Conformation and Packing Structure of Conjugated Molecules under External Constraints: From Solvated Proteins to Reconstructed Surfaces

Dissertation

to obtain the academic degree

Doctor Rerum Naturalium (Dr. rer. nat.)

of the Naturwissenschaftliche Fakultät II Chemie, Physik und Mathematik Martin-Luther-Universität Halle-Wittenberg

by Laura Katharina Scarbath-Evers

born May 17, 1987 in Filderstadt

First Reviewer

Prof. Dr. Daniel Sebastiani Martin-Luther-Universität Halle-Wittenberg Institut für Chemie - Theoretische Chemie Von-Dankelmann-Platz 4 06120 Halle (Saale) Germany

Second Reviewer

Univ.-Prof. Dr. Arthur Ernst Johannes Kepler Universität Linz Institut für theoretische Physik Altenberger Strasse 69 4040 Linz Austria

Day of public defense: October 1, 2019

This dissertation has been created under the supervison of Prof. Daniel Sebastiani at the Institute of Chemistry of the Martin-Luther University Halle-Wittenberg.

Acknowledgment

First, I would like to to express my deepest gratitude to Prof. Dr. Daniel Sebastiani, my PhD advisor, for his continuous support throughout this work and for the numerous interesting discussions. I am also thankful to the current and former members of the Sebastiani group for their support and the friendly atmosphere they provided.

My gratitude also goes to my collaboration partners from the SFB102 project; in particular, I would like to thank Prof. Widdra for his support and many discussions that gave me invaluable insight into experimental scanning tunneling microscopy. I am indebted to Dr. René Hammer for providing me with the experimental STM images which he recorded and interpreted as part of his doctoral thesis. Moreover, I am grateful to Erik Schreck and Momchil Ivanov for the many fruitful discussions and helpful practical suggestions.

Further, I would like to acknowledge my collaboration partners from the photosensor project, in particular Prof. Jörg Matysik, Dr. Chen Song, and Susanne Altmayer for their interesting experiments and discussions. I would also like to thank Sascha Jähnigen, who collaborated with me on the computational part of the project for his support. Moreover, I am grateful to Paul Ahlert, who worked with me on the mutagenesis study and carried out the MD simulations of the mutant.

Many thanks go to Prof. Rinke, for the warm welcome in his group as well as the continuous support during my research stay at Aalto Finland. I am particularly indebted to Dr. Milica Todorovic for the guidance and the many fruitful discussions that helped me a lot with my research project.

I am extremely grateful to Dr. Dorothea Golze for her support, help, and personal advise.

Last but not least I would like to thank Christian Dressler, Dr. Felix Hoffmann, and Dr. Volker Deringer for proof-reading my thesis.

Publications

Below are the publications I have co-authored during this thesis together with a description of the individual contributions of all authors.

 Laura K. Scarbath-Evers, Sascha Jähnigen, Hossam Elgabarty, Chen Song, Rei Narikawa, Jörg Matysik, and Daniel Sebastiani, "Structural heterogeneity in parent ground-state structure in AnPixJg2 revealed by theory and spectroscopy", Phys. Chem. Chem. Phys., 2017, 19, 13882-13894.

I prepared the computational setup, conducted and analyzed the simulations, and wrote the manuscript.

Sascha Jähnigen contributed to the design of the study, the project development and by discussions.

Hossam Elgabarty helped with the computational set-up and took part in discussions.

Chen Song carried out the NMR measurements and analysis of the experimental NMR data and wrote the experimental results part of the manuscript.

Rei Narikawa prepared the protein samples.

Jörg Matysik advised the experimental part and contributed in discussions.

Daniel Sebastiani advised the computational part of the project, contributed in discussions and edited the manuscript.

 Laura K. Scarbath-Evers, Milica Todorović, Dorothea Golze, René Hammer, Wolf Widdra, Daniel Sebastiani and Patrick Rinke, "Gold diggers: Altered reconstruction of the gold surface by physisorbed aromatic oligomers", Phys. Rev. Materials 2019, 3

In this project I contributed in designing the computational study, performed all calculations, analyzed the computational results, and wrote the manuscript.

Milica Todorović advised and helped with the computational setup, the interpretation of the results and wrote parts of the manuscript.

Dorothea Golze participated by software implementations that accelerated the calculations and contributed to discussions.

René Hammer carried out the experimental part of the project and analyzed the experimental data.

Wolf Widdra designed and advised the experimental part of the project, contributed through discussions and wrote the experimental part of the manuscript. D. Sebastiani contributed through discussions regarding the technical realization in the initial phase of the project.

Patrick Rinke advised the computational part of the study, wrote parts of the manuscript and participated in the interpretation of the data.

3. L. K. Scarbath-Evers, R. Hammer *, D. Golze, Martin Brehm, Daniel Sebastiani, Wolf Widdra *"From Flat to tilted: Oligothiophene Growth beyond the Monolayer"*, At the stage of submission.

In this project, I prepared and conducted the computational part, analyzed and interpreted the computational data in particular in the context of the experimental results, and wrote the manuscript.

René Hammer conducted all experimental measurements in the project, analyzed and interpreted the experimental results † .

Dorothea Golze helped with the computational setup, advised the computational part of the project, contributed through many discussions, and wrote parts of the manuscript.

Martin Brehm helped at the initial stage of the computational project.

Daniel Sebastiani advised the computational part of the project and contributed through discussions.

Wolf Widdra initiated the project and defined the scientific direction of the project. Moreover, he advised the experimental part of the project, took part in the interpretation of the experimental data and contributed through discussions regarding the combination of theoretical and experimental results.

^{*}Both authors contributed equally.

[†]The experimental studies have conducted as part of the PhD thesis of R. Hammer¹

Table of Contents

1. Introduction

2.	The	oretica	I Description of Organic, Semiconducting Molecules	17
	2.1.	Born-0	Oppenheimer Approximation and Wave Function Based Methods	17
		2.1.1.	The Hartree Fock Ansatz	18
		2.1.2.	Electron Correlation	21
	2.2.	Densit	y Functional Theory (DFT)	22
		2.2.1.	Formal Derivation of DFT: The Hohenberg-Kohn Theorems	22
		2.2.2.	The Kohn-Sham Approach	26
		2.2.3.	Exchange Correlation Functionals	28
		2.2.4.	Physical Meaning of the Eigenvalues in Kohn-Sham DFT	29
		2.2.5.	Band Gaps of Solids within DFT	30
		2.2.6.	Non-covalent Interactions in Kohn-Sham DFT	30
		2.2.7.	DFT under Periodic Boundary Conditions	34
		2.2.8.	Density Functional Theory for Metallic Systems	37
		2.2.9.	Basis Sets	39
	2.3.	Hybrid	d Quantum Mechanics/Molecular Mechanics (QM/MM) Approa-	
		ches		43
	2.4.	Time	Evolution of Large Systems: Classical Molecular Dynamics	44
		2.4.1.	Newton's and Hamilton's Equations of Motion	46
		2.4.2.	Molecular Dynamics under Periodic Boundary Conditions: Trea-	
			ting Non-bonded Interactions	47
		2.4.3.	Modelling Systems under Biological Conditions: Temperature	•
			and Pressure Control	50
3.	Stru	ctural	Heterogeneity in AnPixJg2 Wild Type and Mutant	55
	3.1.	Introd	uction	55
	3.2.	Struct	ural Heterogeneity in the AnPixJg2 Wild Type	61
		3.2.1.	Structure of the Second GAF Domain of AnPixJ	61
		3.2.2.	Interaction between the Chromophore and its Binding Pocket $\ .$	68
		3.2.3.	Discussion of the Two-state Model for the Conformation of the	
			Chromophore	73
	3.3.	Part 2	: One Mutant Three Sub-states	77
		3.3.1.	A Third Sub-State of PCB in the TYR302PHE Mutant	77
		3.3.2.	Sub-State 3 exhibits an Altered Hydrogen Bond Pattern $\ . \ . \ .$	80
		3.3.3.	Mutation of TYR302 Opens a New Gate for Solvent	85
		3.3.4.	Discussion	89
		3.3.5.	Conclusion and Outlook	93

4.	Olig	othiophenes on Nano-structured Surfaces	95
	4.1.	Introduction	95
	4.2.	Gold Diggers	101
		4.2.1. The Clean Au(100) Surface $\ldots \ldots \ldots$	101
		4.2.2. Adsorption on an Ill-preferred Site: Molecule Alters Surface	
		Reconstruction \ldots	102
		4.2.3. The Costs to Displace a Ridge	105
		4.2.4. Discussion	106
		4.2.5. Conclusion	108
	4.3.	From Flat to tilted: Oligothiophene Growth beyond the Monolayer	109
		4.3.1. Characterization of the Chiral Domains	109
		4.3.2. Lateral Position of the Second Layer Molecules	111
		4.3.3. Molecular Orientation in the Second Layer	112
		4.3.4. Transition to the Bulk Structure	116
		4.3.5. Conclusion	118
5.	Sum	imary	119
Δ	Δnn	endix	123
А.	Δ 1	Appendix to Project 1	123
	11.1.		120
		A 1.2 Data Analysis of the MD Trajectory	120
		A 1.3 Supplementary Material to: One Mutant Three Sub States	124
	1 2	A.1.5. Supplementary Material to. One Mutant Three Sub-States	120
	A.2.	A 2.1 Calculation of STM Imagor	127
		A.2.2. Supplementary Information to: "Cold Diggarge Altered Basen	121
		A.2.2. Supplementary information to. Gold Diggers. Aftered Recon-	,190
		A 2.2 Supplementary Information to: "Erorn Elet to Tilted, Oligethia	129
		phene Growth beyond the Monolayer"	133
B	Refe	erences	139
			100
С.	Aca	demic Curriculum Vitae	167

1. Introduction

Functional, conjugated molecules play ubiquitous roles in in nature and industry. For instance, employed as chromophores in photosensory proteins they act as photoswitches and control the response to light in living organisms.^{2–5} For this purpose, the chromophore reversely switches between two conformal states that absorb light of different wavelengths: one biologically inactive "dark" stable state and one biologically active photoproduct state. In particular the familiy of cyanobacteriochromes (CBCRs), photosensory proteins that control light-dependent processes in cyanobacteria,^{6–8} have drawn attention during the last years. Despite the simple architecture of their photosensory module^{6,8} they exhibit an outstanding variety of photocycles that cover the whole range of the visible spectrum^{9–13} which makes them auspicious candidates for applications in bio-imaging, or optogenetics, i.e. the capacity to control cellular activities by genetically encoded light sensitive proteins.¹⁴ In material science, on the other hand, organic semiconducting molecules have emerged as promising materials for opto-electronic devices^{15–19} due to their structural variety^{20–22} and the possibility of fine-tuning their electronic properties by rational synthetic design.²³

To advance the application of organic, conjugated molecules as molecular switches or in opto-electronic devices, computational chemistry can make important contributions by providing atomistic models and by establishing structure-property relations. However, the structure and therefore the properties of chromophores in proteins or organic semiconductors are strongly steered by the the immediate chemical environment as well as the surrounding conditions, such as temperature and pressure. Those have to be explicitly included in the modeling approach to obtain realistic results. Depending on the chemical environment (e.g. a protein, a surface etc.), different theoretical approaches are required. The method of choice to study the time evolution of a system under ambient conditions is molecular dynamics (MD). Within this framework, the system is propagated in time by numerically solving Newtons equations of motion. The advantage of MD is that it yields dynamical information, such as the mobility of atomic subgroups of the system or hydrogen bond interactions, as well as equilibrium properties of the system in a well-defined thermodynamic ensemble,²⁴ for instance under constant pressure and constant temperature. In case of proteins, system sizes of millions of atoms²⁵ and slow structural rearrangement processes which are at least on time scales of µs-ms,^{26,27} require calculations at the force field level of theory. Within this framework, atoms and bonds are treated classically as balls and sticks that are subject to a previously parametrized empirical potential. The neglect of any quantum effects considerably speeds up the computation and enables the description of systems consisting of millions of atoms^{28,29} on the ms time scale.³⁰ Naturally, such classical approaches inherently fail to describe events that are driven by electronic effects, for instance, charge transfer between a molecule and a surface. To model such systems,

1. Introduction

density functional theory (DFT), which explicitly describes electrons by solving the electronic Schrödinger equation, is commonly the method of choice. It relies on the systems electronic density instead of the many-body wave function which significantly reduces the complexity. Together with advances in electronic structure algorithms^{31,32} and hardware improvements, it allows nowadays the treatment of systems consisting of several hundreds of atoms at affordable computational efforts.

This thesis explores structure and properties of organic, functional molecules under the influence of two two fundamentally different chemical environments: embedded in a protein under biological conditions and in contact with nano-structured metal surfaces. Using different theoretical approaches ranging from force field molecular dynamics on the one hand to density functional theory calculations on the other hand, the following scientific questions will be addressed:

- 1. Structure and dynamics of a chromophore in a photosensory protein (**Project** I)
- 2. Single molecule adsorption and layer assembly of an organic semiconductor on the reconstructed Au(100) surface (**Project II**)

In **Project I** of this thesis, the dark stable state of the second GAF domain of the CBCR AnPixJ (AnPixJg2), which is depicted in Fig. 1.1 **a** is investigated by force field MD. AnPixJg2 is a representative of the red/green CBCRs, which exhibit a red light absorbing "dark" stable state and a green light absorbing photoproduct state. Like all CBCRs, it employs a bilin-derived chromophore that is attached to the protein via a thioether bridge involving a cystein residue.^{8,33} In the particular case of AnPixJg2, the chromophore is phycocyanobilin (PCB) which is depicted in Fig. 1.1 **b**. Formation of the photoproduct state involves reversible $Z \rightarrow E$ isomerization of the methine bridge between *C*-ring and *D*-ring of the chromophore and rotation of the *D*-ring.^{6,33,34}

A peculiar feature that is common among many photosensory proteins is structural heterogeneity, i.e. the existence of two or more structural sub-populations in either the "dark" stable state or the photoproduct state or both.^{11,35–45,45–48} The frequent occurrence of structural heterogeneity suggests that the latter is a unifying property of photosensory proteins. Moreover, the existence of conformational sub-populations could be a possible explanation for the strong variance of the photoisomerization quantum yield among subclasses of CBCRs^{35–37,49} and functionally related photosensory proteins^{50,51} despite structurally similar chromophores. These considerations strongly motivate the investigation of the parent ground state of AnPixJg2 by force field MD that is carried out in **Project I**. While crystallographic experiments often capture only one conformation of the system, molecular dynamics simulations under ambient conditions and of sufficient simulation time allow to characterize different structural sub-states.

The first part of **Project I** (Chapter 3.2) focuses on the investigation of structural heterogeneity in the AnPixJg2 dark stable state. By analysis of a $\approx 1 \,\mu$ s molecular dynamics simulation, two sub-states of the chromophore are identified that are stable on time scales of $10 \,\text{ns} - 500 \,\text{ns}$. The interconversion between these two sub-states is

accompanied by changes in the hydrogen bond pattern between the chromophore and the protein.

Apart from their structurally similar chromophore, all sub-families of the CBCR exhibit highly conserved key residues in the proximity of the chromophore which points towards specific function of these amino acids being essential for the color tuning mechanism. A versatile theoretical tool to probe the function of specific amino acid residues is *in-silico* mutagenesis, in which the respective amino acid is replaced by another amino acid with different physico-chemical properties and the mutated system is investigated using molecular dynamics. In combination with experimental mutagenesis studies, a structural model can be established and verified by calculation and comparison of experimental observables (e.g. NMR shifts).⁵² As stand-alone tool, *in-silico* mutagenesis can efficiently screen various mutants and guide experimental efforts by reducing the pool of possible mutants down to a few, potentially interesting ones.

The second part of **Project I** (Chapter 3.3) explores the role of a conserved tyrosine residue, TYR302, by such an *in-silico* mutagenesis procedure. In the AnPixJg2 dark stable state the hydroxyl group of the side chain of TYR302 is involved in frequent hydrogen bonds with the *C*-ring propionate group of PCB as illustrated in Fig. 1.1 c. Notably, in all CBCR sub-families, this tyrosine residue interacts directly with either the *B*-ring or *C*-ring propionate group of the chromophore³³ which makes a detailed investigation of the role of TYR302 promising. To study the role of the hydrogen bond, TYR302 is replaced by phenylalanine, an amino acid whose non-polar side chain is of similar size but cannot be involved in any hydrogen bonding (see Fig. 1.1 d). Subsequently, the conformation and flexibility of the chromophore as well as its interactions with protein and solvent are investigated by a ≈ 1 µs molecular dynamics simulation.

Project II of this thesis addresses the structure of an oligothiophene, α -sexithiophene (α -6T), on the reconstructed Au(100) surface from single molecule adsorption to bilayer assembly. α -6T is a prototypical organic semiconductor with good hole-conducting properties^{53–58} which serves as model system for polythiophenes and has also already been employed successfully in opto-electronic devices such as thin-film transistors.^{58–61} As shown in Fig. 1.2 **a**, it consists of six thiophene units that are connected via the C_{α} atom of each unit. Due to the rod-like shape of the molecule, its electron density is anisotropically distributed. As a result, the magnitude of charge transport through the molecular layer as well as charge injection into the electrode surface are strongly influenced by the orientation of the α -6T molecules in the organic layers and with respect to the electrode.^{62–66}

The applicability of metallic gold in a wide range of fields, from catalysing oxidation reactions^{67–71} to electrodes in organic field emitting transistors, has steered research towards the investigation of organic-gold interfaces.^{72–81} From a scientific point of view, the low index gold surfaces show complex and large scale surface reconstruction^{82–88} which makes the characterization of the low index gold surfaces intriguing and challenging at the same time. For instance, reconstruction of the Au(100) surface results in a c(28 × 48) supercell that features a contracted, quasi-hexagonal top layer and reconstruction rows running along the [110] direction (Fig. 1.2 b).⁸⁴ That height

1. Introduction



Figure 1.1. – a: Structure of AnPixJg2. b: The chromophore phycocyanobilin (PCB) which is employed in AnPixJg2. c: Interaction between the C-ring propionate group of PCB and a conserved tyrosine residue, TYR302. d: The amino acids tyrosine and phenylalanine. Both are similar in size, however, phenylalanine possesses a non-polar side chain without hydroxyl group that cannot participate in hydrogen bond interactions.

corrugation of the top layer offers various adsorption sites for organic molecules. These adsorption sites can be broadly classified into on-ridge and in-valley of the reconstruction rows and differ in the adsorption strength of the molecule. The preference for a specific adsorption environment makes it possible to employ the surface corrugation as template and enforce an adsorption pattern that is commensurate with the reconstruction pattern of the underlying surface.^{76,77} A successful application of this *patterning concept* requires a profound understanding of the molecule-metal interactions at different adsorption sites. Moreover, it assumes that the reconstruction pattern remains unaffected by molecule-surface interactions.

While density functional theory could aid understanding adsorption processes on such corrugated surfaces in principle, the sheer unit cell sizes of reconstructed surfaces have so far limited theoretical studies to small molecules on idealized substrates. In the first part of **Project II** (Section 4.2), the single molecule adsorption of α -6T on two relevant adsorption sites, referred to as *in-valley* and *on-ridge* of the reconstructed Au(100) surface, is studied by DFT calculations. In order to explicitly account for the surface reconstruction, a 20 × 5 supercell model of the Au(100) surface⁸⁹ is employed.

Its system size of 546 atoms makes a computational treatment still feasible while retaining the physical properties of the reconstructed Au(100) surface such as two-dimensional lateral contraction, rotation, and differently buckled qualitative surface areas.^{82,90}

In the second part of **Project II** (Chapter 4.3), the structural assembly of an α -6T bilayer on Au(100) is investigated by combining hybrid quantum mechanics/molecular mechanics (QM/MM) calculations and experimental scanning tunneling microscopy (STM) measurements conducted by our collaborators.¹ While monolayer formation of α -6T on various metal surfaces is well understood, $^{80,81,91-95}$ the larger molecule-surface distance and the presence of the isolating monolayer makes STM studies of the second molecular layer highly challenging. Therefore, a detailed knowledge of the morphology of the second layer of α -6T on gold is lacking so far. On the other hand, the presence of the first layer also leads to a decoupling of the electronic states of the metal and the molecules in the second layer. As a result, the individual electronic properties of the molecules remain preserved and are visible in the STM image.⁹⁶ STM images derived from electronic structure calculations can aid in assigning those distinct features in the experimental STM image to the relative orientation of the individual molecules in the second layer. The second part of **Project II** presented in this thesis covers an extensive structure search and subsequent calculation of STM images of the relaxed α -6T bilayer structures on gold. By comparing the features in the experimental and theoretical STM image, a growth model for α -6T on gold could be established.



Figure 1.2. – **a**: The investigated molecule, α -sexithiophene (α -6T) **b**: Computed STM image of the 20×5 Au(100) model surface.⁸⁹ The reconstructed surface exhibits a complex height modulation including reconstruction ridges (white) and valleys (dark) that offer different adsorption sites of varying adsorption strength.

The thesis has the following structure: Chapter 2 covers the theoretical methodology used. In Chapter 3 and 4 the results from Project I and Project II, respectively, are presented. Chapter 5 contains concluding remarks and an outlook. Details regarding the employed methods and the data analysis can be found in the Appendix.

2. Theoretical Description of Organic, Semiconducting Molecules

2.1. Born-Oppenheimer Approximation and Wave Function Based Methods

Within this chapter important aspects of theoretical methods applied to organic, semiconducting molecules will be reviewed. In general, the choice of the method is determined by the systems size and required accuracy. In quantum theory, the energy of a system is given by the expectation value of the Hamiltonian, i.e. $E = \langle \hat{H} \rangle = \langle \Psi | \hat{H} | \Psi \rangle$ and an exact description of a system could therefore be achieved by solving the Schrödinger Equation. For a system of N electrons and M nuclei the stationary, non-relativistic Schrödinger equation reads⁹⁷

$$\hat{H}\Psi_{i}\left(\mathbf{x}_{1}, \mathbf{x}_{2}, ..., \mathbf{x}_{N}, \mathbf{R}_{1}, \mathbf{R}_{2}, ..., \mathbf{R}_{M}\right) = E_{i}\Psi_{i}\left(\mathbf{x}_{1}, \mathbf{x}_{2}, ..., \mathbf{x}_{N}, \mathbf{R}_{1}, \mathbf{R}_{2}, ..., \mathbf{R}_{M}\right).$$
(2.1)

With Ψ being the (normalized) many-body wave function of the system and \hat{H} the Hamiltonian which in atomic units is defined as:⁹⁷

$$\hat{H} = \underbrace{-\frac{1}{2} \sum_{i=1}^{N} \nabla_{i}^{2} - \sum_{A=1}^{M} \frac{1}{m_{A}} \nabla_{A}^{2}}_{1} - \underbrace{\sum_{i=1}^{N} \sum_{A=1}^{M} \frac{Z_{A}}{\|\mathbf{r}_{i} - \mathbf{R}_{A}\|}}_{3}}_{+ \underbrace{\sum_{i=1}^{N} \sum_{j>i}^{N} \frac{1}{\|\mathbf{r}_{i} - \mathbf{r}_{j}\|}}_{4} + \underbrace{\sum_{A=1}^{M} \sum_{B>A}^{M} \frac{Z_{A}Z_{B}}{\|\mathbf{R}_{A} - \mathbf{R}_{B}\|}}_{5}}_{5}$$
(2.2)

with M_A being the mass of nucleus, and Z_A and Z_B being the atomic number of nucleus A and B. Term 1 and 2 denote the kinetic energy operator of electron and nuclei respectively, term 3 describes the Coloumb attraction between nuclei and electrons while term 4 and 5 treat the repulsive interactions between the electrons and the nuclei, respectively.

Unfortunately, for systems beyond the hydrogen atom that equation is analytically not tractable due to the many-body interactions. An elegant solution was offered by Born and Oppenheimer⁹⁸ that exploits the different time scales of the movement of electrons and nuclei. While the fastest nuclear motions, vibrations, happen at time scales of $\approx 10^{-12}$ s - 10^{-13} s, electronic excitations are as fast as 10^{-15} s due to the at

2. Theoretical Description of Organic, Semiconducting Molecules

least three orders of magnitude smaller mass of the electrons^{*}. This mass disparity allows to neglect the motion of the nuclei and solve the electronic Schrödinger equation for electrons moving in the field of M stationary point charges, reducing Eq. (2.2) to:⁹⁷

$$\hat{H} = -\frac{1}{2} \sum_{i=1}^{N} \nabla_{i}^{2} - \sum_{i=1}^{N} \sum_{A=1}^{M} \frac{Z_{A}}{\|\mathbf{r}_{i} - \mathbf{R}_{A}\|} + \sum_{i=1}^{N} \sum_{j>i}^{N} \frac{1}{\|\mathbf{r}_{i} - \mathbf{r}_{j}\|} = \hat{T}_{el} + \hat{V}_{e-e} + \hat{V}_{e-N} \quad (2.3)$$

 \hat{T}_{el} denotes the kinetic energy operator of the electrons, \hat{V}_{e-N} is the interaction between electrons and nuclei, and \hat{V}_{e-e} is the electron-electron interaction. Expression (2.3) is also denoted as electronic Hamiltonian.⁹⁷ By decoupling nuclear and electronic motion, the Born-Oppenheimer approximation thus leads to an electronic Schrödinger equation that depends only parametrically on the nuclear coordinates. However, for systems with more than two electrons that electronic Schrödinger equation does still not possess an analytical solution. To approximate the many-electron problem, different approaches have emerged that can be broadly classified into wave function based methods and density based methods.

2.1.1. The Hartree Fock Ansatz

Wave function based methods employ the wave function of the N-electron system to solve the electronic Schrödinger equation. The most simple model for a many-electron system is a system of N non-interacting particles; in that case the Hamiltonian of the system can be written as a sum of the individual one-particle Hamiltonians:¹⁰⁰

$$\hat{H}_{tot} = \sum_{i=1}^{N} \hat{h}_i \tag{2.4}$$

and the corresponding N-particle wave function factorizes to a product of N oneparticle wave functions, commonly denoted as $Hartree \ product$:¹⁰⁰

$$\Psi(\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_n) = \chi_i(\mathbf{x}_1)\chi_j(\mathbf{x}_2)\chi_k(\mathbf{x}_N)$$
(2.5)

 $\chi(\mathbf{x})$ is denoted as spin orbital of the electron and describes the spatial distribution and the spin of the electron:¹⁰⁰

$$\chi(\mathbf{x}) = \begin{cases} \psi(\mathbf{r}) \, \alpha(\omega) \\ \text{or} \\ \psi(\mathbf{r}) \, \beta(\omega) \end{cases}$$
(2.6)

 $\psi(\mathbf{r})$ denotes the spatial wavefunction, while $\alpha(\omega)$ and $\beta(\omega)$ denote the - orthonormalspin wave functions. In the picture of the Hartree product *ansatz*, correlation effects are *a priori* absent, i.e. the position of one particle is completely independent of the position of the other particles. Moreover, it does not treat the electrons as indistinguishable and violates the anti-symmetry principle.¹⁰⁰ The anti-symmetry principle requires

^{*}The Proton-to-electron mass ratio is 1836.15267389⁹⁹

2.1. Born-Oppenheimer Approximation and Wave Function Based Methods

that exchange of spatial and spin coordinates of two arbitrary electrons changes the sign but not the absolute value, of the wave function. Both, the indistinguishability of the electrons and the anti-symmetry principle can be elegantly incorporated by writing the N-body wave function as a $N \times N$ determinant, usually denoted as Slater determinant and illustrated below:¹⁰⁰

$$\Psi(\mathbf{x}_1, \mathbf{x}_2) = (N!)^{-\frac{1}{2}} \det \begin{vmatrix} \chi_i(\mathbf{x}_1) & \chi_j(\mathbf{x}_1) \cdots & \chi_k(\mathbf{x}_1) \\ \chi_i(\mathbf{x}_2) & \chi_j(\mathbf{x}_2) \cdots & \chi_k(\mathbf{x}_2) \\ \vdots & \vdots & \vdots \\ \chi_i(\mathbf{x}_N) & \chi_j(\mathbf{x}_N) \cdots & \chi_k(\mathbf{x}_N) \end{vmatrix}$$
(2.7)

Exchanging the coordinates of the arbitrary electrons corresponds to interchanging two rows which changes the sign of the determinant, and hence the wave function, which fulfills the anti-symmetry principle. Moreover, two electrons with the same spin occupying the same orbital would manifest in two equal (or, more general: linear dependent) rows and hence a zero determinant.

Using a Slater determinant as N-body wave function explicitly accounts for the exchange but inherently excludes correlation. In other words, the motion of electrons with parallel spin is correlated, since the probability that two electrons with parallel spin occupy the same orbital is zero. That area of zero probability for finding another electron with parallel spin around an electron is commonly denoted as Fermi hole.¹⁰¹ While the motion of two electrons with parallel spin is correlated, the motion of two electrons with parallel spin is correlated, the motion of two electrons with parallel spin is correlated within that formalism.

The Hartree Fock approach tackles the complex problem of an electron moving in the field of N-1 interacting electrons by approximating it with the much easier problem of a single particle moving in a homogeneous potential created by the average interaction of the N-1 particles. Starting from a single Slater Determinant composed of N spin orbitals, the ground state wave function Ψ_0 is obtained by minimizing the ground state energy with respect to the spin orbitals which leads to the Hartree Fock equations:⁹⁷

$$\hat{f}\chi_i(\mathbf{x}_i) = \epsilon\chi(\mathbf{x}_i), \quad i = 1, ..., N$$
(2.8)

with the Fock operator, \hat{f}_i , consisting of the one-particle kinetic energy operator, an effective potential $v^{HF}(i)$ acting on the particle and the potential arising from the nuclei:⁹⁷

$$\hat{f}_i = -\frac{1}{2}\nabla_i^2 + v_i^{HF} - \sum_{j=1}^M \frac{Z_j}{\|\mathbf{r}_i - \mathbf{R}_j\|} = \hat{h}_i + v_i^{HF}$$
(2.9)

The effective potential $v^{HF}(i)$ is commonly denoted Hartree-Fock potential and can

be further separated into the Coulomb and Exchange term:⁹⁷

$$v^{HF}(\mathbf{x}_{1}) = \sum_{j=1}^{N} \left(\hat{J}_{j}(\mathbf{x}_{1}) - \hat{K}_{j}(\mathbf{x}_{1}) \right)$$
(2.10)

The Coulomb operator, \hat{J} , accounts for the classical electrostatic interaction between the electron at position \mathbf{x}_1 and the average charge distribution of an electron in spin orbital χ_j .⁹⁷

$$\hat{J}_j(\mathbf{x}_1) = \int \|\chi_j(\mathbf{x}_2)\|^2 \frac{1}{r_{12}} d\mathbf{x}_2$$
(2.11)

Evaluation of Eq. (2.11) requires only knowledge about χ_i at position \mathbf{r}_i which makes the Coulomb operator a local operator. Unlike the Coulomb operator that can be assigned to the electrostatic potential of a classical charge distribution at position \mathbf{r}_2 , the exchange operator \hat{K} has no such classical analogue and can be only understood in terms of its effect on a spin orbital χ_i .⁹⁷

$$\hat{K}_{j}(\mathbf{x}_{1})\chi_{i}(\mathbf{x}_{1}) = \int \chi_{j}^{*}(\mathbf{x}_{2})\frac{1}{r_{12}}\chi_{i}(\mathbf{x}_{2})d\mathbf{x}_{2}\chi_{j}(\mathbf{x}_{1})$$
(2.12)

Operation of $\hat{K}_j(\mathbf{x}_1)$ on χ_i basically exchanges the coordinates of the spin orbitals χ_i and χ_j . $\chi_i(\mathbf{x}_1)$ contains the spatial and the spin information of the electron; evaluation of Eq. (2.12) therefore requires integration over the spatial and the spin coordinates. The spin function is independent of the spatial coordinates, \mathbf{r}_2 , and the integral over the spin functions can thus be evaluated separately. Due to the orthogonality of the spin functions, i.e. $\langle \alpha(\omega), \beta(\omega) \rangle = 0$, Eq. (2.12) is zero for two electrons with opposite spin.⁹⁷ As a result, exchange interaction occurs only between electrons of parallel spin.¹⁰² Moreover, the exchange operator is non-local, since evaluation of $\hat{K}_j(\mathbf{x}_1) \chi_i(\mathbf{x}_1)$ requires integration of $\chi_i(\mathbf{x}_2)$ over the whole space.⁹⁷

It is important to emphasize that the sum in Eq. (2.10) also includes j = i and can therefore be rewritten as:⁹⁷

$$\langle \chi_{i}(\mathbf{x}_{1}) | v^{HF}(\mathbf{x}_{1}) | \chi_{i}(\mathbf{x}_{1}) \rangle = \langle \chi_{i}(\mathbf{x}_{1}) | \left(\hat{J}_{i}(\mathbf{x}_{1}) - \hat{K}_{i}(\mathbf{x}_{1}) \right) | \chi_{i}(\mathbf{x}_{1}) \rangle$$

$$+ \langle \chi_{i}(\mathbf{x}_{1}) | \sum_{j \neq i}^{N} \left(\hat{J}_{j}(\mathbf{x}_{1}) - \hat{K}_{j}(\mathbf{x}_{1}) \right) | \chi_{i}(\mathbf{x}_{1}) \rangle$$

$$= \underbrace{\iint \| \chi_{i}(\mathbf{x}_{2}) \|^{2} \frac{1}{r_{12}} \| \chi_{i}(\mathbf{x}_{1}) \|^{2} d\mathbf{x}_{1} d\mathbf{x}_{2}}_{1} - \underbrace{\iint \| \chi_{i}(\mathbf{x}_{2}) \|^{2} \frac{1}{r_{12}} \| \chi_{i}(\mathbf{x}_{1}) \|^{2} d\mathbf{x}_{1} d\mathbf{x}_{2}}_{2} \qquad (2.13)$$

$$+ \langle \chi_{i}(\mathbf{x}_{1}) | \sum_{j \neq i}^{N} \left(\hat{J}_{j}(\mathbf{x}_{1}) - \hat{K}_{j}(\mathbf{x}_{1}) \right) | \chi_{i}(\mathbf{x}_{1}) \rangle$$

Term one in Eq. (2.13) represents the Coulomb integral for i = j and expresses the, physically meaningless, self-interaction, namely the repulsion of an electron by itself. Fortunately, for i = j term 2, the exchange term, is equal to the Coulomb term and therefore they cancel each other. As a fundamental consequence, the self-interaction problem is inherently circumvented in Hartree Fock theory. In approximate density functional theory, however, this is not the case which affects detrimentally the computed properties, like band gaps (see Section 2.2.4).

Since the effective potential, v^{HF} , depends on the set of spin orbitals that are yet to determined, Eq. (2.8) must be solved self-consistently. Starting from an initial guess the orbitals are re-evaluated iteratively until the energy change is below a certain threshold criterion. The resulting eigenfunctions are denoted as Hartree-Fock orbitals and the corresponding eigenvalues ϵ_i are the orbital energies.⁹⁷ It is important to illustrate that the total minimum energy of the system is not equal to the sum of the orbital energies.¹⁰³ By simply adding up the orbital energies the coulomb and exchange interactions are counted twice as apparent from the following expression:

$$\sum_{i=1}^{N} \epsilon_{i} = \sum_{i=1}^{N} \langle \chi_{i}(\mathbf{x}_{i}) | \hat{h}_{i} | \chi_{i}(\mathbf{x}_{i}) \rangle + \sum_{i=1}^{N} \sum_{j=1}^{N} \langle \chi_{i}(\mathbf{x}_{i}) | \hat{J}_{j}(\mathbf{x}_{i}) - \hat{K}_{j}(\mathbf{x}_{i}) | \chi_{i}(\mathbf{x}_{i}) \rangle$$
(2.14)

This double-counting of the electron-electron interactions can be prevented by simply multiplying the Coulomb and exchange term by 1/2 which results in the correct expression for the total energy E_0 :¹⁰³

$$E_{0} = \sum_{i=1}^{N} \langle \chi_{i}(\mathbf{x}_{1}) | \hat{h} | \chi_{i}(\mathbf{x}_{i}) \rangle + \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} \langle \chi_{i}(\mathbf{x}_{i}) | \hat{J}_{j}(\mathbf{x}_{i}) - \hat{K}_{j}(\mathbf{x}_{i}) | \chi_{i}(\mathbf{x}_{i}) \rangle$$
(2.15)

Nevertheless the Hartree Fock orbital energies have a sound physical justification. According to Koopman's theorem they can be linked to the energy required for the process of adding or subtracting an electron to that respective orbital while neglecting any response of the other orbitals to that process (frozen orbital approximation).¹⁰⁴ The ionization potential of an occupied orbital χ_c can therefore be defined within that frozen orbital approximation as:¹⁰³

$$IP = -\epsilon_c \tag{2.16}$$

while the electron affinity of an unoccupied orbital ϕ_r is given accordingly:¹⁰³

$$\mathbf{EA} = -\epsilon_r \tag{2.17}$$

2.1.2. Electron Correlation

While the Hartree Fock approach treats exchange in an exact manner, electron correlation is inherently absent and the Hartree Fock energy E_{HF} is not the ground state energy E_0 of the N-electron system. The difference between E_0 and E_{HF} is defined as the correlation energy:¹⁰⁵

$$E_C^{HF} = E_0 - E_{HF} (2.18)$$

Due to the variational principle,¹⁰⁶ that guarantees that the true ground state energy E_0 is always lower than E_{HF} , E_C must be of negative value. The main contribution to E_C is dynamical correlation, a short range effect which describes the situation that due to the average potential electrons move too close to each other and in turn get instantaneously repelled leading to an exaggerated electron repulsion term. The other contribution, the static correlation arises from the neglect of nearly degenerate electron configurations due to the description of the ground state by merely a single Slater determinant.⁹⁷

An logical step to incorporate correlation effects is to use a linear combination of several Slater determinants instead of a single one.¹⁰⁷ This idea is pursued by configuration interactions (CI) and coupled cluster (CC) approaches in which the exact wave function is described as sum of the ground state Hartree-Fock determinant (reference determinant) and determinants obtained by excitation of electrons from the ground state into the unoccupied Hartree-Fock orbitals^{*}. Inclusion of all excited determinants, which would in principle yield the exact ground state wave function, is in practice not tractable for most but the smallest systems; therefore the expansion is usually truncated after the double and triple excitations.¹⁰⁷ Due to its high accuracy CCSD(T), which includes double and approximated triple excitations, has become the gold standard in computational chemistry. Unfortunately its high scaling of M^7 with the number of basis functions it is not applicable to large systems of many hundreds of atoms.

2.2. Density Functional Theory (DFT)

2.2.1. Formal Derivation of DFT: The Hohenberg-Kohn Theorems

Unlike in wave function based methods, where the information of the total system is contained in the many body wave function, Density Functional Theory (DFT) defines any property of the system as a unique functional of the systems ground state density. In its formal representation it includes the exact many-body interactions of that system and is valid for any system of interacting particles. The density of the system is defined by:¹⁰⁹

$$n(\mathbf{r_1}) = N \int \dots \int \left\| \Psi(\mathbf{x_1}, \mathbf{x_2}, \dots, \mathbf{x_N}) \right\| d\sigma_1 d\mathbf{x_2}, \dots, d\mathbf{x_N}$$
(2.19)

and corresponds the probability of finding any of the N electrons with arbitrary spin in that volume element $d\mathbf{r_1}$. Consequently, integration of $n(\mathbf{r})$ over the whole space yields the total number of electrons in the system:¹⁰⁹

$$N = \int n(\mathbf{r}) d\mathbf{r} \tag{2.20}$$

^{*}A profound treatment of correlated methods can be found in Ref. 108.

The foundation of Density Functional Theory (DFT) have been laid by the famous Hohenberg-Kohn theorems, stated and proved by Walther Kohn and Pierre Hohenberg in 1964.¹¹⁰ The first Hohenberg-Kohn theorem states that "the external potential $V_{ext}(\mathbf{r})$ is (to within a constant) a unique functional of $n(\mathbf{r})$ ".¹¹⁰ In order to proof the first Hohenberg Kohn theorem, one has therefore to show that the map from $V_{ext}(\mathbf{r})$ to $n(\mathbf{r})$ is a bijection.

In the following the proof of the First Hohenberg-Kohn theorem will be briefly sketched for the case of non-degenerate ground states; an extension to the degenerate ground state can be found in Ref. 111 and Ref. 112.



Figure 2.1. – Visualization of the maps from the set of external potentials, \mathcal{V} onto \mathcal{G} (map A) and subsequently onto \mathcal{N} (map B). The First Hohenberg Kohn theorem states, that A and B are injective, i.e. the maps indicated by dashed lines do not exist. The Figure was adapted from Ref. 111.

The basic idea behind the proof of the First Hohenberg-Kohn is visualized in Fig. 2.1. Starting point is the set of external potentials, \mathcal{V} , that differ by more than a constant^{*}:

$$\mathcal{V} = \{ v_{ext} : \exists \text{ corresponding, non-degenerate } \Psi_0, v'_{ext} \neq v_{ext} + const , \forall v'_{ext} \in \mathcal{V} \}$$
(2.21)

Solution of Schrödinger's equation would then correspond to a map from \mathcal{V} onto a set \mathcal{G} that contains all ground state wave functions that differ by more than a global phase, ϕ :¹¹¹

$$\mathcal{G} = \{\Psi_0 : \exists \text{ corresponding } v_{ext}, \Psi'_0 \neq \exp^{i\phi} \Psi_0 \forall \Psi'_0 \in \mathcal{G}\}$$
(2.22)

This map corresponds to the red arrows in Fig. 2.1. Note that by definition the set \mathcal{G} contains only those wavefunctions that are constructed from an external potential, i.e. the map from \mathcal{V} to \mathcal{G} is assumed to be surjective.

The set \mathcal{N} of the ground state densities that arise from an antisymmetric wave

^{*}A constant shift in v_{ext} would merely add a constant to the eigenvalue ϵ_0 but not affect Φ_0

function can be defined accordingly:

$$\mathcal{N} = \left\{ n_0 : n_0(\mathbf{r}) = N \int \dots \int |\Psi_0(\mathbf{x_1}, \mathbf{x_2}, \dots, \mathbf{x_N}) d\sigma_1 d\mathbf{x_2}, \dots, d\mathbf{x_N}, \Psi_0 \in G \right\}$$
(2.23)

Similar to \mathcal{G} , \mathcal{N} obtains a priori only ground state densities that are constructed from an antisymmetric wave function, i.e. the map $\mathcal{G} \to \mathcal{N}$ is per definition surjective^{*}.

In order to prove the first Hohenberg Kohn theorem one has therefore to show, that the map from \mathcal{V} to \mathcal{N} is injective, which is done in the following for the map A and B, respectively.

The proof for A uses reductio ad absurdum, i.e. assumes that there are indeed two external potentials v_{ext} and v'_{ext} that differ by more than just a constant but still yield the same ground state.¹¹⁰ Starting from the electronic Schrödinger equation:

$$\hat{H} = \hat{T} + \hat{V}_{ext} + \hat{V}_{e-e} = -\frac{\hbar}{2m} \sum_{i=1}^{N} \nabla_i^2 + \sum_{i=1}^{N} v_{ext}(\mathbf{r_i}) + \sum_{i=1}^{N} \frac{e^2}{\|\mathbf{r_i} - \mathbf{r_j}\|}$$
(2.24)

inserting of the two external potentials, v_{ext} and v'_{ext} , leads to:¹¹¹

$$\left(-\frac{\hbar}{2m}\sum_{i=1}^{N}\nabla_{i}^{2}+\sum_{i=1}^{N}v_{ext}(\mathbf{r_{i}})+\sum_{i=1}^{N}\frac{e^{2}}{\|\mathbf{r_{i}}-\mathbf{r_{j}}\|}\right)|\Psi_{0}\rangle=\epsilon_{0}|\Psi_{0}\rangle$$
(2.25)

$$\left(-\frac{\hbar}{2m}\sum_{i=1}^{N}\nabla_{i}^{2}+\sum_{i=1}^{N}v_{ext}^{\prime}(\mathbf{r_{i}})+\sum_{i=1}^{N}\frac{e^{2}}{\|\mathbf{r_{i}}-\mathbf{r_{j}}\|}\right)|\Psi_{0}\rangle=\epsilon_{0}^{\prime}|\Psi_{0}\rangle.$$
(2.26)

Subtracting both equations from each other yields:

$$\sum_{i=1}^{N} \left(v_{ext}(\mathbf{r}_{i}) - v'_{ext}(\mathbf{r}_{i}) \right) = \epsilon_{0} - \epsilon'_{0}.$$

$$(2.28)$$

The expression on the left side of Eq. (2.28) is dependent on the spatial coordinates, whereas the right side is a constant. Hence, any two potentials that differ by more than just a constant would lead to a contradiction since a variation in the spatial coordinates leads to non-constant values on the left side whereas the right side remains unaffected.¹¹¹

Following a similar reasoning for B, it can be shown that the two wave functions, $\Phi_0 \in \mathcal{G}$, and $\Phi'_0 \in \mathcal{G}$, can impossibly lead to the same ground state density. For \hat{H}

^{*}The requirement that the ground state density has to arise from an antisymmetric wave function is called N-representability and it has been proven, that any non-negative function can be expressed by such an antisymmetric N-body wave function according to Eq. 2.19.^{113,114}

Conversely, for the assumption that each antisymmetric wave function is constructed from an external potential, which is known as V-representability, such a formal proof does not exist.

acting on Φ_0 and Φ'_0 one obtains according to the variational principle:¹⁰⁶

$$E_{0} = \langle \Phi_{0} | \hat{H} | \Phi_{0} \rangle < \langle \Phi_{0}^{'} | \hat{H} | \Phi_{0}^{'} \rangle$$

$$(2.29)$$

Since \hat{H} and \hat{H}' differ merely in their external potential \hat{H} can be expressed in terms of \hat{H}' as $\hat{H} = \hat{H}' + V - V'$ and Eq. (2.29) becomes

$$E_{0} < \langle \Phi_{0}^{'} | \hat{H}^{'} + V - V^{'} | \Phi_{0}^{'} \rangle = E_{0}^{'} + \langle \Phi_{0}^{'} | V - V^{'} | \Phi_{0}^{'} \rangle$$
(2.30)

Assuming $n_0(\mathbf{r})$ is the corresponding ground state density to both, Φ_0 and Φ'_0 , Eq. (2.30) can be rewritten as:¹¹¹

$$E_0 < E'_0 + \int n_0 \left(\mathbf{r}\right) \left[v\left(\mathbf{r}\right) - v'\left(\mathbf{r}\right)\right] d^3\mathbf{r}$$
 (2.31)

Performing the same procedure for Φ'_0 yields:¹¹¹

$$E_{0}^{'} = \langle \Phi_{0}^{'} | \hat{H}^{'} | \Phi_{0}^{'} \rangle < \langle \Phi_{0} | \hat{H}^{'} | \Phi_{0} \rangle = E_{0} + \langle \Phi_{0} | V^{'} - V | \Phi_{0} \rangle$$
$$E_{0}^{'} < \int n_{0} \left(\mathbf{r} \right) \left[v^{'} \left(\mathbf{r} \right) - v \left(\mathbf{r} \right) \right] d^{3}\mathbf{r}$$
(2.32)

The sum of Eq. (2.31) and Eq. (2.32) leads to the contradiction:

$$E_0 + E'_0 < E_0 + E'_0 \tag{2.33}$$

As a result of the first Hohenberg Kohn Theorem, any ground state observable can be obtained from the density, as demonstrated below for the ground state density:

$$E[n] = \langle \Psi_0[n] | \hat{H} | \Psi_0[n] \rangle = \langle \Psi_0[n] | \hat{T} + \hat{W} + \hat{V}_{ext} | \Psi_0[n] \rangle$$
(2.34)

$$= F[n] + \int v_{ext}(\mathbf{r})n(\mathbf{r}) \qquad (2.35)$$

 $F[n] = \langle \Psi_0[n] | \hat{T} + \hat{W} | \Psi_0[n] \rangle$ is commonly denoted universal functional, or Hohenberg-Kohn functional.¹¹⁵

Due to the one-to-one correspondence between $|\Psi_0\rangle$ and n_0 the minimum principle holds also for E[n], and all densities $n'_0 \in N$ that are not the ground state density n_0 yield a higher energy:¹¹⁵

$$E[n_0] < E[n_0'] \tag{2.36}$$

Eq. 2.36 is also called the *second Hohenberg-Kohn Theorem*. It states that, given the universal functional, F[n], is known, the ground state density and energy can be obtained by minimizing the systems total energy with respect to the density:¹¹⁵

$$\frac{\partial}{\partial n(\mathbf{r})} \left\{ E[n] - \mu \left(\int n(\mathbf{r}) - N \right) \right\} \bigg|_{n(\mathbf{r}) = n_0(\mathbf{r})} = 0$$
(2.37)

2.2.2. The Kohn-Sham Approach

The ansatz by Hohenberg and Kohn formally enables to obtain the ground state observables of the system directly from the ground state density instead of the N-body wave function which reduces the complexity from 3^N to 3 spatial coordinates. Moreover, the ground state energy functional gives direct access to the systems ground state energy via Eq. (2.37). However, the universal functional F[n] as well as the energy functional $E_0[n]$ for N interacting particles is unknown which impedes practical application of the Hohenberg-Kohn formulation of DFT.

The problem with the unknown energy functional $E_0[n]$ has been ingeniously circumvented in the Kohn-Sham approach by replacing the interacting N particle system by an auxiliary system of N non-interacting particles with the same ground state density^{116*}. For that purpose, the exact energy functional is decomposed into a part arising from the non-interacting auxiliary system and a comparably small remainder which contains all many-body effects.¹¹²

$$E[n] = T[n] + W[n] + V_{ext}[n] + V_{N-N} = T_s[n] + E_{Hartree}[n] + E_{xc}[n] + V_{ext}[n] + V_{N-N}$$
(2.38)

 $T_s[n]$ is the single particle kinetic energy functional of n and in its magnitude comparable to T[n]. Unfortunately, the exact dependence of $T_s[n]$ on n is not known. By reintroducing the concept of orbitals, $T_s[n]$ can be written in terms of single-particle orbitals ϕ_i and the kinetic energy of the total system can be evaluated as the sum over the kinetic energy of the individual, non-interacting particles:¹¹²

$$T_{s}[n] = T_{s}[\{\phi_{i}[n]\}] = -\frac{\hbar^{2}}{2m} \sum_{i=1}^{N} \int \phi_{i}^{*}(\mathbf{r}) \nabla^{2} \phi(\mathbf{r}) d^{3}r$$
(2.39)

The electron-electron interaction W[n] in the left side of Eq. 2.38 is replaced by the Coulomb energy of the electron density interacting with itself, denoted as Hartree energy $(E_{Hartree})$:¹¹²

$$E_{Hartree}[n] = \frac{1}{2} \int \frac{n(\mathbf{r})n(\mathbf{r}')}{\|\mathbf{r} - \mathbf{r}'\|} d^3r d^3r'$$
(2.40)

 $V_{ext}[n]$ incorporates the electrostatic interactions between the electrons and an exter-

^{*}The requirement, that for any interacting particle system with ground state density n_0 a non-interacting N particle system with the same density n_0 can be found is denoted as *non-interacting v-representability* and not rigorously proven.¹¹² For a more detailed discussion on the non-interacting v-representability see Ref. 117.

nal potential, including the electron-nuclei interactions:¹¹²

$$V_{ext}[n] = \int V_{ext}(\mathbf{r})n(\mathbf{r})d^3r \qquad (2.41)$$

 V_{N-N} denotes the electrostatic, constant interactions between the stationary nuclei. The density of the auxiliary non-interacting system, which by definition is the same of the original N-body system, can be expressed as:¹¹²

$$n(\mathbf{r}) = \sum_{i=1}^{N} \|\psi_i(\mathbf{r})\|^2$$
(2.42)

Decomposition of the total energy according to the right side of Eq. (2.38) has the advantage that the kinetic energy term as well as the long range Hartree term can be evaluated exactly while only the much smaller remainder, $E_{xc}[n]$, needs to be approximated. $E_{xc}[n]$, denoted as exchange-correlation functional comprises the differences in the kinetic energy and the electron-electron interaction between the original N-body system and the auxiliary system of N non-interacting particles:¹¹²

$$E_{xc}[n] = \underbrace{T[n] - T_s[\{\phi_i[n]\}]}_{T_c} + W[n] - E_{Hartree}[n]$$

$$(2.43)$$

Since $T_s[\{\phi_i[n]\}]$ is expressed in terms of single-particle orbitals, i.e. a single Slater determinant, all effects that arise from the anti-symmetric principle, most notably exchange, are innately included in $T_s[\{\phi_i[n]\}]$. Hence, the difference between T[n] and $T_s[\{\phi_i[n]\}], T_c$, stems exclusively from correlation effects that are absent in the single particle kinetic energy.

The energy of the ground state is then obtained by minimizing Eq. 2.38 with respect to the wave functions:¹¹²

$$\frac{\partial E_{KS}}{\partial \psi_i^*(\mathbf{r})} = \frac{\partial T_s}{\partial \psi_i^*(\mathbf{r})} + \left[\frac{\partial E_{ext}}{\partial n(\mathbf{r})} + \frac{\partial E_{Hartree}}{\partial n(\mathbf{r})} + \frac{\partial E_{xc}}{\partial n(\mathbf{r})}\right] \frac{\partial n(\mathbf{r})}{\partial \psi_i^*(\mathbf{r})}$$
(2.44)

under the constraint that the wave functions remain orthonormal. Inserting expression (2.39), and (2.42) into Eq. (2.44) leads to the Kohn-Sham equations:¹¹²

$$\left(\hat{H}_{KS} - \epsilon_i\right)\psi_i(\mathbf{r}) = 0 \tag{2.45}$$

with H_{KS} as effective Hamiltonian, and $V_{KS}(\mathbf{r})$ as the effective potential:

$$\hat{H}_{KS} = -\frac{1}{2}\nabla^2 + V_{KS}\left(\mathbf{r}\right) \tag{2.46}$$

$$V_{KS} = V_{ext} \left(\mathbf{r} \right) + \frac{\partial E_{Hartree}}{\partial n \left(\mathbf{r} \right)} + \frac{\partial E_{xc}}{\partial n \left(\mathbf{r} \right)}$$
(2.47)

2. Theoretical Description of Organic, Semiconducting Molecules

Since V_{KS} depends already on the density, Eq. (2.45), which yields the one-particle orbitals and subsequently the ground state density and energy, needs to be solved self-consistently. So far, no approximations have been made and, given an exact expression of V_{XC} , the Kohn-Sham approach would result in the exact ground state density and energy of the system. Unfortunately, from the three contributions to V_{KS} only the Hartree energy and external potential energy are known whereas the exact form of E_{xc} remains unknown and needs to be approximated. Finding a good enough approximation of E_{xc} is the major challenge in DFT; some of the most common approaches are discussed in the next section.

2.2.3. Exchange Correlation Functionals

Since the emerge of Kohn-Sham DFT, tremendous effort has been spent on finding a suitable approximation for E_{xc} . Starting point of that quest is the Local (Spin) Density Approximation (L(S)DA). LDA is based on the model of a homogeneous electron gas, namely the electron density is assumed to be constant everywhere. Within LDA, the exchange-correlation functional takes the following form:¹¹⁸

$$E_{xc}^{LDA}\left[n\right] = \int n(\mathbf{r})\epsilon_{xc}(n(\mathbf{r})d\mathbf{r}$$
(2.48)

 $\epsilon_{xc}(n(\mathbf{r}))$ is defined as the exchange-correlation energy per particle of a homogeneous electron gas of density $n(\mathbf{r})$ and can be separated into an exchange and correlation part:¹¹⁸

$$\epsilon_{xc}(n(\mathbf{r})) = \epsilon_x(n(\mathbf{r})) + \epsilon_c(n(\mathbf{r}))$$
(2.49)

Advantageously, an exact analytic expression has been derived for the exchange part by Bloch and Dirac^{118}

$$\epsilon_x = -\frac{3}{4} \sqrt[3]{\frac{3n(\mathbf{r})}{\pi}} \tag{2.50}$$

while highly accurate results for the correlation part have been provided by Monte-Carlo simulations.¹¹⁹

Aside from being a fundament on which more sophisticated expressions for E_{xc} have been developed, LDA has been widely employed in solid state physics and despite its simple form provided good results. Unlike in solids, where the assumption of a homogeneous electron density distribution still has some justification, LDA fails to describe molecular properties, like bond lengths and binding energies, due to the rapidly changing electron density in molecular systems.¹¹⁸ Logically, the consecutive step is to include also variations of the electron density in form of the density gradient into E_{xc} . This approach led to the emergence of the so-called Generalized Gradient Approximation (GGA) of the exchange-correlation functional which in their general form can be written as^{118}

$$E_{xc}^{GGA}[n] = \int f(n(\mathbf{r}, \nabla n(\mathbf{r})))$$
(2.51)

Usually, $E_{xc}^{GGA}[n]$ is further divided into its exchange and correlation part, respectively¹¹⁸

$$E_{xc}^{GGA}[n] = E_x^{GGA}[n] + E_c^{GGA}[n]$$
(2.52)

and both terms are treated individually. Unlike LDA functionals which stem all from the physical picture of the homogeneous electron gas, there is a multitude of inherently different GGA functionals that are not always derived from a physical model but rather designed to match certain boundary conditions.¹¹⁸ Compared to L(S)DA functionals, GGA functionals tend to improve total energies of atoms,¹²⁰ binding energies¹²¹ and atomization energies,¹²⁰ as well as structural energy differences, energy barriers, and phase transitions^{122,123} significantly.

Since the main contributions to E_{xc} emanate from exchange, the most natural approach to obtain the best approximation for exact exchange-correlation functional is to replace a fraction of E_x by the accessible exact exchange energy of a Slater determinant. Exchange-correlation functionals, which incorporate a fractional amount of exact exchange are commonly referred to as hybrid functionals. Based on perturbation theory^{124,125} it could be shown that a fraction of 20 % – 25 % is reasonable in terms of accuracy. When using the PBE exchange-correlation functional and 25 % of exact exchange, one obtains the well-known PBE0 functional¹²⁴

$$E_{xc}^{PBE0} = E_{xc}^{PBE} + 0.25 \left(E_x^{HF} - E_x^{GGA} \right)$$
(2.53)

2.2.4. Physical Meaning of the Eigenvalues in Kohn-Sham DFT

As stressed in Section 2.1.1, in Hartree-Fock theory the eigenvalues ϵ_i can be linked to either an electron attachment to or an electron removal of the respective orbital by Koopmans theorem.¹⁰⁴ In DFT, the Slater determinant build from the single particle Kohn-Sham orbitals does not coincide with the many-body ground state wave function. Therefore, the corresponding eigenvalues in the Kohn-Sham formalism cannot simply be assigned to any physical property. One exception is the eigenvalue of the highest occupied state of a finite system^{*}, which equals the negative of the ionization potential, -IE. The reason is that the highest eigenvalue of a bound system dominates the asymptotic long-range behavior of the density which by definition equals the density of the true, interacting N-particle system which, in turn, implies the eigenvalue to be exact.¹³⁰ Notably, the equality between the highest eigenvalue and -IE holds only for highest eigenvalues stemming from the exact V_{xc} whereas all current exchange-correlation functionals that approximate V_{xc} fall short to deliver the correct ionization potential.¹³⁰ Reason for this failure is the self interaction error,

^{*}Nevertheless, the Kohn-Sham eigenvalues have shown promising results when employed in qualitative MO schemes.^{126–129}

2. Theoretical Description of Organic, Semiconducting Molecules

i.e. the physically meaningless interaction of an electron with itself that is inherently circumvented in Hartree Fock theory where Coulomb and Exchange term cancel each other for i = j. While the exact exchange-correlation functional is expected to account for the self-interaction in a similar fashion, in approximate Kohn-Sham DFT the Coulomb term is not equaled out by the exchange term for i = j and the repulsion of the electron by itself leads to an overestimation of the HOMO energy which, in turn, does not longer correspond to the correct ionization potential of the system.

2.2.5. Band Gaps of Solids within DFT

The fundamental band gap, G, of an organic semiconductor is defined as the difference between the ionization potential, IP, and the electron affinity, EA:

$$G = EA - IP \tag{2.54}$$

Since the fundamental band gap G of an organic semiconductor is an essential property in their use as opto-electronic material, an accurate prediction of that band gap by DFT would be highly desirable. Unfortunately, current LDA and GGA functionals underestimate the band gap of semiconducting solids by about 40 %.¹³¹ The failure to produce the correct fundamental band gap does not stem exclusively from the approximated exchange correlation functional but is inherently present in exact DFT. As pointed out by Perdew and Levy, the derivation of the exact exchange correlation energy, E_{xc} , with respect to the occupation exhibits discontinuities for integer electron numbers¹³¹ and the Kohn-Sham band underestimates G by an amount equal to that derivative discontinuity:¹³¹

$$G = \epsilon_{N+1}(N) - \epsilon_N(N) + C \tag{2.55}$$

with N being the total number of electrons and C the derivative discontinuity which is undetermined. As a consequence, even in exact DFT, the Kohn-Sham gap differs from the fundamental gap by a constant, C. Approximate LDA and GGA functionals that suffer from spurious interactions of the electron with itself additionally overestimate the HOMO energy and hence the ionization potential and therefore produce too small band gaps.

2.2.6. Non-covalent Interactions in Kohn-Sham DFT

Dispersion interactions play a dominant role in structural self-assembly, dynamical properties and spectroscopic signatures of a vast number of biological and chemical systems. Originating from interactions between multipoles induced by instantaneous charge density fluctuations, they are a pure quantum mechanical, attractive force¹³² and strongly non-local.¹³³ LDA and GGA exchange-correlation functionals that merely depend on the local density and its gradient respectively, account only for the local contributions to the dispersion energy but inherently fail to capture the dispersion interactions at large distances.^{133,134} As a result, LDA and GGA functionals describe

non-covalent interactions arising from dispersion effects like $\pi - \pi$ interactions,¹³⁵ hydrogen bonding in extended systems,^{136,137} and the physisorption of organic molecules¹³⁸ only poorly or even qualitatively wrong. For instance, current GGA functionals fall even short in describing non-covalent interactions between the most simple model system for π -conjugated molecules, namely two benzene rings. The left panel of Fig. 2.2 depicts binding energy curves for two benzene rings in a stacked configuration, that is mainly stabilized by dispersion interactions, for the LDA functional and the PBE functional as a prominent representative for the GGA functionals. The latter one not only fails to predict the binding energy accurately but also describes the interaction between the stacked benzene molecules as purely repulsive.¹³⁵ The lack of inherently included dispersion interaction in current GGA functionals also challenges the theoretical description of the adsorption structure and stability of organic molecules on inorganic surfaces, due to the significant role of non-covalent interactions in adsorption processes. The right panel of Fig. 2.2 depicts computed adsorption energy curves for a prototypical aromatic molecule, 3,4,9,10-perylene-tetracarboxylic acid dianhydride, (PTCDA), on Ag(111) in comparison with the experimental adsorption height and estimated adsorption energy (yellow cross).¹³⁸ It can be seen that the PBE functional predicts no visible minimum and therefore describes the interaction between the molecule and the metal surface even qualitatively wrong.¹³⁸ While LDA seems to perform better in both discussed cases, this apparently superior performance is in fact randomly and arises from error cancellation.¹³⁸



Figure 2.2. – Left: Binding energy curves for a stacked benzene dimer calculated with CCSD(T) as gold standard as well as LDA and the PBE functional (taken from Ref. 135). Right: Adsorption energy curve of the prototypical aromatic molecule PTCDA on Ag(111) for LDA (orange dash-dotted line) and the PBE functional (blue dotted line) only as well as different dispersion corrections; estimated adsorption energy and experimental adsorption distance are indicated by a yellow cross (taken from Ref. 138).

Considering their huge influence on structure and properties of extended systems, it is therefore not surprising that significant effort has been spent on the inclusion of dispersion interactions into the Kohn-Sham DFT scheme. The most straightforward approach to account for dispersion interactions relies on additive *a posteriori* corrections:¹³⁵

$$E_{\rm tot} = E_{\rm DFT} + E_{\rm Disp} \tag{2.56}$$

2. Theoretical Description of Organic, Semiconducting Molecules

 $E_{\rm DFT}$ is the energy obtained with the approximated functional and is expected to include all other non-covalent interactions except dispersion, whereas $E_{\rm Disp}$ is commonly referred to as dispersion correction. Due to their simplicity, pairwise methods that focus merely on interactions between two atoms and neglect all many-body effects have become a popular approach to model dispersion correction. Considering two sufficiently separated neutral molecules, the interaction energy stems purely from dispersion and decays with the inverse sixth power of the intermolecular distance:¹³⁵

$$E_{disp} = -\frac{C_6^{AB}}{R_{AB}^6}$$
(2.57)

 C_{AB} are the system dependent dispersion coefficients for the interaction between atom A and atom B. Widely used pair-wise approaches are the DFT-D family of Grimme *et al.*,¹³⁹⁻¹⁴¹ the DFT+vdW approach by Tkatchenko and Scheffler^{138,142} and the method by Becke and Johnson (BJ),^{143,144} of which the former two are employed in this work. Within the original DFT-D approach, the dispersion energy is given by:¹³⁹

$$E_{Disp}^{DFT-D} = -s_6 \sum_{A}^{N-1} \sum_{B=A+1}^{N} \frac{C_6^{AB}}{R_{AB}^6} f_{damp}(R_{AB})$$
(2.58)

 s_6 denotes a scaling factor and $f_{damp}(R_{AB})$ a damping function which has the form:¹³⁹

$$f_{damp}(R) = \frac{1}{1 + \exp^{-\alpha(R/R_0 - 1)}}$$
(2.59)

with R_0 denoting the averaged van der Waals radii.¹³⁹ The damping function ensures a correct transition from the non-local dispersion energy at long ranges introduced by the dispersion correction to the exchange and correlation energy within the DFT framework at short ranges. For distances smaller than R_0 it decays to zero and therefore prevents any effects of the long-range dispersion correction on covalent bonds. The pairwise dispersion coefficients are combined according to:¹³⁹

$$C_6^{AB} = 2 \frac{C_6^A C_6^B}{C_6^A + C_6^B} \tag{2.60}$$

 C_6^A and C_6^B are the respective atomic coefficients, which, unlike the molecular coefficients, cannot be derived accurately from e.g. molecular polarizabilities.¹⁴⁵ By assuming the additivity of molecular coefficients, the respective atomic coefficients are thus obtained by minimizing the root mean square deviation between the accurate molecular coefficients and molecular coefficients composed additively of atomic coefficients.¹⁴⁵

Despite its significant improvement of the treatment of dispersion interactions compared to uncorrected Kohn Sham DFT, the DFT-D method still has a number shortcomings, most notably the lack of consistent atomic parameters for elements beyond Ne, systematic errors for molecules containing third row elements, and inconsistencies when calculating atomization energies.¹⁴⁰ Subsequent refinement of the DFT-D approach led to the DFT-D2¹⁴⁰ and finally DFT-D3¹⁴¹ method. The latter one was extended to 94 elements and offers a description of dispersion effects for heavy elements that is consistent with that of the lighter elements. Differences to the previous DFT-D and DFT-D2 corrections can be found in the expression of the total dispersion energy which takes a non-additive three body term into account:¹⁴¹

$$E_{disp} = E^{(2)} + E^{(3)} \tag{2.61}$$

In the two body contribution, $E^{(2)}$, it additionally includes the 8th-order dispersion term:¹⁴¹

$$E^{(2)} = \sum_{AB} \left(\frac{C_6^{AB}}{r_{AB}^6} f_{damp,6}(r_{AB}) + s_8 \frac{C_8^{AB}}{r_{AB}^8} f_{damp,6}(r_{AB}) \right)$$
(2.62)

The three body term, $E^{A,B,C}$ involving the atoms A, B, and C is obtained from third order perturbation theory:¹⁴¹

$$E^{ABC} = \frac{C_9^{ABC} \left(3\cos\theta_a\cos\theta_b\cos\theta_c + 1\right)}{\left(r_{AB}r_{BC}r_{AC}\right)^3} \tag{2.63}$$

 θ_a, θ_b , and θ_c denote the internal angles between the three atoms. The $C_9^{A,B,C}$ coefficients are usually approximated by a geometric mean:¹⁴¹

$$C_9^{ABC} \approx -\sqrt{C_6^{AB} C_6^{BC} C_6^{AC}} \tag{2.64}$$

What makes the D3 method superior to the previous DFT-D methods is that it explicitly treats the effect of covalent bond formation in the molecule on the dispersion coefficients by making the reference dispersion coefficients dependent on the coordination number of the respective atom in the molecule.¹⁴¹ Moreover, it is less empirical than the D2 method since the cut-off radii employed in the damping function are now calculated explicitly from first principles for all pairs of the 94 elements while the atomic dispersion coefficients are computed from TD-DFT calculations.¹⁴¹ While the DFT-D3 correction significantly improves the description of non-covalent interactions between molecules it is known to overestimate the adsorption energy of organic molecules on inorganic surfaces as visible in Fig. 2.2 b.¹³⁸

In the DFT+vdW scheme by Tkatchenko and Scheffler,¹⁴² the missing van der Waals interactions are accounted for by adding a pairwise $C_6 R^{-6}$ term to the DFT energy similar to the DFT-D approach.¹⁴² The difference to the DFT-D approach is that the C_6 coefficients and vdW radii are derived from the mean field ground state electron density. The influence of an atom's chemical environment on the C_6 coefficient is taken into account by defining an effective volume of the atom in the molecule or solid referenced to the free atom *in vacuo*:¹⁴²

$$C_{6AA}^{\text{eff}} \approx \left(\frac{V_A^{\text{eff}}}{V_A^{\text{free}}}\right)^2 C_{6AA}^{\text{free}}$$
(2.65)

2. Theoretical Description of Organic, Semiconducting Molecules

The effective volume, V^{eff} , is directly related to the polarizability of the atom in the respective chemical environment¹⁴⁶ and therefore accessible from a Hirshfeld partitioning^{147,148} of the electronic densities.

An extension to the DFT+vdW scheme is the vdW^{surf} method¹³⁸ targeted for the interaction between molecules and surfaces. The vdW^{surf} method¹³⁸ employs the Lifshitz-Zaremba-Kohn (LZK) theory,^{149,150} which describes the interaction between an atom and a surface, in order to determine reduced C₆ coefficients that include non-local screening effects. Both, the pairwise TS vdW correction,¹⁵¹ and the vdW^{surf} dispersion correction¹⁵² have been employed successfully to describe organic molecules on inorganic surfaces. As visible in the right panel of Fig. 2.2, the DFT+vdW and the vdw^{surf} approach lead both to a better prediction of the adsorption energy albeit the former method slightly overestimates the adsorption height.

It is important to keep in mind that in both, the DFT-D and the DFT+vdW approaches the dispersion energy is added *a posteriori* on the DFT total energy and does not directly affect the electron density and hence the molecular properties. However, in geometry optimization the additional dispersion term in the total energy affects the forces and therefore leads to a different minimum structure of the system.

2.2.7. DFT under Periodic Boundary Conditions

As opposed to single molecules in vacuum, the vast number of atoms in a solid makes it theoretical description challenging. Luckily, in crystalline solids the nuclei are arranged in a periodically repeated pattern and thus in practical DFT calculations only the the smallest repeat unit of the periodic pattern has to calculated under periodic boundary conditions in order to describe the whole periodic system.

The structure of a crystalline solid is completely defined by its Bravais lattice and basis. The Bravais lattice denotes the set of translations, $\{R_n\}$, with¹⁵³

$$R_n = l\mathbf{a}_1 + m\mathbf{a}_2 + n\mathbf{a}_3 , \ l, m, n \in \mathcal{N}$$

$$(2.66)$$

with $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3$ being the lattice constants of the Bravais lattice. The primitive unit cell of the system is the smallest unit required to produce the crystal structure by periodic translations through all Bravais lattice vectors.¹⁵³ The basis of a crystal denotes the type and position of the atoms that fill the primitive unit cell. An important primitive unit cell of a crystal is the Wigner-Seitz cell, which possesses the same symmetry as the Bravais lattice. Per definition, the Wigner-Seitz cell of a given lattice point contains all points in space that are closer to that specific lattice point than to any other lattice point.

Due to the translational symmetry of the crystal lattice, the potential generated by the nuclei, $V(\mathbf{r})$, is also periodic, i.e.

$$V\left(\mathbf{r}\right) = V\left(\mathbf{r} + \mathbf{R}_n\right) \tag{2.67}$$

with \mathbf{R}_n being an arbitrary Bravais lattice vector of the system. A theoretical treatment of periodic solids within the framework of DFT therefore requires the solution of Schrödinger's equation of a single electron under the influence of a periodic potential. For a single electron subject to a periodic potential, Bloch's theorem states that the eigenfunctions of the stationary Schrödinger equation are plane waves modulated by a function, $u_k(\mathbf{r})$, that has the same periodicity as the potential, $V(\mathbf{r})^{154}$

$$\phi_k\left(\mathbf{r}\right) = \exp^{i\mathbf{k}\mathbf{r}} u_k\left(\mathbf{r}\right) \tag{2.68}$$

 $\phi_k(\mathbf{r})$ is often referred to as *Bloch wave*. Unlike for a free particle, the wave vector \mathbf{k} of the Bloch wave cannot be associated with the momentum of the electron; instead, the product $\mathbf{k} \cdot \hbar$ can be understood as *crystal momentum*.¹⁵⁵ Since $u_k(\mathbf{r})$ is strictly lattice periodic, it can be expanded in a Fourier series and the wave function becomes¹⁵⁶

$$\psi_k\left(\mathbf{r}\right) = \sum_{\mathbf{G}} C_{\mathbf{k}-\mathbf{G}} \exp^{i(\mathbf{k}-\mathbf{G})}$$
(2.69)

G is a reciprocal lattice vector and fulfils $\exp^{\mathbf{G}\cdot\mathbf{R}_{n}} = 1$. Note that $\psi_{k}(\mathbf{r})$ is periodic in reciprocal space, i.e. wave vectors that differ only by a reciprocal lattice vector, **G** are identical:¹⁵⁶

$$\psi\left(\mathbf{k} + \mathbf{G}\right) = \psi\left(\mathbf{k}\right) \tag{2.70}$$

As a consequence, the electron energy, $\epsilon(\mathbf{k})$ is also periodic in reciprocal space:

$$\epsilon \left(\mathbf{k} \right) = \epsilon \left(\mathbf{k} + \mathbf{G} \right) \tag{2.71}$$

Solving Schrödinger's equation for a given \mathbf{k} results in a countable number of independent solutions. Those solutions are labelled by the quantum number n, which is denoted the band index. In total, each state of a particle in a periodic potential is characterized by three quantum numbers: the wave vector, \mathbf{k} , the band index, n, and the spin (in case of spin unrestricted systems). The band structure of a quasi-free single particle in a zero potential with periodicity $\mathbf{R}_n = \mathbf{a}$ and a single particle in a rectangular periodic potential of height V_0 , width b, and infinitely extended periodicity a+b (Kronig-Penney Model¹⁵⁷) is presented in Fig. 2.3. Due to the relation $\epsilon_n(\mathbf{k}) = \epsilon_n(\mathbf{k} + \mathbf{G})$, every lattice point of the reciprocal lattice can serve as origin of $\epsilon(k)$. In the one-dimensional case of a quasi-free particle this leads to parabolas centered at $0, \pm 2n\pi/a$ as shown in Fig. 2.3 **a**. As a result, the range $-\pi \leq k < \pi$, already contains the complete information about the system. This specific range is denoted the first Brillouin zone of the system and comprises only those vectors \mathbf{k} that are closer to $\mathbf{k} + \mathbf{G}_{\mathbf{m}} = 0$ than to any other lattice point.¹⁵⁸ The point k = 0 is called Γ -point. It can be seen that the band structure of the quasi-free electron shows degeneracies at the Γ -point and the borders of the first Brillouin zone.

The band structure of an electron in a one-dimensional, non-vanishing potential, is presented in Fig. 2.3 c. The solid black line represents the energy curve at k = 0. Instead of degeneracies the parabola now features discontinuities at the borders of the first Brillouin zone, i.e. for these energy values no eigenvalues exist. Those regions are the so called "forbidden bands", while the energy values for which eigenvalues exist are often denoted "allowed bands".¹⁵⁸ For a (closed-shell) system consisting of N periodically repeated unit cells, the number of occupied bands is $n = \frac{N}{2}$. With

2. Theoretical Description of Organic, Semiconducting Molecules

increasing n - and therefore with increasing system size - a broadening of the allowed bands can be observed (Fig. 2.3 c) which emphasizes the concept of extended energy bands in periodic solids in contrast to the sharp energy levels in isolated systems.



Figure 2.3. – a-b Band structure of the free electron in the repeated zone scheme (a) and the reduced zone scheme comprising only the first Brillouin zone (b). c: Gap opening for an electron in a one-dimensionally periodic potential. The figure was taken from Ref. 158.

2.2.7.1. k-Point Sampling in the First Brillouin Zone

In periodic DFT calculations commonly expressions of the form¹⁵⁹

$$\bar{g} = \frac{V_{cell}}{\left(2\pi\right)^3} \int_{BZ} g\left(\mathbf{k}\right) d\mathbf{k}$$
(2.72)

have to be integrated over the first Brillouin zone. In practice, only a finite number of k-points can be evaluated. Therefore, a grid of discrete **k** values has to be specified and the integral is approximated by a sum over those k-points. A careful selection of the k-points is crucial to obtain the optimal balance between accuracy and computational effort: sampling of too few k-points will lead to inaccurate results whereas taking into account too many k-points will slow down the calculation tremendously or even make it infeasible. A convenient approach to sample the first Brillouin zone is the Monkhorst-Pack scheme.¹⁶⁰ For a general point **R** on a lattice generated by the translation vectors $\mathbf{t}_1, \mathbf{t}_2, \mathbf{t}_3$ in real space

$$\mathbf{R} = R_1 \mathbf{t}_1 + R_2 \mathbf{t}_2 + R_3 \mathbf{t}_3 \tag{2.73}$$

the corresponding reciprocal lattice vectors are defined by¹⁶⁰

$$\mathbf{b}_1 = \frac{2\pi}{V} \mathbf{t}_2 \times \mathbf{t}_3 , \ \mathbf{b}_2 = \frac{2\pi}{V} \mathbf{t}_3 \times \mathbf{t}_1 , \ \mathbf{b}_3 = \frac{2\pi}{V} \mathbf{t}_1 \times \mathbf{t}_2$$
(2.74)
with v as the volume of the unit cell. Within the Monkhorst-Pack approach an equidistant grid of k-points, \mathbf{k}_i is now generated according to¹⁶⁰

$$\mathbf{k}_{prs} = u_p \mathbf{b}_1 + u_r \mathbf{b}_2 + u_s \mathbf{b}_3 \tag{2.75}$$

 u_r is defined by¹⁶⁰

$$u_r = \frac{(2r - q - 1)}{2q} , \ r = 1, 2, 3, ..., q , \ q \in \mathbb{N}$$
(2.76)

The integer q defines the number of k-points in each direction. Therefore, in practical DFT applications q has to be simply specified for all three directions in order to generate a k-point mesh of q^3 distinct points that are equidistant in the Brillouin zone. This simple and straightforward approach makes the Monkhorst-Pack scheme a widely used method for k-point sampling that is implemented in most solid state codes.

2.2.8. Density Functional Theory for Metallic Systems

2.2.8.1. Finite Temperatures

In the following the intricacies in describing metallic systems with DFT will be briefly discussed. As mentioned in Section 2.2.7, an electron in a periodic solid can be characterized by the quantum numbers **k**, n and σ (for spin unrestricted systems) with **k** being the wave vector, n the band index, and σ the spin of the electron. For a crystal consisting of N unit cells, there are N allowed k-values in the first Brillouin zone and taking into account to the Pauli Exclusion principle, each band can therefore contain 2N electrons.¹⁶¹ At T = 0 K the states are successively filled in a bottom-up procedure until all electrons are accommodated. The highest completely filled band is denoted as valence band, whereas a band that is only partially filled is denoted as conducting band. For electrons in a one-dimensional potential, as exemplarily discussed in section 2.2.7, the energy between the highest occupied and the lowest unoccupied state is called the Fermi energy, E_F .¹⁶² In three-dimensional **k**-space, i.e. with $\mathbf{k} = (k_x, k_y, k_z)^T$, many **k** correspond to one energy and one obtains surfaces of equal energy. Correspondingly, the energy surface that separates occupied from unoccupied states is denoted as Fermi surface.¹⁶² The concept of the Fermi energy allows to distinguish metals from insulators. Given that at T = 0 K the highest occupied band is completely filled and the successive band is completely empty, the Fermi energy is between the highest occupied and lowest unoccupied state and the material is not conducting. The size of the band gap between the valence band and the empty conducting band defines, whether the material is insulating or semiconducting: For large energy gaps between valence and conduction band the material is an insulator whereas for small energy gaps (around $\approx 1 \,\mathrm{eV^{163}}$) excitations of electrons from the valence to the - previously empty - conduction band at finite temperatures are possible which makes the material semiconducting.¹⁶² In case of close lying bands, neighboring bands can overlap which results in partially filled conducting bands and hence metallic properties. For odd-valency solids (i.e. the number of valence electrons is 1,3,5,...) the

2. Theoretical Description of Organic, Semiconducting Molecules

conducting band is necessarily half filled and the material conducting.

The half filled band of metals leads to problems when numerically evaluating integrals over all occupied states in the first Brillouin zone. The probability of an energy level, E, being occupied is given by the Fermi function¹⁶²

$$F(E) = \frac{1}{\exp^{\left(\frac{E-E_f}{k_B T}\right)} + 1}$$
(2.77)

 E_F is the Fermi energy, T is the temperature of the system and k_B denotes the Boltzmann constant. In the limit of $T \to 0$, F(E) becomes a step function¹⁶²

$$\theta \left(E - E_F \right) = \begin{cases} 1 & \text{if } E < E_F \\ 0, & \text{if } E > E_F \end{cases}$$

$$(2.78)$$

i.e. F(E) exhibits a discontinuity at $E = E_F$. Numerical integration would therefore require a very fine k-point grid which considerably increases the computational effort. One option is therefore to introduce an electronic temperature $T \neq 0$ K which leads to a smooth distribution function. This approach is called "Fermi smearing". Another way to increase convergence with the number of k-points is to replace the step function by a Gaussian-type broadening function¹⁶⁴

$$F(E_{n,\mathbf{k}}) = \frac{1}{2} \left(1 - \operatorname{erf}\left(\frac{E_{n,\mathbf{k}} - E_f}{\Delta}\right) \right)$$
(2.79)

with Δ being the width of the Gaussian.

2.2.8.2. Relativistic Effects

A central aspect in then description of metals is the inclusion of relativistic effects. While in light elements relativistic effects are negligible, in heavier elements they strongly influence physical and chemical properties. Relativistic effects are one reason for the different behavior of the elements in the fifth period (Rb-Xe) compared to the heavy elements in the subsequent, sixth period (Cs-Rn).¹⁶⁵ Other famous examples are the color of gold,¹⁶⁶ the liquidity of mercury¹⁶⁷ and the lead battery in which the main amount of its standard voltage stems from relativistic effects.¹⁶⁸

The relativistic analogue to the non-relativistic Schrödinger equation is given by the four component Dirac equation, which works on a four component wave function, Ψ ,¹⁶⁹ with

$$\Psi = \begin{pmatrix} \phi \\ \chi \end{pmatrix} \tag{2.80}$$

 ϕ will henceforth be denoted large component and χ will be denoted small component

of Φ .¹⁶⁹ The four component Dirac equation is¹⁶⁹

$$V\phi + c\sigma \cdot \mathbf{p}\chi = E\phi \tag{2.81}$$

$$c\sigma \cdot \mathbf{p}\phi + \left(V - 2c^2\right)\chi = E\chi \tag{2.82}$$

 \mathbf{p} is the momentum of the electron, V the external potential, and c the speed of light in vacuum.

In the approach of van Lenthe *et al.*, the four component Dirac equation is first reduced to an effective two component form by eliminating χ , i.e. expressing χ in ϕ by rearranging Eq. 2.81.

$$\chi = \frac{1}{2c^2 + E - V} \cdot c\sigma \cdot \mathbf{p}\phi \tag{2.83}$$

The Hamiltonian, \hat{H}^{esc} , acts now solely on ϕ and is subsequently expanded in $1/(2c^2 - V)$ which yields (in zeroth order):¹⁶⁹

$$\hat{H}^{esc} \approx V + \sigma \cdot \mathbf{p} \frac{c^2}{2c^2 - V} \sigma \cdot \mathbf{p}$$
(2.84)

The Hamiltonian defined by eq. 2.84 is denoted as zeroth order regular approximated (ZORA) Hamiltonian and core piece of the widely employed ZORA (Zeroth Order Regular Approximation) approach.¹⁶⁹ The ZORA approach is also implemented in the electronic structure code FHI-AIMS¹⁷⁰ which is used in this work and provides results that are in excellent agreement with benchmark results from linearized augmented plane wave ((L)APW) calculations¹⁷⁰ that treat the core states fully relativistic.^{171,172}

2.2.9. Basis Sets

Due to the reintroduction of the single-particle orbitals in Kohn-Sham DFT, an appropriate basis has to be chosen to represent the orbitals. Popular choices are Gaussian type basis functions, which are particularly suited for gas phase molecules and available in all standard quantum chemistry codes or a plane wave basis, which is commonly employed for periodic solids. Gaussian type orbitals (GTO) are generally written as:¹⁷³

$$\eta^{GTO} = N x^l y^m z^n \exp^{[-\alpha r^2]} \tag{2.85}$$

with N being the normalization factor that guarantees $\langle \eta_{\nu} | \eta_{\nu} \rangle = 1^*$. α is the orbital exponent that can be adjusted in order to render the basis function either more diffuse or compact. The sum L = l + m + n defines the type of the orbital, i.e. s-type orbital (L = 0), p-type orbital (L = 1), d-type orbital (L = 2) etc. From a physical point of view, Gaussian type functions are not exact eigenfunctions of the hydrogen atom. In particular, a Gaussian function does not exhibit the correct behavior at r = 0 (a cusp) nor does it decay exponentially in the tail regions for $r \to \infty$.¹⁷³ However, the use of

^{*}the basis functions are not necessarily orthonormal, i.e. $\langle \eta_{\nu} | \eta_{\mu} \rangle = 0$ does not hold for $\nu \neq \mu$

2. Theoretical Description of Organic, Semiconducting Molecules

Gaussian type basis functions allows a very efficient and analytical evaluation of the four-center-two-electron integrals in the Coulomb term (and Hartree Fock exchange term). One therefore combines linearly N Gaussian type functions to one *contracted Gaussian function* (CGF) in order to accurately approximate the correct behavior at r = 0 and $r \to \infty$:¹⁷³

$$\eta_{\tau}^{CGF} = \sum_{a}^{N} d_{a\tau} \eta_{a}^{GTO} \tag{2.86}$$

The contraction coefficients $d_{s\tau}$ can be adjusted to mimic as much as possible the cusp at r = 0 as well as the correct asymptotic behavior which requires at least three GTO.¹⁷³ Increasing the number of basis functions improves the accuracy of the description of the atomic orbitals and subsequently leads to the so called "double-zeta" basis set (two times the minimum number of basis functions), "triple-zeta" basis set (three times the minimum number of basis functions) and "quadruple-zeta" basis set (four times the minimum number of basis functions).¹⁷⁴ In order to reduce the computational costs, attempts have been made to express only the valence orbitals involved in chemical bonding in a large basis set whereas the core electrons are described by a small basis without significant loss of accuracy. Basis sets of that kind are known as split-valence basis sets.¹⁷⁴ In order to allow for a higher polarization of the electronic density, so-called polarization functions, higher angular momentum orbitals, can be added which improves the quality particularly in bonding regions.¹⁷⁴ The advantages of Gaussian type orbitals are the usually relatively small basis that is necessary to obtain good results as well as the chemically intuitive picture of a molecule being assembled of, slightly distorted, atoms and the molecular orbitals can be obtained from a linear combination of the atomic orbitals. The latter point, however, is also disadvantageous since it requires the orbitals to be centered at the respective atom and hence dependent in the spatial coordinates of that atom. Moreover, calculations with atomically centered orbitals are prone to basis set superposition errors especially when studying weakly interacting molecular systems or adsorption processes. Due to the incomplete basis set, interacting molecules are additionally stabilized by employing the extra basis functions of the other molecule which artificially enhances the intermolecular interactions.^{175,176}

A somewhat complementary approach is the usage of plane waves as basis functions which are typically employed for extended, periodic systems. As stressed in section 2.2.7, according to Blochs theorem, in an infinite periodic system the wave function can be represented as a product of a periodic function $u_i(\mathbf{r})$, and a wavelike part:¹⁵⁴

$$\phi_i\left(\mathbf{r}\right) = \exp^{i\mathbf{k}\mathbf{r}} u_i\left(\mathbf{r}\right) \tag{2.87}$$

 $u_i(\mathbf{r})$ obeys the periodicity of the underlying periodic potential, i.e. $u_i(\mathbf{r} + \mathbf{L}) = u_i(\mathbf{r})$ with \mathbf{L} being the lattice vector of the crystal unit cell. $u_i(\mathbf{r})$ can be expanded into a Fourier series

$$u_i\left(\mathbf{r}\right) = \sum_{\mathbf{G}} c_{i,G} \exp^{i\mathbf{G}\mathbf{r}} \tag{2.88}$$

G is the reciprocal lattice vector and fulfils the following relation: $\mathbf{G} \cdot \mathbf{L} = 2\pi m m \in \mathbb{N}$. Inserting Eq. (2.88) in Eq. (2.87) yields the wave function written as a sum of plane waves:

$$\phi_i(\mathbf{r}) = \sum_{\mathbf{G}} c_{i,G} \exp^{i(\mathbf{k}+\mathbf{G})\mathbf{r}}$$
(2.89)

While Bloch's theorem allows in theory to expand the wave function in plane waves, in practice plane wave basis sets impede efficient calculations. Due to the orthogonality of the core and valence wave functions the core wave functions exhibit one (or more) radial nodes.¹⁷⁷ A very large number of plane waves are therefore required to account for the rapidly varying wave function in the core region. In solids the main contribution to the physical properties stems from the valence electrons. An elegant approach towards more computational efficiency is therefore to approximate the ionic potential and the core electrons by an effective potential, the pseudopotential, such that the corresponding pseudo wave functions possess no radial nodes in the core region.¹⁷⁷ A comparison of the - oscillating - real wave function and the smoothly decaying pseudo wave function is given in Fig. 2.4. Furthermore, it is required that the scattering properties of the pseudo wave functions match those of the ion and core electrons of the valence wave function.¹⁷⁷ The general form of a pseudopotential is given by:¹⁷⁷

$$V_{NL} = \sum_{l} \left| \ln \right\rangle V_{l} \left\langle \ln \right| \tag{2.90}$$

Application of V_{NL} on the electronic wavefunction leads to a decomposition of the wave function into spherical harmonics, $|\text{lm}\rangle$.¹⁷⁷ V_l is the pseudopotential for angular momentum l. Pseudopotentials described by Eq. 2.90 are referred to as semilocal since they are nonlocal in angular coordinates (i.e. V_l is angular dependent) but local in radial coordinates (i.e. V_l is independent of the radial coordinates).¹⁷⁸ Given that for all angular moments the same potential is used, i.e. V_l is independent of l, the resulting pseudopotential is termed "local".

For a wide usage, the pseudo potential for an atom should be independent of the chemical environment. Such a transferability is provided by norm-conserving pseudopotentials, which meet the following requirements¹⁸⁰

- i the norm must be conserved, i.e. the integral from 0 to r of the pseudo charge density and the real charge density must be the same beyond a certain core cutoff radius, $r > r_c$
- ii the logarithmic derivatives of the real and pseudo wave function and their first energy derivative must be identical beyond that cutoff, r_c
- iii real and pseudo energy eigenvalues of the valence states must be the same
- iv beyond the core cutoff radius r_c , the real and pseudo wave function must coincide

Popular norm-conserving pseudopotentials are the Goedecker-Teller-Hutter pseudopotentials,^{181–183} that have also been employed in this work. They are separable in

2. Theoretical Description of Organic, Semiconducting Molecules



Figure 2.4. – Schematic representation of the all-electron potential, V and the corresponding wave function ϕ (black solid line) in comparison to the pseudopotential, V^{PS} , and the pseudo wave function, ϕ^{PS} . While ϕ exhibits oscillations in the core region, ϕ^{PS} decays smoothly to zero. The figure was taken from Ref. 179

a local and non-local part from which the former one is given by:¹⁸³

$$V_{loc}^{PP}(r) = -\frac{Z_{eff}}{r} erf\left(\alpha^{PP}r\right) + \sum_{i=1}^{4} C_i^{PP}\left(\sqrt{2}\alpha^{PP}r\right)^{2i-2} \exp\left(-\left(\alpha^{PP}r\right)^2\right).$$
 (2.91)

with 183

$$\alpha^{PP} = \frac{1}{\sqrt{2}r_{loc}^{PP}} \tag{2.92}$$

The non-local part is given by 183

$$V_{nl}^{PP}\left(\mathbf{r},\mathbf{r}'\right) = \sum_{lm} \sum_{ij} \langle \mathbf{r} | p_i^{lm} \rangle h_{ij}^l \langle p_j^{lm} | \mathbf{r}' \rangle$$
(2.93)

The projectors, $\langle \mathbf{r} | p_i^{lm} \rangle$ are Gaussian type projectors which are defined by:¹⁸³

$$\langle \mathbf{r} | p_i^{lm} \rangle = N_i^l Y^{lm} \left(\hat{r} \right) r^{l+2i-2} \exp\left(-\frac{1}{2} \left(\frac{r}{r_l} \right)^2 \right)$$
(2.94)

 Y^{lm} are spherical harmonics, and N_i^l the respective normalization constants. The advantage of the GTH pseudopotentials is that they possess an analytic form in both, real and Fourier space,¹⁸³ which is why they are often termed *dual space* pseudopotentials. Moreover, the use of Gaussian functions as projectors enables a fast convergence in real and Fourier space. Since relativistic, all-electron wavefunctions are employed as reference in the parametrization, the relativistic effects of the core electrons are

already accounted for when using pseudopotentials.¹⁸³

2.3. Hybrid Quantum Mechanics/Molecular Mechanics (QM/MM) Approaches

Despite the good scalability of current Kohn-Sham implementations, the vast size of some systems makes it still impossible to treat them on a full quantum mechanical level. Luckily, in many of these cases it is sufficient to consider only the electronic properties of a relevant sub-part of the system, e.g. a small region where bond formation or breaking occurs, whereas the surrounding can be treated on a classical level of theory without any significant loss of accuracy. An elegant solution for such situations are hybrid quantum mechanical/molecular mechanic (QM/MM) approaches, as first introduced by Warshel and Levitt,¹⁸⁴ in which the relevant subsystem is treated on the quantum mechanical level of theory (QM part) whereas the surrounding larger part of the system is described classically by force field approaches (MM part). The interaction between the QM and MM part is not straightforward to access and several schemes have been developed to couple the two regions.^{185–188} Commonly used approaches are additive coupling schemes in which the total energy is partitioned into an energy term of the MM part, an energy term of the QM part and the QM/MM coupling term:¹⁸⁸

$$E_{QM-MM} = E_{MM} + E_{QM} + E_{QM-MM}$$
(2.95)

In case that the QM and MM regions are not connected by covalent bonds the interaction between the QM and MM system can be written as sum of van der Waals interactions, E_{QM-MM}^{vdW} , and electrostatic interactions, E_{QM-MM}^{eS} .¹⁸⁸

$$E_{QM-MM} = E_{QM-MM}^{ES} + E_{QM-MM}^{vdW}$$
(2.96)

The van der Waals interactions are usually modelled by a Lennard-Jones potential while electrostatic effects can be treated on different levels of sophistication. In the most simple approach, the mechanical embedding, the electrostatic interactions are treated merely on the MM level of theory. Partial charges are assigned to the atoms of the QM regions and electrostatic interactions between the QM and the MM region are described by a Coulomb potential. The electronic density of the QM system is evaluated for an isolated system and therefore no polarization effects due to the MM environment are taken into account.¹⁸⁸

In the more advanced electrostatic embedding scheme, polarization effects are accounted for by incorporating the charged atoms of the MM part into the QM Hamiltonian¹⁸⁹

$$h_{i}^{QM-MM} = h_{i}^{QM} - \sum_{j}^{M} \frac{e^{2}Q_{j}}{4\pi\epsilon_{0} \|\mathbf{r}_{i} - R_{j}\|}$$
(2.97)

with h_i^{QM} being the one-electron operator of the QM isolated system. \mathbf{r}_i and \mathbf{R}_j are the positions of electron i and the MM atom J with partial charge Q_j . A common

problem of the polarization embedding is overpolarization at the QM-MM boundary. In standard MM approaches the spatially extended charge distribution of atoms is approximated by point charges. Electrons near the QM-MM boundary get trapped in this electrostatic potential of the point charges which leads to enhanced interaction between the electrons near the QM-MM boundary and the atoms in the MM region. As a result, electronic density is accumulated at the boundary whereas the inner QM region suffers from a depletion of electronic density.¹⁹⁰ A possible way to circumvent that electron spill-out is to spread out the MM charges by replacing the point charges by a Gaussian distributed charge density.¹⁹⁰

Further advances in the embedding of QM and MM region include also polarization effects of the MM region by the electronic density of the QM region (polarization scheme).^{191,192} Such polarization effects are of particular importance in QM/MM calculations of adsorbate/metal interfaces in which the metallic surface is treated on a classical level of theory. Since metals are conducting, the electrostatic potential within the surface must be zero (or constant in case of an external, static potential). When a molecule adsorbs on a metal surface, the electrons within the surface instantaneously redistribute in order to counterbalance the charge density distribution of the molecule such that the electrostatic potential within the surface remains zero. The mutual polarization of the electron density of the molecule and the surface electrons plays a substantial role in structure formation of organic molecules on conducting surfaces^{193,194} and neglecting those polarization effects can lead to qualitatively wrong results. Polarization effects can be incorporated by the method of image charges in which the response of the surface to the charge distribution of the adsorbate is modeled by discrete image charges within the metal. Siepmann and Sprik¹⁹⁵ incorporated such an image charge approach for metal/adsorbate systems into force field calculations. Within the Siepman-Sprik scheme the image charges are modeled as Gaussians centered at the metal atoms. In order to account for the mutual polarization of metal and adsorbate the magnitude of the image charges is determined in a self-consistent fashion. The QM/MM implementation of the Siepman and Sprik scheme^{196,197} is employed in this work to describe the structural assembly of α -sexithiophene layer on gold.

2.4. Time Evolution of Large Systems: Classical Molecular Dynamics

While hybrid QM/MM approaches significantly decrease the computational cost by reducing the size of the system that is treated quantum mechanically, the time scales of molecular dynamic simulations within the QM/MM framework are still restricted to a few nanoseconds.¹⁹⁸ In proteins, dynamical processes like collective motions or folding occur at time scales of tens of nanoseconds to microseconds and milliseconds, respectively.^{26,27} Such time scales are only accessible within a purely classical, force field based framework. In this approach, the atoms of the system are approximated by hard spheres with predefined partial charges that evolve in time according to Newtons laws of motion. In classical molecular dynamics the potential energy of the system

is derived from empirical potential energy functions, usually termed force fields. In common force fields, like the CHARMM force field¹⁹⁹ employed in this work, the potential energy is composed additively of the bonded and non-bonded interactions and takes the following form:¹⁹⁹

$$U(\mathbf{r}) = \sum_{i=1}^{N} U_{bonded}(\mathbf{r}) + \sum_{i=1}^{N^2} U_{nonbonded}(\mathbf{r})$$
(2.98)

The bonded interactions, U_{bonded} , can be further separated into two-body terms (bonds), three body terms (angles) and four body terms (dihedral angles). In standard force fields, bonds and angles are represented as harmonic potential whereas dihedrals are usually modelled using periodic functions:¹⁹⁹

$$U_{bonded} = \sum_{Bonds} k_b \left(\mathbf{r}_{i,j} - \mathbf{r}_0\right)^2 + \sum_{Angles} k_\theta \left(\theta - \theta_0\right)^2 + \sum_{Dihedrals} k_\Phi \left[1 + \cos\left(n \cdot \Phi - \delta\right)\right] + \sum_{Urey-Bradley} k_{ub} \left(\mathbf{r}_{i,k} - \mathbf{r}_{ub}\right)^2 + \sum_{Impropers} k_\omega \left(\omega - \omega_0\right)^2$$
(2.99)

 \mathbf{k}_b and \mathbf{k}_{θ} are the force constants of the bonds and angles within the harmonic approximation whereas \mathbf{r}_0 and θ_0 are the equilibrium distance and angle respectively. The third term captures the contribution from the dihedral angles. A dihedral, or torsion angle, is defined for four atoms as the angle between the two planes that are spanned by the first three atoms and the last three atoms. Term 4, the Urey-Bradley cross term, accounts for interdependence of angles and bonds: For an angle formed by the atoms i,j,k it describes the potential energy change upon angle bending as a function of the distance between the atoms i-k²⁰⁰.

The nonbonded term that accounts for non-covalent interactions is composed of the long-range electrostatic interactions and the short range vdW interactions which are commonly modeled by a Lennard-Jones potential:²⁰¹

$$U_{\text{nonbonded}} = \underbrace{\sum_{i} \sum_{j>i} \frac{\mathbf{q}_{i} \mathbf{q}_{j}}{4\pi\epsilon_{0} r_{ij}}}_{U_{\text{Coulomb}}} + \underbrace{\sum_{i} \sum_{j>i} 4\epsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6} \right]}_{U_{\text{vdW}}}$$
(2.100)

 q_i and q_j in the Coulomb term are the - previously assigned - partial charges of atom i and j and $r_{i,i}$ defines the distance between those two atoms. The Lennard-Jones parameters ϵ and σ define the depth of the potential as well as the particle-particle distance of zero energy.²⁰¹

All those parameters of the bonded term $\{k_b, r_0, k_{\theta}, \theta_0, k_{phi}, \Phi, k_{ub}, r_{ub}, k_{\omega}, \omega \}$, as well as of the non-bonded term $\{q_i, \epsilon_i \sigma_i\}$ have to be parametrized prior to the actual molecular dynamics simulation. In that parametrization process the molecule is decomposed in different, repeating building blocks ("atom types"), e.g. CH_2 , CH_3 , or functional groups like carboxylate or carbonyl groups, which are assumed to be independent of their actual chemical environment.²⁰² For each atom type within such a building block, the parameters are determined individually, usually from quantum mechanical gas phase calculations of small model compounds. The advantage of that building block approach is that it reduces the effort of force field parametrization on finding suitable parameters for small model compounds. Moreover, once the building blocks are parametrized, they can be broadly applied to a variety of molecules.²⁰² As a result, multiple strands of force fields have emerged addressing different kinds of chemical systems, i.e. the AMBER²⁰³ and CHARMM¹⁹⁹ force field targeted for large molecules like proteins and polymers or, on the other hand, the MM potentials developed by Allinger *et al.*^{204,205} and the OPLS potential²⁰⁶ intended for small molecules and simulations of the condensed phase, respectively.

2.4.1. Newton's and Hamilton's Equations of Motion

After choosing and parametrizing the force field potential, U, the time evolution of the system within that potential can be described using Newtons Second law

$$-\nabla_{i}U\left(\mathbf{r}_{i}\right) = \mathbf{F}\left(\mathbf{r}_{i}\right), = m_{i} \cdot \frac{d^{2}\mathbf{r}_{i}}{dt^{2}}$$

$$(2.101)$$

or, in terms of the particles momentum, $\mathbf{p}_i = m \cdot \dot{\mathbf{r}_i}^{207}$

$$\mathbf{F}\left(\mathbf{r}_{i}, \dot{\mathbf{r}}_{i}\right) = \frac{d\mathbf{p}_{i}}{dt}$$
(2.102)

with U being the potential energy function and m_i and r_i denoting mass and position of particle *i*, respectively. For a system of N particles, its time evolution, or, in other words, its microscopic state at time *t*, is fully characterized by 6N variables. That vector { $\mathbf{r}_i, ..., \mathbf{r}_n, \dot{\mathbf{r}_1}, ..., \dot{\mathbf{r}_n}$ } is an element of the so-called phase space of the system and the form of the potential *U* determines the part of the phase space that is accessible to the system.²⁰⁷

Propagating $\{\mathbf{r}_i, ..., \mathbf{r}_n, \dot{\mathbf{r}_1}, ..., \dot{\mathbf{r}_n}\}$ according to Eq. (2.101) yields the trajectory of the system in that phase space.

A drawback of the Newtons equation of motion is that its validity is restricted to Cartesian Coordinates. Since many problems can be solved more conveniently in non-Cartesian coordinates, Newtons equation of motion is commonly expressed within the Hamiltonian formalism of classical mechanics, which is valid for arbitrary coordinate systems²⁰⁸

$$\mathcal{H}(\mathbf{r}_{1},...,\mathbf{r}_{n},\mathbf{p}_{1},...,\mathbf{p}_{n}) = \sum_{i=1}^{N} \frac{\mathbf{p}_{i}^{2}}{2m} + V(\mathbf{r}_{1},...,\mathbf{r}_{n})$$
(2.103)

Starting from Eq. (2.103) Hamilton's equations of motion can be obtained for a set of generalized coordinates, $\{\mathbf{q}_{\alpha}\}$, and their conjugate momenta, $\{\mathbf{p}_{\alpha}\}$.²⁰⁸

$$\dot{q}_{\alpha} = \frac{\partial \mathcal{H}}{\partial p_{\alpha}} \qquad \qquad \dot{p}_{\alpha} = -\frac{\partial \mathcal{H}}{\partial q_{\alpha}} \qquad (2.104)$$

Solving Hamilton's equation of motion requires the evaluation of 6N first-order differential equations. A commonly used approach for numerically solving Eqn. (2.104) and obtaining the positions and velocities of the particles at time $t + \Delta$ is the velocity Verlet algorithm. In Cartesian coordinates the velocity Verlet is given by:²⁰⁹

$$\mathbf{r}_{i}(t+\Delta) = \mathbf{r}_{i}(t) + \Delta \mathbf{v}_{i}(t) + \frac{\Delta^{2}}{2m_{i}}\mathbf{F}_{i}(t)$$
(2.105)

$$\mathbf{v}_{i}(t+\Delta) = \mathbf{v}_{i}(t) + \frac{\Delta}{2m_{i}}\left(\mathbf{F}_{i}(t) + \mathbf{F}_{i}(t+\Delta)\right)$$
(2.106)

 $\mathbf{v}_i(t)$ and $\mathbf{r}_i(t)$ are the velocity and position of the *i*-th particle at time *t*. While the initial positions of the system are usually provided by experimental crystal structures, determining initial velocities is not so straightforward and needs more consideration. A common approach is to treat the values of the velocities as Gaussian random variable and draw them from a Maxwell-Boltzmann distribution evaluated for the particular temperature, T, of the system.²⁰⁹

The size of the time step, Δ , determines the maximum simulation time that can be achieved with a certain computational effort and it is therefore desirable to make Δ as large as possible. Its maximum size is limited by the time scale on which the fastest process in the system occurs. In biological systems the determining process is usually the stretching of a C-H bond which occurs on the time scale of ≈ 10 fs and, as a rule of thumb, Δ should be chosen ten times smaller, namely 1 fs.²⁰⁹ Since the stretching of a C-H bond is usually not of interest, efficient algorithms like the SHAKE²¹⁰ algorithm have been developed for constraining the C-H bond lengths. With fixed C-H bonds time steps of 2 fs are possible which leads to a two-fold speed-up in computation time.²¹¹

2.4.2. Molecular Dynamics under Periodic Boundary Conditions: Treating Non-bonded Interactions

A major challenge in the simulation of condensed phase systems is the correct treatment of boundary effects. In a finite system, molecules at the boundary experience different forces due to the absence of neighbour molecules than molecules in the center of the simulation box. Due to the small size of the model system, those boundary effects become dominant and bulk properties are not described in an adequate manner.²¹² A convenient approach in preventing boundary effects is to study the system under periodic boundary conditions. Therefore, the simulation box is replicated into all directions which enables interactions of particles at the boundaries with particles

2. Theoretical Description of Organic, Semiconducting Molecules

from the image boxes.²¹² While solving the problem of boundary effects, a major drawback of this approach is that, in theory, all non-bonded interactions have to be computed for an infinite number of particles. Especially evaluation of the long-range electrostatic interactions has proven to be a difficult endeavor. The electrostatic energy in atomic units is given as:²¹³

$$V^{zz} = \frac{1}{2} \sum_{n=0}^{\prime} \left(\sum_{i=1}^{N} \sum_{j=1}^{N} \frac{z_i z_j}{|\mathbf{r}_{i,j} + \mathbf{n}|} \right)$$
(2.107)

With z_i, z_j being the charges of atom *i* and atom *j* respectively. **n** is the multiple of the lattice vector L, i.e. for cubic boxes $\mathbf{n} = \{n_x L, n_y L, n_z L\}$. Note that when summing over **n**, the interaction of atom *i* with itself (i = j) is omitted for $\mathbf{n} = 0$, as indicated by a prime.²¹³ For neutral systems, this sum over **n** is not absolutely convergent which means that the result depends on the order of summation.²¹³ A direct sum approach therefore requires a specification of the order of the summation. Usually the boxes are summed up in roughly spherical layers, starting with $(\mathbf{n} = (0, 0, 0))$ followed by $|\mathbf{n}| = 1$ $(\mathbf{n} = (\pm L, 0, 0), \mathbf{n} = (0, \pm L, 0), \mathbf{n} = (0, 0, \pm L))$ and so forth. For practical applications the infinite sum $\sum_{n=0}^{\prime}$ has to be truncated after a certain cut-off. Unfortunately, convergence of the sum turns out to be slow and a large cut-off is required to obtain results of acceptable accuracy. Since the computational effort scales cubically with the cut-off, the direct sum method has a limited applicability in molecular dynamics simulations.

The problem of slow convergence within the direct sum approach is elegantly by passed by the Ewald summation technique.^{214,215} Within this approach, the long-range potential of the point charges is screened by placing a spatially extended charge distribution of equal magnitude and opposite sign at the location of the point charge. This charge distribution is usually described by a Gaussian distribution of width κ that is centered at the position of the original point charge:²¹³

$$\rho_i^z(\mathbf{r}) = z_i \kappa^3 \frac{\exp -\kappa^2 r^2}{\pi^{\frac{3}{2}}}$$
(2.108)

The advantage of using a Gaussian shaped screening charge is that the Poisson equation can be readily evaluated for a Gaussian charge distribution. Consequently, the added Gaussian distributed charge that screens the δ -distributed point charges has to be compensated by a similar Gaussian distributed charge of opposite sign which



Figure 2.5. – Visualization of the Ewald sum method. (a): δ-distributed charge density of the original point charges as well as the charge density of the screening charges described by a Gaussian distribution. (b): Gaussian-distributed compensating charges that are subtracted from the screening charges. The figure was adapted from Ref. 216

is schematically illustrated in Fig. 2.5. In total, one obtains for the potential:²¹³

$$V^{zz} = \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} \sum_{|n|=0}^{r} z_{i} z_{j} \frac{\operatorname{erfc}\left(\kappa \|\mathbf{r}_{ij} + \mathbf{n}\|\right)}{\|\mathbf{r}_{ij} + \mathbf{n}\|} + \frac{1}{2} \frac{1}{\pi L^{3}} \sum_{i=1}^{N} \sum_{j=1}^{N} \sum_{\mathbf{k}\neq0}^{N} z_{i} z_{j} \left(\frac{4\pi^{2}}{k^{3}}\right) \exp\left(\frac{-k^{2}}{4\kappa^{2}}\right) \cos\left(\mathbf{k}\mathbf{r}_{ij}\right) - \left(\frac{\kappa}{\pi^{\frac{1}{2}}}\right) \sum_{i=1}^{N} z_{i}^{2} + \left(\frac{2\pi}{3L^{3}}\right) \left\|\sum_{i=1}^{N} z_{i}\mathbf{r}_{i}\right\|^{2}$$
(2.109)

The first term in Eq. 2.109 captures the electrostatic potential of the original point charge and the screening charge which is now of short range. The second term, which sums over the reciprocal vectors $\mathbf{k} = 2\pi \mathbf{n}/L^2$, converges fast in reciprocal space.²¹³ The third term is called the "self term" and subtracts the - physically meaningless - interaction of the cancelling distribution at \mathbf{r}_i with itself.²¹³ The fourth term is commonly denoted as dipole term, or surface term.²¹³ When summing over the image boxes in roughly spherical layers, the relative permittivity of the medium surrounding the sphere must be considered,^{217,218} i.e. a sphere surrounded by vacuum will experience a dipolar layer that is absent for a sphere surrounded by a conductor. Subtraction of the dipole term, as in Eq. 2.109, allows to employ the Ewald sum method for a sphere surrounded by vacuum.²¹³

2. Theoretical Description of Organic, Semiconducting Molecules

Compared to the Ewald sum approach for electrostatic interactions, handling of van der Waals interactions within the framework of periodic boundary conditions is far less challenging. Due to their short range, van der Waals interactions get usually truncated after a specified cut-off. In the straightforward approach the van der Waals interactions are simply set to zero for particle distances larger than the cut-off. Such an abrupt truncation of the van der Waals interactions leads to discontinuities in the energy and its derivatives which can result in problems during structural minimization.²¹⁹

A more sophisticated approach relies on a smoothing function that is switched on after a specified distance, r_{switch} , and leads to a continuous decay of the van der Waals interactions to zero until the specified cut-off, $r_{cut-off}$. A common type of a switching function^{*} is given by:²¹⁹

$$S(r) = \begin{cases} 0, & r > r_{cut-off} \\ \frac{\left(r^2 - r_{cut-off}^2\right)^2 \cdot \left(2r^2 + r_{cut-off}^2 - 3r_{switch}^2\right)}{\left(r_{cut-off}^2 - r_{switch}^2\right)^3}, & r_{cut-off} \ge r > r_{switch} \\ 1, & r \le r_{switch} \end{cases}$$
(2.110)

The advantage of such an expression for the switching function is that its first derivative is a continuous function in the whole range $r \ge r_{cut-off}$ which avoids the aforementioned problems in structural relaxation.²¹⁹

2.4.3. Modelling Systems under Biological Conditions: Temperature and Pressure Control

A usual goal in molecular dynamics simulations is to derive macroscopic properties by studying the time evolution of a microscopic system under certain macroscopic conditions (total energy, pressure, temperature). Such a collection of microscopic states that possess a mutual set of macroscopic properties is called a statistical ensemble.²²⁰ Common statistical ensembles are the microcanonical ensemble (*NVE*), the canonical ensemble (*NVT*), and the isothermic-isobaric ensemble (*NPT*). In the *NVE* ensemble, a collection of microscopic states of constant particle number, *N*, constant volume, *V*, and constant total energy, *E*, is investigated. In a similar manner, the *NVT* and *NPT* ensembles comprise either a constant particle number, N, a constant volume, V, and a constant temperature, *T*, or a constant particle number, *N*, constant pressure, *P* and constant temperature, T^{\dagger} . The *NPT* ensemble is of particular interest in molecular modelling, since chemical and biological processes usually take place under constant pressure and temperature. Investigation of chemically and biologically relevant systems by molecular dynamics therefore requires suitable methods for temperature and pressure control.

The temperature of the system can be linked to the time average of its kinetic

^{*}A similar type of switching function is also implemented in NAMD, the software package employed in this work

[†]A comprehensive description of those three mentioned ensembles is provided in Ref. 221.

 $energy^{222}$

$$\langle \mathcal{K}_{NVT} \rangle = \frac{3}{2} N k_b T \tag{2.111}$$

Consequently, an intuitive approach to control the temperature would be to scale the velocities of the system by a factor λ . For a N-particle system of temperature T(t) at time t, the difference in temperature due to the scaling of the velocities is expressed as:²²²

$$\Delta T = T_{new} - T(t) = \frac{1}{2} \sum_{i=1}^{N} \frac{2}{3} \frac{m_i \left(\lambda_i v_i\right)^2}{Nk_b} - \frac{1}{2} \sum_{i=1}^{N} \frac{2}{3} \frac{m_i \left(v_i\right)^2}{Nk_b}$$
(2.112)

$$\Delta T = \left(\lambda^2 - 1\right) T(t) \tag{2.113}$$

$$\lambda = \sqrt{\frac{T_{new}}{T(t)}} \tag{2.114}$$

Note that in Eq. (2.112) relation $\langle \mathcal{K} \rangle = 1/2m_i v_i^2$ has been used for the kinetic energy with m_i as the mass of particle *i*. In order to keep the system at a required temperature, T_{req} , the velocities are multiplied at each time step by a factor $\lambda = \sqrt{T_{req}/T(t)}$. A drawback of the simple velocity rescaling method is that it does not capture the fluctuations in kinetic energy of the system and the ensemble generated is no NVTensemble.²²² A more subtle approach is offered by the Berendsen thermostat.²²³ The underlying idea is to couple the system to an external bath held at the required temperature that exchanges thermal energy with the system at a rate τ . Thereby the temperature is slowly adjusted to the desired temperature and the rescaling of the velocities to match the desired temperature does occur continuously within a certain time range. The scaling parameter, λ can be expressed as:²²²

$$\lambda^2 = 1 + \frac{\Delta t}{\tau} \left(\frac{T_{bath}}{T(t)} \right) \tag{2.115}$$

The appropriate value of the parameter τ depends on the size of the time step, Δt , of the simulation. Like the simple velocity rescaling, the Berendsen thermostat does not reproduce the correct thermodynamic properties of a NVT ensemble.²²² Moreover, in both methods, internal conversion and exchange of thermal energy between components of the system is affected.²²² This leads to problems like the "hot solvent cold solute" phenomenon,²²⁴ or even more peculiar artefacts like the "flying ice cube".²²⁵ In the former case the system is formally at the specified temperature but the solvent molecules are at significantly higher temperature than the solute.²²⁴ In the latter case, the equipartition theorem is violated and the thermal energy is accumulated in translational and rotational degrees of freedom instead of being distributed equally over all degrees of freedom. As a result, the system freezes into one configuration that propagates with unnaturally high speed through space.²²⁵

A complementary approach to control temperature is the Langevin thermostat which employs a stochastic coupling of the system to a heat bath via dissipative

2. Theoretical Description of Organic, Semiconducting Molecules

and random forces.²²⁶ In detail, transfer of thermal energy between the systems components is accomplished by randomly occurring collisions between solvent and solute that change the kinetic energy and hence the temperature of the latter. Stochastic coupling is introduced by adding two force terms to Newtons equation of motion which leads to the generic Langevin equation that describes the motion of a solvated particle^{226,227}

$$m_i \dot{v}_i = \mathbf{f}(\mathbf{r}_i) - \gamma v_i - \sqrt{\frac{2\gamma k_B T}{m_i}} R(t)$$
(2.116)

 m_i and v_i denote mass and velocity of the particle. Additionally to the force $\mathbf{f}(\mathbf{r}_i)$ the particle experiences a dissipative force, γv_i , due to the viscosity of the surrounding solvent whose magnitude depends on the friction coefficient, γ . The last term accounts for the fluctuations in the particles kinetic energy due to random collisions. The stochastic nature of these collisions is described by a Gaussian process of mean zero, R(t), whose variance is determined by the friction coefficient and the desired temperature.²²⁸

Compared to temperature control, the process of controlling the pressure during a molecular dynamics simulations is more challenging and requires a change of the volume of the simulation cell during the simulation. A well-established approach for pressure control is the extended-system method which includes the pressure as as an additional degree of freedom in form of piston of mass W.²²⁸ The volume of the simulation cell is then adjusted during the simulation run by that piston in order to maintain the required pressure, P_{ext} . The equations of motion within the extended system approach are:²²⁸

$$\dot{\mathbf{r}}_i = \frac{\mathbf{p}_i}{m_i} + \frac{1}{3} \frac{\dot{V}}{V} \mathbf{r}_i \tag{2.117}$$

$$\dot{\mathbf{p}}_i = \mathbf{f}_i - \frac{1}{3} \frac{\dot{V}}{V} \mathbf{p}_i \tag{2.118}$$

$$\ddot{V} = \frac{1}{W} \left(P(t) - P_{ext} \right)$$
 (2.119)

Here, $\mathbf{p}_i, \mathbf{r}_i$, and m_i are momentum, position and mass of the *i*-th particle, V is the volume of the cell, P(t) and P_{ext} are the pressure at time t and the required pressure, respectively. Solving the equations of motion within the extended system approach generates trajectories in the isobaric-isoenthalpic (NPH). A drawback of the extended system approach is that the degree of coupling strongly depends on the mass W of the piston and a poor choice of W might affect the dynamical properties of the system.²²⁸ A large mass leads to a slow, and potentially inefficient, adjustment of the box volume to the external pressure while a small mass results in rapid and high oscillation of the box volume. In order to treat these oscillations, a partial damping of the piston motion can be achieved by introducing random collisions described by the Langevin equation.²²⁸ The relation between the second time derivative of the volume and the

piston mass is then modified to^{228}

$$\ddot{V} = \frac{1}{W} \left(P(t) - P_{ext} \right) - \gamma \dot{V} + R(t)$$
(2.120)

with γ being the collision frequency and R(t) a Gaussian process describing the random force acting on the piston. R(t) has a mean of zero, while the variance is given by²²⁸

$$\langle R(0)R(t)\rangle = \frac{2\gamma k_b T\delta(t)}{W}$$
(2.121)

 k_b is the Boltzmann constant and T the temperature of the system. The additional term, R(t), therefore couples the piston degree of freedom to a heat bath which should formally produce NPT ensembles. However, due to the weak coupling, additional temperature control methods, as described above, have to be employed in order to guarantee an efficient exchange of heat.

3. Structural Heterogeneity in AnPixJg2 Wild Type and Mutant

3.1. Introduction

Numerous biological processes in living organisms depend on the illumination conditions of the environment. Response to light is regulated by photosensory proteins, large biomolecules that contain a chromophore which is able to undergo reversible photoisomerization between two states with different structure and spectroscopic properties.²²⁹ These two states are commonly denoted as dark stable state and photoproduct state. Most of the time, the chromophore adopts the conformation of the biologically inactive dark stable state. Excitation of light of a wavelength that matches the absorption maximum of the dark stable state leads to a fast photoisomerization of the chromophore and relaxation in the biologically active photoproduct state, followed by a biological reaction.²²⁹ Reconversion from the metastable photoproduct state to the dark stable state occurs either fast by excitation with light of appropriate wavelength (i.e. the wavelength that matches the absorption maximum of the dark conversion.²³⁰

The photosensory proteins to be discovered first were phytochromes (Phy), which typically respond to red/far-red light.²³¹ Their multiple functional subclasses can be found in plants,²³² bacteria,^{233,234} cyanobacteria,^{235–237} algal species, and fungi.²³⁸

A recently discovered class of photosensory proteins are the cyanobacteriochromes (CBCRs) which mediate light-dependent processes like phototaxis^{239–241} and chromatic acclimation,^{7,242} and, so far, have been found exclusively in cyanobacteria.^{6,231} Regarding their general photochemistry, cyanobacteriochromes and phytochromes show striking similarities. Both classes employ structurally similar bilin-derived tetrapyrroles; an overview over the different chromophores is given in Fig. 3.1. Covalent linkage of the chromophore to the protein occurs via a thioether bridge between a conserved cysteine residue and the side chain of the A-ring of the chromophore.^{8,231} The reversible photoconversion mechanism between the dark stable state and the metastable photoproduct state involves Z to E isomerization of the C15-C16 double bond and *vice versa* accompanied by a rotation of the bilin *D*-ring of the chromophore.^{6,40,243-245}

Despite the structural similarities of their chromophores, cyanobacteriochromes and phytochromes strongly differ in their absorption properties. Canonical phytochromes are mostly restricted to the red and far-red part of the solar spectrum, whereas CB-CRs, exhibit an outstanding spectral variety of photocycles covering the whole range of the visible spectrum.^{6,9,10,33,246} The plethora of different photocycles is remarkable considering the much simpler architecture of CBCRs compared to phytochromes: while

3. Structural Heterogeneity in AnPixJg2 Wild Type and Mutant



Figure 3.1. – Structure of different chromophores employed in phytochromes and cyanobacteriochromes. The figure was taken from Ref.231

most phytochromes possess a N-terminal photosensory core which contains the highly conserved PAS (Per-ARNT-Sim)-GAF (cGMP phosphodiesterase/adenylatecyclase/ FhlA)-PHY (phytochrome-specific) domain module,^{231,243,244,247} CBCRs require only one specific, structural domain, the GAF domain.^{6,8} Depending on their absorption properties and their color tuning mechanism, CBCRs can be divided into several subclasses: red/green CBCRs, green/red CBCRs, and the two-Cys CBCRs which can be further classified into the insert-Cys CBCRs and DXCF CBCRs. Red/green CBCRs exhibit a dark stable state with an absorption maximum in the red region of the visible spectrum and a photoproduct state with an absorption maximum in the green region of the visible spectrum.^{9,33} For green/red CBCRs, on the other hand, the absorption properties are reverted, i.e. they possess a dark stable state absorbing light in the green region and a photoproduct state absorbing light in the red region of the visible spectrum.^{7,248} In two-Cys CBCRs reversible formation and cleavage of a second thioether bridge between a cysteine residue and the chromophore mediates the spectral shifts between the dark stable state and the photoproduct state. $^{10,33,47,249-253}$ In DXCF CBCRs, a subgroup of the dual-Cys CBCRs, the second cysteine residue that transiently binds to the chromophore is located within a conserved DXCF motiv of the GAF domain that is unique to this CBCRs subgroup,^{12,47,249,250} while in insert-Cys CBCRs the second Cys residue is part of an inserted sequence.^{10,254} Regarding the chromophore, canonical phytochromes as well as red/green and green/red CBCRs employ both phycocyanobilin (PCB) (depicted in Fig. 3.1). The DXCF CBCRs from the group of the dual-Cys CBCRs, on the other hand, initially bind PCB but reduce it

to PVB which shifts the absorption maximum of their photoproduct state to shorter wave lengths into the spectral regions between teal and orange.^{47,245,250}

Apart from their different absorption properties, phytochromes and the individual subclasses of CBCRs also vary in their photoisomerization quantum yield. Phytochromes have been found to show unexpectedly low $Pr \rightarrow Pfr$ photoisomerization quantum yields $(10\%-15\%)^{50,51}$ compared to photoreceptors like rhodopsin $(65\%)^{255}$ and even smaller fluorescence quantum yields $(\leq 3\%)$.^{256–258} The photoisomerization quantum yield of CBCRs depends strongly on the subclasses: the CBCR NpR6012g4, for instance, which is a member of the red/green CBCRs subgroup shows comparably high quantum yields of 40\%-50\% for the forward and reverse photocycle,^{36,38} while the quantum efficiency of the green/red CBCR RcaE from *Fremyella diplosiphon* is with $\approx 10\%$ significantly lower.³⁷ Both representatives employ the same chromophore, PCB, which excludes a simple relation between the employed chromophore and the quantum yield.

Interestingly, the dark stable state of both, NpR6012g4 and RcaE, as well as the photoproduct state of NpR6012g4 have been reported as heterogeneous,^{35–38} i.e. more than one conformation of the chromophore can be observed in each state. Structural heterogeneity of the chromophore in the dark stable state or in the photoproduct state is not restricted to red/green or green/red CBCRs but includes some also blue/green absorbing members of the DXCF family^{47,48} and even representatives of plant^{259,260} and bacterial^{40–46} phytochromes. The frequent occurrence of conformational sub populations in phytochromes and CBCRs implies that structural heterogeneity might be an innate feature of such photosensory proteins with implications on their spectral tuning properties.

Unfortunately, crystallographic methods are often insensitive to the subtle structural differences between such sub populations which makes the use of other techniques like nuclear magnetic resonance (NMR) spectroscopy, Raman spectroscopy, or molecular dynamics (MD) simulations necessary to unravel the whole structural diversity of chromophores in photosensors.

In this chapter, structure and dynamics of AnPixJg2 in solution are characterized on an atomistic level based on an extensive molecular dynamics study of nearly 1 µs. AnPixJg2, the second GAF domain of AnPixJ, is a representative of the red/green cyanobacteriochromes with an absorption maximum at 648 nm in the red absorbing Pr and 543 nm in the green absorbing Pg state, respectively.⁹

X-ray experimental studies³³ have confirmed a 5Z- 10Z-15Z conformation of PCB in the Pr form of AnPixJg2, which is similar to that of PCB in canonical and cyanobacterial phytochromes, like Cph1 (see Fig. 3.1). However, the step-wise photoconversion mechanism comprising a 15Z to 15E isomerization and three intermediates leads to a different photoproduct.^{34,261}

Whereas, during the years, phytochromes like Cph1 have been subject to extended experimental^{49,262–265} and theoretical studies,^{266–268} fewer results exist for AnPixJg2, mainly due to its later discovery. So far, only the crystal structure of the Pr state of AnPixJg2 is known,³³ but no crystallographic data of the P_g state is available to date that could relate the absorption properties of PCB in the P_g state to its structural



Figure 3.2. – a: Quaternary structure of AnPixJg2 in the crystallographic unit cell. Two molecules are arranged in a dimeric manner, in which the C-terminal α-helix of each molecule interacts with the N-terminal a-helix of the other molecule.
b: Overall structure of the second GAF domain of AnPixJ taken from the protein data bank (PDB ID 3W2Z)³³

properties and interactions with the chromophore binding pocket.^{*} In the attempt to explain the large shift in the absorption maximum in the photoproduct state, several hypotheses have been discussed. An early hypothesis aimed to explain the hypsochromic shift in the P_q state by a conversion of PCB into PVB, i. e. a reduction of the methine bridge between the A-ring and the B-ring which interrupts the conjugated π -system (see Fig. 3.1). This hypothesis has been ruled out out by results from denaturation analysis.⁹ Moreover, deprotonation of the chromophore during photoconversion has been discussed but could be excluded based on Resonance Raman Spectrocopy studies that reveal the same protonation state of the chromophore in the dark stable state and the photoproduct state.²⁶¹ The absence of any chemical modification implies that the shift of the absorption maximum in the photoproduct state is controlled by interactions between chromophore and its environment. One assumption is that the blue shifted absorption maximum is induced by an increased solvation in the chromophore binding pocket in the P_q state. The access of water into the binding pocket in both states is steered by a bulky tryptophan residue, TRP289. In the P_r state TRP289 is kept in place by interactions with the D-ring of PCB and therefore shields the chromophore binding pocket from the solvent molecules.²⁶¹ Isomerization of the C-D methine bridge and rotation of the D-ring upon photoconversion, however, weaken the interaction between Trp289 and PCB, and water penetrates the binding pocket.

On the other hand, mutagenesis studies²⁴⁶ on the structurally related red-green cyanobacteriochrome NpR3784 show that the absorption spectrum of the photopro-

^{*}Currently, the crystal structure of the parent ground state and the photoproduct state of Slr1393, a red/green CBCR that is closely related to AnPixJg2 has been resolved.²⁶⁹ Results from QM/MM calculations²⁷⁰ explain the color tuning mechanism by an effective shortening of the conjugation length of the chromophore due to a strongly twisted geometry in the P_g state. The strong structural resemblance hints towards a similar mechanism for AnPixJg2.

duct is stable against mutation of that Trp residue. Instead they claim that in the P_g state the *D*-ring is trapped in a strongly twisted geometry between two conserved phenylalanine residues. That strongly twisted geometry distorts the conjugation of the chromophore and shifts the absorption maximum to higher frequencies.²⁴⁶ Since NpR3784 and AnPixJ are structurally closely related, a similar mechanism could be responsible for shifted absorption maximum in the P_g state of AnPixJ.

Part I of this chapter is based on preliminary, shorter MD studies that indicated the existence of more than one sub-state²⁷¹ and therefore motivated this extensive MD study of almost 1 µs. In the following, the P_r state of AnPixJg2 is investigated by force field based MD. Starting from the crystal structure, the effects of solvation and temperature on the structure of protein and chromophore are characterized. Subsequently, the interaction between PCB, surrounding amino acids, and solvent are discussed.

The results from molecular dynamics simulations reveal the existence of two structural sub-states of the Pr state of AnPixJg2 from which only one was visible in X-ray data. These two sub-states are stable on time scales of 10 ns - 500 ns and can be distinguished by their conformation of the D-ring of the chromophore. The finding of structural heterogeneity in the Pr form of AnPixJg2 is supported by data from ¹³C MAS NMR experiments.²⁷² In both states, the chromophore is stabilized by three persistent hydrogen bond networks involving adjacent amino acids. The long simulation time of $\approx 1 \, \mu s$ allows to observe reversible interconversion between these two sub-states that can be linked to a change in the solvation pattern of the chromophore binding pocket. Furthermore, a strong structural reorganization of the β 3 sheet can be observed that occurs on a time scale of ≈ 600 ns. The duration of this structural reorganisation demonstrates the necessity for long MD simulations (on a time scale of several 100 ns to a µs) in order to sufficiently capture the rearrangement of large areas of the protein. Moreover, the dynamic behavior of a tryptophan residue, TRP289, is analyzed, which was believed to shield the chromophore binding pocket from solvent.²⁶¹ Conversely, this tryptophan residue undergoes interconversion between three distinct conformations which is in contradiction to its role as gate keeper the hydration model.²⁶¹

After a characterization of the overall structure and dynamics of the parent ground state (P_r state) of AnPixJg2, the influence of individual, conserved amino acids in the chromophore binding pocket on the conformation of the chromophore is elucidated. An elegant approach to isolate the contributions of individual amino acid residues on the properties of the chromophore lies in the targeted mutation of the respective amino acid against a structurally and chemically different one and a comparison of the absorption spectrum of wild type and mutant. In order to understand the influence of the mutation on an atomistic level, mutagenesis experiments are usually combined with computational *in-silico* mutagenesis studies. In Part II of Chapter 3.2 the effect of specific conserved amino acids on the structural and dynamical properties is elucidated by *in-silico* mutagenesis of the tyrosine residue TYR302 and a subsequent molecular dynamics simulation of 1.2 µs. TYR302 is highly conserved among CBCRs³³ indicating its relevance in the function of AnPixJg2 as photosensory protein. As visible in Fig. 3.3 **a**, tyrosine possesses a polar phenol side chain which can act as hydrogen bond



Figure 3.3. – a: Chromophore-protein interactions in the AnPixJg2 wild type. The tyrosine residue TYR302 and the C-ring propionate side chain of PCB are engaged in stable hydrogen bond interactions. b: The TYR302PHE mutant. Mutation of the TYR302 residue to phenylalanine, which which is of equal size but cannot undergo hydrogen bonding with its side chain, interrupts all interactions with PCB.

donor and acceptor. In AnPixJg2 the phenol group of TYR302 is involved in stable hydrogen bonding with the propionate group of the *C*-ring of PCB which suggests a potential role of TYR302 in stabilizing the conformation of the chromophore in the binding pocket. By mutating TYR302 against a phenylalanine residue (Fig. 3.3) which comprises a non polar benzyl side chain this, potentially relevant, hydrogen bond interaction with the chromophore is interrupted.

The results from molecular dynamics simulation reveal that mutation of TYR302 against PHE leads to the formation of a third sub-state of PCB in which both, the A-ring and the D-ring are oriented towards the β -facial side of the chromophore. In this new orientation the A-ring is able to participate in persistent hydrogen bonds with adjacent amino acids that stabilize sub-state III and most likely prevent interconversion back into sub-state I and II. Possible events triggering the structural transition into sub-state III are discussed.

3.2. Part I: Structural Heterogeneity in a Parent Ground-state Structure of AnPixJg2 Revealed by Theory and Spectroscopy

In this section the red absorbing, dark stable state (Pr state) of the second GAF domain of the cyanobacteriochrome AnPixJ (AnPixJg2) is investigated by a molecular dynamics simulation of almost 1 µs duration. The presented results reveal two distinct conformational isoforms of the chromophore, from which only one was known from crystallographic experiments. The interconversion between both isoforms is accompanied by alterations in the hydrogen bond pattern between the chromophore and the protein and the solvation structure of the chromophore binding pocket. The existence of sub-states in the Pr form of AnPixJg2 is supported by results from experimental ¹³C MAS NMR spectroscopy. These finding are consistent with the observation of structural heterogeneity in other cyanobacteriochromes and phytochromes.

Adapted from the author's work²⁷² by permission of the PCCP Owner Societies.

3.2.1. Structure of the Second GAF Domain of AnPixJ

Fig. 3.2 depicts the quaternary structure of the Pr form of AnPixJg2 in the crystallographic unit cell (a), as well as a detailed view on the secondary structural elements of a single AnPixJg2 molecule (b). The β 3-sheet comprises the residues ILE285 -TRP289 and shields the chromophore binding pocket from the surrounding solvent. The residues PRO215-ARG231 and GLU364-THR388 belong to the N-terminal α 1helix and the C-terminal α 5-helix, respectively. It can be seen from Fig. 3.2, that in the crystallographic unit cell two molecules of AnPixJg2 adopt a dimeric arrangement in which the C-terminal α -helix is stabilized by interactions with α helical bundles from the other molecule.³³ Since in this work only one AnPixJg2 molecule is simulated, these stabilizing interactions are absent in the simulation leading to unrealistic structural fluctuations of these segments.²⁶⁸ Therefore the residues PRO215-ARG231 and GLU364-THR388 were excluded from any structural analysis.

Fig. 3.4 **a** shows the root mean square deviations (rmsd) of the residues VAL232-TRP363 of the protein backbone of AnPixJg2 aligned to the backbone atoms of the crystal structure. After ≈ 75 ns the protein backbone rmsd reaches a plateau region with an average value of 2 Å. Therefore, the first 100 ns have been considered as part of the equilibration process and the trajectory from 100 ns – 940 ns has been used for statistical analysis. Notable structural differences between the crystal structure and the simulated structure are found for the residues ILE285-GLU290 which are part of the β 3-sheet that shields the chromophore binding pocket from the surrounding solvent. Fig. 3.4 **b**, depicts the backbone rmsd of the residues ILE285-GLU290, which form that β 3-sheet. It shows that after 50 ns a strong increase in the rmsd towards a plateau region occurs. During the following 600 ns the rmsd strongly fluctuates around this plateau region, denoted as plateau region 1. After a simulation time of about 650 ns, however, it jumps back to lower values. At this lower plateau region, which is denoted as plateau region 2 in Fig. 3.4 **b**, the rmsd remains stable for the

3. Structural Heterogeneity in AnPixJg2 Wild Type and Mutant



Figure 3.4. – **a** root mean square deviations (rmsd) of the protein backbone atoms. **b**: rmsd of the backbone atoms of the β -sheet that shields the chromophore binding pocket from the solvent (ILE285-GLU291). **c**: Structural snapshots extracted at the two plateau regions of the β -sheet rmsd.

rest of the simulation time as indicated by a smaller variance of the individual values.

To visualize the structural changes that correspond to the two plateau regions, simulated structures at simulation times corresponding to that different plateau regions, namely at 150 ns and 700 ns have been extracted and superimposed on the crystal structure. The resulting structural images are presented in Fig. 3.4 c with the crystal structure being depicted in gray and the simulated structures at 150 ns and 700 ns in green and blue, respectively. In the crystal structure the β 3-sheet is positioned slightly below the pyrrole ring plane of the chromophore and close to the β^2 sheet. As visible from the simulated structure at $150 \,\mathrm{ns}$, depicted in green in Fig. 3.4 c, plateau region 1 in the rmsd corresponds to a structure in which β 3-sheet is arranged above the pyrrole ring plane and in much closer proximity to the α 4-helix atop of the chromophore binding pocket. This leads to an enlarged gap between the β 2-sheet and the β 3-sheet. A simulated structure of the protein after 700 ns (blue) corresponding to the stable plateau region 2 in the rmsd is depicted in Fig. 3.4 c. The β 3-sheet in this extracted structure is now again arranged in a closer proximity to the β^2 -sheet and thus more similar to the crystal structure but exhibits a remarkably higher degree of disorder.



Figure 3.5. – Different conformational arrangements of TRP289 during the simulation in correlation to distinct regions in the rmsd of the TRP289 side chain. a: rmsd of the heavy atoms of TRP289. Distinct plateau regions can be observed that are highlighted by green, red, and yellow rectangles. b: Structural conformation of TRP289 in the crystal structure (green) and extracted after 250 ns (red) and 400 ns (dark yellow). The structural snapshots correspond to the rmsd plateau regions highlighted in green, red, and yellow, respectively.

The overall higher disorder in this region includes also the side chain of the tryptophan residue TRP289, which was supposed to shield the chromophore binding pocket according to the hydration $model^{261}$ (see Chapter 3.1). The heavy atom rmsd of TRP289, depicted in Fig. 3.5 a shows several jumps between plateau regions. Snapshots of the different conformations of the indole side chain that correspond to the three different regions in the rmsd are presented in Fig. 3.5 b. The first region, highlighted by a green rectangle, corresponds to the conformation of TRP289 in the crystal structure with the indole nitrogen pointing towards the A-ring of PCB. This conformation is adopted during the first 240 ns of the simulation run as well as during the time from approx. 266 ns - 291 ns. The structural arrangement of TR289 that corresponds to the second plateau region in the rmsd, which is highlighted by a red rectangle, is represented by the red colored TRP289 molecule in Fig. 3.5 b. In this conformation, the TRP289 side chain adopts an out-of-plane arrangement with the indole nitrogen pointing towards the α 4-helix. The third region in the rmsd is highlighted by a vellow rectangle and correlates with a 180° rotation of the indole ring of TRP289 as represented by the vellow colored structure in Fig. 3.4 b extracted at 400 ns. This highly dynamical behavior of TRP289 is in contradiction to its function as a "gatekeeper"²⁶¹ that shields the chromophore binding pocket from the solvent as proposed in the hydration model.

3.2.1.1. Conformation and Flexibility of the Chromophore

In the following, the conformation and the flexibility of the chromophore during the simulation will be discussed. The rmsd of the non-hydrogen atoms of the individual rings of PCB is shown in Fig. 3.6. While the rmsd for the A-ring and the B-ring are stable, the rmsd of the C-ring and, most notably, of the D-ring exhibit jumps between several plateau regions. In total, four different plateau regions can be distinguished out of which region 1 and 3 in the D-ring rmsd show an average height of 1.3 Å and 1.2 Å, respectively, while region 2 and 4 in the D-ring rmsd exhibit both average values of 0.33 Å.



Figure 3.6. – rmsd of the non-hydrogen atoms of all four rings of PCB. While the *A*-ring and the *B*-ring exhibit stable rmsd, for the *C*-ring and, most notably, for the *D*-ring two different plateau regions can be distinguished.

Fig. 3.7 shows the structure of the chromophore averaged over the simulation times that correspond to the different plateau regions of the rmsd, namely from $100 \,\mathrm{ns}$ – 145 ns (**a**), 150 ns-700 ns (**b**), 730 ns-890 ns (**a**), 897 ns-910 ns (**b**), and 915 ns-940 ns(a). The structures presented in Fig. 3.7 a thus correlate with the lower plateau region, while the structures depicted in Fig. 3.7 b can be assigned to the higher plateau region. The averaged structures reveal that the jumps in the rmsd correlate with the position of the D-ring with respect to the pyrrole ring plane. In the crystal structure of PCB as well as in the conformation adopted during the first 145 ns of the simulation time, the D-ring is arranged above the pyrrole ring plane on the α -face of the chromophore. However, from 150 ns to approx. 710 ns it adopts a β -facial disposition, i.e. is arranged below the pyrrole ring plane. After 710 ns, the D-ring performs a backward rotation above the pyrrole ring plane for roughly 180 ns, after which it switches back to its downward oriented position on the β -face of the chromophore for a time period of 13 ns. From approx. 910 ns until the end of the simulation time the D-ring re-adopts its α -facial disposition. This leads to the assumption that at least two conformational substates of PCB in the Pr state exist which can be distinguished by their conformation of the *D*-ring. These sub-states will further be denoted as sub-state I and sub-state II. Hence, Fig. 3.7 **a** shows the chromophore in its sub-state I, which was detected by X-ray experiments. In sub-state I the D-ring adopts an α -facial disposition whereas in sub-state II it is rotated below the pyrrole ring plane on the β -face of the chromophore (see Fig. 3.7 b).

In order to characterize both states separately, the evaluated trajectory was divided into several parts according to the sub-state of the chromophore. For sub-state I, the trajectory from 100 ns to 145 ns, 730 ns to 890 ns, and 915 ns to 940 ns was evaluated. For sub-state II the trajectory from 150 ns to 700 ns and 897 ns to 910 ns was analyzed.



Figure 3.7. – Average structure of the tetrapyrrole skeleton of PCB in its different conformational states. From 100 ns - 145 ns, 730 ns - 890 ns and 915 ns - 940 ns the *D*-ring adopts an α -facial disposition (sub-state I), while from 150 ns - 700 ns and 897 ns - 910 ns it is oriented towards the β -face of the chromophore (sub-state II). The average structure comprises only the pyrrole rings (dark gray), while the side chains (lighter gray) have been excluded from analysis and are only depicted in an exemplary manner.

In order to quantify the downward rotation of the *D*-ring in sub-state II, the dihedral angle distributions of the single and double bond of all methine bridges have been evaluated separately for both conformational states. Fig. 3.8 **a** depicts a timeline of the dihedral angle around the methine bridge single bonds between the *A*-ring and the *B*-ring (A-B single bond) and between the *C*-ring and the *D*-ring (C-D single bond). The C-D single bond dihedral comprises the atoms C13-C14-C15-C16 whereas the A-B single bond dihedral includes the atoms C4-C5-C6-N22 (see Fig. 3.1). For a better visualization, a sketch of the *A-B* and *C-D* single bond dihedral angles is also presented in Fig. 3.8 **c**.

Unlike the A-B single bond dihedral which fluctuates around the same value during the whole simulation time, the C-D single bond dihedral jumps between two distinct average values reflecting the two different conformations of the *D*-ring during the simulation time. This behavior is confirmed by the histograms of the dihedral values (Fig. 3.8 b), which show one Gaussian shaped distribution for the A-B single bond dihedral but two clearly separated distributions for the C-D single bond dihedral. The

value of the crystal structure's dihedral, the mean value, and the standard deviation of the dihedral angles of sub-state I and II are listed in Tab. 3.1.



Figure 3.8. – Distributions of the dihedral angles around the A-B and C-D single bond. a: The timeline shows that the C-D single bond torsion angle jumps between two distinct values, whereas the A-B single bond dihedral fluctuates around a stable value. b: The histogram of the distribution of the dihedral around the C-D single bond shows two separated maxima clearly indicating two distinct values for the dihedral in sub-state I and sub-state II.c: Sketch of the A-B and C-D-single bonds.

With 22° for sub-state I and 18° for sub-state II, the mean value of the A-B single bond dihedral of the simulated structure is shifted towards higher values compared to that of the crystal structure (11°). This implies a slightly more upward rotated A-ring in both structural sub-states leading to a decreased planarity of the A-ring with respect to the inner ring plane. Sub-state I and sub-state II show the same flexibility of the A-B single bond dihedral as well as the A-B double bond dihedral which can be derived from similar standard deviations of 11° and 10° (A-B single bond dihedral), and 12° (A-B double bond dihedral)respectively.

In comparison to the crystal structure, the C-D double bond dihedral of sub-state I is reduced (13° versus 26° in the crystal structure) indicating a higher planarity of the D-ring with regard to the inner ring plane. In sub-state II, the average value of the C-D double bond dihedral is -5° and thus slightly less twisted than in sub-state I. With a mean value of -38° , the C-D single bond dihedral of sub-state II strongly differs from those of sub-state I and the crystal structure (44° and 36°, respectively) confirming a 82° rotation of the D-ring below the pyrrole ring plane in sub-state II. This clearly shows that the downward orientation of the D-ring in sub-state II arises mainly from a rotation around the C-D single bond and not from a twisting around the double bond of the C-D methine bridge.

To gain a detailed picture of the overall flexibility of the chromophore in both structural sub-states, the root mean squared fluctuations (rmsf) of the chromophore's

Table 3.1. – Average value and standard deviations of the dihedral angles (in degree) of the A-B and C-D methine bridge of the crystal structure (c.s.) and both sub-states of the simulated structure. The A-B single and double bond dihedral angles include the atoms C4=C5-C6-N22 and N21-C4=C5-C6 respectively, whereas the C-D single and double bond dihedral angles comprise the atoms C13-C14-C15-C16 and C14-C15=C16-N24 respectively

dihedral	c. s.	sub-state I		sub-state II	
		mean	sigma	mean	sigma
A-B single bond	11	22	11	18	10
A-B double bond	0.4	1	12	-4	12
C-D single bond	36	44	12	-38	10
C-D double bond	26	13	10	-5	11

heavy atoms were calculated separately for sub-state I and sub-state II and are presented in Fig. 3.9 **a** and **b**, respectively. In both sub-states the *B*-ring and the methine bridge between *A*-ring and *B*-ring (C5) as well as between *B*-ring and *C*-ring (C10) possess the smallest rmsf values and thus form the most rigid part of the chromophore. Furthermore, in both structural sub-states, the oxygen atoms of the *B*-ring propionate chain exhibit a high flexibility as indicated by large rmsf values. Differences between two structural sub-states can be found regarding the *D*-ring. Compared to sub-state I, the *D*-ring in sub-state II exhibits an enhanced conformational flexibility particularly regarding its carbonyl group and its ethyl side chain.



Figure 3.9. - Color coded root mean square fluctuations (rmsf) of the non hydrogen atoms of PCB in sub-state I (a) and sub-state II (b). In both sub-states the inner rings are the most rigid part of the chromophore, as indicated by the small rmsf values. Compared to sub-state I, the *D*-ring in sub-state II shows higher rmsf values and thus an increased flexibility.

3.2.1.2. Support of the Two-state Model by ¹³C MAS NMR

The ¹³C MAS NMR data strongly support the two-state model for the Pr chromophore, in which the *D*-ring carbonyl group (C19) is centrally involved in the formation

3. Structural Heterogeneity in AnPixJg2 Wild Type and Mutant

of the distinct rotameric conformations of the *D*-ring. As shown in Fig. 3.10 **a**, the ¹³C signal of this carbon atom occurring at ≈ 175 ppm has an asymmetrical shape and is a superposition of several components. Band-fitting analysis yields two peak maxima at 173.9 ppm and 175.4 ppm with a 0.55 : 0.45 distribution. The signal splitting at C19 ($|\Delta\delta^{C}| = 1.5$ ppm) is most likely caused by the local differences in the chemical environment of the *D*-ring and thus suggests the presence of a heterogeneous mixture of *D*-ring isomers, e.g., with different rotameric structures in solution. This assumption is also consistent with a similar two-component shape observed for the ¹⁵N Pr signal of the *D*-ring pyrrole nitrogen atom.²⁷³ Moreover, in AnPixJ, the chromophore heterogeneity seems not to be unique for the Pr dark state, but also apparent for its green-absorbing photoproduct, in which the C19 signal can be fitted to at least three components with a $\Delta\delta^{C}$ gap of 2.3 ppm (Fig. 3.10 **b**).



Figure 3.10. – ¹³C signal splitting of the *D*-ring carbonyl signal observed in solid-state NMR.²⁷² The MAS NMR data were obtained from holoproteins of AnPixJg2 generated by in vitro assembly with uniformly ¹³C- and ¹⁵N-labeled PCB chromophore as Pr (a) and Pg (b) photostates. Only the spectral region of 160 ppm - 190 ppm characteristic for carbonyl and carboxylate resonances is presented (for full spectra with assignments, see Ref.273). A Voigt function with a mixed Lorentzian/Gaussian ratio of 1:1 was applied to fit the two experimental spectra.²⁷⁴ Fitting curves are shown in black (dotted). Deconvolution produced two Pr fits for C19 signal (indicated as a and b) with the relative integral areas (shaded as blue) of 0.55 and 0.45 (calculated as setting overall area to unit). As Pg, three Voigt fits for C19 (a, b, and c) were generated with the relative integral areas of 0.29, 0.32, and 0.39 (shaded as blue). c: Numbering of the C-atoms of PCB.

3.2.2. Interaction between the Chromophore and its Binding Pocket

Tab. 3.2 shows the hydrogen bond occurrence of the interactions between PCB in its different conformational states and the amino acids of the chromophore binding pocket; a detailed picture of the hydrogen bond networks between PCB and the chromophore binding pocket is presented in Fig. 3.12. In both structural sub-states, frequent hydrogen bonding occurs between the carboxylate side chain of ASP291 and the pyrrole nitrogens of the *B*-ring and the *C*-ring of the chromophore, whereas the *A*-ring nitrogen exhibits only negligible interactions with ASP291. With occurrences between 0.79 (sub-state I) and 0.82 (sub-state II) for the *B*-ring and between 0.71 (sub-state I) and 0.68 (sub-state II) for the *C*-ring, both structural sub-states reveal no significant difference in their interaction with ASP291. Two stable hydrogen bond networks that include the propionate chains of the *B*-ring and *C*-ring can be observed for both sub-states of the chromophore, which will further be denoted as network *B* and *C* respectively. Both hydrogen bond networks are presented in Fig. 3.12 **a**. Network *C*, colored in blue, involves ARG301, TYR302, and the *C*-ring carboxylate group of PCB. One oxygen of the carboxylate group is hydrogen bonded to one hydrogen of the guanidinium group while the other carboxylate oxygen interacts simultaneously with the ϵ nitrogen of the aspartate side chain and the hydroxyl group of TYR302.

Fig. 3.11 depicts the time series of this hydrogen bond interaction between ARG301, TYR302 and both oxygen of the carboxylate group of PCB. Except for a few rotations, the carboxylate group remains rigid indicating a high overall stability of hydrogen bond network C. Moreover, the structural sub-states show no difference in the flexibility of this propionate group.



Figure 3.11. – Timeline depicting the interactions of the propionate oxygens of the *C*-ring with ARG301 and TYR302 (network C) and the propionate oxygens of the *B*-ring with PHE319, HIS322, and HIS318 (network B). Except a few rotations the propionate oxygens of the *C*-ring always interact with the same amino acid, whereas the carboxylate group of the *B*-ring is subject to continuous rotations leading to permanently changing hydrogen bond partners for the carboxylate oxygens.

Network B, which is colored in green, involves the propionate group of the B-ring of PCB, the side chains of two histidine residues, HIS322 and HIS318, and the backbone of a phenylalanine residue, PHE319. In that hydrogen bond network one oxygen of the B-ring propionate group interacts with HIS318 and HIS322 and the other carboxylate oxygen with the backbone nitrogen of PHE319. The hydrogen bond to PHE319 shows similar occurrences for both structural sub-states, namely 0.64 for sub-state I and 0.60 for sub-state II. Considerable hydrogen bond interactions can be observed between

3. Structural Heterogeneity in AnPixJg2 Wild Type and Mutant

the δ nitrogen of HIS322 and the chromophore in both conformational sub-states, with occurrences between 0.54 and 0.65 for sub-state I and sub-state II, respectively. The stability of the hydrogen bond interactions between the δ nitrogen of HIS318 and PCB is rather moderate: for sub-state I the occurrence is 0.18 whereas for sub-state II the occurrence is slightly increased (0.28). Hence, in sub-state II the interactions between the chromophore and both, HIS318 and HIS322 are enhanced compared to sub-state I. As depicted in the time series graph (Fig. 3.11, right panel) the *B*-ring propionate group is subject to continuous rotations during the simulation. This indicates, that network B is more fluctuating and weaker compared to network *C*.



Figure 3.12. – Interactions between PCB, key residues of the chomophore binding pocket, and a persistent water. a Hydrogen bond network B (green) and C (blue) involving the B and C ring carboxylate groups. b: Hydrogen bond interactions between the B/C-ring and ASP291 as well as the D-ring and TYR352. The A-ring does not directly interact with adjacent amino acid residues but is hydrogen bonded to a persistent water in the binding pocket (colored in turquoise), which interacts also with TRP289 and ASP291.

Based on crystallographic data, hydrogen bond formation between the hydroxyl group of TYR352 and the carbonyl group of the *D*-ring has been suggested.³³ For substate I only moderate hydrogen bond interactions with an occurrence of 0.18 have been observed during the molecular dynamics simulation. For sub-state II no interaction between PCB and TYR352 takes place due to an increased distance between TYR352 and the downward oriented D-ring. In the crystal structure the side chain of TRP289 is in close proximity to the A-ring with the indole nitrogen pointing towards the A-ring carbonyl oxygen. Therefore, hydrogen bond interactions between the carbonyl group of the A-ring of PCB and the side chain of TRP289 have been proposed.³³ However, no such hydrogen bond interactions between the carbonyl group of the A-ring and the indole nitrogen of TRP289 could be observed during the simulation; in fact, the A-ring does not form hydrogen bonds with any amino acid of the chromophore binding pocket. This seems to be at odds with the high rigidity of the A-ring as derived from the calculation of the RMSF values. One possible explanation for this puzzle is that the A-ring is instead involved in hydrogen bond interactions with solvent molecules which prevents the hydrogen bond formation with amino acids in the chromophore

binding pocket.

	1		1 4 4			
	donor acceptor		sub-state			
			Ι	II		
A-ring	PCB pyrrole N	ASP291	0.02	0.06		
	$\mathrm{TRP289}$	PCB CO	0.00	0.00		
	central water	PCB CO	0.39	0.39		
$B ext{-ring}$	PCB pyrrole N	ASP291	0.79	0.82		
	HIS322 δ -N	PCB COO ⁻	0.54	0.65		
	HIS318 δ -N	PCB COO ⁻	0.18	0.28		
	PHE319 N_{main}	$PCB COO^-$	0.64	0.60		
$C ext{-ring}$	PCB pyrrole N	ASP291	0.71	0.68		
	ARG301 ϵ -N	PCB COO ⁻	0.83	0.79		
	ARG301 NH2 (H2)	PCB COO-	0.47	0.49		
	TYR302 OH	$PCB COO^-$	0.67	0.66		
$D ext{-ring}$	TYR352	PCB CO	0.18	0.00		
	PCB pyrrole N	water	0.55	0.11		
	water	PCB CO	0.12	0.04		
water-aa	TYR352	water	0.09	0.53		
	$\mathrm{TRP289}$	central water	0.5^{2}	0.5^{2}		
	central water	ASP291	0.54	0.54		

Table 3.2. – Hydrogen bond occurrence between sub-state I and substate II of PCB, amino acids (aa) of the chromophore binding pocket, and water ¹

 $^1\,\mathrm{Calculated}$ from $100\,\mathrm{ns}\text{-}940\,\mathrm{ns}$ (whole production run)

 2 Calculated for a simulation time until 290.5 ns

In the following, that hypothesis is pursued by analyzing the overall amount of water during the whole simulation time in an area of 4 Å around the heavy atoms of the pyrrole skeleton of PCB without consideration of the side chains. Fig. 3.13 **a** depicts the median number of water molecules in the investigated area calculated for 10 ns blocks (blue line) in correlation with the backbone rmsd of the β 3 sheet region of the protein (residue ILE285 to GLU290) that shields the chromophore binding pocket from the surrounding solvent (grey line). During the first 50 ns, which is within the equilibration time, the median amount of water in the chromophore binding pocket lies between 4 and 7 molecules and is thus rather small compared to later simulation times. Likewise, the backbone rmsd of the proteins outer loop regions exhibits only slight deviations from the crystal structure during the first 50 ns goes in hand with a continuous increase in the number of water molecules occupying the chromophore

binding pocket. After approx. 150 ns the amount of water in the chromophore binding pocket reaches a stable value. During the first 150 ns the overall median amount of water in this area is 7 molecules whereas for the time from 150 ns to 940 ns an overall median of 10 molecules can be found confirming a higher degree of solvation of the chromophore binding pocket during that time.

Investigations on the persistence of the solvent in the chromophore binding pocket reveal the existence of a central water molecule. This water first enters the chromophore binding pocket after 1 ns and positions itself between the A-ring and TRP289 leading to an enlarged distance between the carbonyl oxygen of the A-ring and the indole nitrogen of TRP289. Hence, any interaction between the A-ring and TRP289 is prevented by the presence of this central water molecule (see Fig. 3.12 b). Instead, a hydrogen bond network between the central water, the carboxylate group of ASP291, the indole nitrogen of TRP289 and the carbonyl group of the A-ring is formed. After approx. 45 ns this water molecule leaves the chromophore binding pocket, however, 5 ns later another water molecule takes the place of the first water. Further replacements of the persistent water occur after approx. 125 ns, after 714 ns, after 761 ns and finally after 795 ns. The long duration between the exchanges of this central water reflect its stability in the chromophore binding pocket. During the whole simulation run, this central water is involved in hydrogen bond interactions with the carbonyl group of the A-ring, the side chain of TRP289, and the carboxylate group of ASP291. Hydrogen bond analysis reveals an average occurrence of 0.54 for the interaction between the persistent water molecule and the carboxylate group of ASP291 and 0.39 for the interaction with the carbonyl group of the A-ring. Due to the highly dynamical behavior of TRP289, hydrogen bonding between the indole nitrogen of the TRP289 side chain and the pyrrole water can only be established during the first 290 ns of the simulation run with an occurrence of 0.5.

The interaction of surrounding solvent molecules with TYR352, the pyrrole nitrogen, and the carbonyl oxygen of the D-ring strongly depends on the structural sub-state of PCB. Fig. 3.13 b depicts the hydrogen bond occurrence between TYR352 as donor and water calculated for blocks of 10 ns in comparison to the rmsd of the D-ring. During the first 145 ns only a few hydrogen bond interactions between water and TYR352 as hydrogen bond donor can be observed. However, shortly before the transition of PCB from structural sub-state I to sub-state II, the hydrogen bond interactions between TYR352 as donor and water as acceptor drastically increase and are reduced again with the transition of PCB from sub-state II to sub-state I. The overall occurrence for hydrogen bonding between TYR352 as donor and water is 0.53 for sub-state II of PCB and 0.09 for sub-state I of PCB. The opposite behavior can be observed for the interaction between the *D*-ring and the solvent. While PCB adopts its structural sub-state I, the hydrogen bond interactions between the pyrrole nitrogen and water are considerably increased (0.55 occurrence) compared to sub-state II (0.11)occurrence). Irrespective of the conformational sub-state of PCB, the carbonyl group of the *D*-ring is considerably less hydrated as derived from hydrogen bond occurrences of 0.12 for sub-state I and 0.04 for sub-state II.


Figure 3.13. – **a**: Median amount of water molecules in the chromophore binding pocket (blue line), taken for blocks of 10 ns in comparison to the backbone rmsd of the $\beta 3$ sheet (grey line). After a strong jump in the rmsd the amount of water in the chromophore binding pocket is continuously increasing. **b**: Hydrogen bond occurrence for the interactions between TYR352 as a donor and water (blue line) in comparison with the rmsd of the *D*-ring (grey line). The number of hydrogen bond events strongly correlates with the structural sub-state of PCB.

3.2.3. Discussion of the Two-state Model for the Conformation of the Chromophore

In this section, a characterization of the structure and dynamics of the second GAF domain of AnPixJ (AnPixJg2) on a long time scale of almost 1 µs is presented. Results from structural investigations on the chromophore by means of rmsd calculations reveal the existence of a second structural sub-state which was not visible in earlier X-ray measurements.³³ These two structural sub-states, denoted as sub-state I and sub-state II, differ in the facial disposition of the *D*-ring. In sub-state I the *D*-ring is arranged at the α -face of the chromophore, similar to the crystal structure, whereas in sub-state II the D-ring adopts a β -facial disposition (see Fig. 3.7 b). Compared to sub-state I, the average downward rotation of the *D*-ring around the C-D single bond in sub-state II is 82°. This two-state hypothesis is corroborated by results from MAS NMR (see section 3.2.1.2) which report a splitting of the *D*-ring carbonyl signal. A similar two-component shape has also been found for the ¹⁵N signal of the pyrrole nitrogen of the D-ring.²⁷⁵ These signal splittings most likely stem from differences in the chemical environment of the *D*-ring as it is the case in heterogeneous mixtures. Moreover, the structural heterogeneity of AnPixJg2 seems to include also the Pg state, since a signal splitting of the *D*-ring carbonyl group can also be observed for the green absorbing photoproduct state (depicted in Fig. 3.10 b) The existence of at least two structural sub-states in AnPixJg2 is congruent with the observation of structural heterogeneity in several other CBCRs^{35–38,47,48} and phytochromes.^{40–42,44,45,45,46,259} These combined findings suggest that structural heterogeneity of the chromophore might be a common feature of all functionally related photoreceptors.

3.2.3.1. Comparison of the Two Structural Sub-states

The comparison of the intramolecular flexibility of the chromophore (Fig. 3.9) reveals, that in both structural sub-states the *B*-ring is the most rigid part of the molecule, whereas the A-ring and the C-ring exhibit a slightly enhanced flexibility. The comparably high rigidity of the B-ring can be explained by its restricted conformational space due to sterical and π -stacking interactions with a HIS residue at the β -face (HIS293) and HIS318 and HIS322 located at the α -face. In both structural sub-states the B-ring and the C-ring are further confined by a stable hydrogen bond network with ASP291. In sub-state I and sub-state II, the oxygens of the propionate side chain of the C-ring are found to be more rigid than those of the B-ring propionate side chain. This indicates the existence of stable hydrogen bonds between the carboxylate group and residues of the chromophore binding pocket that restrain the propionate side chain. The hydrogen bond analysis reveals indeed a static hydrogen bond pattern in which one oxygen of the C-ring propionate chain is hydrogen bonded to one hydrogen of the guanidinium group of ARG301 while the other carboxylate oxygen is simultaneously engaged in a hydrogen bond with the hydrogen of the ϵ -nitrogen and the hydroxyl group of TYR302 (hydrogen bond network C). Likewise, the propionate side chain of the *B*-ring forms a hydrogen bond network involving the backbone nitrogen of a nearby phenylalanine residue, PHE319, and the side chains of two histidine residues, HIS322 and HIS318, respectively (hydrogen bond network B). However, due to continuous rotations of the ring B carboxylate group, hydrogen bond network B is more dynamic than hydrogen bond network C, as further indicated by generally lower hydrogen bond occurrences for network B (see Tab. 3.2).

Whereas no significant difference between the two structural sub-states can be found regarding the mobility of the A/B/C-ring, the dynamical behavior of the *D*-ring is indeed correlated with the structural sub-state of PCB. In sub-state II the flexibility of the *D*-ring is increased which, in particular, is manifested in a highly dynamical carbonyl group and ethyl side chain.

Due to the different orientations of the *D*-ring in both sub-states, its interaction with the chromophore binding pocket and the solved molecules is strongly dependent on the sub-state. In sub-state I, the *D*-ring is moderately hydrogen bonded to the hydroxyl group of TYR352 (0.18 occurrence), whereas in sub-state II every hydrogen bond interaction is prevented by the increased distance between the hydroxyl group of TYR352 and the rotated *D*-ring. A similar behavior is found for the hydrogen bond interactions between the *D*-ring and solvent. In sub-state I, the pyrrole nitrogen shows considerable hydrogen bond interactions with the surrounding water molecules (0.55 occurrence) whereas in sub-state II, the pyrrole nitrogen exhibits strongly decreased solvent interactions (only 0.11 occurrence). Regardless of the conformational state, the carbonyl group of the *D*-ring is significantly less involved in interactions with the solvent compared to the pyrrole nitrogen. The reduced interactions between the solvent and the *D*-ring in sub-state II can be explained by the change in the chemical environment. In sub-state I, the pyrrole nitrogen and the carbonyl group of the D-ring are oriented towards polar residues. This hydrophilic environment supports the solvent accessibility of the D-ring. In sub-state II, however, the D-ring is rotated below the pyrrole ring plane with the carbonyl group and the pyrrole nitrogen pointing inside the hydrophobic cavity which seriously impedes any hydrogen bonding with the solvent. The absence of hydrogen bond interactions reduces the confinement of the D-ring, which naturally explains its higher flexibility compared to sub-state I.

3.2.3.2. Hydration of the Chromophore Binding Pocket

In contrast to earlier assumptions based on X-ray data, no hydrogen bond interaction occurred between the carbonyl group of the A-ring and the indole nitrogen of TRP289 during the molecular dynamics simulation. The absence of any hydrogen bond is explained by the presence of a persistent water molecule in the center of the binding pocket, which was missing in the crystal structure. The existence of a central water molecule located on the α -face of the bilin is well-known in phytochromes^{46,243,247,267,276} but has been rejected for AnPixJg2 based on results from X-ray experiments.³³ A recent molecular dynamics study proposed the existence of such a central water but could not claim its persistence due to a limited simulation length of about 15 ns.^{261} The results from molecular dynamics simulations presented in this chapter clearly show the persistence of a central water molecule in the chromophore binding pocket of AnPixJg2 during about 1 µs. Its position on the α facial side of PCB between the A-ring and TRP289 prevents any interactions between the chromophore and TRP289. Unlike the pyrrole water in phytochromes, which is hydrogen bonded to the pyrrole nitrogens of PCB, the central water molecule in AnPixJg2 forms a hydrogen bond network with the carboxylate group of ASP291, the indole nitrogen of TRP289, and the carbonyl group of the A-ring (see Fig. 3.12) b.

The overall solvation of the chromophore binding pocket is strongly correlated with structural changes in the β 3-sheet of the protein as depicted in Fig. 3.13. As derived from the structural snapshots presented in Fig. 3.4 c, those structural changes observed in the rmsd correspond to an increased distance between the β 3-sheet and the β 2-sheet. This enlarged distance is accompanied by an increase in solvation. The comparably larger gap between the β 2-sheet and the β 3-sheet opens an additional way for the solvent molecules inside the chromophore binding pocket. This finding is in contradiction with the hypothesis, that the color tuning of the Pg state of AnPixJg2 is accomplished by an increased solvation of the chromophore binding pocket (hydration model).²⁶¹ Likewise, a high flexibility of TRP298 was observed that is inconsistent with the hydration model which assigns TRP298 the role of a rigid "gate keeper" that shields the chromophore binding pocket from the solvent. The assumption that the highly dynamic behavior of the TRP289 side chain excludes any shielding effects has also been pointed out by results from previous, preliminary studies²⁷¹ in which TRP298 was *in-silico* mutated against a small, aliphatic residue (valine) and that did not show a stronger hydrated binding pocket. The discrepancy of our findings and the "hydration model" can easily be explained by the shorter duration of the earlier molecular dynamics simulation,²⁶¹ which did not cover the time scale required for the structural changes in the protein as well as the transition of TRP289 from

its initial conformation into one of the other distinct conformations. Interestingly, a highly mobile behavior of TRP298 including the dynamic transition between different conformations was also reported recently for NpR6012g4, a related CBCR³⁹.

3.2.3.3. Possible Events Promoting the Transition between the Conformational Sub-states

In the following possible events that might trigger the formation of sub-state II will be discussed. The first transition from sub-state I to sub-state II occurs after a simulation time of approx 150 ns. 100 ns before this conformational transition, the β 3-sheet moves away from the β 2 sheet, thus reducing the shielding of the binding pocket from the solvent (see Fig.3.4 **b** and **c**). The high flexibility of TRP289, which switches between three different conformational orientations of the indole ring (see Fig.3.5), impedes the interactions with PCB by hydrogen bonding and π -stacking. This leads to a weaker steric confinement of PCB and presumably facilitates the transition from sub-state I into conformational sub-state II.

Moreover, the structural reorganization of the β 3-sheet after 50 ns leads to an increase of the solvent molecules in the chromophore binding pocket. As a result, more hydrogen bond events between TYR352 and the solvent occur. During the time from 100 ns - 145 ns, when PCB is in sub-state I, only few hydrogen bond events between TYR352 as a hydrogen bond donor and the solvent can be observed (0.09)occurrence). The decrease in TYR352-solvent interactions most likely stems from competing hydrogen bond events between TYR352 and PCB and, most notably, the overall reduced amount of water in the chromophore binding pocket. The increasing number of solvent molecules in the proximity of the pyrrole skeleton leads to more interactions between TYR352 and water, which weakens the interaction between TYR352 and PCB and promotes the transition from sub-state I to sub-state II. For PCB in its conformational sub-state II, the hydrogen bonding between TYR352 as donor and the solvent is significantly increased (0.53 occurrence). This behavior can be explained by the absence of any hydrogen bonding between PCB and TYR352 and a better accessibility of TYR352 due to the downward rotation of the *D*-ring. With the transition from sub-state II back to sub-state I, a strong decrease in the interactions between TYR352 and the solvent can be observed (see Fig. 3.13, b). This decrease is presumably caused by competing interactions between TYR352 and PCB as well a steric screening of the hydroxyl group of TYR352 due to the *D*-ring. In summary, the conformational space of the *D*-ring is determined by the interaction with TYR352, which is strong enough to induce a - frequently interrupted - immobilization of the ring. This immobilization is then regularly broken by water molecules invading the binding pocket, which solvate TYR352 and allow the ring to move.

The rearrangement of the β 3-sheet and the subsequently increased hydration of the chromophore binding pocket which presumably facilitate the formation of sub-state II take place on a time scale of ~ 100 ns. The second sub-state was not observed in previously reported simulations of this system²⁶¹ because of the significantly shorter simulation times. This leads to the conclusion, that long simulation times are necessary to capture structural heterogeneity in photosensory proteins.

Conclusion

The molecular dynamics simulations reveal two stable isoforms of the Pr form of AnPixJg2 with a different chromophore configuration which interconvert on timescales of 10 ns-500 ns as estimated from a $\approx 1 \,\mu$ s MD simulation. The transition between these two sub-states is accompanied by a change in the aqueous solvation of Tyr352 and the *D*-ring of the chromophore. Analysis of the dynamical behavior of the protein exhibits strong structural changes in the β 3-sheet followed by an increased solvation of the binding pocket and a high flexibility of TRP289, which is in disagreement with the previously discussed "hydration model".²⁶¹ An additional important finding is the existence of a persistent water molecule in the chromophore binding pocket which is involved in a hydrogen bond network including the *A*-ring, ASP291, and TRP289.

Intriguingly, structural heterogeneity has previously been reported in other phytochromes and cyanobacteriochromes.^{35–38,40–42,44,45,45–48,259} A compelling hypothesis is that such structural heterogeneity has some evolutionary advantage. This is related to the question of the specific roles played by the two conformational sub-states in the photoresponse of the protein, a matter that requires further investigation.

3.3. Part 2: One Mutant Three Sub-states

In this section, the impact of a conserved amino acid, TYR302, that is involved in a stable hydrogen bond network with the chromophore, on the structure of the chromophore is elucidated by *in silico* mutation against phenylalanine, a non-polar amino acid of equal size. The absence of TYR302 triggers the formation of an additional, third structural sub-state which differs in its orientation of the A-ring. Persistent hydrogen bond interactions between then A-ring and adjacent amino acids in this third sub-state prevent any conversion back into sub-state I and II. These findings illustrate the importance of the conserved TYR302 residue on the conformation of the chromophore and its ability to dynamically interconvert between sub-state I and sub-state II.

3.3.1. A Third Sub-State of PCB in the TYR302PHE Mutant

In the following, the impact of mutating the tyrosine residue TYR302 against phenylalanine (TYR302PHE) on the structure of PCB will be discussed. Fig. 3.14 shows the rmsd of the PCB rings A-D of the TYR302PHE mutant (grey) in comparison to those of the wild type (black). Similar to the wild type, the TYR302PHE mutant undergoes a structural transition into sub-state II as visible by a jump in the D-ring rmsd after ≈ 67 ns to a plateau region, region 2. Unlike in the wild type, however, the D-ring rmsd remains stable at plateau region 2, indicating that no interconversion between the two sub-states takes place at the simulated time scale. Interestingly, the A-ring rmsd of the TYR302PHE mutant also exhibits such a jump after ≈ 160 ns to a plateau region, region 2', while the A-ring rmsd of the wild type remains stable during the whole simulation. The presence of a second plateau in the A-ring rmsd points towards

3. Structural Heterogeneity in AnPixJg2 Wild Type and Mutant



Figure 3.14. – rmsd of the individual pyrrole rings A-D of PCB in the wild type (black) and the TYR302PHE mutant (grey). In both systems, the chromophore adopts two structural conformations as visible in two different plateau regions of the D-ring rmsd (1 and 2). The TYR302PHE mutant exhibits a third structural sub-state as apparent from a strong jump in the A-ring rmsd to plateau region 2['] that is absence in the wild type (highlighted by a red square).

a third structural sub-state of the chromophore in the TYR302PHE mutant, in which the A-ring adopts a different conformation than in the crystal structure.

In order to characterize the chromophore conformations that can be associated with the different plateau regions in the A-ring and D-ring rmsd, the average structures of the chromophore corresponding to those plateau regions have been computed. The average structure I, that corresponds to a time range from 30 ns - 66 ns and thus to plateau region 1 in both, the A-ring and D-ring rmsd, is depicted in 3.15 **a**. As expected, the conformation of the A-ring and the D-ring is similar to that in the crystal structure with both rings being oriented above the pyrrole ring plane and on the α -face of the chromophore.

Fig. 3.15 **b** depicts structure II which is on average adopted by the chromophore during the time from 70 ns - 160 ns and corresponds to plateau region 1 in the A-ring rmsd and to plateau region 2 in the D-ring rmsd. The jump in the D-ring rmsd is indeed associated with a downward rotation of the D-ring with respect to the pyrrole ring plane and the average structure II looks similar to sub-state II in the wild type.

Structure III averaged over a time range from 170 ns to 1200 ns, that corresponds to plateau 2 and 2' in the *D*-ring and *A*-ring rmsd, respectively, is presented in Fig. 3.15 c. It shows that the jump in the *A*-ring rmsd to plateau 2' is caused by a downward rotation of the *A*-ring with respect to the pyrrole ring plane. Both, the *A*-ring and the *D*-ring are now oriented towards the β -face of the chromophore. This structural substate, further denoted as sub-state III, does not occur in the wild type and therefore must be induced by the mutation.

In the following, the three sub-states of the TYR302PHE mutant will be characte-



Figure 3.15. – Average structures of the PCB pyrrole rings in the TYR302PHE mutant (the propionate chains, depicted in a lighter gray, have been excluded from analysis). During the first 66 ns the chromophore adopts the same conformation as in the crystal structure (a) but undergoes a transition into sub-state II shortly after 66 ns (b). At ≈ 170 ns transition into a third sub-state occurs in which the *A*-ring and the *D*-ring point towards β -face of the chromophore. Sub-state III has not been observed in the wild type (c).

rized according to their dihedral angle distribution, the planarity of the A-ring and the D-ring with respect to the B-C-reference plane, and their relative flexibility.

Fig. 3.16 depicts the histograms of the *C-D*-single bond and the *A-B* double bond dihedral distributions of the wild type (black) and the TYR302PHE mutant (grey).* In wild type and mutant the different orientation of the *D*-ring in sub-state I and II is reflected in two clearly separated distributions of the *C-D*-single bond dihedrals, while the *C-D*-double bond is barely affected. Hence, the downward rotation of the *D*-ring arises from a rotation around the C–D single bond and not from a twisting around the double bond of the C–D methine bridge. Mean and standard deviation of the *C-D*-single bond dihedral distributions for both sub-states are presented in Tab. 3.3. For both, wild type and mutant, the *C-D* dihedral angle fluctuates around similar average values: In sub-state I, the average *C-D*-single bond dihedral angle is 44° for the wild type and 47° for the mutant while for sub-state II the average *C-D*-single bond dihedral angle is -38° and -43° for wild type and mutant respectively. Hence, the structural sub-state II found in the mutant is identical to sub-state II found in the wild type.

The existence of a third structural sub-state in the mutant is also reflected in the distributions of the A-B-double bond dihedral as well as the A-B-single bond dihedral presented in Fig. 3.16 which exhibit both two peaks. Unlike the C-D-single bond dihedral distribution, those peaks are in a closer proximity leading to partially overlapping distributions. Hence, the downward rotation of the A-ring in sub-state III is less pronounced compared to the downward rotation of the D-ring in sub-state II. The average A-B-single bond and A-B-double bond dihedral angles can be found in Tab. 3.3. Contrary to the downward rotation of the D-ring, which stems mainly from a rotation around the C-D-single bond dihedral, the conformational change of the A-ring in sub-state III arises equally from a rotation around the A-B-single bond

^{*}For TYR302PHE the third state dominates during the simulation. For a better visibility of the two other states therefore only the first 400 ns have been considered for the histograms.

3. Structural Heterogeneity in AnPixJg2 Wild Type and Mutant



Figure 3.16. – a: Distribution of the double bond and single bond dihedrals around the A-B and C-D methine bridge in wild type (black) and mutant (grey). In both systems, the rotation of the D-ring in sub-state II occurs mainly around the C-D single bond as apparent from two separated C-D single bond distributions. Orientation of the A-ring in sub-state III of the mutant, however, involves a rotation of the A-B-single bond and a twisting of the A-B-double bond as visible from two overlapping peaks in both distributions.
b: Visualization of the double bond dihedral (yellow) and single bond dihedral (red)

dihedral and a twisting of the A-B-double bond dihedral. This is reflected in similar changes in both dihedral angles: the A-B-double bond dihedral angle increases from -6° in sub-state II to -37° in sub-state III, whereas the A-B-single bond dihedral angle changes from 14° to -14° . In total, transition from sub-state II to III leads to a rotation of the A-ring of roughly 30° .

The β -facial disposition of the A-ring in sub-state III also influences the planarity of the pyrrole rings of chromophore. Fig. 3.17 depicts the planarity of the A-ring with respect to the B-C ring plane expressed in terms of the scalar product of their normal vectors for wild type (black) and mutant (grey) and averaged over blocks of 10 ns. For the wild type, the A-ring is almost in plane with the B-C rings as indicated by a planarity between 0.9 and 1. For the mutant, however, transition into sub-state III leads to a loss in planarity between the A-ring and the B/C-rings from $\approx 0.9 - 1$ to ≈ 0.6 .

3.3.2. Sub-State III Exhibits an Altered Hydrogen Bond Pattern

After characterizing the conformation of the three sub-states PCB adopts in the TYR302PHE mutant, their interaction with the chromophore binding pocket during



Figure 3.17. – Planarity of the *A*-ring in wild type (black) and mutant (gray) expressed by the relative orientation of the *A*-ring plane (yellow) and the *B*-*C* ring plane (red). In the wild type, both planes are almost parallel to each other as indicated by a planarity of $\approx 0.9 - 1$. In the mutant, transition into sub-state III leads to a significant reduction of the planarity (≈ 0.6 instead of $\approx 0.9 - 1$).

the simulation will be compared to that of the chromophore's sub-states in the wild type. Tab. 3.4 shows a comparison of the hydrogen bond interactions between PCB and amino acids of the chromophore binding pocket in wild type and mutant with respect to the structural sub-states. In general, sub-state I and II of wild type and mutant show a comparable behavior. For both, the B-ring is moderately hydrogen bonded to a tyrosine residue, TYR352, in sub-state I, albeit with an occurrence of 0.11 the interaction is slightly decreased for the mutant compared to the wild type (0.18). Similar to the wild type, this hydrogen bond interaction is absent in sub-state II, due to the enhanced distance between the tyrosine residue and the downward rotated ring.^{*} In the wild type, the C-ring propionate chain is involved in a stable hydrogen bond network including the side chains of an arginine residue, ARG301, and the tyrosine residue, TYR302. In the TYR302 mutant this tyrosine is replaced with a phenylalanine residue, whose non-polar side chain cannot participate in any hydrogen bonds. While the hydrogen bonds between ARG301 and the chromophore are still intact in the mutant, the absence of the third partner, TYR302, in this hydrogen bond network seems to slightly weaken the interactions between ARG301 and the chromophore. This is reflected in a decrease in the hydrogen bond occurrence between the ARG301 NH-group and PCB in its sub-state I and II from 0.83 /0.79 (sub-state I/sub-state II in the wild type) to 0.68/0.64 (sub-state I/sub-state II in the mutant).

Hydrogen bond network B which involves the B-ring propionate chain, the side chains of two histidine residues, HIS322, HIS318, as well as the backbone nitrogen of the adjacent phenylalanine residue, PHE319, remains intact and is even of slightly enhanced overall strength after the mutation.

In the wild type, the conformation of the tetrapyrrole rings of the chromophore is

^{*}An extensive discussion of the hydrogen bond pattern in the wild type can be found in Section 3.2.2.

Table 3.3. – Mean and standard deviations of the dihedral angles (in degree) of the A-B and C-D methine bridge of the crystal structure (c.s.) and both sub-states of the simulated structure for wild type and mutant. The A-B single and double bond dihedral angles include the atoms C4=C5-C6-N22 and N21-C4=C5-C6 respectively, whereas the C-D single and double bond dihedral angles comprise the atoms C13-C14-C15-C16 and C14-C15=C16-N24 respectively

system	dihedral	sub-state I		sub-state II		sub-state III	
		mean	sigma	mean	sigma	mean	sigma
wild	A-B single bond	22	11	18	10	-	-
\mathbf{type}	A-B double bond	1	12	-4	12	-	-
	C-D single bond	44	12	-38	10	-	-
	C-D double bond	13	10	-5	11	-	-
TYR302PHE	A-B single bond	20	11	14	11	-14	10
	A-B double bond	1	12	-6	13	-37	10
	C-D single bond	47	13	-43	9	-40	8
	C-D double bond	14	11	-14	12	-6	10

stabilized by strong hydrogen bonds between the *B*-ring and *C*-ring nitrogens and the carboxylate group of an aspartate residue in the center of the chromophore binding pocket, ASP291. Similarly, for sub-state I and II in the mutant, persistent hydrogen bonds between ASP291 and the *B*-ring and *C*-ring can be found. In the mutant, the occurrence for the hydrogen bond between the *B*-ring pyrrole nitrogen and ASP291 is 0.73 and 0.8 for sub-state I and II, respectively and therefore similar to the occurrence in the wild type (0.79/0.82 for sub-state I/II). Likewise, with 0.71/0.72 for sub-state I/sub-state II the hydrogen bond occurrence between the *C*-ring pyrrole nitrogen and ASP291 is comparable to that in the wild type (0.71/0.68 for sub-state I/II).

While for sub-state I and II wild type and mutant show a similar hydrogen bonding behavior, the hydrogen bond pattern of sub-state III of the mutant is altered compared to the wild type. The most significant difference is observable for the A-ring. An overview over the differences in the hydrogen bond interactions in sub-state II and sub-state III is presented in Fig. 3.18. In sub-state I and II (the latter is depicted in Fig. 3.18 **a**), the A-ring is not involved in significant hydrogen bond interactions with amino acids of the chromophore binding pocket. In sub-state III, however, the orientation of the A-ring towards the β -face of the chromophore enables interactions between the carbonyl oxygen of the A-ring and the backbone nitrogen of an adjacent threonine group, THR292 as visible in Fig. 3.18 **b**. Remarkably, in state III the A-ring is also involved in hydrogen bonds with ASP291. This interaction, which is negligible in sub-state II (0.06 occurrence), slowly increases during the time the chromophore is in sub-state III. As a consequence, the interaction between the *B*-ring and ASP291 drops from 0.73/0.8 in sub-state I/II to only 0.31 in sub-state III.

In the following, the conformational fluctuations of the chromophore in its three substates will be discussed. Fig. 3.19 depicts the color coded root mean square fluctuation



Figure 3.18. – Hydrogen bond interaction of the the A/B/C-ring nitrogen atoms and adjacent amino acids in sub-state II (a) and sub-state III (b). In sub-state III the A-ring replaces the B-ring as hydrogen bond partner with aspartate (ASP291). The orientation of the A-ring is further stabilized by hydrogen bond interactions with THR292 which explains the long life time of sub-state III.

(rmsf) values of the individual sub-states in the wild type and the TYR302PHE mutant. Blue colored atoms are those with the smallest rmsf value and hence the highest rigidity while red colored atoms exhibit high rmsf values and therefore a high flexibility. Similar to the wild type, the inner rings (*B*-ring and *C*-ring) of PCB in the TYR302PHE mutant are the most rigid ones which is due to similarly strong hydrogen bond interactions with the aspartate residue, ASP291. Compared to the *D*-ring, which shows an enhanced flexibility in sub-state II and III, the *A*-ring in sub-state III is much more rigid. This rigidity can be explained by stable hydrogen bonds between the *A*-ring pyrrole nitrogen and ASP291 as well as between the *A*-ring carbonyl group and the backbone nitrogen of the adjacent threonine residue, THR292 in sub-state III (see Fig. 3.18 **b** and Tab. 3.4).

Differences between wild type and mutant can be found for the C-ring propionate group. In the wild type, the conformation of that propionate group is stabilized by hydrogen bonds with the side chain of an arginine residue, ARG301 and the adjacent tyrosine residue, TYR302 (hydrogen bond network C) which is reflected in stable hydrogen bonds. The missing interaction with TYR302 in the TYR302PHE mutant leads to an enhanced dynamics of that propionate group which is most pronounced for sub-state III. In the crystal structure and the simulated structure of the wild type, that C-ring propionate group adopts a stretched configuration³³ which remains stable during the whole simulation. In order to investigate the flexibility of the C-ring propionate group in the mutant, the rmsd of the carbon atoms of the propionate group was computed and compared to structural snapshots.

Fig. 3.20 shows the rmsd of the C-ring propionate carbons (highlighted by a blue rectangle) as well structural snapshots of PCB in the TYR302PHE mutant (starting structure and simulated structure at 540 ns). It reveals, that the higher flexibility

3. Structural Heterogeneity in AnPixJg2 Wild Type and Mutant



Figure 3.19. – Color coded, atom wise root mean square fluctuations (rmsf) of PCB in the individual sub-states in wild type (a) and TYR302PHE mutant (b). In both systems, the *D*-ring is more flexible in sub-state II. Compared to the wild type the mutant exhibits an enhanced flexibility of the *C*-ring propionate oxygens. Unlike the *D*-ring in sub-state II and III, the *A*-ring exhibits a higher rigidity in sub-state III which is explained by the formation of stable hydrogen bonds with adjacent amino acids which are absent in case of the *D*-ring.

of the C-ring propionate group in the mutant is not only restricted to rotations of the carboxylate oxygen atoms, but affects the total conformation of the propionate group. During the first ≈ 175 ns the C-ring propionate side chain adopts the same, stretched conformation as in the starting structure presented in Fig. 3.20. At ≈ 176 ns, shortly after the transition to sub-state III, however, the rmsd jumps to a plateau region. That plateau region corresponds to a downward bent conformation of the C-ring propionate side chain as presented by the structural snapshot extracted at $540 \,\mathrm{ns.}$ During the remaining simulation time, the C-ring propionate chain jumps between the stretched conformation, which is constantly adopted in the wild type, and the downward bent conformation below the tetrapyrrole ring plane. However, the hydrogen bonding behavior seems to be in general only little affected by this highly dynamic behavior and the strength of the hydrogen bonds between the C-ring propionate oxygen atoms and ARG301 is only slightly decreased compared to the wild type (see Tab. 3.4). This, on the first glance surprising, result can be explained by the structure of the ARG301 side chain which consists of a straight aliphatic chain of three carbon atoms followed by a guanidinium group. Being sterically less demanding and more agile than the tyrosine side chain it can easily adjust to the conformational changes of the propionate side chain which becomes also visible in the structural snapshot taken at $540 \,\mathrm{ns}$ (see Fig. 3.20).



Figure 3.20. – rmsd of the *C*-ring propionate carbon atoms of the TYR302PHE mutant (highlighted by a blue rectangle) depicting several jumps between two plateau regions. Structural snapshots of the starting structure of the TYR302PHE mutant and at ≈ 540 ns, which correspond to the two plateau regions, reveal that in the starting structure the propionate chain adopts a stretched configuration while the conformational changes displayed in the rmsd correspond to a downward bent conformation of the propionate side chain.

3.3.3. Mutation of TYR302 Opens a New Gate for Solvent

Now, the impact of mutating TYR302 on the solvation of the chromophore will be discussed. Fig. 3.21 a shows the median amount of water calculated for blocks of 10 ns^{*} around the chromophore skeleton together with the rmsd of the β 3-sheet backbone. In the wild type, the degree of solvation of the chromophore is correlated with strong configurational changes of the β 3-sheet. A restructuring of the β 3-sheet opens a channel for water molecules into the chromophore binding pocket which increases the number of water molecules around the chromophore (compare chapter 3.2.1). For the mutant, no such correlation seems to exist: After $\approx 70 \,\mathrm{ns}$ the rmsd jumps to a low plateau region, region 2, were it remains for roughly 650 ns. The corresponding conformational changes of the β 3-sheet are visualized by two structural snapshots extracted at 100 ns (blue) and 200 ns (green) in comparison to the crystal structure (grey). It is visible that except of a higher degree of disorder, the conformation of the β 3-sheet in both snapshots does not deviate significantly from that in the crystal structure. Only after ≈ 750 ns a strong jump in the rmsd to plateau region 3 occurs. As apparent from snapshot 3 extracted after 900 ns, this jump can be assigned to an arrangement of the β 3-sheet in much closer proximity to the α 4-helix which opens a potential gap for water to enter the binding pocket. However, the graph displaying the median amount of water around the chromophore displays an increase of solvent molecules already after $\approx 170 \, \text{ns}$; the solvation of the chromophore binding pocket is

^{*}For a detailed description on how the median amount of water was computed see Chapter A.1.2

Table 3.4. – Occurrence and existence of hydrogen bond pattern between PCB and the chromophore binding pocket (CBP) in the AnPixJg2 wild type and its mutant TYR302PHE. In the third structural sub-state of the mutant, the orientation of the A-ring is stabilized by stable hydrogen bonds with ASP291 and THR292 that are absent in sub-state I and II.

	PCB	CBP	wild type		TYR302PHE		
			Ι	II	Ι	II	III
D-ring	CO	TYR352 OH	0.18	0	0.11	0.0	0.0
	COO^{-}	ARG301 ϵ -N	0.83	0.79	0.68	0.64	0.74
$C ext{-ring}$	COO	ARG301 NH2	0.47	0.49	0.49	0.45	0.56
	COO^{-}	TYR302 OH	0.67	0.66	-	-	-
	pyrrole N	ASP291 COO	0.71	0.68	0.71	0.72	0.72
	COO^{-}	HIS318 δ -N	0.18	0.28	0.28	0.32	0.53
B-ring	COO	HIS322 δ -N	0.54	0.65	0.62	0.74	0.7
_	COO^{-}	PHE319 N_{main}	0.64	0.60	0.58	0.59	0.5
	pyrrole N	ASP291 COO	0.79	0.82	0.73	0.8	0.31
A- ring	pyrrole N	ASP291 COO	0.02	0.06	0.06	0.16	0.52
	CO	THR292 N_{main}	0	0	0	0	0.41

therefore independent of the strong conformational changes of the β 3-sheet. Instead, the median amount of water around the chromophore's tetrapyrrole rings seems to be correlated with its structural sub-state. Fig. 3.21 b depicts the median amount of water in comparison to the rmsd of the D-ring (grey) and the A-ring (black). In sub-state I (30 ns - 66 ns), a median of six water molecules can be found around the chromophore pyrrole rings. After transition into sub-state II, the graph depicting the median amount of water jumps shortly to higher values. In total, the median number of water molecules for sub-state II is 7 and thus slightly increased compared to sub-state I. The strongest increase of water around the chromophore can be observed directly after transition into sub-state III ($\approx 166 \,\mathrm{ns}$) which corresponds to the jump in the A-ring rmsd. The median number of water molecules rises almost instantaneously and the total median amount of water in sub-state III is with 10 molecules significantly higher than in the other two sub-states. That strong increase of water molecules directly after the transition to sub-state III indicates potential structural changes in amino acid residues next to the chromophore that are somehow correlated to the interconversion from sub-state II to sub-state III and that might open another channel for water molecules into the binding pocket.

Therefore, the conformation and dynamical features of the amino acid residues in close proximity to the chromophore were investigated. A particularly interesting behavior was found for a histidine residue close to the *B*-ring of the chromophore, HIS293, which is conserved among red/green CBCRs.³³ In the crystal structure and



Figure 3.21. – a: Median amount of water around the PCB skeleton calculated for blocks of 10 ns (blue line) versus the backbone atom rmsd of the β-3 sheet (gray line). As apparent from the structural snapshots extracted at 100 ns and 200 ns the amount of water increases before the restructuring of the β3 sheet that excavates the binding pocket. b: Median amount of water around the PCB skeleton (blue line) in comparison to the rmsd of the A-ring (dark gray line) and D-ring (light grey line) suggesting a correlation between the increased solvation of the chromophore and the transition into sub-state III.

the simulated structure of the wild type, the imidazole side chain of that histidine residue is arranged directly below and parallel to the ring plane of the PCB *B*-ring. In the TYR302PHE mutant, however, that histidine residue exhibits a highly dynamical behavior. Fig. 3.22 **a** shows the rmsd of the HIS293 side chain (red) in comparison to the rmsd of the *A*-ring (gray). At about 160 ns the rmsd of HIS293 exhibits a strong jump to a plateau region. From the inset in Fig. 3.22 **a** it is visible that this jump occurs directly before the transition from sub-state II to sub-state III. Two structural snapshots extracted at 160 ns and 165 ns reveal the conformation of HIS293 directly before and after the structural change that led to a jump in the rmsd. During the time until 160 ns, the side chain of HIS293 was arranged directly below the *B*-ring with the imidazole ring plane approximately parallel to the *B*-ring plane. Shortly after 160 ns, the imidazole ring approaches in a directed movement the *A*-B methine bridge with the imidazole ring plane perpendicular to the *A*-ring / *B*-ring plane. This movement

3. Structural Heterogeneity in AnPixJg2 Wild Type and Mutant



Figure 3.22. – a: Comparison of the A-ring rmsd of PCB in the TYR302PHE mutant and the rmsd of the adjacent HIS293 residue revealing a correlation between the transition of PCB into sub-state III and an increased mobility of the HIS293 side chain during which it changes its conformation from parallel (0 ns - 160 ns) to orthogonal with respect to the B-ring plane. b: rmsd of the HIS293 side chain in the wild type, depicting a stable configuration parallel to the B-ring plane of PCB during the whole simulation.

is almost instantaneously followed by the rotation of the A-ring towards the β -face of the chromophore. During the remaining simulation time, HIS293 exhibits a highly dynamical behavior and continuously switches between conformations in which its side chain is arranged below the B-ring with the imidazole ring plane parallel to the B-ring plane and conformations in which its side chain is arranged between the A-B methine bridge with the imidazole ring plane perpendicular to the B-ring plane.

Fig. 3.22 **b** shows the rmsd of HIS293 in the wild type. The rmsd remains stable indicating that no strong conformational changes occur. During the whole simulation time, the side chain adopts a similar conformation as in the crystal structure, i.e. it remains below the *B*-ring with the imidazole ring plane parallel to the *B*-ring plane as demonstrated by a representative snapshot extracted at 150 ns. This implies that the dynamical behavior of HIS293 is induced by the mutation of TYR302 residue.

The increased mobility of the HIS293 side chain starts shortly before transition of

the chromophore into sub-state III. This structural transition of the chromophore, in turn, is accompanied by a strong increase in solvation of the chromophore binding pocket (see Fig. 3.21 b). One obvious assumption is therefore that the highly dynamical behavior of HIS293, i.e. the jumps between several conformations of the side chain, opens a new gate for the solvent molecules, most likely the area below the B-ring plane that was covered before by the HIS293 side chain.

3.3.4. Discussion

Mutagenesis of the TYR302 residue to phenylalanine leads to the formation of an additional, third structural sub-state, sub-state III, which was not observed in the wild type. Sub-state III is characterized by a rotation of both, the *D*-ring and the *A*-ring, towards the β -facial side of the chromophore. The transition to sub-state III occurred from sub-state II after a simulation time of roughly 160 ns. Unlike the wild type, which exhibits several interconversions between sub-state I and sub-state II, no such interconversion can be observed for the mutant during a simulation time of 1.2 µs. Instead, transition from sub-state I to sub-state II occurs at a rather short simulation time of ≈ 66 ns after which the chromophore stays in sub-state II until the last transition into sub-state III.

The structural transition from sub-state II to sub-state III is accompanied by a change in the hydrogen bond pattern. In sub-state I and II in wild type and mutant, the A-ring does not undergo significant hydrogen bond interactions with adjacent amino acid residues (see Tab. 3.4). Instead, it interacts with a persistent solvent in the binding pocket. In sub-state III however, the A-ring nitrogen is hydrogen bonded to the aspartate residue ASP291 (0.52 occurrence, see Tab. 3.4). As a result of that newly formed hydrogen bond, the interaction between the B-ring nitrogen and that aspartate residue gets significantly weakened (from 0.73/0.8 occurrence in sub-state I/II to 0.31 occurrence in sub-state III). Moreover, the downward rotated orientation of the A-ring in sub-state III facilitates hydrogen bond interactions between the carbonyl group of the A-ring and the backbone nitrogen of a nearby three nine residue (0.41)occurrence). These stable hydrogen bond interactions with two amino acid residues constrain the A-ring in its β -facial orientation and, most likely, impede a conversion from sub-state III back to sub-state II. Interestingly, during the whole simulation time, no conformation was observed in which the A-ring is oriented towards the β -face of the chromophore (downward rotated) while the D-ring is oriented towards the α -face of the chromophore (upwards rotated).

A question that consecutively comes in mind is how the rotation of the A-ring affects the spectral properties of the chromophore. Recent studies have shown that the absorption spectrum in red/green CBCRs is mainly governed by the effective conjugation length of the chromophore.²⁷⁰ As depicted in Fig. 3.17, the planarity of the A-ring with respect to the B-C ring plane significantly decreases after transition into sub-state III. However, as visible in Fig. 3.1, unlike the B/C/D-ring, the A-ring carbons are saturated and therefore the A-ring contributes only slightly to the conjugation of the chromophore. It would therefore be not surprising if the spectral properties of the sub-state III show only modest differences compared to those of sub-

3. Structural Heterogeneity in AnPixJg2 Wild Type and Mutant

state II. In order to answer this question exhaustively, however, subsequent QM/MM studies combined with experimental UV/VIS studies are necessary.

The mutation of TYR302 against phenylalanine also influences the conformation and dynamical behavior of the *C*-ring propionate group. In the wild type, this propionate group is involved in stable hydrogen bonds with TYR302 and adopts a static, stretched conformation during the whole simulation time. The lacking hydrogen bond interaction in the mutant leads to a changed conformation of the propionate group after the transition into sub-state III: first a stable, downward bent conformation with the carboxylate oxygen atoms pointing below the *B*-*C* ring plane; after ≈ 650 ns the propionate group adopts again its previous, stretched conformation and shortly afterwards starts to fluctuate between those two configurations. Hence, its stretched conformation is mainly imposed by the hydrogen bond interactions with the bulky phenol side chain of TYR302 which acts like an anchor due to the steric demand of its side chain. It can therefore be assumed that one function of the conserved TYR302 residue lies in anchoring the rigid, stretched conformation of the *C*-ring carboxylate group.

The newly adopted, downward bent conformation of the C-ring propionate group does not influence the conformation of the B-ring propionate, which maintains the bent conformation that was also found in the wild type. Interestingly, in the structurally related phytochrome CPh1 and CBCR TePixJ, respectively, one of the propionate groups of the chromophore is always bent while the other is stretched.³³ In all cases, the conformation of the propionate side chains are fixed by stable hydrogen bond networks:³³ CPh1 and TePixJ comprise a stretched B-ring propionate chain which is hydrogen bonded to an arginine residue and a bent C-ring propionate side chain which interacts with a histidine residue. In AnPixJg2, the bent B-ring propionate side chain is hydrogen bonded to two histidine residues and a phenylalanine residue (network B, see Fig. 3.12) while the stretched C-ring propionate side chain is involved in a hydrogen bond network with an arginine residue and the tyrosine residue, TYR302. It was previously assumed that the arginine residue is responsible for stabilizing the conformation of the stretched propionate side chain.³³ However, solution NMR structures from a related red/green CBCR NpR6012g4³⁹ reveal multiple configurations for the C-ring propionate chain together with considerable hydrogen bonds with the arginine residue (6/10 structures) but only modest interactions with the tyrosine residue (1/10) which makes a stabilizing contribution of arginine questionable. In agreement with these results obtained for NpR6012g4, the dynamical switching of the C-ring propionate chain between a stretched and downward bent conformation in the investigated TYR301PHE mutant of AnPixJg2 clearly shows, that not arginine, but tyrosine is responsible for stabilizing the stretched propionate side chain of PCB in AnPixJg2.

3.3.4.1. Possible Events Correlating with the Transition into Sub-state III

Mutating the TYR302 residue also seems to influence the dynamics of a histidine residue, HIS293, which is located nine residues apart. In particular, as discussed in Section 3.3.3, a strong increase in the mobility of the HIS293 side chain can be observed directly before the transition into sub-state III which suggests that both events are



Figure 3.23. - Color coded root mean square fluctuations (rmsf) of the protein residues 293-301 of wild type (0-940 ns, a) and the TYR302PHE mutant (0-160 ns, b) including the residue HIS293 and the mutated residue TYR302. Unlike expected, in the wild type the residues adjacent to HIS293 as well as the HIS293 side chain exhibit a slightly increased flexibility compared to the mutant.

correlated. In the following several possible reasons why the mutation of TYR302 alters the mobility of HIS293 will be evaluated. The first hypothesis that naturally comes in mind, is that the mutation of TYR302 increases the dynamical behavior of the whole region of the protein, i.e. the residues 302-293. The secondary structure of the region comprises two β -turns and a helix on which the HIS293 residue is located; a higher degree of disorder and dynamics within that region could destabilize the arrangement of the HIS293 side chain with respect to the chromophore. A quantification and comparison of the degree of mobility within the region TYR302-HIS293 can be obtained from the computed, color coded rmsf values depicted in Fig. 3.23 (blue: stable, red: dynamic). Surprisingly, the wild type exhibits slightly increased fluctuations within that regions; in particular the helix region is more mobile compared to the mutant. Therefore, the highly dynamical behavior of the HIS293 residue cannot simply be explained by an overall increased mobility of that region induced by the mutation.



Figure 3.24. – **a**: Timeline of the vertical distance between the HIS293 imidazole ring plane and the *B*-ring plane for wild type (black) and TYR302PHE mutant (gray) revealing a larger distance for the latter. **b**: Histogram of the HIS293-*B*-ring ring plane distance (0-940 ns for the wild type, 0-150 ns for the mutant). The median distance between HIS293 and the *B* ring is 0.6 Å larger than in the wild type, suggesting weaker $\pi - \pi$ interactions.

3. Structural Heterogeneity in AnPixJg2 Wild Type and Mutant

Subsequently, the arrangement of the HIS293 imidazole ring plane with respect to the *B*-ring is compared between wild type and mutant. Fig. 3.24 **a** depicts a timeline of the vertical distance between the HIS293 imidazole ring plane and the ring plane of the *B*-ring for the wild type (black) and the TYR302PHE mutant (grey) during the first 200 ns and 150 ns of the simulation respectively^{*}. In both systems, the vertical distance at the beginning is equal and roughly 4.5 Å, however, in the wild type this distance decreases during the first 50 ns and remains at median distances of ≈ 3.7 Å. In the mutant however, the distance between the ring planes remains at the initial, higher values (median distance: 4.35 Å).

Aside from larger HIS293-*B*-ring distances, an increased solvation of the HIS293 nitrogen atoms as well as the *C*-ring propionate oxygen atoms can be observed. The median amount of water in an area of 3.5 Å around the HIS293 nitrogen atoms as well as the *C*-ring propionate oxygen atoms is presented in Tab. 3.5. In the wild type, the median amount of solvent molecules around the HIS293 nitrogens and the *C*-ring propionate oxygen atoms is 3 and 4 molecules, respectively, while in the mutant the number of solvent molecules is increased to 4 and 6 molecules.

Table	3.5. –	Median	amount	of	water	around	the	HIS293	nitrogens	and	the	C-ring
		propiona	ate oxyge	n at	oms in	wild typ	be and	d TYR30	2PHE mut	ant.	The	mutant
		exhibits	an increa	sed	solvati	on of th	is are	a as app	arent from	a hi	gher	median
		amount	of solven	t m	olecule	S.						

	wild	type	TYR302PHE		
	HIS293 C-ring		HIS293	C-ring	
median no. water	3	4	4	6	

These results point towards the possibility that in the wild type, the interaction between TYR302 and solvent molecules in form of hydrogen bonds prevents the water from entering the area between the *C*-ring carboxylate groups and the HIS293 side chain. Moreover, TYR302 acts as hydrogen bonding partner for the *C*-ring carboxylate group and thereby blocks one hydrogen bond site for the interaction with water. In the absence of TYR302, water penetrates the area between the *C*-ring and the HIS293 residue and undergoes hydrogen bond interactions with the vacant site which leads to an increased solvation. The increased hydration of the HIS293 nitrogen atoms impedes π -stacking interactions between the HIS293 imidazole ring and the *B*-ring and fosters a higher mobility of the HIS293 side chain. So far, however, these considerations are at the stage of an early hypothesis and require further, targeted investigations to be confirmed.

^{*}After 160 ns the movement of HIS293 as well as interconversion from sub-state II and sub-state III already take place in the TYR302PHE mutant, therefore only the first 150 ns were evaluated. For better comparison, the timeline of the wild type depicts only the first 200 ns.

3.3.5. Conclusion and Outlook

Molecular dynamics simulations reveal, that mutation of TYR302 against a phenylalanine residue triggers the formation of a third sub-state in AnPixJg2 which is characterized by the β -facial orientation of the A-ring and the D-ring. Hydrogen bond interactions involving the A-ring and adjacent amino acids, ASP291 and THR292, preserve that sub-state III on long time scales and prevent conversion back into sub-state I/II. These findings suggest, that TYR301 plays an important role in stabilizing the sub-states I and II in the AnPixJg2 dark stable state. Moreover, the absence of hydrogen bonds with the C-ring propionate chain upon mutation of TYR301 induces the adoption of a downward bent conformation of that propionate chain instead of the stretched conformation found in the AnPixJg2 wild type and other CBCRs and phytochromes,³³ demonstrating that TYR302 is responsible for stabilizing the stretched conformation of the propionate chain instead of ARG301 as previously assumed.

A particularly unexpected behavior is also observed for a histidine residue, HIS293, which is located nine residues apart from the mutated tyrosine residue. That histidine exhibits a highly increased dynamical behavior that correlates with the transition into sub-state III. Remarkably, the mobility of the protein backbone in the respective area, i.e. residue HIS293-TYR302, is not increased. Instead, the mutant features a larger vertical distance between the HIS293 imidazole ring plane and the *B*-ring plane which is accompanied by an increased solvation of the HIS293 nitrogens and the *C*-ring carboxylate group. One consideration is that in the wild type TYR302 hinders the solvent molecules in entering the area between the *C*-ring propionate and HIS293 side chain by engaging the solvent molecules in hydrogen bonds. The absence of TYR302 leads to an increased solvation of that area and the increased hydration of the HIS293 imidazole ring hampers the π - π -interactions with the *B*-ring and increases the flexibility of the HIS293 side chain.

In order to confirm that hypothesis and entangle the relationship between the mutagenesis of TYR302, the mobile histidine residue, and the additional structural sub-state of the chromophore, further effort is required. Future work should be directed towards the particular reasons behind the mobility of that histidine residue as well as the role of that histidine residue in the formation of the third sub-state. One approach to explore a potential causality between the two events (increased mobility of HIS2393 and formation of the third sub-state) is to constrain this histidine residue to its position in the crystal structure and investigate in a molecular dynamics simulation run whether the third sub-state is realized.

In summary the results emphasize the impact of the conserved tyrosine residue, TYR302, on the conformation of the chromophore. These findings stimulate, currently already ongoing, experimental UV/VIS studies that address the impact of the mutation and the third structural sub-state on the absorption properties of AnPixJg2.

4. Oligothiophenes on Nano-structured Surfaces

4.1. Introduction

During the last decades, organic semiconducting materials of small molecular weight have drawn significant attention, mainly due to their inexpensive production, biocompatibility,^{277,278} structural variety,^{20–22} and the potential of tailoring the molecules by rational design.²³ An overview over prominent representatives of small-molecular weight organic semiconductors is provided in Fig. 4.1 a. They all consist of an extended π -electron system that is able to delocalize charge and exhibits optical gaps of \approx $1.5 \,\mathrm{eV} - 3 \,\mathrm{eV}$ which allows light absorption or emission in the visible range.²⁷⁹ Unlike inorganic semiconductors which exhibit continuous energy bands, in their crystal structure, organic semiconducting molecules are only weakly bound via van der Waals forces²⁸⁰ and therefore discrete molecular orbitals are preserved. The existence of discrete energy levels in organic semiconductors (OSC) results in a charge transport mechanism that is inherently different from that in inorganic semiconductors. In OSC, charge is transported from one molecule to the other via a hopping mechanism. The magnitude of charge transport is therefore mainly governed by two factors, namely the transfer integral^{63,281} and the reorganisation energy.^{62,281–283} The transfer integral is associated with the splitting of the HOMO and the LUMO, whereas the reorganisation energy is the energy required for structural reordering when a charge is transported through the molecule.²⁸⁴

Due to the anisotropy of the electron density in π -conjugated molecules, both factors are strongly steered by local chemical interactions 62,63 and hence orientation and packing of the molecules on the electrode surface⁶⁴ and in the organic blend.^{65,66} The dependence of the charge transfer on the molecular orientation in the device can be rationalized by the example of different types of organic field effect transistors (OFETS). A schematic overview of the structure of a conventional and a vertical OFET (VOFET) is given in Fig. 4.1.²⁸⁵ As major building blocks OFETS comprise a source, a drain, (V_D) , and a gate electrode, V_G . The gate electrode is separated from the source and the drain by a dielectric material. The organic semiconducting material is processed on top of the dielectric and in contact with the source and drain electrode. When a voltage is applied between gate and source electrode, the resulting electric field induces an accumulation of charge carriers in a channel-like area at the semiconductor/isolator interface between source and drain.²⁸⁵ Application of a second voltage between source and drain electrode leads to migration of the charge carrier to the drain electrode. In conventional OFETs, source and drain electrode are arranged perpendicular to the gate electrode, hence the direction of charge flow is parallel

4. Oligothiophenes on Nano-structured Surfaces



Figure 4.1. – **a**: Overview over prominent organic semiconducting molecules. **b**-**c**: Building scheme of conventional OFETs (**b**) and vertical OFETs (**c**). Whereas in conventional OFETs source and drain electrodes (V_D) are arranged perpendicularly with respect to the gate electrode, (V_G), in vertical OFETs they are both stacked parallel to the gate electrode. Due to the different direction of the charge carrier migration, different molecular orientations are required for maximum performance. The figures were taken from Ref. 285.

to the semiconductor/isolator interface. In VOFETs, however, source and drain are stacked parallel to the gate electrode which leads to a charge migration perpendicular to the interface. Since in π -conjugated molecules the majority of charge transport occurs in π - π -stacking direction,²⁸⁵ an optimum performance of conventional OFETs requires a standing molecular orientation while in VOFET's horizontal alignment of the semiconducting molecules is beneficial. Depending on the exact nature of the opto-electronic device, different molecular orientations of the organic semiconductors with respect to the electrode surface are therefore required in order to yield maximum performance.

Among the variety of organic semiconducting materials, oligothiophene based OSC stand out due to their high charge carrier mobilities.^{53,54,56–58,61,286} The molecule investigated in this work, α -sexitiophene (α -6T), is depicted in Fig. 4.2 c. α -6T is a prominent representative of oligothiophene based OSC and has already found application in opto-electronic devices, such as organic field emission transistors (OFETS).^{60,287} It consists of six thiophene units that are connected at the C_{α} atom. Two adjacent thiophene units can be either in a *trans* configuration, in which the two sulfur atoms are aligned in an alternating fashion, or in a *cis* configuration, in which the two sulfur atoms point in the same direction. The *all-trans* configuration of α -6T, as depicted in Fig. 4.2 c, is the energetically preferred one.²⁸⁸

Part I of Chapter 4 addresses the investigation of single α -6T molecules on the reconstructed Au(100) surface by atomistic *ab initio* calculations. Gold has a wide range of industrial applications, from catalysing oxidation reactions^{67–71} to electrodes in organic field emitting transistors. This versatility has fueled interest in the properties of organic-gold interfaces.^{72–81} What makes the low index surfaces of gold (Au(100),^{82–84} Au(110),^{85,86} and Au(111)^{87,88}) special are their large scale reconstructions that all have characteristic height modulations in their top layer and offer various adsorption sites for organic molecules of different adsorption strength. These



Figure 4.2. – **a**: STM image of the reconstructed Au(100) surface exhibiting reconstruction rows running along the [110] direction (the unit cell is highlighted in blue).⁸⁴ **b**: Adsorption schematic of α -6T on the reconstructed Au(100) surface. In scenario A the molecule migrates into a valley when initially placed on an ill-favoured *on-ridge* position. In scenario B the molecule digs itself into its preferred *in-valley* site by pushing the ridge aside. **c**: Sketch of α -sexithiophene (α -6T). Carbon atoms bonded to sulfur atoms are denoted as C_{α}. The image is taken from the authors work.²⁸⁹

characteristic patterns could be used as templates to guide molecular assembly on the surface.^{76,77} This *patterning concept* assumes that distinct adsorption sites exist on these surfaces, that organic molecules will migrate almost exclusively to energetically preferred adsorption sites and that the molecule-surface interaction does not affect the reconstruction used for patterning. Molecules that interact strongly with the gold surface violate at least one of these assumptions and the patterning concept breaks down.^{72–75,290–292} For weak molecular adsorption, on the other hand, molecular assembly obeys the topography of the surface.

A deeper understanding of the interaction between organic molecules and reconstructed surfaces can be gained from *ab initio* calculations which provide atomistic insights into properties such as adsorption geometries and heights, adsorption energies, charge transfer, and band gaps. However, the high complexity of the reconstructed noble metal surfaces results in large unit cells that makes theoretical investigations challenging. Theoretical studies of organic adsorbates on metal surfaces have thus mainly focused on the adsorption of small molecules on ideal surfaces as model systems.¹⁵² At the example of the reconstructed Au(100) surface, the work presented in this thesis goes beyond the idealized description of gold surfaces and explicitly takes into account the effect of the large-scale surface reconstruction on the adsorption properties of α -6T. Both, surface and molecule, are described on the DFT level of theory, which provides information about the electronic structure of molecule and surface and allows the investigation of charge transfer processes.

A large-scale STM image of the Au(100) surface is presented in Fig. 4.2. The reconstruction leads to a contracted, quasihexagonal top layer with reconstruction rows running along the [110] direction.⁸⁴ The resulting complex height modulation of the topmost layer gives rise to a variety of different adsorption sites that can be broadly

classified according to their height into *on-ridge* and *in-valley*, with an energetic preference for the latter. For weak molecular adsorption, one might expect that a molecule placed in an unfavorable *on-ridge* site would migrate into a more favorable valley site, as schematically shown in scenario A in Fig. 4.2b. Instead, the results presented in this chapter show that the molecule descends vertically and displaces the gold ridge as depicted in scenario B in Fig. 4.2b. This behavior is incompatible with the concept of well-defined adsorption sites and forces a reconsideration of patterning strategies for organic films on metal substrates.

While Part I of Chapter 4 focuses on the initial state of structural assembly on surfaces, the single molecule adsorption, Part II addresses structure formation of α -6T on Au(100) beyond the first monolayer. The promising properties of α -6T as semiconducting material have brought the focus of research on the investigation of molecular assembly of α -6T on electrode surfaces. Various scanning tunneling microscopy studies have elucidated structural assembly and domain formation of α -6T monolayers on metal surfaces.^{80,81,91-95} They found, that the interactions between the molecule and the metallic substrate enforce a completely flat adsorption geometry of α -6T which maximizes the contact area between metal surface and molecule.

In the *all-trans* configuration of α -6T, the alternated alignment of the sulfur atoms results in a C_{2h} symmetry of the molecule and makes the latter prochiral which is illustrated in Fig. 4.3 d: When α -trans-6T adsorbs flat with its mirror plane (σ_h) parallel to the substrate surface two enantiomeric forms can be distinguished according to the orientation of the sulfur atoms of the outermost thiophene rings. If the outermost sulfur atoms point to the right, the enatiomeric form is denoted R-enantiomer and vice versa. For an achiral substrate, both enantiomers are initially formed with equal probability. However, when considering monolayer assembly of prochiral molecules, the clustering of molecules with the same chirality is preferred²⁹³ leading to formation of domains consisting exclusively of the S enantiomer (further denoted as 1S) or the R-enantiomer (1R). Such a formation of enantiomeric pure domains due to organizational chirality has been reported for α -6T on Ag(100),⁹⁵ Ag(110),²⁹⁴ Au(100),⁸⁰ and Au(111).⁸¹

Another consequence of the enforced flat alignment on metal surfaces is that the structure of α -6T in a monolayer is incommensurate with the structural arrangement of the molecule in the organic crystal. In its bulk structure α -6T crystallizes in two phases, a high temperature phase (HT)^{296,297} and a low temperature (LT) phase^{295,297} which is depicted in Fig. 4.3 **a**. In both structures, the molecules are packed in a staggered herringbone fashion that is characteristic for many rod-shaped molecules. Investigations on α -6T thick films reveal that in molecular layers far away from the metal surface α -6T eventually adopts its herringbone bulk structure.^{298,299} So far, the structural transition from the flat monolayer, that is incommensurate to any crystal plane of the α -6T crystal, to the herringbone bulk structure is not yet understood. For its application in opto-electronic devices, the morphology of α -6T in the transition area between flat monolayer and staggered bulk structure is of fundamental importance since its opto-electronic properties are strongly steered by the packing structure directly on the electrode surface⁶⁴ and within the adjacent molecular layer.

The packing structure in the second layer is determined by two competing inter-



Figure 4.3. – a: Low temperature crystal structure of α-trans-6T.²⁹⁵ The four molecules in the unit cell adapt a herringbone structure and are oriented along the [103] direction. b: Different scenarios for the structure of α-trans-6T in the second layer. One ellipse corresponds to one thiophene unit as visualized in c. d: S-enantiomeric and R-enantiomeric configuration after planar adsorption on a flat surface.

actions: the interaction with the metallic surface and the interaction between the molecules in the second layer. The key question in characterizing that structural transition zone is which interaction dominates. Provided that the structure in the second layer is still mainly governed by the interactions with the metallic surface, one would assume a flat second layer which leads to a brick-like structure for the bilayer as described by structure I in Fig. 4.3 **b**. Such a structure was suggested from force field based molecular dynamics³⁰⁰ for the second layer of α -6T on Ag(111).

If the first molecular layer already sufficiently screens the influence of the substrate, molecule-molecule interactions become competitive and the transition to the staggered bulk structure starts already in the second layer. In that case the particular surface of the α -6T crystal that is in contact with the flat first layer decides the structural alignment of the molecules in the second layer. Morphology and crystal growth of organic molecules on substrates are mainly steered by two competing factors, namely the strength of the molecule-surface interactions and the formation energy of the plane of the organic crystal that is in direct contact with the substrate, denoted as crystal contact plane.²⁹⁸ For α -6T, the crystal plane with the lowest formation energy is the (001) surface,³⁰¹ which leads to a quasi-standing orientation of the molecules, followed by the (010) surface³⁰¹ resulting in a quasi-horizontal alignment of the long molecules parallel to the surface plane.

For organic molecules of rod-like geometry, like α -6T, the following general growth behavior has been observed:²⁹⁸ When evaporated on surfaces that interact only weakly with the adsorbate, like the alkali halide (001) surface,³⁰² a quasi-standing molecular geometry can be observed which corresponds to the crystal plane with the lowest formation energy. On inorganic surfaces with stronger molecule-surface interactions, a quasi-horizontal alignment is preferred because it increases the contact area between molecule and surface.^{303–305} In case of α -6T, such a molecular arrangement is realized by the (010) contact plane and leads to a tilted molecular arrangement similar to structure II depicted in Fig. 4.3 **b**.

4. Oligothiophenes on Nano-structured Surfaces

Since the interaction between the metallic surface and the molecules in the second layer are expected to be still of considerable magnitude, a quasi-standing alignment of α -6T on top of the flat monolayer can be ruled out. Instead, it appears realistic to assume a quasi-horizontal alignment of the second layer in case the molecules in the second layer are already sufficiently screened from the metallic surface. In such a quasi-horizontal alignment the molecules in adjacent molecular rows could be either tilted in the same direction as in scenario B or alternately in the opposite direction which leads to a staggered structure described by structure III (Fig. 4.3 b). Such a staggered C-structure on a flat wetting layer was proposed for p-sexiphenyl (p-6P) on CuO,³⁰⁶ another prominent rod-shaped organic semiconductor. Similar adsorption and crystallization behavior of α -6T and p-6P therefore suggest structure C as third possibility for the growth of α -6T on Au(100).

In part II of Chapter 4, the structure of the second layer of α -6T on Au(100) is elucidated by combining scanning tunneling microscopy (STM) data with atomistic calculations. To tackle the huge size of the system consisting of 720 gold atoms (and 528 C/S/H atoms) a hybrid IC-QM/MM method as described in Section 2.3 is chosen that treats the surface on a force-field level of theory while the α -6T molecules are described quantum mechanically. Such an approach allows to focus on the interactions between the first and the second molecular layer while still retaining the role of the surface as a constraint that enforces a flat configuration of the first layer. Subsequently to an extended structure search, STM images of the molecular layer are computed using the Tersoff-Hamann approach.³⁰⁷ The calculated STM images help to link pronounced features visible in the experimental STM image to the local morphology of the molecules in the second layer. The results show that already in the second layer the α -6T molecules overcome the influence of the gold surface and adopt that staggered alignment described by structure III. Based on the found structure, a general growth mechanism for α -6T on gold is discussed.

4.2. Gold Diggers: Altered Reconstruction of the Gold Surface by Physisorbed Aromatic Oligomers

In this chapter electronic structure theory calculations for the adsorption of α -sexithiophene (α -6T) on the 100 surface of gold are presented. The density-functional theory calculations show that α -6T prefers to adjust an energetically unfavorable adsorption site by modifying the gold surface over seeking out more favourable adsorption sites. Molecular adsorption results in a complex charge transfer pattern, with more charge transfer in more stable sites. These results challenge the current paradigm that weakly interacting (e.g. physisorbed) molecules perceive metal surfaces as rigid templates with preordained adsorption sites.

Reprinted with permission from the author's work²⁸⁹ (C2019 American Physical Society).

4.2.1. The Clean Au(100) Surface

An overview over the structure of the clean, reconstructed Au(100) surface is represented in Fig 4.4a. As visible in the STM image, the reconstruction rows with an apparent height of approximately 0.6 Å are oriented parallel to the [110] direction. Perpendicular to the reconstruction rows, elongated areas of slight depressions can be perceived that give rise to a rhombic lattice with a $c(28 \times 48)$ super cell⁸⁴ which is indicated by the blue rectangle. Due to this complex pattern, various different adsorption sites exist that can be broadly classified according to their height into on-ridge and in-valley of the reconstruction rows in either the dark or bright areas of the rhombic pattern. Fig. 4.4 b shows the calculated STM image of the employed Au(100) (20×5) model⁸⁹ computed at a bias voltage of 1V in comparison to the experimentally recorded one.⁸⁴ Similar to the real Au(100) surface, the distance between two consecutive valley in the (20×5) model comprises 5 gold atoms. Furthermore, both STM images exhibit a broader depression running in [110] direction which is less deep than the reconstruction valleys. From the large scale STM image it is visible that the broader depression occurs periodically. At the crossing point of the broad depression and the reconstruction rows the latter ones are shifted one atomic row; this effect is also visible in the computed STM image. Small differences between the experimental STM image and the calculated STM image can be found with respect to the angle between the [100] direction and the broader depression in [110] direction: while in the experimental STM image this angle is roughly 90%, the angle in the calculated STM image deviates slightly from 90%.

A color-coded height image of the 20×5 DFT-model surface is presented in Fig. 4.4 **d** with the ridges of the reconstruction rows depicted in white and the reconstruction valleys depicted in brown. The computed maximum corrugation height for the clean surface is 0.67 Å, in good agreement with experimental findings⁸⁴ The two different adsorption environments *on-ridge* and *in-valley* of the reconstruction rows are encircled in red and blue, respectively.



Figure 4.4. – a: Large scale atomically resolved STM image of the clean Au(100) surface (500 pA, 1 V, 90 K).⁸⁴ The red rectangle depicts the unit cell of the reconstructed Au(100) surface.⁸⁴ The height variation between the valleys (dark brown) and the ridges (white) corresponds to 60 pm. b: Calculated STM image of the bare (20×5) model surface. c: Experimental STM image of the Au(100) surface at atomic resolution displaying in detail the reconstruction rows and valleys. d: color coded 1×2 supercell of the bare (20×5) model surface (brown: valley, white: ridge). The investigated positions are encircled in red (r) and blue (v).

4.2.2. Adsorption on an III-preferred Site: Molecule Alters Surface Reconstruction

In the following, the effect of molecular adsorption on the corrugation of the Au(100) surface is discussed. The two different adsorption environments on-ridge (r) and invalley (v) depicted in Fig. 4.4 **d** were investigated by optimizing the structure of an α -6T molecule placed either on top or in the valley of a reconstruction row. Figure 4.5 **b** depicts the color coded height image of the Au(100) surface after adsorption when the molecule was initially placed in a valley. Compared to the clean surface, the adsorption amplifies the maximum corrugation height to 0.77 Å but does not change the overall reconstruction. The corrugation pattern shows that the molecular adsorption makes the valley slightly asymmetric as indicated by the different height of the adjacent gold atoms on the slope depicted in the height profile in Fig. 4.5 **b**.

For on-ridge adsorption the molecule was expected to slide down into the valley as depicted in scenario A in Fig. 4.2 **b**. Surprisingly, scenario B occurred, in which the adsorbed molecule induced an inversion of the surface reconstruction pattern by pushing the reconstruction ridge aside and creating its own *in-valley* environment at the original on-ridge adsorption site. Figure 4.5 **c** shows the color-coded corrugation of the surface after adsorption of the molecule at the r-position, with the adsorption position of the molecule indicated in red. It can be seen that the surface now features



Figure 4.5. – a-c: Color coded 1 × 2 supercells of the clean model surface (a) and the surface after *in-valley* (v) adsorption (b) and *on-ridge* (r) adsorption (c) and the corresponding height profiles across the adsorption site (indicated by a bar). The investigated positions are encircled in red and blue. Adsorption of the molecule on the ill-favored r-position induces a translation of the reconstruction row and the formation of a reconstruction valley directly underneath the molecule which exhibits a higher symmetry than the reconstruction valley after v-adsorption.

a valley at the adsorption position. Conversely, the initial valley on the surface has turned into a reconstruction ridge. In the following this induced valley-like adsorption environment will be denoted r-to-v position.

Figure 4.5 **a-c** depicts height profiles of the bare surface and the surface after adsorption on the v-position and the r-to-v-position. Despite their considerable similarity, the v-position and the r-to-v-position are not identical. After v-adsorption, the reconstruction valley exhibits a higher asymmetry as the r-to-v-position as manifested in different heights of the adjacent gold atoms that form the slope of the reconstruction row. However, due to the shallow potential energy surface⁸⁹ of the reconstructed gold surface, their adsorption energies are almost indistinguishable, differing by only

Table 4.1. – №	Median adsorption height (Δ h $[{ m \AA}]$) between S and Au atoms, median torsion
a	ingle*, adsorption energy E_{ads} , and charge transfer δQ for $\alpha\text{-}6T$ adsorbed on
t	he reconstructed Au surface. For the adsorption positions v-fix and r-fix, the
g	old atoms were kept fixed during DFT geometry optimization. The calculated
c	harge transfer is based on Mulliken population analysis. ³⁰⁸

Site	Δh S-Au	Δ h C _{α} -Au	Δ h C _{β} -Au	$\phi_{med}[^{\circ}]$	$E_{ads}[eV]$	$\delta \mathbf{Q} \ [e^-]$
V	3.25	3.34	3.16	179	4.34	0.44
r-to-v	3.18	3.22	3.18	177	4.31	0.45
v-fix	3.31	3.45	3.22	177	4.23	0.40
r-fix	3.47	3.49	3.27	168	3.79	0.32

*the median torsional angle of the gas phase molecule is 165°

4. Oligothiophenes on Nano-structured Surfaces

0.03 eV (see Tab. 4.1). The malleability of the gold surface gives rise to a plethora of adsorption environments without distinct and clearly identifiable adsorption sites. That absence of well-defined adsorption sites on the Au(100) surface connects to theoretical studies regarding CO adsorption on small Au clusters that observed a dynamical reordering of surface gold atoms.³⁰⁹



Figure 4.6. – **a-b:** Structure of the molecule adsorbed *in-valley*. The molecule adsorbs completely flat with the C_{α} atoms above the valley (**b**) and, compared to the gas phase molecule depicted in **d**, exhibits a considerably more planar structure that resembles the one of the cationic α -6T, depicted in **e**. **f**: plot of the interring torsional angles between the thiophene units (indicated exemplarily by a red line for the 3-4 torsional angle in **c**.)

After describing the impact of the molecule on the surface structure, the molecular structure will be characterized upon adsorption in terms of adsorption height, median torsion angle, and charge transfer, as summarized in Tab. 4.1. Following the almost equal adsorption energies in the v-position and the r-to-v-position, α -6T exhibits a similar, completely flat adsorption geometry in both positions as represented in Fig. 4.6 **b** for the v-position. The C_{α} atoms are positioned directly above the reconstruction valley and therefore at the largest distance to the surface (3.34 Å and 3.22 Å for v-position and r-to-v position respectively). Sulfur atoms are found mostly on-top of gold atoms with a median adsorption height between the sulfur and the gold atoms of 3.25 Å for the v-position and 3.18 Å for the r-to-v-position.

The more planar adsorption geometry compared to α -6T in the gas-phase is also manifested in larger torsion angles between the thiophene units. Fig. 4.6 **f** depicts a plot of the interring torsional angles of the neutral gas phase molecule (black), the singly charged cation (gold) and the molecule after v-adsorption (blue) and after r-to-v adsorption (blue). For the gas phase molecule the angles are between 161°-166° with a median of 165°. After adsorption on both positions, however, the torsional angles are significantly enlarged and between 175°-180°, with a median of 177° and 179° for the v-position and the r-to-v position respectively. In fact, such planar torsional angles resemble more those of the singly charged α -6T cation, depicted in Fig. 4.6 **e** and suggests adsorption-induced charge transfer. This assumption is confirmed by calculations of the Mulliken charges³⁰⁸ before and after adsorption that reveal a charge transfer from the molecule to the surface for both adsorption positions. The magnitude of the charge transfer is $0.44e^-$ and $0.45e^-$ for the v-position and the r-to-v-position, respectively.



Figure 4.7. – Potential energy interpolation for the pathways I and II as obtained from DFT calculations. a: In pathway I the molecule creates its own valley. b: In pathway II the molecule slides off a gold ridge, whose structure is frozen. c: Color coded visualization of the Mulliken charge difference for the starting structure on-ridge of the fixed surface (r-fix-position) and the final structure in-valley of the fixed surface (v-fix-position). For the molecule adsorbed on the r-fix-position, a smaller magnitude of charge transfer is found, as visible in the brighter red color of the sulfur atoms that corresponds to a lower positive charge.

4.2.3. The Costs to Displace a Ridge

The observation of scenario B inevitably leads to the question why moving many gold atoms out of the way is energetically cheaper for the molecule than sliding down the ridge into the valley. To find the real adiabatic pathways for either process would require complex minimum energy path (MEP) calculations,³¹⁰ which are computationally not tractable for large systems like the one presented here. The pathways were therefore approximated by performing a linear interpolation between the coordinates of the starting and the final structure for the scenarios A and B, respectively. For scenario B, in which the molecule displaces the ridge, the starting (r) and end geometry (r-to-v) was taken and the coordinates then linearly interpolated to sample 9 structures in between. For scenario A, a constrained DFT geometry relaxation for the molecule adsorbed on the r-position was performed. During that structural relaxation the atoms of the gold surface were kept fixed to prevent any restructuring of the surface and only the molecule was allowed to adjust (r-fix-position). Since the molecule stayed in this local minimum instead of sliding down into the valley, a similar constrained DFT geometry relaxation was performed with the molecule placed on the v-position (v-fix-position). Subsequently, the optimized geometries of the molecule on the r-fix and v-fix position were taken as starting and final geometry and a linear interpolation between those coordinates was performed. At each of these intermediate structures for pathway I and II, single point DFT calculations were conducted without further structure relaxation to obtain the binding energy of α -6T (more information can be found in the appendix).

The resulting energy curves are presented in Fig. 4.7 **a** and **b**. Pathway I exhibits no energy barrier. Surface and molecule gain energy by deforming the gold surface. This process is self-propelled and does not need to be activated. Conversely, sliding the molecule down into the valley costs an initial energy of $\sim 0.3 \text{ eV}$. Once this barrier is overcome, the molecule gains a total of 0.44 eV when it reaches the valley. This is much less than the $\sim 2.0 \text{ eV}$ gained in pathway I, indicating that the gold surface will relax at endpoint r-to-v to accommodate the arrival of the molecule. The lateral diffusion of the molecule is certainly a possible alternative process, but expected to be kinetically hindered at low temperatures and accompanied by a response of the gold surface akin to pathway I in scenario B.

The resulting DFT picture of an energetic preference for α -6T in a valley like adsorption environment agrees with our collaborator's experimental findings for α -6T adsorption at very low coverages (0.015 monolayers).¹ The STM image depicted in Fig. 4.8a illustrates that the α -6T molecules, represented as featureless, bright rods, lie well-separated from each other and exclusively in valley positions.

A possible reason for the v-position preference over the r-position could be charge transfer. Therefore, the magnitude of charge transfer was investigated for both positions based on a Mulliken population analysis.³⁰⁸ Since the initial ridge position is not stable against further surface deformation, the surface atoms were constrained and charge transfer upon α -6T adsorption on the constrained valley and the constrained ridge position (v-fix and r-fix) was compared. The Mulliken charge difference associated with adsorption on the r-fix and v-fix position is visualized in Fig. 4.7 c. In both cases, the surface has a higher negative charge after adsorption, most notably around the molecule, as indicated by the blue, ring-like pattern. Moreover, the sulfur atoms become positively charged after adsorption, indicating that the electron transfer from the molecule to the surface proceeds mainly by depleting sulphur orbitals with aromatic carbons being less affected. The total charge transfer from the molecule to the surface amounts to $0.40 e^{-1}$ for the molecule on the v-fix-position, but only $0.32 e^{-1}$ for the molecule on the r-fix position. Hence, the amount of charge transferred is reduced by $\approx 20\%$ for the adsorption on the r-fix-position, which can also be inferred from the lighter red color of the sulphur atoms. This smaller magnitude of charge transfer is accompanied by larger adsorption distances of the sulphur (3.47 Å versus 3.31 Å for the r-fix-position and the v-fix-position) and a less planar conformation of the molecule on the r-site of the fixed surface as expressed by smaller median torsion angles of 169° (see Tab. 4.1).

4.2.4. Discussion

In the following, the results obtained within the DFT framework are compared with experimental observations. While in experiment organic film formation can be viewed as a sequential process of more and more molecules arriving on the gold surface, DFT calculations are constrained to a finite coverage due to periodic boundary conditions. In the present case, one molecule per 20×5 unit cell amounts to a coverage of

24%. When applying periodic boundary conditions, the low energy pathway (I) of an adsorption-induced displacement of the reconstruction ridges becomes a concerted shift of the complete Au reconstruction pattern across the (periodically treated) surface. In contrast, in experiment the adsorption of a single molecule is a non-periodic event. Any adsorption induced ridge displacements will be local to the adsorption site of the molecule. Such local distortions will create defects such as stacking faults or domain walls further afield. The formation of these defects will cost energy that would have to be added to the adsorption energy of a single molecule for pathway I (scenario B) in a more complete model of the reconstruction process. The fact that the STM image (depicted in Fig. 4.8 a for 0.014 ML coverage) does not reveal any local ridge distortions does not rule out their presence, but it indicates that the additional energy cost of ridge distortion and defect creation may add a barrier to pathway I that exceeds that of pathway II (sliding off the ridge). This finding may change with molecule size. In fact, it can be deduced that for the adsorption of longer oligothic phene molecules (α -nT with n>6) ridge displacements will become favorable, since the gain in relaxation energy scales linearly with molecular size, while the defect creation likely scales sublinearly. This assumption is supported by the recent report of a rigid ridge shift around a single adsorbed polymer molecule for poly-3-hexyl-thiophene on Au(001).³¹¹



Figure 4.8. – **a**: Low coverage adsorption of α -6T on Au(100). The molecules adsorb exclusively in *valley* positions.¹ **b**: α -6T monolayer growth on two substrate terraces. Two regions (2 and 4) with different adsorption morphology can be distinguished on each terrace.⁸⁰ **c** apparent height scan along the white line in panel **b**.

Another route towards overcoming the energy barrier for defect formation is to increase the molecular coverage to a dense monolayer. In that case, the concerted action of the molecules in self-assembled domains can even accomplish a temperature-facilitated partial lifting of the surface reconstruction as observed in STM measurements conducted at elevated temperatures.⁸⁰ Figure 4.8 b depicts a large scale STM image of α -6T at monolayer coverage at 387 K.⁸⁰ Aside from large monolayer domains separated by a substrate step edge (area 1 and 3) a stripe-like smaller region exhibits a different structure on each substrate terrace (area 2 and 4). Fig. 4.8 c depicts the apparent height profile of the domains extracted along the white line. It shows that the molecules in the two stripes 2 and 4 are approximately 0.07 nm \pm 0.01 nm lower in height than the molecules in the large domains that are on the same substrate terrace,

i.e. regions 1 and 3, respectively. 0.07 nm is exactly the aforementioned corrugation height of the reconstructed Au(100) surface compared to the unreconstructed Au(100) surface. This strongly indicates that in the areas 2 and 4 the α -6T molecules have modified the gold surface and lifted reconstruction. Moreover, the α -6T film structure in the areas 2 and 4 significantly differs from that in the areas 1 and 3. In both structures, α -6T molecules are packed in molecular rows, but in the areas 1 and 3 the long molecular axis of the molecules is oriented almost perpendicular to the direction of the molecular rows, whereas in areas 2 and 4 it is at a 30° angle.⁸⁰ This rotated structure strongly resembles the adsorption pattern of α -6T on Ag(100),⁹⁵ a surface that does not have any reconstruction.

Moreover, for the adsorption of styrene on Au(111) an adsorption-induced modification of the periodicity in the Au(111) reconstruction even at very low temperatures was proposed.⁷⁸ An extension of the surface herringbone reconstruction was also observed for the physisorption of dense layers of α -6T⁸¹ as well as Azure-A⁷⁹ on Au(111); in the latter case this effect was traced back to the cationic nature of the adsorbate, which enables stronger molecule-surface interactions combined with the presence of chloride anions that are known to affect the reconstruction of gold surfaces.³¹² These findings suggest that the unusual behavior of α -6T in Au(100) is not restricted to this particular molecule, but could be generalized to other, weakly adsorbing molecules.

4.2.5. Conclusion

In this section, the first comprehensive first principles investigation of aromatic oligomers on large, reconstructed Au surfaces is presented. The results reveal that α -6T does not move off energetically ill-favored on-ridge positions, but instead moves the ridge out of the way to create a preferred valley in which the molecule is then accommodated. While for weak adsorbates metal surfaces have been seen as immutable constraints that could be used to steer molecular adsorption, these findings illustrate that the picture is more complicated. The Au surface must instead be understood as soft and deformable even when only in contact with physisorbed molecules, and the notion of distinct adsorption sites does not apply.
4.3. From Flat to tilted: Oligothiophene Growth beyond the Monolayer

In this work, domain formation and local morphology of α -sexithiophene (α -6T) on Au(100) beyond monolayer coverage is investigated by combining high resolution scanning tunneling microscopy experiments with hybrid image-charge quantummechanics/molecular mechanics calculations. α -6T is a prominent organic semiconductor with real-life application whose opto-electronic properties are strongly determined by its packing structure in the device. The results presented in this section reveal, that already in the second layer the molecules abandon the flat adsorption structure of the monolayer and adopt a tilted conformation. Although that observed tilted arrangement resembles the orientation of α -6T in the bulk, the observed morphology does not yet correspond to a well-defined surface of the α -6T bulk structure. Instead, the second layer acts as a structural transition zone between the flat arrangement in the monolayer and the staggered orientation in the bulk structure.

4.3.1. Characterization of the Chiral Domains



Figure 4.9. – a: 1.5 monolayer coverage ($U_T = 2.2 V$, 80 K) on four substrate terraces (I-IV).¹ Two domains can be observed on terrace II and IV respectively, that differ in their growth direction with respect to the [1-10] direction.¹ b and c: High resolution STM image of the two monolayer domains. Both domains consist of exclusively of either the S-enantiomer (b) or the R-enantiomer (c) of α -6T and differ in their orientation with respect to the [110] direction (characterized by the angle α) and [1-10] direction (characterized by the angle β). Based on those angles, the chirality of the second layer can be determined.

Fig. 4.9 **a** shows a large scale STM image of a 1.5 monolayer coverage of α -6T on Au(100) on four different substrate terraces (I-IV) with terraces II-IV being sufficiently resolved for analysis. On terrace II and III growth of two different, densely packed monolayer domains of α -6T can be observed which are further denoted as 1S (terrace II) and 1R (terrace III). The α -6T monolayer grows in a Frank-van-der-Merwe mode, i.e. with increasing deposition the formation of a full monolayer instead of three

dimensional islands can be observed. The molecules adsorb flat and in well-ordered molecular rows with the long molecular axis (LMA) approximately parallel to the [110] direction and almost perpendicular to the direction of the molecular rows.

The two monolayer domains differ in their growth direction of the molecular rows with respect to the [110] direction as well as the orientation of the long molecular axis of the individual molecules with respect to the [1-10] direction. The orientation of the individual molecules and the direction of the molecular rows can be characterized by two angles α and β , respectively, as visualized in the two insets in Fig. 4.9 **b** and **c**. In the 1S domain, the angle α enclosed by the [110] direction and the LMA of the molecules (blue dotted line) and the angle β between the growth direction of the molecular rows (indicated by a white arrow) and the [1-10] direction are both < 0. In the 1R domain, however, the angle α as well as the angle β are both > 0.

Fig. 4.9 **b** shows a high resolution image of such a 1S domain (the corresponding large scale STM image can be found in Ref. 1). The molecules are clearly resolved and comprise six bright protrusions that correspond directly to the six thiophene units of the molecule and are arranged in a zig-zag pattern. Interestingly, all molecules in domain 1S exhibit the same orientation of the zig-zag pattern that can be assigned to an adsorbed α -6T molecule in the S-enantiomeric form.⁹⁵ Hence, domain 1S contains exclusively α -6T molecules in their S orientation. The molecularly resolved STM image of a 1R domain is shown in Fig. 4.9 c. Compared to Fig. 4.9 b, the zig-zag patterns of all molecules are exclusively oriented into the opposite direction and therefore domain 1R comprises only α -6T molecules in the R-enantiomeric form.

In both domains, the surface unit cell contains one molecule and similar unit cell vectors of |a| = 25.9 Å and |b| = 6.5/6.8 Å (1S/1R). The different orientation of the molecular rows in the two domains, however, is reflected in different cell angles of the surface unit cell as presented in Tab. 4.2. In the 1S domain the unit cell angle is smaller than 90° (roughly 87°), whereas in the 1R domain the unit cell angle is larger than 90° (95°). That deviation of the cell angle from 90° leads to a small translation of the molecules in the adjacent row. As a result the terminal C_{α} atoms of the molecule in one row are placed between two adjacent molecules in the next row, which is highlighted by a red rectangle in the inset of Fig 4.9 b. This small translation of the molecular rows within one layer, further denoted as $\Delta_{\text{row}}^{\text{intra}}$, facilitates a closer contact between molecules of two adjacent rows and hence increases the packing density, as discussed for α -6T on Au(111).⁸¹

On terrace II and IV growth of island-like domains of α -6T on top of the monolayer can be observed. With an apparent height of ≈ 3.1 Å, the domains that grow on the monolayer consist only of one layer of molecules, indicating a continuation of the layer-by-layer growth mode (Frank-van-der-Merwe growth) with the second layer.¹ On terrace II, highlighted by a blue rectangle enclosing terrace I and II, the growth behavior of the second layer domain at step edges and domain borders of the monolayer can be studied. It is visible that the growth of the bilayer on terrace II stops at the I/II and II/III step edges. In addition, the bilayer grows only on the 1R domain (monolayer) of terrace II and does not extend beyond the domain border.

Similar to the first layer, the molecules in the second layer assemble in two enantiomeric pure domains that can be distinguished by the characteristic angles α and β that describe the orientation of the individual molecules and that of the molecular rows. For the second layer domain on terrace II, both angles are similar to that of the underlying 1R domain, implying a 2R second layer on top of a 1R monolayer (2R@1R bilayer structure). On terrace IV, the orientation of the molecular rows and individual molecules in the second layer domain is equal to that of the 1S domain on terrace II; evidently that second layer is of S chirality (2S). Unlike to the second layer domain on terrace II, the chirality of the underlying monolayer domain cannot be determined.

Table 4.2. – Measured lattice constants [Å] and angle [°] of the unit cell of monolayer of α -6T on Au(100)¹ in comparison with lattice constants of the α -6T-(010) terminated surface.

Structure	a	b	γ
1S	25.9	6.5	87 ± 1
$1\mathrm{R}$	25.9	6.8	95 ± 1
$6T(010)^{313}$	44.708	6.029	89.4



Figure 4.10. – a: STM image of a 2S domain on a 1S domain (2S@1S).¹ The image indicates a shift of the second layer molecular rows in the lateral direction parallel to the [110] direction (green lines) and along the long molecular axis of the individual molecules (red and blue line) with respect to molecules in the first layer. b: Visualization of the lateral shift. c: View along the molecular rows of the computed 2S@1S structure confirming the lateral shift observed in the STM images (≈ 2.5 Å). d: View along the LMA of the computed 2S@1S structure exhibiting a median shift of 2.3 Å along the LMA of the molecules.

4.3.2. Lateral Position of the Second Layer Molecules

After characterizing the global structure of the bilayer domains, the local morphology and structural features of the individual molecules in the different domains will be discussed in the following by the example of a high resolution image of a 2S domain presented in Fig. 4.10 **a**. It can be seen that the orientation of the individual molecules as well as the molecular rows in the underlying monolayer is equal to that of the second layer. Hence, Fig. 4.10 **a** depicts a 2S@1S bilayer structure. Snapshots of the corresponding computed 2S@1S bilayer are presented in Fig. 4.10 **b**-d.

As discussed in the previous section, two molecules in two consecutive rows in the monolayer are not perfectly aligned but slightly translated by a shift $\Delta_{\text{row}}^{\text{intra}}$ which is indicated by a red rectangle in Fig. 4.10 b. As a consequence, molecules in the second layer that were located directly above the molecules in the first layer should exhibit as well such a translation in row direction compared to the molecules in the adjacent row in the first layer. This expected position is visualized by a dark green line in Fig. 4.10 b. Comparing the positions of the molecules in the first and in the second layer of two adjacent rows, however, it becomes visible that they are directly aligned and lack the expected translation in row direction. The actual position is indicated by a bright green line in Fig. 4.10 **a** and Fig. 4.10 **b**. This implies that the molecules in the second layer are not directly on top of the molecules in the first layer but laterally shifted.

This lateral shift, further abbreviated as Δ_{lat} was also confirmed and quantified by the performed QM/MM calculations. Fig. 4.10 **c** shows a view along the molecular rows of the computed 2S@1S structure. The lateral shift of the long molecular axis of the second layer molecules is clearly visible and leads to a brick-like arrangement of the molecules in the second layer. With 2.5 Å, the median lateral shift is in agreement with the experimentally found one of roughly 3 Å (see Tab.4.3).

Aside from the the lateral shift, Δ_{lat} , the molecules in the second layer are also translated in direction of the LMA as indicated by the red and blue dotted lines in Fig. 4.10 **a**. The red dotted line marks the end of one molecular row and the beginning of the next molecular row in the first layer, whereas the blue dotted line indicates the end of one molecular row in the second layer. It can be seen that the blue line is shifted with respect to the red line. Hence, the molecules in the second layer are not exactly on top of the molecules in the first layer but are shifted $\approx 3 \text{ Å}$ along their LMA with respect to the molecules in the first layer.

The computed 2S@1S structure is presented in Fig. 4.10 **d** which depicts the molecules with their LMA parallel to the paper plane. The end of one row in the first and second layer is highlighted by red and blue circles, respectively. It can be seen that, in agreement with the experiment, the upper layer molecules are shifted along their LMA which is indicated by the blue arrow. The median shift along the LMA of the upper layer molecules is ≈ 2.3 Å and will further be denoted as Δ_{LMA} . The shifts Δ_{LMA} and Δ_{lat} are also present in the bilayer of opposite chirality, 2R@1R (see Tab. 4.3) which indicates that those two shifts are universal, characteristic features of the α -6T second layer regardless of the chirality of the domains.

4.3.3. Molecular Orientation in the Second Layer

The investigation of the molecular orientation with respect to the surface normal reveals that the molecules in the second layer are arranged as in scenario III, see

Table 4.3. – Computed bilayer structures, their observation in the experiment, their relative energy and structural features: $\Delta E \ [kJ \ mol^{-1}]$ is the energy difference per molecule α -6T of the respective structure with respect to the 2S@1S IIIstructure ($\Delta E = E_{structure} - E_{2S@1S-C}$), Δ_{LMA} [Å] is the shift along the LMA, Δ_{lat} [Å] is the shift of the molecules in lateral direction, and ϕ [°] is the tilt around the LMA of the second layer molecules.

structure		in Fig. 4.9	ΔE	$\Delta_{\text{lat}}[\text{\AA}]$	$\Delta_{\rm LMA}$	ϕ
2S@1S	III	yes	0.0	2.5	2.3	9-17
2S@1S	II	no	0	2.2	0.2	10 - 12
2R@1R	III	yes	2	1.9	0.8	9-17
α -6T bulk	II	-	-	2.7	1.2	33

Figure 4.3 **b**. This conclusion can be derived from high-resolution STM measurements in combination with computational structure search and calculations of the electronic density discussed in the following.

The high resolution STM image of a 2R island displayed in Figure 4.11 **a** shows a distinct protrusion pattern for the molecules in the second layer. Details of the molecular electronic structure are visible since the first layer decouples the electronic structure of the second layer from the surface. This is a known effect exploited previously by using NaCl films for imaging molecular orbitals.⁹⁶ On the contrary, high resolution images of the monolayer display the molecules as featureless, bright rods. since the direct contact with the metal surface leads to a perturbation and broadening of the molecular electronic states. One molecule in the second layer is represented by six droplet-shaped, well separated protrusions of different brightness followed by a blurry extension. In Figure 4.11 **a**, the protrusions are indicated by blue marks and the blurry extensions by a black line. The six droplet-shaped protrusions are arranged in three pairs. The brightness pattern of the protrusions and the blurry extension extend the periodicity of the STM from one to two molecules per unit cell which are denoted in the following as M1 and M2. This is the first indication that molecules are staggered as in scenario III. For structures I and II, the unit cell should contain only one molecule.

In order to assign the protrusion pattern in the experimental STM image to structural motifs in the molecules, the probability density ($\|\Phi_{\text{HOMO}}\|^2$) of the highest occupied molecular orbital (HOMO) is computed. The distribution of the electron density of the HOMO of an isolated α -6T molecule is displayed in Figure 4.11 **b** and exhibits several nodal planes: one within the molecular plane and eleven perpendicular to the latter passing through the center of the thiophene units and through the interring bridges. In the following, the six thiophene units for molecules M1 are denoted with T1-T6 and for molecule M2 with T1'-T6' (see Figure 4.11 **b**). For each thiophene unit a pair of droplet-shaped lobes can be observed. For molecule M1, the density distribution of units T2, T4 and T6 resembles in symmetry and shape the protrusion pattern in the experimental STM, while for molecule M2 the resemblance is found for units T1', T3' and T5'. The other thiophene units are not resolved in the experimental image.

Before discussing why only three out of six thiophene units appear in the STM

4. Oligothiophenes on Nano-structured Surfaces



Figure 4.11. – a: High resolution STM image of the 2R second layer measured at a bias voltage of -1.35 V showing a protrusion pattern and two molecules in the unit cell labeled as M1 and M2.¹ b: Computed density of the HOMO for an isolated α-6T molecule. The molecule is oriented as in the 2R@1R structure with the sulfur of the first thiophene unit (T1, T1') pointing to the right. c: Computationally optimized 2R@1R structure exhibiting a staggered arrangement as in scenario III. d and e: Top-view of the optimized bilayer and corresponding color-coded profile of the relative height of the atoms in the second layer (e). The lowest atom is set to 0. g: Calculated STM image of the 2R@1R structure at an isovalue of 10⁻⁴ (d) and 10⁻⁶ (e).

image, the electronic structure of a thiophene monomer is discussed in more detail. Thiophene has six p electrons: four from the C atoms and two from the sulfur atom. The first three occupied states have π -character and are occupied by the p electrons. To gain insight in the buildup of the molecular π orbitals, the molecular orbitals of a single thiophene have been computed and plotted (see Appendix A.2.3.3). As shown in Figure A.7 in the Appendix, the p orbital of sulfur is not contributing to the HOMO. The same is true in the oligomer, which is discussed in terms of of symmetry arguments in Ref. 314. This is in agreement with the plot of $\|\Phi_{\text{HOMO}}\|^2$ for the isolated α -6T molecule (Figure 4.11b), which shows a reduced density at the sulfur atoms. The alternating pattern of visible and non-visible thiophene units in the STM image indicates height differences between the C_{β} atoms of adjacent units. The comprehensive computational structure search found indeed a tilted, staggered

arrangement as shown in Figure 4.11 c. The computed tilt angles ϕ are 9° – 17° and the height profile of the molecules in the second layer is presented in Figure 4.11 e. For molecule M1, the C_{β} atoms of the units T2, T4, and T6 point upwards and the sulfur atoms downwards, whereas the opposite holds for the units T1, T3 and T5. Molecule M2 is tilted in the opposite direction and C_{β} atoms of T1', T3', and T5' are directed away from the surface. Since the HOMO has the largest density at the C_{β} atoms and due to the strong decrease in the tunneling probability with increasing distance, the HOMO is only visible for the units where the C_{β} atoms point towards the STM tip. This is directly confirmed by the computed STM image shown in Figure 4.11 **f**. The units with C_{β} atoms tilted upwards appear as a bright pair of droplet-shaped lobes. The other units are significantly darker.

The protrusions in the experimental image that have been assigned to the units T6 and T1' appear on average darker than the other protrusions. This variation in brightness can be observed to some extent also in the computational STM images. Some of the T6 and T1' units are bent towards the surface as evident from the height profile in Figure 4.11 e. As a result, they are noticeably darker in the STM. This becomes more obvious when plotting the computed STM at a lower isovalue (10^{-6}) to facilitate comparison to experiment, see Figure 4.11 g. As in the experimental image, the units T2, T4, T3' and T5' dominate in brightness. Furthermore, they superimpose the shape of the lobes originating from the thiophene units where the C_{\beta} atoms point downwards. In addition, the lobes of the terminal units T1 and T6' have no longer a distinct shape resembling the blurry extension in Figure 4.11 a.

In summary, the following four points are strong evidence that the molecules in the second layer are arranged in a staggered configuration (scenario III): i) There are two molecules in the unit cell. ii) The computed HOMO density of the α -6T molecule shows droplet-shaped lobes organized in pairs, which is also observed in the experiment. For a flat structure six pairs are expected. However, only three out of six appear in the STM. iii) The structure optimization yields a staggered arrangement and iv) the corresponding computed STM image resembles strongly the experimental image, in particular, when plotted at lower currents. Note that the electronic structure of the metal is not explicitly accounted for in our computational model, which confirms that the second layer is decoupled from the metallic surface.

The computational optimization procedure yields also a staggered structure for the 2S@1S bilayer. The structural features are similar to 2R@1R, see Table 4.3. Interestingly, the computational structure search also resulted in a 2S@1S structure with a configuration as in scenario II, i.e., the molecules in the second layer are always tilted in the same direction, see Figure A.8 (Appendix A.2.3.3). Both structures, 2S@1S-II and 2S@1S-III, are equal in energy. For 2R@1R, a structure of type II has not been obtained. Furthermore, there is no experimental evidence for equally tilted arrangements for any of the bilayers.¹ It might be indeed that structures II and III are similar in energy. In fact, scenario II resembles more closely the α -6T bulk structure than III, as will be discussed in Chapter 4.3.4. However, the computational model applies periodic boundary conditions and is too small to capture domain effects. Even though the wave functions in the second layer reflect those of molecular orbitals, the hybridization with the surface is not zero. Both effects might tip the energy balance



towards structure III.

Figure 4.12. – Relaxed 2R@1R (a), and 2S@1S III structures (b), which are all approx. equal in energy. c: α -6T-(010) surface. The tilt angle ϕ and the vertical translation, Δz of the molecules are highlighted by a green rectangle and two yellow lines, respectively. e and f: Visualization of the shifts Δ_{lat} (e) and Δ_{LMA} (f) in the α -6T bulk structure.

4.3.4. Transition to the Bulk Structure

In the following, the transition from the flat monolayer into the herringbone structure that α -6T adopts in the crystal is discussed, based on the structural motifs found for the second layer. The structure of the second layer of α -6T on Au(100) already begins to resemble the packing in the (010)-terminated molecular crystal, which becomes apparent from a comparison of the computed 2R@1R and 2S@1S bilayer structures depicted in Fig. 4.12 **a** and **b** in comparison structural features of the α -6T (010)-oriented bulk structure as presented in Fig. 4.12 **c**-**e**.

The atomistic (010)-oriented α -6T bulk structure is presented in Fig. 4.12 c. The AB stacking of the molecules is visualized by different colors, blue for the B layer, orange for the A layer, with the molecules in the second row being depicted in a lighter color. Within one layer, the first molecule is shifted in z-direction compared to the second one. This vertical translation is highlighted by two yellow lines in Fig. 4.12 c and will be further denoted as Δz . Moreover, the molecules are tilted by $\phi \approx 33^{\circ 295}$ with respect to the (010) plane. All molecules in one layer are tilted in the same direction, similar to structure II depicted in Fig. 4.3.

Fig. 4.12 **d** and **e** depict the structural motifs of the α -6T (010)-terminated crystal in more detail. In addition to the molecular tilt that was already visible in Fig. 4.12 **c**, a lateral shift, Δ_{lat} , of ≈ 2.7 Å and a shift along the LMA Δ_{LMA} of ≈ 1.2 Å can be observed. The lateral translation, Δ_{lat} , of the computed and experimentally measured bilayer structures already resembles that of the α -6T molecules in the bulk structure, namely 1.9 Å for the computed 2R@1R structure, 2.5 Å for the computed 2S@1S structure, and ≈ 3 Å for the experimentally measured 2S@1S structure (see Table 4.2). The molecular shift along the LMA, Δ_{LMA} , of the measured and computed 2S@1S bilayer structures amounts to ≈ 3 Å and ≈ 2.3 Å and is therefore significantly increased compared to the bulk structure (≈ 1.2 Å). This enhanced shift can be explained by the presence of the metallic surface: a lateral shift of ≈ 2.3 Å - 3 Å locates the outermost thiophene unit directly above the gap between two molecular rows in the first layer, enabling some of the molecules to bend the first thiophene unit downwards and thereby to maximise the interaction with the surface. Such a bent conformation of the first thiophene unit of some molecules has been observed to some extend in the computed structures (see Fig. A.9 in Appendix A.2.3.4) and occurs most likely also in the experimentally determined structure as derived from the brightness variation in the STM image of the 2R structure (see Fig. 4.11 and Section 4.3.3).

Despite the appearance of characteristic structural features of the α -6T (010) oriented bulk structure in the 2*R* and 2*S* layers, the arrangement of the molecules in the second layer does not yet coincide with that of the molecules in the α -6T bulk structure. The structural differences comprise the lattice constants of α -6T in the second layer and the α -6T crystal, the vertical translation, Δz , and the orientation and magnitude of the molecular tilt, ϕ .

For the the surface unit cell of the second α -6T layer, the experiment measured increased lattice constants compared to the high temperature bulk structure, namely $|a| \approx 51.8$ Å and $|b| \approx 6.5$ Å/6.8 Å for the experimentally determined 2R@1R structure and 2S@1S structure, respectively.¹ In comparison, the lengths of the surface unit cell vectors of the α -6T (010) surface are |a| = 44.708 Å and |b| = 6.029 Å²⁹⁵ (see Tab. 4.2). Such an enlarged surface unit cell, that has also been observed for a similar, rod-shaped like molecule,³¹⁵ can be explained by the flat monolayer structure that is incommensurate with any possible contact plane of the α -6T molecular crystal. The lattice mismatch between the unit cell of the flat monolayer and the α -6T (010) oriented cell enforces a coherently strained surface unit cell of the consecutive layer and leads to elongated surface unit cells in α -6T thin films. In the transition to thicker films, that strain energy associated with the elongation of the unit cell will accumulate and reach a critical point where the formation of dislocations becomes energetically more favorable. By creation of misfit dislocations, the lattice constant of the α -6T molecules in the film will eventually approach that of the α -6T bulk structure.

The flat monolayer structure is also responsible for the absence of the vertical translation, Δz , in the computed bilayer structures and the experimental observations, since the planar conformation of the molecules in the monolayer enforces a similar adsorption height of the molecules in the second layer.

Compared to the (010)-oriented α -6T surface, the tilt angle, ϕ , of the molecules in the second layer of the bilayer structures differs in magnitude and orientation. In the (010)-oriented bulk structure, the molecules in one layer are tilted in equal direction, similar to structure II depicted in Fig. 4.3 **b**, while the measured and computed 2R@1R structure and 2S@1S structure exhibit a staggered arrangement similar to structure III in Fig. 4.3 (b). Moreover, the tilt angle of the molecules in the second layer amounts to 9° – 16° which significantly reduced compared to the tilt angle of the molecules in the bulk structure (33°). The alternately tilted arrangement of the second layer molecules in the 2S@1S and 2R@1R structure III resembles a hybrid structure of layer A and B of the α -6T (010)-plane respectively, does not yet fully

4. Oligothiophenes on Nano-structured Surfaces

correspond to the molecular packing in the bulk.

To reach an ordered α -6T bulk structure, a misfit dislocation has to be created after a critical layer thickness, as in case of the lattice constant. This misfit dislocation comprises a vertical translation that shifts every second molecule by $|b| - \Delta z$. As a result, the molecules within one layer will adopt equally tilted configurations and be vertically shifted by a Δz that is similar to that of the bulk structure. With progressing layer thickness and thus decreasing influence of the metallic surface, we also expect the magnitude of the tilt angle to increase and converge towards that of α -6T on the bulk structure.

Interestingly, for p-sexiphenyl (p-6P), a rod-shaped organic semiconductor with a similar packed bulk structure,^{316–319} a staggered second layer structure on top of a flat monolayer was found on $Cu(110)^{306,315}$ and Ag(001).³²⁰ An identical growth behavior on metallic substrates might therefore be expected for all similar, rod-like organic molecules: While on non-metallic surfaces such rod-shaped molecules exhibit either a quasi-standing or a tilted molecular geometry, corresponding to a well-defined crystal plane, the comparably stronger interactions with the metallic surface enforce a flat lying monolayer that is not necessarily commensurate with any contact plane of the organic crystal. Upon increasing molecular coverage the influence of the metallic substrate gets successively screened, the molecule-molecule interaction in the consecutive layer becomes dominant and the molecular structure progressively merges into a well-defined crystal plane of the organic bulk.

4.3.5. Conclusion

The investigation of the second layer morphology of α -6T on Au(100) by high-resolution STM measurements in synergy with hybrid QM/MM calculations reveals two chiral domains, 2R and 2S, that share three characteristic structural motifs. The first two motifs comprise a lateral shift and a shift in direction of the long molecular axis of the molecules in the second layer with respect to the molecules in the first layer. As third characteristic motif, a tilt of the molecules around the long molecular axis was observed, that led to a staggered arrangement of the molecules in the second layer. While these three characteristic structural features lead to a strong resemblance to packing in the (010)-oriented α -6T bulk structure, none of the found structures coincides with a well-defined plane of the α -6T crystal. Hence, the second layer should be regarded as a structural transition zone where the attempt of the α -6T molecules to adsorb with a specific contact plane on the flat monolayer has not yet outbalanced the effect of the metal surface. However, in the transition to thicker films, we expect the α -6T molecules to adopt progressively the well-ordered arrangement in the (010)-terminated bulk structure.

Together with reports of a similar bilayer structure for the related rod-shaped organic semiconductor p-6P on Cu(110)^{306,315} and Ag(001)³²⁰ our findings for α -6T on Au(100) encourage the assumption of a unified growth mechanism that could be generalized to all similar rod-shaped organic semiconductors on arbitrary metal surfaces.

5. Summary

This thesis investigates the structural and dynamical features of organic, functional molecules under different constraints applying force field based molecular dynamics, density functional theory calculations, and hybrid quantum mechanics/molecular mechanics approaches.

Project I addresses the conformational fluctuations of the chromophore phycocyanobilin (PCB) in the second GAF domain of the cyanobacteriochrome AnPixJ, AnPixJg2. AnPixJg2 belongs to the larger class of photosensory proteins which trigger biochemical processes in response to electromagnetic radiation. For this purpose, they incorporate chromophores that are able to reversibly switch between two conformational states with different absorption properties:²²⁹ a biologically inactive state, usually denoted "dark stable state" and a biologically active state, the photoproduct state.

Based on force field molecular dynamics simulations, two structural sub-states of the chromophore in the dark stable state of AnPixJg2 are characterized in **Project** I, from which only one was previously known from crystallographic data. These two sub-states differ in the orientation of the *D*-ring. In the structural sub-state I, which is equivalent to the conformation of the chromophore in the crystal structure, the D-ring is oriented above the tetrapyrrole ring plane, while in the structural sub-state II the D-ring is rotated below the tetrapyrrole ring plane. Both sub-states interconvert on time scales of 10 ns - 500 ns. This theoretical finding is corroborated by NMR data, which report a signal splitting in the ${}^{13}C$ MAS NMR suggesting the presence of a heterogeneous mixture of *D*-ring isomers.²⁷² Moreover, investigations on the solvation of the chromophore binding pocket and the flexibility of amino acids adjacent to the chromophore contributed to the understanding of the mechanism behind the different absorption properties of dark stable state and photoproduct state. Previously, two contrary mechanism were discussed, the "hydration model"²⁶¹ and the "trapped twist model".³²¹ The former one explains the shift in the absorption maximum by a change in the solvation pattern that is mediated by the presence/absence of a bulky tryptophane residue that shields the chromophore binding pocket.²⁶¹ The latter one correlates the different absorption properties in the photoproduct state with a strongly twisted geometry of the chromophore that distorts the conjugation of the system.³²¹ The results from force field molecular dynamics simulations presented in this thesis reveal a highly dynamical behavior of this tryptophane residue that is in contradiction to a potential role in shielding the chromophore binding pocket. Moreover, quantification of the amount of solvent molecules shows a considerably hydrated chromophore binding pocket in the dark stable state which rules out the hydration model. These findings are in agreement with currently published results stating that the absorption spectrum is exclusively dominated by the effective conjugation length in the chromophore²⁷⁰

5. Summary

In summary, the findings presented in the first part of **Project I** demonstrate the ability of molecular dynamics simulations to access and characterize structural sub-populations which cannot be observed in crystallographic measurements due to crystallographic confinement, i.e. low temperature, and the absence of solvent molecules. Together with findings of structural heterogeneity in other, functionally related photosensory proteins,^{11,35–38,47} these results further accumulate evidence that structural heterogeneity is a unifying feature in CBCRs and related photosensory proteins.

In the second part of **Project I**, the effect of a conserved tyrosine residue on the conformation of the chromophore has been elucidated by *in silico* mutation against a non-polar phenylalanine residue. An extensive force field molecular dynamics simulation on the us time scale revealed a third structural sub-state, sub-state III, which was not observed in the wild type and therefore, most likely, has been induced by the mutation. In sub-state III, both, the A-ring and the D-ring, are oriented below the tetrapyrrole ring plane. Persistent hydrogen bond interactions between the A-ring and adjacent amino acids that are established in sub-state III, stabilize the downward rotated conformation of the A-ring on the time scale of 1 µs and prevent conversion back into sub-state II or sub-state I. Mutation of TYR302 also affects the conformation and rigidity of the C-ring propionate chain which persistently adopts a stretched, stable configuration in the wild time. Transition into sub-state III in the mutant, however, leads to an increased disorder of the C-ring propionate chain configuration which dynamically fluctuates between a stretched configuration and a downward bent configuration. This dynamical behavior of the C-ring propionate side chain in the absence of TYR302 illustrates that TYR302 is responsible for stabilizing the conformation of the propionate side chain, and not the adjacent arginine residue as previously assumed.³³ The results presented in the second part of $\mathbf{Project} \ \mathbf{I}$ highlight the role of TYR302 in stabilizing the structural sub-states I and II in the parent ground state of AnPixJg2. Since this tyrosine residue is highly conserved among the different subgroups of phytochromes and CBCRs³³ the findings presented in this thesis are also relevant for other CBCRs and phytochromes.

While **Project I** focuses on the conformation of a chromophore under the influence of the soft and flexible constraints provided by the protein environment, **Project II** elucidates structural assembly of organic, semiconducting molecules on well defined, nano-structured surfaces. In the first part of **Project II**, single molecule adsorption of the oligothiophene α -sexithiophene (α -6T) on the reconstructed Au(100) surface is investigated by all-electron DFT calculations.

Due to its large scale reconstruction, the Au(100) surface exhibits a complex height modulation consisting of reconstruction rows which offer different adsorption sites. These sites, which can be broadly classified into on ridge and in valley of the reconstruction rows, differ with respect to the adsorption strength of the molecule. The computational results obtained in **Project II** show that α -6T prefers to adsorb in valley of the reconstruction rows. Moreover, the analysis of charge transfer processes between molecule and surface suggests that this preference is mainly caused by an increased magnitude of charge transfer between the molecule adsorbed in-valley of the reconstruction rows compared to more unfavorable sites. Surprisingly, when placed on the ill-favored site on the ridge of the reconstruction rows, the molecule does not slide off the ridge during reconstruction. Instead, it alters the reconstruction pattern by shifting the reconstruction row aside and creating itself the preferred *in-valley* adsorption environment. These findings challenge the established concept of well-defined adsorption sites and illustrate that the gold surface must instead be seen as soft and deformable.

In the second part of **Project II**, structural assembly of two layers of α -6T on Au(100) is investigated by hybrid quantum mechanics/molecular mechanics calculations in combination with experimental scanning tunneling microscopy (STM) measurements. In the first layer, the direct contact with the metal surface strongly disturbs the electronic structure of the adsorbing molecules and as a result the molecules appear as featureless, bright rods in the STM image. In the second layer, however, the underlying first layer leads to a decoupling of the electronic states of molecules and metal and individual features of the molecular orbital density can be studied by STM. The results presented in this thesis demonstrate the potential of theoretical structure searches followed by the calculations of STM images to aid with the interpretation of those features in experimentally measured STM images and link them to the atomistic structure of the molecules on the surface. With this joint experimental and theoretical approach, the growth of α -6T beyond the first molecular layer on the gold surface is studied. The findings obtained in **Project II** show that the molecule-molecule interactions become significant already in the second layer. Due to this increased influence of the molecule-molecule interactions, the molecules abandon the flat adsorption structure found for the monolayer and start to adopt a staggered alignment that resembles the conformation of the molecules in the α -6T crystal. This thorough characterization of the local structure of α -6T layers on the gold surface will open new perspectives for understanding structural assembly of similar rod-shaped organic semiconductors on metal surfaces. In combination with future work targeted on the electronic properties of the α -6T layer on gold, the structural model established in this work will contribute to the elucidation of structure-property relations of oligothiophenes on metallic substrates.

A. Appendix

A.1. Appendix to Project I

A.1.1. Methods

A.1.1.1. Computational Details

The initial starting geometry for the Pr state of AnPixJg2 was taken from the crystallographic structure of AnPixJg2³³ (PDB code: 3W2Z). Hydrogen atoms were added using the VMD tool PSFGEN 1.6.2³²² and the titration state of all ionizable amino acid side chains was set to correspond pH 7.0. For histidine, which has a pKa of 6.0 the protonation state was determined by visual inspection of the chemical environment, which is the usual procedure.^{46,261} Regarding the two HIS residues in the chromophore binding pocket, HIS322 was protonated at the ϵ - and δ -nitrogen and is thus positively charged whereas HIS318 was protonated at the δ -nitrogen and hence is neutral. The carboxyl groups of the propionate side chains of the *B*-ring and the *C*-ring of the chromophore were deprotonated and thus negatively charged. The whole structure was solvated in a box of TIP3P water³²³ and neutralized using the VMD solvate tool.³²²

The MD simulations were carried out with NAMD²²⁶ using periodic boundary conditions (PBC). CHARMM force field parameters^{199,324} were employed to describe the protein whereas the parameters for the chromophore were taken from Kaminski *et al.*⁴⁶ Prior to the final run of 940 ns, an equilibration of the entire system consisting of a minimization run followed by heating to 300 K was carried out. The production run was performed in the NPT ensemble using the combined Nosé-Hoover Langevin piston method^{228,325} as implemented in NAMD. The long-range electrostatic interactions were treated via the particle-mesh-Ewald method³²⁶ whereas for the van der Waals interactions a cut-off of 10 Å in combination with a smoothing function was employed. During the whole run, the bond lengths between the heavy atoms and the hydrogen atoms were fixed using the SHAKE algorithm,²¹⁰ allowing time steps of 2 fs. The resulting trajectory was analyzed using the python packages MDAnalysis,³²⁷ Numpy,³²⁸ and Pandas.³²⁹ The plots were generated using Matplotlib³³⁰ whereas all molecular structures were rendered with VMD³²² and Chimera.³³¹

The crystal structure and the simulated structure were compared by means of root mean square deviations (rmsd) and root mean square fluctuations (rmsf). Hydrogen bond interactions were calculated based on the geometric criterion³³² using a donor-acceptor cut-off distance of 3.5 Å and a cut-off of 20° for the donor-hydrogen-acceptor angle.

A.1.1.2. NMR Experiments

The 1D ¹³C cross-polarization magic-angle spinning (CP/MAS) spectra were recorded using a Bruker AV-750 WB spectrometer (Karlsruhe, Germany) equipped with a 4 mm triple-resonance MAS probe (Bruker). About 1.3 mg holo-AnPixJg2 generated by *in vitro* assembly with uniformly ¹³C- and ¹⁵N-labeled PCB chromophore was used in this study. Both Pr and Pg photostates of the sample were investigated with the remaining fraction of the other state negligible for NMR measurements at 233 K (gas readout temperature) with a MAS rate of 13 kHz (\pm 3 Hz). A detailed description of the NMR acquisition and processing parameters can be found elsewhere.²⁷⁵

A.1.2. Data Analysis of the MD Trajectory

A.1.2.1. Structural Properties: Root Mean Square Deviations (rmsd)

Structural differences between a reference structure (e.g. the structure obtained from crystallographic measurements) and a simulated structure are most commonly expressed in terms of the root mean square deviation (rmsd) of the atomic coordinates. The rmsd is defined as follows:

$$\mathbf{rmsd} = \sqrt{\frac{1}{N} \cdot \sum_{i=1}^{N} \left(u_i - u_i^{\mathrm{ref}}\right)^2}$$
(A.1)

 u_i are the coordinates of the investigated structure and u_i^{ref} are the coordinates of the reference structure, N is the number of atoms in the considered atom group. This means that for each frame of the trajectory an averaged rmsd value for the regarded atom group is obtained.

A.1.2.2. Dynamical Properties: Root Mean Square Fluctuations (rmsf)

The relative flexibility of the individual atoms can be obtained by calculating their root mean square fluctuations (rmsf). The rmsf are the standard deviation of the respective atomic coordinates and are given by:

$$\operatorname{rmsf}_{j} = \sqrt{\frac{1}{T} \sum_{i=1}^{T} \left(\mathbf{u}_{j} \left(t_{i} \right) - \bar{\mathbf{u}}_{j} \right)^{2}}$$
(A.2)

 $\mathbf{u}(t_i)$ are the atomic positions of particle j at time t, $\mathbf{\bar{u}}$ are the atomic coordinates of the averaged structure of the molecule and T is the number of frames during which the fluctuations of the atomic coordinates are observed.

A.1.2.3. Planarity of the Chromophore during the Simulation

In order to characterize the planarity of the chromophore during the simulation time, the dihedral angles between the individual pyrrole rings are computed. Special focus



Figure A.1. – **a-b**: Definition of the single bond and double bond dihedral around the *A-B* and *C-D* methine bridge. **c**: Definition of the *B-C* reference plane and the *A*-ring as well as the *D*-ring plane.

is given to the dihedral angles around the A-B and C-D methine bridge. As visible in Fig. A.1, those dihedral angles can either enclose the double bond (double bond dihedral) or the single bond (single bond dihedral) which are highlighted in blue and red, respectively. Moreover, the planarity of the A-ring and D with respect to the B-C-ring plane is calculated. The B-C-reference plane as well as the A-ring and D-ring plane are visualized in Fig. A.1. In order to measure the degree of planarity during the simulation run, the normal vector of the A-ring plane as well as the D-ring plane has been projected on the normal vector of the B-C-reference plane are in one plane, i.e. totally planar, the projection is 1. If the A-ring (or the D-ring) are tilted with respect to the B-C-reference plane, the projection is < 1.

A.1.2.4. Hydrogen Bond Occurrences

The existence of a hydrogen bond between a donor, further denoted as d, and an acceptor, a, is calculated using a geometric criterion. According to this geometric criterion, hydrogen bonding between a and d occurs, if the distance between d and a is below a certain threshold and the angle between d, the hydrogen atom bound to d, and a is larger than a certain cut-off angle.³³² In this work, a distance criterion of 3.5 Å^* and a cut-off angle of 160° is used.

A.1.2.5. Solvation of the Chromophore Binding Pocket

To evaluate the amount of water in the chromophore binding pocket, the number of solvent molecules in a distance of 3.5 Å around the heavy atoms of the tetrapyrrole rings excluding the side chains was quantified during the simulation time. In order to suppress short-term fluctuations, the trajectory was divided into blocks of 10 ns. For each of the block, the median number of solvent molecules was computed from the trajectory frames within that block.

^{*}U sually a distance criterion of ≈ 3.5 Å is taken, ³³² as derived from the first minimum of the radial distribution function of water ^{333,334}

A.1.3. Supplementary Material to: One Mutant Three Sub-States

A.1.3.1. Sub-State III Exhibits an Altered Hydrogen Bond Pattern

In the following, the rmsd of the of the carbons of the C-ring propionate side chain of wild type and mutant will be compared to each other. In the TYR302PHE mutant the C-ring propionate side chain dynamically switches between a stretched conformation and a downward bent conformation. These two conformations are visible in the the rmsd of the carbon atoms of the C-ring propionate side chain depicted in Fig. A.2 which exhibits jumps between two plateau regions (see Chapter 3.3.2) Fig. A.2 depicts



Figure A.2. – Comparison of the rmsd of the *C*-ring propionate chain carbons in TYR302PHE mutant (**a**) and in the wild type (**b**). While the rmsd in the mutant exhibits distinct plateau regions, the rmsd of the wild type remains stable except small fluctuations from $\approx 720 \text{ ns} - \approx 940 \text{ ns}$. **c**: Structural snapshots of the *C*-ring propionate chain in the wild type's crystal structure (c.s.) and after $\approx 760 \text{ ns}$ which corresponds to the upper region in the rmsd. It can be seen that the fluctuations in the rmsd correspond merely to the relative position of the second carbon atom and the side chain maintains its stretched configuration during the whole simulation time.

the rmsd of the carbon atoms of the *C*-ring propionate side chain of the TYR302PHE mutant, that is also shown in Chapter 3.3.2 (a) in comparison to that of the wild type (b). It can be seen that the rmsd of the wild type remains stable except for small fluctuations from $\approx 720 \text{ ns} - \approx 940 \text{ ns}$. A structural snapshot of the crystal structure and a structure extracted at $\approx 760 \text{ ns}$ corresponding to the upper region are depicted in Fig. A.2 c. Based on those structures it can be seen that the small fluctuations

correspond to the relative position of the second carbon atom (highlighted by a blue rectangle) but the propionate side chain maintains its stretched configuration during the whole simulation time. Thus, the bent configuration does not occur in the wild type and is very likely induced by the mutation.

A.2. Appendix to Project II

A.2.1. Calculation of STM Images

In the following, the Tersoff-Hamann approximation for calculating STM images is briefly discussed. This approach has been used in **Project II** to compute STM images of the Au(100) (20 × 5) model ("Gold Diggers: Altered Reconstruction of the Gold Surface by Physisorbed Aromatic Oligomers", Chapter 4.2) as well as of the α -6T layer on gold ("From Flat to tilted: Oligothiophene Growth beyond the Monolayer", Chapter 4.3). From a theoretical point of view, the tunneling process can be easily understood by the example of a particle subject to a one-dimensional potential barrier of finite width. Within the barrier, which corresponds to the vacuum region between tip and sample, the wave function of the particle is given by³³⁵

$$\psi(z) = \psi(0) \exp(-\kappa z) \tag{A.3}$$

$$\kappa = \frac{\sqrt{2m\left(|U| - E\right)}}{\hbar} \tag{A.4}$$

Due to the exponential decay of the wave function, which results in an exponential dependence of the tunneling current on the tip-sample distance, the STM possesses a high sensitivity to height changes.³³⁵

The first attempt to theoretically describe STM, made by Bardeen,³³⁶ is the transfer hamiltonian method. This approach is based on several assumptions:^{335,336}

- i the tunneling process is treated as one-particle process
- ii direct interactions between the tip and the surface, i.e. a coupling of tip and surface states, are excluded
- iii inelastic processes are neglected

The first and second assumption are fulfilled in the low tunneling regime and at large enough distance between tip and sample.³³⁵ Within Bardeens formalism, the tunneling current, I, is given by:³⁰⁷

$$I = \frac{2\pi}{\hbar} \sum_{\nu,\mu} f(E_{\mu}) \left[1 - f(E_{\nu} + eV) \right] \left| M_{\mu\nu} \right|^2 \delta(E_{\mu} - E_{\nu})$$
(A.5)

f(E) is the Fermi function, E_{μ} denotes the energy of state ϕ_{μ} in the absence of tunneling, V is the applied bias voltage, e the electronic charge, and $M_{\mu\nu}$ the tunneling matrix element between the states of the tip, ϕ_{μ} , and the surface, ϕ_{ν} . In the limit of

A. Appendix

low temperature and at small bias voltage, Eq. (A.5) is simplified to:³⁰⁷

$$I = \frac{2\pi}{\hbar} e^2 V \sum_{\nu,mu} |M_{\mu,\nu}|^2 \delta(E_{\mu-E_F}) \delta(E_{\nu-E_F})$$
(A.6)

In order to solve either Eq. (A.5) or Eq. (A.6), the tunneling matrix element, $|M_{\mu,\nu}|$ has to be evaluated. It was shown by Bardeen, that $|M_{\mu,\nu}|$ can be written as:^{307,336}

$$M_{\mu\nu} = \frac{\hbar^2}{2m} \int \underbrace{\left(\psi_{\mu}^* \nabla \psi_{\nu} - \psi_{\nu} \nabla \psi_{\mu}^*\right)}_{1} d\mathbf{S}$$
(A.7)

Term 1 is the current operator, and the integral has to be taken over any surface, S, which lies entirely within the barrier region and separates the two sides (surface and tip).³⁰⁷ Note that Eq. (A.7) contains the wave function of the tip; unfortunately, the tip states are generally unknown, which impedes a practical application of Bardeen's approach.

Figure A.3 – Schematic of the tunneling geometry within the Hamann-Tersoff approximation. The tip geometry is approximated by a sphere with radius R that is centered at \mathbf{r}_i . The figure was taken from Ref. 307.



Hamann and Tersoff³⁰⁷ successfully provided an approximation for the tip states that is simple to implement and yet delivers results that agree with experiment. Within their approach, they approximated the microscopic structure of the tip by a locally spherical shape with radius R that is centered at \mathbf{r}_i , as schematically shown in Fig. A.3. The tip wave function within that region of interest can then be assumed to have spherical form, i.e. is a s-type wave function ("spherical tip approximation").³⁰⁷ Evaluation of the matrix element, $|M_{\mu,\nu}|$, leads then to the following expression for the tunneling current:³⁰⁷

$$I = 32\pi^{3}\hbar^{-1}e^{2}V\phi^{2}D_{t}(E_{F})R^{2}\kappa^{-4}\exp(2\kappa R)\cdot\sum_{\nu}\left|\phi_{\nu}(r_{0})\right|^{2}\delta(E_{\nu}-E_{F})$$
(A.8)

 ϕ^2 is the work function of the tip which for simplicity is assumed to be equal to that of the surface.³⁰⁷ $D_t(E_F)$ is the density of states per unit volume of the probe tip. κ denotes the minimum inverse decay length for the wave function in vacuum, i.e. $\kappa = -\hbar^2 (2m\phi)^{\frac{1}{2}}.^{307}$

To summarize, the Hamann Tersoff approach models the tip geometry as a geometrical point and, as a result, STM images computed within this approach exhibit only the features of the unperturbed surface. Despite this simple model of the tip geometry, it is usually employed when calculating STM images and has been proven to be extremely well suited in describing phenomena like the large superstructures of reconstructed surfaces, as well as defects or adsorbates.³³⁵

A.2.2. Supplementary Information to: "Gold Diggers: Altered Reconstruction of the Gold Surface by Physisorbed Aromatic Oligomers"

A.2.2.1. Methods

All STM measurements have been performed in ultrahigh vacuum conditions. Au(100)sample preparation by Ar⁺ sputtering and annealing cycles followed the procedure in Ref. [84]. STM and low-energy electron diffraction experiments reveal the $c(28 \times$ 48) Au(100) reconstruction prior to α -6T evaporation onto the surface at 300 K. Calculations were carried out with the electronic structure code $FHI-AIMS^{170}$ at the DFT level. All adsorption structures were obtained by performing geometry optimizations in which the structure was relaxed until forces were smaller than $10^{-3} \text{eV} \text{ Å}^{-1}$. We used the tier1-tight numerical basis sets, the Perdew-Burke-Ernzerhof (PBE) exchange-correlation functional,¹²⁴ the Tkatchenko-Scheffler van der Waals correction $(TS vdw)^{142}$ and the atomic zeroth-order regular approximation $(ZORA)^{337,338}$ to account for relativistic effects. Since the experimentally observed $c(28 \times 48)$ super cell of the bare Au(100) surface is computationally not tractable, the 20×5 model developed by Havu *et al.*⁸⁹ was chosen. With unit cell vectors of |a| = 58.97 Å and |b| = 15.8 Å this model is computationally still feasible while retaining the main physical features of the surface such as two-dimensional lateral contraction, rotation, and differently buckled qualitative surface areas.^{82,90,339} The slab consists of five layers of gold of which the three lowest were kept fixed during structural relaxation. A $1 \times 2 \times 1$ k-point grid and periodic boundary conditions were employed for the surface slab using a bulk lattice constant of 4.150 Å. The latter has been optimized at the PBE level using a $10 \times 10 \times 10$ k-point grid per bulk unit cell. Calculation of the STM image of the bare Au(100) surface was done using the Hamann-Tersoff approach.³⁰⁷ The computed STM images were postprocessed using the scanning probe microsopy software WSXM.³⁴⁰ Data analysis was done using the python packages MDAnalysis,³²⁷ Numpy,³²⁸ and Pandas.³²⁹ Plots were generated using the python package Matplotlib³³⁰ whereas all molecular structures were rendered with VMD³²² and Chimera.³³¹ The computed STM images were postprocessed using the scanning probe microsopy software WSXM.³⁴⁰

A.2.2.2. Comparison of the Tkatchenko-Scheffler Dispersion Correction and the van-der-Waals^{surf} Dispersion Correction

The description of the adsorption structures of organic molecules on metal surfaces relies strongly on the accurate inclusion of van der Waals interactions between molecule and substrate. The DFT+vdW method developed by Tkatchenko *et al.*¹⁴²

A. Appendix

(further abbreviated as TS vdW) accounts for the van der Waals interactions by a pairwise-additive model. Unlike in other pairwise-additive correction schemes, the pairwise coefficients employed in the TS vdW dispersion correction are not fitted to experimental data but obtained from the mean-field ground state electronic density for molecules and solids. The vdW^{surf} dispersion correction¹³⁸ employs the Lifshitz-Zaremba-Kohn (LZK) theory,^{150,341} that describes the interaction between an atom and a surface, in order to determine reduced C₆ coefficients that include non-local screening effects. Both, the pairwise TS vdW dispersion correction,¹⁵¹ and the vdW^{surf} dispersion correction¹⁵² have been employed successfully to describe organic molecules on inorganic surfaces.



Figure A.4. – Comparison of the surface corrugation and the adsorption structure of the r-to-v adsorbed molecule computed with either the TS (left) or the vdW^{surf} dispersion correction (right). Both schemes result in nearly identical surface corrugation patterns and adsorption structures that closely resemble each other.

In order to compare the performance of the TS vdW dispersion correction and the vdW^{surf} dispersion correction for corrugated gold surfaces, structural relaxations for the molecule on the reconstruction ridge and in the reconstruction valley of the Au(100) surface for both dispersion corrections were performed.

A comparison of the adsorption energies and heights with the TS vdW and the vdW^{surf} dispersion correction is presented in Tab. A.1. Due to the screened C₆ coefficients, the vdW^{surf} dispersion correction yields smaller adsorption energies and slightly lower adsorption heights compared to the TS vdW dispersion corrections, which is in agreement with findings for PTCDA on Ag(111).¹³⁸ Apart from the overall lower adsorption heights, the TS vdW and the vdW^{surf} methods yield similar adsorption structures: in both cases the molecule adsorbs completely flat with the C_{α} atoms directly about the reconstruction valley which is also manifested in larger C_{α} - Au distances (see Tab. A.1).

Fig. A.4 shows the surface corrugation after r-to-v adsorption of α -6T when employing either the TS vdW dispersion correction (left) or the vdW^{surf} dispersion

	vdW ^{sur} dispersion correction.						
system	E_{ads} [eV]	E_{ads}	$\Delta h S$ - Au	$\Delta h \ C_{\beta}$ -	$\Delta h C_{\alpha}$ -		
		$[kJ mol^{-1}]$		Au	Au		
v TS	-4.34	-417	3.25	3.16	3.34		
r-to-v, TS	-4.31	-414	3.18	3.18	3.22		
v, vdW^{surf}	-3.77	-362	3.19	3.02	3.19		
r-to-v,	-3.78	-363	3.25	2.97	3.15		
$\mathrm{vdW}^{\mathrm{surf}}$							

 $\label{eq:table_ads} \begin{array}{l} \mbox{Table A.1.} - \mbox{Comparison of the adsorption energies } E_{ads} \mbox{ in } kJ \, mol^{-1} \mbox{ and } eV \mbox{ as well as the median adsorption heights } \Delta h \mbox{ (Å) computed using the TS vdW or the vdW^{surf} dispersion correction.} \end{array}$

correction (right). In both cases the molecule pushed the ridge aside and created itself a valley-like adsorption environment that looks almost indistinguishable from each other. Interestingly, when using the TS vdW dispersion correction, alteration of the surface reconstruction by the molecule placed on the r-site took already place during the pre-optimization using the tier1-light basis. When employing the vdW^{surf} dispersion correction and the tier1-light basis, however, the molecule merely stayed in its local minimum on top of the reconstruction ridge instead of altering the surface reconstruction or migrate into the global minimum in the reconstruction valley. After switching from the tier1-light to the more accurate tier1-tight basis the molecule also started to move the ridge aside and created a reconstruction valley directly underneath itself. This behavior is particularly noteworthy since occasionally calculations of largescale systems are performed using the light basis in order to reduce computational costs, which in this case would lead to a completely different result.

A.2.2.3. Approximating Pathways for Scenario A and B: Linear Interpolation

To investigate the energy barrier for the two scenarios, A, in which the molecule slides down the ridge, and B, in which the molecule moves the ridge aside, both pathways were approximated by performing a linear interpolation between the starting and final coordinates of both processes. For the starting structure of scenario B (pathway I) the molecule was placed on the reconstruction ridge of the previously relaxed gold surface whereas for the final structure the optimized structure after r-to-v adsorption was taken. We then linearly connected the coordinates of each atom of the starting structure and the final structure by a vector $\mathbf{u}_{s,f}$ by subtracting the vector of the starting coordinates from the vector of the final coordinates. The coordinates of the nine intermediate structures were obtained according to:

$$\mathbf{u_i} = \mathbf{u}_s + i \cdot \mathbf{u}_{s,f}, \ i = 0.1, 0.2, ..., 0.9$$
 (A.9)

with \mathbf{u}_s and \mathbf{u}_f being the coordinates of the starting and final structure, respectively, and $\mathbf{u}_{s,f}$ being the vector difference between the coordinates of the starting and final structure. This procedure resulted in 11 structures in total (one starting structure, one final structure and 9 intermediate structures) for which single point calculations with the t1-tight basis were performed.



Figure A.5. – Energy curves for the linear interpolation between the starting and final coordinates of scenario B (a) and the optimization of the molecule placed initially on the reconstruction ridge of the deformable surface (b).

We wanted to assure that the performed linear interpolation which resulted in the barrier-free pathway I, is not merely a repetition of the geometry optimization for the molecule placed initially on the reconstruction ridge of the deformable surface (r-to-v adsorption). Fig. A.5 depicts the energy curves for pathway I (a) and the the corresponding geometry optimization (b). It is visible that the energy curves are completely dissimilar to each other and hence the results from linear interpolation suggest indeed an independent, barrier-free pathway for the molecule-induced restructuring of the surface.

In order to obtain a starting structure for scenario A, the molecular structure of α -6T on the reconstruction ridge of the previously relaxed surface was optimized. During the optimization procedure, the coordinates of all gold atoms were fixed to avoid any restructuring of the surface. The molecule remained on top of the ridge in a local minimum and did not slide down into the valley during the optimization. For the final structure, the structure of the molecule was optimized in the reconstruction valley while keeping the surface coordinates fixed. We then linearly interpolated between those two structures by creating intermediate structures according to the procedure described previously for scenario B^{*}

^{*}We additionally included an intermediate structure between structure 7 and 8 for a better estimation of the barrier height, leading to a total of 10 intermediate structures

A.2.3. Supplementary Information to: "From Flat to Tilted: Oligothiophene Growth beyond the Monolayer"

A.2.3.1. Methods

The unit cell of the (20×5) reconstructed Au(100) surface and the unit surface cell of the α -6T monolayer are incommensurable. Therefore, the hexagonal, flat Au surface is taken as a model system for the reconstructed Au(100) surface. The metallic substrate was modeled by a five-layer slab and laterally by a (18×7) repetition of the unit cell for the α -6T layer using the experimental lattice constant of Au (4.07 Å).³⁴² Periodic boundary conditions were applied in all three dimensions. To decouple the periodic images in z direction, 23 Å of vacuum were added. The molecules have been only absorbed on one side of the slab, hence the slab structure is asymmetric.

All calculations were carried out with the quantum chemistry package CP2K $4.0.^{343}$ In order to tackle the huge system size (up to 1300 atoms per unit cell) a hybrid image charge-quantum mechanics/molecular mechanics (IC-QM/MM) approach as recently implemented in CP2K¹⁹⁶ was chosen. This approach allows to treat the metal atoms on a force field level of theory and the adsorbate quantum mechanically and has already been used successfully to describe the physisorption of different systems on metallic surfaces.^{196,197} For the α -6T molecules the valence electrons were represented by double- ζ plus polarization basis sets of the MOLOPT type.³⁴⁴ In order to describe the interactions between valence and core electrons, norm conserving Goedecker, Teter, and Hutter pseudopotentials (GHT pseudopotentials)^{181–183} were employed. An accurate description of packing structure and band gaps of π -conjugated organic systems requires 20% - 40% exact exchange.³⁴⁵⁻³⁴⁷ The exchange and correlation potential was thus modeled by the Perdew-Burke-Ernzerhof (PBE) functional³⁴⁸ using 25% of exact exchange (PBE0). Dispersion interactions were accounted for by Grimme's D3 correction.¹⁴¹ Furthermore, the well established auxiliary density matrix method (ADMM)³⁴⁹ was employed, which makes a description of large systems (several hundred atoms) with hybrid functionals feasible, together with the orbital transformation method (OT).³² The gold surface was described using an EAM scheme (embedded atom scheme).³⁵⁰ The electrostatic interaction between adsorbate and substrate are accounted for by the image charge approach whereas the dispersion interaction and Pauli repulsion was modeled by a Lennard Jones potential. The Lennard-Jones parameter for Au and for the sexithiophene C. S and H atoms were taken from Heinz *et al.*³⁵¹ and Moreno *et al.*,³⁵² respectively.

All pre-processing and post-processing structural analyses was carried out using the python packages MDAnalyis,³²⁷ Numpy,³²⁸ and Pandas.³²⁹ The plots were generated using the python package Matplotlib³³⁰ whereas all molecular structures were rendered with VMD³²² and Chimera.³³¹

A.2.3.2. Parameter and Method Validation

So far, neither experimental values nor values from theoretical studies are available for the adsorption energy of α -6T on gold. However, the adsorption energy of thiophene on gold has been derived from results from desorption spectroscopy experiments^{353–355} applying the Redhead equation³⁵⁶ and calculated on a DFT level of theory.¹⁵² The experimentally measured values lie in a range of $51 \text{ kJ mol}^{-1} - 75 \text{ kJ mol}^{-1}$ whereas the value from DFT studies is 74 kJ mol^{-1} . In order to benchmark our QM/MM setup that describes the interaction between the molecule and the surfaces by a pairwise Lennard-Jones interaction, the adsorption energy and geometry of a single thiophene molecule on gold has been calculated.

One concern regarding the Lennard-Jones parameters are the various combination rules which alter the result significantly. The well-known Berthelot mixing rules which are adequate for elements of the first two rows fail to describe interactions with heavy elements beyond the first and second row in terms of overestimating the non-covalent interactions.³⁵⁷ The mathematically derived Waldman-Hagler rules were already successfully applied to describe the interaction between unlike atoms in a similar system as ours.³⁵⁸ It was further suggested to use the Berthelot combination rules only for hydrogen and employ the Waldmann-Hagler rules for all heavy atoms.³⁵⁹ We find that employing only the Berthelot rules leads to an overestimation of the adsorption energy of over 50% (113 kJ mol⁻¹). Using the Waldman-Hagler and the Berthelot combination rules as suggested in Ref.,³⁵⁹ we obtain an adsorption energy of $\approx 88 \, \text{kJ} \, \text{mol}^{-1}$ for one thisphene on a hexagonal Au surface which is $\approx 20\%$ higher than the results from full DFT calculations as well as the upper boundary of the experimental values. A deviation of 20% of the adsorption energy can be considered acceptable in case that both competing interactions, the adsorption of the molecule on gold and the interaction between the molecules show a significant difference in their magnitude. Indeed, highly correlated Coupled Cluster calculations of thiophene dimers³⁶⁰ result in interaction energies which are at least five times lower than the adsorption energy of thiophene on gold from our calculations. Due to the large difference between the molecule-molecule and molecule-surface interaction, an overestimation of 20% in the adsorption energy is thus acceptable.

In order to investigate the accuracy of our dispersion corrected DFT approach (DFT-D3 approach) concerning thiophene based aromatic systems, the interaction energy between two stacked thiophene molecules has been calculated and the results have been compared to data from coupled cluster calculations.³⁶⁰ We find that DFT-D3 overestimates the interaction energy about 11 % to a maximum of 16 % but is well able to reproduce the trend. A comparison of the coupled cluster values³⁶⁰ and the computed DFT-D3 values is visualized in Fig. A.6

A.2.3.3. Interpretation of the Experimental and Computed STM Images

In the following, additional material (i.e. orbital plots of a single thiophene and a comparison of the computed STM images of the relaxed bilayer structures) that supports the interpretation of the experimental STM images shown in Chapter 4.3.3 is presented.

Fig. A.7 shows a plot of the highest molecular orbital (HOMO, \mathbf{a}) and the HOMO-1 (\mathbf{b}) of a single thiophene molecule. It can be seen that only the p-orbitals of the carbon atoms contribute to the HOMO but not the p orbitals of the sulfur. This is different for the HOMO-1, which is also located at the sulfur. Since the experimental and



Figure A.6. – Comparison of the interaction energies of a stacked thiophene dimer calculated on the DFT level of theory with results from CCSD(T) calculations from Tsuzuki *et. al.*³⁶⁰



Figure A.7. – Computed highest occupied molecular orbital (HOMO, **a**) and HOMO-1 (**b**) of thiophene. It can be seen that, unlike to the HOMO-1, the π electrons of the sulfur do not contribute to the HOMO.

the computational STM images are measured at a bias voltage that includes only the highest occupied state, i.e. the HOMO, the isosurface plot of the HOMO depicted in Fig. A.7 confirms, that the droplet shaped lobes visible in the STM image can indeed ne assigned to the aromatic carbon atoms while the sulfur atom does not contribute to the STM image.

Fig. A.8 depicts the computed bilayer structures. The extensive structure search also resulted in a tilted 2S@1S structure similar to structure II that was equal in energy (Fig A.8 c). Interestingly, that equally tilted structure II has not been found for the 2R@1R structure. Moreover, the existence of such an equally tilted structure could not be observed by STM measurements for any of the enantiomeric structures. A comparison between the staggered and the equally tilted 2S@1S structures is given in Fig. A.8. Except of the molecules in two consecutive rows being tilted in equal direction, the B-structure comprises the same structural motifs as the staggered



Figure A.8. – Calculated STM images of the staggered 2R@1R and 2S@1S structures (structure III) as well as of the equally tilted 2S@1S structure II that was not observed experimentally.

2S@1S structure III, namely a lateral shift of 2.2 Å and the characteristic, albeit much smaller shift along the LMA of only 0.2 Å. The tilt angles are $10^{\circ} - 12^{\circ}$ and hence narrower in their distribution compared to the investigated structures that are similar to structure III. Notably, the 2S@1S structure II exhibits a considerable translation of two consecutive molecular rows in the second layer similar to the row translation $\Delta_{\rm row}^{\rm intra}$ of the first layer. The row translation $\Delta_{\rm row}^{\rm intra}$ is ≈ 2.3 Å. In comparison, the 2S@1S structure III and 2R@1R structure III comprise only a row translation of 0.8 Å and 0.7 Å, respectively.

The lower panel of Fig. A.8 depicts the corresponding calculated STM images of the three structures. A comparison of the STM images of the 2S@1S structure III (staggered) and the 2S@1S structure II (equally tilted) shows that the equally tilted structure II depicts indeed a unit cell comprising only one α -6T molecules, i.e. the STM pattern of two consecutive molecules is the same, while it is reverted for the 2S@1S structure III.

A.2.3.4. Transition to the Bulk Structure

Fig. A.9 depicts a structural snapshot of the computed 2S@1S bilayer structure; for a better visualization only one molecular row of the second layer is shown. The shift along the LMA, Δ_{LMA} , is clearly visible and enables a downward bent configuration of the first thiophene unit, as highlighted by a red rectangle. Such a downward bent configuration is favorable since it increases the interactions between the molecules in the second layer and the surface, which exceed the molecule-molecule interactions by far, see Section A.2.3.2.



Figure A.9. – Structural snapshot of one molecular row in the second layer of the 2S@1S structure. It can be seen that the first thiophene unit of the molecules in the second layer is bent towards the surface.

B. References

- [1] R. Hammer, Molekulare und elektronische Struktur ultradünner alpha-Sexithiophenschichten auf Edelmetalleinkristalloberflächen, Ph.D. thesis, Naturwissenschaftliche Fakultät II, Martin Luther Universität Halle (2014).
- [2] V. C. Galvão and C. Fankhauser, Sensing the light environment in plants: photoreceptors and early signaling steps, Curr. Opin. Neurobiol. 34, 46 – 53, molecular biology of sensation (2015).
- [3] D. P. Fraser, S. Hayes, and K. A. Franklin, *Photoreceptor crosstalk in shade avoidance*, Curr. Opin. Plant Biol. **33**, 1 7, sI: 33: Cell signalling and gene regulation 2016 (2016).
- [4] M. Jaubert, J.-P. Bouly, M. R. d'Alcalà, and A. Falciatore, Light sensing and responses in marine microalgae, Curr. Opin. Plant Biol. 37, 70 – 77, 37 Physiology and metabolism 2017 (2017).
- [5] Z. Yu and R. Fischer, Light sensing and responses in fungi, Nat. Rev. Microbiol. (2019).
- [6] M. Ikeuchi and T. Ishizuka, Cyanobacteriochromes: a new superfamily of tetrapyrrole-binding photoreceptors in cyanobacteria, Photochem. Photobiol. Sci. 7, 1159–1167 (2008).
- [7] Y. Hirose, R. Narikawa, M. Katayama, and M. Ikeuchi, Cyanobacteriochrome CcaS regulates phycoerythrin accumulation in Nostoc punctiforme, a group II chromatic adapter, Proc. Natl. Acad. Sci. USA 107, 8854–8859 (2010).
- [8] N. C. Rockwell, R. Ohlendorf, and A. Möglich, *Cyanobacteriochromes in full color and three dimensions*, Proc. Natl. Acad. Sci. USA **110**, 806–807 (2013).
- [9] R. Narikawa, Y. Fukushima, T. Ishizuka, S. Itoh, and M. Ikeuchi, A Novel Photoactive GAF Domain of Cyanobacteriochrome AnPixJ That Shows Reversible Green/Red Photoconversion, J. Mol. Biol. 380, 844–855 (2008).
- [10] N. C. Rockwell, S. S. Martin, K. Feoktistova, and J. C. Lagarias, *Diverse two-cysteine photocycles in phytochromes and cyanobacteriochromes*, Proc. Natl. Acad. Sci. USA **108**, 11854–11859 (2011).
- [11] N. C. Rockwell, S. S. Martin, and J. C. Lagarias, *Red/Green Cyanobacteri-ochromes: Sensors of Color and Power*, Biochemistry 51, 9667–9677 (2012).

- [12] N. C. Rockwell, S. S. Martin, and J. C. Lagarias, Mechanistic Insight into the Photosensory Versatility of DXCF Cyanobacteriochromes, Biochemistry 51, 3576–3585 (2012).
- [13] N. C. Rockwell, S. S. Martin, and J. C. Lagarias, *Identification of Cyanobacteriochromes Detecting Far-Red Light*, Biochemistry 55, 3907–3919, pMID: 27295035 (2016).
- [14] Method of the Year 2010, Nature Methods, doi:, Editorial (2010).
- [15] C. W. Tang and S. A. VanSlyke, Organic electroluminescent diodes, Appl. Phys. Lett. 51, 913–915 (1987).
- [16] J. H. Burroughes, D. D. C. Bradley, A. R. Brown, R. N. Marks, K. Mackay, R. H. Friend, P. L. Burns, and A. B. Holmes, *Light-emitting diodes based on conjugated polymers*, Nature **347** (1990).
- [17] T. M. Figueira-Duarte and K. Müllen, Pyrene-Based Materials for Organic Electronics, Chem. Rev. 111, 7260–7314, pMID: 21740071 (2011).
- [18] K. Asadi, M. Li, P. W. Blom, M. Kemerink, and D. M. de Leeuw, Organic ferroelectric opto-electronic memories, Mater. Today 14, 592 – 599 (2011).
- [19] A. N. Sokolov, B. C.-K. Tee, C. J. Bettinger, J. B.-H. Tok, and Z. Bao, Chemical and Engineering Approaches To Enable Organic Field-Effect Transistors for Electronic Skin Applications, Acc. Chem. Res. 45, 361–371, pMID: 21995646 (2012).
- [20] J. E. Anthony, Functionalized Acenes and Heteroacenes for Organic Electronics, Chem. Rev. 106, 5028–5048, pMID: 17165682 (2006).
- [21] A. R. Murphy and J. M. J. Fréchet, Organic Semiconducting Oligomers for Use in Thin Film Transistors, Chem. Rev. 107, 1066–1096, pMID: 17428023 (2007).
- [22] H. Yamada, T. Okujima, and N. Ono, Organic semiconductors based on small molecules with thermally or photochemically removable groups, Chem. Commun. pages 2957–2974 (2008).
- [23] J. Mei, Y. Diao, A. L. Appleton, L. Fang, and Z. Bao, Integrated Materials Design of Organic Semiconductors for Field-Effect Transistors, J. Am. Chem. Soc. 135, 6724–6746, pMID: 23557391 (2013).
- [24] M. E. Tuckerman, Statistical Mechanics: Theory and Molecular Simulation, chapter 3, page 97, Oxford University Press (2010).
- [25] J. R. Perilla, B. C. Goh, C. K. Cassidy, B. Liu, R. C. Bernardi, T. Rudack, H. Yu, Z. Wu, and K. Schulten, *Molecular dynamics simulations of large macromolecular complexes*, Curr. Opin. Struct. Biol. (2015).

- [26] T. Sosnick, L. Mayne, R. Hiller, and S. Englander, The barriers in protein folding, Nat. Struct. Biol. 1 (1994).
- [27] K. Henzler-Wildman and D. Kern, Dynamic personalities of proteins, Nature 450, 964–972 (2007).
- [28] G. Zhao, J. R. Perilla, E. L. Yufenyuy, X. Meng, B. Chen, J. Ning, J. Ahn, A. M. Gronenborn, K. Schulten, C. Aiken, and P. Zhang, *Mature HIV-1 capsid* structure by cryo-electron microscopy and all-atom molecular dynamics, Nature 497 (2013).
- [29] D. E. Chandler, J. Strümpfer, M. Sener, S. Scheuring, and K. Schulten, Light Harvesting by Lamellar Chromatophores in Rhodospirillum photometricum, Biophys. J. 106, 2503 – 2510 (2014).
- [30] D. E. Shaw, R. O. Dror, J. K. Salmon, J. P. Grossman, K. M. Mackenzie, J. A. Bank, C. Young, M. M. Deneroff, B. Batson, K. J. Bowers, E. Chow, M. P. Eastwood, D. J. Ierardi, J. L. Klepeis, J. S. Kuskin, R. H. Larson, K. Lindorff-Larsen, P. Maragakis, M. A. Moraes, S. Piana, Y. Shan, and B. Towles, *Millisecond-scale Molecular Dynamics Simulations on Anton*, In *Proceedings of the Conference on High Performance Computing Networking, Storage and Analysis*, SC '09, pages 39:1–39:11. ACM, New York, NY, USA, doi: (2009).
- [31] V. W. zhe Yu, F. Corsetti, A. García, W. P. Huhn, M. Jacquelin, W. Jia, B. Lange, L. Lin, J. Lu, W. Mi, A. Seifitokaldani, Álvaro Vázquez-Mayagoitia, C. Yang, H. Yang, and V. Blum, *ELSI: A unified software interface for Kohn–Sham electronic structure solvers*, Comput. Phys. Commun. **222**, 267 – 285 (2018).
- [32] J. VandeVondele and J. Hutter, An efficient orbital transformation method for electronic structure calculations, J. Chem. Phys. 118, 4365–4369 (2003).
- [33] R. Narikawa, T. Ishizuka, N. Muraki, T. Shiba, G. Kurisu, and M. Ikeuchi, Structures of cyanobacteriochromes from phototaxis regulators AnPixJ and TePixJ reveal general and specific photoconversion mechanism, Proc. Natl. Acad. Sci. USA 110, 918–923 (2013).
- [34] Y. Fukushima, M. Iwaki, R. Narikawa, M. Ikeuchi, Y. Tomita, and S. Itoh, Photoconversion Mechanism of a Green/Red Photosensory Cyanobacteriochrome AnPixJ: Time-Resolved Optical Spectroscopy and FTIR Analysis of the AnPixJ-GAF2 Domain, Biochemistry 50, 6328–6339 (2011).
- [35] P. W. Kim, L. H. Freer, N. C. Rockwell, S. S. Martin, J. C. Lagarias, and D. S. Larsen, Second-Chance Forward Isomerization Dynamics of the Red/Green Cyanobacteriochrome NpR6012g4 from Nostoc punctiforme, J. Am. Chem. Soc. 134, 130–133 (2012).

- [36] P. W. Kim, L. H. Freer, N. C. Rockwell, S. S. Martin, J. C. Lagarias, and D. S. Larsen, Femtosecond Photodynamics of the Red/Green Cyanobacteriochrome NpR6012g4 from Nostoc punctiforme. 1. Forward Dynamics, Biochemistry 51, 608–618 (2012).
- [37] S. M. Gottlieb, P. W. Kim, N. C. Rockwell, Y. Hirose, M. Ikeuchi, J. C. Lagarias, and D. S. Larsen, Primary Photodynamics of the Green/Red-Absorbing Photoswitching Regulator of the Chromatic Adaptation E Domain from Fremyella diplosiphon, Biochemistry 52, 8198–8208 (2013).
- [38] P. W. Kim, L. H. Freer, N. C. Rockwell, S. S. Martin, J. C. Lagarias, and D. S. Larsen, *Femtosecond Photodynamics of the Red/Green Cyanobacteriochrome NpR6012g4 from Nostoc punctiforme. 2. Reverse Dynamics*, Biochemistry 51, 619–630, pMID: 22148731 (2012).
- [39] S. Lim, Q. Yu, S. M. Gottlieb, C.-W. Chang, N. C. Rockwell, S. S. Martin, D. Madsen, J. C. Lagarias, D. S. Larsen, and J. B. Ames, *Correlating structural* and photochemical heterogeneity in cyanobacteriochrome NpR6012g4, Proc. Natl. Acad. Sci. USA 115, 4387–4392 (2018).
- [40] C. Song, G. Psakis, C. Lang, J. Mailliet, W. Gärtner, J. Hughes, and J. Matysik, Two ground state isoforms and a chromophore D-ring photoflip triggering extensive intramolecular changes in a canonical phytochrome, Proc. Natl. Acad. Sci. USA 108, 3842–3847 (2011).
- [41] P. W. Kim, N. C. Rockwell, S. S. Martin, J. C. Lagarias, and D. S. Larsen, *Heterogeneous Photodynamics of the Pfr State in the Cyanobacterial Phytochrome Cph1*, Biochemistry 53, 4601–4611 (2014).
- [42] P. W. Kim, N. C. Rockwell, S. S. Martin, J. C. Lagarias, and D. S. Larsen, Dynamic Inhomogeneity in the Photodynamics of Cyanobacterial Phytochrome Cph1, Biochemistry 53, 2818–2826, pMID: 24742290 (2014).
- [43] F. Velazquez Escobar, C. Lang, A. Takiden, C. Schneider, J. Balke, J. Hughes, U. Alexiev, P. Hildebrandt, and M. A. Mroginski, Protonation-Dependent Structural Heterogeneity in the Chromophore Binding Site of Cyanobacterial Phytochrome Cph1, J. Phys. Chem. B 121, 47–57, pMID: 27966353 (2017).
- [44] J. R. Wagner, J. Zhang, D. von Stetten, M. Günther, D. H. Murgida, M. A. Mroginski, J. M. Walker, K. T. Forest, P. Hildebrandt, and R. D. Vierstra, Mutational Analysis of Deinococcus radiodurans Bacteriophytochrome Reveals Key Amino Acids Necessary for the Photochromicity and Proton Exchange Cycle of Phytochromes, J. Biol. Chem. 283, 12212–12226 (2008).
- [45] D. von Stetten, M. Gänther, P. Scheerer, D. H. Murgida, M. A. Mroginski, N. Krauß, T. Lamparter, J. Zhang, D. M. Anstrom, R. D. Vierstra, K. T. Forest, and P. Hildebrandt, *Chromophore Heterogeneity and Photoconversion in Phytochrome Crystals and Solution Studied by Resonance Raman Spectroscopy*, Angew. Chem. Int. Ed. 47, 4753–4755 (2008).

- [46] S. Kaminski, G. Daminelli, and M. A. Mroginski, Molecular Dynamics Simulations of the Chromophore Binding Site of Deinococcus radiodurans Bacteriophytochrome Using New Force Field Parameters for the Phytochromobilin Chromophore, J. Phys. Chem. B 113, 945–958 (2009).
- [47] T. Ishizuka, A. Kamiya, H. Suzuki, R. Narikawa, T. Noguchi, T. Kohchi, K. Inomata, and M. Ikeuchi, *The Cyanobacteriochrome, TePixJ, Isomerizes Its Own Chromophore by Converting Phycocyanobilin to Phycoviolobilin*, Biochemistry **50**, 953–961, pMID: 21197959 (2011).
- [48] N. C. Rockwell, S. S. Martin, A. G. Gulevich, and J. C. Lagarias, *Phycoviolobilin Formation and Spectral Tuning in the DXCF Cyanobacteriochrome Subfamily*, Biochemistry **51**, 1449–1463 (2012).
- [49] P. W. Kim, J. Pan, N. C. Rockwell, C.-W. Chang, K. C. Taylor, J. Clark Lagarias, and D. S. Larsen, Ultrafast E to Z photoisomerization dynamics of the Cph1 phytochrome, Chem. Phys. Lett. 549, 86–92 (2012).
- [50] J. M. Kelly and J. C. Lagarias, Photochemistry of 124-kilodalton Avena phytochrome under constant illumination in vitro, Biochemistry 24, 6003–6010 (1985).
- [51] K. C. Toh, E. A. Stojković, I. H. M. van Stokkum, K. Moffat, and J. T. M. Kennisa, Proton-transfer and hydrogen-bond interactions determine fluorescence quantum yield and photochemical efficiency of bacteriophytochrome, Proc. Natl. Acad. Sci. USA 107, 9170–9175 (2010).
- [52] F. Hoffmann, J. Adler, B. Chandra, K. R. Mote, G. Bekçioğlu-Neff, D. Sebastiani, and D. Huster, *Perturbation of the F19-L34 Contact in Amyloid β* (1-40) Fibrils Induces Only Local Structural Changes but Abolishes Cytotoxicity, J. Phys. Chem. Lett. 8, 4740–4745, pMID: 28910107 (2017).
- [53] L. Torsi, A. Dodabalapur, L. J. Rothberg, A. W. P. Fung, and H. E. Katz, Charge transport in oligothiophene field-effect transistors, Phys. Rev. B 57, 2271–2275 (1998).
- [54] F. Garnier, A. Yassar, R. Hajlaoui, G. Horowitz, F. Deloffre, B. Servet, S. Ries, and P. Alnot, *Molecular engineering of organic semiconductors: design of self*assembly properties in conjugated thiophene oligomers, J. Am. Chem. Soc. 115, 8716–8721 (1993).
- [55] N. Koch, Organic Electronic Devices and Their Functional Interfaces, ChemPhysChem 8, 1438–1455 (2007).
- [56] S. Günes, H. Neugebauer, and N. S. Sariciftci, Conjugated Polymer-Based Organic Solar Cells, Chem. Rev. 107, 1324–1338, pMID: 17428026 (2007).
- [57] H. E. Katz and J. Huang, *Thin-Film Organic Electronic Devices*, Annu. Rev. Mater. Res. **39**, 71–92 (2009).

- [58] E. M. Mannebach, J. W. Spalenka, P. S. Johnson, Z. Cai, F. J. Himpsel, and P. G. Evans, *High Hole Mobility and Thickness-Dependent Crystal Structure* in α,ω-Dihexylsexithiophene Single-Monolayer Field-Effect Transistors, Adv. Funct. Mater. 23, 554–564 (2013).
- [59] G. Horowitz, D. Fichou, X. Peng, Z. Xu, and F. Garnier, A field-effect transistor based on conjugated alpha-sexithienyl, Solid State Commun. 72, 381 – 384 (1989).
- [60] G. Horowitz, Organic Semiconductors for new electronic devices, Adv. Mater. 2, 287–292 (1990).
- [61] A. Dodabalapur, L. Torsi, and H. E. Katz, Organic Transistors: Two-Dimensional Transport and Improved Electrical Characteristics, Science 268, 270–271 (1995).
- [62] S. T. Bromley, M. Mas-Torrent, P. Hadley, and C. Rovira, Importance of Intermolecular Interactions in Assessing Hopping Mobilities in Organic Field Effect Transistors: Pentacene versus Dithiophene-tetrathiafulvalene, J. Am. Chem. Soc. 126, 6544–6545, pMID: 15161269 (2004).
- [63] J.-L. Brédas, D. Beljonne, V. Coropceanu, and J. Cornil, Charge-Transfer and Energy-Transfer Processes in π-Conjugated Oligomers and Polymers: A Molecular Picture, Chem. Rev. 104, 4971–5004, pMID: 15535639 (2004).
- [64] G. Heimel, L. Romaner, E. Zojer, and J.-L. Bredas, *The Interface Energetics of Self-Assembled Monolayers on Metals*, Acc. Chem. Res. 41, 721–729, pMID: 18507404 (2008).
- [65] H. Ishii, K. Sugiyama, E. Ito, and K. Seki, Energy Level Alignment and Interfacial Electronic Structures at Organic/Metal and Organic/Organic Interfaces, Adv. Mater. 11, 605–625 (1999).
- [66] I. Avilov, V. Geskin, and J. Cornil, Quantum-Chemical Characterization of the Origin of Dipole Formation at Molecular Organic/Organic Interfaces, Adv. Funct. Mater. 19, 624–633 (2009).
- [67] J. Gong, Structure and Surface Chemistry of Gold-Based Model Catalysts, Chem. Rev. 112, 2987–3054, pMID: 21524063 (2012).
- [68] Y. Zhang, X. Cui, F. Shi, and Y. Deng, Nano-Gold Catalysis in Fine Chemical Synthesis, Chem. Rev. 112, 2467–2505, pMID: 22112240 (2012).
- [69] X. Zhang and Y. Ding, Unsupported nanoporous gold for heterogeneous catalysis, Catalysis Science & Technology 3, 2862–2868 (2013).
- [70] H. Masatake, Chance and Necessity: My Encounter with Gold Catalysts, Angew. Chem., Int. Ed. 53, 52–56 (2014).
- [71] A. Wittstock and M. Bäumer, Catalysis by Unsupported Skeletal Gold Catalysts, Acc. Chem. Res. 47, 731–739, pMID: 24266888 (2014).
- [72] P. Maksymovych, D. C. Sorescu, and J. T. Yates, Gold-Adatom-Mediated Bonding in Self-Assembled Short-Chain Alkanethiolate Species on the Au(111) Surface, Phys. Rev. Lett. 97, 146103 (2006).
- [73] A. Cossaro, R. Mazzarello, R. Rousseau, L. Casalis, A. Verdini, A. Kohlmeyer, L. Floreano, S. Scandolo, A. Morgante, M. L. Klein, and G. Scoles, X-ray Diffraction and Computation Yield the Structure of Alkanethiols on Gold(111), Science **321**, 943–946 (2008).
- [74] G. Nenchev, B. Diaconescu, F. Hagelberg, and K. Pohl, Self-assembly of methanethiol on the reconstructed Au(111) surface, Phys. Rev. B 80, 081401 (2009).
- [75] P. Maksymovych, O. Voznyy, D. B. Dougherty, D. C. Sorescu, and J. T. Yates, Gold adatom as a key structural component in self-assembled monolayers of organosulfur molecules on Au(111), Prog. Surf. Sci. 85, 206 – 240 (2010).
- [76] I. Chizhov, A. Kahn, and G. Scoles, Initial growth of 3,4,9,10perylenetetracarboxylic-dianhydride (PTCDA) on Au(111): a scanning tunneling microscopy study, J. Cryst. Growth 208, 449 – 458 (2000).
- [77] W. Zhao, H. Zhu, H. Song, J. Liu, Q. Chen, Y. Wang, and K. Wu, Adsorption and Assembly of Photoelectronic TiOPc Molecules on Coinage Metal Surfaces, J. Phys. Chem. C 122, 7695–7701 (2018).
- [78] A. E. Baber, S. C. Jensen, E. V. Iski, and E. C. H. Sykes, *Extraordinary Atomic Mobility of Au111 at 80 Kelvin: Effect of Styrene Adsorption*, J. Am. Chem. Soc. **128**, 15384–15385, pMID: 17131999 (2006).
- [79] F. Rossel, P. Brodard, F. Patthey, N. V. Richardson, and W.-D. Schneider, Modified herringbone reconstruction on Au(111) induced by self-assembled Azure A islands, Surf. Sci. 602, L115 – L117 (2008).
- [80] A. Höfer, K. Duncker, M. Kiel, S. Wedekind, and W. Widdra, Adsorption of α -sexithiophene on Au(001): Molecule-induced partial lifting of the substrate reconstruction, Phys. Rev. B 83, 075414 (2011).
- [81] M. Kiel, K. Duncker, C. Hagendorf, and W. Widdra, Molecular structure and chiral separation in α-sexithiophene ultrathin films on Au(111): Low-energy electron diffraction and scanning tunneling microscopy, Phys. Rev. B 75, 195439 (2007).
- [82] D. G. Fedak and N. A. Gjostein, Structure and Stability of the (100) Surface of Gold, Phys. Rev. Lett. 16, 171–172 (1966).

- [83] M. V. Hove, R. Koestner, P. Stair, J. Bibérian, L. Kesmodel, I. BartoŠ, and G. Somorjai, The surface reconstructions of the (100) crystal faces of iridium, platinum and gold: I. Experimental observations and possible structural models, Surf. Sci. 103, 189 – 217 (1981).
- [84] R. Hammer, A. Sander, S. Förster, M. Kiel, K. Meinel, and W. Widdra, Surface reconstruction of Au(001): High-resolution real-space and reciprocal-space inspection, Phys. Rev. B 90, 035446 (2014).
- [85] W. Moritz and D. Wolf, Structure determination of the reconstructed Au(110) surface, Surf. Sci. 88, L29 – L34 (1979).
- [86] D. dos Reis, F. Negreiros, V. de Carvalho, and E. Soares, Geometry of the Au(110)-(1×2) missing-row clean surface: A New LEED and DFT study, Surf. Sci. 604, 568 - 573 (2010).
- [87] C. Wöll, S. Chiang, R. J. Wilson, and P. H. Lippel, Determination of atom positions at stacking-fault dislocations on Au(111) by scanning tunneling microscopy, Phys. Rev. B 39, 7988–7991 (1989).
- [88] J. V. Barth, H. Brune, G. Ertl, and R. J. Behm, Scanning tunneling microscopy observations on the reconstructed Au(111) surface: Atomic structure, long-range superstructure, rotational domains, and surface defects, Phys. Rev. B 42, 9307– 9318 (1990).
- [89] P. Havu, V. Blum, V. Havu, P. Rinke, and M. Scheffler, Large-scale surface reconstruction energetics of Pt(100) and Au(100) by all-electron density functional theory, Phys. Rev. B 82, 161418 (2010).
- [90] P. W. Palmberg and T. N. Rhodin, Surface Structure of Clean Au (100) and Ag (100) Surfaces, Phys. Rev. 161, 586–588 (1967).
- [91] M. Kiguchi, G. Yoshikawa, S. Ikeda, and K. Saiki, Molecular orientation control of sexithienyl thin film on Cu substrates, Surf. Sci. 566-568, Part 1, 603 – 607, proceedings of the 22nd European Conference on Surface Science (2004).
- [92] M. Kiguchi, S. Entani, K. Saiki, and G. Yoshikawa, One-dimensional ordered structure of alpha-sexithienyl on Cu(110), Appl. Phys. Lett. 84, 3444–3446 (2004).
- [93] G. Yoshikawa, M. Kiguchi, S. Ikeda, and K. Saiki, Molecular orientations and adsorption structures of alpha-sexithienyl thin films grown on Ag(110) and Ag(111) surfaces, Surf. Sci. (2004).
- [94] A. J. Mäkinen, J. P. Long, N. J. Watkins, and Z. H. Kafafi, Sexithiophene Adlayer Growth on Vicinal Gold Surfaces, J. Phys. Chem. B 109, 5790–5795, pMID: 16851630 (2005).

- [95] K. Duncker, M. Kiel, A. Höfer, and W. Widdra, Commensurate surface structures and concerted cis-trans-isomerization within ordered monolayers of α -sexithiophene on Ag(001), Phys. Rev. B 77, 155423 (2008).
- [96] J. Repp, G. Meyer, S. M. Stojković, A. Gourdon, and C. Joachim, Molecules on Insulating Films: Scanning-Tunneling Microscopy Imaging of Individual Molecular Orbitals, Phys. Rev. Lett. 94, 026803 (2005).
- [97] W. Koch and M. C. Holthausen, A Chemist's Guide to Density Functional Theory, chapter 1, pages 3–18, John Wiley and Sons (2001).
- [98] M. Born and R. Oppenheimer, Zur Quantentheorie der Molekeln, Ann. Phys. 389, 457–484 (1927).
- [99] "proton-electron mass ratio". The NIST Reference on Constants, Units, and Uncertainty. (2014).
- [100] A. Szabo and N. S. Ostlund, Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory, chapter 2, pages 46–55, Dover Publications, New York (1996).
- [101] W. Koch and M. C. Holthausen, A Chemist's Guide to Density Functional Theory, chapter 2, pages 24–26, John Wiley and Sons (2001).
- [102] A. Szabo and N. S. Ostlund, Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory, chapter 3, page 133, Dover Publications, New York (1996).
- [103] A. Szabo and N. S. Ostlund, Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory, chapter 3, pages 123–127, Dover Publications, New York (1996).
- [104] T. C. Koopmans, Uber die Zuordnung von Wellenfunktionen und Eigenwerten zu den Einzelnen Elektronen Eines Atoms, Physica 1, 104 – 113 (1934).
- [105] P.-O. Löwdin, Correlation Problem in Many-Electron Quantum Mechanics I. Review of Different Approaches and Discussion of Some Current Ideas, pages 207–322, John Wiley & Sons, Ltd, doi: (1959).
- [106] W. Ritz, Über eine neue Methode zur Lösung gewisser Variationsprobleme der mathematischen Physik, Journal für die reine und angewandte Mathematik (Crelle's Journal) (1909).
- [107] T. Helgaker, P. Jorgensen, and J. Olsen, Molecular electronic structure theory, chapter 5, pages 181–196, Wiley (2008).
- [108] T. Helgaker, P. Jorgensen, and J. Olsen, *Molecular Electronic-structure Theory*, Wiley (2008).

- [109] W. Koch and M. C. Holthausen, A Chemist's Guide to Density Functional Theory, chapter 2, page 19, John Wiley and Sons (2001).
- [110] P. Hohenberg and W. Kohn, Inhomogeneous Electron Gas, Phys. Rev. 136, B864–B871 (1964).
- [111] E. Engel and R. M. Dreizler, Density Functional Theory An Advanced Course, chapter 2, pages 11–18, Springer Science & Business Media (2011).
- [112] R. M. Martin, Electronic structure- Basic Theory and practical methods, chapter 7, page 135–151, Cambridge University Press, doi: (2004).
- [113] J. E. Harriman, Orthonormal orbitals for the representation of an arbitrary density, Phys. Rev. A 24, 680–682 (1981).
- [114] T. L. Gilbert, Hohenberg-Kohn theorem for nonlocal external potentials, Phys. Rev. B 12, 2111–2120 (1975).
- [115] W. Koch and M. C. Holthausen, A Chemist's Guide to Density Functional Theory, chapter 4, pages 35–36, John Wiley and Sons (2001).
- [116] W. Kohn and L. J. Sham, Self-Consistent Equations Including Exchange and Correlation Effects, Phys. Rev. 140, A1133–A1138 (1965).
- [117] R. G. Parr and W. Yang, Density Functional Theory of Atoms and Molecules, chapter 7, pages 142–152, Oxford University Press (1989).
- [118] W. Koch and M. C. Holthausen, A Chemist's Guide to Density Functional Theory, chapter 6, pages 70–76, John Wiley and Sons (2001).
- [119] D. M. Ceperley and B. J. Alder, Ground State of the Electron Gas by a Stochastic Method, Phys. Rev. Lett. 45, 566–569 (1980).
- [120] M. Ernzerhof and G. E. Scuseria, Assessment of the Perdew-Burke-Ernzerhof exchange-correlation functional, J. Chem. Phys. 110, 5029–5036 (1999).
- [121] D. C. Patton, D. V. Porezag, and M. R. Pederson, Simplified generalized-gradient approximation and anharmonicity: Benchmark calculations on molecules, Phys. Rev. B 55, 7454–7459 (1997).
- [122] B. Hammer, K. W. Jacobsen, and J. K. Nørskov, Role of nonlocal exchange correlation in activated adsorption, Phys. Rev. Lett. 70, 3971–3974 (1993).
- [123] D. R. Hamann, Generalized Gradient Theory for Silica Phase Transitions, Phys. Rev. Lett. 76, 660–663 (1996).
- [124] J. P. Perdew, M. Ernzerhof, and K. Burke, *Rationale for mixing exact exchange with density functional approximations*, J. Chem. Phys. **105**, 9982–9985 (1996).
- [125] K. Burke, M. Ernzerhof, and J. P. Perdew, The adiabatic connection method: a non-empirical hybrid, Chem. Phys. Lett. 265, 115 – 120 (1997).

- [126] W. Kohn, A. D Becke, and R. G Parr, Density functional theory of electronic structure, J. Phys. Chem. 100, 12974–12980 (1996).
- [127] E. J. Baerends and O. V. Gritsenko, A Quantum Chemical View of Density Functional Theory, J. Phys. Chem. A 101, 5383–5403 (1997).
- [128] R. Stowasser and R. Hoffmann, What Do the Kohn-Sham Orbitals and Eigenvalues Mean?, J. Am. Chem. Soc. 121, 3414–3420 (1999).
- [129] E. J. Baerends, Perspective on "Self-consistent equations including exchange and correlation effects", Theor. Chem. Acc. 103, 265–269 (2000).
- [130] W. Koch and M. C. Holthausen, A Chemist's Guide to Density Functional Theory, chapter 5, pages 49–50, John Wiley and Sons (2001).
- [131] J. P. Perdew and M. Levy, Physical Content of the Exact Kohn-Sham Orbital Energies: Band Gaps and Derivative Discontinuities, Phys. Rev. Lett. 51, 1884– 1887 (1983).
- [132] A. Ambrosetti, N. Ferri, R. A. DiStasio, and A. Tkatchenko, Wavelike charge density fluctuations and van der Waals interactions at the nanoscale, Science 351, 1171–1176 (2016).
- [133] C. Corminboeuf, Minimizing Density Functional Failures for Non-Covalent Interactions Beyond van der Waals Complexes, Acc. Chem. Res. 47, 3217–3224, pMID: 24655016 (2014).
- [134] S. Kristyán and P. Pulay, Can (semi)local density functional theory account for the London dispersion forces?, Chem. Phys. Lett. 229, 175 – 180 (1994).
- [135] G. A. DiLabio and A. Otero-de-la Roza, Noncovalent Interactions in Density Functional Theory, chapter 1, pages 17–20, Wiley-Blackwell, doi: (2016).
- [136] J. C. Grossman, E. Schwegler, E. W. Draeger, F. Gygi, and G. Galli, Towards an assessment of the accuracy of density functional theory for first principles simulations of water, J. Chem. Phys. 120, 300–311 (2004).
- [137] B. Santra, J. c. v. Klimeš, D. Alfè, A. Tkatchenko, B. Slater, A. Michaelides, R. Car, and M. Scheffler, *Hydrogen Bonds and van der Waals Forces in Ice at Ambient and High Pressures*, Phys. Rev. Lett. **107**, 185701 (2011).
- [138] V. G. Ruiz, W. Liu, E. Zojer, M. Scheffler, and A. Tkatchenko, Density-Functional Theory with Screened van der Waals Interactions for the Modeling of Hybrid Inorganic-Organic Systems, Phys. Rev. Lett. 108, 146103 (2012).
- [139] S. Grimme, Accurate description of van der Waals complexes by density functional theory including empirical corrections, J. Comput. Chem. 25, 1463–1473 (2004).

- [140] S. Grimme, Semiempirical GGA-type density functional constructed with a longrange dispersion correction, J. Comput. Chem. 27, 1787–1799 (2006).
- [141] S. Grimme, J. Antony, S. Ehrlich, and H. Krieg, A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu, J. Chem. Phys. 132, 154104 (2010).
- [142] A. Tkatchenko and M. Scheffler, Accurate Molecular Van Der Waals Interactions from Ground-State Electron Density and Free-Atom Reference Data, Phys. Rev. Lett. 102, 073005 (2009).
- [143] A. D. Becke and E. R. Johnson, Exchange-hole dipole moment and the dispersion interaction, J. Chem. Phys. 122, 154104 (2005).
- [144] A. D. Becke and E. R. Johnson, A density-functional model of the dispersion interaction, J. Chem. Phys. 123, 154101 (2005).
- [145] Q. Wu and W. Yang, Empirical correction to density functional theory for van der Waals interactions, J. Chem. Phys. 116, 515–524 (2002).
- [146] T. Brinck, J. S. Murray, and P. Politzer, *Polarizability and volume*, J. Chem. Phys. 98, 4305–4306 (1993).
- [147] E. R. Johnson and A. D. Becke, A post-Hartree–Fock model of intermolecular interactions, J. Chem. Phys. 123, 024101 (2005).
- [148] A. Olasz, K. Vanommeslaeghe, A. Krishtal, T. Veszprémi, C. Van Alsenoy, and P. Geerlings, *The use of atomic intrinsic polarizabilities in the evaluation of the dispersion energy*, J. Chem. Phys. **127**, 224105 (2007).
- [149] E. M. Lifshitz, The Theory of Molecular Attractive Forces between Solids, Sov. Phys. JETP 2, 73 (1956).
- [150] E. Zaremba and W. Kohn, Van der Waals interaction between an atom and a solid surface, Phys. Rev. B 13, 2270 (1976).
- [151] A. Kumar, K. Banerjee, M. Dvorak, F. Schulz, A. Harju, P. Rinke, and P. Liljeroth, *Charge-Transfer-Driven Nonplanar Adsorption of F4TCNQ Molecules on Epitaxial Graphene*, ACS Nano **11**, 4960–4968, pMID: 28467831 (2017).
- [152] R. J. Maurer, V. G. Ruiz, J. Camarillo-Cisneros, W. Liu, N. Ferri, K. Reuter, and A. Tkatchenko, Adsorption structures and energetics of molecules on metal surfaces: Bridging experiment and theory, ProgSurfSc 91, 72–100 (2016).
- [153] P. Hofmann, Solid State Physic an Introduction, chapter 1, pages 2–3, Wiley-VCH (2016).
- [154] F. Bloch, Uber die Quantenmechanik der Elektronen in Kristallgittern, Zeitschrift für Physik 52, 555–600 (1929).

- [155] S. Elliott, The Physics and Chemistry of Solids, chapter 5, page 312, Wiley (1998).
- [156] H. Ibach and H. Lüth, Solid State Physics, chapter 7, pages 131–132, Springer (1996).
- [157] R. d. L. Kronig and W. G. Penney, Quantum Mechanics of Electrons in Crystal Lattices, Proc. R. Soc. Lond. A 130, 499–513 (1931).
- [158] H. Ibach and H. Lüth, Solid-State Physics, chapter 6, pages 137–160, Springer Berlin Heidelberg, Berlin, Heidelberg, doi: (2003).
- [159] D. S. Sholl and J. A. Steckel, Nuts and Bolts of DFT Calculations, chapter 3, page 53, John Wiley & Sons, Ltd, doi: (2009).
- [160] H. J. Monkhorst and J. D. Pack, Special points for Brillouin-zone integrations, Phys. Rev. B 13, 5188–5192 (1976).
- [161] J. Q. Quinn and Y. Kyung-Soo, Solid State Physics Principles and Modern Applications, chapter Elements of Band Theory, page 123, Springer-Verlag Berlin Heidelberg, doi: (2009).
- [162] R. E. Hummel, *Electronic Properties of Materials*, chapter Electrons in a Crystal, pages 63–65, Springer-Verlag New York (2000).
- [163] H. Ibach and H. Lüth, *Halbleiter*, chapter 12, pages 381–464, Springer Berlin Heidelberg, Berlin, Heidelberg, doi: (1999).
- [164] C. L. Fu and K. M. Ho, First-principles calculation of the equilibrium groundstate properties of transition metals: Applications to Nb and Mo, Phys. Rev. B 28, 5480–5486 (1983).
- [165] P. Pyykkö, Relativistic Effects in Chemistry: More Common Than You Thought, Annu. Rev. Phys. Chem. 63, 45–64, pMID: 22404585 (2012).
- [166] P. Pyykkö, Theoretical Chemistry of Gold, Angew. Chem. Int. Ed. 43, 4412– 4456 (2004).
- [167] F. Calvo, E. Pahl, M. Wormit, and P. Schwerdtfeger, Evidence for Low-Temperature Melting of Mercury owing to Relativity, Angew. Chem. Int. Ed. 52, 7583–7585 (2013).
- [168] R. Ahuja, A. Blomqvist, P. Larsson, P. Pyykkö, and P. Zaleski-Ejgierd, *Relativity and the Lead-Acid Battery*, Phys. Rev. Lett. **106**, 018301 (2011).
- [169] E. van Lenthe, The ZORA equation, Ph.D. thesis, Vrije University (1996).
- [170] V. Blum, R. Gehrke, F. Hanke, P. Havu, V. Havu, X. Ren, K. Reuter, and M. Scheffler, Ab initio molecular simulations with numeric atom-centered orbitals, Comput. Phys. Commun. 180, 2175 – 2196 (2009).

- [171] J. Desclaux, A multiconfiguration relativistic DIRAC-FOCK program, Comput. Phys. Commun. 9, 31 – 45 (1975).
- [172] P. Blaha, K. Schwarz, P. Sorantin, and S. Trickey, Full-potential, linearized augmented plane wave programs for crystalline systems, Comput. Phys. Commun. 59, 399 – 415 (1990).
- [173] W. Koch and M. C. Holthausen, A Chemist's Guide to Density Functional Theory, chapter 7, pages 97–98, John Wiley and Sons (2001).
- [174] F. Jensen, Introduction to Computational Chemistry, chapter 5, pages 194–195, Wiley (2007).
- [175] H. B. Jansen and P. Ros, Non-empirical molecular orbital calculations on the protonation of carbon monoxide, Chem. Phys. Lett. 3, 140–143 (1969).
- [176] B. Liu and A. D. McLean, Accurate calculation of the attractive interaction of two ground state helium atoms, J. Chem. Phys. 59, 4557–4558 (1973).
- [177] M. C. Payne, M. P. Teter, D. C. Allan, T. A. Arias, and J. D. Joannopoulos, Iterative minimization techniques for ab initio total-energy calculations: molecular dynamics and conjugate gradients, Rev. Mod. Phys. 64, 1045–1097 (1992).
- [178] L. Kleinman and D. M. Bylander, Efficacious Form for Model Pseudopotentials, Phys. Rev. Lett. 48, 1425–1428 (1982).
- [179] W. E. Pickett, Pseudopotential methods in condensed matter applications, Computer Physics Reports 9, 115 – 197 (1989).
- [180] D. R. Hamann, M. Schlüter, and C. Chiang, Norm-Conserving Pseudopotentials, Phys. Rev. Lett. 43, 1494–1497 (1979).
- [181] S. Goedecker, M. Teter, and J. Hutter, Separable dual-space Gaussian pseudopotentials, Phys. Rev. B 54, 1703–1710 (1996).
- [182] C. Hartwigsen, S. Goedecker, and J. Hutter, *Relativistic separable dual-space Gaussian pseudopotentials from H to Rn*, Phys. Rev. B 58, 3641–3662 (1998).
- [183] M. Krack, Pseudopotentials for H to Kr optimized for gradient-corrected exchange-correlation functionals, Theor. Chem. Acc. **114**, 145–152 (2005).
- [184] A. Warshel and M. Levitt, Theoretical studies of enzymic reactions: Dielectric, electrostatic and steric stabilization of the carbonium ion in the reaction of lysozyme, J. Mol. Biol. 103, 227 – 249 (1976).
- [185] F. Maseras and K. Morokuma, IMOMM: A new integrated ab initio + molecular mechanics geometry optimization scheme of equilibrium structures and transition states, J. Comput. Chem. 16, 1170–1179 (1995).

- [186] M. Svensson, S. Humbel, R. D. J. Froese, T. Matsubara, S. Sieber, and K. Morokuma, ONIOM: A Multilayered Integrated MO + MM Method for Geometry Optimizations and Single Point Energy Predictions. A Test for Diels-Alder Reactions and Pt(P(t-Bu)3)2 + H2 Oxidative Addition, J. Phys. Chem. 100, 19357–19363 (1996).
- [187] P. Sherwood, A. H. de Vries, M. F. Guest, G. Schreckenbach, C. A. Catlow, S. A. French, A. A. Sokol, S. T. Bromley, W. Thiel, A. J. Turner, S. Billeter, F. Terstegen, S. Thiel, J. Kendrick, S. C. Rogers, J. Casci, M. Watson, F. King, E. Karlsen, M. Sjøvoll, A. Fahmi, A. Schäfer, and C. Lennartz, QUASI: A general purpose implementation of the QM/MM approach and its application to problems in catalysis, J. Mol. Struct. THEOCHEM 632, 1 28 (2003).
- [188] H. M. Senn and W. Thiel, QM/MM Methods for Biomolecular Systems, Angew. Chem. Int. Ed. 48, 1198–1229 (2009).
- [189] G. Groenhof, Biomolecular Simulations, chapter 3, pages 43–66, Springer Protocols (2013).
- [190] T. Laino, F. Mohamed, A. Laio, and M. Parrinello, An Efficient Real Space Multigrid QM/MM Electrostatic Coupling, J. Chem. Theory Comput. 1, 1176– 1184, pMID: 26631661 (2005).
- [191] S. W. Rick and S. J. Stuart, Potentials and Algorithms for Incorporating Polarizability in Computer Simulations, chapter 3, pages 89–146, John Wiley & Sons, Ltd, doi: (2003).
- [192] A. Warshel, M. Kato, and A. V. Pisliakov, *Polarizable Force Fields: History, Test Cases, and Prospects*, J. Chem. Theory Comput. **3**, 2034–2045, pMID: 26636199 (2007).
- [193] T. Yokoyama, T. Takahashi, K. Shinozaki, and M. Okamoto, Quantitative Analysis of Long-Range Interactions between Adsorbed Dipolar Molecules on Cu(111), Phys. Rev. Lett. 98, 206102 (2007).
- [194] G. Tomba, M. Stengel, W.-D. Schneider, A. Baldereschi, and A. De Vita, Supramolecular Self-Assembly Driven by Electrostatic Repulsion: The 1D Aggregation of Rubrene Pentagons on Au(111), ACS Nano 4, 7545–7551, pMID: 21080665 (2010).
- [195] J. I. Siepmann and M. Sprik, Influence of surface topology and electrostatic potential on water/electrode systems, J. Chem. Phys. 102, 511–524 (1995).
- [196] D. Golze, M. Iannuzzi, M.-T. Nguyen, D. Passerone, and J. Hutter, Simulation of Adsorption Processes at Metallic Interfaces: An Image Charge Augmented QM/MM Approach, J. Chem. Theory Comput. 9, 5086–5097, pMID: 26583423 (2013).

- [197] D. Golze, J. Hutter, and M. Iannuzzi, Wetting of water on hexagonal boron nitride@Rh(111): a QM/MM model based on atomic charges derived for nanostructured substrates, Phys. Chem. Chem. Phys. 17, 14307–14316 (2015).
- [198] M. T. C. Martins-Costa and M. F. Ruiz-López, Reaching multi-nanosecond timescales in combined QM/MM molecular dynamics simulations through parallel horsetail sampling, J. Comput. Chem. 38, 659–668 (2017).
- [199] J. A. D. MacKerell, D. Bashford, M. Bellott, J. R. L. Dunbrack, J. D. Evanseck, M. J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, D. Joseph-McCarthy, L. Kuchnir, K. Kuczera, F. T. K. Lau, C. Mattos, S. Michnick, T. Ngo, D. T. Nguyen, B. Prodhom, W. E. Reiher, B. Roux, M. Schlenkrich, J. C. Smith, R. Stote, J. Straub, M. Watanabe, J. Wírkiewicz-Kuczera, D. Yin, and M. Karplus, All-Atom Empirical Potential for Molecular Modeling and Dynamics Studies of Proteins, J. Phys. Chem. B 102, 3586–3616, pMID: 24889800 (1998).
- [200] L. Monticelli and D. P. Tieleman, Force Fields for Classical Molecular Dynamics, pages 197–213, Humana Press, Totowa, NJ, doi: (2013).
- [201] M. J. Field, *Molecular mechanics*, chapter 5, page 87, Cambridge University Press, 2 edition, doi: (2007).
- [202] A. Leach, Molecular Modelling: Principles and Applications, chapter 4, pages 168–169, Pearson (2001).
- [203] W. D. Cornell, P. Cieplak, C. I. Bayly, I. R. Gould, K. M. Merz, D. M. Ferguson, D. C. Spellmeyer, T. Fox, J. W. Caldwell, and P. A. Kollman, A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules, J. Am. Chem. Soc. 117, 5179–5197 (1995).
- [204] N. L. Allinger, Conformational analysis. 130. MM2. A hydrocarbon force field utilizing V1 and V2 torsional terms, J. Am. Chem. Soc. 99, 8127–8134 (1977).
- [205] N. L. Allinger, Y. H. Yuh, and J. H. Lii, Molecular mechanics. The MM3 force field for hydrocarbons. 1, J. Am. Chem. Soc. 111, 8551–8566 (1989).
- [206] W. L. Jorgensen, D. S. Maxwell, and J. Tirado-Rives, Development and Testing of the OPLS All-Atom Force Field on Conformational Energetics and Properties of Organic Liquids, J. Am. Chem. Soc. 118, 11225–11236 (1996).
- [207] M. E. Tuckerman, Statistical Mechanics: Theory and Molecular Simulation, chapter 1, pages 5–7, Oxford University Press (2010).
- [208] M. E. Tuckerman, Statistical Mechanics: Theory and Molecular Simulation, chapter 1, pages 18–19, Oxford University Press (2010).
- [209] M. J. Field, Molecular mechanics, chapter 9, pages 173–175, Cambridge University Press, 2 edition, doi: (2007).

- [210] J.-P. Ryckaert, G. Ciccotti, and H. J. C. Berendsen, Numerical Integration of the Cartesian Equations of Motion of a System with Constraints: Molecular Dynamics of n-Alkanes, J. Comput. Phys. 23, 327–341 (1977).
- [211] M. J. Field, *Molecular mechanics*, chapter 11, page 260, Cambridge University Press, 2 edition, doi: (2007).
- [212] M. P. Allen and D. J. Tildesley, Computer Simulations of Liquids, chapter 1, pages 23–25, Oxford Science Publications (1989).
- [213] M. P. Allen and D. J. Tildesley, Computer Simulations of Liquids, chapter 5, pages 156–162, Oxford Science Publications (1989).
- [214] P. P. Ewald, Die Berechnung optischer und elektrostatischer Gitterpotentiale, Ann. Phys. 369, 253–287 (1921).
- [215] E. Madelung, Das elektrische Feld in Systemen von regelmäßig angeordneten Punktladungen, Phys. Z. 19, 524–533 (1918).
- [216] M. P. Allen and D. J. Tildesley, Computer Simulation of Liquids, Clarendon Press, New York, NY, USA (1989).
- [217] S. W. de Leeuw, J. W. Perram, E. R. Smith, and J. S. Rowlinson, Simulation of electrostatic systems in periodic boundary conditions. I. Lattice sums and dielectric constants, Proceedings of the Royal Society of London. A. Mathematical and Physical Sciences 373, 27–56 (1980).
- [218] M. W. Deem, J. M. Newsam, and S. K. Sinha, The h = 0 term in Coulomb sums by the Ewald transformation, J. Phys. Chem. **94**, 8356–8359 (1990).
- [219] M. J. Field, Molecular mechanics, chapter 10, pages 195–197, Cambridge University Press, 2 edition, doi: (2007).
- [220] M. E. Tuckerman, Statistical Mechanics: Theory and Molecular Simulation, chapter 2, pages 61–63, Oxford University Press (2010).
- [221] M. E. Tuckerman, Statistical Mechanics: Theory and Molecular Simulation, chapter 3-5, pages 74–260, Oxford University Press (2010).
- [222] A. R. Leach, Molecular Modelling Principles and Applications, chapter 7, pages 383–384, Pearson Education Limited (2001).
- [223] H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, A. DiNola, and J. R. Haak, *Molecular dynamics with coupling to an external bath*, J. Chem. Phys. 81, 3684–3690 (1984).
- [224] A. Cheng and K. M. Merz, Application of the Nosé-Hoover Chain Algorithm to the Study of Protein Dynamics, J. Phys. Chem. 100, 1927–1937 (1996).

- [225] S. C. Harvey, R. K.-Z. Tan, and T. E. Cheatham, The flying ice cube: Velocity rescaling in molecular dynamics leads to violation of energy equipartition, J. Comput. Chem. 19, 726–740 (1998).
- [226] J. C. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R. D. Skeel, L. Kalé, and K. Schulten, *Scalable molecular dynamics with NAMD*, J. Comput. Chem. **26**, 1781–1802 (2005).
- [227] M. Kubo, R.; Toda and N. Hashitsume, Statistical Physics II: Nonequilibrium Statistical Mechanics, Springer: New York, 2 edition (1991).
- [228] S. E. Feller, Y. Zhang, R. W. Pastor, and B. R. Brooks, Constant pressure molecular dynamics simulation: The Langevin piston method, J. Chem. Phys. 103, 4613–4621 (1995).
- [229] M. A. van der Horst and K. J. Hellingwerf, Photoreceptor Proteins, "Star Actors of Modern Times": A Review of the Functional Dynamics in the Structure of Representative Members of Six Different Photoreceptor Families, Acc. Chem. Res. 37, 13–20 (2004).
- [230] N. C. Rockwell, Y. S. Su, and J. C. Lagarias, *Phytochrome structure and signaling mechanisms*, Annu Rev Plant Biol. 57 (2006).
- [231] N. C. Rockwell and J. C. Lagarias, A Brief History of Phytochromes, ChemPhysChem 11, 1172–1180 (2010).
- [232] H. A. Borthwick, S. B. Hendricks, M. W. Parker, E. H. Toole, and V. K. Toole, A Reversible Photoreaction Controlling Seed Germination, Proc. Natl. Acad. Sci. USA 38, 662–666 (1952).
- [233] R. D. Vierstra and B. Karniol, *Phytochromes in Microorganisms*, In *Handbook of Photosensory Receptors*, chapter 8. Phytochromes in Microorganisms, pages 171–195. Wiley-VCH Verlag GmbH &; Co. KGaA (2005).
- [234] S. J. Davis, A. V. Vener, and R. D. Vierstra, Bacteriophytochromes: Phytochrome-Like Photoreceptors from Nonphotosynthetic Eubacteria, Science 286, 2517–2520 (1999).
- [235] J. Hughes, T. Lamparter, F. Mittmann, E. Hartmann, W. Gärtner, A. Wilde, and T. Börner, A prokaryotic phytochrome, Nature 386, 663 (1997).
- [236] T. Lamparter, Evolution of cyanobacterial and plant phytochromes, FEBS Lett. 573, 1 – 5 (2004).
- [237] K.-C. Yeh, S.-H. Wu, J. T. Murphy, and J. C. Lagarias, A Cyanobacterial Phytochrome Two-Component Light Sensory System, Science 277, 1505–1508 (1997).

- [238] A. Blumenstein, K. Vienken, R. Tasler, J. Purschwitz, D. Veith, N. Frankenberg-Dinkel, and R. Fischer, *The Aspergillus nidulans Phytochrome FphA Represses Sexual Development in Red Light*, Curr. Biol. **15**, 1833 – 1838 (2005).
- [239] R. Narikawa, F. Suzuki, S. Yoshihara, S.-i. Higashi, M. Watanabe, and M. Ikeuchi, Novel Photosensory Two-Component System (PixA-NixB-NixC) Involved in the Regulation of Positive and Negative Phototaxis of Cyanobacterium Synechocystis sp. PCC 6803, Plant Cell Physiol. 52, 2214–2224 (2011).
- [240] S. Yoshihara, F. Suzuki, H. Fujita, X. X. Geng, and M. Ikeuchi, Novel Putative Photoreceptor and Regulatory Genes Required for the Positive Phototactic Movement of the Unicellular Motile Cyanobacterium Synechocystis sp. PCC 6803, Plant Cell Physiol. 41, 1299–1304 (2000).
- [241] J.-Y. Song, H. S. Cho, J.-I. Cho, J.-S. Jeon, J. C. Lagarias, and Y.-I. Park, Near-UV cyanobacteriochrome signaling system elicits negative phototaxis in the cyanobacterium Synechocystis sp. PCC 6803, Proc. Natl. Acad. Sci. USA pages 10780–10785 (2011).
- [242] D. M. Kehoe and A. R. Grossman, Similarity of a Chromatic Adaptation Sensor to Phytochrome and Ethylene Receptors, Science 273, 1409–1412 (1996).
- [243] J. R. Wagner, J. S. Brunzelle, K. T. Forest, and R. D. Vierstra, A light-sensing knot revealed by the structure of the chromophore-binding domain of phytochrome, Nature 438, 325–331 (2005).
- [244] X. Yang, J. Kuk, and K. Moffat, Crystal structure of Pseudomonas aeruginosa bacteriophytochrome: Photoconversion and signal transduction, Proc. Natl. Acad. Sci. USA 105, 14715–14720 (2008).
- [245] K. Anders and L.-O. Essen, The family of phytochrome-like photoreceptors: diverse, complex and multi-colored, but very useful, Curr. Opin. Struct. Biol. 35, 7 - 16 (2015).
- [246] N. C. Rockwell, S. S. Martin, and J. C. Lagarias, *Identification of DXCF cyanobacteriochrome lineages with predictable photocycles*, Photochem. Photobiol. Sci. 14, 929–941 (2015).
- [247] L.-O. Essen, J. Mailliet, and J. Hughes, The structure of a complete phytochrome sensory module in the Pr ground state, Proc. Natl. Acad. Sci. USA 105, 14709– 14714 (2008).
- [248] Y. Hirose, T. Shimada, R. Narikawa, M. Katayama, and M. Ikeuchi, Cyanobacteriochrome CcaS is the green light receptor that induces the expression of phycobilisome linker protein, Proc. Natl. Acad. Sci. USA 105, 9528–9533 (2008).
- [249] N. C. Rockwell, S. L. Njuguna, L. Roberts, E. Castillo, V. L. Parson, S. Dwojak, J. C. Lagarias, and S. C. Spiller, A Second Conserved GAF Domain Cysteine Is

Required for the Blue/Green Photoreversibility of Cyanobacteriochrome Tlr0924 from Thermosynechococcus elongatus, Biochemistry **47**, 7304–7316, pMID: 18549244 (2008).

- [250] G. Enomoto, Y. Hirose, R. Narikawa, and M. Ikeuchi, *Thiol-Based Photocycle of the Blue and Teal Light-Sensing Cyanobacteriochrome Tlr1999*, Biochemistry 51, 3050–3058 (2012).
- [251] E. Burgie, J. Walker, G. P. Jr., and R. Vierstra, A Photo-Labile Thioether Linkage to Phycoviolobilin Provides the Foundation for the Blue/Green Photocycles in DXCF-Cyanobacteriochromes, Structure 21, 88–97 (2013).
- [252] R. Narikawa, G. Enomoto, N. Ni-Win, K. Fushimi, and M. Ikeuchi, A New Type of Dual-Cys Cyanobacteriochrome GAF Domain Found in Cyanobacterium Acaryochloris marina, Which Has an Unusual Red/Blue Reversible Photoconversion Cycle, Biochemistry 53, 5051–5059 (2014).
- [253] N. C. Rockwell, S. S. Martin, and J. C. Lagarias, There and Back Again: Loss and Reacquisition of Two-Cys Photocycles in Cyanobacteriochromes, Photochem. Photobiol. 93, 741–754 (2017).
- [254] S. M. Gottlieb, P. W. Kim, S. C. Corley, D. Madsen, S. J. Hanke, C.-W. Chang, N. C. Rockwell, S. S. Martin, J. C. Lagarias, and D. S. Larsen, *Primary* and Secondary Photodynamics of the Violet/Orange Dual-Cysteine NpF2164g3 Cyanobacteriochrome Domain from Nostoc punctiforme, Biochemistry 53, 1029–1040, pMID: 24437620 (2014).
- [255] J. E. Kim, M. J. Tauber, and R. A. Mathies, Wavelength Dependent Cis-Trans Isomerization in Vision, Biochemistry 40, 13774–13778, pMID: 11705366 (2001).
- [256] V. A. Sineshchekov, S. Frances, and M. J. White, Fluorescence and photochemical characterization of phytochrome in de-etiolated pea mutant lip, J. Photochem. Photobiol., B 28, 47 – 51 (1995).
- [257] K. Anders, D. von Stetten, J. Mailliet, S. Kiontke, V. A. Sineshchekov, P. Hildebrandt, J. Hughes, and L.-O. Essen, Spectroscopic and Photochemical Characterization of the Red-Light Sensitive Photosensory Module of Cph2 from Synechocystis PCC 6803, Photochem. Photobiol. 87, 160–173 (2011).
- [258] M. E. Auldridge, K. A. Satyshur, D. M. Anstrom, and K. T. Forest, Structureguided engineering enhances a phytochrome-based infrared fluorescent protein, J. Biol. Chem. (2011).
- [259] P. Schmidt, T. Gertsch, A. Remberg, W. Gärtner, S. E. Braslavsky, and K. Schaffner, *The Complexity of the Pr to Pfr Phototransformation Kinetics Is an Intrinsic Property of Native Phytochrome*, Photochem. Photobiol. 68, 754–761 (1998).

- [260] C. Song, L.-O. Essen, W. Gärtner, J. Hughes, and J. Matysik, Solid-State NMR Spectroscopic Study of Chromophore-Protein Interactions in the Pr Ground State of Plant Phytochrome A, Molecular Plant 5, 698 – 715 (2012).
- [261] F. Velazquez Escobar, T. Utesch, R. Narikawa, M. Ikeuchi, M. A. Mroginski, W. Gärtner, and P. Hildebrandt, *Photoconversion Mechanism of the Second GAF Domain of Cyanobacteriochrome AnPixJ and the Cofactor Structure of Its Green-Absorbing State*, Biochemistry **52**, 4871–4880 (2013).
- [262] M. A. Mroginski, D. von Stetten, F. V. Escobar, H. M. Strauss, S. Kaminski, P. Scheerer, M. Gänther, D. H. Murgida, P. Schmieder, C. Bongards, W. Gärtner, J. Mailliet, J. Hughes, L.-O. Essen, and P. Hildebrandt, *Chromophore Structure of Cyanobacterial Phytochrome Cph1 in the Pr State: Reconciling Structural and Spectroscopic Data by QM/MM Calculations*, Biophys. J. 96, 4153–4163 (2009).
- [263] T. Rohmer, C. Lang, J. Hughes, L.-O. Essen, W. Gärtner, and J. Matysik, Lightinduced chromophore activity and signal transduction in phytochromes observed by 13C and 15N magic-angle spinning NMR, Proc. Natl. Acad. Sci. USA 105, 15229–15234 (2008).
- [264] A. T. Ulijasz, G. Cornilescu, D. von Stetten, S. Kaminski, M. A. Mroginski, J. Zhang, D. Bhaya, P. Hildebrandt, and R. D. Vierstra, *Characterization of Two Thermostable Cyanobacterial Phytochromes Reveals Global Movements in the Chromophore-binding Domain during Photoconversion*, J. Biol. Chem. 283, 21251–21266 (2008).
- [265] D. J. Heyes, B. Khara, M. Sakuma, S. J. O. Hardman, R. O'Cualain, S. E. J. Rigby, and N. S. Scrutton, Ultrafast Red Light Activation of *italiciSynechocystis*/*italici Phytochrome Cph1 Triggers Major Structural* Change to Form the Pfr Signalling-Competent State, PLoS One 7, e52418 (2012).
- [266] R. A. Matute, R. Contreras, G. Pérez-Hernández, and L. González, The chromophore structure of the cyanobacterial phytochrome Cph1 as predicted by time-dependent density functional theory., J. Phys. Chem. B 112, 16253–6 (2008).
- [267] S. Kaminski and M. A. Mroginski, Molecular Dynamics of Phycocyanobilin Binding Bacteriophytochromes: A Detailed Study of Structural and Dynamic Properties, J. Phys. Chem. B 114, 16677–16686 (2010).
- [268] H. Elgabarty, P. Schmieder, and D. Sebastiani, Unraveling the existence of dynamic water channels in light-harvesting proteins: alpha-C-phycocyanobilin in vitro, Chem. Sci. 4, 755–763 (2013).
- [269] X.-L. Xu, K.-H. Zhao, A. Höppner, and W. Gärtner, Three-dimensional structure of the parental and photoproduct state of GAF3 from Slr1393 of Synechocystis sp. PCC6803, To be published.

- [270] C. Wiebeler, A. G. Rao, W. Gärtner, and I. Schapiro, The Effective Conjugation Length Is Responsible for the Red/Green Spectral Tuning in the Cyanobacteriochrome Slr1393g3, Angew. Chem. Int. Ed. 58, 1934–1938 (2019).
- [271] L. K. Scarbath-Evers, Theoretical Studies on Embedded Chromophores, Master's thesis, Institute of Chemistry, Martin-Luther University Halle-Wittenberg (2015).
- [272] L. K. Scarbath-Evers, S. Jähnigen, H. Elgabarty, C. Song, R. Narikawa, J. Matysik, and D. Sebastiani, *Structural heterogeneity in a parent ground-state structure of AnPixJg2 revealed by theory and spectroscopy*, Phys. Chem. Chem. Phys. **19**, 13882–13894 (2017).
- [273] C. Song, F. Velazquez-Escobar, X.-L. Xu, R. Narikawa, M. Ikeuchi, F. Siebert, W. Gärtner, J. Matysik, and P. Hildebrandt, A Red/Green Cyanobacteriochrome Sustains Its Color Despite a Change in the Bilin Chromophore's Protonation State, Biochemistry 54, 5839–5848, pMID: 26335286 (2015).
- [274] S. D. Bruce, J. Higinbotham, I. Marshall, and P. H. Beswick, An Analytical Derivation of a Popular Approximation of the Voigt Function for Quantification of NMR Spectra, J. Magn. Reson. 142, 57–63 (2000).
- [275] C. Song, R. Narikawa, M. Ikeuchi, W. Gärtner, and J. Matysik, Color Tuning in Red/Green Cyanobacteriochrome AnPixJ: Photoisomerization at C15 Causes an Excited-State Destabilization, J. Phys. Chem. B 119, 9688–9695, pMID: 26115331 (2015).
- [276] E. S. Burgie, A. N. Bussell, J. M. Walker, K. Dubiel, and V. R. D., Crystal structure of the photosensing module from a red/far-red light-absorbing plant phytochrome, Proc. Natl. Acad. Sci. USA 111, 10179–10184 (2014).
- [277] J. Rivnay, R. M. Owens, and G. G. Malliaras, The Rise of Organic Bioelectronics, Chem. Mater. 26, 679–685 (2014).
- [278] J. F. Ponder, A. M. Österholm, and J. R. Reynolds, Conjugated Polyelectrolytes as Water Processable Precursors to Aqueous Compatible Redox Active Polymers for Diverse Applications: Electrochromism, Charge Storage, and Biocompatible Organic Electronics, Chem. Mater. 29, 4385–4392 (2017).
- [279] W. Brütting, *Physics of Organic Semiconductors*, chapter 1, page 1, WILEY-VCH Verlag (2005).
- [280] O. Ostroverkhova, Organic Optoelectronic Materials: Mechanisms and Applications, Chem. Rev. **116**, 13279–13412, pMID: 27723323 (2016).
- [281] D. A. da Silva Filho, E.-G. Kim, and J.-L. Brédas, Transport Properties in the Rubrene Crystal: Electronic Coupling and Vibrational Reorganization Energy, Adv. Mater. 17, 1072–1076 (2005).

- [282] G. R. Hutchison, M. A. Ratner, and T. J. Marks, Hopping Transport in Conductive Heterocyclic Oligomers: Reorganization Energies and Substituent Effects, J. Am. Chem. Soc. 127, 2339–2350, pMID: 15713114 (2005).
- [283] J. Cornil, J.-L. Brédas, J. Zaumseil, and H. Sirringhaus, Ambipolar Transport in Organic Conjugated Materials, Adv. Mater. 19, 1791–1799 (2007).
- [284] C. Wang, H. Dong, W. Hu, Y. Liu, and D. Zhu, Semiconducting π-Conjugated Systems in Field-Effect Transistors: A Material Odyssey of Organic Electronics, Chem. Rev. 112, 2208–2267, pMID: 22111507 (2012).
- [285] J. Yang, D. Yan, and T. S. Jones, Molecular Template Growth and Its Applications in Organic Electronics and Optoelectronics, Chem. Rev. 115, 5570–5603, pMID: 25970808 (2015).
- [286] B. Kan, M. Li, Q. Zhang, F. Liu, X. Wan, Y. Wang, W. Ni, G. Long, X. Yang, H. Feng, Y. Zuo, M. Zhang, F. Huang, Y. Cao, T. P. Russell, and Y. Chen, A Series of Simple Oligomer-like Small Molecules Based on Oligothiophenes for Solution-Processed Solar Cells with High Efficiency, J. Am. Chem. Soc. 137, 3886–3893, pMID: 25736989 (2015).
- [287] B. S. Ong, Y. Wu, P. Liu, and S. Gardner, *High-Performance Semiconducting Polythiophenes for Organic Thin-Film Transistors*, J. Am. Chem. Soc. **126**, 3378–3379, pMID: 15025437 (2004).
- [288] H. Glowatzki, S. Duhm, K.-F. Braun, J. P. Rabe, and N. Koch, *Molecular chains and carpets of sexithiophenes on* Au(111), Phys. Rev. B 76, 125425 (2007).
- [289] L. K. Scarbath-Evers, M. Todorović, D. Golze, R. Hammer, W. Widdra, D. Sebastiani, and P. Rinke, Gold diggers: Altered reconstruction of the gold surface by physisorbed aromatic oligomers, Phys. Rev. Materials 3, 011601 (2019).
- [290] A. Kühnle, L. M. Molina, T. R. Linderoth, B. Hammer, and F. Besenbacher, Growth of Unidirectional Molecular Rows of Cysteine on Au(110)-(1×2) Driven by Adsorbate-Induced Surface Rearrangements, Phys. Rev. Lett. 93, 086101 (2004).
- [291] F. Hiebel, B. Shong, W. Chen, R. J. Madix, E. Kaxiras, and C. M. Friend, Self-assembly of acetate adsorbates drives atomic rearrangement on the Au(110) surface, Nat. Commun. 7, 13139 (2016).
- [292] A. D. Jewell, H. L. Tierney, and E. C. H. Sykes, Gently lifting gold's herringbone reconstruction: Trimethylphosphine on Au(111), Phys. Rev. B 82, 205401 (2010).
- [293] N. Richardson, Two-dimensional crystals: Through the looking glass, Nat. Mater. 5, 91–92 (2005).

- [294] T. Wagner, D. R. Fritz, and P. Zeppenfeld, α -6T on Ag(110): The formation of the wetting layer, Synth. Met. **161**, 2006 2010 (2011).
- [295] G. Horowitz, B. Bachet, A. Yassar, P. Lang, F. Demanze, J.-L. Fave, and F. Garnier, Growth and Characterization of Sexithiophene Single Crystals, Chem. Mater. 7, 1337–1341 (1995).
- [296] T. Siegrist, R. Fleming, R. Haddon, R. Laudise, A. Lovinger, H. Katz, P. Bridenbaugh, and D. Davis, *The crystal structure of the high-temperature* polymorph of α-hexathienyl (α-6T/HT), J. Mater. Res. **10**, 2170–2173 (1995).
- [297] P. Hermet, J.-L. Bantignies, A. Rahmani, J.-L. Sauvajol, and M. R. Johnson, Polymorphism of Crystalline α-Quaterthiophene and α-Sexithiophene: Ab Initio Analysis and Comparison with Inelastic Neutron Scattering Response, J. Phys. Chem. A 109, 4202–4207, pMID: 16833746 (2005).
- [298] C. Simbrunner, Epitaxial growth of sexithiophene and para-hexaphenyl and its implications for the fabrication of self-assembled lasing nano-fibres, Semicond. Sci. Technol. 28 (2013).
- [299] G. Koller, S. Berkebile, J. Ivanco, F. Netzer, and M. Ramsey, *Device relevant organic films and interfaces: A surface science approach*, Surf. Sci. **601**, 5683 5689, wagga Symposium on Surfaces and Interfaces 2006 (2007).
- [300] L. Chen, H. Li, and A. T. S. Wee, One-Dimensional Molecular Chains with Dispersive Electronic States, Nano Lett. 9, 4292–4296, pMID: 19835386 (2009).
- [301] D. Nabok, P. Puschnig, and C. Ambrosch-Draxl, Cohesive and surface energies of π-conjugated organic molecular crystals: A first-principles study, Physical review. B, Condensed matter 77 (2008).
- [302] D.-M. Smilgies and E. J. Kintzel, *Epitaxial orientations of para-sexiphenyl* platelets grown on alkali halide (001) surfaces, Phys. Rev. B **79**, 235413 (2009).
- [303] J. Ivanco, T. Haber, J. Krenn, F. Netzer, R. Resel, and M. Ramsey, Sexithiophene films on ordered and disordered TiO2(110) surfaces: Electronic, structural and morphological properties, Surf. Sci. 601, 178 – 187 (2007).
- [304] T. Haber, J. Ivanco, M. Ramsey, and R. Resel, *Epitaxial growth of sexithiophene* on TiO2(110), J. Cryst. Growth **310**, 101 – 109 (2008).
- [305] M. Koini, T. Haber, S. Berkebile, G. Koller, M. Ramsey, R. Resel, and M. Oehzelt, Growth of sexithiophene crystals on Cu(110) and Cu(110)-(2×1)O stripe phase—The influence of surface corrugation, J. Cryst. Growth **311**, 1364 – 1369 (2009).
- [306] A. J. Fleming, S. Berkebile, T. Ules, and M. G. Ramsey, *Pre-nucleation dynamics of organic molecule self-assembly investigated by PEEM*, Phys. Chem. Chem. Phys. **13**, 4693–4708 (2011).

- [307] D. R. Hamann and J. Tersoff, Theory of the scanning tunneling microscope, Phys. Rev. B (1985).
- [308] R. S. Mulliken, Electronic Population Analysis on LCAO-MO Molecular Wave Functions. I, J. Chem. Phys. 23, 1833–1840 (1955).
- [309] E. C. Beret, L. M. Ghiringhelli, and M. Scheffler, *Free gold clusters: beyond the static, monostructure description*, Faraday Discuss. **152**, 153–167 (2011).
- [310] H. Jónsson, G. Mills, and K. W. Jacobsen, Nudged elastic band method for finding minimum energy paths of transitions, In B. J. Berne, G. Ciccotti, and D. F. Coker, editors, Classical and Quantum Dynamics in Condensed Phase Simulations, pages 385–404, doi: (1998).
- [311] S. Förster and W. Widdra, Structure of single polythiophene molecules on Au(001) prepared by in situ UHV electrospray deposition, J. Chem. Phys. 141, 054713 (2014).
- [312] D. Kolb, Reconstruction phenomena at metal-electrolyte interfaces, Prog. Surf. Sci. 51, 109 – 173 (1996).
- [313] M. Oehzelt, S. Berkebile, G. Koller, J. Ivanco, S. Surnev, and M. Ramsey, Alphasexithiophene on Cu(110) and Cu(110)-(2 x 1)O: an STM and NEXAFS study, Surf. Sci. 603, 412–418 (2009).
- [314] R. Telesca, H. Bolink, S. Yunoki, G. Hadziioannou, P. T. Van Duijnen, J. G. Snijders, H. T. Jonkman, and G. A. Sawatzky, *Density-functional study of the evolution of the electronic structure of oligomers of thiophene: Towards a model Hamiltonian*, Phys. Rev. B 63, 155112 (2001).
- [315] M. Wagner, S. Berkebile, F. P. Netzer, and M. G. Ramsey, Revealing the Buried Metal–Organic Interface: Restructuring of the First Layer by van der Waals Forces, ACS Nano 9, 12070–12078, pMID: 26505912 (2015).
- [316] K. N. Baker, A. V. Fratini, T. Resch, H. C. Knachel, W. Adams, E. Socci, and B. Farmer, Crystal structures, phase transitions and energy calculations of poly(p-phenylene) oligomers, Polymer 34, 1571 – 1587 (1993).
- [317] K. N. Baker, H. C. Knachel, A. V. Fratini, and W. W. Adams, Structural Transformations in Crystalline Oligomers of Polyparaphenylene, MRS Proceedings 134, 497 (1988).
- [318] L. Athouël, G. Froyer, M. Riou, and M. Schott, Structural studies of parasexiphenyl thin films: Importance of the deposition parameters, Thin Solid Films 274, 35 – 45 (1996).
- [319] R. Resel, N. Koch, F. Meghdadi, G. Leising, L. Athouel, G. Froyer, and F. Hofer, A Polymorph Crystal Structure of Hexaphenyl Observed in Thin Films, Cryst. Res. Technol. 36, 47–54 (2001).

- [320] M. Hollerer, S. Pachmajer, D. Lüftner, B. Butej, E.-M. Reinisch, P. Puschnig, G. Koller, M. G. Ramsey, and M. Sterrer, *Growth study of para-sexiphenyl on* Ag(001): From single molecule to crystalline film, Surf. Sci. (2018).
- [321] N. C. Rockwell, S. S. Martin, F. Gan, D. A. Bryant, and J. C. Lagarias, NpR3784 is the prototype for a distinctive group of red/green cyanobacteriochromes using alternative Phe residues for photoproduct tuning, Photochem. Photobiol. Sci. 14, 258–269 (2015).
- [322] W. Humphrey, A. Dalke, and K. Schulten, VMD Visual Molecular Dynamics, J. Mol. Graphics 14, 33–38 (1996).
- [323] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, and M. L. Klein, Comparison of simple potential functions for simulating liquid water, J. Chem. Phys. 79, 926–935 (1983).
- [324] A. D. Mackerell, M. Feig, and C. L. Brooks, Extending the treatment of backbone energetics in protein force fields: Limitations of gas-phase quantum mechanics in reproducing protein conformational distributions in molecular dynamics simulations, J. Comput. Chem. 25, 1400–1415 (2004).
- [325] G. J. Martyna, D. J. Tobias, and M. L. Klein, Constant pressure molecular dynamics algorithms, J. Chem. Phys. 101, 4177–4189 (1994).
- [326] T. Darden, D. York, and L. Pedersen, Particle mesh Ewald: An Nlog(N) method for Ewald sums in large systems, J. Chem. Phys. 98, 10089–10092 (1993).
- [327] N. Michaud-Agrawal, E. J. Denning, T. B. Woolf, and O. Beckstein, MDAnalysis: A toolkit for the analysis of molecular dynamics simulations, J. Comput. Chem. 32, 2319–2327 (2011).
- [328] T. Oliphant, NumPy: A guide to NumPy, USA: Trelgol Publishing (2006).
- [329] pandas: Python Data Analysis Library, Online (2012).
- [330] J. D. Hunter, Matplotlib: A 2D graphics environment, Computing In Science & Engineering 9, 90–95 (2007).
- [331] E. F. Pettersen, T. D. Goddard, C. C. Huang, G. Couch, D. M. Greenblatt, E. Meng, and T. E. Ferrin, UCSF Chimera-a visualization system for exploratory research and analysis., J Comput Chem. 13, 1605–12 (2004).
- [332] D. Chandler and A. Luzar, Hydrogen-bond kinetics in liquid water, Nature 379, 55–57 (1996).
- [333] A. Soper and M. Phillips, A new determination of the structure of water at 25° C, Chem. Phys. 107, 47 – 60 (1986).
- [334] B. Chen, I. Ivanov, M. L. Klein, and M. Parrinello, Hydrogen Bonding in Water, Phys. Rev. Lett. 91, 215503 (2003).

- [335] S. Lounis, Theory of Scanning Tunneling Microscopy, Lecture Notes of the 45th IFF Spring School "Computing Solids - Models, ab initio methods and supercomputing" (2014).
- [336] J. Bardeen, Tunnelling from a Many-Particle Point of View, Phys. Rev. Lett. 6, 57–59 (1961).
- [337] C. van Wüllen, Molecular density functional calculations in the regular relativistic approximation: Method, application to coinage metal diatomics, hydrides, fluorides and chlorides, and comparison with first-order relativistic calculations, J. Chem. Phys. 109, 392 (1998).
- [338] J. van Lenthe, S. Faas, and J. Snijders, Gradients in the ab initio scalar zerothorder regular approximation (ZORA) approach, Chem. Phys. Lett. 328, 107 – 112 (2000).
- [339] P. Heilmann, K. Heinz, and K. Müller, The superstructures of the clean Pt(100) and Ir(100) surfaces, Surf. Sci. 83, 487 – 497 (1979).
- [340] I. Horcas, R. Fernández, J. M. Gómez-Rodríguez, J. Colchero, J. Gómez-Herrero, and A. M. Baro, WSXM: A software for scanning probe microscopy and a tool for nanotechnology, Rev. Sci. Instrum. 78, 013705 (2007).
- [341] E. M. Lifshitz, The Theory of Molecular Attractive Forces between Solids, Sov. Phys. JETP 2, 73 (1956).
- [342] R. W. G. Wyckoff, Second edition. Interscience Publishers, New York, New York Cubic closest packed, ccp, structure, Crystal Structures 1, 7–83 (1963).
- [343] J. Hutter, M. Iannuzzi, F. Schiffmann, and J. VandeVondele, cp2k: atomistic simulations of condensed matter systems, WIREs Comput. Mol. Sci. 4, 15–25 (2014).
- [344] J. VandeVondele and J. Hutter, Gaussian basis sets for accurate calculations on molecular systems in gas and condensed phases, J. Chem. Phys. 127, 114105 (2007).
- [345] M. Kertesz, C. H. Choi, and S. Yang, Conjugated Polymers and Aromaticity, Chem. Rev. 105, 3448–3481, pMID: 16218558 (2005).
- [346] C. H. Choi, M. Kertesz, and A. Karpfen, The effects of electron correlation on the degree of bond alternation and electronic structure of oligomers of polyacetylene, J. Chem. Phys. 107, 6712–6721 (1997).
- [347] U. Salzner, Theoretical Investigation of Excited States of Oligothiophenes and of Their Monocations, J. Chem. Theory Comput. 3, 1143–1157, pMID: 26627433 (2007).
- [348] J. P. Perdew, K. Burke, and M. Ernzerhof, Generalized Gradient Approximation Made Simple, Phys. Rev. Lett. 77, 3865–3868 (1996).

- [349] M. Guidon, J. Hutter, and J. VandeVondele, Auxiliary Density Matrix Methods for Hartree-Fock Exchange Calculations, J. Chem. Theory Comput. 6, 2348– 2364, pMID: 26613491 (2010).
- [350] S. M. Foiles, M. I. Baskes, and M. S. Daw, Embedded-atom-method functions for the fcc metals Cu, Ag, Au, Ni, Pd, Pt, and their alloys, Phys. Rev. B 33, 7983–7991 (1986).
- [351] H. Heinz, R. A. Vaia, B. L. Farmer, and R. R. Naik, Accurate Simulation of Surfaces and Interfaces of Face-Centered Cubic Metals Using 12-6 and 9-6 Lennard-Jones Potentials, J. Phys. Chem. C 112, 17281–17290 (2008).
- [352] M. Moreno, C. Mosé, G. Raos, S. V. Meille, and R. Po, Molecular Modeling of Crystalline Alkylthiophene Oligomers and Polymers, J. Phys. Chem. B 114, 1591–1602, pMID: 20058895 (2010).
- [353] S. M. Wetterer, D. J. Lavrich, T. Cummings, S. L. Bernasek, and G. Scoles, Energetics and Kinetics of the Physisorption of Hydrocarbons on Au(111), J. Phys. Chem. B 102, 9266–9275 (1998).
- [354] G. Liu, J. A. Rodriguez, J. Dvorak, J. Hrbek, and T. Jirsak, Chemistry of sulfurcontaining molecules on Au(111): thiophene, sulfur dioxide, and methanethiol adsorption, Surf. Sci. 505, 295 – 307 (2002).
- [355] A. Nambu, H. Kondoh, I. Nakai, K. Amemiya, and T. Ohta, Film growth and X-ray induced chemical reactions of thiophene adsorbed on Au(111), Surf. Sci. 530, 101 – 110 (2003).
- [356] P. Redhead, *Thermal desorption of gases*, Vacuum **12**, 203 211 (1962).
- [357] W. Song, P. J. Rossky, and M. Maroncelli, Modeling alkane +perfluoroalkane interactions using all-atom potentials: Failure of the usual combining rules, J. Chem. Phys. 119, 9145–9162 (2003).
- [358] W. Janke, Computer Simulation Studies of Polymer Adsorption and Aggregation
 From Flexible to Stiff, Physics Proceedia 68, 69 79 (2015).
- [359] A. Nikitin, Y. Milchevskiy, and A. Lyubartsev, AMBER-ii: New Combining Rules and Force Field for Perfluoroalkanes, J. Phys. Chem. B 119, 14563–14573, pMID: 26498002 (2015).
- [360] S. Tsuzuki, K. Honda, and R. Azumi, Model Chemistry Calculations of Thiophene Dimer Interactions: Origin of π-Stacking, J. Am. Chem. Soc. 124, 12200–12209, pMID: 12371860 (2002).

C. Academic Curriculum Vitae

Laura Katharina Scarbath-Evers

Personal Information

Name Laura Katharina Scarbath-Evers Date and place of 17.05.1987, Filderstadt, Germany birth Nationality German

Education

10/2015-06/2019 (expected)	Doctoral Researcher (PhD) in Computational Chemistry , <i>Martin-Luther University Halle-Wittenberg, Germany</i> , Advisor: Prof. Dr. Daniel Sebastiani, Thesis Title: Conformation and Packing Structure of Conjugated Molecules under the Influence of Hard and Soft Constraints.
06/2018-08/2018	Mobility stay , <i>Group of Computational Electronic Structure Theory, Aalto University, Helsinki, Finland</i> , Topic: Electronic properties of sexithiophene on Au(100), Host: Prof. Patrick Rinke.
04/2017-12/2017	Mobility stay , <i>Group of Computational Electronic Structure Theory, Aalto University</i> , Helsinki, Finland, Topic: Single-molecule adsorption of sexithiophene on Au(100), Host: Prof. Patrick Rinke.
04/2013-08/2015	M.Sc. in Chemistry , <i>Leipzig University, Germany</i> , Advisor: Prof. Dr. Daniel Sebastiani, Thesis Title: Theoretical Studies on Embedded Chromophores.
02/2014-12/2014	Erasmus Exchange, Jagiellonian University Kraków, Poland.
	• Research Project "Identification and quantification of selected biomarkers of en- dothelial dysfunction using HPLC-ESI-MS/MS and FT/MS"
10/2009-12/2012	• Research Project "Anthracycline cytotoxicity studied <i>in vitro</i> by Raman microscopy" B.Sc. in Chemistry , <i>University of Tübingen, Germany</i> , Advisor: Prof. Dr. Reinhold Fink, Thesis Title: Quantenchemische Beschreibung der Struktur und der Oberflächenverstärkten-Raman-Spektroskopie von Mercaptobenzoth- iazoladsorbated auf Gold-(111)-Oberflächen.
	Professional Memberships
Gesellschaft Deutscher	2015-Present

Chemiker

Ludwig-Wucherer Str. 57 – 06108 Halle – Germany ☎ +49 (0)345 55 25 248 • ⊠ laura.scarbath-evers@chemie.uni-halle.de

Languages

German Native speaker English Advanced proficiency

Publications

- L. K. Scarbath-Evers, M. Todorović, D. Golze, R. Hammer, W. Widdra, D. Sebastiani and P. Rinke, "Gold diggers: Altered reconstruction of the gold surface by physisorbed aromatic oligomers", Phys. Rev. Materials 2019, 3
- L. K. Scarbath-Evers, S. Jähnigen, H. Elgabarty, C. Song, R. Narikawa, J. Matysik, and D. Sebastiani, "Structural heterogeneity in parent ground-state structure in AnPixJg2 revealed by theory and spectroscopy", Phys. Chem. Chem. Phys., 2017, 19, 13882-13894.
- L. K. Scarbath-Evers, P. A. Hunt, B. Kirchner, D. R. MacFarlane, and S. Zahn, "Molecular features contributing to the lower viscosity of phosphonium ionic liquids compared to their ammonium analogues", Phys. Chem. Chem. Phys., 2015, 17, 20205-20216.

Halle (Saale), March 19, 2019

Ich erkläre an Eides statt, dass ich die Arbeit selbstständig und ohne fremde Hilfe verfasst, keine anderen als die von mir angegebenen Quellen und Hilfsmittel benutzt und die den benutzten Werken wörtlich oder inhaltlich entnommenen Stellen als solche kenntlich gemacht habe.

I declare under penalty of perjury that this thesis is my own work entirely and has been written without any help from other people. I used only the sources mentioned and included all the citations correctly both in word or content.

Datum/Date

Unterschrift des Antragstellers/ Signature of the applicant