the unit cell, especially useful when treating larger and more complex systems, which disqualify GPW-DFT due to its high computational effort.

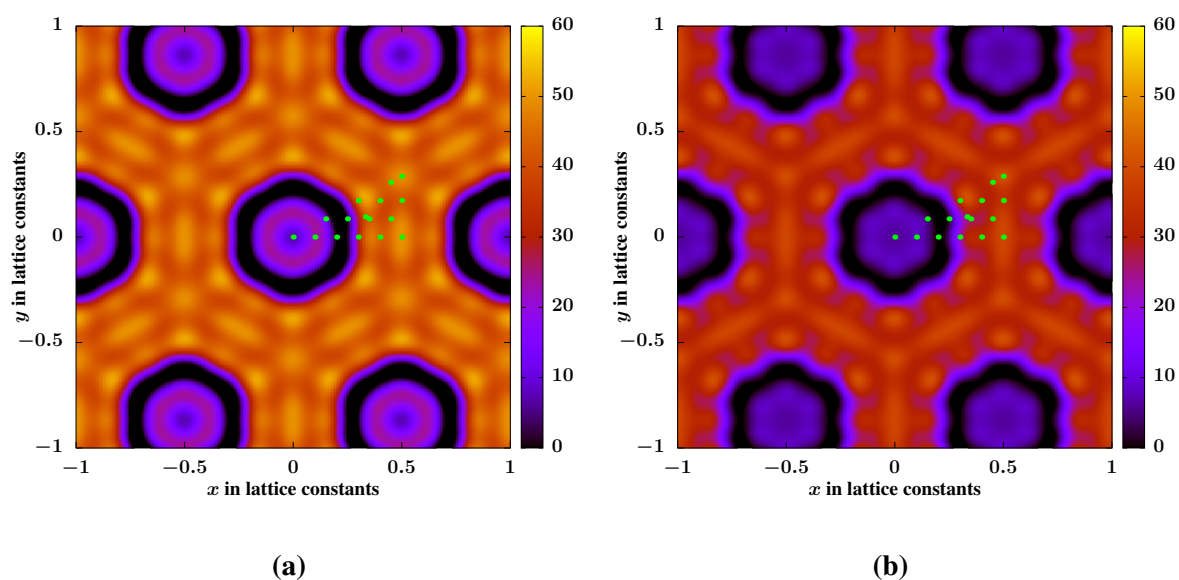


Figure 10.13: Investigations for interlayer displacements of COF-LZU1 following the plane wave interpolation scheme of the obtained values for $E(x,y)$ in Eq (10.2) obtained from both theoretical methods. a) The PES as interpolated from GPW-DFT reference data. b) The PES interpolated from CCM GFN-xTB reference data.

10.6 Overview of the Investigated Species

In the following if not stated otherwise, the COF systems have been investigated for their structural base properties using the standard setup Section 10.4.5. An overview of the investigated COFs is provided by a characterization aiming at the main structural parameters for each species respectively. Here, the treated unit cells are investigated for their symmetry by means of the CS algorithm (compare Chapter 6) showing the number of equivalent atomic environments. For COF-LZU1 as well as TpPa-1-COF, these are then used for charge equilibration over equivalent environments. Furthermore, high symmetry stacking schemes are analyzed by means of GPW-DFT, showing the impact of the theoretical description as well as the used building units. Here, the four distinct types are analyzed namely AA, AA', AB and C type stacking for all COFs. The AA stacking refers to the eclipsed case with no in-plane displacement between adjacent layers. AA' on the other hand refers to initial structures created by an in-plane displacement of 1.0 Å in x and y direction respectively in order to break the symmetry. The case of AB stacking, refers to structures created when translating one layer by one third of the cell vectors (0.333,0.333) as illustrated in Figure 10.12. This type of displacement leads to the alternating overlap of the pore and linker centers resulting in the so called staggered configuration with heavily reduced pore size. C-type stacking refers to the case of displacing one layer half the unit cell along the direction of lattice vector a and therefore is given as the lower corner of the triangle indicated in Figure 10.12. The minima obtained from analysis of the resulting PES are then used for further analysis, as they provide an estimate of the most probable arrangements. Even though energetic differences of geometries are not a target property of the GFN-xTB parameterization, the obtained energy landscape acts as a qualitative estimate as shown in Section 10.5. Due to limited computational effort, the obtained PES can be calculated even for the largest investigated COFs in reasonable time. Furthermore the pore sizes \mathfrak{P} are calculated using the PSD (compare Chapter 7) for various displacement paths generated by means of FRDG (compare Chapter 5). The most probable interlayer shifting for various displacements d is determined based on a comparison of the pore size $\mathfrak{P}(d)$ with the experimentally obtained value for the literature known COFs. Finally, the best description determined by the used methodology is suggested in terms of interlayer displacement as well as GPW-DFT setup.

10.6.1 Planar Species

Investigation of planar COF species build from various standard components reveals insights into the impact of the used description. Here, the suitability of theoretical description due to the structural specifics of the COF species are revealed. The accuracy and limits of applicability are the main target of the conducted study. Comparison with the experiment allows for determination of the most suitable structural model as well as to explore the boundaries for possible interpretation. Extrapolation of the investigated effects allows for the suggestion of the most suitable description for not yet investigated COFs comprised of similar components.

Characteristics: COF LZU-1

The COF-LZU1 is an azine based framework and was the first two dimensional imine linked COF system synthesized by Ding et al. in 2011²⁴¹. Its minimalistic nature simplifies theoretical studies, due to its unit cell consisting of two identical building blocks directly connected with each other by imine bonds as can be seen from its unit cell in Figure 10.14a. The initial crystal structure used for the following calculations was extracted from X-ray powder diffraction data using materials studio²⁴². The *c*-cell parameter of the initial structures was set in reference to the experimentally determined interlayer distance of $r_{\text{inter}} = 3.73 \text{ \AA}$ ²⁴¹.

Investigations of the COFs' stacking behavior using cell optimizations on the standard GPW-DFT level, revealed the slightly shifted AA' stacking to be the energetically most favorable of the investigated high symmetry stackings summarized in Table 10.6. The perfectly eclipsed AA case is determined to be energetically disfavored by $\Delta E \approx 15.32 \text{ kcal/mol}$. A reason is found in the increase of repulsive electrostatic interactions due to identical atoms of adjacent layers directly positioned on top of each other due to suppression of in-plane displacements $d = 0.0 \text{ \AA}$. From comparison with the experiment the obtained deviation in the cell parameters from the literature value shows the AA' stacking to be off by $\Delta c \approx 0.2 \text{ \AA}$, which is also the case for AB-stacking. Calculation of the energy for the optimized configuration shows AB stacking to be energetically disfavored by around $\Delta E \approx 32.0 \text{ kcal/mol}$, indicating a reduced stabilization due to less $\pi - \pi$ overlap regions as exemplary shown in Figure 10.8. The same holds for C-type stacking, indicating largely displaced arrangements to be disfavored even if they fulfill high symmetry properties. From the obtained information the most probable

structure can be estimated to be found for smaller displacements d in the vicinity around the AA configuration. The pore sizes $\mathfrak{P}(d)$ were determined using the PSD workflow as described in Chapter 7 for systems created by the FRDG algorithm as described in Chapter 5 consisting of 40 randomly stacked layers with an assumed in-plane displacements in the range of $d = 0.0 - 5.0 \text{ \AA}$ in intervals of 0.5 \AA . The obtained pore size with respect to the used displacement is illustrated in Figure 10.15a, with the value corresponding to a displacement of $d = 1.5 \text{ \AA}$ showing the best agreement with the experimentally obtained pore size distribution, showing a clear maximum around 1.2 nm.

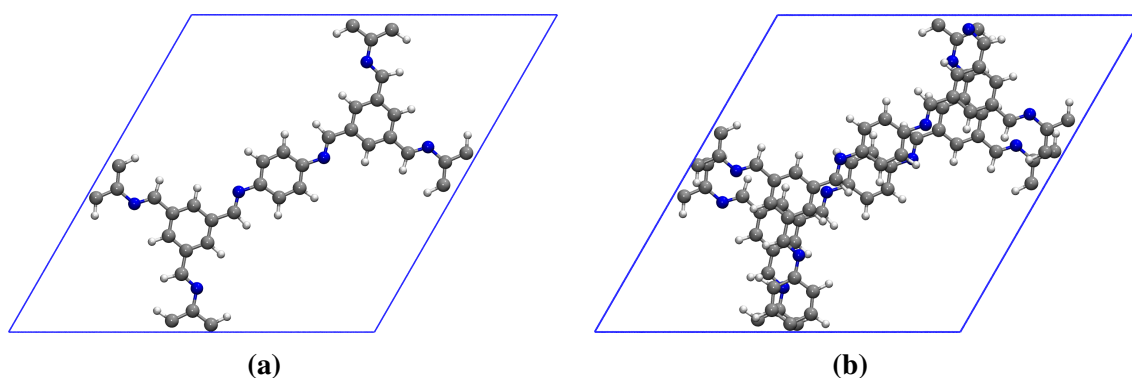


Figure 10.14: Top view of the COF-LZU1 unit cells after cell-optimization on the GPW-DFT level using the standard setup. a) The case of AA stacking. b) The case of AA' stacking, showing the structures to remain planar after optimization.

Table 10.6: Lattice constants and relative energies with respect to the lowest value for the optimized COF-LZU1 unit cells. Calculations were performed on the TZV2P-GTH/PBE-D3 level using GTH-PBE-PPs.

| Stacking | $a = b$ in \AA | c in \AA | ΔE in kcal/mol |
|----------|-------------------------|---------------------|------------------------|
| AA | 22.47 | 8.10 | 15.32 |
| AA' | 22.47 | 7.27 | 0 |
| AB | 22.47 | 7.29 | 32.00 |
| C | 22.47 | 7.65 | 32.28 |

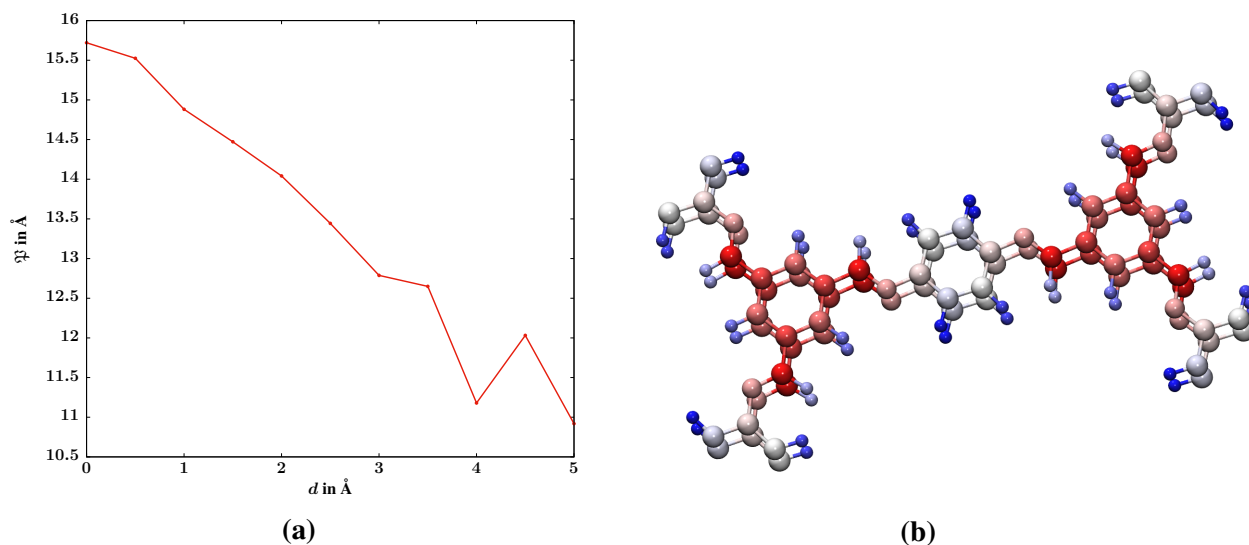


Figure 10.15: Extended analysis of the COFs fundamental structural properties. a) Average pore size $\mathfrak{P}(d)$ obtained using the PSD for FRDG-generated random paths for an in-plane displacement d . The random paths were constructed using 40 layers with interlayer distances in stacking direction of $c = 3.64$ Å as obtained from the optimized AA' unit cell on the GPW-DFT standard level. b) Symmetry analysis with the CS-algorithm conducted for the COF-LZU1 unit cell, resulting in 12 equivalence classes of 6 atoms each, differentiated by color.

The structural analysis of COF-LZU1 shows the eclipsed AA' stacking type unit cell to be the best representation of the average COF-crystal within the chosen level of description. The experimentally determined pore size of 1.2 nm was determined using the PSD algorithm for a random displacement of $d = 3.0$ Å. Randomly generated interlayer displacements shows great impact on the effective pore size \mathfrak{P} as determined by GCMC simulations of adsorption at the example of COF-LZU1¹⁶². The authors showed that an interlayer displacement within the range of $d = 3.0 - 4.5$ Å yields the best agreement with the experimentally obtained value. An interlayer displacement of $d = 3.0$ Å was identified as the closest to the experimentally obtained pore size distribution. The interlayer displacement map shown in Figure 10.13b shows the global energetic minimum for an interlayer displacement of $d \approx 4.5$ Å. This can be thought of as an upper limit of the possible displacement stabilization as it might occur in the real structure. Another side minimum is found for the AA' case located at (0,0) in the displacement map. Considering all indications observed from the basic analysis, a slight interlayer slipping of adjacent layers comprising the pore-tubes is identified as the most probable.

Polarization by Adsorbed Molecules

Further analysis for the systems charge distribution by symmetry considerations using the CS algorithm as described in Chapter 6, reveals the different classes of atoms within the treated unit cell. The obtained discrimination of atoms in symmetry classes is shown in Figure 10.15b and allows the identification of the minimum set of atomic environments. During the adsorption of molecules into the COF pores as e.g. in the case of adsorption measurements for pore size analysis of synthesized COF systems, the guest molecules interact with the framework giving rise to potential charge shifts. To obtain insights on the strength of the resulting framework polarization, the charge distribution for various loadings and types of adsorbed molecules are analyzed. To this end, systems are extracted from GCMC simulations of adsorption isotherms¹⁶². Atomic DDEC6 partial charges²⁴³⁻²⁴⁶ were calculated by partitioning of the electronic density obtained for the optimized AA' unit cells. The reference density is calculated for constructed 2×2 super cells by single point energy calculations using GPW-DFT. For securing higher accuracy, the setup is modified by using $E_{\text{cut}} = 500$ Ry, EPS_DEFAULT = 10^{-14} and EPS_SCF = 10^{-7} . The obtained DDEC6 partial charges are consecutively averaged over the atomic environments illustrated in Figure 10.15b using the CS algorithm. The empty reference framework is then filled with N₂ and CO₂ adsorbed molecules. For various loadings extracted at different pressures simulated in the GCMC-adsorption isotherms, the charge distribution is recalculated and the partial charges \bar{q} are averaged over each equivalence class \mathfrak{F} respectively. The data in the case of CO₂ adsorption is summarized in Table 10.7, showing only minor deviations depending on the pressure P . Similar results are obtained for N₂ adsorption with the data summarized in Table 10.8. In summary, the adsorption of gas molecules showed only small polarization for the frameworks' atoms for the COF-LZU1 super cells. The obtained redistribution of charges was too small to cause any significant impact on the adsorption behavior modeled by GCMC simulations. Consequently, the use of polarizable forcefields for simulating the adsorption process, can be neglected for COF-LZU1. An explanation is found by the neutrally charged species comprising the walls of the pore with electrons participating in covalent bonding. Therefore, the frameworks' insignificant polarization by guest molecules is possibly attributed to the limited flexibility of charge reorganization.

Table 10.7: Calculated DDEC6 net partial charges $\bar{q}(P)$ averaged over each identified equivalence class \mathfrak{F} (the element of is indicated in brackets), for CO₂ loaded framework super-cells for the COF-LZU1 framework as extracted from GCMC calculated isotherms at $T = 183.15$ K in for different values of the environment pressure P in bar.

| \mathfrak{F} | $\bar{q}(0)$ | $\bar{q}(1000)$ | $\bar{q}(15000)$ | $\bar{q}(100000)$ |
|----------------|--------------|-----------------|------------------|-------------------|
| N-1 | -0.299 | -0.299 | -0.299 | -0.297 |
| C-1 | -0.160 | -0.159 | -0.159 | -0.157 |
| C-2 | -0.131 | -0.131 | -0.131 | -0.131 |
| C-3 | -0.131 | -0.130 | -0.130 | -0.131 |
| C-4 | 0.066 | 0.066 | 0.066 | 0.068 |
| C-5 | 0.084 | 0.084 | 0.085 | 0.087 |
| C-6 | 0.212 | 0.213 | 0.213 | 0.216 |
| H-1 | 0.052 | 0.052 | 0.052 | 0.055 |
| H-2 | 0.100 | 0.099 | 0.099 | 0.095 |
| H-3 | 0.100 | 0.100 | 0.100 | 0.097 |
| H-4 | 0.106 | 0.105 | 0.104 | 0.099 |

Table 10.8: Calculated DDEC6 net partial charges $\bar{q}(P)$ averaged over each identified equivalence class \mathfrak{F} (the element of is indicated in brackets), for N₂ loaded framework super-cells for the COF-LZU1 framework as extracted from GCMC calculated isotherms at $T = 77.36$ K in for different values of the environment pressure P in bar.

| \mathfrak{F} | $\bar{q}(0)$ | $\bar{q}(300)$ | $\bar{q}(1000)$ | $\bar{q}(10000)$ |
|----------------|--------------|----------------|-----------------|------------------|
| N-1 | -0.299 | -0.296 | -0.294 | -0.294 |
| C-1 | -0.160 | -0.155 | -0.151 | -0.156 |
| C-2 | -0.131 | -0.131 | -0.131 | -0.130 |
| C-3 | -0.131 | -0.129 | -0.127 | -0.129 |
| C-4 | 0.052 | 0.216 | 0.222 | 0.218 |
| C-5 | 0.066 | 0.069 | 0.071 | 0.069 |
| C-6 | 0.084 | 0.087 | 0.088 | 0.088 |
| H-1 | 0.100 | 0.049 | 0.047 | 0.051 |
| H-2 | 0.100 | 0.091 | 0.086 | 0.093 |
| H-3 | 0.106 | 0.098 | 0.094 | 0.093 |
| H-4 | 0.212 | 0.098 | 0.096 | 0.097 |

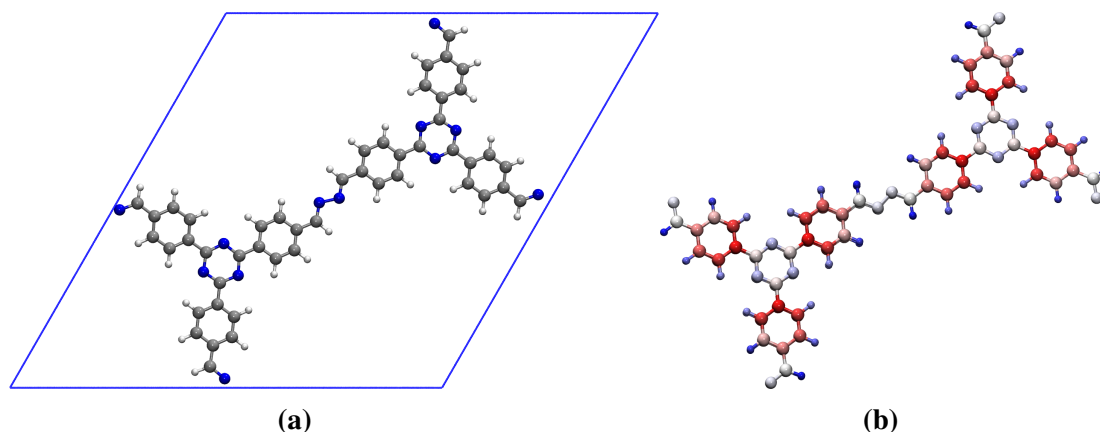
Characteristics: N3-COF

Figure 10.16: Overview of the cell optimized N3-COF unit cell. a) Top view onto the perfectly stacked AA structure after optimization. b) The N3-COF molecular structure after treatment with the CS algorithm, showing the equivalent atomic environments resulting in thirteen equivalence classes of six atoms each, differentiated by color.

The N3-COF (COF N3) is the final stage of an azine based COF platform first investigated by Vyas et al. in 2015²⁴⁷. The basic unit is a C3-linker shown in Figure 10.17a modified by introduction of nitrogens to the central ring in order to eliminate rotation of the phenyl groups, resulting in a completely planar COF system. Rotated phenyls can be systematically suppressed by introducing nitrogen atoms on the central ring as illustrated by the comparison of both cases in Figure 10.17. With three nitrogens at the aryl ring, the resulting linker configuration is perfectly flat as shown in Figure 10.17b, resulting in a significantly larger building unit that can be linked via imine bridges as shown for the unit cell in Figure 10.16a. As a result, the electrostatics of this framework are expected to impact the stabilization of the framework more significantly. The investigated COF-System was constructed based on crystal structures extracted from X-ray powder diffraction data using materials studio²⁴². An overview of the general stacking behavior of the N3-COF by means of GPW-DFT is summarized in Table 10.9. The extreme cases of AB and C type stackings are energetically much more disfavored than the slightly shifted AA' one, showing the preference of slight displacements. Perfect stacking is even more disfavored for the N3-COF than for COF-LZU1 with $\Delta E \approx 21.5$ kcal/mol. The analysis of general displacement configurations shown in Figure 10.18b indicates that slightly smaller displacements are preferable. The obtained unit

cell parameter in stacking direction using the PBE functional $r_{\text{inter,PBE}} = 3.65 \text{ \AA}$ deviates from the experimentally obtained one by $\Delta c = 0.2 \text{ \AA}$. Contrarily, the value obtained for AB stacking seems to clearly fit the experimental observation. However, the AA' configuration is energetically favored by $\Delta E \approx 35.75 \text{ kcal/mol}$ making it the by large preferable stacking scheme. Further analysis of the displacements shows the region of stable structures close to the eclipsed scheme as can be seen from Figure 10.18b. This consequently indicates the possibility of the used method being unsuited to reproduce the structural behaviors. Switching from the PBE to the BLYP functional leads to perfect agreement of interlayer distances for the investigated case as can be seen from the obtained $r_{\text{inter,BLYP}} = 3.46 \text{ \AA}$ compared to the experimental one of $r_{\text{inter,exp}} = 3.43 \text{ \AA}$. Using the CS algorithm leads to identification of 13

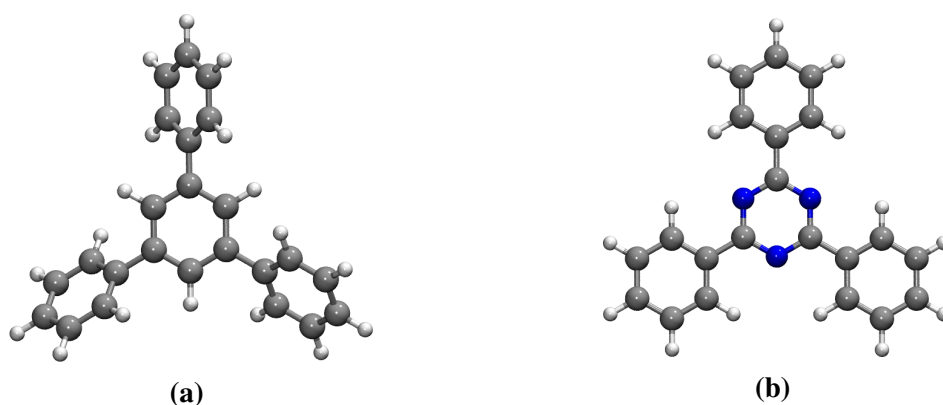


Figure 10.17: Comparison of the C3-linker with its modified version consisting of an aryl ring. a) The C3-linker configuration with phenyl groups attached to the central ring at 1.3.5 position. b) The N3-linker configuration with the central aryl ring introducing 2.4.6 nitrogens.

Table 10.9: Lattice constants and relative energies with respect to the lowest value for the optimized N3-COF unit cells. Calculations were performed on the TZV2P-GTH/PBE-D3 level using GTH-PBE-PPs.

| Stacking | $a = b$ in \AA | c in \AA | ΔE in kcal/mol |
|----------|-------------------------|---------------------|------------------------|
| AA | 29.47 | 8.06 | 21.54 |
| AA' | 29.47 | 7.30 | 0 |
| AB | 29.46 | 6.69 | 35.75 |
| C | 29.41 | 6.59 | 37.50 |

distinct atomic environments comprising this COF. From the displacement map it is visible that

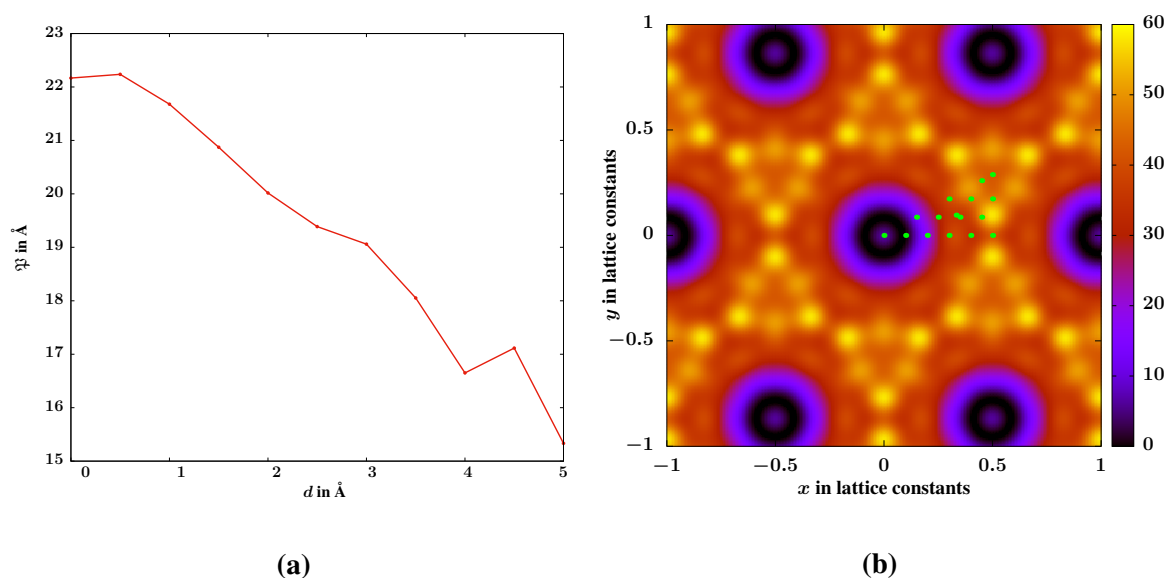


Figure 10.18: Analysis of the interlayer displacement behavior for the N3-COF system. a) Average pore size $\mathfrak{P}(d)$ determined using the PSD for FRDG-generated randompaths using various in-plane displacements d . The randompaths were constructed using the FRDG algorithm for 40 layers with interlayer distances of $c = 3.42 \text{ \AA}$ as obtained from the optimized AA' unit cell. b) The PES determined following Section 10.5.1, the energies corresponding to the respective colors are given in kcal/mol.

interlayer slippings around 0.1 unit cell lengths corresponding to $d \approx 3.5 \text{ \AA}$. Interlayer slippings investigated for their effective pore size show the best agreement with the experimentally obtained value of 2.6 nm for the N3-COF system. Therefore, larger displacements are not only energetically disfavored in the theoretical description, but also show larger discrepancies to the experimentally obtained pore size. Consequently, slight displacements up to $d \approx 3.0 \text{ \AA}$ are expected for this COF system based on the theoretical conducted study. Compared to the COF-LZU1 system, the displacement is less compared to the pore size and thus indicates a structural stabilization due to the aryl compounds.

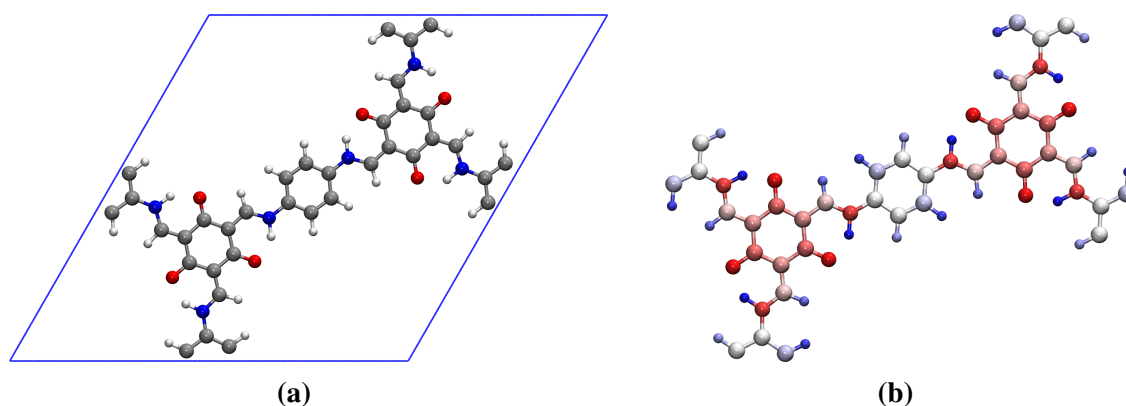
Characteristics: COF-TpPa-1

Figure 10.19: Overview of the cell optimized TpPa-COF unit cell for structural analysis. a) The unit cell after optimization of the AA stacking with the in-plane cell boundaries indicated in blue. b) Symmetry analysis with the CS-algorithm conducted for the periodic TpPa-COF unit cell, resulting in twelve equivalence classes consisting of six atoms each, differentiated by color.

The TpPa-1 COF was first synthesized by Kandambeth in 2012²⁴⁸ in order to show its chemical stability. The basic structural features are shown in Figure 10.19 with the unit cell of the perfectly eclipsed AA stacking shown in Figure 10.19a. The framework is structurally a modified version of COF-LZU-1, constructed by introduction of oxygen atoms to the central phenyl ring in the of bonding positions. The COFs unit cell consists of twelve distinct atomic environments with six atoms each, showing a high level of symmetry in this system illustrated by colors in Figure 10.19b. Again the initial unit-cell structures for one layer are extracted from experimentally obtained X-ray powder diffractograms using materials studio²⁴². Optimization of the resulting system showed the COF-layers to be completely planar in all investigated arrangements. The obtained unit-cell parameters are similar to the case of COF-LZU1 with the largest difference being the experimentally reported interlayer distance of $r_{\text{inter,exp}} = 3.34 \text{ \AA}$. Calculations on the PBE-D3 level showed good agreement with the obtained interlayer distance of $r_{\text{inter,PBE}} \approx 3.43 \text{ \AA}$ and are identical when compared to calculations with the BLYP functional $r_{\text{inter,BLYP}} \approx 3.44 \text{ \AA}$. The introduction of additional oxygen atoms to the linker unit leads to an interesting displacement configuration (0.2,0.1). The energetic minimum of the investigated stackings on GPW-DFT level is again found for the slightly shifted AA' initial structure as shown in Table 10.10. The AB and C configurations are disfavored compared to the AA' case

by around $\Delta E \approx 40$ kcal/mol, which is even higher than for the previously investigated species. An interesting additional feature is observed for the perfectly eclipsed AA stacking which is additionally disfavored by $\Delta E = 10$ kcal/mol when compared to the unmodified COF-LZU1. This energetic shift can be solely attributed to interactions of the oxygen atoms as they are the only structural difference between the two COF-species. Interactions between the oxygen atoms, seem to influence the electrostatics between the layers in the same magnitude than $\pi - \pi$ interactions. Interestingly, the AA structure is even more disfavored when using BLYP optimized unit cells which was not observed in the COF-LZU1 system, clearly indicating the importance of the used description with respect to the oxygen influence. A possible explanation of this effect is found due to the description of the $\pi - \pi$ stackings as discussed in Section 10.4.1 which is described less repulsive in the BLYP case. Consequently it can be concluded that the energetic displacement landscape for TpPa-1-COF is dominated by the electrostatic interactions of the oxygen atoms. Interestingly, the obtained lattice parameters for the shifted structure are identical for both functionals as can be concluded from a comparison with Table 10.11.

Table 10.10: Lattice constants and relative energies with respect to the lowest value for the optimized TpPa-COF unit cells. Calculations were performed on the TZV2P-GTH/PBE-D3 level using GTH-PBE-PPs.

| Stacking | $a = b$ in Å | c in Å | ΔE in kcal/mol |
|----------|--------------|----------|------------------------|
| AA | 22.58 | 8.01 | 25.54 |
| AA' | 22.59 | 6.85 | 0.00 |
| AB | 22.57 | 7.02 | 38.60 |
| C | 22.58 | 7.61 | 42.90 |

Table 10.11: Lattice constants and relative energies with respect to the lowest value for the optimized TpPa-COF unit cells. Calculations were performed on the TZV2P-GTH/BLYP-D3 level using GTH-BLYP-PPs.

| Stacking | $a = b$ in Å | c in Å | ΔE in kcal/mol |
|----------|--------------|----------|------------------------|
| AA | 22.71 | 7.77 | 34.62 |
| AA' | 22.70 | 6.87 | 0.00 |
| AB | 22.71 | 6.89 | 40.64 |
| C | 22.71 | 7.42 | 49.17 |

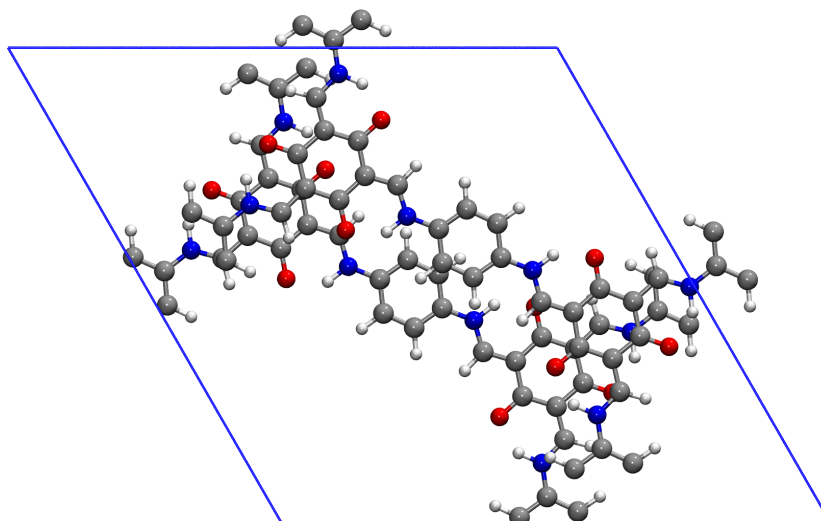


Figure 10.20: The optimized unit-cell structure of TpPa-COF, for the initially AA' structure showing the overlap of oxygen atom and phenyl rings. The same configuration is obtained when (0.2,0.1) is chosen as initial displacement.

Screening the energetics of various displacements, lead to the observation of two distinct minima in the potential energy landscape illustrated in Figure 10.21b. A minimum was observed for the eclipsed AA' structure as well as for the (0.2,0.1) shifted structure which both resulted in one oxygen atom of the displaced layer directly located in the center of an adjacent ring. The structure after cell optimization on the GFN-xTB level as used for the displacement screening is shown in Figure 10.20. The obtained interlayer distances of the AA' structures were used for the generation of random paths to obtain further insight in the displacement behavior of the TpPa-1-COF system. The pore sizes \mathfrak{P} for various random displacements, showed best agreement with the experimentally obtained value of 1.25 nm for displacements in the range of $d = 2.5 - 3.0 \text{ \AA}$ as can be concluded from Figure 10.21a. Comparison of the obtained displacement range with the observed potential energy landscape shows the second minimum for alternating oxygen-phenyl stacking for the optimized structure for (0.2,0.1) displacements corresponding to a total initial displacement of $d \approx 5.0 \text{ \AA}$. However, the final structure is only shifted by $d = 2.0 \text{ \AA}$, as it closely resembles the result of the optimized AA' unit cell. The fact that such a large interlayer shift occurs during the optimization indicates the resulting minimum to be dominant in the near vicinity. Further investigation of the resulting structure reveals the special arrangement of the optimized structure Figure 10.20. Calculation of \mathfrak{P} using the vdW-radii of hydrogen as minimum distance from pore wall atoms is not

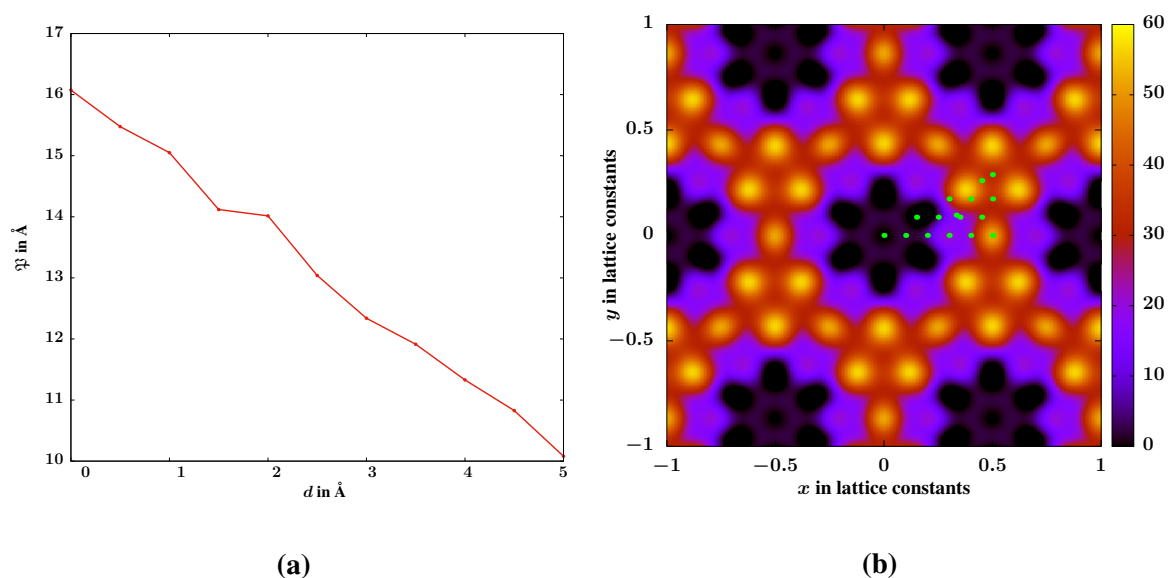


Figure 10.21: Analysis of the interlayer displacement behavior for the TpPa-COF system. a) Average pore size $\mathfrak{P}(d)$ determined using the PSD for FRDG-generated random paths using average in-plane displacements d . The randompaths were constructed for 40 layers with interlayer distances of $c = 3.44 \text{ \AA}$ as obtained from the optimized AA' unit cell. b) The PES determined following Section 10.5.1, the energies corresponding to the respective colors are given in kcal/mol.

suitable due to oxygen atoms reaching into the pore and therefore determining \mathfrak{P} as calculated with PSD. However, the walls of the TpPa-COF consist of oxygen atoms, to account for this, the obtained pore size has to be corrected by $r_{\text{vdW,O}} = 1.54 \text{ \AA}$. Recalculating the \mathfrak{P} with the appropriate vdW-radius, results in a shift to overall smaller values. Displacements larger than the observed minima in Figure 10.21b are unlikely to occur, due to the stability of the resulting oxygen phenyl arrangements, as shown in Figure 10.20.

Polarization by Adsorbed Molecules

Polarization by guest molecules was determined to be insignificant for COF-LZU1 in Section 10.6.1 and associated to the limited interaction of atoms at the pore walls with the gas molecules. In the case of the TpPa-COF, the situation is different due to oxygen atoms ranging into the pore, increasing interaction due to their lone pairs. In analogy to the previously performed study for COF-LZU1, the charge distribution for various loadings of CO_2 was analyzed by calculation of DDEC6 partial charges for structures created based on optimized AA' shifted structures for 2×2 super-cells. The density was obtained from single point energies using

GPW-DFT with $E_{\text{cut}} = 500$ Ry, $\text{EPS_DEFAULT} = 10^{-14}$ as well as $\text{EPS_SCF} = 10^{-7}$. Averaged partial charges \bar{q} are determined by the CS algorithm averaging over the equivalence classes \mathfrak{F} respectively. The data in the case of CO_2 adsorption is summarized in Table 10.12 showing larger deviations on oxygen atoms than observed for all other classes in the case of COF-LZU1, increasing with the applied pressure P . However, recalculation of the frameworks adsorption at the given point of the isotherm by GCMC simulations revealed again no significant changes. In summary, polarization by adsorbate molecules has no significant influence on the frameworks adsorption behavior as simulated by means of GCMC simulations using the DREIDING²⁴⁹ forcefield¹⁶². The framework was used with a constant charge distribution as well as fixed bond distances and atomistic positions during the GCMC simulations. No significant change was observed due to polarization by guest molecules, however the framework contraction due to increased pressure would mark the natural follow up for investigation. Modification of the adsorption isotherm with respect to the experiment might rather be achieved by use of a force field describing dynamically changing bonds introducing possible layer curving and allowing for the occurrence of additional adsorption sites by adjusting the "exposed" surface area.

Table 10.12: Calculated DDEC6 net partial charges $\bar{q}(P)$ averaged over each identified equivalence class \mathfrak{F} (the element of is indicated in brackets), for CO_2 loaded framework super-cells for the TpPa-COF framework as extracted from GCMC calculated isotherms at $T = \text{K}$ in for different values of the environment pressure P in bar.

| \mathfrak{F} | $\bar{q}(0)$ | $\bar{q}(1000)$ | $\bar{q}(7500)$ | $\bar{q}(10000)$ |
|----------------|--------------|-----------------|-----------------|------------------|
| O-1 | -0.484 | -0.513 | -0.517 | -0.496 |
| N-1 | -0.234 | -0.228 | -0.230 | -0.228 |
| C-1 | -0.243 | -0.244 | -0.246 | -0.244 |
| C-2 | -0.169 | -0.164 | -0.165 | -0.165 |
| C-3 | -0.159 | -0.152 | -0.154 | -0.155 |
| C-4 | 0.121 | 0.114 | 0.108 | 0.124 |
| C-5 | 0.204 | 0.204 | 0.204 | 0.206 |
| C-6 | 0.365 | 0.369 | 0.369 | 0.368 |
| H-1 | 0.094 | 0.075 | 0.063 | 0.093 |
| H-2 | 0.114 | 0.100 | 0.089 | 0.112 |
| H-3 | 0.119 | 0.106 | 0.106 | 0.113 |
| H-4 | 0.273 | 0.264 | 0.264 | 0.268 |

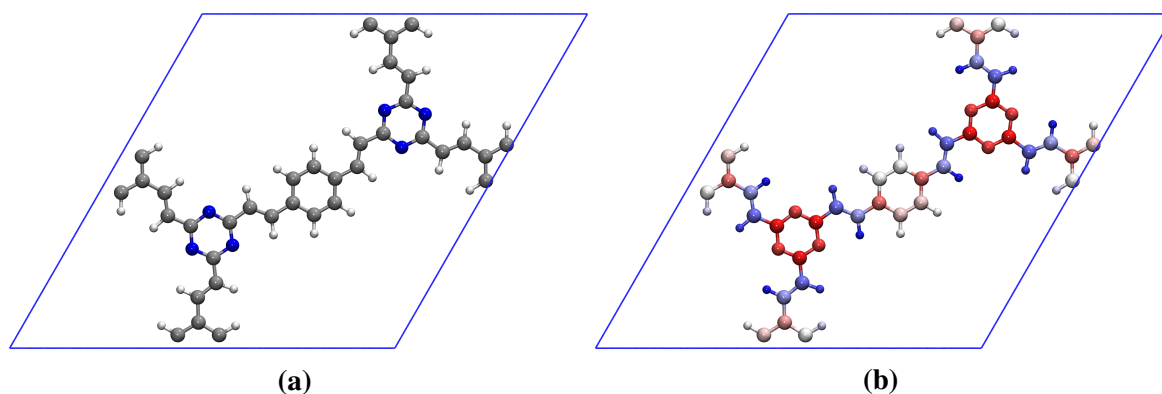
Characteristics: V-COF

Figure 10.22: Overview of the cell optimized V-COF unit cell for structural analysis. a) The unit cell after optimization of the AA stacking with the boundaries of the hexagonal unit cell indicated in blue. b) The unit cell after symmetry analysis with the CS-algorithm, resulting in eleven equivalence classes consisting of six atoms each, differentiated by color. The unit cell with the atoms colored corresponding to their atomic environment as identified by the CS-algorithm.

As an additional system of reference, the V-COF was constructed on the basis of the previous investigations in order to combine the observed effects within a hypothetical framework shown in Figure 10.22. The COF was constructed by modification of the COF-LZU1 unit cell with the aryl unit used for stabilization of the N3-COF. The resulting unit cell for one layer is shown in Figure 10.22a and shows a high level of symmetry in the chosen configuration consisting of eleven distinct atomic environments as can be seen in Figure 10.22b. The dominant interaction between the layers is expected to arise from the aryl groups as the nitrogen bridges in COF-LZU1 are replaced by simple carbon connections. The N3-rings were incorporated in the designed framework, to test the previous observations potentially favoring smaller displacements by favoring the eclipsed stacking motif. Consequently, the effective pore size \mathfrak{P} is maximized due to increased stability of the AA' stacking, without changing the system's unit cell nor the number of atoms. Alternative ways of achieving a favored eclipsed stacking include mainly additional attachments of functional groups ranging into the pore cylinder, counteracting the increase of \mathfrak{P} .

Analysis of the stacking behavior of this COF showed the same energetic trends as the LZU1-COF system when using the standard GPW-DFT description. From the obtained data summarized in Table 10.13 it becomes clear that strongly displaced stackings such as AB and

Table 10.13: Lattice constants and relative energies with respect to the lowest value for the optimized V-COF unit cells. Calculations were performed on the TZV2P-GTH/PBE-D3 level using GTH-PBE-PPs.

| Stacking | $a = b$ in Å | c in Å | ΔE in kcal/mol |
|----------|--------------|----------|------------------------|
| AA | 22.50 | 8.15 | 16.75 |
| AA' | 22.51 | 7.27 | 0.00 |
| AB | 22.48 | 6.91 | 25.49 |
| C | 22.50 | 7.43 | 25.20 |

C are disfavored by $\Delta E \approx 25$ kcal/mol, which is the smallest energetic difference between the stackings that was observed in the study so far. Furthermore perfectly eclipsed stacking is disfavored by $\Delta E \approx 15$ kcal/mol as expected from the previous investigations. However, as the interactions of aryl rings are much higher weighted in this case, the resulting energetic difference can be expected to be of similar magnitude as the one of phenyls. This can be further justified as the energetic differences observed for AA stacking in COF-LZU1 are identical in magnitude when compared to the V-COF system.

Analysis of the obtained energetic landscape shows, the AA' initial structure to result in the most favorable arrangement within the investigated displacement grid on the GFN-xTB level. Therefore, the global minimum as can be seen in Figure 10.23b is located at the origin, without the occurrence of any significant local minimum observed within the displacement grid. This indicates the possibility of this framework to provide highly ordered eclipsed structures with potentially minimal slipping effects. The pore size corresponding to a random displacement of $d = 0.5$ Å within the generated random path, was determined as $\mathfrak{P}(0.5) \approx 1.56$ nm, which would lead to the potentially highest observed value within the augmentations of COF-LZU1. Only slight displacements are of significant, therefore the assumption of a constant interlayer distance obtained from the AA' optimized DFT unit cells within the generated displacement paths is most justified and re-optimization can be excluded as a major source of error within the path generation. The conceptualization of the hypothetical V-COF nicely utilizes the trends observed within the previous theoretical analysis of planar COF-species and exemplifies the prediction of larger pore sizes based on theoretical observations.

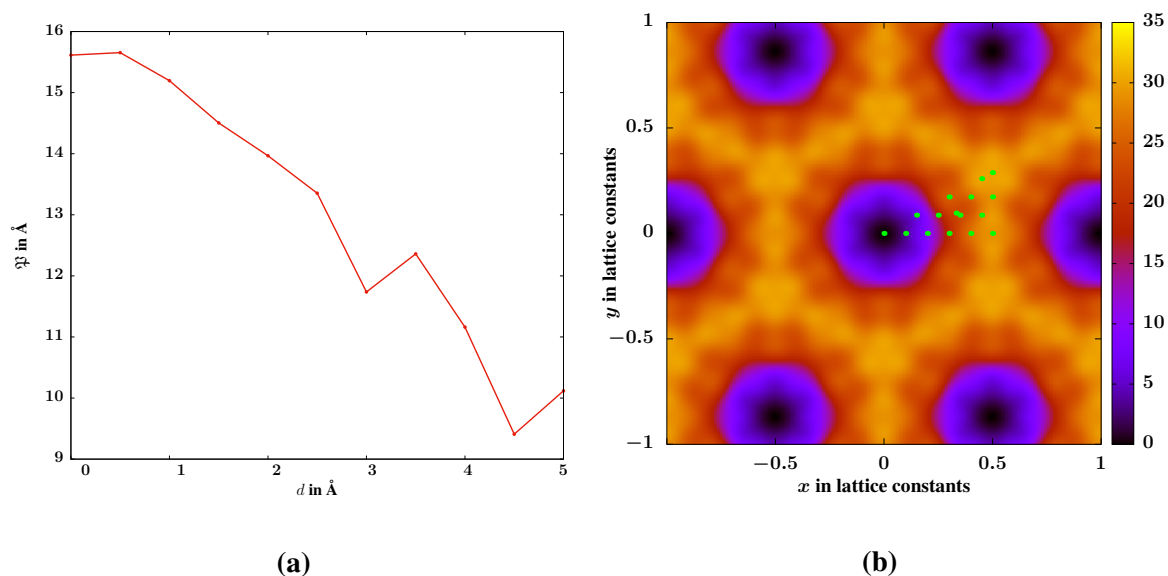


Figure 10.23: Analysis of the interlayer displacement behavior for the V-COF system. a) Average pore size $\mathfrak{P}(d)$ determined using the PSD for FRDG-generated random paths for various in-plane displacements d . The random paths were constructed for 40 layers with interlayer distances of $c = 3.63 \text{ \AA}$ as obtained from the optimized AA' unit cell. b) The PES determined following Section 10.5.1, the energies corresponding to the respective colors are given in kcal/mol.

General Observations for the Planar Species

For planar COFs with symmetric components a treatment of possible intra-molecular orientations is not necessary, allowing for a direct investigation of stacking effects. The interactions between layers for planar systems are dominated by $\pi-\pi$ -interactions and substituted aromatic compounds. The main difficulty for describing the system in accordance with the experiment was found in the correct description of interlayer interactions with the resulting c -lattice constant of the most probable stacking as control parameter. The used standard for GPW-DFT description was chosen after consideration of the five main influences namely: basis set, PW-cutoff and rel-Cutoff, vdW-correction and functional choice. The analysis revealed the main influence as the description of non-covalent interactions Section 10.4.4 as well as the pure functional choice Section 10.4.1, resulting in the standard description Section 10.4.5 used in the following.

The analysis of planar COF systems with the used standard GPW-DFT description shows good agreement with the experimentally obtained values. For all investigated systems the AA'

configuration was determined as the most stable, however the stability under displacement differed depending on the utilized aromatic compounds. Subsequently, the impact of using PBE or BLYP as XC-functional could be shown to be influential on the obtained c -lattice parameters for optimized unit cells. Possible effects of the functional choice can be attributed to the weighting of repulsive interactions arising from various aromatic rings. The choice of the PBE functional for description is less suited than BLYP, for system comprised of heterocyclic aromatic compounds as shown in Section 10.6.1. Furthermore, the energetic differences between the obtained structures differ for O-Phenyl interactions as concluded from Section 10.6.1.

Studies by means of GFN-xTB seem to underestimate the repulsive character of interactions between aromatic rings, leading to smaller c -lattice constants than in the GPW-DFT calculations as shown in Section 10.5. Consequently, interlayer distances are smaller, however the qualitative energetic behavior for displaced two layered systems after cell optimization, is similar to the GPW-DFT case. This can be attributed to GFN-xTB using the D3 dispersion correction as well as non-covalent interactions being a target property of the used parameterization (compare Section 4.3). Calculation of the averaged pore sizes \mathfrak{P} using PSD for FRDG-generated random paths using displacements in the range of $d = 0.0 - 5.0 \text{ \AA}$, hints the range of probable displacements for the COF models. Using the PW-interpolated PES (compare Section 10.5.1), allows for an estimated assessment of the most probable interlayer displacements d , securing the estimates stemming from comparison of $\mathfrak{P}(d)$ with the experimental values. For the case of the TpPa-COF, the O-Phenyl interactions could be furthermore identified as crucial for the obtained stability, as during optimizations interlayer displacements of around $r_{\text{disp}} \approx 5.0 \text{ \AA}$ are overcome without the occurrence of a local minimum satisfying the convergence criteria.

Finally, the insights gained from studying the interactions within COF-LZU1 and the N3-COF systems allowed for the conceptualization of the hypothetical V-COF system. The calculation of PW-interpolated PES showed a single global minimum for the AA' initial configuration, making this framework the most stable under displacements. Calculation of \mathfrak{P} for the obtained interlayer displacement $d_{AA'}$ lead to the largest effective pore size within the series of COF-LZU1 modifications.

10.6.2 Non-planar Species

For the creation of large pore COFs within an azide based building scheme a prominent approach is the insertion of building units acting as bridges between the linkers increasing the pore size while potentially adding structural functionality to the framework. Linkage between the building blocks is easily provided by connecting phenyl groups to increase the pore size of the constructed systems. In many cases these groups are rotated out of plane and the treatment of non-planar COFs becomes necessary. As a precursor for the later investigated PPB-TAB-COF series, effects arising due to non-planar arrangements are discussed at the example of the DAQ-BTA-COF and its modified version with additional methoxy groups TpOMe-DAQ-COF. In the following, both COFs are referred to as COF-NP and COF-NP-m to simplify readability. (-m indicates methoxy attachments)

In the case of the investigated non-planar COF-systems additional effects arise due to rotated aromatic rings. Out of plane arrangements further lead to different types of interactions as will be investigated at the example of methoxy groups within COF TpOMe-DAQ leading to additional stabilization. For this reason the description of these systems has to be adjusted accordingly with the main considerations as always being the cell parameters as well as pore sizes. The previously used standard description fails to reproduce the experimentally obtained interlayer distance indicating the necessity of additional modifications within the used description.

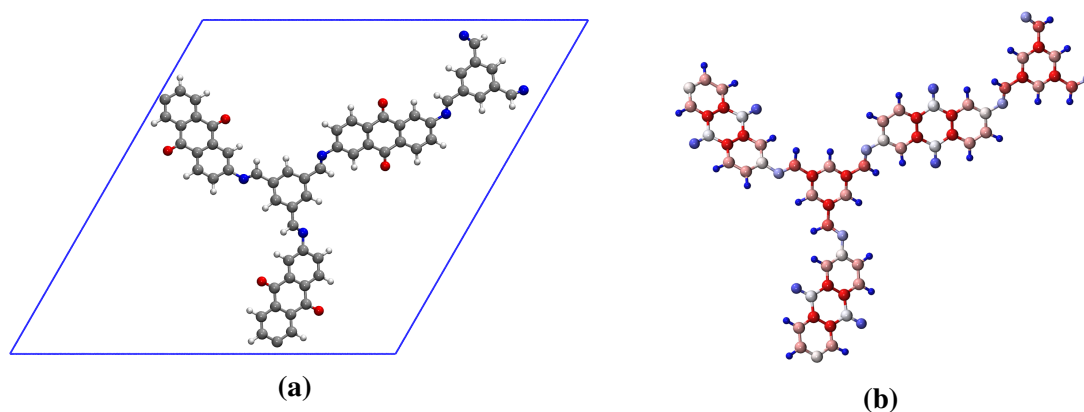


Figure 10.24: Overview of the cell optimized NP-COF unit cell for structural analysis. a) Top view onto the AA unit cell with the cell boundaries indicated in blue. b) The unit cells' atoms colored in correspondence to the fifteen equivalent environments comprised of six atoms each, identified by the CS-algorithm.

Characteristics: NP-COF

The investigated non-planar DAQ-BTA-COF (NP-COF) was used for functionalization of graphene sheets as super capacitors in 2015 by Li and coworkers²⁵⁰. However, within this thesis it acts as example system to illustrate non-planar description with its structural features are shown in Figure 10.24. The hexagonal unit cell of the NP-COF shown in Figure 10.24a is comprised of the building unit used in COF-LZU1 and the DAAQ bridge unit connected via imine linkages. Analysis of the COFs unit cell shows the fifteen atomic environments within the treated unit cell. Structural data for this COF-system is not provided in the literature, therefore, the NP-m-COF unit-cell parameters as obtained by X-ray powder diffraction are used as initial values. The experimentally reported interlayer distance of NP-m-COF is given as $r_{\text{inter}} \approx 3.7 \text{ \AA}$ which is expected to be similar in the case of the NP-COF. Analysis of the general stacking behavior by means of GPW-DFT was conducted on the standard level for the four main stacking schemes by means of cell-optimizations. The results summarized in Table 10.14 show the AA' configuration to be the energetically most favorable by $\Delta E \approx 14.9 \text{ kcal/mol}$ when compared to the AA case. The shifted structures of the AB and C-type stackings are energetically even more disfavored by $\Delta E \approx 68.3 \text{ kcal/mol}$ and $\Delta E \approx 55.0 \text{ kcal/mol}$. The AA' stacked structure perfectly reproduces the experimentally obtained value of 3.7 \AA obtained for the NP-m-COF, indicating smaller displacements to be favorable.

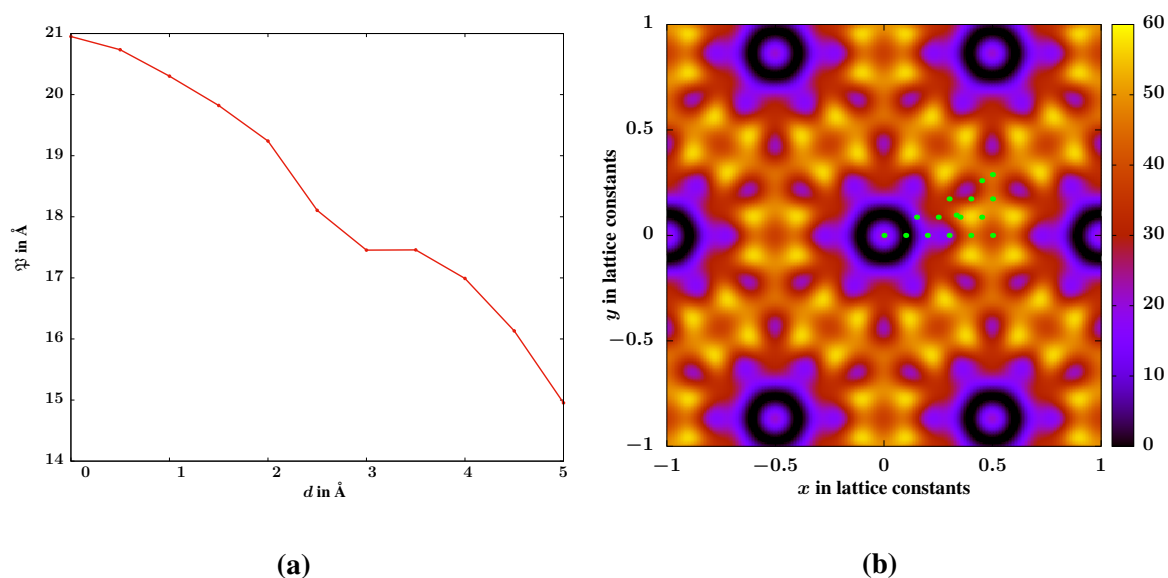


Figure 10.25: Displacement behavior for the NP-COF system. a) Average pore size $\mathfrak{P}(d)$ determined by PSD for FRDG-generated random paths of 40 layers for various in-plane displacements d . Interlayer distances of $c = 3.84$ Å extracted from the optimized AA' unit cell were assumed. b) The PES determined following Section 10.5.1, with energies in kcal/mol indicated by color.

Table 10.14: Lattice constants and relative energies with respect to the most stable configuration for adj-COF unit cells, cell-optimized using the GPW standard setup.

| Stacking | $a = b$ in Å | c in Å | ΔE in kcal/mol |
|----------|--------------|----------|------------------------|
| AA | 29.53 | 7.97 | 14.92 |
| AA' | 29.62 | 7.67 | 0.00 |
| AB | 29.53 | 7.13 | 68.31 |
| C | 29.67 | 7.36 | 55.00 |

Analysis of the corresponding displacement map using GFN-xTB, reveals the global minimum located for slightly displaced structures. A second minimum is found for the (0.4,0.2) displaced initial structure. The obtained minima, contain structures with alternately stacked imines, partially reproducing the geometries of the previously observed structures for the (0.1,0) case. In summary slightly displaced structures, as shown for the AA' optimized unit cells, are energetically the most stable and therefore dominate the stacking behavior. Displacements between $d = 2.0 - 4.0$ Å between adjacent layers are to be expected. Pore sizes calculated for the corresponding displacements as shown in Figure 10.25a, are given in the range of 20.0 Å. For higher entropic environments given by increasing temperatures, the obtained stacking schemes are expected to drift more reducing the effective pore size.

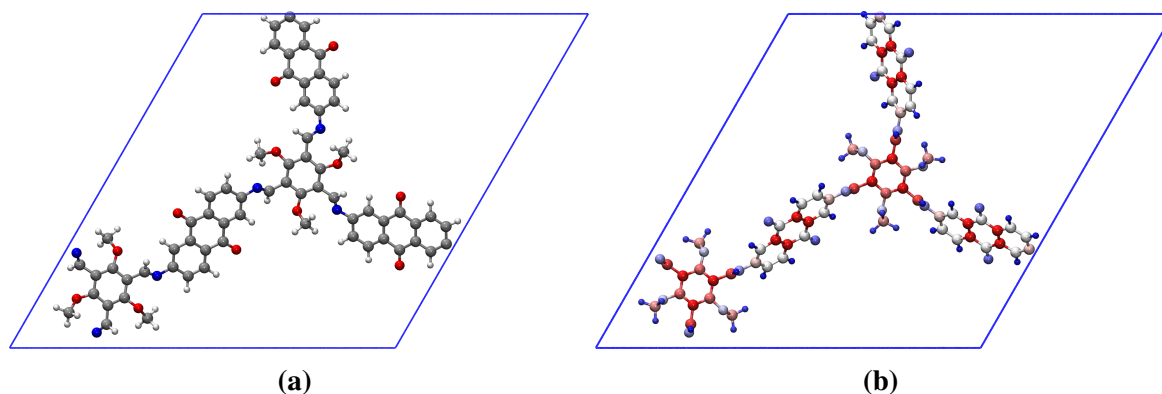


Figure 10.26: Overview of the cell optimized NP-m-COF unit cell for structural analysis. a) The unit cell after optimization of the AA stacking with the boundaries of the hexagonal unit cell indicated in blue. b) The unit cell with the atoms colored corresponding to their atomic environment as identified by the CS-algorithm showing 19 different equivalence classes. (A descriptor cutoff of 3.0 Å was used.)

Characteristics: NP-m-COF

The COF TpOMe-DAQ was first synthesized by Halder et. al. in 2018²⁵¹ to show the applicability of COFs as super-capacitors. The framework's chemical stability was further increased by the formation of interlayer hydrogen bonds due to additionally attached methoxy groups. The observed effect of hydrogen bonding as a new type of interlayer interaction was investigated specifically for its capability of stabilizing the COF against thermal instability. This can be directly seen from its energetic behavior under displacements. The basic structure of the modified COF is illustrated in Figure 10.26a, showing the additional methoxy attachments to the rings in the Tp-linker segment. The methoxy groups are oriented out of plane creating additional interlayer interactions as arising due to their systematic re-orientation towards the methoxy groups of the following layer. Furthermore, methoxy groups are observed to orient themselves towards the imines in subsequent layers when slightly displaced structures are used. An analysis of the standard displacement cases by means of GPW-DFT based cell optimizations revealed the AA and AA' configurations to be the most favorable with a relative energy difference of $\Delta E \approx 0.3$ kcal/mol as summarized in Table 10.15. This is an exceptional result unprecedented in the conducted study and can directly be attributed to the methoxy groups for their role as functional anchors. This phenomenon is further supported by the fact that the displaced AB and C type structures result in energies that are even more disfavored

Table 10.15: Lattice constants and relative energies with respect to the lowest value for the optimized vdW-COF unit cells. Calculations were performed on the TZV2P-GTH/PBE-D3 level using GTH-PBE-PPs.

| Stacking | $a = b$ in Å | c in Å | ΔE in kcal/mol |
|----------|--------------|----------|------------------------|
| AA | 29.46 | 7.94 | 0.31 |
| AA' | 29.46 | 7.87 | 0.00 |
| AB | 29.53 | 7.67 | 86.23 |
| C | 29.43 | 7.25 | 73.13 |

then in the standard case without methoxy groups leading to the highest differences with $\Delta E \approx 86$ kcal/mol and $\Delta E \approx 73$ kcal/mol. A crucial observation made for this COF system lies in the fact, that even though the AA' type stacking as the most favorable, the obtained interlayer distance $r_{\text{inter}} = 3.9$ Å in stacking direction, deviates from the experimentally obtained value by 0.2 Å. This is even more the case for AA-type stacking where the deviation is close to 0.3 Å showing the structural discrepancies. However, an analysis of the corresponding displacements, reveals only one global minimum in the potential energy displacement map shown in Figure 10.27b. The minimum energy obtained by analysis with GFN-xTB is identified for zero displacement between the layers in accordance with the conducted DFT analysis. Furthermore, the result is supported by the accordance of the obtained pore size for $d = 0.0$ Å displacement between subsequent layers as shown in Figure 10.27a, with a value of approximately 23.5 Å obtained by the PSD algorithm. This result is of great significance as possible statistical effects due to FRDG-generated random paths are negligible.

The results of the conducted study show great consistency within the applied methods and are in overall good accordance with the experimental observation. This allows the conclusion that methoxy attachments lead to improved stability by favoring eclipsed stacking and in the observed structures as well as slightly increased pore sizes.

General Observations for the non-Planar Species

The analysis of non-planar species shows the general applicability of the previously obtained setup for the planar COF systems. Generally the occurrence of different types of aromatic interactions shows differences in the used theoretical description as well as the difficulty of determining the most suitable structural model for investigations. The rather simple example

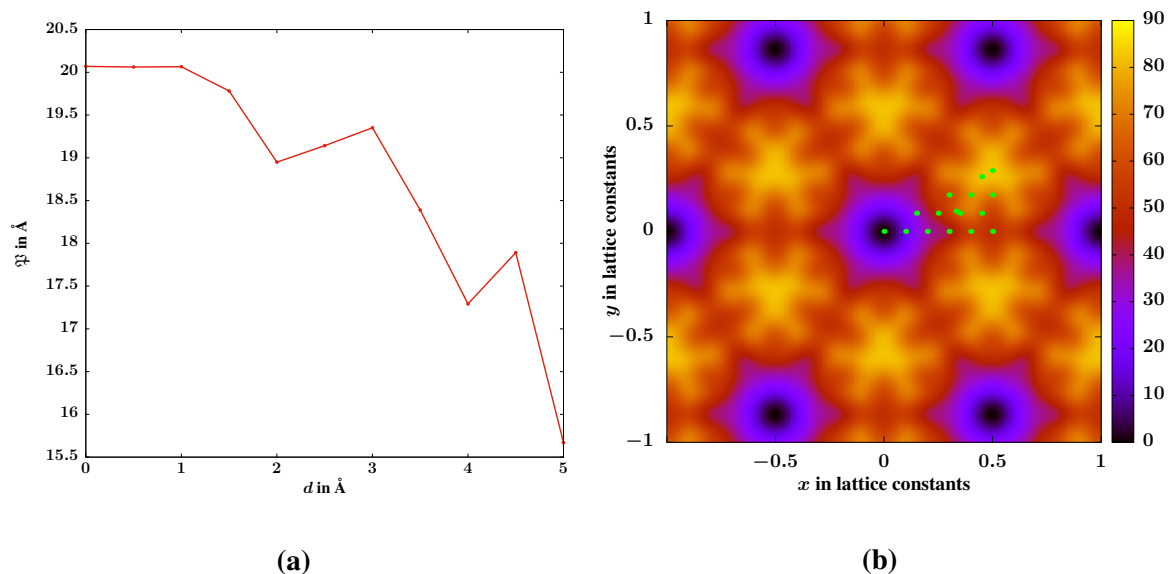


Figure 10.27: Analysis of the interlayer displacement behavior for the NP-m-COF system. a) Average pore size $\mathfrak{P}(d)$ determined using the PSD for FRDG-generated random paths using average in-plane displacements d . The randompaths were constructed for 40 layers with interlayer distances of $c = 3.94 \text{ \AA}$ as obtained from the optimized AA' unit cell. b) The PES determined following Section 10.5.1, the energies corresponding to the respective colors are given in kcal/mol.

case of the NP-COF and its modified NP-m-COF version with additional methoxy groups showed great agreement with the experimental results. Furthermore, the pore size obtained for the minimal displaced NP-m-COF indicated the perfectly eclipsed structure to be the most probable by means of PSD-FRDG analysis. A similar but slightly larger displacement was found to be optimal for the unmodified NP-COF as well as a generally more flat displacement map Figure 10.25b. A larger pore size does not necessarily indicate a higher thermal stability, however, it could be shown that the only minimum in the conducted study of displacements was found for the eclipsed stacking. This result further indicates increased stability which showed the effects of methoxy attachments due to the newly introduced systematic reorientation between the groups of subsequent layers. The example of methoxy groups acting as "anchors" for eclipsed structures acts as intriguing example for stabilizing interlayer interactions provided by additional attachments to standard COF backbones in two dimensional COFs, potentially offering insights into the design of highly crystalline large pore systems.

10.7 Insights Gained for the description of COF-systems

The description of even the simplest two-dimensional COF systems was shown to be highly adventurous when the focus of interest is the determination of the most probable structural arrangement as well as the corresponding theoretical description. A possible workflow for the systematic determination of the most probable COF-structure as well as the corresponding theoretical description is suggested in the following.

- Initially a geometric analysis of the pore size for perfectly eclipsed structures gives insights in the general trends for the modeled system. The resulting properties are the target parameters for the AA configuration and \mathfrak{P}_{AA} .
- Secondly, the AA' configurations energetic difference ΔE with respect to the AA case is determined. The obtained target unit-cell parameters alongside the slightly modified $\mathfrak{P}_{AA'}$ are then compared for their plausibility with the experimental data.
- With the gathered data further analysis of the shifted structures is justified if AA' is energetically more favorable, indicating layer shifting to be energetically more favorable. If the lattice parameters are far off the experimental target values a switch in the functional or dispersion correction depending on the present interactions is worth consideration. Especially, if the geometrically obtained pore size $\mathfrak{P}_{AA'}$ fits the experiment, but the cell vector $c_{AA'}$ does not.
- Further investigations of the pore size with random structures using the cell parameters obtained for the AA' case allows for an estimate of the range of probable displacements d by comparing $\mathfrak{P}(d)$ with the experimental value \mathfrak{P}_{exp} . The probability of a given displacement can then be asserted by probing the energetic landscape for shifts within the hexagonal unit-cell. In this work a PW-interpolation as described in Section 10.5.1 fulfills this purpose and reveals the most probable vicinities for energetic minima.
- From the obtained pore size \mathfrak{P} as well as the position of energetic minima $E_{min}(n_1, n_2)$ within the displacement maps, the most probable structural arrangement within the framework of the presented methods is determined. Recalculating unit cells within the

10.7 Insights Gained for the description of COF-systems

determined vicinity highly increases the possibility of hitting the (or at least one of the) dominant structures for the observed experimental conditions.

- If the target properties are far off the experimentally determined ones, e.g. $\Delta\mathfrak{P} > 20\%$, the optimization of AB and C type stacking on DFT-level might reveal possible minima for other high symmetry schemes. If lattice parameters, especially the c -lattice constant, are far off the experimental value with the structures leading to the desired pore size \mathfrak{P} , improving the theoretical description by consideration of higher accuracy in the setup (compare Section 10.4.2 and Section 10.4.3) or more suitable dispersion treatment (compare Section 10.4.4) is necessary.

Of course the suggested algorithm will not work for every system and an automation of the work-flow faces severe difficulties depending on the level of individuality in a given COF system. The general applicability of the presented methods also depends on the level of accuracy required for determination of the target properties as well as the conducted measurements. Follow up investigations benefit greatly from an accurate description of the most probable layer to layer arrangement. However, possible improvements include three-layered and bigger systems directly considering the interplay of different interlayer shifts, as they are expected from the experiment. Determination of the pore size for a fully DFT optimized random path would be the most sophisticated way of constructing a model in correspondence to the experiment. A possible way of achieving this goal lies in the training of a Neural Network for calculation of extended systems on DFT level. Comparison of the total energy for optimized displacement paths could be used to generalize the procedure from determination of the most probable stacking of adjacent layers, to the most probable displacement paths corresponding to a given displacement d . Calculation of the ensemble average over a multitude of FRDG-generated paths for a given displacement d would allow for a more sophisticated comparison.

The presented study provides the reader with the necessary framework for determining a suitable theoretical description based on comparison with experimental data. Consideration of the dominant interactions for the example COFs gives an idea for identification of necessary configurations and trends to be considered for e.g. generating a suitable training set for a neural network. Furthermore, theoretical analysis of the COFs' component interaction

10 Theoretical Analysis for Covalent Organic Frameworks

within a given unit-cell allows the design of hypothetical frameworks maximizing the desired target properties by combining the most suitable interaction types. This conceptualization has been exemplarily shown for the V-COF maximizing Φ by favoring the eclipsed configuration. The same concepts are used in the following study of large-pore COFs (Chapter 11), using stabilization of the eclipsed structures by methoxy-attachments for highly non-planar COFs, similar to the investigated example case for the non-planar species.

Chapter 11

Theoretical Investigations of Interlayer Interactions for Large-Pore COFs

The design of covalent organic frameworks (COFs) for catalytic applications requires large pores. The real stacking behavior of two-dimensional COF systems is decisively influenced by interlayer slipping and crucial for the effective pore size Φ ⁵⁹. Systematically influencing the stacking behavior as well as slipping effects becomes especially important for COF design aiming at tuning Φ as main target property⁵¹ due to high required crystallinity and stacking order. Furthermore, for applications at room temperature, the stability towards thermally induced interlayer displacements d is of significant importance as it potentially leads to pore blocking. The general possibility of rationalizing the design of COFs, by theoretical investigations is illustrated by studying the phenanthridine based PP-TAB-framework and its methoxy modifications as published in "Interlayer Interactions as design Tools for Large-Pore COFs"⁶⁰. Attachment of methoxy groups has been shown to potentially lead to high stability in a COF system by Halder et. al.²⁵¹. For the example PP-TAB COF-series both target properties are tuned by interlayer interactions arising from methoxy-group attachments to the basic framework. The attachments act as molecular "anchors" adding stability by favoring the eclipsed AA' stacking scheme. Using the whole repertoire of methods introduced in Chapter 10, the exploitable interlayer interactions for large-pore design shown at the example of the isorecticular PP-TAB-COF-series are elucidated.

11 Theoretical Investigations of Interlayer Interactions for Large-Pore COFs

With the insights gained from the creation of the theoretical model for structural analysis, the effects were further studied experimentally by synthesis of the even larger PP-TAPB series of COFs. Their corresponding pore size was determined to be of the magnitude of the so far largest synthesized imine based COFs. The theoretical investigations conducted show good agreement with the experimental findings and further simplify the interpretation of the observation for both COF-series, illuminating the path for future predictions by theoretical investigations.

With the methodology proven to be applicable a second type of interlayer interactions is studied by theoretical means aiming at prediction of its suitability for the creation of large pore COFs. Here, the effects of fluorination for the phenylphenanthridine component in the PP-TAB series is investigated. The theoretical prediction reveals the phenomenon of T-stacked fluorophenyls in large pore COFs as stabilizing interlayer interaction. Furthermore, the study aims at a systematic increase of the T-shape stacking phenomenon as a tool for decreasing layer buckling and increasing pore stabilization. Due to the observed proclivity of linkers to stack in T-shaped configuration, theoretical investigations reveal the potential of a second class of interactions for large pore design. Combined with the previously investigated methoxy attachments, eclipsed configuration and therefore T-shaped stacking can be systematically induced, potentially maximizing the previously observed effects and consequently the pore size β .

Based on the two separated studies the limitations of the theoretical predictions as well as potential alternatives for suitable interlayer interactions are discussed in an attempt to provide possibilities for rationalizing the design large-pore COFs by theoretical investigations.

11.1 Categorization of the PP-TAB COFs

All COFs constructed throughout this study, are based on the same PP-TAB basic system, which is then modified by various add-ons. The basic system is investigated by explicitly targeting its building blocks for their respective contribution to overall layer interactions, for rationalization of the following analysis of full COF systems. The PP-TAB system consists of a phenylphenanthridine based building block in combination with diamine units connected to

the previously introduced C3-linker units.

Interlayer interactions are studied by adding methoxy-groups to the COFs central unit as well as various linker alignments of the phenantridine units. The COF can be categorized, in the sense of the nomenclature introduced in this work, as non-planar, non-symmetric (in the sense of full D_{6h} symmetry) and two-dimensional. A multitude of difficulties arises due to the COFs complexity, hindering the direct comparison of target properties with the experiment. The used models especially consider the uncertainty of interlayer displacements as well as structural orientations between adjacent layers. The system itself is comprised of multiple components, leaving many uncertainties in their expected arrangements due to their individual orientation. This makes consideration of the resulting multitude of conformers necessary when analyzing interlayer displacement effects and stacking schemes. Furthermore, the large pore COFs of the investigated series, consist at minimum of 279 atoms per unit-cell as shown for the PP-TAB system in Figure 11.1. Consequently, calculations considering two-layered unit-cells, quickly lead to the limits of computational affordability. However, investigations of these systems are required for interpretation of interlayer-interactions for structural clarification in two-dimensional COF systems (compare Section 10.6). Therefore, full analysis of all conformers has to be avoided, making identification of the most probable structures necessary before full unit-cells are considered. Even when only eclipsed configurations are considered, the explicit calculation of optimized unit-cells for all possible conformers by means of GPW-DFT, is neither affordable nor yields a unique optimal structure.

In the following analysis, the eclipsed case is treated representatively for all displacements as it is the targeted stacking scheme for the investigated large-pore systems. Here, two layers are assumed to stack directly with small to no in-plane displacement, leading to interaction of identical components between adjacent layers. The total interaction is provided as a sum over multiple contributions, which can be investigated in separation from each other as well as in small combinations of “blocks” as well as pore “walls”. The most probable eclipsed structure is constructed by hierarchical analysis of interactions, from gas phase components over periodic compounds to full unit-cell arrangements. In a first step the interaction of the identifiable segments comprising the pore walls leads to an overview of probable two component arrangements in the eclipsed case. An overview of the components specific

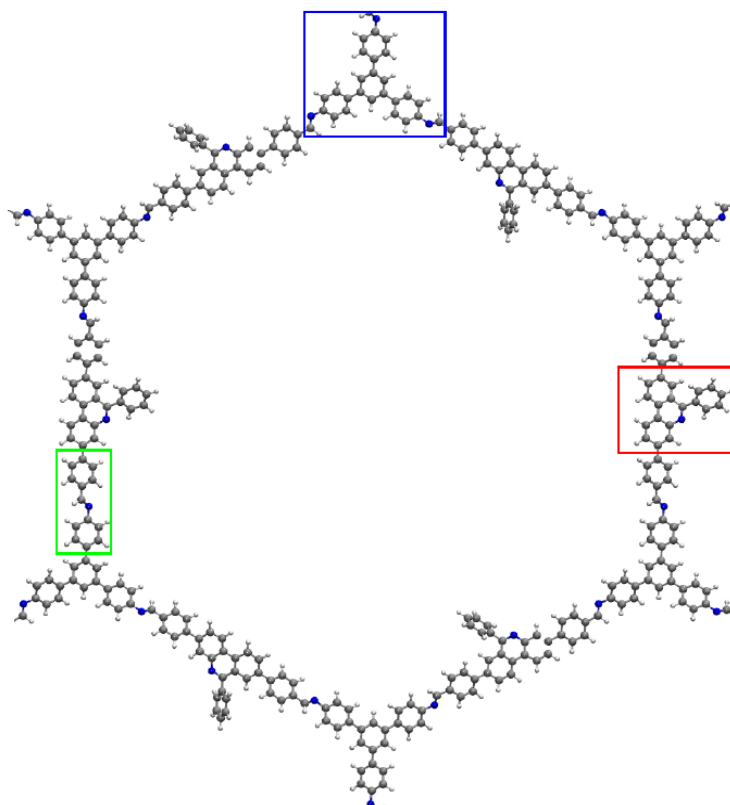


Figure 11.1: Exemplary unit-cell structure for one layer of the unmodified PP-TAB-COF system, forming hexagonal pores, illustrating the general properties of the system. The components comprising the COF are indicated by their respective colors, in blue the C3-unit, in green the imine linker (IL) and red the phenylphenanthridine (PP) component.

properties is of principal interest for all follow up investigations. In a second step, the degrees of freedom as well as theoretical description is adjusted to the case of interacting components in the COF system, by constructing auxiliary “blocks”. From the obtained insights, the combination of components is further investigated by treating the interaction of auxiliary “walls”, reducing the computational costs while modeling the main effects to be considered. In a final step the results of all three levels of investigation are combined to construct a set of the most probable structures for the full unit-cell treatment. Furthermore the obtained eclipsed configuration is used as initial for the generation of interlayer displacement effects following the methodology described in Section 10.5.1. The most probable stacking schemes are illustrating by PW-interpolated energetic displacement maps. Comparison with the experimentally obtained pore size and interlayer displacements, the COFs thermal stability

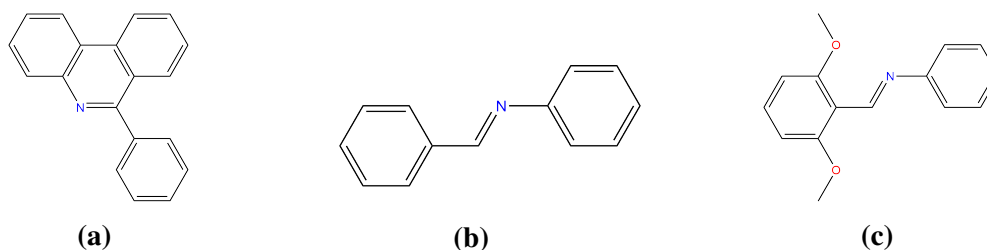


Figure 11.2: The Lewis structures of the used building blocks for the PPD-TAB COF as arising within the complete framework. a) The phenanthridine linker (PP), b) the imine linker (IL), c) the methoxy-imine-linker (MIL).

as well as its most probable stacking scheme is provided. The pore size is determined by means of FRDG (compare Chapter 5) generated displacement paths using the PSD algorithm presented in Chapter 7.

11.2 Component Analysis by means of DFT

Interactions between adjacent COF layers can be approximately described by the interactions of the respective components. An overview of the three main building units, of the investigated COF-series, is shown in Figure 11.2. The components' geometries are optimized in the gas-phase, using DL-find interfaced via CHEMSHELL¹⁶³. Energies are obtained by means of DFT using TURBOMOLE¹⁷³ on the PBE-D3/def2-TZVPP level^{73,158,252}.

The optimized compounds are summarized in Figure 11.3, showing the optimized geometries for the phenyl-phenanthridine compound (PP) shown in Figure 11.3a, imine-linker (IL) Figure 11.3b and its modified version given by the methoxy-imine linker (mIL) Figure 11.3c. The methoxy groups are attached to the imine linkers in meta position to the imine bonds, aiming at a reduction of the imine-bonds dipole moment⁵⁷. The C3-backbone is identical to the already investigated compound in Section 10.6.1 showing high flexibility by rotating its components out of plane due to phenyl-phenyl interactions. Initial optimization of the isolated components' geometries by means of DFT allows the use of molecular electrostatic potentials (MEP) to gain an overview of the underlying electrostatics for estimating the most probable arrangements. The MEPs are constructed by calculation of an iso-surface of the electron density $\rho(r_{\text{iso}}) = \text{const.}$. The MEP calculated as the force acting on a positive probe charge q^+ is calculated at the positions r_{iso} on the surface and mapped onto the generated density

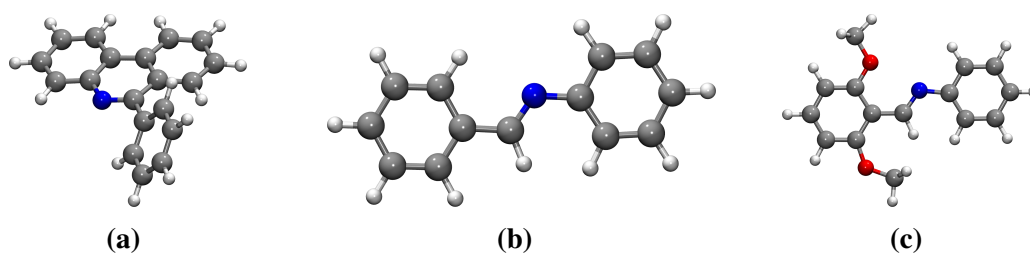


Figure 11.3: Structural overview of the building blocks for the PPD-TAB COF after geometry optimization. a) The phenantridine linker (PP), b) the imine linker (IL), c) the methoxy-imine-linker (MIL).

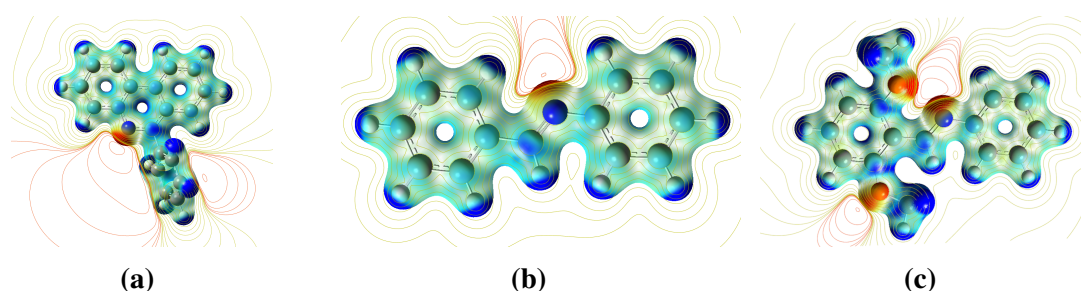


Figure 11.4: The MEP mapped onto an iso-surface of the density $\rho(r_{\text{iso}}) = 0.04 \text{ e/Bohr}^3$, calculated with GAUSSIAN for the three components. a) The phenantridine linker (PP), b) the imine linker (IL), c) the methoxy-imine-linker (MIL). Red areas of the mapped MEP correspond to nucleophilic whereas blue ones correspond to electrophilic regions. The electrostatic potential lines are drawn within the plane with normal vector $(0,0,1)$, to show the potentials impact in the immediate molecular vicinity. Yellow colored potential lines correspond to repulsive, whereas red corresponds to attractive forces on positive charges.

surface with its strength indicated by color. The potential is plotted in the range of -0.02 a.u. to 0.4 a.u. with colors ranging from red over green to blue.

For the targeted eclipsed stacking scheme the displacement d between adjacent layers is small as exemplary shown in the AA' case in Chapter 10. Therefore, the electrostatics of identical components with their counterpart in adjacent layers mainly determine the interlayer interactions. Consequently, probable arrangements can be analyzed by investigation of two-component interactions. For two-dimensional COFs the stacking of layers can be estimated by arranging components in order to minimize repulsive electrostatic interactions. Identification of the most probable component arrangement within the investigated COF-series is motivated by two main considerations:

- In the case of eclipsed stacking the pore walls are constructed by parallel oriented and

slightly displaced layers allowing only limited rotations with respect to each other.

- The component arrangement with the highest interaction energy E_{int} provides the most stability to the COF-system favoring its occurrence.

In a first step vacuum components are analyzed for their stacking behavior by geometry-optimization of two components placed on top of each using a gas phase DFT treatment. The optimized combined structures are then compared with their isolated counter parts, in order to determine the amount of guaranteed stabilization. The stabilization is estimated by calculation of the interaction energy for components 1 and 2 following

$$E_{\text{int}} = E_{12} - E_1 - E_2 \quad , \quad (11.1)$$

with the energy of the combined system E_{12} and of the isolated components E_1 and E_2 .

11.2.1 Phenylphenanthridine-Combinations

A main feature of the asymmetric building blocks is the importance of their orientation with respect to each other after their combination. Combining two identical Phenylphenanthridine (PP)-components, results in four main orientations as shown in Figure 11.5. The possible PP-stacking orientations are distinguishing in S (same-side) arrangements and D (different-side) arrangements depending on the relative orientation of phenyl linkers with respect to each other. Furthermore, the position of the phenyl attachment of the central ring is distinguished by 1 and 0. The nomenclature for possible combinations is explained at the example of the S10 as well as D10 configurations in Figure 11.5a and Figure 11.5b. Furthermore, parallel and T-shaped stacking of the pore ranging phenyls is distinguished. The parallel alignment is exemplary illustrated for the S10 configuration shown in Figure 11.5c whereas the T-shaped orientation, indicated in the nomenclature by an additional -T, is shown in Figure 11.5d.

Optimizations were performed on the def2-TZVPP/PBE-D3 level with the interaction energies calculated following Eq (11.1) leading to Table 11.1. The results show the S arrangements to be favored with the S10 combination showing the largest interaction energy with $E_{\text{int}} \approx 14.8$ kcal/mol followed by the S11 configuration. The T-shape arrangements are in both cases disfavored when compared to the parallel stacked ones. In the case of

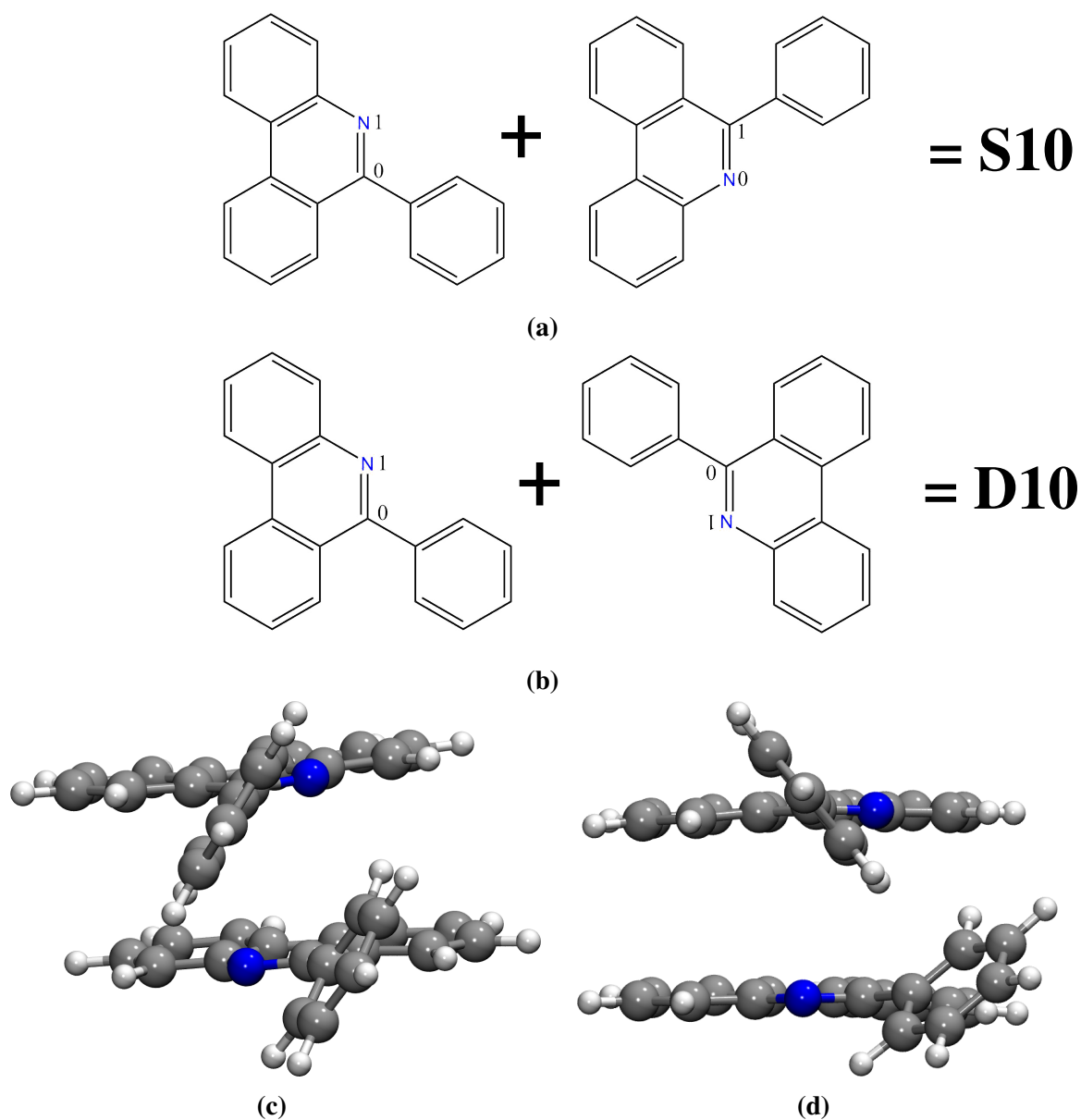


Figure 11.5: Optimized geometries for the S10 configurations explaining the nomenclature. a) Scheme presenting the used nomenclature for the combination of PP-components resulting in the S10 configuration. b) Scheme presenting the used nomenclature for the combination of PP-components resulting in the D10 configuration. c) Parallel configuration of the PP linkers. (S10) d) T-shape configuration for the PP-linkers. (S10-T)

S10, the corresponding T-shape stacking is disfavored by $\Delta E_{\text{int}} \approx 1.4$ kcal/mol. Finally both D-configurations are shown to be significantly less stable by $\Delta E_{\text{int}} \approx 2.0$ kcal/mol, indicating increased sparsity in their expected occurrence.

Table 11.1: Interaction energies determined by Eq (11.1) for the investigated combinations of phenylphenanthridine linkers.

| Arrangement | E_{int} in [kcal/mol] |
|-------------|--------------------------------|
| S10 | 14.83 |
| S10-T | 13.30 |
| S11 | 14.08 |
| S11-T | 13.59 |
| D11 | 12.48 |
| D10 | 12.20 |

11.2.2 Imine-linker-Combinations

For possible arrangements of imine-linkers within the COF's unit cell, two main orientations are expected and are further referred to as alternating (alt) and parallel (para) orientation, referencing the alignment of imine bonds with respect to each other. Both investigated versions of the imine linker are illustrated after optimization in Figure 11.6. Four distinct component combinations are possible, enumerated following the nomenclature "simple" for the unmodified imine linkers and "methoxy" for the modified version. Geometries were optimized at the PBE-D3/def2-TZVPP-level for the isolated components as well as their combinations respectively. The optimized combinations for the unmodified linkers in the parallel case are shown in Figure 11.6a whereas the alternating imines are shown in Figure 11.6b. Additionally the modified methoxy versions after optimization are illustrated for parallel imines in Figure 11.6c and alternating imines in Figure 11.6d. Interaction energies E_{int} calculated following Eq (11.1) are summarized in Table 11.2. The obtained interaction energies indicate, the alternating arrangement of imines as generally more favorable. Comparison of simple and methoxy structures, reveals an increase in E_{int} by attachment of additional methoxy-groups onto the linking sites, leading to $\Delta E_{\text{para}} \approx 0.74$ kcal/mol for the parallel and $\Delta E_{\text{alt}} \approx 1.54$ kcal/mol for the alternating orientation. Furthermore, the observed energetic difference due to imine orientations is reduced from $\Delta E_{\text{simple}} \approx 3.5$ kcal/mol to $\Delta E_{\text{methoxy}} \approx 2.7$ kcal/mol.

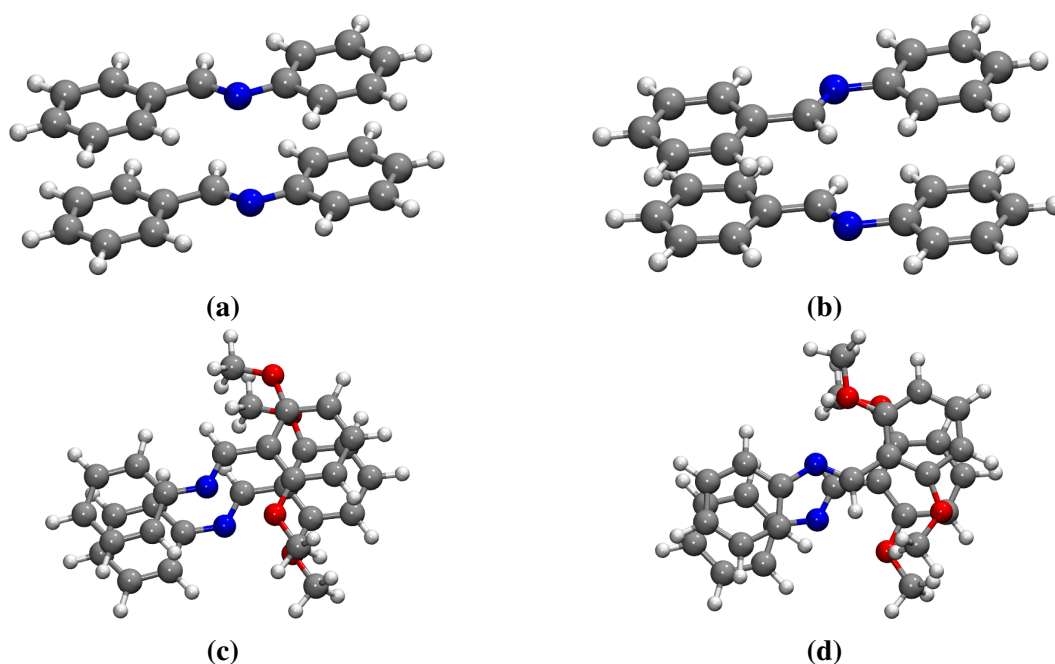


Figure 11.6: Starting orientations for the optimizations of a) parallel IL (simple-para), b) alternating IL (simple-alt), c) parallel mIL (methoxy-para), d) alternating mIL (methoxy-alt).

Table 11.2: Interaction energies determined by Eq (11.1) for the investigated combinations of imine linkers.

| Arrangement | E_{int} in [kcal/mol] |
|--------------|--------------------------------|
| simple-para | 7.68 |
| simple-alt | 11.20 |
| methoxy-para | 9.24 |
| methoxy-alt | 11.94 |

11.2.3 Block Analysis with GPW-DFT

Investigation of the the optimized two component systems for the ILs presented in Figure 11.6, reveals strong rotation out of the initially parallel arrangement. This behavior is not expected within real crystal structures due to restriction of the components rotation by covalent connection with adjacent units. To find a more suitable description PBCs allow for the treatment of "block" structures, suppressing possible rotations out of the simulated alignment. A more suitable model of the interactions within a real COF is expected for blocks of components with restrictions added by additional phenyl groups that are then connected by C-C triple bonds over the periodic boundaries. The created model mimics not only the parallel alignment, but

also the components' immediate vicinity dominated by aromatic interactions. The resulting model scheme is illustrated in Figure 11.7a with periodic extensions indicated by arrows. Cell optimizations were performed for the constructed blocks with PBCs in all spacial dimensions. Energies are calculated using the GPW-DFT approach as implemented in CP2K, utilizing the PBE-D3/TZV2P-GTH level of theory with corresponding PBE-GTH pseudopotentials describing the two respective S electrons. The treated PP-blocks were constructed for all previously studied component orientations, utilizing an orthorhombic unit-cell with cell vectors $a = 25.0 \text{ \AA}$, $b = 22.5 \text{ \AA}$ and $c = 7.5 \text{ \AA}$. The molecules are initially aligned in y and stacked in z direction. The x direction was chosen ad-hoc to guarantee no interactions between the simulated blocks as shown in Figure 11.7a and is therefore excluded from optimization. The blocks' nomenclature follows the previously introduced scheme as presented in Figure 11.5 and Figure 11.6 respectively. The D10 structure after optimization is exemplary shown in Figure 11.7b. Comparison of the obtained energies after optimization for the phenantridine-

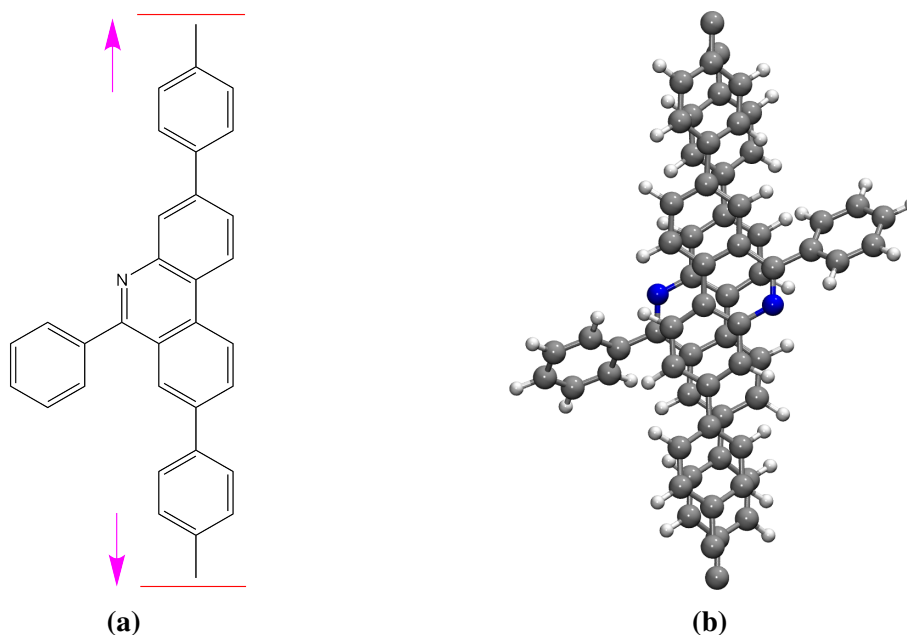


Figure 11.7: Setups utilized for the conducted calculations of PP-blocks. a) The schematic idea illustrated by the blocks Lewis structure for the phenantridine components. Red bars indicate the cell boundaries whereas pink arrows indicate the periodic extension via C-C triple bonds. b) The top view onto the exemplary molecular arrangement within an optimized unit-cell for the D10 configuration.

units as summarized in Table 11.3, allows for the identification of the most stable conformer. The interaction energies E_{int} in the periodic case are not uniquely defined by Eq (11.1), as the

11 Theoretical Investigations of Interlayer Interactions for Large-Pore COFs

exact reference is not clear. Consequently, the obtained energies are sorted according to their relative deviation to the lowest obtained value. The configurations are sorted in reference to the component that was lowest in energy, with the relative deviations from the lowest value ΔE .

Table 11.3: Interaction energies determined by Eq (11.1) for the investigated combinations of phenylphenanthridine linkers.

| Arrangement | ΔE in [kcal/mol] |
|-------------|--------------------------|
| D10 | 0 |
| S10-T | 1.167 |
| S11-T | 1.168 |
| S10 | 0.905 |
| S11 | 5.143 |
| D11 | 6.639 |

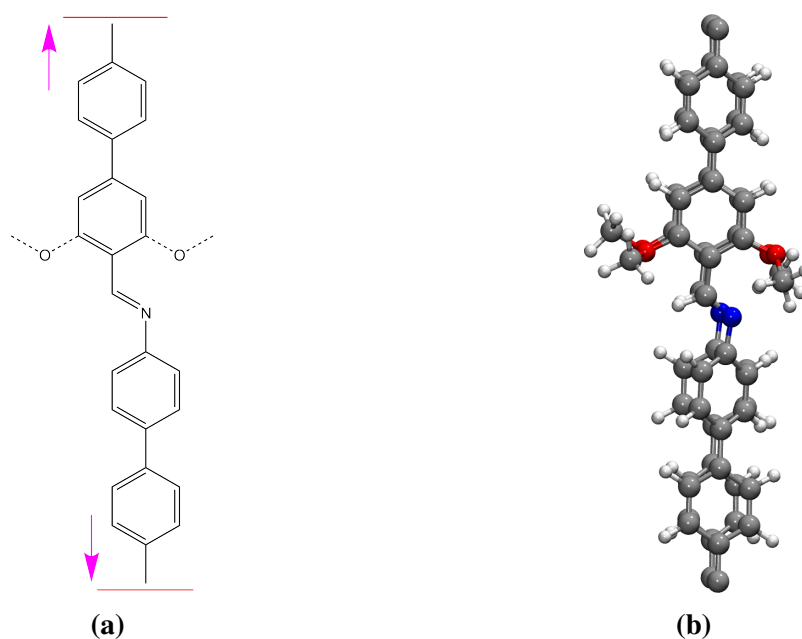


Figure 11.8: Setups utilized for the conducted study of IL-blocks. a) the corresponding scheme for the methoxy-imine-linkers, unmodified structures were investigated analogous, the cell boundaries are indicated in red, whereas pink arrows symbolize periodic continuation. b) The top view onto an optimized example unit-cell for the methoxy-para configuration.

Analogous to the previous PP-blocks the imine linkers and its modified versions were analyzed by optimization of the constructed blocks as shown in Figure 11.8. Energies were calculated on the standard level using GPW-DFT and an orthorhombic unit-cell with cell vectors as $a = 30 \text{ \AA}$, $b = 22.5 \text{ \AA}$, $c = 7.5 \text{ \AA}$. The molecules were aligned in y direction and the a cell-vector was constrained during the optimizations avoiding possible interaction of periodic images in the x -direction. The schematic situation is shown in Figure 11.8a. An exemplary structure for the methoxy-para block arrangement after optimization is shown in Figure 11.8b. The relative energy differences between the alt E_{alt} and para E_{para} arrangements are calculated following

$$\Delta E = E_{\text{para}} - E_{\text{alt}} \quad , \quad (11.2)$$

with the resulting energy differences summarized in Table 11.4. In the case of unmodified imine linkers, the para arrangement is favored by $\Delta E \approx 12.3 \text{ kcal/mol}$. Contrarily, in the methoxy modified version the para arrangement is favored by $\Delta E \approx 5.7 \text{ kcal/mol}$ indicating a significant reduction of the total energetic difference between both arrangements.

Table 11.4: Interaction energies determined by Eq (11.1) for the investigated combinations of phenylphenanthridine linkers.

| Arrangement | ΔE in [kcal/mol] |
|-------------|--------------------------|
| simple | -12.271 |
| methoxy | 5.755 |

T-shape analysis

The stability of T-shaped phenyl groups allows for an interesting consideration, due to the possibility of further restricting interlayer displacements and therefore increasing the crystallinity of the COF-system. The effect of T-shaped stacking in phenyls is a well studied field in the literature^{253–256}. GPW-DFT optimized block systems in T-shape configurations are shown in Figure 11.9. With the case of S10-T shown in Figure 11.9a and S11-T shown in Figure 11.9b. Total energies calculated for the optimized unit-cells were compared to the parallel phenyl

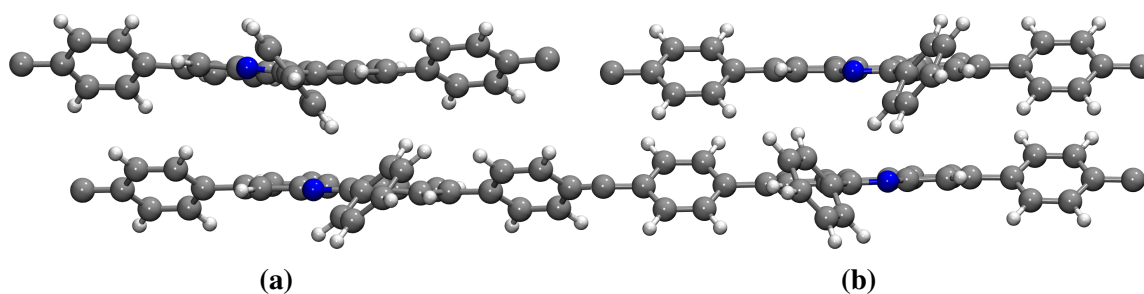


Figure 11.9: Side view onto the optimized geometries for blocks considering T-shaped phenyl linkers after cell-optimization. a) The optimized S11-T configuration. b) The optimized S10-T configuration.

Table 11.5: Energetic differences between both S- and S-T-block arrangements calculated following Eq (11.3).

| Arrangement | ΔE in [kcal/mol] |
|-------------|--------------------------|
| S10-T/S10 | 0.26 |
| S11-T/S11 | 49.87 |

stack following

$$\Delta E = E_{S-T} - E_S \quad , \quad (11.3)$$

with S indicating the respective same side and S-T its corresponding T-shape arrangement. The results are summarized in Table 11.5 showing both T-shaped stacked configurations to be clearly favored in the S11 case by $\Delta E \approx 50$ kcal/mol. However, in the S10 case both configurations are almost energetically on par by differing less than $\Delta E \approx 0.3$ kcal/mol.

The components investigated in Section 11.2 showed the S-T arrangements to be energetically disfavored over parallel stacked S arrangements, however, for the constructed blocks the restriction of in-plane rotations within the model yields different results. For the investigations, both structures are positioned directly on top of each other before optimization, with the phenyl linkers rotated accordingly. The optimized S10-structures resulted only in slight shifts in direction of the b -lattice vector, corresponding to a shifting of length $|r_a| \approx 0.80$ Å. Both configurations are therefore considered equally stable which is significantly different from the study of components in the vacuum case. In contrast, the S11-T structures arrange with a bigger displacement amongst themselves $|r_b| \approx 2.0$ Å. This rearrangement as well as the

T-shaped orientation of the linker groups resulted in a significant energy difference between the structures $\Delta E_T \approx 50.00$ kcal/mol. The S11-T arrangement is therefore more stabilizing than the parallel S11 configuration. The overall lowest energy of all investigated conformers was found for S10-T which is therefore considered the most stable of the PP-configurations following the block analysis data summarized in Table 11.3.

11.2.4 Wall Analysis with GPW-DFT

For an even more detailed analysis of interlayer interactions a combination of components, as they are found in the synthesized PP-TAB-COFs pore walls, is conducted. Each system investigated resembles a possible orientation of the compounds within a given pore wall.

Always two layers of the artificial "walls" were investigated at a time revealing the interlayer interactions in the almost eclipsed configuration. The schematic design of the constructed pore walls for one distinct layer are shown in Figure 11.10. Considering the component sequence of IL-PP-IL of the compounds results in a multitude of possible orientations, where different numbers of attached methoxy groups are considered for modeling the pore walls within the PP-TAB-COF-series. Furthermore, as two IL-compounds are found in each system, a total of three main types of imine bond orientations are distinguished. These consist of bonds orientation in parallel (para), alternatingly (alt) and the occurrence of both (mixed). In order to identify the different orientations a nomenclature in analogy to the previously conducted block analysis is used. As shown in Figure 11.10a always two IL linkers referred to by the previously introduced abbreviations are attached to one PP unit which's combination is referred to following the scheme presented in Figure 11.5. The "walls" are aligned in x -direction and always two are stacked in z direction. The initial unit-cell was chosen as orthorhombic with cell vectors [$a = 33.0$ Å, $b = 25.5$ Å, $c = 7.5$ Å]. The b -vector was fixed during the optimization process to secure sufficiently large distances between the periodic images in y direction.

Optimizations were performed for unmodified IL-linkers summarized in structures referred to as "simple". Furthermore, the modified mIL linker combinations are referred to as "methoxy" configurations. The exemplary structure of an optimized unit-cell for the S10-configuration is shown in Figure 11.10b.

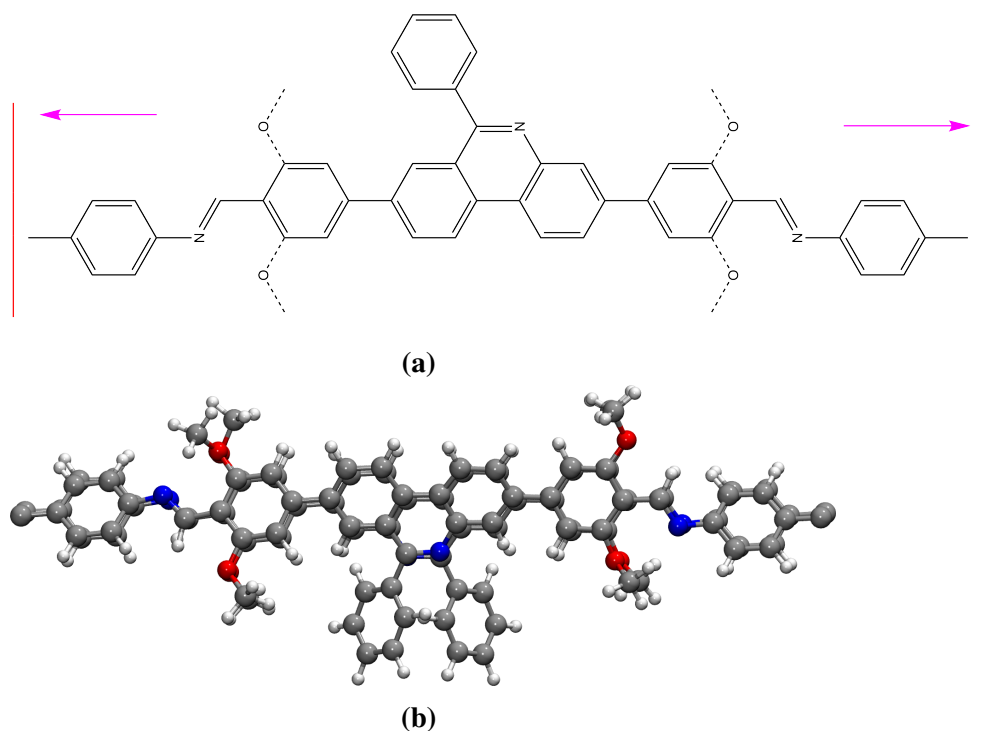


Figure 11.10: Utilized setup for the performed band calculations with a) the schematic idea of the constructed bands, dashed lines indicate possible attachment sites for methoxy-groups. The red bars indicate the cell boundaries, where the periodic extension via the C-C triple bonds is performed in direction of the pink arrows. b) Top view onto an example unit-cell for the methoxy-para-S10 configuration after optimization.

Additionally, the influence of methoxy groups is analyzed by creating walls with only one group attached for each IL. These structures are consequently referred to as "1-methoxy" and are discussed in detail in Section 11.2.5, where the nomenclature is further extended.

When analyzing the results for the band analysis and comparing simple and methoxy structures as shown in Section G.2, three main trends can be observed:

- Using mILs for construction of the pore wall disfavors D arrangements of the PP component when comparing with the S-configurations by at least $\Delta E \approx 13.01$ kcal/mol as seen from the data summarized in Table G.4. In the case of unmodified IL linkers this is not the case and no preferred PP orientation can be identified (compare Table G.3).
- For walls constructed using ILs, alternating imine bonds are more stable than parallel arranged ones for the dominant S10 structure $\Delta E_{\text{alt-para}} \approx 4.71$ kcal/mol. This is not the case when mILs are used, as the differences between alt and para arrangements become

significantly smaller $\Delta E_{\text{alt-para}} \approx 0.5$ kcal/mol. The observed structural effect can be attributed to the initial motivation of reducing the inherent dipole moment of the imine linker by attaching methoxy groups, reducing the associated repulsion, especially for parallel orientations²⁵⁷.

- The case of mixed schemes of imine bond orientations for the S-configurations is a lot less stable when using ILs than for mLs. The energetic differences are determined to be at least $\Delta E_{\text{mIL}} \approx 1.0$ kcal/mol and $\Delta E_{\text{IL}} \approx 5.5$ kcal/mol.

The observations made during the analysis of the constructed wall models lead to the conclusion that methoxy groups increases the structural order of the associated COFs by favoring S arrangements for the PP components as well as disfavoring mixed orientations of the IL-bonds. Experimentally, an increase in crystallinity and thermal stability could be observed for the synthesized dPP-TAB-COF⁶⁰. Furthermore, an increase in the effective pore size was observed which indicates a higher proclivity towards eclipsed stacking types.

These effects can be partially explained by the observations made for the components arrangements so far, however, further insight into the thermal stability requires additional analysis of in-plane displacements d between two layers.

11.2.5 Influences of the Methoxy Groups

The stabilizing role of methoxy groups for eclipsed layer stacking is investigated for the various positions of attachment as well as inter-methoxy interactions. A suitable way is provided by considering imine-linker combinations with only one methoxy group, offering an intermediate between the previously investigated cases. The construction of pore walls based on these linkers, makes extension of the previously introduced nomenclature necessary for a unique identification. The nomenclature distinguishes between various possibilities of methoxy attachments with respect to each other, the two main cases are illustrated in Figure 11.11. Antiparallel positioning (anti) of the methoxy groups is shown in Figure 11.11a, whereas parallel positioned (para) is shown in Figure 11.11b. Each of the main attachment schemes offers the possibility to study a specific effect in more detail. For parallel attachment non-covalent methoxy-methoxy interactions are studied, whereas for anti-parallel arrangements

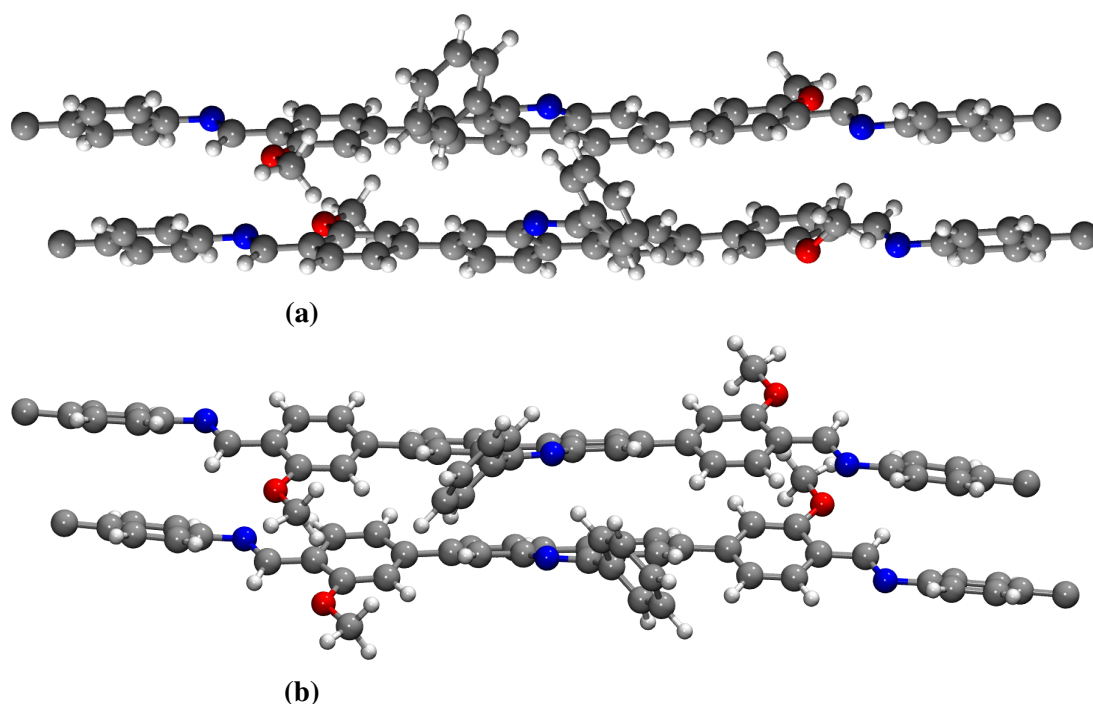


Figure 11.11: Illustration of the parallel and antiparallel arrangement of the methoxy-groups, with a) the 1-methoxy-para-S10-anti configuration after cell optimization and b) the corresponding 1-methoxy-para-S10-para configuration after cell optimization.

the impact on the imine bridges can be estimated. The intermediate situation with each arrangement present for one of the two layers was not considered, because it hinders the interpretation of the associated effects.

The results of the analysis summarized in Table G.5 show two major trends for the investigated structures:

- Comparing both attachment positions shows parallel positioned methoxy-groups favored over their anti-parallel counterparts for all investigated configurations. This clearly indicates favorable non-covalent interlayer interactions between adjacent methoxy-groups as already studied for the mNP-COF in Section 10.6.2 as expected from the literature²⁵¹.
- Investigation of trends for the PP-components within the anti and para series it becomes obvious that parallel arranged methoxy groups favor S-configurations. The least stable S is still more stable than the most stable D configuration. Contrarily, in the case of antiparallel attachment orientations, D configurations are not significantly disfavored

over S-configurations. Additionally, the S11 and S10 configurations with parallel imine groups, with an energetic difference of only $\Delta E = 0.627$ kcal/mol, are determined as the most stable, with the energy difference to the next compound being larger than $\Delta E \approx 11.0$ kcal/mol, indicating a clear preference for S configurations.

Possible reasons for these observations are found by the limited flexibility in displacements provided by the parallel oriented imine bonds as well as T-shaped phenyl-linker orientations. Both configurations disfavor displacements between the layers, allowing for stabilizing methoxy interactions by systematic orientation due to closely located attachment sites. This assumption is further supported when investigating the corresponding arrangements after optimization. Both structures clearly show systematic reorientation as well as distances smaller than 2.5 Å between the CH₃ head and the oxygen of the underlying group placing the interaction at the edge of hydrogen bond distances.

11.2.6 Analysis of Methoxy Group Orientations

Further study of the inter-methoxy interactions between adjacent layers reveals the possibility of two main orientations. The groups can be oriented in parallel or anti-parallel with respect to each other. The situations are shown in Figure 11.13, where the initial and final structures are illustrated for both distinct cases showing the additional configurations to be considered. The nomenclature ordered refers to all methoxy groups being oriented in parallel as shown in Figure 11.12b whereas unordered refers to anti-parallel orientations as shown in Figure 11.12a. The wording is chosen in order to avoid confusion with the previously introduced nomenclature for the attachment position (compare Section 11.2.5). To obtain a suitable overview, optimizations for the whole unit-cells are conducted for the dPP-TAB-COF using the standard setup for GPW-DFT calculations. The initial unit-cell for the following investigations is constructed based on the methoxy-alt-S10 eclipsed AA' arrangement for the dPP-TAB-COFs (compare Section 11.3.1).

During the performed optimizations the methoxy groups remain close to their initial orientation indicating that local minima for both main orientations. Systematic arrangement of all methoxy sites within the analyzed unit-cell allows an estimate of the stabilizing character for the specific interaction type. The systematic orientations of all methoxy groups are referred

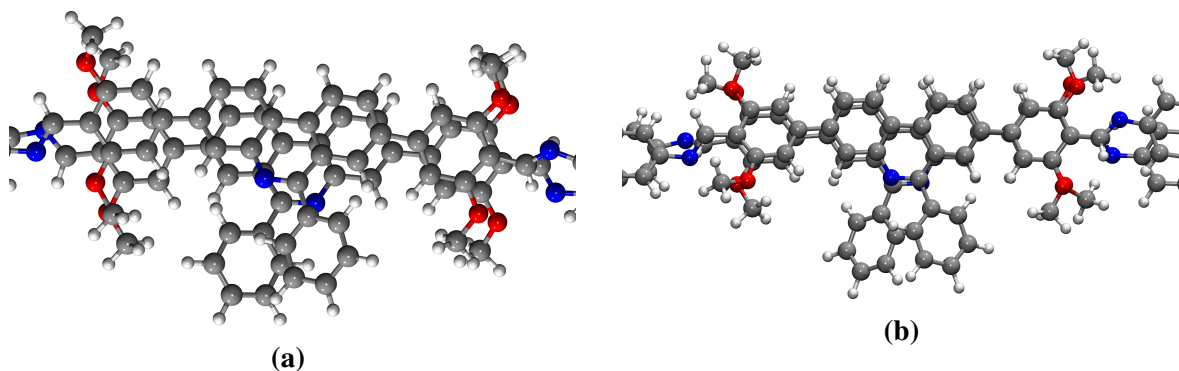


Figure 11.12: The situation when focussing on one pore wall, showing the systematic orientations for the optimized mPP-TAB-COF unit-cells using GPW-DFT on the standard level. a) The situation in the case of ordered methoxy groups. b) The situation in the case of unordered methoxy groups.

Table 11.6: Energetic differences $\Delta E_{m\text{-order}}$ between ordered and unordered structures following Eq (11.4) for various levels of description.

| Description | $\Delta E_{m\text{-order}}$ in [kcal/mol] |
|-------------|---|
| PBE-D3 | 31.38 |
| PBE-D2 | 28.24 |
| BP-D2 | 11.90 |
| GFN-xTB | 3.14 |

to as ordered and unordered. The case of ordered methoxy groups is analyzed using in plane oriented initial structures as shown in Figure 11.13a, with the CH_3 headgroup of both groups remaining closely to the initial structure after optimization. The case of unordered methoxy groups is investigated using a slightly rotated initial structure as shown in Figure 11.13c which further increased the interlayer interaction by tilting of the CH_3 -head towards the oxygen atom of the adjacent group. The energetic differences $\Delta E_{m\text{-order}}$ are obtained by

$$\Delta E_{m\text{-order}} = E_o - E_u \quad , \quad (11.4)$$

with the energies of the optimized structures for the unordered E_u and ordered E_o cases respectively. Additionally, calculations considering PBE-D2 and BP-D2 description are performed for otherwise identical setups. Each description is used for optimization of ordered and unordered initial methoxy-orientations with the results summarized in Table 11.6.

The obtained data show ordered methoxy arrangements to be clearly favored for all uti-

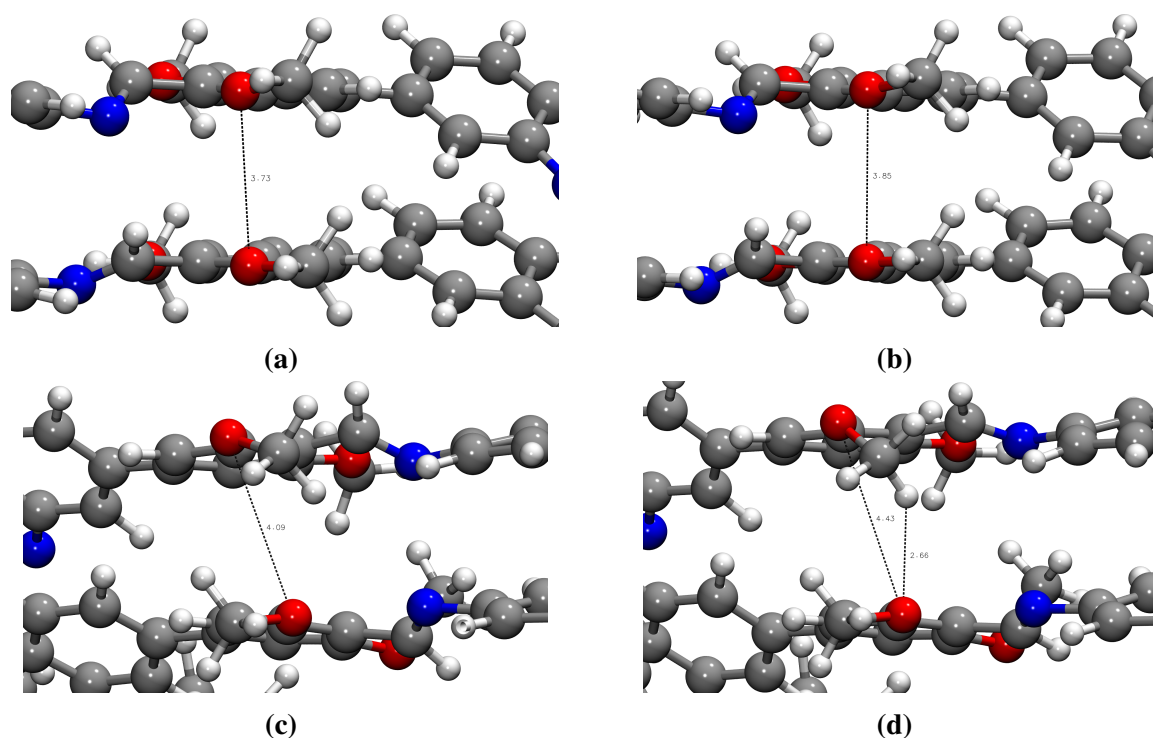


Figure 11.13: The initial and final structures for the methoxy orientation study on the TZV2P-GTH/PBE-D3-level, a) initial orientation for the ordered structures, b) final structure after cell optimization, c) initial orientation for the unordered arrangement, d) final structure after cell optimization.

lized descriptions ($\Delta E \approx 31.0$ kcal/mol for the standard setup). Comparison of the relative displacements between the layers reveals the ordered cases to be more shifted due to less tilted methoxy groups. Surprisingly, the systematic orientation of methoxy groups favoring their inter-methoxy interactions in the unordered case, shows less stabilizing impact than in-plane oriented ones in the ordered case. A possible reason is provided by the proclivity towards larger displacements for ordered structures due to repulsive interactions between the oxygens.

Contrarily to the previously investigated case of the mNP-COF no systematic out of plane re-orientation of adjacent methoxy groups forming hydrogen like bonding structures are observed. However, oxygen-oxygen interactions as arising in the planar oriented methoxy groups are favored due to interlayer displacements.

11.3 Stability Analysis of the PP-TAB-COF Series

The stability of the investigated large pore PP-TAB-COF-series can be investigated by analyzing the energetic landscape for two-layered unit-cells at various interlayer displacements. The unit-cell for the generated COFs is sampled as explained in Section 10.5.1, using CCM implemented GFN-xTB for generation of $E(x, y)$ reference data interpolated by PWs. The obtained energy landscape is compared to GPW-DFT reference calculations for dPP-TAB-COF at the example of a displacement path from AA' \rightarrow AB stacking. Here, 20 structures are cell-optimized using GFN-xTB along the path from eclipsed to staggered arrangement. Subsequently, single point energy calculations were performed on the GPW-DFT standard level for the previously optimized arrangements. The resulting energy profiles are summarized in Figure 11.14 showing qualitatively similar energetic behavior. Consequently an analysis

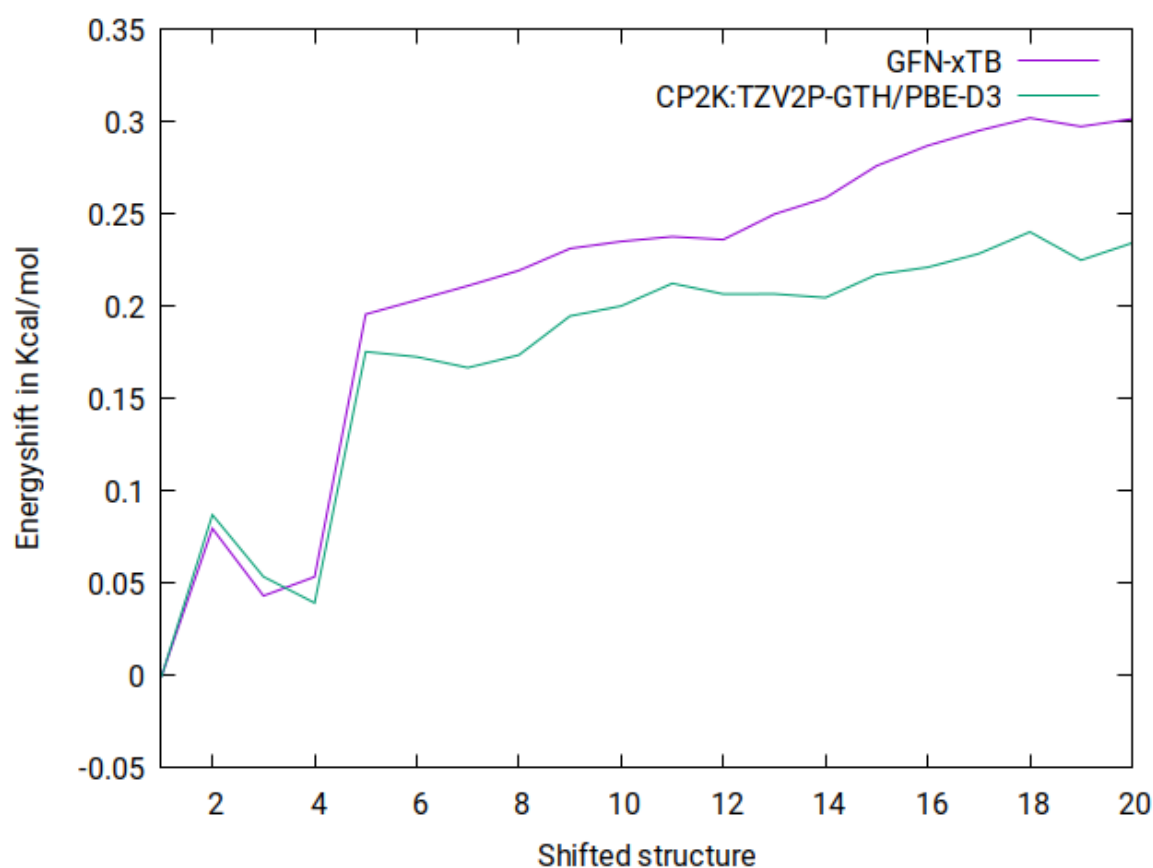


Figure 11.14: The relative energy ΔE with respect to the eclipsed case of the dPP-TAB-COF for twenty distinct structures along the displacement path AA' \rightarrow AB. The fraction of the displacement vector is indicated by the number of the respective structure on the x -axis.

for the full displacement map using GFN-xTB can be expected to yield a sufficiently accurate description. This can be attributed to the GFN-xTB parameterization for non covalent interactions as well as the utilized DFT-D3 dispersion correction for both methods.

11.3.1 Construction of the Eclipsed Structure

Based on the previously conducted analysis of structural arrangements due to component orientations the most probable structures for the PP-TAB-COF, mPP-TAB-COF and dPP-TAB-COF are constructed. The orientations of the PP units as well as the appropriate orientation of imine bonds is to be adjusted for each COF respectively. The S10 configuration of the PP linker was identified to as the most stable throughout the whole study making it the natural choice for further investigations. Additionally, alternating orientation of imine groups is chosen as it is energetically on par with parallel orientations as shown for the case of unmodified ILs in Figure 11.15a and favored in the case of mILs shown in Figure 11.17a. It could be shown that parallel attached methoxy groups are favored for every IL-PP-IL combination using once methoxy modified linkers, making it the natural choice for the mPP-TAB-COF. The resulting structure for the eclipsed configuration is shown in Figure 11.16a. Furthermore parallel oriented methoxy groups are also assumed in the case of the dPP-TAB-COF due to stabilizing effects found in the previous study (compare Section 11.2.6).

Theoretical Investigation: (R. Schuldt)

For all three investigated versions of the PP-TAB-COF-series the initial structures for sampling the unit-cell, are generated by displacing the COMs of both layers by multiples of the unit-cell vectors. The configurations are constructed based on the most probable eclipsed unit-cell arrangements determined for each COF species as summarized in Section 11.3.1.

For the unmodified PP-TAB-COF the global minimum is determined to be located at the AB position indicating generally higher displacements between the layers to be energetically favored. The total difference between eclipsed and staggered cases determined by means of GFN-xTB is given as $\Delta E_{AA' - AB} \approx 220.0$ kcal/mol. Adding one methoxy group generates a local minimum for the eclipsed configuration, indicating that a mixture of both stacking types to be present in a real structure. The relative energetic differences between

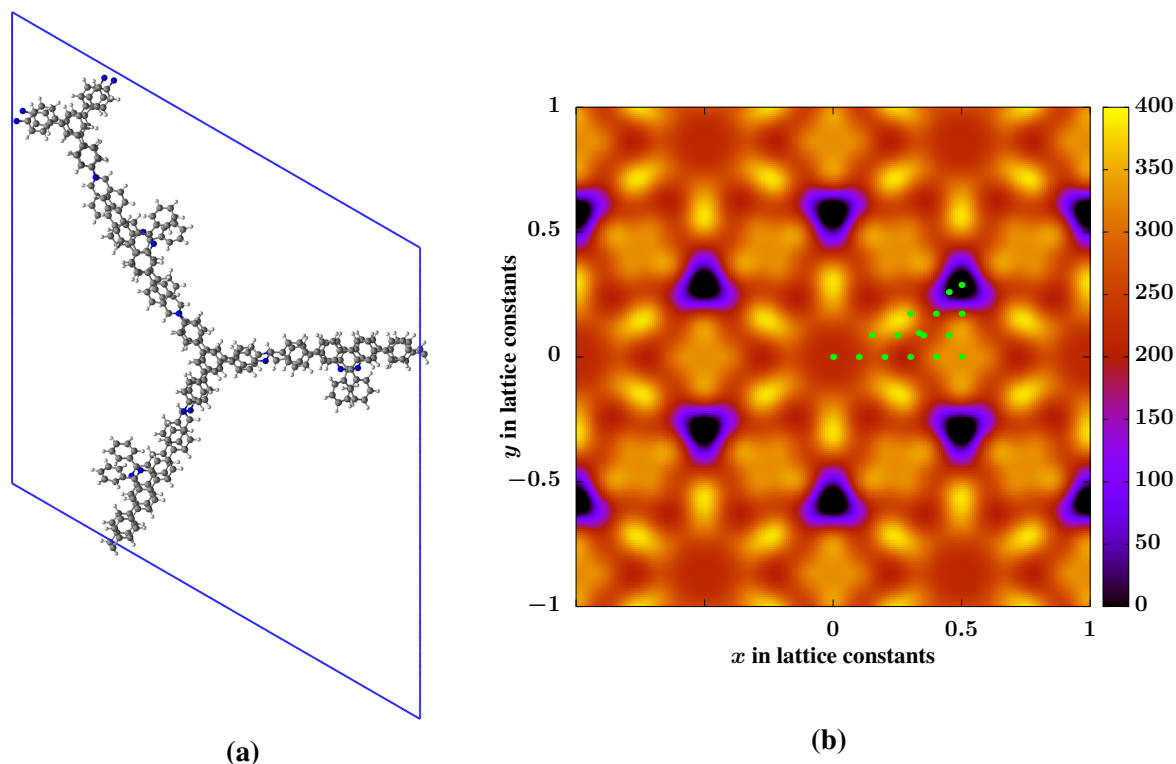


Figure 11.15: Illustration of the most probable eclipsed configuration as well as the potential energy surface determined by PW interpolation for the PP-TAB-COF system. a) The optimized unit-cell after GPW-DFT optimization on the standard level. b) The PES determined following Section 10.5.1, the energies corresponding to the respective colors are given in kcal/mol.

both configurations is determined as $\Delta E_{AA' - AB} \approx 400$ kcal/mol indicating the preferred AB stacking. However, a barrier of $\Delta E_{\text{barrier}} \approx 100$ kcal/mol was found around the eclipsed configuration indicating a slight stabilization of the eclipsed configuration should the layers arrange in AA' stacking. In the case of dPP-TAB-COF, the global minimum is located at the eclipsed position, without the occurrence of side minima, indicating the AA' configuration to be the dominant stacking scheme. The energy difference between both configurations is now $\Delta E_{AA' - AB} \approx 210.0$ kcal/mol in favor of the eclipsed stacking. Furthermore, the AB configuration corresponds to the global maximum for the dPP-TAB-COF showing a clear tendency of the resulting COF to generally favor only slight displacements. In summary, the successive attachment of methoxy groups to the basic structure leads to increased stabilization of the eclipsed stacking scheme for the PP-TAB-COF-series within the used theoretical model. Stabilization of the eclipsed configuration leads to a systematic increase of the effective pore

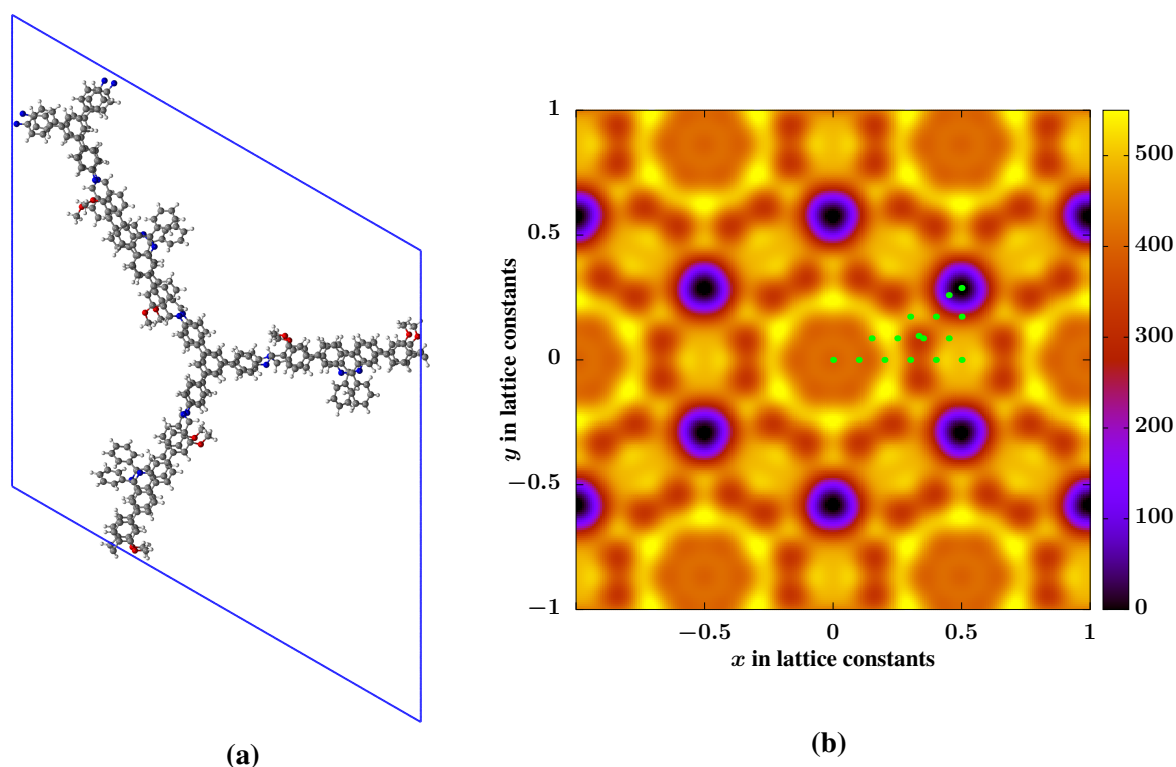


Figure 11.16: Illustration of the most probable eclipsed configuration as well as the potential energy surface determined by PW interpolation for the mPP-TAB-COF system. a) The optimized unit-cell after GPW-DFT optimization on the standard level. b) The PES determined following Section 10.5.1, the energies corresponding to the respective colors are given in kcal/mol.

size \mathfrak{B} with methoxy groups acting as molecular "anchors". Furthermore, the stability towards thermally induced displacements d between adjacent layers is increased, allowing the COFs to retain their crystallinity while increasing the environments temperatures.

Experimental Findings: (S. Emmerling)

Experimental heat measurements in combination with XPRD-measurements conducted in the range of $\Delta T = 20 - 200^\circ\text{C}$ were conducted to investigate the thermal behavior of both synthesized isorecticular series. The thermally most stable structure is found for dPP-TAB-COF which showed no significant structural changes except reversible thermal expansion between the layers up to $T = 200^\circ\text{C}$. The conducted heat measurements showed a clear trend for phase transitions from eclipsed to staggered arrangements with increased environmental temperature for the COFs mPP-TAB and dPP-TAPB that are expected to crystallize in eclipsed

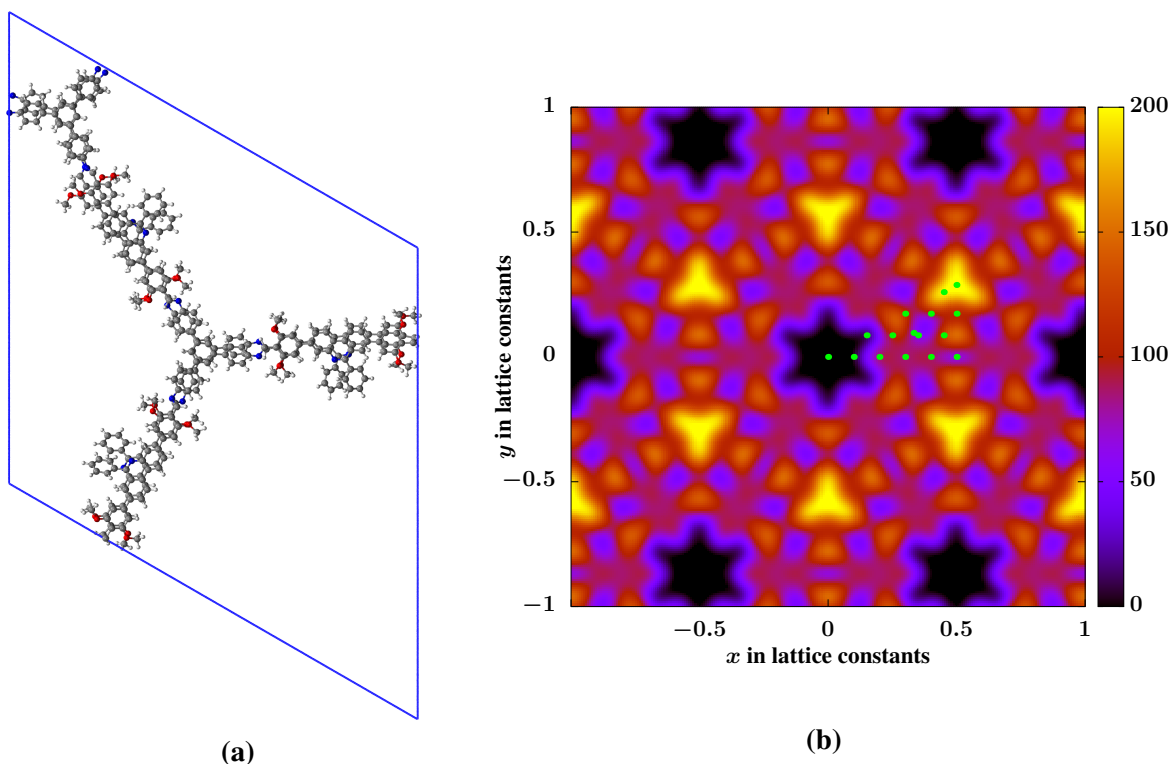


Figure 11.17: Illustration of the most probable eclipsed configuration as well as the potential energy surface determined by PW interpolation for the dPP-TAB-COF system. a) The optimized unit-cell after GPW-DFT optimization on the standard level. b) The PES determined following Section 10.5.1, the energies corresponding to the respective colors are given in kcal/mol.

AA' configuration. The PP-TAB, PP-TAPB and mPP-TAPB systems however showed an increase in layer buckling and reduced crystallinity during the heat measurements as staggered AB like arrangements are already favored over the eclipsed cases. The experimentally observed temperatures for the transitions from AA' \rightarrow AB determined are $T \approx 140^\circ\text{C}$ for dPP-TAPB and $T \approx 120^\circ\text{C}$ for the mPP-TAB COF⁶⁰.

11.3.2 Pore Size Analysis using PSD

For the constructed systems the pore sizes \mathfrak{P} is determined using FRDG generated random paths for the experimentally assumed displacements d between the layers. In experiments a pore size of $\mathfrak{P}_{\text{exp}}^{\text{mPP-TAB}} \approx 4.8$ nm for the methoxy-containing mPP-TAB and dPP-TAB-COF was found which indicates an increase in pore size, supporting the results of the stability analysis, showing nearly perfect AA' eclipsed stacking as the favored scheme. Comparison

11.3 Stability Analysis of the PP-TAB-COF Series

with the methoxy free PP-TAB system, reveals a smaller pore size given as $\mathfrak{P}_{\text{exp}}^{\text{PP-TAB}} \approx 4.2$ nm, corresponding to a reduced pore size indicating the systems proclivity towards staggered and therefore more AB-like stacking schemes. However, an intermediate situation between AA' and AB stacking is assumed corresponding to more shifted AA' stacking which can be attributed to larger interlayer shifts due to low stabilization in the interaction regions.

In the series of larger PP-TAPB-COFs the twice methoxy modified version dPP-TAPB shows the largest pore size with $\mathfrak{P}_{\text{exp}}^{\text{dPP-TAPB}} \approx 5.8$ nm, while for PP-TAPB and mPP-TAPB smaller pore sizes of $\mathfrak{P}_{\text{exp}}^{\text{PP-TAPB}} \approx 4.8$ nm and $\mathfrak{P}_{\text{exp}}^{\text{mPP-TAPB}} \approx 4.9$ nm are found⁶⁰. Here, the observation for both species allows for the conclusion that a more staggered arrangement is observed, whereas the dPP-TAPB-COF maximizes its pore size due to eclipsed AA' stacking.

The experimentally obtained $\mathfrak{P}_{\text{exp}}$ are crosschecked for further interpretation by means of the geometric pore analysis using the PSD algorithm (compare Chapter 7). Using the PSD algorithm for FRDG random displacement paths constructed based on structures from X-ray powder diffractograms and provides an alternative way of determining the effective pore size. The geometrically obtained pore size $\mathfrak{P}_{\text{PSD}}$ is determined by the biggest sphere able to fit through the spanned pore arrangement for each layer averaged over all twenty layers as determined by consideration of five adjacent ones. The obtained results summarized in Table 11.7 perfectly follow the trend of the experiment showing the increase in pore size from $\mathfrak{P}_{\text{PSD}}^{\text{PP-TAPB}} \approx 4.05$ nm for PP-TAPB to $\mathfrak{P}_{\text{PSD}}^{\text{dPP-TAPB}} \approx 5.90$ nm for dPP-TAPB and, thus, further support the interpretation of AA' stackings being favored by methoxy attachments.

Table 11.7: Calculated pore sizes $\mathfrak{P}_{\text{PSD}}$ based on the PSD geometric pore analysis in comparison with experimental values $\mathfrak{P}_{\text{exp}}$ for the investigated PP-TAB and PP-TAPB species.

| COF-species | $\mathfrak{P}_{\text{exp}}$ in Å | $\mathfrak{P}_{\text{PSD}}$ in Å |
|--------------------|----------------------------------|----------------------------------|
| PP-TAB (AA' → AB) | 42.2 | 40.5 |
| dPP-TAB (AA') | 48.0 | 46.2 |
| PP-TAPB (AA' → AB) | 47.5 | 45.6 |
| dPP-TAPB (AA') | 58.0 | 59.2 |

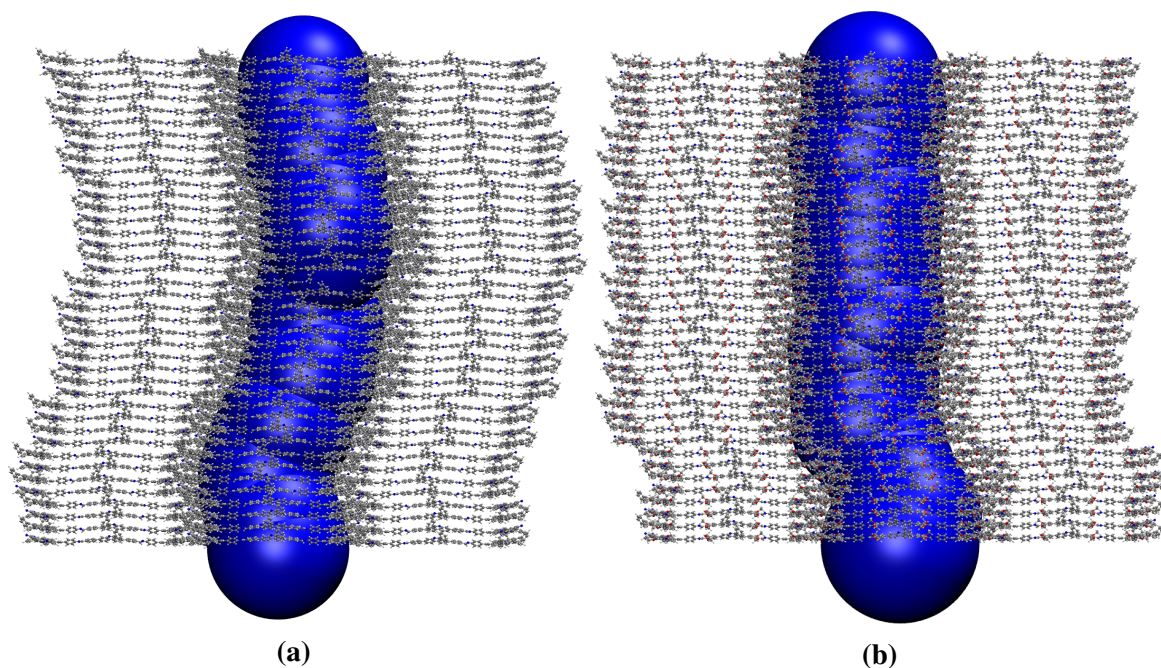


Figure 11.18: Illustration of the FRDG-generated random paths used for comparison with the experimental values for the PP-TAB-COF-series. Always two layers generated based on the assumed stacking scheme were randomly displaced using an interlayer displacement of $d = 3.0 \text{ \AA}$. Blue spheres indicate the pore size of the twenty considered layer-pairs used for the determination of $\mathfrak{P}_{\text{PSD}}$. a) The model situation in the case of the unmodified PP-TAB-COF, assuming stronger shifts between the layers simulating an admixture of AA' and AB stacking. b) The model situation in the case of the twice methoxy modified dPP-TAB-COF, assuming AA' stacking.

Layer Buckling and Displacements

During heat measurements the observed unit-cells determined from XPRD measurements showed a decrease in the in-plane cell vectors as well as a reduction of the interlayer distances indicating layer buckling as well as shifting. From the conducted experiments for mPP-TAB and dPP-TAPB-COFs both effects are observed without initiation period, allowing for the interpretation, that shifting does not require initial buckling and buckling does not require initial shifting. For further analysis of buckling and shifting, the optimized hexagonal unit-cells used for the screening of the energetics of interlayer displacements in Section 11.3 are used to elucidate the relation between interlayer-displacement and in-plane lattice parameters $a = b$. Reduction of the in-plane lattice vectors' lengths $a = b$ for optimized geometries after displacement, indicates a reduction of the layers extension. This is only possible when

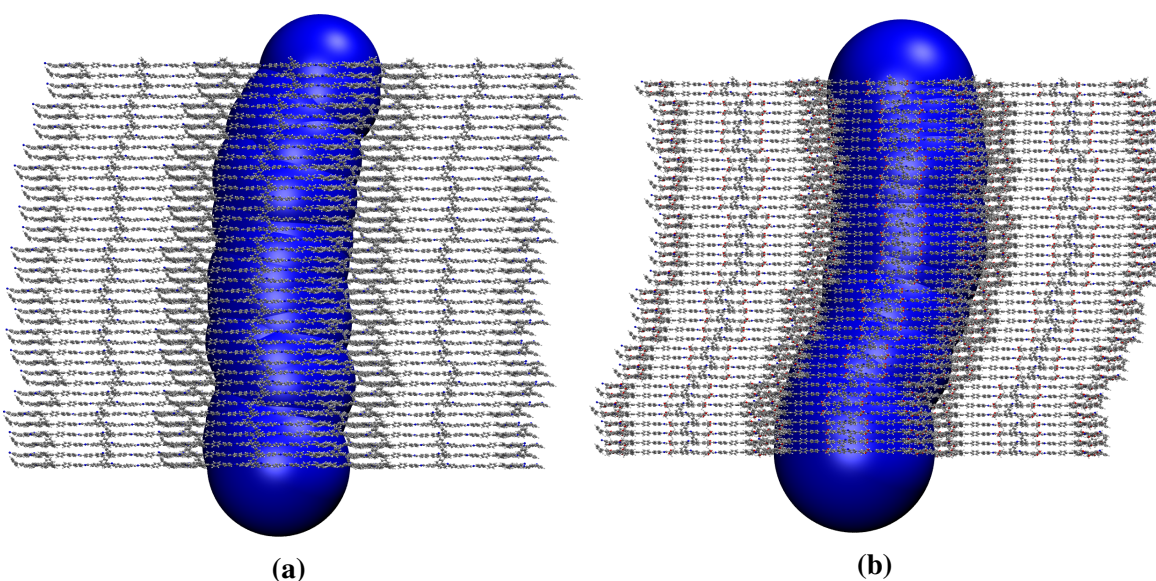


Figure 11.19: Illustration of the FRDG-generated random paths used for comparison with the experimental values for the PP-TAPB-COF-series. Always two layers generated based on the assumed stacking scheme were randomly displaced using an interlayer displacement of $d = 3.0 \text{ \AA}$. Blue spheres indicate the pore size of the twenty considered layer-pairs used for the determination of \mathfrak{P} . a) The model situation in the case of the unmodified PP-TAB-COF, assuming stronger shifts between the layers simulating an admixture of AA' and AB-like stacking. b) The model situation in the case of the twice methoxy modified dPP-TAPB-COF, assuming AA' stacking.

layer buckling occurs as in-plane bond distances are not affected by interlayer interactions as can be concluded from the investigated planar systems Section 10.6.1. An example structure showing both effects as observed in the PES screening for dPP-TAB-COF for the optimized (0,2,0)-configuration is shown in Figure 11.20. The data summarized in Table 11.8 show the evolution of the lattice constants for the dPP-TAB-COF after cell optimization for the displacements used for PES interpolation. The lattice constants were extracted from the GFN-xTB optimized unit-cells used for validation of the methods applicability as presented in Section 11.3. The smaller the length of in-plane cell vectors $a = b$, the higher the estimated layer amount of layer buckling.

The data reveal a clear connection between displacement d along the connection from AA' to C type stacking and the reduction of the in-plane lattice constants $a = b$ as well as a the related c -lattice constant. Furthermore, the obtained results for shifted initial structure favor the experimentally estimated interpretation of buckling being favored by increased interlayer

displacements, which is associated with a reduction of the cell vector lengths. In contrast to the observations for planar species made in Section 10.6.1, buckling is seemingly favored for non-planar systems as investigated in this study. Further data for the mPP-TAB-COF are summarized in Table G.8 showing similar trends.

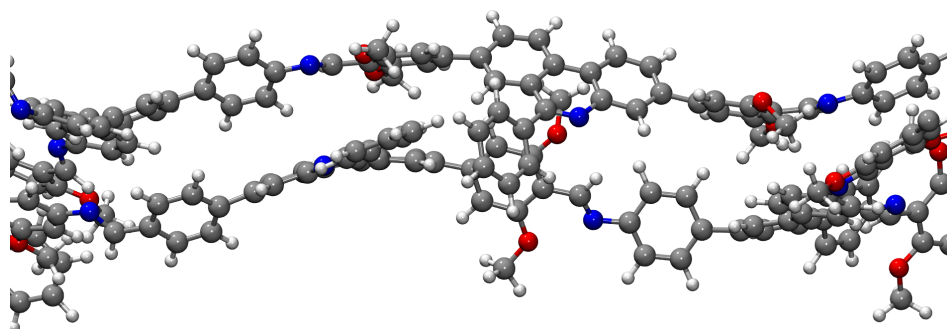


Figure 11.20: The side view onto an example structure showing layer buckling for the GFN-xTB optimized two-layered unit-cell for the dPP-TAB-COF system with (0.2,0) displacements.

Table 11.8: Lattice constants extracted from the optimized unit-cells as obtained during the sampling of the displacement energy maps for the dPP-TAB-COF system.

| x in $ a $ | y in $ b $ | a_{opt} in Å | c_{opt} in Å |
|------------|------------|-----------------------|-----------------------|
| 0 | 0 | 58.98 | 7.43 |
| 0.1 | 0.1 | 55.86 | 7.15 |
| 0.1 | 0 | 57.67 | 7.11 |
| 0.2 | 0.1 | 56.99 | 6.44 |
| 0.2 | 0.2 | 57.71 | 6.40 |
| 0.2 | 0 | 57.70 | 6.71 |
| 0.277 | 0.111 | 56.23 | 6.12 |
| 0.3 | 0.1 | 57.86 | 4.60 |
| 0.3 | 0.2 | 57.34 | 5.66 |
| 0.3 | 0.3 | 59.08 | 5.70 |
| 0.3 | 0 | 55.11 | 6.35 |
| 0.333 | 0.333 | 59.13 | 5.80 |
| 0.4 | 0.1 | 57.88 | 6.11 |
| 0.4 | 0.2 | 58.47 | 4.93 |
| 0.4 | 0 | 54.27 | 6.45 |
| 0.5 | 0 | 51.83 | 6.36 |

11.4 Fluorinated-Linker Design

The previous study of interlayer interactions for the PP-TAB-COF-series revealed the explanatory power of the theoretical machinery used by explaining the working mechanism of the methoxy modifications for the PP-TAB-COF series. During the analysis of the phenylphenanthridine (PP) linkers, T-shaped stacking of the phenyl groups ranging into the pore was observed as an alternative to the parallel case. However, both configurations were determined to be energetically similar with T-shaped stacking being favored in the S10 case by only $\Delta E_{\text{T}}^{\text{S10}} = 0.26$ kcal/mol for the conducted block analysis (compare Section 11.2.3). Contrarily, in the S11 case the energetic differences between both configurations was determined to be $\Delta E_{\text{T}}^{\text{S11}} \approx 50.0$ kcal/mol showing a clear favoring of the T-shape even though the S10-T remains the energetically most stable configuration. The T-shape phenomenon indicates a strong interlayer interaction and its usability for stabilization is analyzed by using fluorinated phenylphenanthridine (F-PP) units as shown in Figure 11.21, which are known to interact via much larger quadrupole stabilization increasing the T-shape stabilization. Consequently, the effect of fluorinated linkers stacked in T-shape is investigated for its ability to increase stability of the eclipsed stacking types. The geometry of the F-PP component optimized on the PBE-D3/def2-TZVPP level is shown in Figure 11.21a, whereas its overall electrostatics are illustrated in Figure 11.21b showing similar arrangements as in the unfluorinated case. However, the molecular electrostatic potential V_C in the near vicinity of the fluorinated phenyl group is acting much stronger on a potential probe charge q^+ as can be seen from the higher number of potential lines in its immediate vicinity. This is caused by the charge redistribution within the compound inverting the polarity of the quadrupole while increasing the charge differences due to pulling electrons from the rings center outwards to the fluorine atoms.

Analysis of the various configurations for the F-PP units in the gas phase gives a first insight into favorable interlayer interactions in the eclipsed case. For this purpose geometry optimizations are conducted using TURBOMOLE on the PBE-D3/def2-TZVPP level with the stability of each arrangement estimated by calculation of the corresponding interaction energies following Eq (11.1). The nomenclature of the data summarized in Table 11.9, follows the scheme introduced for the unmodified PP-linkers in Section 11.2.1.

The stabilizing character of the components interactions is estimated by calculation of E_{int}

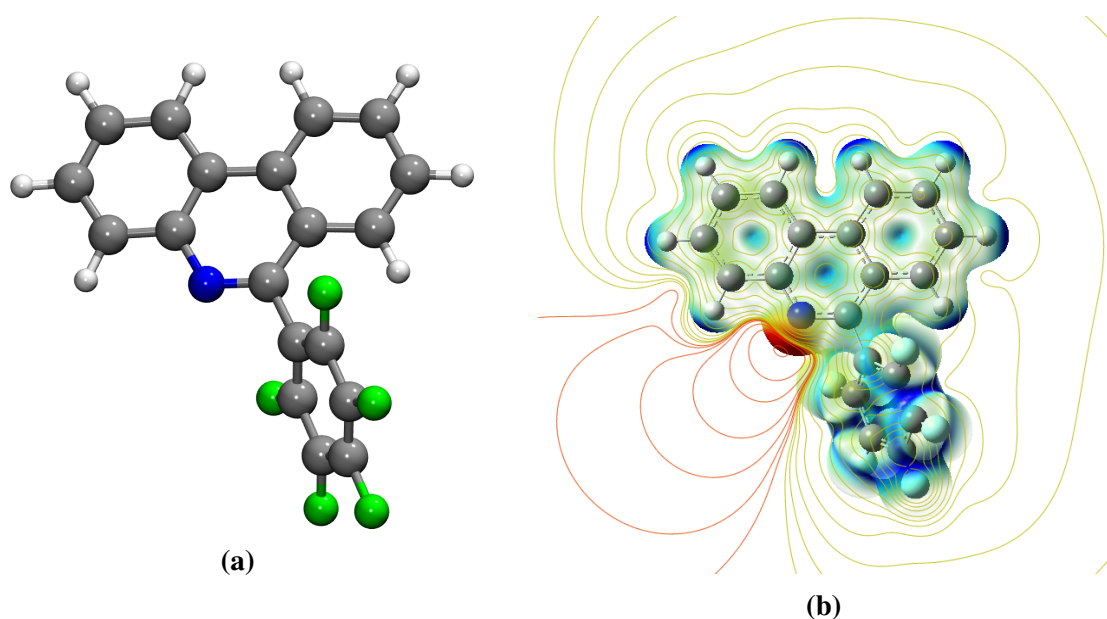


Figure 11.21: Overview of the F-PP-linkers structural properties. a) The geometrical arrangement showing the increased volume of the fluorinated phenyl group. b) The electrostatics illustrated by mapping V_C onto the density iso-surface for $\rho_{\text{iso}} = 0.05$ a.u., with the colors indicating the electrostatic potential arising from the surrounding.

Table 11.9: Interaction energies E_{int} calculated following Eq (11.1), for the main two component F-PP-linker arrangements. Calculations were performed on the PBE-D3/def2-TZVPP level.

| Component Arrangement | E_{int} in kcal/mol |
|-----------------------|------------------------------|
| F-S10 | 14.04 |
| F-S11 | 15.02 |
| F-S10-T | 10.78 |
| F-S11-T | 11.62 |
| F-D10 | 12.76 |
| F-D11 | 12.45 |

with the data summarized in Table 11.9. From the summarized data, the F-S11 configuration is identified to be the energetically most favorable followed by F-S10 with the optimized geometry shown in Figure 11.22. Comparing the energetics of parallel (F-S10) and T-shaped (F-S10-T) arrangements, shown in Figure 11.22a and Figure 11.22b respectively, reveals parallel stacking to be favored. In the case of F-S11 arrangements the observed in plane displacements after optimization is larger than expected in the case of the full COF in the eclipsed case, which makes the slightly less energetically favorable F-S10 arrangement the most probable.

However, the whole analysis of fluorinated components showed strong out of plane rotations reducing the overall comparability to the full COF situation. These observations are accounted for by a separate analysis for the constructed F-PP-COF system which is described in the following section.

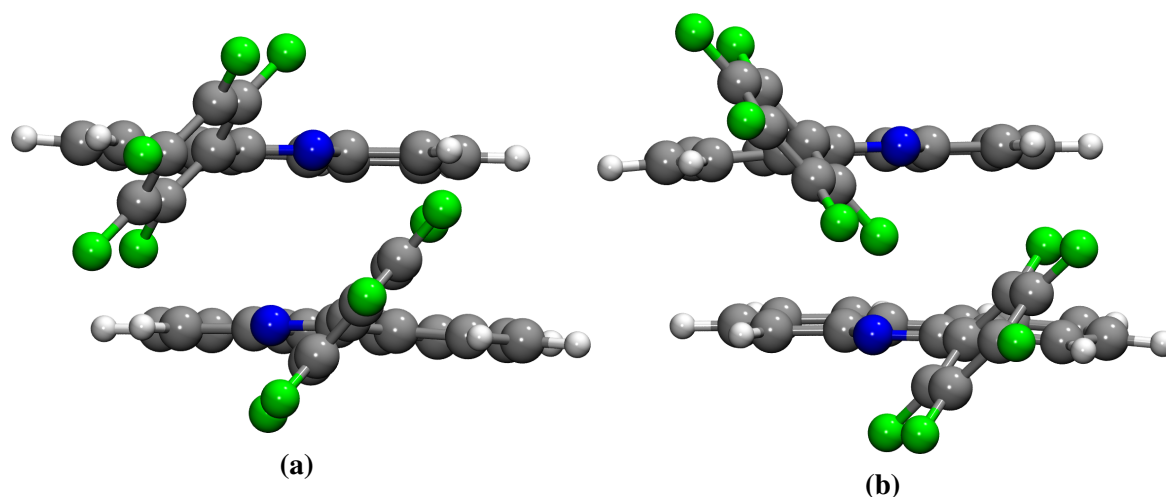


Figure 11.22: GFN-xTB optimized two component arrangements for the S10-configuration, illustrating both orientations for the fluorinated phenyl groups. a) The case of parallel stacked fluorophenyl linkers (F-S10). b) The case of T-shaped fluorophenyl linkers (F-S10-T).

11.4.1 T-Shaped Analysis for the F-PP-COF System

As the gas-phase component analysis revealed significant drawbacks for interpretation due to strong out of plane rotated F-PP linkers, the effects of T-shaped phenyl orientation are investigated by construction of a minimal COF. The chosen design is based on the previously introduced C3-building unit, linked via direct imine bonds to the PP-units. The two-layered unit-cells for the eclipsed structures summarized in Figure 11.23, are constructed based on the previous analysis with alternating imine bonds as well as S10 arrangements for the PP-linkers. To interpret the impact of fluorination on the overall stability of T-shaped stacking, two unit-cells based on PP as well as F-PP components are investigated respectively. The unit-cells after optimization on the GPW-DFT standard level for the PP-COF system is illustrated in Figure 11.23a whereas the F-PP-COF system is shown in Figure 11.23b.

The impact of the PP-linkers phenyl arrangements are investigated by creating two distinct

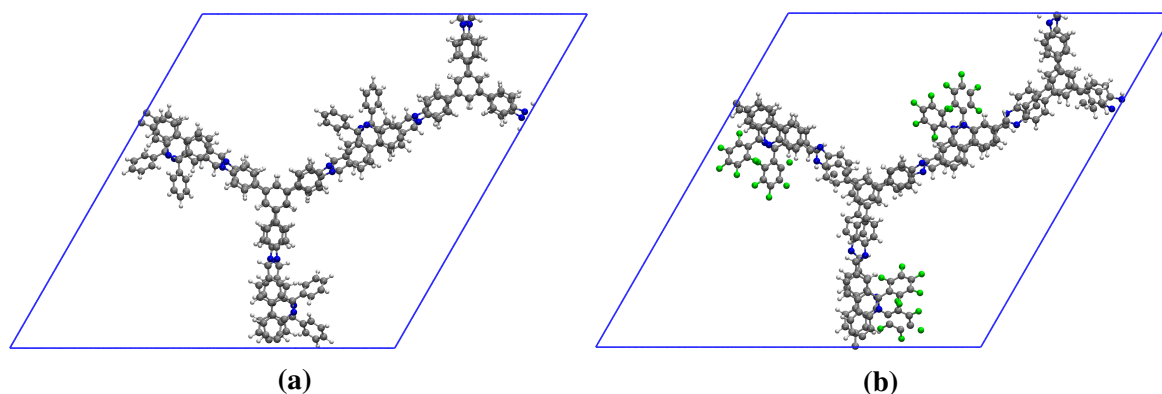


Figure 11.23: Example structures showing the T-shaped arrangements within unit-cells for the investigated PP-COFs, optimized by means of GPW-DFT on the standard level (compare Section 10.4.5). a) The optimized PP-COF unit-cell in the eclipsed configuration. b) The optimized F-PP-COF unit-cell in the eclipsed configuration.

configurations with all three phenyl pairs within the unit-cell oriented either in parallel or T-shape. Cell-optimizations were performed for parallel and T-shape oriented linkers for both systems with energies calculated by means of GPW-DFT using the standard setup. Both phenyl orientations remain stable with only slight readjustments. Therefore, the main molecular rearrangements during optimization are limited to layer buckling and slight adjustments of the interlayer distance for both investigated COF species. The stabilization was calculated in analogy to Eq (11.3) for the whole unit-cells. The data summarized in Table 11.10 obtained from GPW-DFT, indicates the T-shaped orientation to be slightly favored for the unmodified PP-COF by $\Delta E_T \approx 13.28$ kcal/mol. This means an energetic difference of $\Delta E_T \approx 4.4$ kcal/mol per T-shaped PP-component predicting both orientations to be equally probable. Furthermore, analysis of the unit-cell parameters after optimization as summarized in Table 11.11, revealed smaller in-plane lattice parameters for the optimized para arrangement for both investigated COFs. This result suggests T-shaped stacking for PP components as a possibility for reducing layer buckling which is associated with a smaller in-plane unit-cell size as shown in Section 11.3.2. Analogous investigation for the F-PP-COF reveals T-shaped stacking to be favored by $\Delta E_T \approx 180.0$ kcal/mol corresponding to an average stabilization of $\Delta E_T \approx 60.0$ kcal/mol per T-shape oriented F-PP-linker combination. Reference calculations on the GFN-xTB level are summarized in Table G.1 and show qualitatively similar results for the energetic differences. This result indicates the underlying interactions to be weighted similarly

Table 11.10: Energy differences ΔE calculated following Eq (11.3), for T-shaped stacking F-PP-linkers for the F-PP-TAB-COF arrangements using GPW-DFT with the standard setup.

| Component Arrangement | ΔE_T in kcal/mol |
|-----------------------|--------------------------|
| PP-linker | 13.28 |
| F-PP-linker | 180.07 |

Table 11.11: Unit-cell parameters as obtained after optimization by means of GPW-DFT using the standard setup, for the investigated PP-COFs as obtained for parallel and T-shaped initial orientations.

| Component Arrangement | $a = b$ in Å | c in Å |
|-----------------------|--------------|----------|
| S10 | 42.10 | 8.04 |
| S10-T | 44.55 | 8.01 |
| F-S10 | 41.68 | 8.08 |
| F-S10-T | 43.86 | 8.18 |

for both methods. Here, the energetic differences ΔE_T determined for F-PP-COFs by means of GFN-xTB are determined to be of approximately identical magnitude to the GPW-DFT case for the optimized unit-cells.

In summary, GPW-DFT identifies S10-T orientations to be preferable when compared to S10 for the investigated PP-COF as well as F-PP-COF systems. In the case of the F-PP-COF, the reorientation of one F-PP component system from parallel to T-shape is favored by $\Delta E \approx 60.0$ kcal/mol. Furthermore, optimizations of the unit-cells with T-shape initial orientation showed an increase of the in-plane lattice constants, indicating a reduction of layer buckling. Both observations give rise to the assumption, that T-shaped fluorinated linkers are not only energetically favorable but also lead to an increase in pore size by maximizing the layer extension in the eclipsed case.

11.5 Stability Analysis of the F-PP-TAB COF Series

The main target property of the introduced F-PP components is the maximization of \mathfrak{F} , by increased stabilization of the eclipsed stacking motif and reduction of interlayer displacements d . For further analysis of the possible impact of fluorinated linkers onto the stability of COFs the displacement behavior of the fluorinated isorecticular PP-TAB-COF-series is investigated due to the possibility of direct comparison. Based on the previously conducted studies the most probable arrangements for the eclipsed AA' configuration are constructed also for the fluorinated case assuming the corresponding interactions to be qualitatively similar to the unfluorinated series.

11.5.1 Theoretical Investigation of Displacements for F-PP-TAB-COFs

The most probable eclipsed configuration is constructed based on the components analysis in analogy to the unmodified versions of the PP-TAB-COF series. The used ILs as well as the modified mIL versions are stacked in the alternating arrangement (compare Section 11.3.1), as the immediate surroundings are identical in the F-PP series. The F-S10-T configuration was chosen based on the previous investigation of the unfluorinated series showing S10 to be the generally most stable configuration. Furthermore, T-shaped stacking was assumed for all systems due to the results of the previous analysis for the minimal example F-PP-COF (compare Section 11.4.1) showing it to be energetically favorable by an astonishing $\Delta E \approx 180.0$ kcal/mol using GPW-DFT with the standard setup. Comparison of the data obtained from DFT shown in Table 11.10 reveals identical observation as the GFN-xTB data summarized in Table G.1 showing the direct transferability between the methods. Due to the observed behavior in the previous study for the F-PP-COF, GFN-xTB is chosen for the further analysis of displacement behaviors.

The GFN-xTB optimized unit-cells for the eclipsed (AA') configurations constructed for the F-PP-TAB COF is shown in Figure 11.24a, for the singly methoxylated F-mPP-TAB-COF it is shown in Figure 11.24a and for the doubly methoxylated F-dPP-TAB-COF shown in Figure 11.24a. For all three investigated versions of the F-PP-TAB-COF-series, the initial structures for sampling of the unit cell, are generated by displacing the center of masses (COMs) of both layers by multiples of the unit-cell vectors with respect to each other.

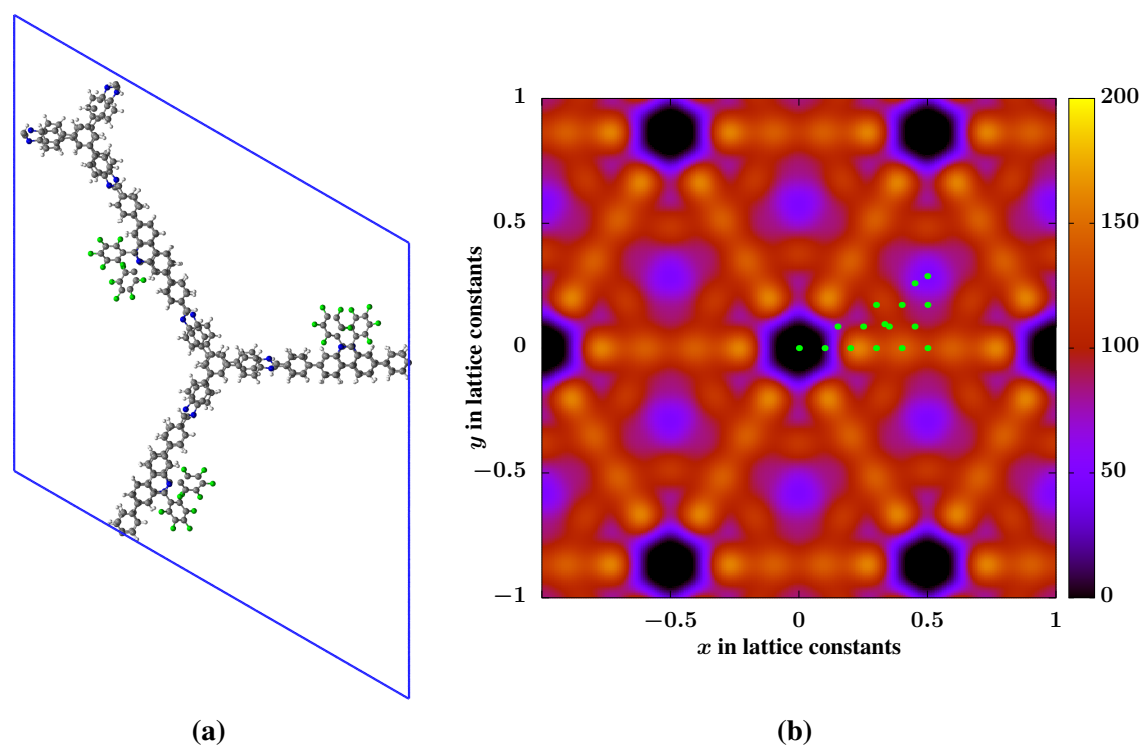


Figure 11.24: Optimized unit-cell and Potential energy surfaces determined by PW interpolation for the F-dPP-TAB-COF species. a) Top view onto the unit-cell. b) The PES determined following Section 10.5.1 with colors indicating the relative energies in kcal/mol.

The configurations were constructed based on the most probable unit-cell arrangements determined for each COF species as summarized in Section 11.5.1. The previously conducted theoretical study, predicts T-shape orientations to occur for eclipsed configurations for all three investigated species, allowing for its usage as a design tool for large pore COFs by favoring the formation of highly ordered eclipsed configurations. Consequently, it is expected that substituting F-PP components into the isoreticular PP-TAB-COF series will further increase its stability. A stability analysis is conducted for the fluorinated species in analogy to the PP-TAB-COF-series as previously illustrated in Section 11.3. The potential energy displacement maps for the F-PP-TAB species were obtained as described in Section 10.5.1. The PES obtained for the F-PP-TAB-COF shows its global minimum for the eclipsed AA' configuration with the occurrence of a side minimum for the staggered AB configuration. However, the global minimum region around the AA' structure is quite enlarged indicating the possibility of a multitude of rearrangements within its immediate vicinity. The result has to be interpreted with caution as the (0,0) position is constructed using T-shape arrangements for all three

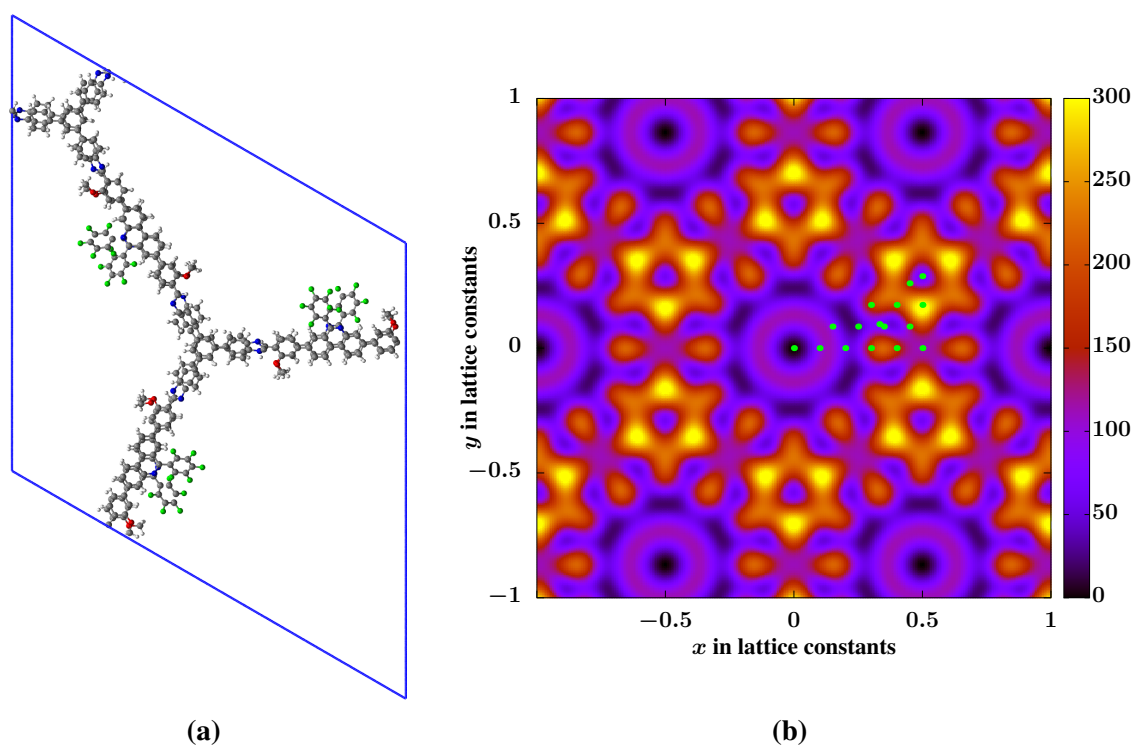


Figure 11.25: Optimized unit-cell and Potential energy surfaces determined by PW interpolation for the F-dPP-TAB-COF species. a) Top view onto the unit-cell for the singly methoxy-modified F-mPP-TAB-COF with cell boundaries indicated in blue. b) The PES determined for the singly methoxy-modified mPP-TAB-COF.

species respectively, leading to a major consideration for possible interpretations and structural prediction for the F-PP-TAB-COF-series. However, T-shaped stacking can only occur for eclipsed configurations build by combining F-PP-linkers in the S configuration. Consequently, the eclipsed configuration might be energetically the most favorable for all three species, but does not necessarily occur for all species in the experiment. Specifically in the case of the unmodified F-PP-TAB-COF an extended side minimum for a displaced configuration at (0.1,0) as well as a local minimum in the staggered AB configuration is observed, indicating a concurrence between both major stacking schemes. The situation is similar to the previously discussed mPP-TAB-COF shown in Figure 11.16b. In the case of F-mPP-TAB-COF, the global minimum is in the eclipsed configuration with relatively high energetic differences for small displacements within the immediate surroundings, leading to barrier of around $\Delta E \approx 100$ kcal/mol. Furthermore, the secondary side minimum is energetically highly separated by a barrier of about $\Delta E_{\text{barrier}} \approx 300.0$ kcal/mol. Consequently, based on both observations the eclipsed stacking is already determined to be strongly favored due to the

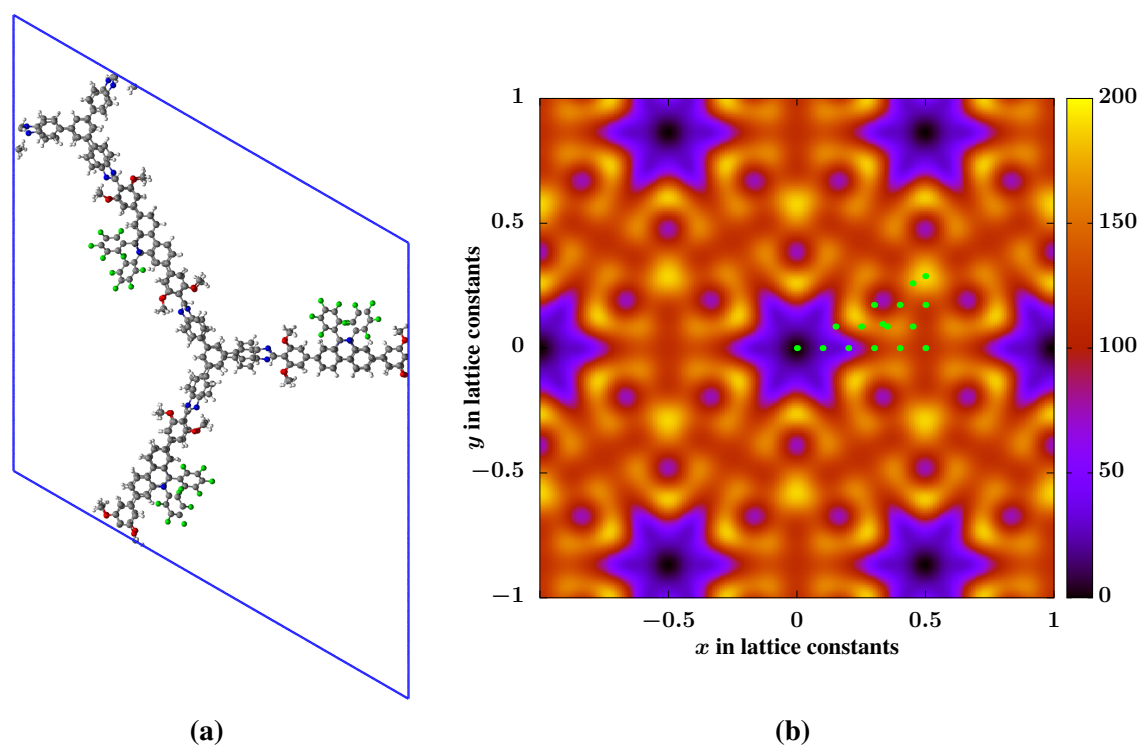


Figure 11.26: Optimized unit-cell and Potential energy surfaces determined by PW interpolation for the F-dPP-TAB-COF species. a) Top view onto the unit-cell for the twice methoxy-modified F-dPP-TAB-COF with cell boundaries indicated in blue. b) The PES determined for the twice methoxy-modified dPP-TAB-COF.

stabilizing interactions accumulating from methoxy groups as well as fluorinated phenyls.

This COF is expected to provide enough stability for observation of T-shaped stacking acting as a promising tool for modifying a COFs stability. This prediction is reasoned by methoxy groups being known to already favor the eclipsed configuration, allowing for the possible reorientation of the F-PP linkers during the crystallization process. In the case of F-dPP-TAB-COF, no secondary minimum is observed besides the global one located at the eclipsed AA' configuration at (0,0). This COF is expected to be of even increased stability when compared to the dPP-TAB-COF due to T-shape stabilization in accordance with the previously conducted study (compare Section 11.3).

For interpretation of the resulting predictions for stability by displacements of layers with respect to each other it is important to note, that only a small number of component orientations for the COFs are considered. All configurations are initially constructed using the F-S10 orientation, which is chosen based on the analysis of the AA' structures. Therefore, possible interactions between F-PP components and phenyl groups within subsequent layers,

which are of significance for larger interlayer displacements, might be neglected. Furthermore, the possibility of flaws in the utilized description has to be considered as no benchmark possibilities to higher level theory exists.

11.6 Lessons Learned for the Design of Large-Pore COFs

The presented studies showed the possibility of rational design of two-dimensional COFs targeting large effective pore sizes by using interlayer interactions at the example of the unmodified PP-TAB-COF-series and the fluorinated F-PP-TAB-COF-series. A suitable model description for the targeted eclipsed AA' configuration was constructed based on a hierarchical component analysis allowing for the determination of the most probable structural orientation. GPW-DFT was used to provide insight into the analysis of structural arrangements for PP-TAB-COFs, refining the experimentally obtained structural models. From the obtained insights the underlying effects for favorable interlayer interactions were obtained by investigating the impact of methoxy groups on the components and their resulting stability, allowing for an explanation of the experimentally observed effects. The role of molecular "anchors" of methoxy groups as estimated from the experiment could be explained by the conducted analysis of interlayer displacement behavior. Furthermore, the connection between interlayer displacements and possible layer buckling was observed for the investigated PP-TAB as well as PP-TAPB species, securing the experimental estimate⁶⁰. The methods introduced in Chapter 10 were successfully applied for investigation of a newly synthesized COF. Based on the analysis of the COFs components, their large-pore character could be explained and the resulting thermal behavior was explainable by the subsequent study of in-plane interlayer displacements. Additionally, experimentally obtained pore sizes and the interpretation of layer buckling as the result of interlayer shifts, were cross-validated by means of PSD analysis for FRDG generated random paths.

The observation of the T-shape phenomenon during the investigation of interactions arising between PP-components, the fluorinated F-PP component was investigated. Gas phase calculations showed no increase in interaction energies and identified T-shaped arrangements to be even disfavored when compared to parallel ones. However, T-shaped stacking was

11.6 Lessons Learned for the Design of Large-Pore COFs

observed to be favored for the investigated F-PP-COF, where in-plane movement and rotation of the adjacent layers are limited. In summary it can be concluded that gas phase component interactions are only comparable to the full COF case, when the alignment is not changed significantly during optimization.

Motivated by specifically favoring T-shaped stacking, the effects of substitution of PP-components by fluorinated F-PP ones were investigated. In the framework of modification of components, fluorination of the PP-linkers marks adjustment of the second building unit, offering a potential alternative or addition to the previous approach of modifying the first building unit (the ILs).

The impact of T-shaped stacking onto the overall stability in a model COF system was analyzed by means of GPW-DFT on the standard level, predicting it as the favored stacking scheme for the F-PP-COF system. T-shaped arrangements were determined to be energetically favored by $\Delta E_{T\text{-para}} \approx 60.0$ kcal/mol for each F-PP linker. Based on this indications, the utilized eclipsed structures for investigation of the PP-TAB-COF-series were directly modified by fluorinating the PP-linker units, creating the most probable eclipsed structures for the F-PP-TAB-COF (compare Section 11.5.1). The stability analysis by means of PW-interpolation of the created PES, revealed the eclipsed T-shape configuration to be the energetically most favorable for all three F-PP-TAB-COF systems. The global minimum was found in the eclipsed (0,0) configuration for all fluorinated systems. However, broadening of the potential sink as well as side minima being diminished only after attachment of additional methoxy groups, leads to the conclusion that the stabilizing effects of F-PP units in T-shape alone are not sufficient to significantly increase Φ . However, the phenomenon of T-shaped stacking in F-PP linker arrangements is estimated to be significant for F-mPPb-TAB-COF as well as F-dPP-TAB-COF, where it favors additional layer extension as well as stability of the eclipsed AA' configuration therefore increasing Φ .

Based on the investigations in Chapter 10 for maximizing Φ by favoring the eclipsed AA' stacking scheme, the manipulation of the third building unit namely the C3-linker, marks the natural step for completing the study. Investigations leading to synthesis of the N3-COF system showed that substitution of nitrogen into the central ring of the C3-linker leads to planarity of the corresponding linker unit²⁴⁷. Doing so reduces out of plane rotation of the

11 Theoretical Investigations of Interlayer Interactions for Large-Pore COFs

adjacent phenyl groups and leads to increased planarity of the resulting COF. Consequently, a reduction of the COFs proclivity towards layer buckling due to increased planarity as well as reduction of in-plane displacements can be expected due to stabilizing interactions of the N3-linker as previously shown for the V-COF system (compare Section 10.6.1).

Therefore, increased planarity of the COF as well as aryl-aryl interactions in the eclipsed configuration might offer an alternative way of increasing the effective pore size, stacking order and crystallinity by interlayer interactions. The indication of a suitable follow up analysis by the presented methodology is given by the results for the N3-COF in Section 10.6.1 as well as the presented study for the PP-TAB-COF series within this chapter. Recently, a COF-system with exceptionally large pore size based on the N3-linker was successfully synthesized and further indicates the plausibility of increasing pore sizes by use of N3-units due to aryl-aryl interactions as well as enforced planarity²⁵⁸.

Chapter 12

Summary

Part I of this thesis, gave a detailed overview of the theoretical formalisms used for the conducted studies. Starting from the fundamentals of density functional theory and its generalization for periodic systems, the redefinition of the polarization as a bulk property was reviewed. Additionally, the extension of the DFT formalism by treating external electric fields using the Berry-phase approach was explained. Furthermore, a detailed description on the hierarchical improvements of the extensively used density functional tight binding methods was provided. The mathematical basic considerations for the developed analysis tools and a detailed description of their respective work flows was provided in Part II. Well founded on the previous reviews and descriptions, the methods' application was reported in Part III. Here, the methods were firstly applied on the structural investigation of SAM-metal interfaces and subsequent simulations of field evaporation. Secondly, an investigation targeting the theoretical description of COFs and the applicability of the created tools was conducted for an in-depth characterization of planar and non-planar example systems. Based on GPW-DFT analysis and the constructed tools, interlayer interactions potentially maximizing the pore size in two-dimensional COF systems were analyzed. Insights gained on the role of specific components were used to confirm the experimental observations of increased stability due to methoxy group attachments. Furthermore, the potential of T-shaped fluorinated linker arrangements for stabilization of COF pores was predicted. The results obtained by employing the theoretical apparatus and designed analysis tools, reported for the four respective fields of investigation, are summarized in the following.

12.1 Investigations of Surface Models for Self-Assembled Monolayers on Metal Substrates

During the initial investigation of SAMs on metal substrates, a suitable workflow for surface-model creation was explained for investigations based on GPW-DFT and CCM implemented GFN-xTB. The in-plane lattice parameters used for surface construction, were extracted from cell-optimizations of fcc-bulk structures for the respective metals. From the optimized bulks, surface-models were created by increasing the c lattice constant of optimized fcc structures yielding idealized (111) surfaces. The detailed analysis for gold-SAM interfaces, revealed a surface coverage of $\rho_{3 \times 2}^{-1} \approx 21.55 \text{ \AA}^2$ per molecule for 3×2 surfaces with two SAM-molecules per unit cell as best representation of the experimentally obtained values $\rho_{\text{exp}}^{-1} \approx 21.5 \text{ \AA}^2$. The analysis of surface-models for C10-SAM, NH₂-SAM as well as F-SAM molecules on gold was reported and revealed fundamental insights into the characteristics of the investigated species. Based on geometric analysis, the minimal SAM-system was shown to require a gold surface of 2×2 atoms. Differences in the structural arrangements of SAM-molecules on surfaces arising from modification of the backbone as well as head-group were investigated by means of geometry optimizations. For the C10-SAM and NH₂ systems, high molecular tilting with respect to the surface normal was observed for both modeled systems sizes. In the case of F-SAM systems, molecular tilting is slightly reduced $\phi_{3 \times 2}^{2F} \approx 20.7^\circ$ due to the fluorinated backbone introducing more repulsive interactions. Calculation of interaction energies contributed by periodic images for unit cells of 3×2 systems, revealed high stabilization in the case of the C10-SAM with $E_{2\text{C10,int}}^{\text{PER}} \approx -75.9 \text{ kcal/mol}$ and even higher for the NH₂-SAM, where $E_{2\text{NH}_2,\text{int}}^{\text{PER}} \approx -97.2 \text{ kcal/mol}$ could be determined. In the latter case, the phenomenon of molecular pair formation due to intermolecular hydrogen bonds between the head groups was observed. Contrarily, for F-SAMs the interactions with periodic images determined as $E_{2\text{F,int}}^{\text{PER}} \approx -11.7 \text{ kcal/mol}$ showed way less stabilizing interactions. Furthermore, analysis of the intermolecular interactions between the two SAM-molecules within the unit cell showed repulsive interactions for all species. However, interaction energies determined as $E_{\text{C10,int}}^{\text{MOL}} \approx 9.81 \text{ kcal/mol}$ and $E_{\text{NH}_2,\text{int}}^{\text{MOL}} \approx 3.75 \text{ kcal/mol}$ revealed only little influence for C10 and NH₂-SAM systems. Contrarily, for the F-SAM system $E_{\text{F,int}}^{\text{MOL}} \approx 34.51 \text{ kcal/mol}$ was obtained

showing larger destabilization due to increased repulsion between the molecular backbones. In summary, the surface analysis revealed similarities between C10 and NH₂-SAM systems, with the main difference being pair formation by hydrogen bonding. Contrarily, for F-SAM systems, fundamental differences due to high intermolecular repulsion were observed.

12.2 Field Evaporation of Self-Assembled-Monolayers

Based on gas phase calculations, the impacts of electric fields onto the general description within the DFT formalism have been shown at the example of H₂. Incremental increase of the applied field \mathcal{E} was observed to systematically suppress the bond barrier until the critical field strength $\mathcal{E}_{\text{crit}}$ favors the separation of H atoms. Similar observations were made for periodic systems, however, fundamental differences in the description such as divergent bond profiles in the unit cells' center in field direction, were observed. A generally applicable framework for compensation of SCF-convergence issues in various scenarios due to density distortion was developed. Initial investigations of the minimal 2×2 reference systems comprised of one SAM molecule revealed general trends for the description when incrementally increasing the field strength.

During the increase of \mathcal{E} , charge separation within the molecule was reported via calculation of Hirshfeld atomic partial charges, reflecting the increase of the system's polarization \mathbf{P} . Based on the initial investigations, the methodology was used for analysis of field desorption of SAMs on gold surfaces. Initially, the setup was tested by evaporation of the (111) gold surface, yielding an evaporation field strength of $\mathcal{E}_{\text{crit}}^{\text{DFT}} = \mathcal{E}_{\text{crit}}^{\text{GFN-xTB}} = 0.065$ a.u. for both utilized methods nicely corresponding to the experimental value of $\mathcal{E}_{\text{crit}}^{\text{exp}} = 0.055 - 0.06$ a.u.²⁰⁸. During the follow-up investigations of reference systems constructed for a 2×2 surface for C10-SAM and its modified versions, Hirshfeld partial charge calculations indicated charge transfer from the head group to the bulk for increasing \mathcal{E} for all three SAM-species. The charge separation was observed to be largest immediately before the fragmentation event, leading to the separation of ionized fragments from the molecules. The charge separation in the molecules were further investigated by calculation of the bulk polarization, yielding an increase of the field parallel component of \mathbf{P} during increase of \mathcal{E} . The simulations showed

12 Summary

largely similar change of properties for C10 and NH₂-SAM systems during the increase of \mathcal{E} . These observations correspond well to similarities in their surface-models reported in the previous surface analysis.

During the evaporation of the first fragment, a complete alignment of the remaining molecules with the field axis was observed for all three investigated species using the GPW-DFT model. The overall field strengths determined for the fragments were unexpectedly high and didn't correspond well to the experimentally observed ones. The minimal systems for 2×2 surfaces were utilized to illustrate the underlying effects. However, the obtained fragmentation patterns were unusable due to the oversimplified description with only one molecule explicitly optimized yielding a too "rigid" model, where intermolecular interactions arose only in the form of periodic images.

The more sophisticated 3×2 systems with two SAM-molecules per unit cell showed systematic increase of description quality by structural rearrangements as well as a modeled SAM density of $\rho_{3 \times 2}^{-1} \approx 21.55 \text{ \AA}^2$ nicely corresponding to the experimental value of $\rho_{\text{exp}}^{-1} \approx 21.5 \text{ \AA}^2$. The evaporation fields for these systems were significantly lower when compared to the smaller 2×2 ones. Furthermore, both utilized methods lead to evaporation events, however field strengths were significantly higher in the GFN-xTB calculations for all systems as exemplified for the first fragment of the F-SAM series CF₃⁺, that evaporated at $\mathcal{E}^{\text{DFT}} = 0.03 - 0.04$ a.u. in the GPW-DFT case and $\mathcal{E}^{\text{GFN-xTB}} = 0.055 - 0.06$ a.u. using GFN-xTB. Only in the case of F-SAM, real fragmentation could be observed, which was attributed to intra-molecular C-C bonds being described stronger than the S-metal bond. Furthermore, the flexibility of electron polarization is strongly impacted by the diffuse character of the describing basis functions, which seemed necessary for the description of the C10-SAM as well as NH₂-SAM. The obtained fragmentation patterns for the F-SAM differed from previous investigations using isolated molecule chains on a metal atom¹⁸ as well as cluster simulations¹⁹. The conducted calculations yield higher evaporation field strengths, e.g., for the first CF₃⁺-fragment the surface model yields $\mathcal{E}^{\text{surf}} = 0.035 - 0.04$ a.u. whereas $\mathcal{E}^{\text{iso}} \approx 0.015 - 0.02$ a.u. was observed for isolated molecules. Larger fragments as observed in gas phase calculations, e.g., C₂F₄⁺ did not occur in the surface model. An explanation was found in the alternately staggered arrangement of segments for two SAM-F molecules, leading to alternating evaporation of

identical fragments form both molecules.

The conducted studies of field evaporation were successfully conducted using PBCs and revealed general effects of the solid state DFT description within external fields. The reported results showed clear weaknesses in the modeling of atom-probe tomography via static simulations, which might be compensated in the future by more dynamic models, e.g., molecular dynamics simulations. Furthermore, the description of charge reorganization revealed flexibility in the density as crucial for modeling during increase of the electric field, the use of diffuse functions and consideration of all electrons is, therefore, determined as promising for improving the quality of description.

12.3 Theoretical Analysis for Covalent Organic Frameworks

Within this thesis, fundamental aspects of simulating the structural properties of COFs by means of GPW-DFT and GFN-xTB are investigated. The setup for GPW-DFT calculations was obtained by systematic investigation of various impacts as observed for COF-LZU1 and the N3-COF. The choices of XC-functional, basis-set, PW-cutoff as well as dispersion correction were analyzed. The resulting standard setup using TZV2P-PBE-GTH/PBE-DFTD3 was chosen based on the observed tendencies in relation to computational effort.

Additionally, non-covalent interactions between two-layered systems were investigated by cell-optimization for two-layered systems in perfectly eclipsed (AA) and imperfectly eclipsed (AA') configuration. The influence of pair-wise dispersion corrections and non-local functionals on the obtained structural model was shown to vary the obtained interlayer distances in the range of $\Delta r_{\text{iner}} \approx 0.3 \text{ \AA}$. By comparison with GFN-xTB it was further shown that aromatic interlayer interactions are modeled less repulsive as could be seen from reduction of the *c*-lattice vector by $\Delta c \approx 1.0 \text{ \AA}$. Energetic trends were shown to be qualitatively similar for both methods. This was illustrated by sampling displacements between two modeled layers within the unit cell assuming D_6 symmetry. PW-interpolation of the obtained energetic landscape with both methods revealed GFN-xTB to yield generally flatter PESs, as energetic differences were less pronounced.

By investigation of literature-known planar COFs with the established setup and specifically

12 Summary

created tools, crucial aspects for modeling COF structures were determined. The decisive impact of interlayer-interactions on the obtained stacking schemes and the related stability was shown. For COF-LZU1, N3-COF and TpPa-COF species the conducted study yielded the AA' stacking as the preferred scheme. Energy differences to staggered AB and C arrangements were connected to the nature of interlayer interactions arising from aromatic compounds within the COF-structures. Furthermore, relative energetic differences between the stacking motifs as calculated using PBE and BLYP could be related to their respective weighting of aromatic interactions. The calculations showed better agreement with experimental values using the BLYP-functional for heterocyclic aromatic compounds. Furthermore, BLYP was shown to describe oxygen-phenyl interactions more stable when compared to PBE. For all COFs, effective pore sizes \mathfrak{P} were calculated using the PSD for FRDG-generated random paths, constructed for displacements ranging from 0.0 to 5.0 Å. The resulting $\mathfrak{P}(d)$ allowed for identification of the most likely displacement by comparison with experimental values $\mathfrak{P}_{\text{exp}}$. For $\mathfrak{P}(d_{\text{ref}}) \approx \mathfrak{P}_{\text{exp}}$ direct reference to the experimental values was established. The obtained values proved consistent with the near vicinity of global minima obtained from PW interpolated PES for the investigated species. The observations made for COF-LZU1 and N3-COF systems allowed the design of the hypothetical V-COF by substitution of atoms within COF-LZU1. The target property for V-COF design was maximization of \mathfrak{P} due to favored AA' configuration by avoiding imine-imine in favor of aryl-aryl interactions. The displacement analysis revealed the global minimum for the eclipsed configuration AA' with minimal displacement, leading to an increase in \mathfrak{P} when compared to the unmodified COF-LZU1 unit cell.

Based on two-dimensional planar COF-systems, the general applicability of the chosen GPW-DFT setup, CCM implemented GFN-xTB as well as the suitability of the designed tools was shown. Investigations of the planar COF-LZU1, N3-COF, TpPa-COF as well as the hypothetically created V-COF, resulted in a systematic workflow for determination of the most probable COF-structure by comparison with the experiment. The main considerations can be summarized as follows:

The nature of the interlayer-interactions was estimated by analysis of the AA configuration. Subsequently, the energetic difference ΔE between the AA and AA' configuration alongside

12.4 Theoretical Investigations of Interlayer Interactions for Large-Pore COFs

the corresponding pore sizes \mathfrak{P}_{AA} and $\mathfrak{P}_{AA'}$ are used to validate the plausibility of the structural model. Deviations from the experimentally obtained value of the c lattice constant indicate the XC-functional or dispersion treatment to be unsuited. Comparing $\mathfrak{P}_{PSD}(d)$ for FRDG-generated displacement paths constructed for various displacements d with the experimentally obtained \mathfrak{P}_{exp} , allowed the estimation of the most probable structural arrangement. Investigation of the energetic minima $E_{min}(n_1, n_2)$ as obtained from PW interpolated displacement analysis secured the previous estimate for the most probable d . Optimization of AB and C type stacking on GPW-DFT-level was shown to exclude other high symmetry schemes. Finally, the decisiveness of interlayer interactions was shown at the example of the artificially designed V-COF.

Furthermore, the study of non-planar COFs revealed the stabilizing character of methoxy-groups by favoring the AA eclipsed configuration while simultaneously disfavoring AB and C type stacking. These were shown to be energetically disfavored by $\Delta E > 70$ kcal/mol in the case of the mNP-COF, making them even more improbable when compared to $\Delta E > 50$ kcal/mol for the unmodified NP-COF.

12.4 Theoretical Investigations of Interlayer Interactions for Large-Pore COFs

The potential impact of interlayer interactions on the obtained pore size for two-dimensional large-pore COFs was investigated by means of GPW-DFT as well as CCM implemented GFN-xTB. At the example of the newly synthesized isorecticular PP-TAB-COF series, the possible reasons for successive pore size increase for the singly and doubly methoxylated mPP-TAB and dPP-TAB-COFs could be revealed. Eclipsed stacking schemes maximizing \mathfrak{P} were investigated for their stability. The multitude of possible unit-cell conformers for this series of highly non-planar and anti-symmetric COF species, was compensated by a hierarchical analysis of component arrangements. Here, the interactions of two-component arrangements from gas phase over “blocks” to “walls”, revealed general trends for the components’ orientations. The most probable eclipsed AA’ configuration was constructed for alternatingly stacked imine bonds (alt) as well as same side (S) oriented phenylphenanthridine (PP) linkers. Furthermore, the interactions between methoxy groups were found to be favored when attached in parallel as

12 Summary

well as oriented in “order” with the adjacent ones. The investigations of interlayer interactions for the AA' configuration of the PP-TAB-COF series revealed the role of methoxy groups as molecular “anchors”. Analysis of the displacement behavior showed the global energetic minimum for the staggered AB configuration for the unmodified PP-TAB-COF. Similarly for the singly methoxylated mPP-TAB-COF, however, a local minimum is found for the AA' configuration, explaining the experimentally observed slightly increased β . For the doubly methoxylated dPP-TAB-COF, the global minimum is found for the AA' configuration showing the general trend towards larger β when attaching additional methoxy-groups. Furthermore, the decrease in length of the in-plane lattice vectors while shifting layers with respect to each other was shown to cause layer buckling. The effect of layer buckling could be connected to displacements d and was minimal for AA' configurations. The experimentally obtained β_{exp} could be validated by using the obtained structural model by calculation of β_{PSD} for FRDG random paths. Treating the isorecticular PP-TAB-COF series, the previously introduced methodology was proven to be successful in explaining the experimentally observed effects.

The obtained insights into the components' behavior showed the phenomenon of T-shaped stacking to be energetically favored within blocks and walls with PP-linkers arranged in S10-configuration. However, the energetic differences between parallel and T-shaped stacking was shown to be insignificant throughout the PP-TAB-COF series indicating both orientations to appear with equal probability.

Building on the phenomenon of T-shaped stacking, the impact of F-PP-linkers on the PP-TAB-COF series was revealed by component analysis in analogy to the PP linkers. Even though gas phase calculations only revealed S-arrangements to be favored, investigation of the T-shape phenomenon for the F-PP-COFs showed an astonishing energetic difference between parallel and T-shape oriented linkers. The results obtained by means of GPW-DFT as well as GFN-xTB, identify T-shaped stacking to be favorable by $\Delta E_{\text{T-para}} \approx 60.0$ kcal/mol for each F-PP-linker within the simulated unit cell. As a side effect, F-PP linkers in T-shape were shown to reduced layer buckling, consequently, leading to an additional increase of β . The most probable AA' eclipsed structure was estimated based on the previously investigated PP-TAB-COF series by substitution of PP with F-PP components. Analyzing the energetic landscape by displacement of two layers, the AA' configuration was determined as global

12.4 Theoretical Investigations of Interlayer Interactions for Large-Pore COFs

minimum for all three investigated F-PP-TAB-COFs. However, the number of local side minima decreased with additional methoxy groups. The effects of fluorinated linkers could be shown to be only relevant when T-shaped orientation is assumed. The global minima were not found for AA' configurations when parallel oriented linkers were assumed showing the importance of the T-shape interaction. The conducted studies, for the isorecticular PP-TAB-COF and F-PP-TAB-COF series, illuminate a possible path for rationalization of large-pore COF synthesis by aid of theoretical investigation of interlayer interactions.

Appendix A

Theory

A.1 Separation of the Time Dependence

For simplification, the wave function is usually separated into components depending only on the time and space, respectively. This is achieved by the application of a product ansatz for the wave function of form

$$\Psi(t, \mathbf{r}) = \Phi(\mathbf{r})f(t) \quad . \quad (\text{A.1})$$

After inserting Eq (A.1) in Eq (2.1) and division by $f(t)$ and $\Phi(\mathbf{r})$ respectively, one obtains the form

$$i\hbar \frac{1}{f(t)} \frac{\partial}{\partial t} f(t) = \frac{1}{\Phi(\mathbf{r})} \hat{\mathbf{H}}\Phi(\mathbf{r}) \quad . \quad (\text{A.2})$$

Separating the t and \mathbf{r} dependent parts via introduction of a common constant E leads to the equations Eq (A.3a) and eq.Eq (A.3b), respectively. The constant E can be further identified to be the systems' energy characterizing the solution for a given Ψ :

$$E = i\hbar \frac{1}{f(t)} \frac{\partial}{\partial t} f(t) \qquad -\frac{iE}{\hbar} f(t) = \frac{\partial}{\partial t} f(t) \quad (\text{A.3a})$$

A Theory

$$E = \frac{1}{\Phi(\mathbf{r})} \hat{\mathbf{H}}\Phi(\mathbf{r}) \qquad \hat{\mathbf{H}}\Phi(\mathbf{r}) = E\Phi(\mathbf{r}) \qquad (\text{A.3b})$$

The solution of Eq (A.3a) is thus given as $f(t) = Ae^{-iEt/\hbar}$ independent of the system of interest, whereas for equation Eq (A.3b) the exact form of the systems Hamiltonian $\hat{\mathbf{H}}$ has to be known. However, this treatment holds only for time-independent Hamiltonians, as otherwise the separation of time is not possible.

A.2 Born-Oppenheimer Approximation

A brief look at the Hamiltonians shape for the molecular problem yields two main parts namely $\hat{\mathbf{H}}_{\text{el}}$ the electronic Hamiltonian and $\hat{\mathbf{T}}_{\text{N}}$ the nuclear kinetic energy operator. As $\hat{\mathbf{H}}_{\text{el}}$ is hermitian, its eigenfunctions $\phi_n(\mathbf{r}, \mathbf{R})$ are a complete set of orthonormal functions, where n marks the electronic state corresponding to energy E_n . Given this properties, they can be used for a basis set expansion of any molecular wave function $\Psi_k(\mathbf{r}, \mathbf{R})$. Thus Ψ can be expressed as

$$\Psi_k(\mathbf{r}, \mathbf{R}) = \sum_{n=1}^{\infty} \phi_n(\mathbf{r}, \mathbf{R})\chi_{n,k}(\mathbf{R}) \qquad (\text{A.4})$$

with $\chi_{n,k}$ being the expansion coefficients. Using Eq (A.4) in Eq (2.1) leads to the following separation of terms

$$\sum_n (\hat{\mathbf{T}}_{\text{N}}\phi_n\chi_{n,k} + \chi_{n,k}E_n\phi_n) = \mathcal{E}_k \sum_n \phi_n\chi_{n,k} \qquad (\text{A.5})$$

The coordinate-dependence has been dropped for visibility reasons. It was used that $\hat{\mathbf{H}}_{\text{el}}$ only acts on the electronic wave function. Using the orthonormality of the electronic wave functions $\langle \phi_m | \phi_n \rangle = \delta_{mn}$ leads, after multiplication of Eq (A.5) with $\langle \phi_m |$ from the left, to

$$\sum_n \langle \phi_m | \hat{\mathbf{T}}_{\text{N}} | \phi_n \rangle \chi_{n,k} + \delta_{mn}E_n\chi_{n,k} = \mathcal{E}_k \sum_n \delta_{mn}\chi_{n,k} \qquad (\text{A.6})$$

This can be further simplified by executing the δ operation to obtain the decoupled Schrödinger equation of the core movement to a specific electronic state, with energy E_m and wave

function ϕ_m , given as

$$\sum_n \langle \phi_m | \hat{\mathbf{T}}_N | \phi_n \rangle \chi_{n,k} + E_m \chi_{m,k} = \mathcal{E}_k \chi_{m,k} \quad . \quad (\text{A.7})$$

With the product rule for the Laplace-operator acting on two functions f and g given as

$$\nabla^2(fg) = f \nabla^2 g + 2 \nabla f \nabla g + g \nabla^2 f \quad . \quad (\text{A.8})$$

The first term of Eq (A.7) can be expressed as

$$\sum_n \langle \phi_m | \hat{\mathbf{T}}_N | \phi_n \chi_{n,k} \rangle = \frac{\hbar^2}{2M_k} \left(\underbrace{\langle \phi_m | \phi_n \rangle}_{\delta_{mn}} \nabla^2 + 2 \langle \phi_m | \nabla | \phi_n \rangle \nabla + \langle \phi_m | \nabla^2 | \phi_n \rangle \right) \chi_{n,k} \quad , \quad (\text{A.9})$$

leading to a separation of terms into

$$\sum_n \langle \phi_m | \hat{\mathbf{T}}_N | \phi_n \chi_{n,k} \rangle = \hat{\mathbf{T}}_N \chi_{m,k} + \sum_n \left(2 \langle \phi_m | \nabla | \phi_n \rangle \nabla + \langle \phi_m | \nabla^2 | \phi_n \rangle \right) \chi_{n,k} \quad . \quad (\text{A.10})$$

The so called non-adiabatic terms are given as

$$\hat{\mathbf{T}}'_{m,n} = -\frac{\hbar^2}{2M_k} \langle \phi_m | \nabla | \phi_n \rangle \nabla \quad (\text{A.11})$$

$$\hat{\mathbf{T}}''_{m,n} = -\frac{\hbar^2}{2M_k} \langle \phi_m | \nabla^2 | \phi_n \rangle \quad (\text{A.12})$$

leading to the final form of equation Eq (A.7) given as

$$\sum_n (2\hat{\mathbf{T}}'_{m,n} + \hat{\mathbf{T}}''_{m,n}) \chi_{n,k} + (\hat{\mathbf{T}}_N + E_m) \chi_{m,k} = \mathcal{E}_k \chi_{m,k} \quad . \quad (\text{A.13})$$

This new form is a system of n coupled differential equations for the nuclear wave functions $\chi_{n,k}$, where the energies of the electronic problem, act as the external potential. In quantum chemistry, usually the coupling terms are dropped in the Born-Oppenheimer-approximation leading to

$$(\hat{\mathbf{T}}_N + E_m) \chi_{m,k} = \mathcal{E}_k \chi_{m,k} \quad . \quad (\text{A.14})$$

A Theory

As a result one obtains a system of decoupled equations corresponding to an electronic state m , representing a huge simplification in terms of computational efforts. This approximation is justified if the electronic wave function changes only slowly with the variation of the core coordinates, as the coupling terms include the first $\hat{\mathbf{T}}'_{m,n}$ and second $\hat{\mathbf{T}}''_{m,n}$ derivatives of the electronic wave function after the nuclear coordinates. Meaning they only contribute significantly, when small shifts of the nucleic positions lead to a large change in the electronic wave function.

A.3 Variational Principle

The variational principle states that the energy E_{trial} corresponding to an arbitrary electronic trial wave function Ψ_{trial} , is always higher than the energy E_0 of the electronic ground state Ψ_0 . This can be explained by expanding the ground state wave function in a complete set of a basis functions ϕ_α , resulting in

$$\Psi_{\text{trial}} = \sum_{\alpha} c_{\alpha} \phi_{\alpha} \quad , \quad (\text{A.15})$$

with expansion coefficients c_{α} . Here the exact formulation is reached for a complete basis set, resulting in the condition $\Psi_{\text{trial}} = \Psi_0$. Thus considering the expectation value for a Hamiltonian calculated by using a trial wave function, results in the trial energy:

$$\langle \hat{\mathbf{H}} \rangle = \langle \Psi_{\text{trial}} | \hat{\mathbf{H}} | \Psi_{\text{trial}} \rangle = E_{\text{trial}} \quad . \quad (\text{A.16})$$

The ground state energy is defined as the lowest energy a system can achieve, thus any wave function describing a different state, leads to an expectation value corresponding to a higher energy. Using orthonormal eigenfunctions of the Hamiltonian as basis functions, fulfilling the relation

$$\hat{\mathbf{H}} \phi_{\alpha} = E_{\alpha} \phi_{\alpha} \quad , \quad (\text{A.17})$$

one obtains for the trial wave function

$$\hat{\mathbf{H}}|\Psi_{\text{trial}}\rangle = \sum_{\alpha} c_{\alpha} E_{\alpha} \quad . \quad (\text{A.18})$$

Using this result as well as the orthonormality constraint $\langle\phi_{\alpha}|\phi_{\beta}\rangle = \delta_{\alpha\beta}$, leads to

$$\langle\Psi_{\text{trial}}|\hat{\mathbf{H}}|\Psi_{\text{trial}}\rangle = \sum_{\alpha} c_{\alpha} c_{\beta}^{*} E_{\alpha} \underbrace{\langle\phi_{\alpha}|\phi_{\beta}\rangle}_{=\delta_{\alpha\beta}} = \sum_{\alpha} |c_{\alpha}|^2 E_{\alpha} = \sum_{\alpha} E_{\alpha} \geq E_0 \quad . \quad (\text{A.19})$$

Here it has to be noted, that the equality holds only if the expanded trial function reproduces Ψ_0 . However, in the usual case Eq (A.19) would lead to a sum of energy eigenvalues $E_{\alpha} > E_0$ belonging to arbitrary states ϕ_{α} and thus results in an energy eigenvalue larger than E_0 .

A.4 Theorems for Gaussian-Functions

For Gaussian functions of form $f(\mathbf{r}) = e^{-a\mathbf{r}^2}$ multiple identities are known from mathematics, greatly simplifying their handling.

Product of Gaussians

The product of two Gaussians is again a Gaussian, this can be easily shown by for the general case of two shifted Gaussians by explicit calculation via

$$f_1(\mathbf{r})f_2(\mathbf{r}) = N_1N_2e^{-\alpha_1(\mathbf{r}-\mathbf{R}_1)^2}e^{-\alpha_2(\mathbf{r}-\mathbf{R}_2)^2} \quad . \quad (\text{A.20})$$

Using the rules for exponential calculus one obtains with $F(\mathbf{r}) = f_1(\mathbf{r})f_2(\mathbf{r})$ the formulation

$$F(\mathbf{r}) = N_P e^{-\alpha_P(\mathbf{r}-\mathbf{R}_P)^2} \quad . \quad (\text{A.21})$$

Here the new variables are defined as

$$N_P = N_1N_2e^{-(\alpha_1\mathbf{R}_1^2)-(\alpha_2\mathbf{R}_2^2)-(\alpha_P\mathbf{R}_P)}, \quad \alpha_P = \alpha_1 + \alpha_2 \quad \text{and} \quad \mathbf{R}_P = \frac{\alpha_1\mathbf{R}_1 + \alpha_2\mathbf{R}_2}{\alpha_1 + \alpha_2} \quad . \quad (\text{A.22})$$

Integral over a Gaussian

The integral over a Gaussian function $f(x) = e^{-a(x-b)^2/c^2}$ is given by the simple relation

$$\int_{-\infty}^{\infty} f(x)dx = \int_{-\infty}^{\infty} e^{-ax^2+bx} dx = \sqrt{\frac{\pi}{a}} e^{\frac{b^2}{4a}} \quad . \quad (\text{A.23})$$

Fourier Transformation of a Gaussian

The Fourier transform $F(\mathbf{k})$ of a Gaussian function $f(\mathbf{r}) = Ne^{-\frac{1}{2}\mathbf{r}^2}$ is again of Gaussian shape in the variable \mathbf{k} . The underlying relation can be stated as follows:

$$F(\mathbf{k}) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{1}{2}\mathbf{r}^2} e^{-i\mathbf{k}\mathbf{r}} d\mathbf{r} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{1}{2}\mathbf{r}^2 - i\mathbf{k}\mathbf{r}} d\mathbf{r} \quad (\text{A.24})$$

Using the relation for the Gaussian integral Eq (A.23) one obtains

$$\int_{-\infty}^{\infty} e^{-ar^2-ikr} dr = \frac{\pi}{a} e^{\frac{(-ik)^2}{4a}} , \quad (\text{A.25})$$

leading to the final formulation for the Fourier transform as

$$F(\mathbf{k}) = \sqrt{\frac{\pi}{a}} e^{-\frac{1}{4a} \mathbf{k}^2} , \quad (\text{A.26})$$

which is again a Gaussian in the variable \mathbf{k}

A.5 Hohenberg-Kohn Theorem I

The basic condition for the density describing the total number of particles in a system is trivially incorporated within its definition. The density is defined as the absolute value of the wave function squared, which on its own is defined to be normalized

$$\int \rho(\mathbf{r})d\mathbf{r} = \int |\Psi(\mathbf{r})|^2 d\mathbf{r} = N \langle \Psi | \Psi \rangle = N \quad . \quad (\text{A.27})$$

To see that the density also unambiguously defines the external potential, it is suitable to assume the existence of two separate external potentials, defined by the same density. If two different potentials exist for the same density, it also means that they correspond to two different Hamiltonians. This means as well that two different ground-state wave functions leading to the same density exist. For the ground-state wave function the equality

$$E_0 = \langle \Psi | \hat{\mathbf{H}} | \Psi \rangle \quad (\text{A.28})$$

holds. After applying the variational theorem, this means that every other wave function leads to an energy strictly greater than the respective Hamiltonians ground state energy E_0 . Given these basic conditions, one obtains for two different Hamiltonians the relations:

$$\begin{aligned} E'_0 < \langle \Psi | \hat{\mathbf{H}}' | \Psi \rangle &= \langle \Psi | \hat{\mathbf{H}}' + (\hat{\mathbf{H}} - \hat{\mathbf{H}}) | \Psi \rangle \\ &= \langle \Psi | \hat{\mathbf{H}} | \Psi \rangle + \langle \Psi | \hat{\mathbf{H}}' - \hat{\mathbf{H}} | \Psi \rangle \\ &= E_0 + \int \rho(\mathbf{r})(v'_{\text{ext}}(\mathbf{r}) - v_{\text{ext}}(\mathbf{r}))d\mathbf{r} \end{aligned} \quad (\text{A.29})$$

$$\begin{aligned} E_0 < \langle \Psi' | \hat{\mathbf{H}} | \Psi' \rangle &= \langle \Psi' | \hat{\mathbf{H}} + (\hat{\mathbf{H}}' - \hat{\mathbf{H}}) | \Psi' \rangle \\ &= \langle \Psi' | \hat{\mathbf{H}}' | \Psi' \rangle + \langle \Psi' | \hat{\mathbf{H}} - \hat{\mathbf{H}}' | \Psi' \rangle \\ &= E'_0 + \int \rho(\mathbf{r})(v_{\text{ext}}(\mathbf{r}) - v'_{\text{ext}}(\mathbf{r}))d\mathbf{r} \end{aligned} \quad (\text{A.30})$$

Addition of the two equations leads to the paradox statement

$$E'_0 + E_0 < E_0 + E'_0 \quad , \quad (\text{A.31})$$

which proves that there can only be one external potential corresponding to a specific density.

A.6 Parsevals Theorem

The theorem of Parseval states for two Riemann integrable functions $A(x)$ and $B(x)$ with forms

$$A(x) = \sum_{n=-\infty}^{\infty} a_n e^{inx} \quad \text{and} \quad B(x) = \sum_{n=-\infty}^{\infty} b_n e^{inx} \quad (\text{A.32})$$

it follows that

$$\sum_{n=-\infty}^{\infty} a_n b_n^* = \frac{1}{2\pi} \int_{-\pi}^{\pi} A(x) B(x)^* dx \quad . \quad (\text{A.33})$$

In other terms, an infinite sum over the weights of square roots of one is equal to a gauched integration of the same over a whole circle.

A.7 Plancherets Theorem

The theorem of Plancheret states that the inner product of two functions is equivalent to the product of the respective Fourier transforms

$$\int_{-\infty}^{\infty} f(x) g(x)^* dx = \int_{-\infty}^{\infty} F(k) G(k)^* dk \quad (\text{A.34})$$

A.8 Mean Value Theorem for Lattice Periodic Functions

Suppose a function $F(\mathbf{k})$ that is periodic with respect to a lattice with lattice vectors \mathbf{R} and can therefore be expanded in a Fourier series. The integral over the full \mathbf{k} -space is given as

$$I = \int F(\mathbf{k})d\mathbf{k} \quad . \quad (\text{A.35})$$

As the function $F(\mathbf{k})$ is lattice periodic, it can be written as

$$F(\mathbf{k}) = \sum_j F_j e^{i\mathbf{k}\mathbf{R}_j} \quad , \quad (\text{A.36})$$

where j runs over possible choices of lattice vectors given as $\mathbf{R}_j = \mathbf{a}_1 n_{1,j} + \mathbf{a}_2 n_{2,j} + \mathbf{a}_3 n_{3,j}$. The exponent in A.36 is typically interpreted as oscillation with wave vector \mathbf{k}_j and, thus, the summation is in principle executed over weighted plane waves of different crystal momentum. Inserting A.36 in A.35 leads to

$$I = \int \sum_j F_j e^{i\mathbf{k}\mathbf{R}_j} d\mathbf{k} \quad , \quad (\text{A.37})$$

which can be easily evaluated, as the integral over the complex exponential function is defined as

$$\delta(x - x_0) = \int e^{i(x-x_0)} dx \quad , \quad (\text{A.38})$$

As a consequence, only the term with vanishing exponent contributes. With this relation in mind, the value of the integral can be evaluated by

$$I = \sum_j F_j \int e^{i\mathbf{k}\mathbf{R}_j} d\mathbf{k} = \sum_j F_j \delta(i\mathbf{k}\mathbf{R}_j) = F_0 \quad , \quad (\text{A.39})$$

where the argument of δ can only become zero if \mathbf{R}_j becomes zero. This is usually assumed as the vector \mathbf{R}_0 , where all lattice coefficients are zero. It can therefore be assumed that only the coefficient F_0 of the Fourier-series contributes to the integral.

Appendix B

Modern Theory of Polarization

B.1 Position Operator and Polarization

Let $T(u)$ be a one dimensional translation operator that shifts a reference particle by a length u . In order to illustrate the effects of the relation $[T(L), x] \neq 0$ on the calculation of the polarization, it is suitable to firstly calculate commutation relation explicitly for a trial wavefunction. Recasting the effects of both operators on a trial function respectively, leads to

$$\hat{x}\Phi(x) = x \cdot \Phi(x) \quad , \quad (\text{B.1})$$

for the position operator, and

$$\hat{\mathbf{T}}(u)\Phi(x) = \Phi(x + u) \quad , \quad (\text{B.2})$$

for the translation operator. This results in the commutation relation

$$\begin{aligned} [\hat{\mathbf{T}}(L), \hat{x}]\Phi(x) &= \hat{\mathbf{T}}(L)\hat{x}\Phi(x) - \hat{x}\hat{\mathbf{T}}(L)\Phi(x) \\ &= \hat{\mathbf{T}}(L)x\Phi(x) - \hat{x}\Phi(x + L) \\ &= x\Phi(x + L) - (x + L)\Phi(x + L) = L\Phi(x + L) \neq 0 \quad . \end{aligned} \quad (\text{B.3})$$

B Modern Theory of Polarization

As a result, the polarization defined as the expectation value of the charge density multiplied with the position operator is not unique and well defined as can be seen by evaluating the integral over the unit cell, given in the one dimensional case as

$$P_x = \frac{1}{L} \int_0^L x \rho(x) dx \quad . \quad (\text{B.4})$$

B.1.1 Example Systems

Consider the density $\rho(x) = e\delta(x) - e\delta(x + \frac{L}{4})$ for a one dimensional system consisting of two charges with opposite sign. The polarization is then defined by

$$P_x = \frac{e}{L} \int_0^L x \left(\delta(x) - \delta\left(x + \frac{L}{4}\right) \right) dx = \frac{e}{L} \left(0 - \frac{L}{4} \right) = -\frac{1}{4}e \quad , \quad (\text{B.5})$$

whereas the same arrangement in different boundary positions given as $\rho(x) = -e\delta(x + \frac{L}{8}) + e\delta(x + \frac{7L}{8})$ would lead to

$$P_x = \frac{e}{L} \int_0^L x \left(\delta\left(x + \frac{L}{8}\right) - \delta\left(x + \frac{7L}{8}\right) \right) dx = \frac{e}{L} \left(\frac{L}{8} - \frac{7L}{8} \right) = -\frac{3}{4}e \quad . \quad (\text{B.6})$$

Notably for a centrosymmetric lattice, a change of sign can be obtained by using the inversion symmetric charge-configurations

$$\rho_+(x) = e \left(\delta\left(x + \frac{L}{2}\right) - \delta\left(x + \frac{3L}{2}\right) \right) \rightarrow P_+ = -\frac{1}{2}e \quad , \quad (\text{B.7})$$

and

$$\rho_-(x) = -e \left(\delta\left(x + \frac{L}{2}\right) + \delta\left(x + \frac{3L}{2}\right) \right) \rightarrow P_- = \frac{1}{2}e \quad , \quad (\text{B.8})$$

respectively. Whereas the generalization of this result in the 3d case, leads to the definition of the polarization quantum of the form

$$P_q = \frac{e\mathbf{a}}{\Omega} \quad , \quad (\text{B.9})$$

B.2 Derivation of the Berry Phase for Adiabatic Change of State

where \mathbf{a} is a lattice vector and Ω is the unit cell volume. For a non polar lattice, this will result in a total polarization of

$$\mathbf{P}^{\text{NP}} = \frac{\mathbf{P}_q}{2} + n\mathbf{P}_q \quad , \quad (\text{B.10})$$

where n runs over all integers.

B.2 Derivation of the Berry Phase for Adiabatic Change of State

Here, the derivation of the Berry-phase following the original paper by Berry¹³⁰ provided in the review by Xiao et. al.²⁵⁹, is illustrated. The model system is described by the time dependent Schrödinger equation as

$$\hat{\mathbf{H}}(\mathbf{R}(t))\Psi_n(t) = i\hbar\frac{\partial}{\partial t}\Psi_n(t) \quad , \quad (\text{B.11})$$

where each contribution has a parametric time dependence. The Hamiltonians eigenfunctions $|n(\mathbf{R})\rangle$ are further defined for every instant $\mathbf{R} = \mathbf{R}(t)$, by the equation

$$\hat{\mathbf{H}}(\mathbf{R})|n(\mathbf{R})\rangle = \epsilon_n(\mathbf{R})|n(\mathbf{R})\rangle \quad . \quad (\text{B.12})$$

with the basis functions $|n(\mathbf{R})\rangle$ to the energy eigenvalues ϵ_n . If a system is changed adiabatically, it remains at any given time step t in an eigenstate of $\hat{\mathbf{H}}(t)$. Suppose a state initially given as $|n(\mathbf{R}(0))\rangle$ evolves with time and is given as $|n(\mathbf{R}(t))\rangle$ at t . The state $\Psi_n(t)$ is defined only within a phase factor and can, thus, be chosen as

$$\Psi_n(t) = \underbrace{e^{i\gamma(t)}}_B \cdot \underbrace{e^{-\frac{i}{\hbar}\int_0^t \epsilon_n(\mathbf{R}(t'))dt'}}_D \cdot |n(\mathbf{R}(t))\rangle \quad . \quad (\text{B.13})$$

The phase that can be chosen freely in Eq (B.13) is divided into two parts, where the dynamical phase is given as D and the Berry-phase B. For further analysis it is important to summarize some side facts resulting from the described system's properties.

B Modern Theory of Polarization

1. The Hamiltonian is explicitly time dependent $\partial_t \hat{\mathbf{H}} \neq 0 \rightarrow \partial_t |n(\mathbf{R}(t))\rangle \neq 0$.
2. The Berry-phase $\gamma(t)$ is non integratable, cannot be written as a function of \mathbf{R} and is not single valued under a complete continuation around a circuit.
3. The derivative of the integral appearing in D can be evaluated by the fundamental theorem of calculus.

With these considerations in mind, the Berry-phase can be calculated by substituting Eq (B.13) in Eq (B.11). For the time derivative, the product rule leads to a sum of three distinct terms given as

$$\begin{aligned} \partial_t \Psi_n(t) &= e^{i\gamma(t)} e^{-\frac{i}{\hbar} \int_0^t \epsilon_n(\mathbf{R}(t')) dt'} \cdot i\dot{\gamma}(t) |n(\mathbf{R}(t))\rangle \\ &\quad - e^{i\gamma(t)} e^{-\frac{i}{\hbar} \int_0^t \epsilon_n(\mathbf{R}(t')) dt'} \cdot \frac{i}{\hbar} \epsilon_n |n(\mathbf{R}(t))\rangle \\ &\quad + e^{i\gamma(t)} e^{-\frac{i}{\hbar} \int_0^t \epsilon_n(\mathbf{R}(t')) dt'} \cdot \frac{\partial}{\partial t} (|n(\mathbf{R}(t))\rangle) \quad , \end{aligned} \quad (\text{B.14})$$

where $\partial_t |n(\mathbf{R})\rangle$ can be reformulated using the chainrule as

$$\frac{\partial}{\partial t} |n(\mathbf{R}(t))\rangle = \frac{\partial \mathbf{R}}{\partial t} \frac{\partial}{\partial \mathbf{R}} |n(\mathbf{R})\rangle = \dot{\mathbf{R}} \cdot \frac{\partial}{\partial \mathbf{R}} |n(\mathbf{R})\rangle \quad (\text{B.15})$$

for the lefthand side, one directly obtains the result by using Eq (B.12), as

$$\hat{\mathbf{H}} \Psi_n(t) = \epsilon_n e^{i\gamma(t)} \frac{i}{\hbar} e^{-\frac{i}{\hbar} \int_0^t \epsilon_n(\mathbf{R}(t')) dt'} |n(\mathbf{R}(t))\rangle \quad . \quad (\text{B.16})$$

Identifying the common phase factors on both sides, the total expression given as

$$\epsilon_n |n(\mathbf{R})\rangle = i\hbar \left(i\dot{\gamma} - \frac{i}{\hbar} \epsilon_n + \dot{\mathbf{R}} \frac{\partial}{\partial \mathbf{R}} \right) |n(\mathbf{R})\rangle \quad , \quad (\text{B.17})$$

is obtained. It leads to the cancellation of the energy dependent terms after multiplication with $|n(\mathbf{R}(t))\rangle$ from the left, and after addition of $i\dot{\gamma}$ to the expression

$$\dot{\gamma} = -i \langle n(\mathbf{R}) | \frac{\partial}{\partial \mathbf{R}} |n(\mathbf{R})\rangle \dot{\mathbf{R}} \quad . \quad (\text{B.18})$$

B.2 Derivation of the Berry Phase for Adiabatic Change of State

From the time derivative, the total phase change for state n can be calculated by integrating over a closed path C in the parameter t , leading to the expression

$$\gamma_n(C) = i \oint_C \frac{\partial}{\partial \mathbf{R}} |n(\mathbf{R})\rangle dt \frac{d\mathbf{R}}{dt} = \oint_C \mathcal{A}_n(\mathbf{R}) d\mathbf{R} \quad , \quad (\text{B.19})$$

where the Berry connection \mathcal{A}_n is given as

$$\mathcal{A}_n(\mathbf{R}) = i \langle n(\mathbf{R}) | \frac{\partial}{\partial \mathbf{R}} |n(\mathbf{R})\rangle \quad , \quad (\text{B.20})$$

This is also referred to as Berry vector potential. From Eq (B.19) it follows, that the overall phase change along a closed path only depends on the geometry of the path itself, but not on how it changes in time.

Appendix C

Tight Binding Methods

C.1 Reformulation of the DFT Energy Expression for DFTB Approximation

From the Kohn-Sham equations describing non interacting electrons in an external potential provided by the remaining electrons, an expression based on the energies of the independent electronic energies and a correction term is used to derive the DFTB Hamiltonian. The KS equations are given as

$$\left(-\frac{1}{2}\nabla^2 + V_{\text{eff}}(\mathbf{r})\right)\Psi_i = \epsilon_i\Psi_i \quad , \quad (\text{C.1})$$

where the effective potential $V_{\text{eff}}(\mathbf{r})$ is depending on the density $\rho = \sum_i n_i\Psi_i^*\Psi_i$ and is comprised of

$$V_{\text{eff}} = V_{\text{ne}} + V_{\text{ee}} + V_{\text{xc}} \quad , \quad (\text{C.2})$$

C Tight Binding Methods

with the nuclear-electron V_{ne} , electron-electron V_{ee} and exchange correlation V_{xc} potentials, as introduced in Section 2.4. The kinetic energy T_S of the non interacting system can then be written after reformulating Eq (C.1) as

$$T_S = -\frac{1}{2} \sum_i n_i \langle \Psi_i | \nabla^2 | \Psi_i \rangle = \sum_i n_i (\epsilon_i - \langle \Psi_i | V_{\text{eff}}(\mathbf{r}) | \Psi_i \rangle) \quad . \quad (\text{C.3})$$

For the total DFT energy expression as introduced in Eq (2.26) it follows

$$E_{\text{DFT}}[\rho(\mathbf{r})] = T_S[\rho(\mathbf{r})] + E_{ee}[\rho(\mathbf{r})] + E_{ne}[\rho(\mathbf{r})] + E_{xc}[\rho(\mathbf{r})] + E_{nn} \quad (\text{C.4})$$

$$= \sum_i n_i \langle \Psi_i | \frac{1}{2} \nabla^2 | \Psi_i \rangle + \underbrace{\frac{1}{2} \int \frac{\rho(\mathbf{r})\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} \mathbf{d}\mathbf{r}}_{\sum_i n_i \langle \Psi_i | V_{ee} | \Psi_i \rangle} + \langle \Psi_i | V_{ne} | \Psi_i \rangle + E_{xc} + E_{nn} \quad . \quad (\text{C.5})$$

identification of the already present contributions in the effective potential as well as introduction of an additional exchange correlation term lead to a tight-binding like formulation as presented in Section 4.1. Reformulation of Eq (C.5) leads to the expression

$$E_{\text{DFT}}[\rho(\mathbf{r})] = \sum_i n_i \langle \Psi_i | -\frac{1}{2} \nabla^2 + \underbrace{V_{ne} + V_{ee} + V_{xc}}_{V_{\text{eff}}} | \Psi_i \rangle - \frac{1}{2} \int \frac{\rho(\mathbf{r})\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} \mathbf{d}\mathbf{r} - \sum_i n_i \langle \Psi_i | V_{xc} | \Psi_i \rangle + E_{xc} + E_{nn} \quad (\text{C.6})$$

$$= \sum_i n_i \epsilon_i - \frac{1}{2} \int V_{ee}[\rho(\mathbf{r})]\rho(\mathbf{r}) \mathbf{d}\mathbf{r} - \int V_{xc}[\rho(\mathbf{r})]\rho(\mathbf{r}) \mathbf{d}\mathbf{r} + E_{xc} + E_{nn} \quad , \quad (\text{C.7})$$

highlighting the connection to the non interacting system of electrons.

C.2 Derivation of the DFTB-Energy Expression

In the literature it is often stated that within DFTB formalism, the electron density is given as sum of an atomic reference and small fluctuations following $\rho = \rho_0 + \Delta\rho = \sum_i n_i \Psi_i^* \Psi_i$. However the substitution process and the subsequent reformulation of Eq (C.7) are explicitly

illustrated in the following. Substitution of $\rho \rightarrow \rho_0 + \Delta\rho$ in Eq (C.7) leads to:

$$\begin{aligned}
 E_{\text{DFTB}}[\rho_0 + \Delta\rho] &= \sum_i n_i \langle \Psi_i | -\frac{1}{2} \nabla^2 + V_{\text{ne}} + \int' \frac{(\rho'_0 + \Delta\rho')}{|\mathbf{r} - \mathbf{r}'|} \mathbf{d}\mathbf{r}' + V_{\text{xc}}[\rho_0 + \Delta\rho] | \Psi_i \rangle \\
 &\quad - \frac{1}{2} \int' \int \frac{(\rho(\mathbf{r}) + \Delta\rho)(\rho(\mathbf{r}') + \Delta\rho')}{|\mathbf{r} - \mathbf{r}'|} \mathbf{d}\mathbf{r} \mathbf{d}\mathbf{r}' - \sum_i n_i \langle \Psi_i | V_{\text{xc}}[\rho_0 + \Delta\rho] | \Psi_i \rangle \\
 &\quad + E_{\text{xc}}[\rho_0 + \Delta\rho] + E_{\text{nn}} \quad , \tag{C.8}
 \end{aligned}$$

where the terms that depend on $\Delta\rho$ in the first line of Eq (C.8) can be extracted leading to a contribution only depending on ρ_0 . The obtained expression is given as

$$\begin{aligned}
 E_{\text{DFTB}}[\rho_0 + \Delta\rho] &= \sum_i n_i \langle \Psi_i | \underbrace{-\frac{1}{2} \nabla^2 + V_{\text{ne}} + \int' \frac{\rho'_0}{|\mathbf{r} - \mathbf{r}'|} \mathbf{d}\mathbf{r}' + V_{\text{xc}}[\rho_0]}_{\hat{\mathbf{H}}_0} | \Psi_i \rangle \\
 &\quad + \int' \int \frac{\Delta\rho'(\rho_0(\mathbf{r}) + \Delta\rho)}{|\mathbf{r} - \mathbf{r}'|} \mathbf{d}\mathbf{r}' \mathbf{d}\mathbf{r} + \int V_{\text{xc}}[\rho] \rho(\mathbf{r}) \mathbf{d}\mathbf{r} \\
 &\quad - \int V_{\text{xc}}[\rho_0] \rho(\mathbf{r}) \mathbf{d}\mathbf{r} - \underbrace{\sum_i n_i \langle \Psi_i | V_{\text{xc}}[\rho_0 + \Delta\rho] | \Psi_i \rangle}_{\int V_{\text{xc}}[\rho] \rho(\mathbf{r}) \mathbf{d}\mathbf{r}} \\
 &\quad - \frac{1}{2} \int' \int \frac{(\rho(\mathbf{r}')\rho(\mathbf{r}) + \Delta\rho(\mathbf{r})\rho(\mathbf{r}') + \rho(\mathbf{r})\Delta\rho(\mathbf{r}') + \Delta\rho(\mathbf{r})\Delta\rho(\mathbf{r}'))}{|\mathbf{r} - \mathbf{r}'|} \mathbf{d}\mathbf{r} \mathbf{d}\mathbf{r}' \\
 &\quad + E_{\text{xc}}[\rho_0 + \Delta\rho] + E_{\text{nn}} \quad , \tag{C.9}
 \end{aligned}$$

where multiple operations have been performed simultaneously. These consist of:

1. Writing the product of the density in the Coulomb term explicitly.
2. Extracting the Coulomb term depending on the fluctuations in the first line.
3. Introducing an additional exchange potential solely depending on ρ_0 .
4. Extracting $V_{\text{xc}}[\rho_0 + \Delta\rho]$ from the first line expression.

With these operations performed, the identification of equivalent terms can be used to rewrite and simplify Eq (C.9). The identification of the first line in Eq (C.9) as the energy of the independent reference system given as $E_0 = \sum_i n_i \epsilon_i$, leads to the first crucial simplification. In

C Tight Binding Methods

a second step, it becomes obvious that the exchange potential integrals $\int V_{xc}\rho(\mathbf{r})\mathbf{d}\mathbf{r}$ cancel each other as they were artificially introduced in the first place. The newly obtained Coulomb term depending on the density fluctuation $\Delta\rho$ can be identified within the mixed contributions of the full coulomb term as indicated in blue color. After applying all simplifications, the final expression is given as

$$\begin{aligned}
 E_{\text{DFTB}}[\rho_0 + \Delta\rho] &= \sum_i n_i \langle \Phi_i | \hat{\mathbf{H}}_0 | \Phi_i \rangle \\
 &- \frac{1}{2} \int' \int \frac{\rho_0(\mathbf{r}')\rho_0(\mathbf{r})}{|\mathbf{r}' - \mathbf{r}|} \mathbf{d}\mathbf{r}\mathbf{d}\mathbf{r}' + \frac{1}{2} \int' \int \frac{\Delta\rho(\mathbf{r}')\Delta\rho(\mathbf{r})}{|\mathbf{r}' - \mathbf{r}|} \mathbf{d}\mathbf{r}\mathbf{d}\mathbf{r}' \\
 &- \int V_{xc}[\rho_0(\mathbf{r})]\rho(\mathbf{r})\mathbf{d}\mathbf{r} + E_{xc}[\rho_0 + \Delta\rho] + E_{\text{nn}} \quad . \quad (\text{C.10})
 \end{aligned}$$

The term $E_{xc}[\rho_0 + \Delta\rho]$ can now be expanded in a Taylor series around the reference density ρ_0 as

$$\begin{aligned}
 E_{xc}[\rho_0 + \Delta\rho] &= E_{xc}[\rho_0] + \int \left[\frac{\delta E_{xc}}{\delta \rho} \right]_{\rho_0} \Delta\rho \mathbf{d}\mathbf{r} + \frac{1}{2} \int' \int \left[\frac{\delta^2 E_{xc}}{\delta \rho \delta \rho'} \right]_{\rho_0, \rho'_0} \Delta\rho \Delta\rho' \mathbf{d}\mathbf{r}\mathbf{d}\mathbf{r}' \\
 &+ \frac{1}{6} \int'' \int' \int \left[\frac{\delta^3 E_{xc}}{\delta \rho \delta \rho' \delta \rho''} \right]_{\rho_0, \rho'_0, \rho''_0} \Delta\rho \Delta\rho' \Delta\rho'' \mathbf{d}\mathbf{r}\mathbf{d}\mathbf{r}'\mathbf{d}\mathbf{r}'' + \dots \quad (\text{C.11})
 \end{aligned}$$

In the most sophisticated version of DFTB, the Taylor series is used up to third order and plugging Eq (C.11) into the reformulated energy expression Eq (C.10) leads to the full DFTB3 energy expression Eq (4.19).

C.3 Compressed Atomic Basis Functions

The basis functions used for the expansion in atomic orbitals, are determined by solving a system of KS-equations with an additional confinement potential. The KS-orbitals describing the valence electrons are then expanded in a minimal LCAO basis set whereas the core electrons are described by pseudo potentials (compare Section 2.10). The orbitals are then determined by solving the KS-problem in an additional confining potential leading to compressed atomic orbitals^{149,152,260}. TB formalism is an approximation to KS-DFT depending on the choice of an appropriate input density, leading to an optimization problem for the reference density, in

order to minimize the error with respect to the ground state density. The compressed atomic orbitals are determined by the modified KS-problem as

$$\left[-\frac{1}{2}\nabla^2 + v^{\text{eff}}[\rho^{\text{atom}}] + \left(\frac{r}{r_c}\right)^p \right] \phi_i = e_i \phi_i \quad , \quad (\text{C.12})$$

with the additional compression radius r_c which is typically chosen as two times the atomic radius of the reference atom. The exponent parameter chosen as $p = 2$ has little influence on the overall result as was shown in¹⁵².

C.4 Cyclic Cluster Model

The CCM model is based on a Γ -point ($\mathbf{k} = 0$) approach which is suitable for systems consisting of large unit-cells, as described in Section 2.9. The cyclic cluster model is widely used for semi empirical and DFT applications alike and its concepts are described in detail in multiple publications^{261–266}. The CCM combines the advantages of a periodic implementation within a finite model approach. The translational symmetry is conserved by the Bloch theorem Eq (2.71), whereas the use of BvK-boundary conditions Eq (3.9) avoids the occurrence of edge effects. A Wigner-Seitz-Super-Cell (WSSC) is constructed for each atom respectively, determining the field of interaction for the electrostatics. The WSSC is constructed by translating the atoms of the primitive unit cell by the primitive lattice vectors.

When a corresponding atom is outside of the WSSC, a periodic image is created for the explicit treatment. For the correct treatment of long range interactions, the Ewald summation method has been used. A review about the method and related tasks is given in²⁶⁷. For each atom within the cyclic cluster, the integrals are solved within the interaction space defined by the corresponding WSSC. As the system has to fulfill the BvK-Boundary conditions Eq (3.9), which leads to identical wave functions for translation by a primitive lattice vector, resulting in a formal cyclic cluster¹⁴⁴.

Appendix D

SAM-Basics

D.1 SAM Densities for Au (111)

SAM-densities calculated following Eq (8.2) with the surface area calculated as Eq (8.3). The in plane lattice vector of gold was calculated to be $a_{\text{Au}} = 2.88 \text{ \AA}$. The lengths of the unit cell vectors are given as integer multiples of the lattice parameter as $a = n_a \cdot a_{\text{Au}}$ and $b = n_b \cdot a_{\text{Au}}$. A suitable model to fit experimentally obtained data can be obtained from the following tables.

Table D.1: Inverse densities ρ_{SAM}^{-1} for one SAM molecule per unit cell, depending on the cell size in integer multiples n_a, n_b of the in-plane cell vectors for gold $a = b = 2.88 \text{ \AA}$.

| $n_a \cdot n_b$ | 1 | 2 | 3 | 4 |
|-----------------|-------|--------|--------|---------|
| 1 | 7.200 | 14.366 | 21.549 | 28.732 |
| 2 | 14.40 | 28.732 | 43.098 | 57.465 |
| 3 | 21.60 | 43.098 | 64.648 | 86.197 |
| 4 | 28.80 | 57.465 | 86.197 | 114.930 |

Table D.2: The inverse densities ρ_{SAM}^{-1} for two SAM molecules per unit cell, depending on the cell size given in integer multiples n_a, n_b of the in plane cell vectors for gold $a = b = 2.88 \text{ \AA}$.

| $n_a \cdot n_b$ | 1 | 2 | 3 | 4 |
|-----------------|-------|--------|--------|--------|
| 1 | 3.600 | 7.1831 | 10.774 | 14.366 |
| 2 | 7.200 | 14.366 | 21.549 | 28.732 |
| 3 | 10.80 | 21.549 | 32.324 | 43.098 |
| 4 | 14.40 | 28.732 | 43.098 | 57.465 |

D.2 SAM Densities for Cu and Pt (111)

SAM-densities calculated using Eq (8.2), with the surface area calculated following Eq (8.3). The lattice constants used for calculation of the reference area, are chosen to be $a_{\text{Pt}} = 2.78 \text{ \AA}$ for platinum and $a_{\text{Cu}} = 2.55 \text{ \AA}$ for copper. The lengths of the unit cell vectors are given as integer multiples of the lattice parameter as $a = n_a \cdot a_{\text{Au}}$ and $b = n_b \cdot a_{\text{Au}}$. A suitable model to fit the experimental data, can be chosen in comparison with the following tables.

Table D.3: The inverse densities ρ_{SAM}^{-1} for one SAM molecule per unit cell, depending on the cell size given in integer multiples n_a, n_b of the in plane cell vectors for platinum $a = b = 2.78 \text{ \AA}$.

| $n_a \cdot n_b$ | 1 | 2 | 3 | 4 |
|-----------------|--------|--------|--------|---------|
| 1 | 6.693 | 13.385 | 20.078 | 26.771 |
| 2 | 13.385 | 26.771 | 40.157 | 53.543 |
| 3 | 20.078 | 40.157 | 60.236 | 80.315 |
| 4 | 26.771 | 53.543 | 80.315 | 107.087 |

Table D.4: The inverse densities ρ_{SAM}^{-1} for two SAM molecule per unit cell, depending on the cell size given in integer multiples n_a, n_b of the in plane cell vectors for platinum $a = b = 2.78 \text{ \AA}$.

| $n_a \cdot \cdot \cdot n_b$ | 1 | 2 | 3 | 4 |
|-----------------------------|--------|--------|--------|--------|
| 1 | 3.346 | 6.693 | 10.039 | 13.386 |
| 2 | 6.692 | 13.386 | 20.079 | 26.772 |
| 3 | 10.039 | 20.079 | 30.118 | 40.158 |
| 4 | 13.385 | 26.772 | 40.158 | 53.544 |

Table D.5: The inverse densities ρ_{SAM}^{-1} for one SAM molecule per unit cell, depending on the cell size given in integer multiples n_a, n_b of the in plane cell vectors for copper $a = b = 2.55 \text{ \AA}$.

| $n_a \cdot \cdot \cdot n_b$ | 1 | 2 | 3 | 4 |
|-----------------------------|--------|--------|--------|--------|
| 1 | 5.631 | 11.263 | 16.894 | 22.525 |
| 2 | 11.263 | 22.525 | 33.788 | 45.050 |
| 3 | 16.894 | 33.788 | 50.682 | 67.576 |
| 4 | 22.526 | 45.051 | 67.576 | 90.102 |

Table D.6: The inverse densities ρ_{SAM}^{-1} for two SAM molecule per unit cell, depending on the cell size given in integer multiples n_a, n_b of the in plane cell vectors for copper $a = b = 2.55 \text{ \AA}$.

| $n_a \cdot \cdot \cdot n_b$ | 1 | 2 | 3 | 4 |
|-----------------------------|--------|--------|--------|--------|
| 1 | 2.816 | 5.631 | 8.446 | 11.262 |
| 2 | 5.631 | 11.262 | 16.893 | 22.525 |
| 3 | 8.446 | 16.893 | 25.340 | 33.787 |
| 4 | 11.262 | 22.525 | 33.787 | 45.050 |

D.3 Molecular Length and Tilting Angle

Table D.7: Overview of the calculated effective molecular lengths as well as tilting angles for the studied 2×2 SAM-surface arrangements after geometry-optimization.

| | d in Å following Eq (8.1) | $\phi_{2 \times 2}$ in ° |
|--------------------|-----------------------------|--------------------------|
| Au C10 | 11.57 | 26.5 |
| Pt C10 | 11.00 | 28.4 |
| Cu C10 | 12.69 | 54.6 |
| Au NH ₂ | 15.06 | 13.4 |
| Pt NH ₂ | 13.44 | 27.7 |
| Cu NH ₂ | 13.40 | 35.1 |

Table D.8: Overview of the calculated effective molecular lengths as well as tilting angles for the studied 3×2 SAM-surface arrangements after geometry-optimization.

| | d in Å following Eq (8.1) | $\phi_{3 \times 2}$ in ° |
|----------------------|-----------------------------|--------------------------|
| Au 2 C10 | 11.57 | 28.3 |
| | 10.49 | 22.4 |
| Pt 2 C10 | 12.06 | 27.0 |
| | 10.45 | 11.5 |
| Cu 2 C10 | 12.05 | 25.5 |
| | 11.81 | 24.8 |
| Au 2 NH ₂ | 15.49 | 29.7 |
| | 15.30 | 28.5 |
| Pt 2 NH ₂ | 14.92 | 21.8 |
| | 14.16 | 21.0 |
| Cu 2 NH ₂ | 15.64 | 28.8 |

D.4 Interaction Energies for SAMs on Gold (111) Surfaces

Table D.9: Estimates for the interaction energies E_{int} for SAM molecules on the gold (111) surface.

| Metal - SAM | E_{int} following Eq (11.1) in kcal/mol |
|-------------------|--|
| C10 | -79.2 |
| F | 71.25 |
| NH ₂ | -78.47 |
| 2 C10 | -29.76 |
| 2 F | 6.58 |
| 2 NH ₂ | -134.45 |

Table D.10: Estimates for the interaction energies $E_{\text{int}}^{\text{PER}}$ for SAM molecules on the gold (111) surface.

| SAM | $E_{\text{int}}^{\text{PER}}$ following Eq (8.5) in kcal/mol |
|-------------------|--|
| C10 | -49.25 |
| F | -22.75 |
| NH ₂ | -57.38 |
| 2 C10 | -75.85 |
| 2 F | -11.71 |
| 2 NH ₂ | -97.15 |

Table D.11: Estimates for the interaction energies $E_{\text{int}}^{\text{MOL}}$ for SAM molecules on the gold (111) surface.

| SAM | $E_{\text{int}}^{\text{MOL}}$ following Eq (8.6) in kcal/mol |
|-------------------|--|
| 2 C10 | 9.81 |
| 2 F | 34.51 |
| 2 NH ₂ | 3.75 |

D.5 Comparison of Interaction Energies

Table D.12: Estimates for the interaction energies E_{int} its corresponding sulfur-metal bond E_{SM} contribution as well as the interaction arising from the periodic treatment of the SAMs $E_{\text{per}}^{\text{SAM}}$.

| Metal - SAM | E_{int} following Eq (11.1) in a.u. | $E_{\text{per}}^{\text{SAM}}$ following Eq (8.5) in a.u. |
|------------------------|--|--|
| Au - C10 | 0.0785 | 0.1262 |
| Pt - C10 | 0.1036 | 0.1801 |
| Cu - C10 | 0.1580 | 0.1314 |
| Au - NH ₂ | 0.0914 | 0.1251 |
| Pt - NH ₂ | 0.1213 | 0.1860 |
| Cu - NH ₂ | 0.1910 | 0.1314 |
| Au - F | 0.0363 | 0.1275 |
| Pt - F | 0.0281 | 0.1701 |
| Cu - F | 0.0117 | 0.1214 |
| Au - 2 F | 0.0187 | 0.1767 |
| Pt - 2 F | -0.0287 | 0.3129 |
| Cu - 2 F | -0.1923 | 0.2586 |
| Au - 2 NH ₂ | 0.1662 | 0.1820 |
| Pt - 2 NH ₂ | 0.1668 | 0.3157 |
| Cu - 2 NH ₂ | 0.2402 | 0.2619 |
| Au - 2 C10 | 0.1209 | 0.0918 |
| Pt - 2 C10 | 0.1327 | 0.3484 |
| Cu - 2 C10 | 0.2079 | 0.2787 |

Appendix E

Field Evaporation

E.1 Fragmentation Patterns - GFN-xTB

Table E.1: Summary of the obtained fragmentation patterns for the investigated C10-SAM systems.

| molecules per unit cell | fragment | \mathcal{E}_{sim} in a.u. |
|-------------------------|--------------------------------------|------------------------------------|
| 1 | $\text{C}_{10}\text{H}_{21}^+$ | 0.07 - 0.08 |
| 1 - (S-fix) | H^+ | 0.09 - 0.10 |
| | $\text{C}_{10}\text{H}_{20}^+$ | 0.09 - 0.10 |
| 2 | $\text{SC}_{10}\text{H}_{21}^+$ (I) | 0.03 - 0.04 |
| | $\text{SC}_{10}\text{H}_{21}^+$ (II) | 0.05 - 0.06 |

Table E.2: Summary of the obtained fragmentation patterns for the investigated NH_2 -SAM systems.

| molecules per unit cell | fragment | \mathcal{E}_{sim} in a.u. |
|-------------------------|--------------------------------------|------------------------------------|
| 1 | $\text{C}_{11}\text{NH}_{24}^+$ | 0.075 - 0.08 |
| 2 | $\text{C}_{11}\text{NH}_{24}^+$ (I) | 0.035 - 0.04 |
| | $\text{C}_{11}\text{NH}_{24}^+$ (II) | 0.04 - 0.045 |

Table E.3: Summary of the obtained fragmentation patterns for the investigated F-SAM systems.

| molecules per unit cell | fragment | \mathcal{E}_{sim} in a.u. |
|-------------------------|---|------------------------------------|
| 1 | CF_3^+ | 0.06 |
| | $\text{SC}_9\text{F}_{12}\text{H}_4^+$ | 0.075 - 0.08 |
| 2 | CF_3^+ (II) | 0.055 - 0.06 |
| | $\text{SC}_{10}\text{F}_{17}\text{H}_4^+, \text{SC}_9\text{F}_{14}\text{H}_4^+$ | 0.065 - 0.07 |

E.2 Charge Redistribution for the H₂-Example**Table E.4:** Distribution of the Hirschfeld partial charges for the geometry optimized H₂ example system. Calculations were performed utilizing the SCC-implementation using a TZV2PX-MOLOPT-GTH basis set.

| \mathcal{E} in a.u. | q_{H_1} in a.u. | q_{H_2} in a.u. |
|-----------------------|--------------------------|--------------------------|
| 0.000 | -0.000 | 0.000 |
| 0.005 | -0.008 | 0.008 |
| 0.010 | -0.025 | 0.025 |
| 0.015 | -0.047 | 0.047 |
| 0.020 | -0.072 | 0.072 |
| 0.025 | -0.097 | 0.097 |
| 0.030 | -0.116 | 0.116 |
| 0.035 | -0.136 | 0.136 |
| 0.040 | -0.162 | 0.162 |
| 0.045 | -0.183 | 0.183 |
| 0.050 | -0.203 | 0.203 |
| 0.055 | -0.224 | 0.224 |
| 0.060 | -0.251 | 0.251 |
| 0.065 | -0.272 | 0.272 |
| 0.070 | -0.293 | 0.293 |
| 0.075 | -0.314 | 0.314 |
| 0.080 | -0.336 | 0.336 |
| 0.085 | -0.357 | 0.357 |
| 0.090 | -0.379 | 0.379 |
| 0.095 | -0.401 | 0.401 |
| 0.100 | -0.423 | 0.423 |
| 0.105 | -0.445 | 0.445 |
| 0.110 | -0.468 | 0.468 |
| 0.115 | -0.491 | 0.491 |
| 0.120 | -0.514 | 0.514 |
| 0.125 | -0.538 | 0.538 |
| 0.130 | -0.564 | 0.564 |
| 0.135 | -0.591 | 0.591 |
| 0.140 | -0.621 | 0.621 |
| 0.145 | -0.664 | 0.664 |
| Evaporation -H- | - | - |
| 0.150 | - | - |

Appendix F

COF-Basics

F.1 GPW-DFT Benchmarks

N-COF

Benchmark calculations were performed on the standard level for the cell optimized two layered N-COF system in eclipsed configuration (AA). The hexagonal unit-cell is spanned by cell vectors of length $a = b = 22.470 \text{ \AA}$ and $c = 8.101 \text{ \AA}$. (PBE-D3/TZV2P-GTH, PBE-GTH-PPs, $E_{\text{cut}} = 300 \text{ Ry}$ and $E_{\text{rel}} = 50 \text{ Ry}$).

Table F.1: Single point energies E_{el} calculated for various values of the PW-cutoff E_{cut} .

| E_{cut} in Ry | E_{el} in a.u. |
|------------------------|-------------------------|
| 50 | -549.82647 |
| 100 | -556.68630 |
| 150 | -556.76855 |
| 200 | -556.80129 |
| 250 | -556.80097 |
| 300 | -556.80093 |
| 350 | -556.79986 |
| 400 | -556.79985 |
| 450 | -556.79974 |
| 500 | -556.79969 |

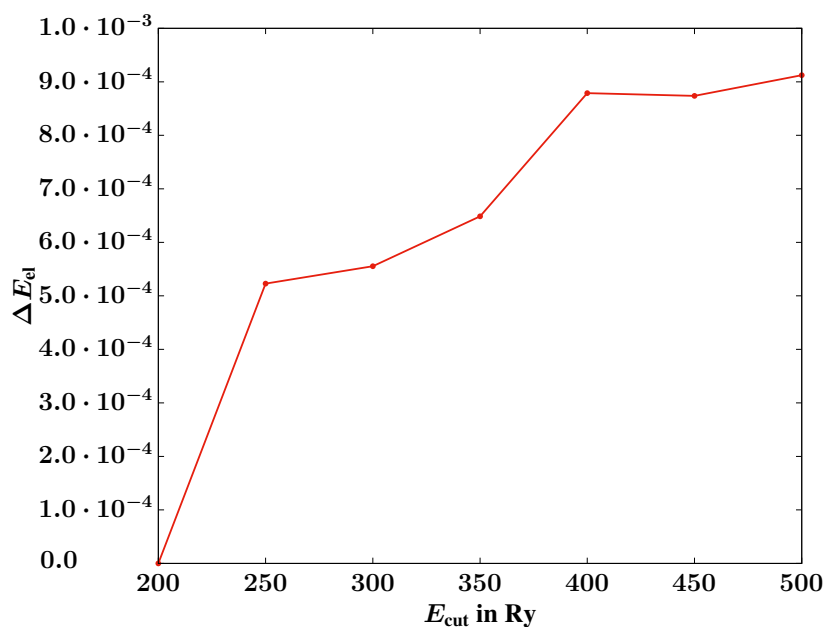
Table F.2: Cell parameters after optimization using the BLYP functional with corresponding PPs.

| Stackings | a in Å | c in Å | ΔE_{el} in kcal/mol |
|-----------|----------|----------|-----------------------------|
| AA | 22.52 | 7.87 | 15.58 |
| AA' | 22.55 | 6.94 | 0.00 |
| AB | 22.48 | 7.12 | 28.36 |
| C | 22.52 | 7.58 | 31.24 |

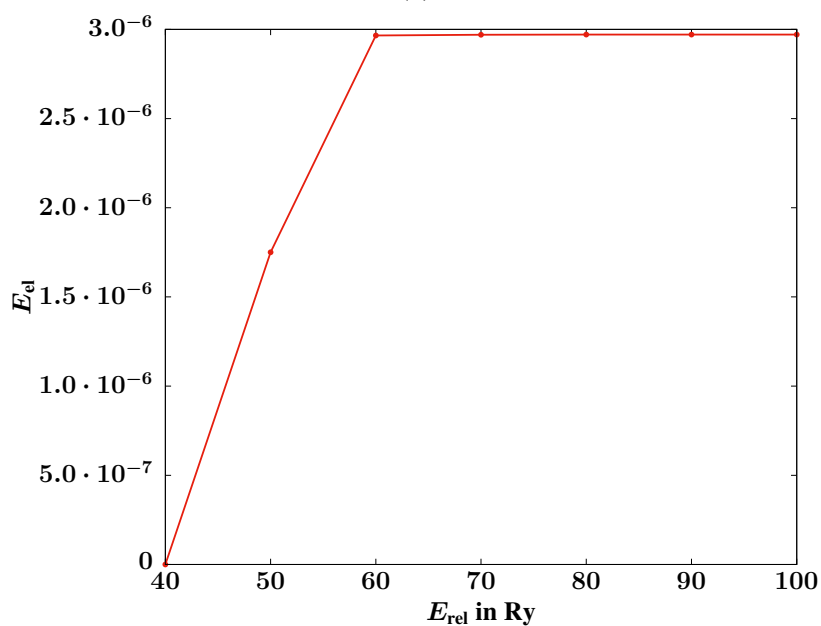
Table F.3: Single point energies E_{el} obtained for various values of E_{rel} using a PW-cutoff of $E_{cut} = 300.0$ Ry.

| E_{rel} in Ry | E_{el} |
|-----------------|------------|
| 10.00 | -557.00112 |
| 20.00 | -556.80906 |
| 30.00 | -556.80102 |
| 40.00 | -556.80095 |
| 50.00 | -556.80093 |
| 60.00 | -556.80093 |
| 70.00 | -556.80093 |
| 80.00 | -556.80093 |
| 90.00 | -556.80093 |
| 100.00 | -556.80093 |

Illustrations of the benchmarks for determination of the cutoffs' impact on the accuracy of single point energy calculations for the two-layered N3-COF example system.



(a)



(b)

Figure F.1: Convergence tests for the relevant cutoff values during setup creation. a) Total energy over the maximum energy of the PW-expansion, b) Influence of the maximum energy used for the integration of Gaussian functions on the finest grid for $E_{\text{cut}} = 300$ Ry.

The relative cutoff of the finest integration grid E_{rel} has to be adjusted accordingly. The results of the convergence tests are illustrated in Figure F.1. The total energies illustrated in Figure F.1a indicate that increasing the PW-cutoff above $E_{\text{cut}} = 300$ Ry doesn't lead to significant improvements with the relative change being around $1.0 \cdot 10^{-4}$ ha. The main argument is given by the increased calculation times for higher cutoffs which make productive runs non-economic. However, the increase to $E_{\text{cut}} = 500$ Ry is considered for high accuracy single point calculations, where the additional accuracy is expected to be necessary (compare Section 10.6.1). Furthermore, the relative cutoffs influence, bench marked for $E_{\text{cut}} = 300$ Ry, is insignificant above $E_{\text{rel}} = 60$ Ry as can be concluded from the obtained total energies shown in Figure F.1b. The relative difference of the order of $1.0 \cdot 10^{-6}$ ha arising due to E_{rel} is significantly lower than for E_{cut} , which is to be expected.

Table F.4: Single point energies E_{el} calculated for increasing basis set accuracy using $E_{\text{cut}} = 300.0$ Ry and $E_{\text{rel}} = 60$ Ry.

| Basis Set | E_{el} in a.u. |
|-------------------|-------------------------|
| SZV-GTH | -549.26165 |
| DZVP-GTH | -556.80093 |
| TZVP-GTH | -557.22062 |
| TZV2P-GTH | -557.34912 |
| QZV2P-GTH | -557.38559 |
| TZV2P-MOLOPT-GTH | -557.41034 |
| TZV2PX-MOLOPT-GTH | -557.48660 |

Table F.5: Single point energies E_{el} calculated for increasing basis set accuracy using $E_{\text{cut}} = 300.0$ Ry and $E_{\text{rel}} = 60$ Ry.

| Basis Set | E_{el} in a.u. |
|-------------------|-------------------------|
| SZV-GTH | -404.20424 |
| DZVP-GTH | -409.74310 |
| TZVP-GTH | -410.03295 |
| TZV2P-GTH | -410.12373 |
| QZV2P-GTH | -410.15448 |
| TZV2P-MOLOPT-GTH | -410.17294 |
| TZV2PX-MOLOPT-GTH | -410.22841 |

N3-COF

Benchmark calculations performed on the standard level for one optimized layer of the N3-COF in a hexagonal unit-cell with cell vectors $a = b = 29.469 \text{ \AA}$ and $c = 7.276 \text{ \AA}$. (PBE-D3/TZV2P-GTH, PBE-GTH-PPs, $E_{\text{cut}} = 300 \text{ Ry}$ and $E_{\text{rel}} = 50 \text{ Ry}$).

Table F.6: Single point energies E_{el} calculated for various values of E_{cut} .

| E_{cut} in Ry | E_{el} in a.u. |
|------------------------|-------------------------|
| 50.00 | -409.48973 |
| 100.00 | -409.82314 |
| 150.00 | -409.74998 |
| 200.00 | -409.74365 |
| 250.00 | -409.74313 |
| 300.00 | -409.74310 |
| 350.00 | -409.74300 |
| 400.00 | -409.74277 |
| 450.00 | -409.74278 |
| 500.00 | -409.74274 |

Table F.7: Single point energies E_{el} obtained for various values of E_{rel} using a PW-cutoff of $E_{\text{cut}} = 300.0 \text{ Ry}$.

| E_{rel} in Ry | E_{el} |
|------------------------|-----------------|
| 10.00 | -410.29980 |
| 20.00 | -409.74590 |
| 30.00 | -409.74322 |
| 40.00 | -409.74313 |
| 50.00 | -409.74313 |
| 60.00 | -409.74313 |
| 70.00 | -409.74313 |
| 80.00 | -409.74313 |
| 90.00 | -409.74313 |
| 100.00 | -409.74313 |

F.2 PW-Fit Coordinates

Table F.8: For the plane wave fit, a grid of 16 points was used in order to span the fundamental triangle comprising the unit-cell. The initial positions of the cell optimizations are determined by shifts from the origin in multiples of the lattice vectors.

| Position in [\mathbf{a}] | Position in [\mathbf{b}] |
|------------------------------|------------------------------|
| 0 | 0 |
| 0.333 | 0.333 |
| 0.277 | 0.111 |
| 0.1 | 0 |
| 0.2 | 0 |
| 0.3 | 0 |
| 0.4 | 0 |
| 0.5 | 0 |
| 0.1 | 0.1 |
| 0.2 | 0.1 |
| 0.3 | 0.1 |
| 0.4 | 0.1 |
| 0.2 | 0.2 |
| 0.3 | 0.2 |
| 0.4 | 0.2 |
| 0.3 | 0.3 |

Appendix G

Analysis of Large Pore COFs (PP-TAB series)

G.1 T-shape analysis

Table G.1: Energy differences ΔE calculated following Eq (11.3), for T-shaped stacking F-PP-linkers for the F-PP-TAB-COF arrangements using GFN-xTB.

| Component Arrangement | ΔE in kcal/mol |
|-----------------------|------------------------|
| Flourinated | 179.65 |
| Unflourinated | 3.03 |

Table G.2: Energy differences ΔE calculated following Eq (11.3), for T-shaped stacking F-PP-linkers for the F-PP-TAB-COF arrangements using GFN-xTB.

| Component Arrangement | ΔE in kcal/mol |
|-----------------------|------------------------|
| S10-90grad-F | -146.703238218 |
| S11-90grad-F | -146.701675612 |
| D10-F | -146.700511910 |
| D11-F | -146.702662672 |
| S10-F | -146.700414118 |
| S11-F | -146.704693902 |

G.2 Band Analysis Data (GPW-DFT)

The analysis of the constructed bands as explained in Section 11.2.4 was conducted using GPW-DFT with the standard setup as explained in. Geometries as well as the cell parameters a and c were optimized. The analysis was performed for all three cases consisting of bands constructed of simple imine linkers the single as well as its the double methoxy modified versions.

Table G.3: Relative energies ΔE with respect to the most stable arrangement after cell optimization for the investigated bands using the single modified methoxy-imine linkers following the nomenclature of Figure 11.10a, using the standard setup for GPW-DFT.

| Component-arrangement | ΔE in [kcal/mol] |
|-----------------------|--------------------------|
| simple-alt-s10 | 0 |
| simple-alt-s11 | 0.03 |
| simple-mixed-s10 | 5.48 |
| simple-mixed-s11 | 5.68 |
| simple-para-s11 | 0.6 |
| simple-para-s10 | 4.71 |
| simple-alt-d10 | 1.94 |
| simple-alt-d11 | 2.03 |
| simple-mixed-d10 | 0.51 |
| simple-mixed-d11 | 0.57 |
| simple-para-d11 | 3.37 |
| simple-para-d10 | 2.18 |

Table G.4: Relative energies ΔE with respect to the most stable arrangement after cell optimization for the investigated bands using the single modified methoxy-imine linkers following the nomenclature of Figure 11.10a, using the standard setup for GPW-DFT.

| Component-arrangement | ΔE in [kcal/mol] |
|-----------------------|--------------------------|
| methoxy-alt-s10 | 0 |
| methoxy-alt-s11 | 21.71 |
| methoxy-mixed-s10 | 1.04 |
| methoxy-mixed-s11 | 1.83 |
| methoxy-para-s11 | 4.42 |
| methoxy-para-s10 | 0.5 |
| methoxy-para-d11 | 18.04 |
| methoxy-para-d10 | 17.16 |
| methoxy-alt-d10 | 13.01 |
| methoxy-alt-d11 | 34.83 |
| methoxy-mixed-d10 | 15.3 |
| methoxy-mixed-d11 | 17.55 |

Table G.5: Relative energies ΔE with respect to the most stable arrangement after cell optimization for the investigated bands using the single modified methoxy-imine linkers following the nomenclature of Figure 11.10a, using the standard setup for GPW-DFT.

| Component-arrangement | ΔE in [kcal/mol] |
|-------------------------|--------------------------|
| 1-methoxy-para-s10-para | 0 |
| 1-methoxy-para-d11-anti | 12.98 |
| 1-methoxy-alt-d11-para | 13.51 |
| 1-methoxy-alt-d11-anti | 25.6 |
| 1-methoxy-alt-s10-anti | 23.24 |
| 1-methoxy-alt-s10-para | 10.26 |
| 1-methoxy-alt-s11-anti | 24.58 |
| 1-methoxy-alt-s11-para | 8.78 |
| 1-methoxy-para-s11-anti | 14.47 |
| 1-methoxy-para-s11-para | 3.17 |
| 1-methoxy-para-s10-anti | 19.09 |
| 1-methoxy-para-d11-para | 13.17 |
| 1-methoxy-alt-d10-anti | 24.00 |
| 1-methoxy-alt-d10-para | 11.80 |
| 1-methoxy-para-d10-anti | 10.02 |
| 1-methoxy-para-d10-para | 14.12 |

G.3 Data for the full COF-analysis

Table G.6: Lattice constants and relative energies with respect to the lowest value for the optimized PP-TAB-COF unit-cells. Calculations were performed on the TZV2P-GTH/PBE-D3 level using GTH-PBE-PPs.

| Component Arrangement | $a = b$ in Å | c in Å |
|-----------------------|--------------|----------|
| PP-TAB-para-s10 | 58.524 | 7.26 |
| PP-TAB-para-s11 | 58.577 | 7.31 |
| PP-TAB-alt-s10 | 58.233 | 7.13 |
| PP-TAB-alt-s11 | 58.145 | 7.21 |
| dPP-TAB-alt-s10 | 58.198 | 7.25 |
| dPP-TAB-alt-s11 | 58.242 | 7.32 |
| dPP-TAB-para-s10 | 58.779 | 7.47 |
| dPP-TAB-para-s11 | 58.608 | 7.37 |

G.4 Data Points for the Stability Analysis

Table G.7: Total energies E_{tot} calculated for optimized two-layered unit cells of the PP-TAB-COF system, for various displacements between the layers.

| x in $ a $ | y in $ b $ | $E_{\text{tot}}(x, y)$ in a.u. | ΔE in kcal/mol |
|--------------|--------------|--------------------------------|------------------------|
| 0 | 0 | -779.247 | 219.628 |
| 0.1 | 0.1 | -779.109 | 306.224 |
| 0.1 | 0 | -779.226 | 232.805 |
| 0.2 | 0.2 | -779.031 | 355.170 |
| 0.2 | 0 | -779.194 | 252.886 |
| 0.277 | 0.111 | -779.164 | 271.711 |
| 0.3 | 0.1 | -779.186 | 257.906 |
| 0.3 | 0 | -779.267 | 207.077 |
| 0.4 | 0.1 | -779.094 | 315.637 |
| 0.4 | 0.2 | -779.209 | 243.473 |
| 0.4 | 0 | -779.093 | 316.264 |
| 0.5 | 0 | -779.076 | 326.932 |
| 0.3 | 0.3 | -779.375 | 139.306 |
| 0.3 | 0.2 | -779.285 | 195.782 |
| 0.333 | 0.3333 | -779.597 | 0.000 |

Table G.8: Total energies E_{tot} calculated for optimized two-layered unit cells of the mPP-TAB-COF system, for various displacements between the layers.

| x in $ a $ | y in $ b $ | $a_{\text{opt}} = b_{\text{opt}}$ | c_{opt} | $E_{\text{tot}}(x, y)$ in a.u. | ΔE in kcal/mol |
|--------------|--------------|-----------------------------------|------------------|--------------------------------|------------------------|
| 0 | 0 | 58.87 | 7.218 | -874.051 | 396.585 |
| 0.1 | 0.1 | 58.19 | 6.325 | -873.948 | 461.219 |
| 0.1 | 0 | 58.54 | 6.915 | -874.04 | 403.488 |
| 0.2 | 0.1 | 56.71 | 6.386 | -873.893 | 495.732 |
| 0.2 | 0.2 | 57.50 | 4.846 | -873.962 | 452.433 |
| 0.2 | 0 | 57.57 | 6.672 | -873.928 | 473.769 |
| 0.277 | 0.111 | 57.82 | 5.382 | -873.907 | 486.946 |
| 0.3 | 0.1 | 56.21 | 4.213 | -874.379 | 190.762 |
| 0.3 | 0.2 | 58.46 | 4.730 | -873.997 | 430.471 |
| 0.3 | 0.3 | 55.76 | 4.120 | -874.618 | 40.7880 |
| 0.3 | 0 | 57.28 | 4.641 | -874.04 | 403.488 |
| 0.333 | 0.333 | 55.05 | 4.093 | -874.683 | 0.000 |
| 0.4 | 0.1 | 51.94 | 6.153 | -873.911 | 484.436 |
| 0.4 | 0.2 | 55.95 | 4.317 | -874.204 | 300.576 |
| 0.4 | 0 | 58.36 | 5.049 | -873.896 | 493.849 |
| 0.5 | 0 | 58.14 | 5.208 | -873.896 | 493.849 |

Table G.9: Total energies E_{tot} calculated for optimized two-layered unit cells of the dPP-TAB-COF system, for various displacements between the layers.

| x in $ a $ | y in $ b $ | $E_{\text{tot}}(x, y)$ in a.u. | ΔE in kcal/mol |
|--------------|--------------|--------------------------------|------------------------|
| 0 | 0 | -968.840 | 0.000 |
| 0.1 | 0.1 | -968.818 | 13.805 |
| 0.1 | 0 | -968.837 | 1.882 |
| 0.2 | 0.1 | -968.681 | 99.773 |
| 0.2 | 0.2 | -968.667 | 108.559 |
| 0.2 | 0 | -968.707 | 83.458 |
| 0.277 | 0.111 | -968.678 | 101.656 |
| 0.3 | 0.1 | -968.781 | 37.023 |
| 0.3 | 0.2 | -968.689 | 94.753 |
| 0.3 | 0.3 | -968.534 | 192.017 |
| 0.3 | 0 | -968.705 | 84.713 |
| 0.333 | 0.333 | -968.507 | 208.960 |
| 0.4 | 0.1 | -968.647 | 121.109 |
| 0.4 | 0.2 | -968.674 | 104.166 |
| 0.4 | 0 | -968.693 | 92.243 |
| 0.5 | 0 | -968.735 | 65.888 |

Table G.10: Total energies E_{tot} were calculated for optimized two-layered unit-cells of the F-PP-TAB-COF system, for various displacements between the layers.

| x in $ \mathbf{a} $ | y in $ \mathbf{b} $ | $E_{\text{tot}}(x, y)$ in a.u. | ΔE in kcal/mol |
|-----------------------|-----------------------|--------------------------------|------------------------|
| 0 | 0 | -917.180 | 0.000 |
| 0.333 | 0.333 | -917.090 | 56.054 |
| 0.277 | 0.111 | -917.053 | 79.399 |
| 0.1 | 0 | -917.168 | 7.577 |
| 0.2 | 0 | -916.968 | 132.645 |
| 0.3 | 0 | -916.924 | 160.434 |
| 0.4 | 0 | -916.938 | 151.991 |
| 0.5 | 0 | -917.013 | 104.407 |
| 0.1 | 0.1 | -917.046 | 84.119 |
| 0.2 | 0.1 | -917.032 | 92.473 |
| 0.3 | 0.1 | -917.061 | 74.354 |
| 0.4 | 0.1 | -916.957 | 139.489 |
| 0.2 | 0.2 | -916.958 | 139.344 |
| 0.4 | 0.2 | -917.084 | 59.837 |
| 0.3 | 0.3 | -917.042 | 86.651 |

Table G.11: Total energies E_{tot} were calculated for optimized two-layered unit-cells of the F-mPP-TAB-COF system, for various displacements between the layers.

| x in $ \mathbf{a} $ | y in $ \mathbf{b} $ | $E_{\text{tot}}(x, y)$ in a.u. | ΔE in kcal/mol |
|-----------------------|-----------------------|--------------------------------|------------------------|
| 0 | 0 | -1011.997 | 0.000 |
| 0.333 | 0.333 | -1011.880 | 73.601 |
| 0.277 | 0.111 | -1011.876 | 76.226 |
| 0.1 | 0 | -1011.853 | 90.538 |
| 0.2 | 0 | -1011.898 | 62.389 |
| 0.3 | 0 | -1011.727 | 169.256 |
| 0.4 | 0 | -1011.726 | 170.022 |
| 0.5 | 0 | -1011.811 | 116.636 |
| 0.1 | 0.1 | -1011.853 | 90.467 |
| 0.3 | 0.1 | -1011.881 | 72.702 |
| 0.4 | 0.1 | -1011.814 | 115.095 |
| 0.2 | 0.2 | -1011.736 | 164.141 |
| 0.3 | 0.2 | -1011.677 | 200.906 |
| 0.4 | 0.2 | -1011.488 | 319.270 |
| 0.3 | 0.3 | -1011.806 | 120.291 |

Table G.12: Total energies E_{tot} calculated for optimized two-layered unit-cells of the F-dPP-TAB-COF system, for various displacements between the layers.

| x in $ \mathbf{a} $ | y in $ \mathbf{b} $ | $E_{\text{tot}}(x, y)$ in a.u. | ΔE in kcal/mol |
|-----------------------|-----------------------|--------------------------------|------------------------|
| 0 | 0 | -1106.796 | 0.000 |
| 0.333 | 0.333 | -1106.501 | 185.330 |
| 0.277 | 0.111 | -1106.623 | 108.631 |
| 0.1 | 0 | -1106.754 | 26.245 |
| 0.2 | 0 | -1106.726 | 43.836 |
| 0.3 | 0 | -1106.624 | 108.067 |
| 0.4 | 0 | -1106.594 | 126.919 |
| 0.5 | 0 | -1106.613 | 114.696 |
| 0.2 | 0.1 | -1106.531 | 166.281 |
| 0.3 | 0.1 | -1106.481 | 197.861 |
| 0.4 | 0.1 | -1106.583 | 133.598 |
| 0.2 | 0.2 | -1106.630 | 104.204 |
| 0.3 | 0.2 | -1106.619 | 111.206 |
| 0.4 | 0.2 | -1106.599 | 123.521 |
| 0.3 | 0.3 | -1106.518 | 174.437 |
| 0.1 | 0.1 | -1106.732 | 40.220 |

List of Figures

| | | |
|-----|---|----|
| 1.1 | Schematic representation of an extended surface model as utilized for the simulation of field evaporation of SAMs. The direction of the applied electric field \mathcal{E} is indicated in orange. | 3 |
| 1.2 | Exemplary illustration of an eclipsed COF-LZU1-pore network filled with CO ₂ | 5 |
| 2.1 | Schematic representation of the lattice description in real space in two dimensions with cell vectors \mathbf{a}_1 and \mathbf{a}_2 | 35 |
| 3.1 | Schematics of two lattice arrangements illustrating the ambiguous dipole definition. | 59 |
| 3.2 | Schematic representation of a cross section along the x -direction showing the resulting total potential $V_{\text{tot}} = V_c + V_{\mathcal{E}}(x)$ | 70 |
| 3.3 | Schematic representation of a periodic potential superimposed with an external field in x direction. Depicted are multiple unit cells showing that the potential in adjacent cells gets displaced by a constant offset. | 71 |
| 3.4 | Schematic representation for possible molecular arrangements within the unit cell in field direction. | 72 |
| 3.5 | The configurations of the HOMO orbital E_i^{H} and LUMO orbital E_i^{L} for both spin states in a) the stable case and b) the unstable case after the field has reached the bond-breaking threshold. | 75 |
| 3.6 | Schematics of possible charge shifts in a homo-nuclear dimer. | 76 |

List of Figures

| | | |
|-----|---|-----|
| 3.7 | Schematic representation of a hypothetical bond profile, for a homo-nuclear dimer, with the bond profile without electric field depicted in black, the modified profile with finite \mathcal{E} is presented in orange. | 77 |
| 4.1 | Schematic representation of the charge treatment for the hierarchically improved DFTB-family of models. | 93 |
| 5.1 | Examples for possible approaches for random path generation. Ordered displacement a) without continuity condition at the boundary, b) with continuity at the boundary due to inherent symmetry. | 99 |
| 5.2 | Periodic displacement with continuity at the boundary due to the utilized optimization via the displacement function Eq (5.4), for $c = 1.2$ | 102 |
| 5.3 | Schematic workflow of the Optimization algorithm used for creation of a periodic displacement path by minimizing the difference function $D(x)$ | 103 |
| 5.4 | Visualization of two examples of FRDG-generated displacement paths for 15 layers of COF-LZU1. The components of the two-dimensional displacement path are illustrated with the view aligned with a) the molecular axis, b) perpendicular to the molecular axis. | 104 |
| 6.1 | Schematic representation of the situation utilized for the atomistic environmental description by RDF. a) Only considering the atomic position of atoms within the cutoff radius R_c around the reference atom a indicated in blue. b) Considering 3 different atom types indicated by their respective colors. An analogous procedure is used for the ADF ¹⁷⁰ | 108 |
| 6.2 | Exemplary arbitrarily chosen radial distribution function considering atoms at the delta-functions distances indicated in purple. The convolutions with a Gaussian function of width 0.01 and 0.05 are indicated in green and blue respectively. | 109 |
| 6.3 | Overview of the methods for expansion of the convoluted RDF a) The first ten Chebychev polynomials renormalized onto the interval $[0, 5]$, b) Expansion of the convoluted radial distribution function with 100 polynomials. | 112 |

| | | |
|-----|---|-----|
| 6.4 | Schematic workflow of the steps in the symmetrization algorithm for determination of the charge-equilibrated system. | 115 |
| 6.5 | Visualization of the symmetry arrangement for the coronene molecule. a) Basic structure of coronene, carbon atoms in gray, hydrogen atoms in white. b) The four obtained equivalence classes for the coronene molecule differentiated by color. | 116 |
| 6.6 | Illustration of the CS treatment for the N-coronene molecule. a) Basic structure of N-coronene, carbon atoms in gray, hydrogen atoms in white and nitrogen atoms in blue b) Equivalence classes for the N-coronene molecule differentiated by color. Using a value of $R_{\text{cut}} = 4.0 \text{ \AA}$ resulted in a total of 12 equivalence classes. c) Increase to $R_{\text{cut}} = 10 \text{ \AA}$, resulting in a total of 20 equivalence classes. | 118 |
| 6.7 | Visualization of the symmetry arrangement for the COF-LZU1 unit-cell. a) Basic structure of the COF-LZU1 unit-cell, carbon atoms are indicated in gray, hydrogen atoms in white and nitrogen atoms in blue. b) The eleven equivalence classes for the COF-LZU1 unit cell differentiated by color. . . . | 119 |
| 6.8 | Example of the CS application on a two layered COF-LZU1 system. a) The structural arrangement of layers within the treated unit cell. b) The obtained eleven equivalence classes for the eclipsed system, differentiated by color. . . | 120 |
| 6.9 | Overview of the multilayer treatment for the shifted COF-LZU1 unit cell. a) The structural alignment of the treated unit cell. b) Visualization of the symmetry arrangement for the shifted COF-LZU1 with $R_{\text{cut}} = 4.0 \text{ \AA}$, thus including the adjacent layer environments, and resulting in a total of 64 equivalence classes. c) Visualization of the equivalence classes for the shifted layers of the COF-LZU1 unit-cell after separated treatment, resulting again in the 11 equivalence classes. | 122 |
| 7.1 | Schematic representation of the described pore size definition, the red circle indicates the pore surface resulting from the circle with largest radius r_{max} fitting inside the pore. | 124 |

List of Figures

| | | |
|-----|--|-----|
| 7.2 | Scheme of an eclipsed stacking type, where the simplest definition of the pore size indicated by the red circle with radius r_{\max} around the pores center is still applicable. | 125 |
| 7.3 | Schematic representation of the situation for shifted layers. The original pore size as resulting from eclipsed stacking is indicated by a dashed orange circle. The geometric pore size for shifted layers, considering point like atoms, is indicated by a red circle with radius r_{\min} . The even smaller effective pore size considering the atoms repulsion of radius r_{real} , is drawn in dashed red with the van-der-Waals radius r_{vdW} around the edge atoms indicated in blue. | 127 |
| 7.4 | The reduced pore size resulting from shifted layers, when calculated from the pores center K_0 for L_2 is indicated in red. The corrected pore size after relocating the center of the constructed sphere to K_{shift} is represented in green. The determined maximum radius determined by consideration of the pores' center is indicated by r_{\max} . However, the maximum diameter of a sphere truly able to propagate through the whole layer arrangement, is indicated by d_{cross} | 127 |
| 7.5 | Schematic work flow of the PSD algorithm. | 129 |
| 7.6 | Exemplary situations of the N-COF unit cell in perfectly eclipsed stacking. a) Top view onto the unit cell, showing the sphere of a one layered pore. b) Extended system with the pore size determined for every individual layer, indicated in blue. | 130 |
| 7.7 | Relevant pore volume, indicated by the overlapping spheres in blue, for a system of 10 stacked and randomly displaced layers of the N3-COF. | 131 |
| 8.1 | Scheme of a SAM with the substrates-surface, consisting of a headgroup, a carbon chain and if required a functional group. | 137 |
| 8.2 | Gas phase optimized geometries on the def2-TZVPP/B3LYP-D3 level for saturated a) decanethiol C10-SAM, b) perfluoro-decanethiol F-SAM and c) amino-undecanethiol NH_2 -SAM. | 138 |
| 8.3 | Gas phase optimized geometries on the def2-TZVPP/B3LYP-D3 level for saturated a) decanethiol C10-SAM, b) perfluoro-decanethiol F-SAM and c) amino-undecanethiol H-SAM. | 140 |

| | | |
|------|---|-----|
| 8.4 | Illustration of the utilized systems for surface construction with a) top view onto the optimized fcc-bulk gold structure in $x - y$ plane, b) view onto the crystal structure of fcc gold after cell optimization, c) reconstructed surface utilizing a three layered 4×4 initial structure. | 143 |
| 8.5 | Schematics of the considerations made for creation of the smallest C10-SAM systems. | 151 |
| 8.6 | The unit cell arrangement after optimization for one molecule for the C10-SAM on a 2×2 gold surface. a) Top view along the surface normal with the cell boundaries indicated in blue. b) side view of the unit-cell, with the periodic images in transparent colors. | 153 |
| 8.7 | Side view onto the unit cell for one decanethiol molecule on a three layered 3×2 gold surface, revealing the increased tilting of the SAM molecules towards the surface. | 153 |
| 8.8 | The unit cell arrangement after optimization for one decanethiol molecule on a two layered 3×2 gold surface. Top view onto the unit cell of decanethiol after optimization, with the cell boundaries in blue. Side view of the decanethiol unit cell, with the periodic images in transparent colors, the cell boundaries are indicated in blue. | 155 |
| 8.9 | The unit cell arrangement after optimization for three C10 molecules a two layered 3×3 gold surface. Top view onto the unit cell of three C10-SAM molecules per unit cell after optimization, with the cell boundaries in blue. b) Side view onto the optimized unit cell including three SAM-molecules. . . . | 156 |
| 8.10 | Schematics of the considerations made for creation of the smallest F-SAM systems. | 159 |
| 8.11 | Unit cell for one F-SAM molecule on a two layered 2×2 gold surface. a) Top view onto the unit cell after optimization, with the cell boundaries in blue. b) Side view of the unit cell, with the periodic images in transparent colors. . . . | 159 |
| 8.12 | Overview of the F-SAM arrangement in the low density regime. | 160 |

List of Figures

- 8.13 Unit cell arrangement for two SAM molecules per unit cell on a two layered 3×2 gold surface. a) Top view onto the unit cell of the perfluoro-decanethiol unit cell after optimization, with the cell boundaries in blue. b) Side view of the perfluoro-decanethiol unit cell, with the periodic images in transparent colors. 161
- 8.14 Schematics of the chemical structure for the NH_2 -SAM molecules attached on a metal substrate. 163
- 8.15 Unit cell arrangement for one NH_2 -SAM molecule on a 2×2 gold surface. a) Top view onto the unit cell of the amino-undecanethiol unit cell after optimization. The cell boundaries in the $x - y$ plane, are depicted with blue lines. b) Side view of the amino-undecanethiol unit cell, with the periodic images in transparent colors. 164
- 8.16 Unit cell arrangement for one NH_2 -SAM molecule on a two layered 3×2 gold surface. a) Top view onto the unit cell of the NH_2 -SAM after optimization. The cell boundaries in the $x - y$ plane, are depicted with blue lines. b) Side view of the unit cell, with the periodic images in transparent colors. 165
- 8.17 Zoom onto the hydrogen bond arrangement in the functional groups for two NH_2 -SAM molecules on a 3×2 gold unit cell, illustrating the formation of molecule pairs. The distance between the donor nitrogen and acceptor hydrogen in angstrom is indicated by a dashed black line. 165
- 8.18 Periodic contribution to the interaction energies in a.u. following Eq (8.5) per SAM molecule, for various SAM-densities. Energies for systems with one molecule per unit cell are indicated by dots. 168
- 9.1 Exemplary bond profile considering the electronic energy E_{el} , calculated for H_2 in the gas phase. Single point energy calculations were performed on the B3LYP-D3/def2-TZVPP level, for an applied field of $\mathcal{E} = 0.04$ a.u. in displacement direction, every 0.01 \AA 173

| | | |
|------|---|-----|
| 9.2 | Bond profiles of H ₂ in vacuum calculations, performed with the TURBO-MOLE program package for various field strengths in displacement direction. The single point energy calculations were performed on the B3LYP-D3/def2-TZVPP level every 0.02 Å. | 174 |
| 9.3 | Bond profiles, obtained by singlepoint energy calculations for displacements r in field direction oriented in parallel to the molecular axis. Results for closed shell calculations are indicated in yellow, open-shell results in dashed turquoise and results considering an applied field of $\mathcal{E} = 0.04$ a.u. in green. | 176 |
| 9.4 | Equilibrium distance r_{eq} of H ₂ , calculated by geometry optimization for incrementally increased field strengths. | 177 |
| 9.5 | Total polarization P calculated at the equilibrium distance for H ₂ for incrementally increased field strength. | 178 |
| 9.6 | Schematics of the experimental APT setup. | 180 |
| 9.7 | Schematic workflow of the loops necessary for the most complicated calculations during the evaporation process. | 186 |
| 9.8 | Setup for the evaporation of gold, a) The reorganization of the surface after relaxation of the uppermost layer at $\mathcal{E} = 0$ a.u., b) Desorption process during the field step $\mathcal{E} = 0.05 - 0.06$ a.u.. | 187 |
| 9.9 | Stages of the desorption of C10-SAM molecules. | 190 |
| 9.10 | Stages of the desorption process for two C10-SAM molecules. | 192 |
| 9.11 | Stages of the desorption process for one NH ₂ -SAM molecule. | 194 |
| 9.12 | Stages of the desorption process for two NH ₂ -SAM molecules. | 197 |
| 9.13 | Stages of the desorption process for one F-SAM molecule. | 199 |
| 9.14 | Initial stages of the desorption process for two F-SAM molecules. | 201 |
| 9.15 | Stages of the continued simulation process. a-b) Simultaneous evaporation of CF ₃ ⁺ from chain II followed by CF ₂ ⁺ from chain I during $\mathcal{E} = 0.04 - 0.045$ a.u. . c) Relaxed structure after removing the fragments at $\mathcal{E} = 0.03$ a.u. . d) The evaporation of CF ₂ ⁺ from chain II during the field step $\mathcal{E} = 0.04 - 0.045$ a.u. . e) Relaxed structure in a field of $\mathcal{E} = 0.045$ a.u., f) Full desorption of the remaining molecules during the field step $\mathcal{E} = 0.045 - 0.05$ a.u. | 203 |

List of Figures

- 9.16 Exemplary situations for the evaporation of gold in a field applied perpendicular to the surface. a) Shows the minor reorganization of the top layer in a field of strength $\mathcal{E} = 0.01$ a.u., b) The situation during the field step $\mathcal{E} = 0.055 - 0.06$ a.u. where two gold atoms evaporate simultaneously. 208
- 9.17 The initial geometry of the simulated 2×2 C10-SAM system at $\mathcal{E} = 0$ 209
- 9.18 Stages of the desorption process for one C10-SAM molecule using GFN-xTB. 210
- 9.19 Stages of the desorption process for two C10-SAM molecules using GFN-xTB. 211
- 9.20 Optimized initial geometry of the smallest NH₂-SAM system after geometry optimization at $\mathcal{E} = 0$ a.u. with GFN-xTB. 212
- 9.21 Stages of the desorption process for one NH₂-SAM molecule using GFN-xTB. 213
- 9.22 Stages of the desorption process for two NH₂-SAM molecules using GFN-xTB. 214
- 9.23 Stages of the desorption process for one F-SAM molecule using GFN-xTB. . . 216
- 9.24 Stages of the desorption process for one C10-SAM molecule using GFN-xTB. 217
-
- 10.1 Overview of the electrostatics for a benzene ring as calculated by means of DFT on the B3LYP/TZVP level using the GAUSSIAN program package. The shown surface was created by mapping the MEP onto an iso surface of the electron density ρ_{el} , created for a density isovalue of $\rho = 0.02$. Red corresponds to a MEP value of -0.02 a.u. and blue with 0.2 a.u. . a) Top view onto the molecule with the in-plane potential lines indicating the hexagonal symmetry. b) Side view onto the molecule with the electrostatic potential arising perpendicular to the molecular axis. 227
- 10.2 The three main types of interactions arising between two benzene molecules depending on their stacking scheme. a) The sandwich stacking scheme. b) The parallel displaced stacking scheme. c) The T-shape stacking scheme. . . 228
- 10.3 Top view onto V_C mapped onto a density iso surface of ρ_{el} for the shifted COF-LZU1 system. Structures were obtained by cell-optimization on the PBE-D3/TZV2P-GTH level with corresponding GTH-PPs for all elements. . . 229

| | | |
|-------|---|-----|
| 10.4 | Molecular arrangement of two benzene rings forming a T-shape arrangement with the MEP mapped onto an iso value of ρ_{el} . The arrangement shows opposite sign charges orienting towards each other after optimization by means of DFT on the B3LYP/TZVP level. | 230 |
| 10.5 | Overview of the relaxed unit cells for the N3-COF for both GGA functionals. a) Top view using the PBE functional, b) Top view using the BLYP-functional, c) Side view using the PBE functional, d) Side view using the BLYP functional. | 232 |
| 10.6 | Deviation ΔE_{el} from the obtained minimum value $E_{\text{el}}^{\text{TZV2PX-MOLOPT}}$ for each basis set respectively. | 234 |
| 10.7 | Side view onto the exemplary AA'-Stacking type investigated for COF-LZU1, with d being the in-plane displacement and Δr_{inter} the interlayer distance. . . | 236 |
| 10.8 | Cell-optimized structure showing the extended layers utilizing PBE/TZV2P-GTH with DFT-D2 correction. | 237 |
| 10.9 | Structures after cell-optimizations using GTH-PPs and a TZV2P-GTH basis set. a) The side view after optimization with the vdW-DF1 (DRSLL) functional. b) The side view after optimization with the vdW-DF2 (LMKLL) functional, c) One optimized layer of the extended system using the RVV10 non-local functional, with the unit-cell boundaries indicated in blue. The second layer is excluded for reasons of visibility. | 239 |
| 10.10 | Unit cells of the N3-COF optimized using GFN-xTB. a) The case of AA stacking leading to $a = b \approx 29.11 \text{ \AA}$ and $c \approx 6.97 \text{ \AA}$, b) The case of AA' stacking $a = b \approx 29.13 \text{ \AA}$ and $c \approx 6.46 \text{ \AA}$ | 241 |
| 10.11 | Side view onto the cells of the COF-LZU1 optimized using GFN-xTB. a) The case of AA stacking leading to $a = b \approx 22.23 \text{ \AA}$ and $c \approx 7.06 \text{ \AA}$, b) The case of AA' stacking $a = b \approx 22.23 \text{ \AA}$ and $c \approx 6.51 \text{ \AA}$ | 241 |

List of Figures

10.12 Schematic representation of the assumptions entering in the PW fit of the energy landscape. The treatment of the hexagonal unit cell shown in red with cell vectors \mathbf{a} and \mathbf{b} was further simplified by using the properties of the D_6 symmetry, which leaves only the green triangle \mathfrak{D} , spanning only 1/12 of the unit cell, for calculation. The remaining part of the unit cell is constructed by mirroring the interpolated triangle \mathfrak{D} as indicated by \mathfrak{D}' and rotation by multiples of $\frac{\pi}{3}$ 244

10.13 Investigations for interlayer displacements of COF-LZU1 following the plane wave interpolation scheme of the obtained values for $E(x, y)$ in Eq (10.2) obtained from both theoretical methods. a) The PES as interpolated from GPW-DFT reference data. b) The PES interpolated from CCM GFN-xTB reference data. 245

10.14 Top view of the COF-LZU1 unit cells after cell-optimization on the GPW-DFT level using the standard setup. a) The case of AA stacking. b) The case of AA' stacking, showing the structures to remain planar after optimization. . . . 248

10.15 Extended analysis of the COFs fundamental structural properties. a) Average pore size $\mathfrak{P}(d)$ obtained using the PSD for FRDG-generated random paths for an in-plane displacement d . The random paths were constructed using 40 layers with interlayer distances in stacking direction of $c = 3.64 \text{ \AA}$ as obtained from the optimized AA' unit cell on the GPW-DFT standard level. b) Symmetry analysis with the CS-algorithm conducted for the COF-LZU1 unit cell, resulting in 12 equivalence classes of 6 atoms each, differentiated by color. 249

10.16 Overview of the cell optimized N3-COF unit cell. a) Top view onto the perfectly stacked AA structure after optimization. b) The N3-COF molecular structure after treatment with the CS algorithm, showing the equivalent atomic environments resulting in thirteen equivalence classes of six atoms each, differentiated by color. 252

| | | |
|-------|--|-----|
| 10.17 | Comparison of the C3-linker with its modified version consisting of an aryl ring. a) The C3-linker configuration with phenyl groups attached to the central ring at 1.3.5 position. b) The N3-linker configuration with the central aryl ring introducing 2.4.6 nitrogens. | 253 |
| 10.18 | Analysis of the interlayer displacement behavior for the N3-COF system. a) Average pore size $\mathfrak{P}(d)$ determined using the PSD for FRDG-generated randompaths using various in-plane displacements d . The randompaths were constructed using the FRDG algorithm for 40 layers with interlayer distances of $c = 3.42 \text{ \AA}$ as obtained from the optimized AA' unit cell. b) The PES determined following Section 10.5.1, the energies corresponding to the respective colors are given in kcal/mol. | 254 |
| 10.19 | Overview of the cell optimized TpPa-COF unit cell for structural analysis. a) The unit cell after optimization of the AA stacking with the in-plane cell boundaries indicated in blue. b) Symmetry analysis with the CS-algorithm conducted for the periodic TpPa-COF unit cell, resulting in twelve equivalence classes consisting of six atoms each, differentiated by color. | 255 |
| 10.20 | The optimized unit-cell structure of TpPa-COF, for the initially AA' structure showing the overlap of oxygen atom and phenyl rings. The same configuration is obtained when (0.2,0.1) is chosen as initial displacement. | 257 |
| 10.21 | Analysis of the interlayer displacement behavior for the TpPa-COF system. a) Average pore size $\mathfrak{P}(d)$ determined using the PSD for FRDG-generated random paths using average in-plane displacements d . The randompaths were constructed for 40 layers with interlayer distances of $c = 3.44 \text{ \AA}$ as obtained from the optimized AA' unit cell. b) The PES determined following Section 10.5.1, the energies corresponding to the respective colors are given in kcal/mol. | 258 |

List of Figures

- 10.22 Overview of the cell optimized V-COF unit cell for structural analysis. a) The unit cell after optimization of the AA stacking with the boundaries of the hexagonal unit cell indicated in blue. b) The unit cell after symmetry analysis with the CS-algorithm, resulting in eleven equivalence classes consisting of six atoms each, differentiated by color. The unit cell with the atoms colored corresponding to their atomic environment as identified by the CS-algorithm. 260
- 10.23 Analysis of the interlayer displacement behavior for the V-COF system. a) Average pore size $\mathfrak{P}(d)$ determined using the PSD for FRDG-generated random paths for various in-plane displacements d . The random paths were constructed for 40 layers with interlayer distances of $c = 3.63 \text{ \AA}$ as obtained from the optimized AA' unit cell. b) The PES determined following Section 10.5.1, the energies corresponding to the respective colors are given in kcal/mol. 262
- 10.24 Overview of the cell optimized NP-COF unit cell for structural analysis. a) Top view onto the AA unit cell with the cell boundaries indicated in blue. b) The unit cells' atoms colored in correspondence to the fifteen equivalent environments comprised of six atoms each, identified by the CS-algorithm. 265
- 10.25 Displacement behavior for the NP-COF system. a) Average pore size $\mathfrak{P}(d)$ determined by PSD for FRDG-generated random paths of 40 layers for various in-plane displacements d . Interlayer distances of $c = 3.84 \text{ \AA}$ extracted from the optimized AA' unit cell were assumed. b) The PES determined following Section 10.5.1, with energies in kcal/mol indicated by color. 266
- 10.26 Overview of the cell optimized NP-m-COF unit cell for structural analysis. a) The unit cell after optimization of the AA stacking with the boundaries of the hexagonal unit cell indicated in blue. b) The unit cell with the atoms colored corresponding to their atomic environment as identified by the CS-algorithm showing 19 different equivalence classes. (A descriptor cutoff of 3.0 \AA was used.) 267

| | | |
|-------|---|-----|
| 10.27 | Analysis of the interlayer displacement behavior for the NP-m-COF system. | |
| | a) Average pore size $\mathfrak{P}(d)$ determined using the PSD for FRDG-generated random paths using average in-plane displacements d . The random paths were constructed for 40 layers with interlayer distances of $c = 3.94 \text{ \AA}$ as obtained from the optimized AA' unit cell. b) The PES determined following Section 10.5.1, the energies corresponding to the respective colors are given in kcal/mol. | 269 |
| 11.1 | Exemplary unit-cell structure for one layer of the unmodified PP-TAB-COF system, forming hexagonal pores, illustrating the general properties of the system. The components comprising the COF are indicated by their respective colors, in blue the C3-unit, in green the imine linker (IL) and red the phenylphenanthridine (PP) component. | 276 |
| 11.2 | The Lewis structures of the used building blocks for the PPD-TAB COF as arising within the complete framework. a) The phenanthridine linker (PP), b) the imine linker (IL), c) the methoxy-imine-linker (MIL). | 277 |
| 11.3 | Structural overview of the building blocks for the PPD-TAB COF after geometry optimization. a) The phenanthridine linker (PP), b) the imine linker (IL), c) the methoxy-imine-linker (MIL). | 278 |
| 11.4 | The MEP mapped onto an iso-surface of the density $\rho(r_{\text{iso}}) = 0.04 \text{ e/Bohr}^3$, calculated with GAUSSIAN for the three components. a) The phenanthridine linker (PP), b) the imine linker (IL), c) the methoxy-imine-linker (MIL). Red areas of the mapped MEP correspond to nucleophilic whereas blue ones correspond to electrophilic regions. The electrostatic potential lines are drawn within the plane with normal vector (0,0,1), to show the potentials impact in the immediate molecular vicinity. Yellow colored potential lines correspond to repulsive, whereas red corresponds to attractive forces on positive charges. . | 278 |

List of Figures

- 11.5 Optimized geometries for the S10 configurations explaining the nomenclature. a) Scheme presenting the used nomenclature for the combination of PP-components resulting in the S10 configuration. b) Scheme presenting the used nomenclature for the combination of PP-components resulting in the D10 configuration. c) Parallel configuration of the PP linkers. (S10) d) T-shape configuration for the PP-linkers. (S10-T) 280
- 11.6 Starting orientations for the optimizations of a) parallel IL (simple-para), b) alternating IL (simple-alt), c) parallel mIL (methoxy-para), d) alternating mIL (methoxy-alt). 282
- 11.7 Setups utilized for the conducted calculations of PP-blocks. a) The schematic idea illustrated by the blocks Lewis structure for the phenantridine components. Red bars indicate the cell boundaries whereas pink arrows indicate the periodic extension via C-C triple bonds. b) The top view onto the exemplary molecular arrangement within an optimized unit-cell for the D10 configuration. 283
- 11.8 Setups utilized for the conducted study of IL-blocks. a) the corresponding scheme for the methoxy-imine-linkers, unmodified structures were investigated analogous, the cell boundaries are indicated in red, whereas pink arrows symbolize periodic continuation. b) The top view onto an optimized example unit-cell for the methoxy-para configuration. 284
- 11.9 Side view onto the optimized geometries for blocks considering T-shaped phenyl linkers after cell-optimization. a) The optimized S11-T configuration. b) The optimized S10-T configuration. 286
- 11.10 Utilized setup for the performed band calculations with a) the schematic idea of the constructed bands, dashed lines indicate possible attachment sites for methoxy-groups. The red bars indicate the cell boundaries, where the periodic extension via the C-C triple bonds is performed in direction of the pink arrows. b) Top view onto an example unit-cell for the methoxy-para-S10 configuration after optimization. 288

| | | |
|-------|---|-----|
| 11.11 | Illustration of the parallel and antiparallel arrangement of the methoxy-groups, with a) the 1-methoxy-para-S10-anti configuration after cell optimization and b) the corresponding 1-methoxy-para-S10-para configuration after cell optimization. | 290 |
| 11.12 | The situation when focussing on one pore wall, showing the systematic orientations for the optimized mPP-TAB-COF unit-cells using GPW-DFT on the standard level. a) The situation in the case of ordered methoxy groups. b) The situation in the case of unordered methoxy groups. | 292 |
| 11.13 | The initial and final structures for the methoxy orientation study on the TZV2P-GTH/PBE-D3-level, a) initial orientation for the ordered structures, b) final structure after cell optimization, c) initial orientation for the unordered arrangement, d) final structure after cell optimization. | 293 |
| 11.14 | The relative energy ΔE with respect to the eclipsed case of the dPP-TAB-COF for twenty distinct structures along the displacement path $AA' \rightarrow AB$. The fraction of the displacement vector is indicated by the number of the respective structure on the x -axis. | 294 |
| 11.15 | Illustration of the most probable eclipsed configuration as well as the potential energy surface determined by PW interpolation for the PP-TAB-COF system. a) The optimized unit-cell the after GPW-DFT optimization on the standard level. b) The PES determined following Section 10.5.1, the energies corresponding to the respective colors are given in kcal/mol. | 296 |
| 11.16 | Illustration of the most probable eclipsed configuration as well as the potential energy surface determined by PW interpolation for the mPP-TAB-COF system. a) The optimized unit-cell the after GPW-DFT optimization on the standard level. b) The PES determined following Section 10.5.1, the energies corresponding to the respective colors are given in kcal/mol. | 297 |

List of Figures

11.17 Illustration of the most probable eclipsed configuration as well as the potential energy surface determined by PW interpolation for the dPP-TAB-COF system. a) The optimized unit-cell the after GPW-DFT optimization on the standard level. b) The PES determined following Section 10.5.1, the energies corresponding to the respective colors are given in kcal/mol. 298

11.18 Illustration of the FRDG-generated random paths used for comparison with the experimental values for the PP-TAB-COF-series. Always two layers generated based on the assumed stacking scheme were randomly displaced using an interlayer displacement of $d = 3.0 \text{ \AA}$. Blue spheres indicate the pore size of the twenty considered layer-pairs used for the determination of $\mathfrak{P}_{\text{PSD}}$. a) The model situation in the case of the unmodified PP-TAB-COF, assuming stronger shifts between the layers simulating an admixture of AA' and AB stacking. b) The model situation in the case of the twice methoxy modified dPP-TAB-COF, assuming AA' stacking. 300

11.19 Illustration of the FRDG-generated random paths used for comparison with the experimental values for the PP-TAPB-COF-series. Always two layers generated based on the assumed stacking scheme were randomly displaced using an interlayer displacement of $d = 3.0 \text{ \AA}$. Blue spheres indicate the pore size of the twenty considered layer-pairs used for the determination of \mathfrak{P} . a) The model situation in the case of the unmodified PP-TAB-COF, assuming stronger shifts between the layers simulating an admixture of AA' and AB-like stacking. b) The model situation in the case of the twice methoxy modified dPP-TAPB-COF, assuming AA' stacking. 301

11.20 The side view onto an example structure showing layer buckling for the GFN-xTB optimized two-layered unit-cell for the dPP-TAB-COF system with (0.2,0) displacements. 302

| | | |
|-------|--|-----|
| 11.21 | Overview of the F-PP-linkers structural properties. a) The geometrical arrangement showing the increased volume of the fluorinated phenyl group. b) The electrostatics illustrated by mapping V_C onto the density iso-surface for $\rho_{\text{iso}} = 0.05$ a.u., with the colors indicating the electrostatic potential arising from the surrounding. | 304 |
| 11.22 | GFN-xTB optimized two component arrangements for the S10-configuration, illustrating both orientations for the fluorinated phenyl groups. a) The case of parallel stacked fluorophenyl linkers (F-S10). b) The case of T-shaped fluorophenyl linkers (F-S10-T). | 305 |
| 11.23 | Example structures showing the T-shaped arrangements within unit-cells for the investigated PP-COFs, optimized by means of GPW-DFT on the standard level (compare Section 10.4.5). a) The optimized PP-COF unit-cell in the eclipsed configuration. b) The optimized F-PP-COF unit-cell in the eclipsed configuration. | 306 |
| 11.24 | Optimized unit-cell and Potential energy surfaces determined by PW interpolation for the F-dPP-TAB-COF species. a) Top view onto the unit-cell. b) The PES determined following Section 10.5.1 with colors indicating the relative energies in kcal/mol. | 309 |
| 11.25 | Optimized unit-cell and Potential energy surfaces determined by PW interpolation for the F-dPP-TAB-COF species. a) Top view onto the unit-cell for the singly methoxy-modified F-mPP-TAB-COF with cell boundaries indicated in blue. b) The PES determined for the singly methoxy-modified mPP-TAB-COF. | 310 |
| 11.26 | Optimized unit-cell and Potential energy surfaces determined by PW interpolation for the F-dPP-TAB-COF species. a) Top view onto the unit-cell for the twice methoxy-modified F-dPP-TAB-COF with cell boundaries indicated in blue. b) The PES determined for the twice methoxy-modified dPP-TAB-COF. | 311 |
| F.1 | Convergence tests for the relevant cutoff values during setup creation. a) Total energy over the maximum energy of the PW-expansion, b) Influence of the maximum energy used for the integration of Gaussian functions on the finest grid for $E_{\text{cut}} = 300$ Ry. | 359 |

List of Tables

| | | |
|-----|--|-----|
| 7.1 | Components of the determined sphere position \mathbf{K} with respect to the pore center and maximal radii r_{\max} for each layer L_i shown in Figure 7.7 from top to bottom respectively. | 131 |
| 8.1 | The molecular length d_{sd} and the carbon-sulfur bond distance d_{CS} of the three SAM-molecules calculated from their molecular geometry optimized on the B3LYP-D3/def2-TZVPP and PBE-D3/def2-TZVPP level respectively. | 139 |
| 8.2 | The spatial distance d_{sd} and the carbon-sulfur bond distance d_{CS} for the radical version of the three SAM-molecules, calculated from their molecular geometry optimized on B3LYP-D3/def2-TZVPP and PBE-D3/def2-TZVPP, respectively using the UKS formalism for unpaired spin orbitals. | 140 |
| 8.3 | Overview of the deviations of calculated in plane lattice vectors lengths from literature values. The cell vectors were extracted from GPW-DFT optimized fcc bulk structures with constrained hexagonal symmetry. | 144 |
| 8.4 | Overview of the lattice parameters deviations from literature in the GFN-xTB optimized fcc bulk structures. | 145 |
| 8.5 | The inverse densities ρ_{SAM}^{-1} utilized in the investigated models depending on the number of SAM molecules N_{SAM} , and the integer multiples n_a, n_b of the cell vectors $a = b = 2.88 \text{ \AA}$ | 147 |
| 9.1 | Distribution of the Hirschfeld partial charges for the C10-SAM systems during incremental field increase. | 189 |

List of Tables

| | | |
|------|---|-----|
| 9.2 | Development of the total bulk Polarization P_{tot} and its components P_i following Eq (3.23), during increase of the external electric field for one C10-SAM molecule. | 190 |
| 9.3 | Molecular tilting and length for the relaxed geometries for incrementally increased external field for the investigated 2×2 C10-SAM system. | 191 |
| 9.4 | Summary of the obtained fragmentation patterns for the investigated C10-SAM systems. | 192 |
| 9.5 | Distribution of the Hirschfeld partial charges for the NH ₂ -SAM systems during incremental field increase. | 195 |
| 9.6 | Development of the total bulk Polarization P_{tot} and its components P_i following Eq (3.23), during increase of the external electric field for one NH ₂ -SAM molecule. | 195 |
| 9.7 | Molecular tilting and length for the relaxed geometries for incrementally increased external field for the investigated 2×2 NH ₂ -SAM system. | 195 |
| 9.8 | Summary of the obtained fragmentation patterns for the investigated NH ₂ -SAM systems. | 196 |
| 9.9 | Distribution of the Hirschfeld partial charges for the F-SAM systems during incremental field increase. | 200 |
| 9.10 | Development of the total bulk Polarization P_{tot} and its components P_i following Eq (3.23), during increase of the external electric field for one NH ₂ -SAM molecule. | 200 |
| 9.11 | Molecular tilting and length for the relaxed geometries for incrementally increased external field for the investigated 2×2 F-SAM system. | 201 |
| 9.12 | Summary of the obtained fragmentation patterns for the investigated F-SAM systems. | 203 |
| 10.1 | The lattice constants for the optimized AA' unit cells for the two GGA functionals. | 233 |
| 10.2 | The lattice constants for the optimized COF-LZU1 AA' unit cells using the PBE functional corrected by the DFT-D series. | 237 |

| | | |
|-------|---|-----|
| 10.3 | The lattice constants for the optimized N3-COF AA' unit cells using the PBE functional corrected by the DFT-D series. | 238 |
| 10.4 | The lattice constants for the optimized COF-LZU1 AA' unit cells using the PBE functional combined with non-local-vdW functionals. | 239 |
| 10.5 | The lattice constants for the optimized N3-COF AA' unit cells using the PBE functional combined with non-local-vdW functionals. | 239 |
| 10.6 | Lattice constants and relative energies with respect to the lowest value for the optimized COF-LZU1 unit cells. Calculations were performed on the TZV2P-GTH/PBE-D3 level using GTH-PBE-PPs. | 248 |
| 10.7 | Calculated DDEC6 net partial charges $\bar{q}(P)$ averaged over each identified equivalence class \mathfrak{F} (the element of is indicated in brackets), for CO ₂ loaded framework super-cells for the COF-LZU1 framework as extracted from GCMC calculated isotherms at $T = 183.15$ K in for different values of the environment pressure P in bar. | 251 |
| 10.8 | Calculated DDEC6 net partial charges $\bar{q}(P)$ averaged over each identified equivalence class \mathfrak{F} (the element of is indicated in brackets), for N ₂ loaded framework super-cells for the COF-LZU1 framework as extracted from GCMC calculated isotherms at $T = 77.36$ K in for different values of the environment pressure P in bar. | 251 |
| 10.9 | Lattice constants and relative energies with respect to the lowest value for the optimized N3-COF unit cells. Calculations were performed on the TZV2P-GTH/PBE-D3 level using GTH-PBE-PPs. | 253 |
| 10.10 | Lattice constants and relative energies with respect to the lowest value for the optimized TpPa-COF unit cells. Calculations were performed on the TZV2P-GTH/PBE-D3 level using GTH-PBE-PPs. | 256 |
| 10.11 | Lattice constants and relative energies with respect to the lowest value for the optimized TpPa-COF unit cells. Calculations were performed on the TZV2P-GTH/BLYP-D3 level using GTH-BLYP-PPs. | 256 |

List of Tables

| | | |
|-------|--|-----|
| 10.12 | Calculated DDEC6 net partial charges $\bar{q}(P)$ averaged over each identified equivalence class \mathfrak{F} (the element of is indicated in brackets), for CO ₂ loaded framework super-cells for the TpPa-COF framework as extracted from GCMC calculated isotherms at $T = K$ in for different values of the environment pressure P in bar. | 259 |
| 10.13 | Lattice constants and relative energies with respect to the lowest value for the optimized V-COF unit cells. Calculations were performed on the TZV2P-GTH/PBE-D3 level using GTH-PBE-PPs. | 261 |
| 10.14 | Lattice constants and relative energies with respect to the most stable configuration for adj-COF unit cells, cell-optimized using the GPW standard setup. | 266 |
| 10.15 | Lattice constants and relative energies with respect to the lowest value for the optimized vdW-COF unit cells. Calculations were performed on the TZV2P-GTH/PBE-D3 level using GTH-PBE-PPs. | 268 |
| 11.1 | Interaction energies determined by Eq (11.1) for the investigated combinations of phenylphenanthridine linkers. | 281 |
| 11.2 | Interaction energies determined by Eq (11.1) for the investigated combinations of imine linkers. | 282 |
| 11.3 | Interaction energies determined by Eq (11.1) for the investigated combinations of phenylphenanthridine linkers. | 284 |
| 11.4 | Interaction energies determined by Eq (11.1) for the investigated combinations of phenylphenanthridine linkers. | 285 |
| 11.5 | Energetic differences between both S- and S-T-block arrangements calculated following Eq (11.3). | 286 |
| 11.6 | Energetic differences $\Delta E_{m\text{-order}}$ between ordered and unordered structures following Eq (11.4) for various levels of description. | 292 |
| 11.7 | Calculated pore sizes $\mathfrak{P}_{\text{PSD}}$ based on the PSD geometric pore analysis in comparison with experimental values $\mathfrak{P}_{\text{exp}}$ for the investigated PP-TAB and PP-TAPB species. | 299 |

| | | |
|-------|--|-----|
| 11.8 | Lattice constants extracted from the optimized unit-cells as obtained during the sampling of the displacement energy maps for the dPP-TAB-COF system. | 302 |
| 11.9 | Interaction energies E_{int} calculated following Eq (11.1), for the main two component F-PP-linker arrangements. Calculations were performed on the PBE-D3/def2-TZVPP level. | 304 |
| 11.10 | Energy differences ΔE calculated following Eq (11.3), for T-shaped stacking F-PP-linkers for the F-PP-TAB-COF arrangements using GPW-DFT with the standard setup. | 307 |
| 11.11 | Unit-cell parameters as obtained after optimization by means of GPW-DFT using the standard setup, for the investigated PP-COFs as obtained for parallel and T-shaped initial orientations. | 307 |
| D.1 | Inverse densities ρ_{SAM}^{-1} for one SAM molecule per unit cell, depending on the cell size in integer multiples n_a, n_b of the in-plane cell vectors for gold $a = b = 2.88 \text{ \AA}$. | 347 |
| D.2 | The inverse densities ρ_{SAM}^{-1} for two SAM molecules per unit cell, depending on the cell size given in integer multiples n_a, n_b of the in plane cell vectors for gold $a = b = 2.88 \text{ \AA}$. | 348 |
| D.3 | The inverse densities ρ_{SAM}^{-1} for one SAM molecule per unit cell, depending on the cell size given in integer multiples n_a, n_b of the in plane cell vectors for platinum $a = b = 2.78 \text{ \AA}$. | 348 |
| D.4 | The inverse densities ρ_{SAM}^{-1} for two SAM molecule per unit cell, depending on the cell size given in integer multiples n_a, n_b of the in plane cell vectors for platinum $a = b = 2.78 \text{ \AA}$. | 349 |
| D.5 | The inverse densities ρ_{SAM}^{-1} for one SAM molecule per unit cell, depending on the cell size given in integer multiples n_a, n_b of the in plane cell vectors for copper $a = b = 2.55 \text{ \AA}$. | 349 |
| D.6 | The inverse densities ρ_{SAM}^{-1} for two SAM molecule per unit cell, depending on the cell size given in integer multiples n_a, n_b of the in plane cell vectors for copper $a = b = 2.55 \text{ \AA}$. | 349 |

List of Tables

| | | |
|------|---|-----|
| D.7 | Overview of the calculated effective molecular lengths as well as tilting angles for the studied 2×2 SAM-surface arrangements after geometry-optimization. | 350 |
| D.8 | Overview of the calculated effective molecular lengths as well as tilting angles for the studied 3×2 SAM-surface arrangements after geometry-optimization. | 350 |
| D.9 | Estimates for the interaction energies E_{int} for SAM molecules on the gold (111) surface. | 351 |
| D.10 | Estimates for the interaction energies $E_{\text{int}}^{\text{PER}}$ for SAM molecules on the gold (111) surface. | 351 |
| D.11 | Estimates for the interaction energies $E_{\text{int}}^{\text{MOL}}$ for SAM molecules on the gold (111) surface. | 351 |
| D.12 | Estimates for the interaction energies E_{int} its corresponding sulfur-metal bond E_{SM} contribution as well as the interaction arising from the periodic treatment of the SAMs $E_{\text{per}}^{\text{SAM}}$. | 352 |
| E.1 | Summary of the obtained fragmentation patterns for the investigated C10-SAM systems. | 353 |
| E.2 | Summary of the obtained fragmentation patterns for the investigated NH ₂ -SAM systems. | 353 |
| E.3 | Summary of the obtained fragmentation patterns for the investigated F-SAM systems. | 354 |
| E.4 | Distribution of the Hirschfeld partial charges for the geometry optimized H ₂ example system. Calculations were performed utilizing the SCC-implementation using a TZV2PX-MOLOPT-GTH basis set. | 355 |
| F.1 | Single point energies E_{el} calculated for various values of the PW-cutoff E_{cut} . | 357 |
| F.2 | Cell parameters after optimization using the BLYP functional with corresponding PPs. | 358 |
| F.3 | Single point energies E_{el} obtained for various values of E_{rel} using a PW-cutoff of $E_{\text{cut}} = 300.0$ Ry. | 358 |
| F.4 | Single point energies E_{el} calculated for increasing basis set accuracy using $E_{\text{cut}} = 300.0$ Ry and $E_{\text{rel}} = 60$ Ry. | 360 |

| | | |
|-----|---|-----|
| F.5 | Single point energies E_{el} calculated for increasing basis set accuracy using $E_{\text{cut}} = 300.0$ Ry and $E_{\text{rel}} = 60$ Ry. | 360 |
| F.6 | Single point energies E_{el} calculated for various values of E_{cut} | 361 |
| F.7 | Single point energies E_{el} obtained for various values of E_{rel} using a PW-cutoff of $E_{\text{cut}} = 300.0$ Ry. | 361 |
| F.8 | Coordinates for the PW-fit. | 362 |
| G.1 | Energy differences ΔE calculated following Eq (11.3), for T-shaped stacking F-PP-linkers for the F-PP-TAB-COF arrangements using GFN-xTB. | 363 |
| G.2 | Energy differences ΔE calculated following Eq (11.3), for T-shaped stacking F-PP-linkers for the F-PP-TAB-COF arrangements using GFN-xTB. | 363 |
| G.3 | Relative energies ΔE with respect to the most stable arrangement after cell optimization for the investigated bands using the single modified methoxyimine linkers following the nomenclature of Figure 11.10a, using the standard setup for GPW-DFT. | 364 |
| G.4 | Relative energies ΔE with respect to the most stable arrangement after cell optimization for the investigated bands using the single modified methoxyimine linkers following the nomenclature of Figure 11.10a, using the standard setup for GPW-DFT. | 365 |
| G.5 | Relative energies ΔE with respect to the most stable arrangement after cell optimization for the investigated bands using the single modified methoxyimine linkers following the nomenclature of Figure 11.10a, using the standard setup for GPW-DFT. | 365 |
| G.6 | Lattice constants and relative energies with respect to the lowest value for the optimized PP-TAB-COF unit-cells. Calculations were performed on the TZV2P-GTH/PBE-D3 level using GTH-PBE-PPs. | 366 |
| G.7 | Total energies E_{tot} calculated for optimized two-layered unit cells of the PP-TAB-COF system, for various displacements between the layers. | 366 |
| G.8 | Total energies E_{tot} calculated for optimized two-layered unit cells of the mPP-TAB-COF system, for various displacements between the layers. | 367 |

List of Tables

G.9 Total energies E_{tot} calculated for optimized two-layered unit cells of the dPP-TAB-COF system, for various displacements between the layers. 367

G.10 Total energies E_{tot} were calculated for optimized two-layered unit-cells of the F-PP-TAB-COF system, for various displacements between the layers. 368

G.11 Total energies E_{tot} were calculated for optimized two-layered unit-cells of the F-mPP-TAB-COF system, for various displacements between the layers. 368

G.12 Total energies E_{tot} calculated for optimized two-layered unit-cells of the F-dPP-TAB-COF system, for various displacements between the layers. 369

Bibliography

- (1) Kreuzer, H. J. *Chemistry and Physics of Solid Surfaces VIII*; Springer, 1990; pp 133–158.
- (2) Miller, M. K.; Forbes, R. G. *Mater. Charact.* **2009**, *60*, 461–469.
- (3) Bas, P.; Bostel, A.; Deconihout, B.; Blavette, D. *App. Surf. Sci.* **1995**, *87*, 298–304.
- (4) Larson, D. J.; Gault, B.; Geiser, B. P.; De Geuser, F.; Vurpillot, F. *Curr. Opin. Solid State Mater. Sci.* **2013**, *17*, 236–247.
- (5) Fletcher, C.; Moody, M. P.; Haley, D. *J. Phys. D Appl. Phys.* **2019**, *52*, 435305.
- (6) Kelly, T. F.; Larson, D. J. *Annu. Rev. Mater. Res.* **2012**, *42*, 1–31.
- (7) Amouyal, Y.; Schmitz, G. *MRS Bull.* **2016**, *41*, 13–18.
- (8) Zhang, Y.; Hillier, A. C. *Anal. Chem.* **2010**, *82*, 6139–6147.
- (9) Eder, K.; Felfer, P. J.; Gault, B.; Ceguerra, A. V.; La Fontaine, A.; Masters, A. F.; Maschmeyer, T.; Cairney, J. M. *Langmuir* **2017**, *33*, 9573–9581.
- (10) Nickerson, B. Kinetics of Field Evaporation, a Study of Molecules in High Electric Fields. M.Sc. thesis, Dalhousie University, 2015.
- (11) Schwarz, T.; Weikum, E.; Meng, K.; Hadjixenophontos, E.; Dietrich, C.; Kästner, J.; Stender, P.; Schmitz, G. *Sci. Rep.* **2020**, *10*, 20271.
- (12) Meng, K.; Schwarz, T. M.; Weikum, E. M.; Stender, P.; Schmitz, G. *Microsc. Microanal.* **2021**, 1–11, DOI: 10.1017/S143192762101254X.

G Bibliography

- (13) Love, J. C.; Estroff, L. A.; Kriebel, J. K.; Nuzzo, R. G.; Whitesides, G. M. *Chem. Rev.* **2005**, *105*, 1103–1170.
- (14) Onclin, S.; Ravoo, B. J.; Reinhoudt, D. N. *Angew. Chem. Int. Ed.* **2005**, *44*, 6282–6304.
- (15) Stoffers, A.; Oberdorfer, C.; Schmitz, G. *Langmuir* **2012**, *28*, 56–59.
- (16) Mohanty, S. K.; Tolochko, O. *App. Surf. Sci.* **2019**, *494*, 152–161.
- (17) Solodenko, H.; Stender, P.; Schmitz, G. *Microsc. Microanal.* **2021**, 1–10, DOI: 10.1017/S1431927621012654.
- (18) Nickerson, B.; Karahka, M.; Kreuzer, H. *Ultramicroscopy* **2015**, *159*, 173–177.
- (19) Dietrich, C. A.; Schuldt, R.; Born, D.; Solodenko, H.; Schmitz, G.; Kästner, J. *J. Phys. Chem. A* **2020**, *124*, 8633–8642.
- (20) Cote, A. P.; Benin, A. I.; Ockwig, N. W.; O’Keeffe, M.; Matzger, A. J.; Yaghi, O. M. *Science* **2005**, *310*, 1166–1170.
- (21) Huang, N.; Wang, P.; Jiang, D. *Nat. Rev. Mater.* **2016**, *1*, 16068.
- (22) Nagai, A. *Covalent Organic Frameworks*; Jenny Stanford Publishing: Singapore, 2020.
- (23) Diercks, C. S.; Lin, S.; Kornienko, N.; Kapustin, E. A.; Nichols, E. M.; Zhu, C.; Zhao, Y.; Chang, C. J.; Yaghi, O. M. *J. Am. Chem. Soc.* **2018**, *140*, 1116–1122.
- (24) Wang, X.; Chen, L.; Chong, S. Y.; Little, M. A.; Wu, Y.; Zhu, W.-H.; Clowes, R.; Yan, Y.; Zwiijnenburg, M. A.; Sprick, R. S. *Nat. Chem.* **2018**, *10*, 1180–1189.
- (25) Chen, R.; Wang, Y.; Ma, Y.; Mal, A.; Gao, X.-Y.; Gao, L.; Qiao, L.; Li, X.-B.; Wu, L.-Z.; Wang, C. *Nat. Commun.* **2021**, *12*, 1354.
- (26) Biswal, B. P.; Vignolo-Gonzalez, H. A.; Banerjee, T.; Grunenberg, L.; Savasci, G.; Gottschling, K.; Nuss, J.; Ochsenfeld, C.; Lotsch, B. V. *J. Am. Chem. Soc.* **2019**, *141*, 11082–11092.

- (27) Gottschling, K.; Savasci, G.; Vignolo-González, H.; Schmidt, S.; Mauker, P.; Banerjee, T.; Rovó, P.; Ochsenfeld, C.; Lotsch, B. V. *J. Am. Chem. Soc.* **2020**, *142*, 12146–12156.
- (28) Sharma, R. K.; Yadav, P.; Yadav, M.; Gupta, R.; Rana, P.; Srivastava, A.; Zbořil, R.; Varma, R. S.; Antonietti, M.; Gawande, M. B. *Mater. Horiz.* **2020**, *7*, 411–454.
- (29) Huang, N.; Chen, X.; Krishna, R.; Jiang, D. *Angew. Chem. Int. Ed.* **2015**, *54*, 2986–2990.
- (30) Furukawa, H.; Yaghi, O. M. *J. Am. Chem. Soc.* **2009**, *131*, 8875–8883.
- (31) Gao, Q.; Li, X.; Ning, G.-H.; Xu, H.-S.; Liu, C.; Tian, B.; Tang, W.; Loh, K. P. *Chem. Mater.* **2018**, *30*, 1762–1768.
- (32) Wang, Z.; Zhang, S.; Chen, Y.; Zhang, Z.; Ma, S. *Chem. Soc. Rev.* **2020**, *49*, 708–735.
- (33) Nguyen, H. L.; Hanikel, N.; Lyle, S. J.; Zhu, C.; Proserpio, D. M.; Yaghi, O. M. *J. Am. Chem. Soc.* **2020**, *142*, 2218–2221.
- (34) Zhai, L.; Huang, N.; Xu, H.; Chen, Q.; Jiang, D. *Chem. Commun.* **2017**, *53*, 4242–4245.
- (35) Stegbauer, L.; Hahn, M. W.; Jentys, A.; Savasci, G.; Ochsenfeld, C.; Lercher, J. A.; Lotsch, B. V. *Chem. Mater.* **2015**, *27*, 7874–7881.
- (36) Xu, F.; Yang, S.; Chen, X.; Liu, Q.; Li, H.; Wang, H.; Wei, B.; Jiang, D. *Chem. Sci.* **2019**, *10*, 6001–6006.
- (37) Singh, H.; Tomer, V. K.; Jena, N.; Bala, I.; Sharma, N.; Nepak, D.; De Sarkar, A.; Kailasam, K.; Pal, S. K. *J. Mater. Chem. A* **2017**, *5*, 21820–21827.
- (38) Song, Y.; Sun, Q.; Aguila, B.; Ma, S. *Adv. Sci.* **6**, 1801410.
- (39) Li, J.; Zhou, X.; Wang, J.; Li, X. *Ind. Eng. Chem. Res.* **2019**, *58*, 15394–15406.
- (40) Zhao, F.; Liu, H.; Mathe, S. D.; Dong, A.; Zhang, J. *Nanomaterials* **2018**, *8*, 15.
- (41) Song, Y.; Sun, Q.; Aguila, B.; Ma, S. *Adv. Sci.* **2019**, *6*, 1801410.

G Bibliography

- (42) Mendoza-Cortes, J. L.; Pascal, T. A.; Goddard III, W. A. *J. Phys. Chem. A* **2011**, *115*, 13852–13857.
- (43) Zhang, G.; Li, X.; Liao, Q.; Liu, Y.; Xi, K.; Huang, W.; Jia, X. *Nat. Commun.* **2018**, *9*, 2785.
- (44) Liu, S.; Hu, C.; Liu, Y.; Zhao, X.; Pang, M.; Lin, J. *Chem. Eur. J.* **2019**, *25*, 4315–4319.
- (45) Chen, X.; Geng, K.; Liu, R.; Tan, K. T.; Gong, Y.; Li, Z.; Tao, S.; Jiang, Q.; Jiang, D. *Angew. Chem. Int. Ed.* **2020**, *59*, 5050–5091.
- (46) Tong, M.; Lan, Y.; Yang, Q.; Zhong, C. *Chem. Eng. Sci.* **2017**, *168*, 456–464.
- (47) Tong, M.; Lan, Y.; Qin, Z.; Zhong, C. *J. Phys. Chem. C* **2018**, *122*, 13009–13016.
- (48) Ongari, D.; Yakutovich, A. V.; Talirz, L.; Smit, B. *ACS Cent. Sci.* **2019**, *5*, 1663–1675.
- (49) Jin, S.; Furukawa, K.; Addicoat, M.; Chen, L.; Takahashi, S.; Irle, S.; Nakamura, T.; Jiang, D. *Chem. Sci.* **2013**, *4*, 4505–4511.
- (50) Xu, F.; Jin, S.; Zhong, H.; Wu, D.; Yang, X.; Chen, X.; Wei, H.; Fu, R.; Jiang, D. *Sci. Rep.* **2015**, *5*, 8225.
- (51) Fang, Q.; Zhuang, Z.; Gu, S.; Kaspar, R. B.; Zheng, J.; Wang, J.; Qiu, S.; Yan, Y. *Nat. Commun.* **2014**, *5*, 4503.
- (52) Yu, S.-B.; Lyu, H.; Tian, J.; Wang, H.; Zhang, D.-W.; Liu, Y.; Li, Z.-T. *Polym. Chem.* **2016**, *7*, 3392–3397.
- (53) Zhao, C.; Lyu, H.; Ji, Z.; Zhu, C.; Yaghi, O. M. *J. Am. Chem. Soc.* **2020**, *142*, 14450–14454.
- (54) Xu, T.; An, S.; Peng, C.; Hu, J.; Liu, H. *Ind. Eng. Chem. Res.* **2020**, *59*, 8315–8322.
- (55) Weber, J.; Bergstrom, L. *Langmuir* **2010**, *26*, 10158–10164.
- (56) Muralidharan, V.; Hui, C.-Y. *Macromol. Rapid Commun.* **2004**, *25*, 1487–1490.

- (57) Deng, H.; Grunder, S.; Cordova, K. E.; Valente, C.; Furukawa, H.; Hmadeh, M.; Gándara, F.; Whalley, A. C.; Liu, Z.; Asahina, S. *Science* **2012**, *336*, 1018–1023.
- (58) Vyas, V. S.; Vishwakarma, M.; Moudrakovski, I.; Haase, F.; Savasci, G.; Ochsenfeld, C.; Spatz, J. P.; Lotsch, B. V. *Adv. Mater.* **2016**, *28*, 8749–8754.
- (59) Pütz, A. M.; Terban, M. W.; Bette, S.; Haase, F.; Dinnebier, R. E.; Lotsch, B. V. *Chem. Sci.* **2020**, *11*, 12647–12654.
- (60) Emmerling, S. T.; Schuldt, R.; Bette, S.; Yao, L.; Dinnebier, R. E.; Kästner, J.; Lotsch, B. V. *J. Am. Chem. Soc.* **2021**, *143*, 15711–15722.
- (61) Grimme, S.; Bannwarth, C.; Shushkov, P. *J. Chem. Theory Comput.* **2017**, *13*, 1989–2009.
- (62) Kühne, T. D.; Iannuzzi, M.; Del Ben, M.; Rybkin, V. V.; Seewald, P.; Stein, F.; Laino, T.; Khaliullin, R. Z.; Schütt, O.; Schiffmann, F. et al. *J. Chem. Phys.* **2020**, *152*, 194103.
- (63) Levine, I. N.; Busch, D. H.; Shull, H. *Quantum chemistry*; Pearson Prentice Hall Upper Saddle River, NJ, 2009; Vol. 6.
- (64) Hohenberg, P.; Kohn, W. *Phys. Rev.* **1964**, *136*, B864.
- (65) Fermi, E. *Z. Phys.* **1928**, *48*, 73–79.
- (66) Thomas, L. H. The calculation of atomic fields. Mathematical proceedings of the Cambridge philosophical society. 1927; pp 542–548.
- (67) Kohn, W.; Sham, L. J. *Phys. Rev.* **1965**, *140*, A1133.
- (68) Ceperley, D. M.; Alder, B. J. *Phys. Rev. Lett.* **1980**, *45*, 566.
- (69) Perdew, J. P.; Wang, Y. *Phys. Rev. B* **1992**, *45*, 13244.
- (70) Perdew, J. P.; Zunger, A. *Phys. Rev. B* **1981**, *23*, 5048.
- (71) Koch, W.; Holthausen, M. C. *A chemist's guide to density functional theory*; John Wiley & Sons, 2015.

G Bibliography

- (72) Wang, Y.; Perdew, J. P. *Phys. Rev. B* **1991**, *43*, 8911.
- (73) Perdew, J. P.; Burke, K.; Ernzerhof, M. *Phys. Rev. Lett.* **1996**, *77*, 3865.
- (74) Levy, M. *Int. J. Quantum Chem.* **1989**, *36*, 617–619.
- (75) Zhang, Y.; Yang, W. *Phys. Rev. Lett.* **1998**, *80*, 890.
- (76) Perdew, J. P.; Ruzsinszky, A.; Csonka, G. I.; Vydrov, O. A.; Scuseria, G. E.; Constantin, L. A.; Zhou, X.; Burke, K. *Phys. Rev. Lett.* **2008**, *100*, 136406.
- (77) Becke, A. D. *Phys. Rev. A* **1988**, *38*, 3098.
- (78) Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785.
- (79) Colle, R.; Salvetti, O. *Theor. Chim. Acta* **1975**, *37*, 329–334.
- (80) Tao, J.; Perdew, J. P.; Staroverov, V. N.; Scuseria, G. E. *Phys. Rev. Lett.* **2003**, *91*, 146401.
- (81) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 1372–1377.
- (82) Adamo, C.; Barone, V. *J. Chem. Phys.* **1999**, *110*, 6158–6170.
- (83) Vosko, S. H.; Wilk, L.; Nusair, M. *Can. J. Phys.* **1980**, *58*, 1200–1211.
- (84) Slater, J. C. *Phys. Rev.* **1930**, *36*, 57.
- (85) Rebhan, E. *Theoretische Physik: Quantenmechanik*; Spektrum, 2008.
- (86) Bloch, F. *Z. Phys.* **1929**, *52*, 555–600.
- (87) Marzari, N.; Mostofi, A. A.; Yates, J. R.; Souza, I.; Vanderbilt, D. *Rev. Mod. Phys.* **2012**, *84*, 1419.
- (88) VandeVondele, J.; Hutter, J. *J. Chem. Phys.* **2007**, *127*, 114105.
- (89) Kratzer, P.; Neugebauer, J. *Front. Chem.* **2019**, *7*, 106.
- (90) Monkhorst, H. J.; Pack, J. D. *Phys. Rev. B* **1976**, *13*, 5188.

- (91) Froyen, S. *Phys. Rev. B* **1989**, *39*, 3168.
- (92) Phillips, J. C.; Kleinman, L. *Phys. Rev.* **1959**, *116*, 287.
- (93) Blöchl, P. E. *Phys. Rev. B* **1994**, *50*, 17953.
- (94) Stoll, H.; Szentpály, L. V.; Fuentealba, P.; Flad, J.; Dolg, M.; Fraschio, F.; Schwerdtfeger, P.; Igel, G.; Preuss, H. *Int. J. Quantum Chem.* **1984**, *26*, 725–727.
- (95) Hamann, D.; Schlüter, M.; Chiang, C. *Phys. Rev. Lett.* **1979**, *43*, 1494.
- (96) Kresse, G.; Hafner, J. *J. Phys. Condens. Matter* **1994**, *6*, 8245.
- (97) Mi, W.; Zhang, S.; Wang, Y.; Ma, Y.; Miao, M. *J. Chem. Phys.* **2016**, *144*, 134108.
- (98) Marefat Khah, A.; Reinholdt, P.; Olsen, J. M. H.; Kongsted, J.; Haättig, C. *J. Chem. Theory Comput.* **2020**, *16*, 1373–1381.
- (99) Goedecker, S.; Teter, M.; Hutter, J. *Phys. Rev. B* **1996**, *54*, 1703.
- (100) Hartwigsen, C.; Goedecker, S.; Hutter, J. *Phys. Rev. B* **1998**, *58*, 3641.
- (101) Krack, M. *Theor. Chem. Acc.* **2005**, *114*, 145–152.
- (102) Hutter, J.; Iannuzzi, M.; Schiffmann, F.; VandeVondele, J. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2014**, *4*, 15–25.
- (103) VandeVondele, J.; Borstnik, U.; Hutter, J. *J. Chem. Theory Comput.* **2012**, *8*, 3565–3573.
- (104) Iannuzzi, M.; Chassaing, T.; Wallman, T.; Hutter, J. *Chimia Int. J. Chem* **2005**, *59*, 499–503.
- (105) Krack, M.; Parrinello, M. *High Perform. Comput. Chem.* **2004**, *25*, 29–51.
- (106) VandeVondele, J.; Krack, M.; Mohamed, F.; Parrinello, M.; Chassaing, T.; Hutter, J. *Comput. Phys. Commun.* **2005**, *167*, 103–128.
- (107) Łazarski, R.; Burow, A. M.; Sierka, M. *J. Chem. Theory Comput.* **2015**, *11*, 3029–3041.
- (108) Genovese, L.; Deutsch, T.; Goedecker, S. *J. Chem. Phys.* **2007**, *127*, 054704.

G Bibliography

- (109) Genovese, L.; Deutsch, T.; Neelov, A.; Goedecker, S.; Beylkin, G. *J. Chem. Phys.* **2006**, *125*, 074105.
- (110) Cerioni, A.; Genovese, L.; Mirone, A.; Sole, V. A. *J. Chem. Phys.* **2012**, *137*, 134108.
- (111) DiLabio, G. A.; Otero-de-la Roza, A. *Reviews in computational chemistry* **2016**, *29*, 1–97.
- (112) Ortmann, F.; Bechstedt, F.; Schmidt, W. *Phys. Rev. B* **2006**, *73*, 205101.
- (113) Atkins, P. W.; Friedman, R. S. *Molecular quantum mechanics*; Oxford university press, 2011.
- (114) Tao, J.; Perdew, J. P.; Ruzsinszky, A. *Proc. Natl. Acad. Sci.* **2012**, *109*, 18–21.
- (115) Klimeš, J.; Michaelides, A. *J. Chem. Phys.* **2012**, *137*, 120901.
- (116) Dion, M.; Rydberg, H.; Schröder, E.; Langreth, D. C.; Lundqvist, B. I. *Phys. Rev. Lett.* **2004**, *92*, 246401.
- (117) Sabatini, R. Non-local correlation in Density Functional Theory. Ph.D. thesis, 2012.
- (118) Vydrov, O. A.; Van Voorhis, T. *J. Chem. Phys.* **2010**, *133*, 244103.
- (119) Resta, R. *Handbook of Materials Modeling: Methods: Theory and Modeling 2020*, 151–181.
- (120) Mosotti, O. F. *Mem. Mat. Fis. Soc. Ital. Sci.(Modena)* **1850**, *24*, 49–74.
- (121) Clausius, R. *Die Mechanische Behandlung der Electricität*; Springer, 1879; pp 62–97.
- (122) Resta, R. *J. Phys. Condens. Matter* **2010**, *22*, 123201.
- (123) Kittel, C.; McEuen, P. *Introduction to solid state physics*; Wiley New York, 1996; Vol. 8.
- (124) Ashcroft, N. W.; Mermin, N. D. *Festkörperphysik*; Oldenbourg Wissenschaftsverlag, 2001.
- (125) Spaldin, N. A. *J. Solid State Chem.* **2012**, *195*, 2–10.

- (126) Resta, R. *Phys. Rev. Lett.* **1998**, *80*, 1800.
- (127) King-Smith, R.; Vanderbilt, D. *Phys. Rev. B* **1993**, *47*, 1651.
- (128) Kohn, W. *Phys. Rev.* **1964**, *133*, A171.
- (129) Resta, R.; Vanderbilt, D. *Phys. Ferroelectr.* **2007**, 31–68.
- (130) Berry, M. V. *Proc. R. Soc. A: Math. Phys. Eng. Sci.* **1984**, *392*, 45–57.
- (131) Zak, J. *Phys. Rev. Lett.* **1989**, *62*, 2747.
- (132) Marzari, N.; Vanderbilt, D. *Phys. Rev. B* **1997**, *56*, 12847.
- (133) Blount, E. I. *Solid State Phys.*; Elsevier, 1962; Vol. 13; pp 305–373.
- (134) Resta, R.; Baldereschi, A. *Phys. Rev. B* **1981**, *23*, 6615.
- (135) Dal Corso, A.; Posternak, M.; Resta, R.; Baldereschi, A. *Phys. Rev. B* **1994**, *50*, 10715.
- (136) Dal Corso, A.; Mauri, F.; Rubio, A. *Phys. Rev. B* **1996**, *53*, 15638.
- (137) Gonze, X.; Lee, C. *Phys. Rev. B* **1997**, *55*, 10355.
- (138) Umari, P.; Pasquarello, A. *Phys. Rev. Lett.* **2002**, *89*, 157602.
- (139) Kunc, K.; Resta, R. *Phys. Rev. Lett.* **1983**, *51*, 686.
- (140) Kirtman, B.; Ferrero, M.; Rérat, M.; Springborg, M. *J. Chem. Phys.* **2009**, *131*, 044109.
- (141) Nunes, R.; Vanderbilt, D. *Phys. Rev. Lett.* **1994**, *73*, 712.
- (142) Mauri, F.; Galli, G.; Car, R. *Phys. Rev. B* **1993**, *47*, 9973.
- (143) Souza, I.; Íñiguez, J.; Vanderbilt, D. *Phys. Rev. Lett.* **2002**, *89*, 117602.
- (144) Thinius, S. Optische Eigenschaften und Oberflächenreaktivität von Chalkopyrit. Ph.D. thesis, Rheinische Friedrich-Wilhelms-Universität Bonn, 2020.
- (145) Gaus, M. Extension and Parametrization of an Approximate Density Functional Method for Organic and Biomolecules. Ph.D. thesis, Karlsruher Institut für Technologie (KIT), 2011.

G Bibliography

- (146) Elstner, M.; Porezag, D.; Jungnickel, G.; Elsner, J.; Haugk, M.; Frauenheim, T.; Suhai, S.; Seifert, G. *Phys. Rev. B* **1998**, *58*, 7260.
- (147) Slater, J. C.; Koster, G. F. *Phys. Rev.* **1954**, *94*, 1498.
- (148) Seifert, G.; Eschrig, H. *Phys. Status Solidi B* **1985**, *127*, 573–585.
- (149) Seifert, G.; Eschrig, H.; Bieger, W. Z. *Phys. Chem* **1986**, *267*, 8.
- (150) Elstner, M. *Theor. Chem. Acc.* **2006**, *116*, 316–325.
- (151) Elstner, M. *J. Phys. Chem. A* **2007**, *111*, 5614–5621.
- (152) Porezag, D.; Frauenheim, T.; Köhler, T.; Seifert, G.; Kaschner, R. *Phys. Rev. B* **1995**, *51*, 12947.
- (153) Seifert, G.; Porezag, D.; Frauenheim, T. *Int. J. Quantum Chem.* **1996**, *58*, 185–192.
- (154) Gaus, M.; Cui, Q.; Elstner, M. *J. Chem. Theory Comput.* **2011**, *7*, 931–948.
- (155) Seifert, G. *J. Phys. Chem. A* **2007**, *111*, 5609–5613.
- (156) Foulkes, W. M. C.; Haydock, R. *Phys. Rev. B* **1989**, *39*, 12520.
- (157) Elstner, M.; Seifert, G. *Phil. Trans. R. Soc. A* **2014**, *372*, 20120483.
- (158) Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. *J. Chem. Phys.* **2010**, *132*, 154104.
- (159) Hehre, W. J.; Stewart, R. F.; Pople, J. A. *J. Chem. Phys.* **1969**, *51*, 2657–2664.
- (160) Bannwarth, C. N. Development and Application of Efficient Methods for the Computation of Electronic Spectra of Large Systems. Ph.D. thesis, 2018.
- (161) Grimme, S.; Bannwarth, C. *J. Chem. Phys.* **2016**, *145*, 054103.
- (162) Kessler, C.; Schuldt, R.; Emmerling, S.; Lotsch, B.; Kästner, J.; Gross, J.; Hansen, N. *Microporous Mesoporous Mat.* **2022**, 111796.
- (163) Kästner, J.; Carr, J. M.; Keal, T. W.; Thiel, W.; Wander, A.; Sherwood, P. *J. Phys. Chem. A* **2009**, *113*, 11856–11865.

- (164) Behler, J. *J. Chem. Phys.* **2011**, *134*, 074106.
- (165) Behler, J.; Parrinello, M. *Phys. Rev. Lett.* **2007**, *98*, 146401.
- (166) Zaverkin, V.; Kästner, J. *J. Chem. Theory Comput.* **2020**, *16*, 5410–5421.
- (167) Artrith, N.; Urban, A.; Ceder, G. *Phys. Rev. B* **2017**, *96*, 014112.
- (168) Artrith, N.; Urban, A. *Comput. Mater. Sci.* **2016**, *114*, 135–150.
- (169) Miksch, A. M.; Morawietz, T.; Kästner, J.; Urban, A.; Artrith, N. *Mach. Learn.: Sci. Technol.* **2021**, *2*, 031001.
- (170) Artrith, N.; Urban, A.; Ceder, G. *J. Chem. Phys.* **2018**, *148*, 241711.
- (171) Mason, J. C.; Handscomb, D. C. *Chebyshev polynomials*; CRC press, 2002.
- (172) Brunauer, S.; Emmett, P. H.; Teller, E. *J. Am. Chem. Soc.* **1938**, *60*, 309–319.
- (173) Balasubramani, S. G.; Chen, G. P.; Coriani, S.; Diedenhofen, M.; Frank, M. S.; Franzke, Y. J.; Furche, F.; Grotjahn, R.; Harding, M. E.; Hättig, C. *J. Chem. Phys.* **2020**, *152*, 184107.
- (174) Vericat, C.; Vela, M. E.; Corthey, G.; Pensa, E.; Cortés, E.; Fonticelli, M. H.; Ibanez, F.; Benitez, G.; Carro, P.; Salvarezza, R. C. *RSC Adv.* **2014**, *4*, 27730–27754.
- (175) Chaki, N. K.; Aslam, M.; Sharma, J.; Vijayamohanan, K. *J. Chem. Sci.* **2001**, *113*, 659–670.
- (176) Furche, F.; Ahlrichs, R.; Hättig, C.; Klopper, W.; Sierka, M.; Weigend, F. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2014**, *4*, 91–100.
- (177) Metz, S.; Kästner, J.; Sokol, A. A.; Keal, T. W.; Sherwood, P. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2014**, *4*, 101–110.
- (178) Vericat, C.; Vela, M.; Benitez, G.; Carro, P.; Salvarezza, R. *Chem. Soc. Rev.* **2010**, *39*, 1805–1834.
- (179) Häkkinen, H. *Nat. Chem.* **2012**, *4*, 443–455.

G Bibliography

- (180) Andreoni, W.; Curioni, A.; Grönbeck, H. *Int. J. Quantum Chem.* **2000**, *80*, 598–608.
- (181) Chen, Y.; Lu, P.; Zhang, C.; Huang, X.; Yao, F. *J. Power Sources* **2016**, *332*, 30–36.
- (182) Karhanek, D. Self-assembled monolayers studied by density-functional theory. Ph.D. thesis, Universität Wien, 2010.
- (183) Kohn, W.; Lau, K.-H. *Solid State Commun.* **1976**, *18*, 553–555.
- (184) Lang, N.; Kohn, W. *Phys. Rev. B* **1973**, *8*, 6010.
- (185) Zaremba, E.; Kohn, W. *Phys. Rev. B* **1976**, *13*, 2270.
- (186) Bengtsson, L. *Phys. Rev. B* **1999**, *59*, 12301.
- (187) Freysoldt, C.; Mishra, A.; Ashton, M.; Neugebauer, J. *Phys. Rev. B* **2020**, *102*, 045403.
- (188) Truyens, A.-J.; Vekeman, J.; Tielens, F. *Surf. Sci.* **2020**, *696*, 121597.
- (189) Ferrighi, L.; Pan, Y.-x.; Grošnbeck, H.; Hammer, B. *J. Phys. Chem. C* **2012**, *116*, 7374–7379.
- (190) Andersson, M. *J. Theor. Chem.* **2013**, *20*.
- (191) Carrasco, J.; Klimeš, J.; Michaelides, A. *J. Chem. Phys.* **2013**, *138*, 024708.
- (192) Peiretti, L. F.; Quaino, P.; Tielens, F. *J. Phys. Chem. C* **2016**, *120*, 25462–25472.
- (193) Nassoko, D.; Seydou, M.; Goldmann, C.; Chanéac, C.; Sanchez, C.; Portehault, D.; Tielens, F. *Mater. Today Chem.* **2017**, *5*, 34–42.
- (194) Duan, X.; Warschkow, O.; Soon, A.; Delley, B.; Stampfl, C. *Phys. Rev. B* **2010**, *81*, 075430.
- (195) Fallaque-Najar, J.; Morales-Gomero, J.-C.; Timon, V. *Surf. Sci.* **2019**, *689*, 121458.
- (196) Bedford, E.; Humblot, V.; Méthivier, C.; Pradier, C.-M.; Gu, F.; Tielens, F.; Boujday, S. *Chem. Eur. J.* **2015**, *21*, 14555–14561.

- (197) Tielens, F.; Costa, D.; Humblot, V.; Pradier, C.-M. *J. Phys. Chem. C* **2008**, *112*, 182–190.
- (198) Grönbeck, H.; Curioni, A.; Andreoni, W. *J. Am. Chem. Soc.* **2000**, *122*, 3839–3842.
- (199) Meena, S. K.; Goldmann, C.; Nassoko, D.; Seydou, M.; Marchandier, T.; Moldovan, S.; Ersen, O.; Ribot, F.; Chanéac, C.; Sanchez, C. *ACS Nano* **2017**, *11*, 7371–7381.
- (200) Lushtinetz, R.; Oliveira, A. F.; Frenzel, J.; Joswig, J.-O.; Seifert, G.; Duarte, H. A. *Surf. Sci.* **2008**, *602*, 1347–1359.
- (201) Kelly, T. F.; Miller, M. K. *Rev. Sci. Instrum.* **2007**, *78*, 031101.
- (202) Oberdorfer, C.; Eich, S. M.; Schmitz, G. *Ultramicroscopy* **2013**, *128*, 55–67.
- (203) El-Zoka, A.; Langelier, B.; Botton, G.; Newman, R. *Mater. Charact.* **2017**, *128*, 269–277.
- (204) Rigutti, L.; Bonef, B.; Speck, J.; Tang, F.; Oliver, R. *Scr. Mater.* **2018**, *148*, 75–81.
- (205) Stiller, K.; Thuvander, M.; Povstugar, I.; Choi, P.-P.; André, H. O. *MRS Bull.* **2016**, *41*, 35.
- (206) Reddy, S. M.; Saxey, D. W.; Rickard, W. D.; Fougereuse, D.; Montalvo, S. D.; Verberne, R.; Van Riessen, A. *Geostand. Geoanalytical Res.* **2020**, *44*, 5–50.
- (207) Gault, B.; Yang, W.; Ratinac, K. R.; Zheng, R.; Braet, F.; Ringer, S. P. *Langmuir* **2010**, *26*, 5291–5294.
- (208) Gault, B.; Moody, M. P.; Cairney, J. M.; Ringer, S. P. *Atom probe microscopy*; Springer Science & Business Media, 2012; Vol. 160.
- (209) Lohse, M. S.; Bein, T. *Adv. Funct. Mater.* **2018**, *28*, 1705553.
- (210) Kushwaha, R.; Kaleeswaran, D.; Haldar, S.; Chakraborty, D.; Mullangi, D.; Borah, A.; Vinod, C. P.; Murugavel, R.; Vaidhyathan, R. *ACS Appl. Nano Mater.* **2020**, *3*, 9088–9096.

G Bibliography

- (211) Kandambeth, S.; Dey, K.; Banerjee, R. *J. Am. Chem. Soc.* **2018**, *141*, 1807–1822.
- (212) Lan, Y.; Han, X.; Tong, M.; Huang, H.; Yang, Q.; Liu, D.; Zhao, X.; Zhong, C. *Nat. Commun.* **2018**, *9*, 5274.
- (213) Yan, T.; Lan, Y.; Tong, M.; Zhong, C. *ACS Sustain. Chem. Eng.* **2019**, *7*, 1220–1227.
- (214) Deeg, K. S.; Damasceno Borges, D.; Ongari, D.; Rampal, N.; Talirz, L.; Yakutovich, A. V.; Huck, J. M.; Smit, B. *ACS Appl. Mater. Interfaces* **2020**, *12*, 21559–21568.
- (215) Ongari, D.; Talirz, L.; Smit, B. *ACS Cent. Sci.* **2020**, *6*, 1890–1900.
- (216) Ongari, D.; Yakutovich, A. V.; Talirz, L.; Smit, B. *ACS Cent. Sci.* **2019**, *5*, 1663–1675.
- (217) Tuoc, V. N.; Thao, N. T. *J. Phys. Conf. Ser.* **2019**, *1274*, 012010.
- (218) Yang, C.; Yang, Z.-D.; Dong, H.; Sun, N.; Lu, Y.; Zhang, F.-M.; Zhang, G. *ACS Energy Lett.* **2019**, *4*, 2251–2258.
- (219) Tylisanakis, E.; Klontzas, E.; Froudakis, G. E. *Nanoscale* **2011**, *3*, 856–869.
- (220) Martinez, C. R.; Iverson, B. L. *Chem. Sci.* **2012**, *3*, 2191–2201.
- (221) Thakuria, R.; Nath, N. K.; Saha, B. K. *Cryst. Growth Des.* **2019**, *19*, 523–528.
- (222) Carter-Fenk, K.; Herbert, J. M. *Chem. Sci.* **2020**, *11*, 6758–6765.
- (223) Anaya-Plaza, E.; Joseph, J.; Bauroth, S.; Wagner, M.; Dolle, C.; Sekita, M.; Gröhn, F.; Spiecker, E.; Clark, T.; de la Escosura, A. *Angew. Chem. Int. Ed.* **2020**, *59*, 18786–18794.
- (224) Claessens, C. G.; Stoddart, J. F. *J. Phys. Org. Chem.* **1997**, *10*, 254–272.
- (225) Sinnokrot, M. O.; Sherrill, C. D. *J. Am. Chem. Soc.* **2004**, *126*, 7690–7697.
- (226) Arnstein, S. A.; Sherrill, C. D. *Phys. Chem Chem. Phys.* **2008**, *10*, 2646–2655.
- (227) Wheeler, S. E.; Houk, K. *J. Am. Chem. Soc.* **2009**, *131*, 3126–3127.

- (228) Raju, R. K.; Bloom, J. W.; An, Y.; Wheeler, S. E. *Chemphyschem* **2011**, *12*, 3116–3130.
- (229) Meng, D.; Yang, J. L.; Xiao, C.; Wang, R.; Xing, X.; Kocak, O.; Aydin, G.; Yavuz, I.; Nuryyeva, S.; Zhang, L. *Proc. Natl. Acad. Sci.* **2020**, *117*, 20397–20403.
- (230) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Petersson, G. A.; Nakatsuji, H. et al. Gaussian~16 Revision C.01. 2016; Gaussian Inc. Wallingford CT.
- (231) Dennington, R.; Keith, T. A.; Millam, J. M. GaussView Version 6. 2019; Semichem Inc. Shawnee Mission KS.
- (232) Fang, Y. Hydrogen bond, pi-pi stacking, and van der Waals interactions investigated with density functional theory. Ph.D. thesis, Tulane University School of Science and Engineering, 2013.
- (233) Ye, X.; Li, Z.-H.; Wang, W.; Fan, K.; Xu, W.; Hua, Z. *Chem. Phys. Lett.* **2004**, *397*, 56–61.
- (234) Silva, N. J.; Machado, F. B.; Lischka, H.; Aquino, A. J. *Phys. Chem Chem. Phys.* **2016**, *18*, 22300–22310.
- (235) Melikova, S. M.; Voronin, A. P.; Panek, J.; Frolov, N. E.; Shishkina, A. V.; Rykounov, A. A.; Tretyakov, P. Y.; Vener, M. V. *RSC Adv.* **2020**, *10*, 27899–27910.
- (236) Goerigk, L.; Kruse, H.; Grimme, S. *Chemphyschem* **2011**, *12*, 3421–3433.
- (237) Nakano, T. *pi-stacked polymers and molecules*; Springer, 2016.
- (238) Ehrlich, S.; Moellmann, J.; Grimme, S. *Acc. Chem. Res.* **2013**, *46*, 916–926.
- (239) Swart, M.; van der Wijst, T.; Fonseca Guerra, C.; Bickelhaupt, F. M. *J. Mol. Model.* **2007**, *13*, 1245–1257.
- (240) Lee, K.; Murray, É. D.; Kong, L.; Lundqvist, B. I.; Langreth, D. C. *Phys. Rev. B* **2010**, *82*, 081101.

G Bibliography

- (241) Ding, S.-Y.; Gao, J.; Wang, Q.; Zhang, Y.; Song, W.-G.; Su, C.-Y.; ; Wang, W. *J. Am. Chem. Soc.* **2011**, *133*, 19816–19822.
- (242) Meunier, M.; Robertson, S. *Mol. Simul.* **2021**, *47*, 537–539.
- (243) Manz, T. A.; Limas, N. G. *RSC Adv.* **2016**, *6*, 47771–47801.
- (244) Manz, T. A. *RSC Adv.* **2017**, *7*, 45552–45581.
- (245) Limas, N. G.; Manz, T. A. *RSC Adv.* **2016**, *6*, 45727–45747.
- (246) Limas, N. G.; Manz, T. A. *RSC Adv.* **2018**, *8*, 2678–2707.
- (247) Vyas, V. S.; Haase, F.; Stegbauer, L.; Savasci, G.; Podjaski, F.; Ochsenfeld, C.; Lotsch, B. V. *Nat. Commun.* **2015**, *6*, 8508.
- (248) Kandambeth, S.; Mallick, A.; Lukose, B.; Mane, M. V.; Heine, T.; Banerjee, R. *J. Am. Chem. Soc.* **2012**, *134*, 19524–19527.
- (249) Mayo, S. L.; Olafson, B. D.; Goddard, W. A. *J. Phys. Chem.* **1990**, *94*, 8897–8909.
- (250) Li, X.; Wang, Z.; Qiu, Y.; Pan, Q.; Hu, P. *J. Alloys Compd.* **2015**, *620*, 31–37.
- (251) Halder, A.; Ghosh, M.; Khayum M, A.; Bera, S.; Addicoat, M.; Sasmal, H. S.; Karak, S.; Kurungot, S.; Banerjee, R. *J. Am. Chem. Soc.* **2018**, *140*, 10941–10945.
- (252) Weigend, F.; Ahlrichs, R. *Phys. Chem Chem. Phys.* **2005**, *7*, 3297–3305.
- (253) Churchill, C. D.; Wetmore, S. D. *J. Phys. Chem. B* **2009**, *113*, 16046–16058.
- (254) Pitonak, M.; Neogrady, P.; Rezac, J.; Jurecka, P.; Urban, M.; Hobza, P. *J. Chem. Theory Comput.* **2008**, *4*, 1829–1834.
- (255) Rutledge, L. R.; Wetmore, S. D. *J. Chem. Theory Comput.* **2008**, *4*, 1768–1780.
- (256) Butchosa, C.; Simon, S.; Voityuk, A. A. *Org. Biomol. Chem.* **2010**, *8*, 1870–1875.
- (257) Xu, H.; Gao, J.; Jiang, D. *Nat. Chem.* **2015**, *7*, 905–912.

- (258) Mu, Z.; Zhu, Y.; Li, B.; Dong, A.; Wang, B.; Feng, X. *J. Am. Chem. Soc.* **2022**, *144*, 5145–5154.
- (259) Xiao, D.; Chang, M.-C.; Niu, Q. *Rev. Mod. Phys.* **2010**, *82*, 1959.
- (260) Eschrig, H.; Bergert, I. *Phys. Status Solidi B* **1978**, *90*, 621–628.
- (261) Bredow, T.; Evarestov, R.; Jug, K. *Phys. Status Solidi B* **2000**, *222*, 495–516.
- (262) Miró, J.; Deák, P.; Ewels, C.; Jones, R. *J. Phys. Condens. Matter* **1997**, *9*, 9555.
- (263) Janetzko, F.; Köster, A. M.; Salahub, D. R. *J. Chem. Phys.* **2008**, *128*, 024102.
- (264) Noga, J.; Baňacký, P.; Biskupič, S.; Boča, R.; Pelikán, P.; Svrček, M.; Zajac, A. *J. Comput. Chem.* **1999**, *20*, 253–261.
- (265) Deák, P.; Snyder, L. C. *Phys. Rev. B* **1987**, *36*, 9619.
- (266) Deak, P. *Phys. Status Solidi B* **2000**, *217*, 9–21.
- (267) Toukmaji, A. Y.; Board Jr, J. A. *Comput. Phys. Commun.* **1996**, *95*, 73–92.

Erklärung über die Eigenständigkeit der Dissertation

Ich versichere, dass ich die vorliegende Arbeit mit dem Titel *Computational Analysis of Periodic Systems for Covalent Organic Frameworks and Molecules in High Electric Fields* selbständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe; aus fremden Quellen entnommene Passagen und Gedanken sind als solche kenntlich gemacht.

Declaration of Authorship

I hereby certify that the dissertation entitled *Computational Analysis of Periodic Systems for Covalent Organic Frameworks and Molecules in High Electric Fields* is entirely my own work except where otherwise indicated. Passages and ideas from other sources have been clearly indicated.

Stuttgart, den 9. April 2022

Robin Schuldt