

LINEAR RESPONSE IN CONDENSED PHASE:  
SUPRAMOLECULAR STRUCTURE, DYNAMICS, AND CHIRALITY  
FROM FIRST PRINCIPLES

Dissertation

zur Erlangung des Doktorgrades der Naturwissenschaften  
(Dr. rer. nat.)

der

Naturwissenschaftlichen Fakultät II  
Chemie, Physik und Mathematik

der Martin-Luther-Universität  
Halle-Wittenberg

vorgelegt von

Sascha Jähnigen

geb. am 1. März 1989 in Karl-Marx-Stadt, jetzt Chemnitz



1. Gutachter: Prof. Dr. Daniel Sebastiani  
Martin-Luther-Universität Halle-Wittenberg

2. Gutachter: Prof. Dr. Laurence A. Nafie  
Syracuse University

Vorsitzende der Promotionskommission: Prof. Dr. Kirsten Bacia  
Martin-Luther-Universität Halle-Wittenberg

Bearbeitungszeitraum: 1. Oktober 2014 bis 7. März 2019

Tag der Verteidigung: 15. August 2019

Ich erkläre, dass ich diese Arbeit selbständig und ohne fremde Hilfe verfasst habe und andere als die von mir angegebenen Quellen und Hilfsmittel nicht benutzt und die den benutzten Werken wörtlich oder inhaltlich entnommenen Stellen als solche kenntlich gemacht habe. Diese Arbeit wurde keiner anderen Institution oder Einrichtung zur Begutachtung vorgelegt.

Halle, den 7. März 2019

Sascha Jähnigen



*"The dynamic nature of their philosophy laid more stress upon the process through which perfection was sought than upon perfection itself. True beauty could be discovered only by one who mentally completed the incomplete."*

– The Book of Tea by Kakuzo Okakura.



# CONTENTS

ACKNOWLEDGEMENTS v

PUBLICATIONS AND PRESENTATIONS vii

## INTRODUCTION AND THEORY

1	PREFACE	3
2	BASIC THEORY	7
2.1	Equations of Motion . . . . .	7
2.1.1	Exact Factorisation of the Electron-Nuclear Wavefunction . . . . .	7
2.1.2	Born-Oppenheimer Approximation . . . . .	9
2.1.3	Complete Adiabatic Approximation . . . . .	10
2.2	Electronic Structure . . . . .	13
2.2.1	Density Functional Theory . . . . .	13
2.2.2	Time-Dependent Density Functional Theory . . . . .	15
2.2.3	Density Functional Perturbation Theory . . . . .	16
2.3	Molecular Dynamics . . . . .	18
2.3.1	Nuclear Equations of Motion . . . . .	18
2.3.2	<i>Ab Initio</i> Molecular Dynamics . . . . .	19
2.3.3	Force Field Molecular Dynamics . . . . .	20
2.3.4	QM/MM . . . . .	21
2.4	Linear Response for Spectroscopy . . . . .	22
2.4.1	Magnetic Field Perturbation . . . . .	22
2.4.2	Nuclear Velocity Perturbation . . . . .	24
2.4.3	Electric Field Perturbation . . . . .	27

## CONTENTS

I VIBRATIONAL CIRCULAR DICHROISM AND SUPRAMOLECULAR CHIRALITY	
3	THE SAMPLING PROBLEM 31
4	L-ALANINE: ENHANCEMENT OF VCD IN MOLECULAR CRYSTALS 33
4.1	Gauge Invariance in the Crystalline Phase . . . . . 35
4.2	The Origins of VCD Enhancement . . . . . 38
4.2.1	Conformational Freedom . . . . . 38
4.2.2	Local and Supramolecular Scope . . . . . 39
4.3	Spatial Analysis . . . . . 41
4.3.1	Crystal Topology of L-Alanine . . . . . 41
4.3.2	Radially Resolved IR and VCD spectra . . . . . 42
4.3.3	Non-Local Vibrations . . . . . 45
4.4	Conclusion . . . . . 47
5	(S)-LACTIC ACID: NON-LOCAL VCD AND CHIRALITY IN SOLUTION 49
5.1	(S)-Lactic Acid in Aprotic Environment . . . . . 50
5.1.1	Evidence for Aggregation . . . . . 50
5.1.2	Monomer and Dimer . . . . . 51
5.1.3	Trimer and the Breaking of Symmetry . . . . . 55
5.2	Preliminary Considerations on VCD in Water . . . . . 57
5.2.1	Induced VCD – Induced Chirality . . . . . 57
5.2.2	An Adaptive Moment Scaling Scheme (“Background Correction”) . . . . . 58
5.3	(S)-Lactic Acid in Aqueous Solution . . . . . 60
5.3.1	Restructuring of the Hydrogen Bond Network . . . . . 61
5.3.2	Supramolecular VCD in Solution . . . . . 65
5.3.3	Chirality Induction and Solvation Model for Carboxylic Acids . . . . . 67
5.4	Conclusion . . . . . 69
II COLOUR TUNING OF PHYCOCYANOBILIN IN PROTEINS	
6	THE PHYTOCHROME SUPERFAMILY 73
6.1	A Case Study: P <sub>r</sub> States of CphiΔ <sub>2</sub> and AnPixJg <sub>2</sub> . . . . . 77
7	ANPIXJ AND CPHI: PROTEIN STRUCTURE AND DYNAMICS 79



7.1	AnPixJg2 . . . . .	79
7.1.1	Evidence for Microheterogeneity . . . . .	79
7.1.2	Restructuring of the Binding Pocket . . . . .	83
7.1.3	Hydrophobic Interaction and Solvation . . . . .	83
7.2	Cphi $\Delta$ 2 . . . . .	88
8	PCB: SPECTROSCOPIC CHARACTERISATION . . . . .	91
8.1	NMR Spectroscopy . . . . .	91
8.1.1	Sampling Statistical Ensembles . . . . .	91
8.1.2	Size of the QM Part and Convergence . . . . .	92
8.1.3	Charge and Conjugation Pattern of PCB . . . . .	94
8.1.4	Information Conveyed by $^{13}\text{C}$ Chemical Shifts . . . . .	95
8.1.5	$^{15}\text{N}$ and $^1\text{H}$ Chemical Shifts: Sensors of Hydrogen Bonding . . . . .	99
8.1.6	Sampling Protocol for Experimental Conditions . . . . .	102
8.2	UV-Vis and ECD Spectroscopy . . . . .	105
8.2.1	Absorption Spectra from AIMD . . . . .	105
8.2.2	Robustness of Absorption Properties of PCB . . . . .	105
8.2.3	Colour Tuning is Mediated by Ring D . . . . .	108
8.2.4	Differentiation of $P_r$ Sub-States in AnPixJg2 . . . . .	109
8.3	Conclusion . . . . .	112

## BIBLIOGRAPHY

REFERENCES . . . . .	115
LIST OF ABBREVIATIONS . . . . .	145
LIST OF FIGURES . . . . .	149
LIST OF TABLES . . . . .	157

## APPENDIX

A COMPUTATIONAL DETAILS . . . . .	xi
A.1 Vibrational Circular Dichroism and Supramolecular Chirality . . . . .	xi
A.1.1 Preparations . . . . .	xi

## CONTENTS

A.1.2	<i>Ab Initio</i> Molecular Dynamics Simulations . . . . .	xi
A.1.3	NVPT Calculations . . . . .	xii
A.1.4	Calculation of IR and VCD Spectra . . . . .	xii
A.1.5	Static Calculations . . . . .	xiii
A.2	Colour Tuning of Phycocyanobilin in Proteins . . . . .	xiv
A.2.1	Preparations . . . . .	xiv
A.2.2	Force Field Molecular Dynamics Simulations . . . . .	xiv
A.2.3	QM/MM Molecular Dynamics Simulations . . . . .	xiv
A.2.4	Calculation of Nuclear Shieldings . . . . .	xv
A.2.5	UV-Vis and ECD Calculations . . . . .	xvi
A.3	Other . . . . .	xx

## ACKNOWLEDGEMENTS

This work was made possible by the continuous support of my supervisor Daniel Sebastiani whose mentoring and advice helped me throughout the process of obtaining the results presented here and who would always be available to listen to questions and moaning. I am especially grateful for the many workshops, conferences, and travels that I could attend throughout my doctorate. One particular highlight was organising the STC 2018 at home in Halle.

I want to thank Arne Scherrer and Rodolphe Vuilleumier for their guidance and patience in listening to my questions regarding quantum mechanics and NVPT. Sharing with me their enthusiasm and having me as a guest at the ENS in Paris has been a great experience.

Thanks go to my collaborators Chen Song, Susanne Altmayer, Wolfgang Gärtner, and Jörg Matysik for the countless consultations we had in Leipzig or Halle, as well as to Christian Wiebeler and Igor Schapiro for sharing their knowledge on the calculation of absorption spectra of PCB. The entire VOA community shall be acknowledged here for a friendly welcome and great support, especially Sergio Abbate who invited me to give a talk on VOA6 2018 in Brescia. Special thanks go to my colleagues in Halle for many great moments and in particular to Tobias Watermann, Gabriel Kabbe, and Martin Brehm as together we were always struggling to keep the group cluster working.

Contributions to this work from Laura K. Scarbath-Evers and Michael Türk, as well as provided computational resources by the ZiH in Dresden, the SCC in Karlsruhe, the ITZ in Halle, and the HLRN in Berlin are thankfully acknowledged. Furthermore, I want to thank the *Fonds der Chemischen Industrie* for granting me a Kekulé scholarship.

I am very grateful to Axel Straube, who carefully reviewed style and language of this monograph. Finally, I want to thank Saskia Jähnigen, who became the dearest part of my life, for careful proof-reading; not only did she support me with love and empathy, but also with her unlimited expertise and intelligence.



## PUBLICATIONS AND CONFERENCES

Parts of this work have been published in the following peer-reviewed journals:

- [1] Jähnigen, S.; Scherrer, A.; Vuilleumier, R.; Sebastiani, D. Chiral Crystal Packing Induces Enhancement of Vibrational Circular Dichroism *Angew. Chem., Int. Ed.* 2018, *57*, 13344–13348.
- [2] Scarbath-Evers, L. K.; Jähnigen, S.; Elgabarty, H.; Song, C.; Narikawa, R.; Matysik, J.; Sebastiani, D. Structural heterogeneity in a parent ground-state structure of AnPixJg2 revealed by theory and spectroscopy. *Phys. Chem. Chem. Phys.* 2017, *19*, 13882–13894.

This work has been presented at the following events and conferences:

- [a] Jähnigen, S.; Sebastiani, D.; Vuilleumier, R. Supramolecular Chirality and Vibrational Circular Dichroism based on Molecular Dynamics and Nuclear Velocity Perturbation Theory, 12<sup>th</sup> *European Conference on Computational and Theoretical Chemistry*, Perugia (Italy), September 2019, *conference talk*.
- [b] Jähnigen, S.; Scherrer, A.; Sebastiani, D.; Vuilleumier, R. VCD and Superchirality: Realistic Models based on Molecular Dynamics and Nuclear Velocity Perturbation Theory, 17<sup>th</sup> *International Conference on Chiroptical Spectroscopy*, Pisa (Italy), June 2019, *conference poster*.
- [c] Jähnigen, S.; Scherrer, A.; Sebastiani, D.; Vuilleumier, R. VCD and Superchirality: Realistic Models based on Molecular Dynamics and Nuclear Velocity Perturbation Theory, *Lorentz Workshop “VCD in Industry”*, Leiden (Netherlands), April 2019, *conference poster*.
- [d] Jähnigen, S.; Scherrer, A.; Vuilleumier, R.; Sebastiani, D. Chiral Crystal Packing Induces Enhancement of Vibrational Circular Dichroism, 54<sup>th</sup> *Symposium on Theoretical Chemistry*, Halle an der Saale (Germany), September 2018, *conference poster*.

## PUBLICATIONS AND CONFERENCES

- [e] Jähnigen, S.; Scherrer, A.; Vuilleumier, R.; Sebastiani, D. Calculation of Vibrational Circular Dichroism: Correlations in Space and Time, *6<sup>th</sup> International Conference on Vibrational Optical Activity*, Brescia, (Italy), September 2018, *invited conference talk*.
- [f] Jähnigen, S.; Scherrer, A.; Vuilleumier, R.; Sebastiani, D. Electronic Flux and Vibrational Circular Dichroism by means of Nuclear Velocity Perturbation Theory, *53<sup>rd</sup> Symposium on Theoretical Chemistry*, Basel (Switzerland), August 2017, *conference poster*.
- [g] Jähnigen, S.; Scarbath-Evers, L. K.; Ahlert, P.; Sebastiani, D. Structural and Spectroscopic Calculations of Phycocyanobilin (PCB) in the Protein Environment: *Ab Initio* Molecular Dynamics and Spectroscopy, *Photochemisches Symposium*, Leipzig University (Germany), May 2017, *invited conference talk*.
- [h] Scherrer, A.; Jähnigen, S.; Agostini, F.; Vuilleumier, R.; and Sebastiani, D. Vibrational Circular Dichroism in the Condensed Phase from First Principles;  
Scarbath-Evers, L. K.; Jähnigen, S.; Ahlert, P.; Elgabarty, H.; Altmayer, S.; Song, C.; Gärtner, W.; Matysik, J.; Watermann, T.; Sebastiani, D. Capturing the Dynamics of Phycocyanobilin in a Protein Photosensor, *52<sup>nd</sup> Symposium on Theoretical Chemistry*, Bochum (Germany), September 2016, *conference posters*.
- [i] Jähnigen, S.; Scherrer, A.; Vuilleumier, R.; Sebastiani, D. The Dynamical Coupling of Electrons and Nuclei within the Nuclear Velocity Perturbation Theory, *Faraday Discussions on Reaction Rate Theory*, Cambridge (United Kingdom), September 2016, *conference poster*.
- [j] Scherrer, A.; Jähnigen, S.; Agostini, F.; Vuilleumier, R.; Sebastiani, D. Vibrational Circular Dichroism in the Condensed Phase from First Principles, *5<sup>th</sup> International Conference on Vibrational Optical Activity*, Antwerp (Belgium), September 2016, *conference poster*.
- [k] Scherrer, A.; Jähnigen, S.; Vuilleumier, R.; Sebastiani, D. Vibrational Circular Dichroism in the Condensed Phase from First Principles, *51<sup>st</sup> Symposium on Theoretical Chemistry*, Potsdam (Germany), September 2015, *conference poster*.
- [l] Scherrer, A.; Jähnigen, S.; Vuilleumier, R.; Sebastiani, D. Vibrational Circular Dichroism in the Condensed Phase from First Principles, *114<sup>th</sup> Bunsentagung*, Bochum (Germany), May 2015, *conference poster*.

# INTRODUCTION AND THEORY





# I PREFACE

“Supramolecular chemistry is the chemistry of the intermolecular bond.”<sup>a</sup> With this sentence, Jean-Marie Lehn put in a nutshell his pioneering work in the field of non-covalent interactions that established a new and innovative branch of chemistry during the past 40 years.<sup>3</sup> Supramolecular systems are characterised by spatial organisation forming a *superstructure*, and intermolecular connections based, *inter alia*, on coordination, hydrogen bonding, electrostatic interactions, and van der Waals interactions. The important role of supramolecular concepts for systems of biological as well as non-biological origin is indisputable. Ever since, the complexity of chemical systems, possibilities of development, and scientific objectives has multiplied. Interdisciplinary research nowadays pursues an integrated view on processes in the microscopic as well as in the macroscopic sense. Evolution of theoretical modelling has all along tried to keep pace with the increased scope of practical demands and ambitions. Indeed, the recent years have seen the wealth of theoretical methods steadily growing, fuelled by constant extension of affordable computational power. Design *in silico* has become an integral part in modern chemical science, indispensable for a wide range of development and application.

Modern challenges in chemistry need realistic theoretical models that reflect experimental conditions. Especially in the context of supramolecular chemistry, *weak interactions* characterised by shallow potential energy surfaces and highly mobile counterparts have to be incorporated. In contrast to strong interactions being robust and well-determined, they rely on dynamics and entropy. With the 200<sup>th</sup> anniversary of the second law of thermodynamics drawing near, modern physical chemistry discusses free energies and free enthalpies as a matter of course. Nevertheless, to some extent, the idea of molecular dynamics influencing chemical structure and functionality did not yet take root in the general perception. Dynamics introduce anharmonicities that are germane to supramolecular chemistry. For instance, hydrogen bonds exhibit potential energy profiles that strongly deviate from

---

<sup>a</sup>Jean-Marie Lehn in his keynote lecture on occasion of the 20<sup>th</sup> anniversary of the *Friends of the Chemistry Olympiad* association, Leipzig, January 26, 2012. It is a recurring expression that can be found throughout his work.<sup>3</sup> Upon organising the reunion mentioned above, the author of this work had the chance to personally attend one of Lehn’s captivating lectures.

the harmonic approximation.<sup>4</sup> Anharmonicity determines distribution probabilities of molecular degrees of freedom and becomes especially important for coupled motion.

Another cornerstone of molecular and supramolecular chemistry, chirality describes the property of molecular systems to behave like mirror images, called *enantiomers*, that are non-superposable. It has fascinated scientists ever since Pasteur discovered optically active enantiomers of tartaric acid in Paris.<sup>5</sup> Most biological systems are chiral, while at the same time pharmaceuticals usually rely only on one enantiomer as the physiologically active species. This is because (biological) function builds upon the key-lock principle that describes the favourable combination of certain enantiomers to give an active complex.<sup>3</sup> In an achiral environment, molecular enantiomers cannot be differentiated as they exhibit equal physico-chemical properties. The allocation of a *chiral bias*, however, invokes distinguishable *diastereomeric conditions*, because it selects energetically favourable combinations of local enantiomers with external chirality. The bias potential can be a protein, a chiral solvent, but also a chiral electromagnetic wave, such as (circularly) polarised light.

The scope of this work is the development and application of realistic theoretical models for the study of supramolecular assemblies in the condensed phase. All systems presented touch the subject of supramolecular chirality, either explicitly by emulation of chiral spectroscopy or implicitly by studying the intriguing nature of protein structure and functionality.

In the first part, the relation of *vibrational circular dichroism* (VCD) and supramolecular chirality is dealt with. VCD refers to the absorption difference of left- and right-circularly polarised light.<sup>6,7</sup> In contrast to related Raman optical activity (ROA), VCD is measured in the infrared (IR) region of the electromagnetic spectrum and can thus be regarded as “chiral IR spectroscopy”.<sup>8</sup> Bound to molecular vibrations, VCD delivers a large amount of stereochemical information according to which parts of the molecule are involved in the underlying vibrational mode. VCD has become an invaluable tool in the determination of absolute configurations by assigning recorded spectra to structural information obtained from theoretical calculations.<sup>6</sup>

The second part enters the field of tetrapyrrole cofactors of photomorphogenic organisms. It aims for understanding of how a supramolecular protein environment is capable of specifically triggering the *colour tuning of phycocyanobilin* without chemically modifying the chromophore. This is addressed by means of nuclear magnetic resonance (NMR) and UV-Vis spectroscopy. NMR bases on the separation of nuclear spin states in the presence of a strong magnetic field (Zeeman-Effect) that is probed with electromagnetic radiation in the radio frequency region.<sup>9,10</sup> The spectroscopic result is highly dependent on the local electronic structure and geometry of atoms, which has made NMR spectroscopy the most important analytical tool in biochemistry but also in organic and inorganic

synthesis.<sup>11</sup>

Spectroscopic signatures are calculated by means of linear response within the margins of *density functional theory* (DFT).<sup>12</sup> The presented methodology is based on force field and *ab initio* molecular dynamics that aim for establishing realistic conditions regarding both structure and thermodynamics. This treatise begins with a survey on the basic methodology and fundamental theory used in this work. Coming from the time-dependent Schrödinger equation, the concept of exact factorisation of the electron-nuclear wavefunction is outlined, from which the Born-Oppenheimer and the *complete adiabatic* approximation can be inferred. Solution approaches for the electronic and nuclear subsystem are presented subsequently, reducing dimensionality of the electron structure problem and invoking the classical limit for the nuclei. Eventually, the basic formalism of calculating linear response observables for VCD, NMR, and UV-Vis spectroscopy are introduced.

Although the footing of this work is provided by many theoretical achievements from the past years and decades that are well documented in literature, only a small selection can be reproduced here and the interested reader is referred to references for further background. Consequently, introducing the theory of pseudopotentials, plane waves, and localised orbitals<sup>13</sup> has been omitted, while the position operator problem and gauge problem<sup>14-17</sup> are merely mentioned at appropriate position. The calculation of nuclear gradients, thermostats and barostats<sup>18</sup>, the constraints and SHAKE algorithms<sup>18,19</sup>, as well powerful implementations, such as the Gaussian and plane waves (GPW)<sup>20</sup> and the Gaussian-augmented plane waves (GAPW)<sup>21</sup> cannot be discussed here, either.



## 2 BASIC THEORY

The term “first principles” or *ab initio* implies theoretical models formulated from the postulates of quantum physics. However, approximations are made throughout without dropping pretensions of working in an accurate framework. This is necessary not only to bridge theory and experiment within the chemical world by reducing the problems’ dimensionality, but it also renders still a quantum description in the sense that a direct solution of the fundamental equations is carried out. Within this work, this is realised for the electronic equations of motion that are described by density functional theory, which requires approximations regarding the calculation of electronic exchange and correlation. The nuclear equations of motion, in turn, are considered in the classical limit, which is a reasonable approximation for the problems addressed within this work.

The following survey of theory sets a focus on response functions that yield the spectroscopic observables of the following investigations. It does not aim to reproduce a comprehensive description, which can be acquired respectively by means of given references, but highlights the important aspects and novelties of the theory underlying this work. Throughout, equations and definitions are written in atomic units with magnetic properties being formulated in the Gaussian cgs system. For partial derivatives the notation  $\partial_x = \frac{\partial}{\partial x}$  is used in most cases.

### 2.1 Equations of Motion

#### 2.1.1 Exact Factorisation of the Electron-Nuclear Wavefunction

The non-relativistic time-dependent Schrödinger equation (TDSE)<sup>22</sup> in the absence of external potentials is<sup>23</sup>

$$\hat{\mathcal{H}}\Psi(\mathbf{r}, \mathbf{R}, t) = i\partial_t\Psi(\mathbf{r}, \mathbf{R}, t), \quad (2.1)$$

## 2 BASIC THEORY

with  $\Psi(\mathbf{r}, \mathbf{R}, t)$  being the complex time-dependent electron-nuclear wavefunction that solves Equation 2.1 and the Hamiltonian,

$$\begin{aligned}\hat{\mathcal{H}}(\mathbf{r}, \mathbf{R}) &= \sum_i -\frac{1}{2}\nabla_i^2 + \sum_\nu -\frac{1}{2M_\nu}\nabla_\nu^2 + \frac{1}{2}\sum_{i\neq j}\frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} \\ &\quad + \frac{1}{2}\sum_{\nu\neq\mu}\frac{Z_\nu Z_\mu}{|\mathbf{R}_\nu - \mathbf{R}_\mu|} - \sum_{i\nu}\frac{Z_\nu}{|\mathbf{r}_i - \mathbf{R}_\nu|} \\ &= \hat{T}_e + \hat{T}_n + \hat{V}_{ee}(\mathbf{r}) + \hat{V}_{nn}(\mathbf{R}) + \hat{V}_{en}(\mathbf{r}, \mathbf{R}).\end{aligned}\tag{2.2}$$

$\hat{T}_e$  and  $\hat{T}_n$  denote the kinetic energy operators of electrons and nuclei, while  $\hat{V}_{ee}(\mathbf{r})$ ,  $\hat{V}_{nn}(\mathbf{R})$ , and  $\hat{V}_{en}(\mathbf{r}, \mathbf{R})$  represent the potential energy operators of the interaction of electrons, nuclei, and there between.  $M_\nu$  and  $Z_\nu$  are mass and atomic number of the  $\nu$ -th nucleus.  $i$  and  $\nu$  mark the sum over electrons and nuclei, respectively. In this section,  $\mathbf{r}$  and  $\mathbf{R}$  stand for the set of electronic and nuclear degrees of freedom,  $\{\mathbf{r}_1, \dots, \mathbf{r}_n\}$  and  $\{\mathbf{R}_1, \dots, \mathbf{R}_N\}$ .

It has been shown that  $\Psi(\mathbf{r}, \mathbf{R}, t)$  can be *exactly* factorised to give

$$\Psi(\mathbf{r}, \mathbf{R}, t) = \Phi_{\mathbf{R}}(\mathbf{r}, t)X(\mathbf{R}, t),\tag{2.3}$$

where  $X(\mathbf{R}, t)$  forms the nuclear and  $\Phi_{\mathbf{R}}(\mathbf{r}, t)$  the electronic complex wavefunction, the latter with parametric dependence on the nuclear positions.<sup>24,25</sup> Conservation of probability densities of finding nuclei or electrons in a specific configuration at time  $t$  is ensured by a normalisation condition.<sup>26,27</sup> As a result, the equations of motion for electronic and nuclear degrees of freedom can be formulated,<sup>16</sup>

$$[\hat{\mathcal{H}}_e(\mathbf{r}, \mathbf{R}) - \epsilon(\mathbf{R}, t)]\Phi_{\mathbf{R}}(\mathbf{r}, t) = i\partial_t\Phi_{\mathbf{R}}(\mathbf{r}, t),\tag{2.4a}$$

$$\hat{\mathcal{H}}_n(\mathbf{R}, t)X(\mathbf{R}, t) = i\partial_tX(\mathbf{R}, t),\tag{2.4b}$$

with the normalisation condition<sup>26,27</sup> inserted into the electronic equation by means of the Lagrange multiplier  $\epsilon(\mathbf{R}, t)$ , which is interpreted as the time-dependent potential energy surface (TD PES) of the nuclei (see below). The electronic and nuclear Hamiltonians are respectively defined via

$$\hat{\mathcal{H}}_e(\mathbf{r}, \mathbf{R}) = \hat{T}_e + \hat{V}_{ee}(\mathbf{r}) + \hat{V}_{en}(\mathbf{r}, \mathbf{R}) + \hat{V}_{nn}(\mathbf{R}) + \hat{\mathcal{U}}_{en}^{\text{coup}}[\Phi_{\mathbf{R}}, X]\tag{2.5}$$

and

$$\hat{\mathcal{H}}_n(\mathbf{R}, t) = \sum_\nu \frac{[-i\nabla_\nu + \mathbf{A}_\nu(\mathbf{R}, t)]^2}{2M_\nu} + \epsilon(\mathbf{R}, t).\tag{2.6}$$

The operator  $\hat{\mathcal{U}}_{en}^{\text{coup}} [\Phi_{\mathbf{R}}, X]$  exhibits a functional dependence on the electronic and nuclear wavefunction and represents the *non-adiabatic* electron-nuclear coupling,

$$\hat{\mathcal{U}}_{en}^{\text{coup}} [\Phi_{\mathbf{R}}, X] = \sum_{\nu} \frac{1}{M_{\nu}} \left[ \frac{[-i\nabla_{\nu} + \mathbf{A}_{\nu}(\mathbf{R}, t)]^2}{2} + \left( \frac{-i\nabla_{\nu} X}{X} + \mathbf{A}_{\nu}(\mathbf{R}, t) \right) (-i\nabla_{\nu} - \mathbf{A}_{\nu}(\mathbf{R}, t)) \right]. \quad (2.7)$$

$\mathbf{A}_{\nu}(\mathbf{R}, t)$  has the shape of a vector potential,<sup>28</sup>

$$\mathbf{A}_{\nu}(\mathbf{R}, t) = \langle \Phi_{\mathbf{R}}(t) | -i\nabla_{\nu} \Phi_{\mathbf{R}}(t) \rangle_{\mathbf{r}}. \quad (2.8)$$

Interestingly, within this formalism, the (non-adiabatic) effect of electrons on the nuclei is mediated by two elements, a time-dependent vector potential and the TD PES, both resulting from the electronic equation of motion. The latter, in turn, contains the full non-adiabatic coupling terms by virtue of the differential operator  $\hat{\mathcal{U}}_{en}^{\text{coup}} [\Phi_{\mathbf{R}}, X]$ . It addresses the parametric dependence on the nuclear positions of the electronic wavefunction. In spite or just because of this separation being exact, the factorisation can only be a starting point for a variety of approximations of different degrees;<sup>29–31</sup> the probably most important is presented in the next section.

### 2.1.2 Born-Oppenheimer Approximation

In view of the large mass of nuclei compared to that of electrons, the degrees of freedom of nuclei and electrons can be separated. The Born-Oppenheimer (BO) approximation partitions the electron-nuclear wavefunction such that the electronic structure problem is solved with the nuclei kept fixed.<sup>32</sup> The obtained time-independent BO potential energy surface (PES) enters the nuclear equation of motion – containing only nuclear degrees of freedom – that is solved afterwards. As a consequence, the electronic wavefunction turns out stationary and real, retaining the parametric dependence on the positions of the nuclei, and the factorisation of the electron-nuclear wavefunction from Equation 2.3 becomes

$$\Psi(\mathbf{r}, \mathbf{R}, t) = \Phi_{\mathbf{R}}^{\text{BO}}(\mathbf{r}) X(\mathbf{R}, t). \quad (2.9)$$

The transition  $\Phi_{\mathbf{R}}(\mathbf{r}, t) \rightarrow \Phi_{\mathbf{R}}^{\text{BO}}(\mathbf{r})$  implies an *adiabatic approximation* as the electronic system is considered adapting instantaneously to changed nuclear conditions. Upon solution of the electronic structure problem the nuclei appear immobile, hence their corresponding velocity vanishes

## 2 BASIC THEORY

together with all terms of non-adiabatic coupling,

$$\hat{\mathcal{U}}_{en}^{\text{coup}} [\Phi_{\mathbf{R}}, X] = 0. \quad (2.10)$$

The adiabatic wavefunction  $\Phi_{\mathbf{R}}^{\text{BO}}(\mathbf{r})$  solves the stationary Schrödinger equation of the electronic system,

$$[\hat{\mathcal{H}}^{\text{BO}}(\mathbf{r}, \mathbf{R}) - \epsilon^{\text{BO}}(\mathbf{R})] \Phi_{\mathbf{R}}^{\text{BO}}(\mathbf{r}) = 0, \quad (2.11)$$

with the BO Hamiltonian,

$$\hat{\mathcal{H}}^{\text{BO}}(\mathbf{r}, \mathbf{R}) = \hat{T}_e + \hat{V}_{ee}(\mathbf{r}) + \hat{V}_{en}(\mathbf{r}, \mathbf{R}) + \hat{V}_{nn}(\mathbf{R}), \quad (2.12)$$

and BO eigenvalues,  $\epsilon^{\text{BO}}(\mathbf{R})$ , that describe the time-independent BO potential energy surface (PES) whereon the nuclei move. Likewise, the nuclear equation reads

$$[\hat{T}_n + \epsilon^{\text{BO}}(\mathbf{R})] X(\mathbf{R}, t) = i\partial_t X(\mathbf{R}, t). \quad (2.13)$$

Hence, although the BO approximation seems to be an experiential concept, it can be recovered from the exact factorisation. It represents one of the most important concepts of quantum chemistry, particularly because in most applications, solving the stationary electronic structure problem only is sufficient. Nevertheless, means of extending or loosening the approximation are widespread (*e.g.*, *via* the Born-Huang approximation<sup>33</sup>). Beyond-BO effects are commonly treated within perturbation theory, which forms the basis of most spectroscopic applications (and hence of this work). A special case, however, is represented by the complete adiabatic (CA) approximation.

### 2.1.3 Complete Adiabatic Approximation

The importance of the BO ansatz has been proved by a manifold of successful applications where the electronic structure problem is solved in a static picture and consideration of nuclear motion is not necessary.<sup>23,34</sup> Even more, concepts like vibrational analysis or *ab initio* molecular dynamics (see Section 2.3) too take great advantage of the approximation *via* the BO PES.<sup>35</sup> And yet, formally allowing nuclear motion within the BO approximation entails a dichotomy between the formulations of quantum mechanics by Heisenberg and Schrödinger. The adiabatic theorem demands the spatial probability density (of electrons) adapt to a change in nuclear configuration. In the Heisenberg pic-



ture, this requirement is met by the time-independent electronic wavefunction that is found anew for any nuclear configuration using the BO approximation. However, the Schrödinger picture does not convey the time-evolution of electronic probability density as the electronic wavefunction remains time-independent; in this manner it does not adhere to the adiabatic theorem. Nafie showed that adding to the electronic Hamiltonian the dependence on nuclear velocities corresponds to a factorable, thus adiabatic, electron-nuclear wavefunction beyond the BO approximation.<sup>6,36</sup> Labelled as *complete adiabatic approximation*, it neglects any vibronic detail and does neither affect the nuclear kinetics nor nuclear or electronic (BO) energy levels, so that the BO PES is retained. After all, it clears out inconsistencies of the Schrödinger picture regarding the adiabatic theorem.

Coming from exact factorisation, the complete adiabatic approximation can be achieved in a formally exact way by adding corrections to the BO state, as has been shown by Scherrer, Agostini, and co-workers.<sup>16</sup> Considering a real non-degenerate BO ground state, the vector potential in Equation 2.8 vanishes and the electron-nuclear coupling (Equation 2.7) becomes

$$\hat{\mathcal{U}}_{en}^{\text{coup}} [\Phi_{\mathbf{R}}, X] = \sum_{\nu} \left[ \frac{-\nabla_{\nu}^2}{2M_{\nu}} + \frac{-i\nabla_{\nu}X}{M_{\nu}X} \cdot (-i\nabla_{\nu}) \right]. \quad (2.14)$$

The first term on the right-hand side of Equation 2.14 represents second-order non-adiabatic couplings and can be neglected in view of the magnitude of the second term. The  $X$ -dependent part can be approximated to the classical limit, yielding the nuclear velocity  $\dot{\mathbf{R}}_{\nu}(t)$ ,

$$\frac{-i\nabla_{\nu}X(\mathbf{R}, t)}{M_{\nu}X(\mathbf{R}, t)} = \frac{\mathbf{P}_{\nu}(t)}{M_{\nu}} = \dot{\mathbf{R}}_{\nu}(t), \quad (2.15)$$

where  $\mathbf{P}_{\nu}(t)$  denotes the nuclear momentum. In more general terms,<sup>a</sup> a (small) perturbation parameter is introduced that in the following characterises the BO correction term,

$$\lambda_{\nu}(\mathbf{R}, t) = \frac{1}{M_{\nu}} \frac{-i\nabla_{\nu}X(\mathbf{R}, t)}{X(\mathbf{R}, t)}. \quad (2.16)$$

The corrected electronic Hamiltonian from Equation 2.12 reads

$$\hat{\mathcal{H}}_e(\mathbf{r}, \mathbf{R}) = \hat{\mathcal{H}}^{\text{BO}}(\mathbf{r}, \mathbf{R}) + \sum_{\nu} \lambda_{\nu}(\mathbf{R}, t) \cdot (-i\nabla_{\nu}), \quad (2.17)$$

---

<sup>a</sup>The classical limit is not a prerequisite of the BO correction.

## 2 BASIC THEORY

and the electronic wavefunction is augmented with an imaginary correction term,

$$\Phi_{\mathbf{R}}^{\text{CA}}(\mathbf{r}, t) = \Phi_{\mathbf{R}}^{\text{BO}}(\mathbf{r}) + \sum_{\nu, \alpha} i \lambda_{\nu\alpha}(\mathbf{R}, t) \Phi'_{\mathbf{R}, \nu\alpha}(\mathbf{r}). \quad (2.18)$$

$\lambda_{\nu\alpha}$  is the  $\alpha$  Cartesian component of parameter  $\lambda_{\nu}$ , while  $\Phi'_{\mathbf{R}, \nu\alpha}(\mathbf{r})$  denotes the first-order perturbation to the Born-Oppenheimer wavefunction, depending on the  $\alpha$  Cartesian component of the  $\nu$ -th nucleus,

$$\Phi'_{\mathbf{R}, \nu\alpha}(\mathbf{r}) = \sum_{e \neq 0} \frac{d_{e0, \nu\alpha}(\mathbf{R})}{\omega_{e0}(\mathbf{R})} \Phi_{\mathbf{R}}^{\text{BO}(e)}(\mathbf{r}), \quad (2.19)$$

where  $d_{e0, \nu\alpha}(\mathbf{R})$  is the  $\alpha$  Cartesian component of the non-adiabatic coupling vector between the BO ground state (0) and the BO excited state ( $e$ ),  $\langle \Phi_{\mathbf{R}}^{\text{BO}(e)} | \partial_{\nu\alpha} \Phi_{\mathbf{R}}^{\text{BO}} \rangle$ , and  $\omega_{e0}(\mathbf{R})$  the energy difference between them.

The correction term depends implicitly on time *via*  $\lambda_{\nu}(\mathbf{R}, t)$ . In the classical limit this corresponds to the complete adiabatic correction proposed by Nafie.<sup>b</sup> Evidently, the electronic wavefunction now incorporates parametric dependence also on nuclear velocities. In fact, Equation 2.18 represents the lowest-order correction to the BO approximation in time-dependent regime.<sup>16</sup> It is interesting that, even though the correction is inferred from non-adiabatic coupling by means of a sum-over-states ansatz in Equation 2.19, the resulting electronic wavefunction adheres to the adiabatic theorem. It is important to note that due to corrections to the electronic wavefunction being imaginary, the BO ground state density remains unchanged and the nuclear equations are not affected either, but remain determined by the BO PES. Acknowledging this allows for inclusion of the CA approximation into standard (*i.e.*, BO) methodologies, such as *ab initio* molecular dynamics typically performed on a BO PES (see Section 2.3).

---

<sup>b</sup> The formulation of the CA wavefunction by Nafie can be recovered by Taylor expansion of  $\Phi_{\mathbf{R}}^{\text{CA}}(\mathbf{r}, t)$  about its equilibrium position to the first order.<sup>6,37,38</sup>

## 2.2 Electronic Structure

### 2.2.1 Density Functional Theory

The stationary Schrödinger equation of the electronic structure problem (Equation 2.11) contains a many-body wavefunction and is therefore not exactly solvable.<sup>39</sup> Within *density functional theory* (DFT) the central idea is to replace the one multidimensional problem with many one-dimensional problems deploying the one-particle density. Hence, along with the Hartree-Fock (HF) approach, DFT attributes to a mean field approach.<sup>34</sup>

The problem of the multidimensional many-body electronic wavefunction of  $n$  electrons  $\Phi_{\mathbf{R}}(\mathbf{r}) \equiv \Phi_{\mathbf{R}}(\mathbf{r}_1, \dots, \mathbf{r}_n)$  is reformulated in terms of the electronic density,<sup>13</sup>

$$n(\mathbf{r}) = \sum_i \langle \Phi_{\mathbf{R}}(\mathbf{r}_1, \dots, \mathbf{r}_n) | \delta(\mathbf{r} - \mathbf{r}_i) | \Phi_{\mathbf{R}}(\mathbf{r}_1, \dots, \mathbf{r}_n) \rangle. \quad (2.20)$$

The BO Hamiltonian of the electronic structure problem in Equation 2.12 is generalised by replacing the potential energy terms  $\hat{V}_{en}$  and  $\hat{V}_{ee}$  respectively with  $\hat{V}^{\text{ext}}$  and  $\hat{V}^{\text{int}}$ ,

$$\hat{\mathcal{H}}_e(\mathbf{r}) = \hat{T}_e + \hat{V}^{\text{int}}(\mathbf{r}) + \hat{V}^{\text{ext}}(\mathbf{r}) + E_{v\mu}. \quad (2.21)$$

$E_{v\mu}$  represents the interaction energy between nuclei. In the DFT context,  $\mathbf{r}$  does not anymore represent the set of variables of the electronic degrees of freedom,  $\{\mathbf{r}_1, \dots, \mathbf{r}_n\}$ , as in the previous section, but the general position in space.

The formalism of DFT is based on the Hohenberg-Kohn theorems, the first of which states that in a system of interacting particles the particle ground state density  $n_0(\mathbf{r})$  uniquely determines an external potential  $\hat{V}^{\text{ext}}(\mathbf{r})$ , thus the electronic Hamiltonian and all system properties. The second theorem introduces the *universal* density functional for the energy,  $E[n]$ , that is minimised by the ground state density  $n_0(\mathbf{r})$ .<sup>12</sup> Although this concept is in principle exact, the shape of this functional remains unknown.

One problem is faced in terms of a density formulation of the kinetic energy.<sup>13</sup> By introducing the Kohn-Sham (KS) auxiliary system, a model of fictitious independent particles<sup>c</sup> with wavefunctions  $\varphi_i^{\text{KS}}(\mathbf{r})$  is invoked, wherein the total energy  $E_{\text{KS}}$  is determined by the particle's (uncorrelated) kinetic

---

<sup>c</sup>Also referred to as “Kohn-Sham fermions”.

## 2 BASIC THEORY

energy  $T_s[\varphi_i^{\text{KS}}]$  and an effective potential  $V_s[n]$  which can be resolved *via* classical potential energy terms,<sup>40</sup>

$$\begin{aligned} E_{\text{KS}}[n] &= T_s[\varphi_i^{\text{KS}}] + V_s[n] \\ &= T_s[\varphi_i^{\text{KS}}] + E_{\text{H}}[n] + E^{\text{ext}}[n] + E_{v\mu} + E_{\text{xc}}[n], \end{aligned} \quad (2.22)$$

where  $E_{\text{H}}[n]$  is the interaction energy of the electronic charge density (Hartree energy) and  $E^{\text{ext}}[n]$  its interaction energy with the external potential.  $E_{\text{xc}}[n]$  denotes the *exchange-correlation functional* that is determined by the energy difference of the (exact) physical system and the KS system,

$$E_{\text{xc}}[n] = \langle \hat{T} \rangle_{\Phi_{\text{r}}} - T_s[\varphi_i^{\text{KS}}] + \langle \hat{V}^{\text{int}} \rangle_{\Phi_{\text{r}}} - E_{\text{H}}[n]. \quad (2.23)$$

Eventually, differentiation of Equation 2.22 with respect to  $\varphi_i^{\text{KS}}(\mathbf{r})$  yield the KS equations,

$$(\hat{\mathcal{H}}^{\text{KS}}(\mathbf{r}) - \epsilon_i^{\text{KS}}) \varphi_i^{\text{KS}}(\mathbf{r}) = 0, \quad (2.24)$$

with

$$\begin{aligned} \hat{\mathcal{H}}^{\text{KS}}(\mathbf{r}) &= -\frac{1}{2}\nabla^2 + \partial_n E_{\text{H}} + \partial_n E^{\text{ext}} + \partial_n E_{\text{xc}} \\ &= -\frac{1}{2}\nabla^2 + \int d^3r' \frac{n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} + \sum_{\nu} \frac{Z_{\nu}}{|\mathbf{r} - \mathbf{R}_{\nu}|} + \hat{V}_{\text{xc}}(\mathbf{r}), \end{aligned} \quad (2.25)$$

that have to be solved self-consistently for each particle.  $\hat{V}_{\text{xc}}(\mathbf{r})$  represents the exchange-correlation potential yet to be determined.<sup>33</sup>

The endeavour for the universal exchange-correlation functional appears futile, but developing and improving approximated formulations, tailor-made for specific applications, is still subject of broad research.<sup>41,42</sup> The main types of density functionals encompass *local density approximation* (LDA), *generalised gradient approximation* (GGA), and hybrid functionals that partially include the exchange term from Hartree-Fock calculations (*exact exchange*).<sup>33</sup>

In the KS formalism the particles are independent, yet they form a correlated electron density, which is encoded in  $E_{\text{xc}}[n]$ . This was the initiation of many successful applications of DFT in quantum chemistry. A notorious drawback of KS DFT is the limitation regarding the interpretation of issued wavefunctions and eigenvalues as they do not represent the physical system, which is only ensured for

the total energy and particle density.<sup>d</sup> Nevertheless, post-processing, such as in perturbation theory, is well possible.<sup>33</sup>

### 2.2.2 Time-Dependent Density Functional Theory

The general form of the external potential  $V^{\text{ext}}(\mathbf{r})$  in Equation 2.21 allows for extensions that may not have been considered in the DFT formalism, but are in principle allowed.<sup>34</sup> One important example are time-dependent external fields, which lead to formulation of time-dependent density functional theory (TD-DFT). It is based on the *Runge-Gross theorem* – the time-dependent extension of the first Hohenberg-Kohn theorem.<sup>34,43,44</sup> It states that the time-dependent electron density  $n(\mathbf{r}, t)$  is uniquely defined by a time-dependent external potential  $V^{\text{ext}}(\mathbf{r}, t)$ . The updated electronic structure problem can be reformulated by means of time-dependent perturbation theory.<sup>34</sup> Therein, the external potential is written in the form

$$V^{\text{ext}}(\mathbf{r}, t) = V^{\text{ext}}(\mathbf{r}) + \delta V^{\text{ext}}(\mathbf{r}, t), \quad (2.26)$$

where  $V^{\text{ext}}(\mathbf{r})$  is the unperturbed potential (see previous section) and  $\delta V^{\text{ext}}(\mathbf{r}, t)$  the time-dependent external perturbation (*e.g.*, an electric field).

For small perturbations, the linear response of the ground state takes the general form

$$\begin{aligned} \delta n(\mathbf{r}, t) &= \int dt' \int d^3r' \chi(\mathbf{r}, t, \mathbf{r}', t') \delta V^{\text{ext}}(\mathbf{r}', t'), \\ \delta n(\mathbf{r}, \omega) &= \int d^3r' \chi(\mathbf{r}, \mathbf{r}'; \omega) \delta V^{\text{ext}}(\mathbf{r}', \omega), \end{aligned} \quad (2.27)$$

in time and frequency domain, respectively.  $\chi(\mathbf{r}, t, \mathbf{r}'; \omega)$  is the non-local *linear response function* or *electronic susceptibility*. Equations 2.27 require the explicit treatment of electronic interaction, which makes it unsuitable for actual applications. However, switching to the Kohn-Sham auxiliary system takes advantage of the independent particle description, that nevertheless yield the same response of

---

<sup>d</sup>An exception is found for frontier (HOMO, LUMO) orbitals that often exhibit similar spatial properties as the particle density and can thus be considered as physical, too.<sup>33</sup>

the electron density. The linear response function is calculated *via* a Dyson equation<sup>34</sup>

$$\begin{aligned} \chi(\mathbf{r}, \mathbf{r}', \omega) = & \chi^{\text{KS}}(\mathbf{r}, \mathbf{r}', \omega) + \int d^3r_1 \int d^3r_2 \chi^{\text{KS}}(\mathbf{r}, \mathbf{r}_1, \omega) \\ & \times \left[ \frac{1}{|\mathbf{r}_1 - \mathbf{r}_2|} + f_{\text{xc}}(\mathbf{r}_1, \mathbf{r}_2, \omega) \right] \chi(\mathbf{r}_2, \mathbf{r}', \omega), \end{aligned} \quad (2.28)$$

with exchange-correlation kernel  $f_{\text{xc}}(\mathbf{r}, t, \mathbf{r}', t') = \frac{\delta V_{\text{xc}}(\mathbf{r}, t)}{\delta n(\mathbf{r}', t')}$  and the KS response function

$$\chi^{\text{KS}}(\mathbf{r}, \mathbf{r}', \omega) = 2 \sum_{kl} \frac{\varphi_k^*(\mathbf{r}) \varphi_l(\mathbf{r}) \varphi_l^*(\mathbf{r}') \varphi_k(\mathbf{r}')}{\omega - (\epsilon_k - \epsilon_l)}, \quad (2.29)$$

which is merely a product of the KS ground state functions with eigenvalues  $\epsilon$ , whereas the indices  $k$  and  $l$  mark occupied and virtual orbitals alike.

Equation 2.28 represents the KS linear response corrected by response terms that stem from the Hartree and exchange-correlation potentials. The main task is to find the poles of  $\chi(\mathbf{r}, \mathbf{r}', \omega)$ , which is usually done by means of Casida's equations.<sup>45</sup> All the above equations are formally exact, but in a practical TD-DFT calculation, approximations have to be made as the exact nature of the exchange-correlation functional and kernel is unknown. TD-DFT has become a standard tool in the area of calculating electronic excitations for larger systems. It convinces with low computational cost, but at the same time shares with standard DFT the reputation of being artful and unsystematic.<sup>44</sup> Indeed, the result of TD-DFT strongly relies on the quality of the underlying DFT calculation, locality of exchange-correlation, adiabaticity, and soundness of the (non-interacting) KS wavefunction model.<sup>44</sup>

### 2.2.3 Density Functional Perturbation Theory

An important drawback of TD-DFT approaches is the need of occupied as well as unoccupied KS orbitals to compute the linear response function that contains the poorly convergent double sum in Equation 2.29. Another way to evaluate the response function is the use of a variational density functional perturbation theory (DFPT) approach based on the Sternheimer equation.<sup>46-48</sup> In generalised variation, DFPT can even be applied to perturbations that cannot be expressed in Hamiltonian form, for instance when working with a magnetic field in a periodic system.<sup>14,15</sup>

Therein, perturbation of the KS ground state adds to the KS energy in the form

$$E^{\text{tot}}[n] = E^{\text{KS}}[n] + \lambda E^{\text{pert}}[n], \quad (2.30)$$

where  $\lambda$  is the perturbation parameter. The response of the KS states  $\varphi$  and the electron density  $n(\mathbf{r})$  can be expanded around its unperturbed value,

$$\varphi_i(\mathbf{r}) = \varphi_i^{(0)}(\mathbf{r}) + \lambda \varphi_i^{(1)}(\mathbf{r}) + \lambda^2 \varphi_i^{(2)}(\mathbf{r}) + \dots, \quad (2.31a)$$

$$n(\mathbf{r}) = n^{(0)}(\mathbf{r}) + \lambda n^{(1)}(\mathbf{r}) + \lambda^2 n^{(2)}(\mathbf{r}) + \dots. \quad (2.31b)$$

Similarly, total energy  $E^{\text{tot}}[n]$  can be expanded,

$$E^{\text{tot}}[n] = E^{\text{tot}}[\{\varphi_i^{(0)}(\mathbf{r}) + \lambda \varphi_i^{(1)}(\mathbf{r}) + \dots\}] = E^{(0)} + \lambda E^{(1)} + \lambda^2 E^{(2)} + \dots. \quad (2.32)$$

To perturbatively find the minimum of  $E^{\text{tot}}[n]$ , Equation 2.32 is expanded to second power in  $\lambda$ , supplemented by an additional orthogonality constraint on the response orbitals,

$$\langle \varphi_k^{(0)} | \varphi_l^{(1)} \rangle = 0 \quad \forall k, l. \quad (2.33)$$

If  $E^{\text{tot}}[n]$  represents the energy minimum, the first non-vanishing term in the expansion of Equation 2.32 amid variation is  $E^{(2)}$ ,

$$\begin{aligned} E^{(2)} = & \sum_{kl} \left\langle \varphi_k^{(1)} \left| \hat{\mathcal{H}}^{\text{KS}} \delta_{kl} - \langle \varphi_l^{(0)} | \hat{\mathcal{H}}^{\text{KS}} | \varphi_k^{(0)} \rangle \right| \varphi_l^{(1)} \right\rangle \\ & + \frac{1}{2} \int d^3r \int d^3r' K(\mathbf{r}, \mathbf{r}') n^{(1)}(\mathbf{r}) n^{(1)}(\mathbf{r}') \\ & + \sum_k \left\langle \varphi_k^{(1)} \left| \frac{\delta E^{\text{pert}}[n]}{\delta \langle \varphi_k^{(0)} |} + \frac{\delta E^{\text{pert}}[n]}{\delta | \varphi_k^{(0)} \rangle} \right| \varphi_k^{(1)} \right\rangle, \end{aligned} \quad (2.34)$$

with Hartree-exchange-correlation kernel  $K(\mathbf{r}, \mathbf{r}') = \frac{\delta^2(E_{\text{H}} + E_{\text{xc}})[n^{(0)}]}{\delta n(\mathbf{r}) \delta n(\mathbf{r}')}.$ <sup>14,15</sup> Variation of  $E^{(2)}$  in  $\varphi_k^{(1)}$  under the orthogonality constraint from Equation 2.33 delivers a set of Sternheimer equations,

$$- \hat{\mathcal{P}}_e \sum_l \left( \hat{\mathcal{H}}^{\text{KS}} \delta_{kl} - \langle \varphi_l^{(0)} | \hat{\mathcal{H}}^{\text{KS}} | \varphi_k^{(0)} \rangle \right) \hat{\mathcal{P}}_e | \varphi_l^{(1)} \rangle = \hat{\mathcal{P}}_e \hat{\mathcal{H}}_e^{(1)} | \varphi_k^{(0)} \rangle, \quad (2.35)$$

with the perturbation functional written as perturbation Hamiltonian  $\hat{\mathcal{H}}_e^{(1)}$ .  $\hat{\mathcal{P}}_e = \sum_l 1 - |\varphi_l\rangle\langle\varphi_l|$  is a projection operator on the empty orbitals. In contrast to the previous section, indices  $k$  and  $l$  denote only occupied KS states. Hence, the perturbation can be calculated based on knowledge of

occupied states only, which makes DFPT very efficient.<sup>e</sup>

Equation 2.34 simplifies considerably for purely imaginary perturbations (*e.g.*, magnetic fields, nuclear velocities), because of a vanishing first order density  $n^{(1)}(\mathbf{r})$ ; the new Sternheimer equation reads

$$-\hat{\mathcal{P}}_e \sum_l \left( \hat{\mathcal{H}}^{\text{KS}} \delta_{kl} - \langle \varphi_l^{(0)} | \hat{\mathcal{H}}^{\text{KS}} | \varphi_k^{(0)} \rangle \right) \hat{\mathcal{P}}_e |\varphi_l^{(1)}\rangle = \hat{\mathcal{P}}_e \frac{\delta E^{\text{pert}}[n]}{\delta \langle \varphi_k^{(0)} |}. \quad (2.36)$$

It only contains the perturbation operator independent of the electronic density. Therefore, in contrast to Equation 2.35, Equation 2.36 need not be solved self-consistently.<sup>14</sup>

The advantage of this formulation is that the explicit dependence on a perturbation Hamiltonian can be avoided, which enables the calculation of  $E^{\text{pert}}[n]$  based on an arbitrary set of wavefunctions or orbital-dependent perturbation Hamiltonians.<sup>14</sup> Although the Sternheimer equation can be solved analytically, it is often done numerically in a conjugate gradient scheme.<sup>49,50</sup>

## 2.3 Molecular Dynamics

### 2.3.1 Nuclear Equations of Motion

After separation of electronic and nuclear degrees of freedom and application of the adiabatic approximation (*cf.* Section 2.1), the nuclear equations of motion evolve (Equation 2.13). In general terms, they describe the motion of the nuclear wavefunction in a potential  $U(\mathbf{R}) = \epsilon^{\text{BO}}(\mathbf{R})$ .<sup>35</sup> In the  $\hbar \rightarrow 0$  limit, they correspond to the classical laws of Newtonian mechanics,<sup>18</sup>

$$\partial_t \mathbf{P}_\nu(t) = -\nabla_\nu U(\mathbf{R}) = \mathbf{F}_\nu(t), \quad (2.37a)$$

$$\mathbf{P}_\nu(t) = M_\nu \dot{\mathbf{R}}_\nu(t) = M_\nu \partial_t \mathbf{R}_\nu(t), \quad (2.37b)$$

where  $\mathbf{R}_\nu(t)$  is the nuclear position and  $\mathbf{F}_\nu(t)$  the classical force acting on the  $\nu$ -th nucleus.<sup>18</sup>

The idea of molecular dynamics (MD) simulations is the solution of the nuclear many-body problem by integration of the (classical) equations of motion.<sup>18</sup> Bound to time evolution, the *trajectory* of the

<sup>e</sup>There are cases, where explicit knowledge of included virtual orbitals is desirable, such as in the computation of electronic absorption.



system samples the *phase space* of the nuclear system, spanned by variables  $\mathbf{R}_v(t)$  and  $\dot{\mathbf{R}}_v(t)$ . In this way, according to the ergodic hypothesis,<sup>f</sup> it generates a statistical ensemble – after simulations over a long time.<sup>18</sup> The conditions imposed on the (classical) nuclear Hamiltonian determine the type of statistical ensemble,<sup>g</sup> wherein the MD simulations run.<sup>18</sup>

The equations of motion represent differential equations of second order that are numerically integrated, for instance by means of a velocity Verlet algorithm,<sup>51</sup>

$$\mathbf{R}_v(t + \Delta t) = \mathbf{R}_v(t) + \dot{\mathbf{R}}_v(t)\Delta t + \frac{1}{2} \frac{\mathbf{F}_v(t)}{M_v} \Delta t^2, \quad (2.38a)$$

$$\dot{\mathbf{R}}_v(t + \Delta t) = \dot{\mathbf{R}}_v(t) + \frac{\mathbf{F}_v(t) + \mathbf{F}_v(t + \Delta t)}{2M_v} \Delta t. \quad (2.38b)$$

An important prerequisite of any solver for the classical equations of motion in Equation 2.37 is long-time stability that preserves time reversibility and ensures conservation of energy.<sup>18</sup> The inclusion of dynamics into the system widens the scope of any conceivable application that was realised in the static picture before. In statistical ensembles the thermodynamic system does no longer reside in its potential energy minima, but samples the PES according to the programmed temperature, exploring less stable areas that may be of high importance for the chemical problem. Therefore, the role of entropy, especially important in supramolecular chemistry, is incorporated, wherefore probability distributions are based on the *free energy surface* of the molecular system. Depending on how  $U(\mathbf{R})$  is being calculated, system size, as well as accuracy and timescales of the trajectory can be adjusted.

### 2.3.2 *Ab Initio* Molecular Dynamics

*Ab initio* molecular dynamics (AIMD) rely on a PES resulting from quantum-mechanical evaluation of the electronic degrees of freedom, which is often carried out within the adiabatic approximation.<sup>52</sup> The most frequent variant of AIMD, which has also been used in this work, are Born-Oppenheimer molecular dynamics (BOMD) (*cf.* Section 2.1),<sup>35</sup> where the PES is updated *at every timestep* calculating the corresponding BO ground state,

$$U(\mathbf{R}) = \min_{\Phi_{\mathbf{R}}} \left\{ \langle \Phi_{\mathbf{R}}^{\text{BO}}(\mathbf{r}) | \hat{\mathcal{H}}^{\text{BO}}(\mathbf{r}, \mathbf{R}) | \Phi_{\mathbf{R}}^{\text{BO}}(\mathbf{r}) \rangle \right\} = \epsilon^{\text{BO}}(\mathbf{R}), \quad (2.39)$$

<sup>f</sup>In simple terms, the ergodic hypothesis states that averages taken over time equal averages taken over the statistical ensemble.<sup>18</sup>

<sup>g</sup>The most frequent variants are microcanonical (NVE), canonical (NVT), and isobaric-isothermal (NPT) ensembles.<sup>18</sup>

by solving the stationary Schrödinger equation of the electronic structure self-consistently,

$$\hat{\mathcal{H}}^{\text{BO}}(\mathbf{r}, \mathbf{R})|\Phi_{\mathbf{R}}^{\text{BO}}(\mathbf{r})\rangle = \epsilon^{\text{BO}}(\mathbf{R})|\Phi_{\mathbf{R}}^{\text{BO}}(\mathbf{r})\rangle. \quad (2.40)$$

### 2.3.3 Force Field Molecular Dynamics

Force field molecular dynamics (FFMD) represent an ultrafast alternative to AIMD, because they do not require the expensive determination of a quantum-mechanical electronic ground state. In return, they deploy a fully empirical effective potential, for which molecular degrees have been parametrised. The way how this is put into effect strikes the paths of a plethora of *force fields* tailor-made for specific chemical systems and atom types. The parametrisation allows for the study of very large systems and extends the MD timescales by one or two orders of magnitude (ns to  $\mu\text{s}$ ). In this work, the CHARMM22<sup>h</sup> force field is used for MD simulations of solvated proteins comprising more than 50,000 atoms.<sup>54-56</sup> It has the general form

$$\begin{aligned} U(\{\mathbf{R}_\nu\}) = & \sum_{\text{bonds}} K_b(b - b_0)^2 + \sum_{\text{angles}} K_\theta(\theta - \theta_0)^2 + \sum_{\text{dihedrals}} K_\phi(1 + \cos(n\phi - \delta)) \\ & + \sum_{\text{improper}} K_\varphi(\varphi - \varphi_0)^2 + \sum_{\text{UB}} K_{\text{UB}}(S - S_0)^2 \\ & + \sum_{i,j>i} \frac{1}{4\pi\epsilon r_{ij}} \frac{q_i q_j}{|r_{ij}|} + \epsilon_{ij} \left[ \left( \frac{\sigma_{ij}}{r_{ij}} \right)^{12} - 2 \left( \frac{\sigma_{ij}}{r_{ij}} \right)^6 \right]. \end{aligned} \quad (2.41)$$

$K_x$  denotes force constants for bonds ( $x = b$ ), Urey-Bradley (UB) 1,3-distances ( $x = S$ ), bond angles ( $x = \theta$ ), dihedral angles ( $x = \phi$ ), and improper dihedral angles ( $x = \varphi$ ) with  $x_0$  being the respective value of the variable at equilibrium. Intermolecular terms are described *via* Lennard-Jones and Coulomb terms, with Lennard-Jones parameters  $\sigma_{ij}$  and  $\epsilon_{ij}$ , interatomic distances  $r_{ij}$ , and partial charges  $q_i$  and  $q_j$  of atoms  $i$  and  $j$ , respectively.<sup>56</sup>

<sup>h</sup>Chemistry at HARvard Macromolecular Mechanics<sup>53</sup>

## 2.3.4 QM/MM

In some cases, a compromise between large system extents and quantum-mechanical accuracy has to be achieved. For instance, the calculation of spectroscopic properties requires a full quantum description of the electronic structure, but may not do so for the entire macromolecular system. This work contains the study of a cofactor in its protein embedding, which raises exactly such issue. Hybrid QM/MM methods partition the macromolecular system into small subsystems described by means of *ab initio* theories (s1, QM part) and large subsystems bound to force fields (s2, MM part).<sup>57-59</sup> Connection between the subsystems is established either by an *additive* or *subtractive* scheme with the total potential  $U^{\text{QM/MM}}$  and that of the subsystems,  $U^{\text{QM}}$  and  $U^{\text{MM}}$ ,<sup>60</sup>

$$U_{\text{add}}^{\text{QM/MM}} = U^{\text{MM}}(s1) + U^{\text{QM}}(s2) + U^{\text{QM/MM}}(s1/s2), \quad (2.42a)$$

$$U_{\text{sub}}^{\text{QM/MM}} = U^{\text{MM}}(s1 + s2) + U^{\text{QM}}(s1) - U^{\text{MM}}(s2), \quad (2.42b)$$

whereas only the former contains explicit QM/MM interaction terms, while the latter, though being straightforward, renders them only approximate. Covalent bonds between QM and MM part must be treated separately, either by auxiliary linking atoms or other means of saturation of dangling bonds to avoid defects in the QM part description.<sup>61</sup>

All QM/MM calculations in this work are based on the additive scheme; the QM/MM interaction energy can be written as

$$U^{\text{QM/MM}}(\mathbf{r}) = \sum_{\mu \in \text{MM}} q_{\mu} \int d^3r \frac{\rho(\mathbf{r}; \{\mathbf{R}_\nu\})}{|\mathbf{r} - \mathbf{R}_{\mu}|} + \sum_{\mu \in \text{MM}} \sum_{\nu \in \text{QM}} U_{\text{vdW}}(\mathbf{R}_{\mu}, \mathbf{R}_{\nu}), \quad (2.43)$$

where  $U_{\text{vdW}}(\mathbf{R}_{\mu}, \mathbf{R}_{\nu})$  denotes the van der Waals terms according to the force field (*cf.* Equation 2.41) and  $\rho(\mathbf{r}; \{\mathbf{R}_\nu\})$  is the total density of electrons *and* nuclei.<sup>62</sup> Due to singularities in the Coulomb terms stemming from classical point charges, the evaluation of this term can be very costly, especially within the plane wave representation of electrostatic interactions.<sup>20</sup>

Laino and co-workers introduced the Gaussian expansion of the electrostatic potential (GEEP) that mitigates the computational overhead of calculating  $U^{\text{QM/MM}}(\mathbf{r})$  by up to two orders of magnitude.<sup>62</sup> Therein, MM charges are represented as Gaussian distribution (“smeared charges”) to ap-

proximately account for the atomic radius

$$\rho(|\mathbf{r} - \mathbf{R}_\mu|) = \left( \frac{1}{\pi r_{c,\mu}^2} \right)^{\frac{3}{2}} \exp \left( -\frac{|\mathbf{r} - \mathbf{R}_\mu|^2}{r_{c,\mu}^2} \right), \quad (2.44)$$

where  $r_{c,\mu}$  is the width of the Gaussian charge of the classical atom  $\mu$  that corresponds to  $q_\mu$ . The resulting term for the electrostatic potential is then further expanded into Gaussian functions, which yields the mentioned speed-up without loss of accuracy.<sup>63</sup>

## 2.4 Linear Response for Spectroscopy

### 2.4.1 Magnetic Field Perturbation

In this work, the calculation of NMR chemical shifts is based on magnetic field perturbation theory (MFPT), which is handled in the DFPT scheme (Section 2.2).<sup>14,21</sup> The application of a strong external magnetic field triggers a heterogeneous electronic response that strongly depends on the local topology of the molecule. The electronic response itself induces a magnetic field, which superimposes the external field leading to characteristic *chemical shifts*. The magnetic field,

$$\mathbf{B}^{\text{ext}} = \nabla \times \mathbf{A}(\mathbf{r}), \quad (2.45)$$

can be written in terms of a vector potential  $\mathbf{A}(\mathbf{r})$ .<sup>28</sup> The cross product in Equation 2.45 allows  $\mathbf{A}(\mathbf{r})$  to have a *gauge freedom*, which will always yield the same magnetic field. The vector potential is used for the calculation of the induced current as it conveys the presence of the magnetic field to the electronic Hamiltonian (in atomic units),

$$\begin{aligned} \hat{\mathcal{H}}_e(\mathbf{r}) &= \frac{1}{2} \hat{\boldsymbol{\pi}}(\mathbf{r})^2 + \hat{\mathcal{V}}(\mathbf{r}) \\ &= \frac{1}{2} (-i\nabla - \hat{\mathbf{A}}(\mathbf{r}))^2 + \hat{\mathcal{V}}(\mathbf{r}). \end{aligned} \quad (2.46)$$

$\hat{\mathcal{V}}(\mathbf{r})$  indicates the potential energy terms (*cf.* Equation 2.21) and  $\hat{\boldsymbol{\pi}}(\mathbf{r}) = -i\nabla - \hat{\mathbf{A}}(\mathbf{r})$  represents the generalised momentum operator.

It straightforwardly delivers the first-order perturbation Hamiltonian,

$$\hat{\mathcal{H}}_e^{(1)}(\mathbf{r}) = -i \frac{1}{2} \hat{\mathbf{A}}(\mathbf{r}) \cdot \nabla, \quad (2.47)$$

which gives rise to an imaginary correction to the KS ground state wavefunction,

$$\varphi_l(\mathbf{r}) = \varphi_l^{(0)}(\mathbf{r}) + i \varphi_l^{(1)}(\mathbf{r}). \quad (2.48)$$

Due to the gauge freedom of  $\mathbf{A}(\mathbf{r})$ , inversion of Equation 2.45 does not yield a unique definition of the vector potential for a given magnetic field, but choosing a gauge is vital for evaluating the perturbation Hamiltonian (*gauge origin problem*).<sup>14,15</sup> In this work, the magnetic field is evaluated in *virtual cells*<sup>i</sup> that make use of a *continuous set of gauge transformations* (CSGT)<sup>64</sup> and maximally localised Wannier orbitals (MLWO).<sup>65,66</sup> Although with this, the perturbation Hamiltonian becomes *orbital-dependent*, DFPT is capable of handling virtual cells with its implicit perturbation scheme represented by the Sternheimer equation.<sup>15</sup>

As the perturbation Hamiltonian is purely imaginary, it follows the Sternheimer equation of the shape

$$\hat{\mathcal{P}}_e \sum_l \left( \hat{\mathcal{H}}^{ks} \delta_{kl} - \langle \varphi_l^{(0)} | \hat{\mathcal{H}}^{ks} | \varphi_k^{(0)} \rangle \right) \hat{\mathcal{P}}_e | \varphi_l^{(1)} \rangle = \hat{\mathcal{P}}_e \frac{1}{2} \hat{\mathbf{A}}(\mathbf{r}) \cdot \nabla | \varphi_k^{(0)} \rangle, \quad (2.49)$$

with the imaginary first-order orbitals  $\varphi_l^{(1)}$ . Equation 2.49 can be solved without self-consistence and without explicit knowledge of the unoccupied states.

The observable chemical shift is defined via

$$\sigma(\mathbf{R}_\nu) = \frac{\partial \mathbf{B}^{\text{ind}}(\mathbf{R}_\nu)}{\partial \mathbf{B}^{\text{ext}}}, \quad (2.50)$$

where  $\mathbf{B}^{\text{ind}}(\mathbf{R}_\nu)$  is the magnetic field induced by the electronic response. It is calculated from the electronic current density  $\mathbf{j}(\mathbf{r})$  that results from the perturbed states after solving Equation 2.49,

$$\mathbf{B}^{\text{ind}}(r) = \frac{1}{c^2} \int d^3 r' \frac{\mathbf{r}' - \mathbf{r}}{|\mathbf{r}' - \mathbf{r}|^3} \times \mathbf{j}(\mathbf{r}'), \quad (2.51)$$

with speed of light  $c$  (Biot-Savart law).<sup>28</sup>

<sup>i</sup> Sebastiani and co-workers bypassed the *position operator problem*, wherein the position operator is ill-defined under periodic boundary conditions, newly defining a periodic position operator of saw-tooth shape.<sup>14</sup>

## 2.4.2 Nuclear Velocity Perturbation

A major part of this work relates to the calculation of VCD patterns based on AIMD simulations and molecular magnetic moments.<sup>†</sup> To this end, nuclear velocity perturbation theory (NVPT) presents a highly effective framework exploiting the phase space provided by the trajectory.<sup>1,17,67</sup> Coming from the complete adiabatic approximation (see Section 2.1), a DFPT scheme can be set up, wherein the perturbation Hamiltonian to the BO ground state in the classical limit for the nuclei (*cf.* Equation 2.15 and 2.17) reads

$$\hat{\mathcal{H}}_e^{(1)}(t) = -i \sum_{\nu} \dot{\mathbf{R}}_{\nu}(t) \cdot \nabla_{\nu}, \quad (2.52)$$

with  $\dot{\mathbf{R}}_{\nu}(t)$  being the classical velocity of the  $\nu$ -th nucleus as it occurs in AIMD. The complex, time-dependent adiabatic electronic wave function can be transferred into KS states,

$$\varphi_l(\mathbf{r}, t) = \varphi_l^{(0)}(\mathbf{r}) + i \varphi_l^{(1)}(\mathbf{r}, t), \quad (2.53)$$

where  $\varphi^{(1)}(\mathbf{r}, t)$  is the total perturbative correction which roots itself in nuclear velocities. It collates corrections from all  $\alpha$  Cartesian components of the nuclei as weighted sum,

$$\varphi_l^{(1)}(\mathbf{r}, t) = \sum_{\nu, \alpha} \dot{R}_{\nu\alpha}(\mathbf{R}, t) \varphi_{l, \nu\alpha}^{(1)}(\mathbf{r}). \quad (2.54)$$

The correction is of purely imaginary kind, which like in the case of magnetic perturbation substantially simplifies the response term of the electronic system and leads to component-wise Sternheimer equations,

$$-\hat{\mathcal{P}}_e \sum_l \left( \hat{\mathcal{H}}^{\text{KS}} \delta_{kl} - \langle \varphi_l^{(0)} | \hat{\mathcal{H}}^{\text{KS}} | \varphi_k^{(0)} \rangle \right) \hat{\mathcal{P}}_e | \varphi_{l, \nu\alpha}^{(1)} \rangle = \hat{\mathcal{P}}_e \partial_{\nu\alpha} | \varphi_k^{(0)} \rangle \quad \forall \nu, \alpha. \quad (2.55)$$

Their solution requires two response calculations per set of  $\{\nu, \alpha\}$ :

One for the nuclear displacement perturbation,

$$-\hat{\mathcal{P}}_e \sum_l \left( \hat{\mathcal{H}}^{\text{KS}} \delta_{kl} - \langle \varphi_l^{(0)} | \hat{\mathcal{H}}^{\text{KS}} | \varphi_k^{(0)} \rangle \right) \hat{\mathcal{P}}_e | \varphi_{l, \nu\alpha}^{\text{NDP}} \rangle = \hat{\mathcal{P}}_e \frac{\partial \hat{\mathcal{H}}^{\text{KS}}(\mathbf{r})}{\partial R_{\nu\alpha}} | \varphi_k^{(0)} \rangle \quad \forall \nu, \alpha, \quad (2.56)$$

---

<sup>†</sup>VCD can also be calculated by means of magnetic field perturbation according to the previous section, which is omitted here for brevity.<sup>6</sup>

and another for the nuclear velocity perturbation,

$$\hat{\mathcal{P}}_e \sum_l \left( \hat{\mathcal{H}}^{\text{KS}} \delta_{kl} - \langle \varphi_l^{(0)} | \hat{\mathcal{H}}^{\text{KS}} | \varphi_k^{(0)} \rangle \right) \hat{\mathcal{P}}_e | \varphi_{l,\nu\alpha}^{(1)} \rangle = \hat{\mathcal{P}}_e | \varphi_{l,\nu\alpha}^{\text{NDP}} \rangle \quad \forall \nu, \alpha, \quad (2.57)$$

whereas the former, depending on the response density, has to be solved self-consistently.<sup>16,38</sup> The sum over  $\varphi_l$  indicates the general case as this formalism is not restricted to using Kohn-Sham eigenstates, but can also be employed for derived functions such as maximally localised Wannier functions, for which the elements  $\mathcal{H}_{lk}^{\text{KS}}$  may be non-zero.<sup>15</sup>

Recently, Scherrer and co-workers presented the formulation of a *projected* NVPT which maps the nuclear velocities already on the perturbation Hamiltonian.<sup>17</sup> Thus, the prior knowledge of the nuclear velocities in MD simulations substantially reduces the number of necessary response calculations, as the sum over the set of  $\{\nu, \alpha\}$  is taken before the actual solution of the Sternheimer equations:

$$\hat{\mathcal{H}}_e^{\text{PROJ}(1)}(\mathbf{r}, t) = \sum_{\nu, \alpha} \dot{R}_{\nu\alpha}(t) \cdot \frac{\partial \hat{\mathcal{H}}^{\text{KS}}(\mathbf{r})}{\partial R_{\nu\alpha}}. \quad (2.58)$$

Equation 2.56 (and analogously Equation 2.57) simplifies to

$$- \hat{\mathcal{P}}_e \sum_l \left( \hat{\mathcal{H}}^{\text{KS}} \delta_{kl} - \langle \varphi_l^{(0)} | \hat{\mathcal{H}}^{\text{KS}} | \varphi_k^{(0)} \rangle \right) \hat{\mathcal{P}}_e | \varphi_l^{(1)} \rangle = \hat{\mathcal{P}}_e \hat{\mathcal{H}}_e^{\text{PROJ}(1)} | \varphi_k^{(0)} \rangle, \quad (2.59)$$

thereby reducing the number of perturbation calculations from  $6N_\nu$  ( $N_\nu$  denotes the number of nuclei) to two.

In the Heisenberg picture of ergodic AIMD, the isotropic IR and VCD signals at a given frequency  $\omega$  are defined as Fourier transform of time-correlation functions (TCF),<sup>k</sup>

$$\alpha(\omega) = F(\omega) \frac{\beta}{2\pi\omega} \int_{-\infty}^{\infty} dt e^{-i\omega t} C_{\dot{\mu}\dot{\mu}}(t), \quad (2.60a)$$

$$\Delta\alpha(\omega) = F(\omega) \frac{4\beta}{2\pi} \int_{-\infty}^{\infty} dt e^{-i\omega t} C_{\dot{\mu}\mathbf{m}}(t), \quad (2.60b)$$

with time  $t$ , inverse temperature  $\beta = (k_B T)^{-1}$ , and pre-factor  $F(\omega) = \frac{4\pi^2\omega}{3Vn(\omega)}$ , where  $V$  is the volume of the sample and  $n(\omega)$  is the refractive index.<sup>17,68-73</sup>  $C_{\dot{\mu}\dot{\mu}}(t)$  and  $C_{\dot{\mu}\mathbf{m}}(t)$  denote the TCF of the current dipole moment  $\dot{\mu}$  as auto-correlation and cross-correlation with the magnetic dipole

<sup>k</sup>In this work, all spectra are presented as a function of wavenumbers  $\tilde{\nu} = \frac{\omega}{c}$ .

## 2 BASIC THEORY

moment  $\mathbf{m}$ , respectively,<sup>68</sup> having the general form

$$C_{\mathbf{ab}}(t) = \langle \mathbf{a} \cdot \mathbf{b}(t) \rangle. \quad (2.61)$$

The dipole moments entering Equation 2.60 are calculated by means of molecular moments  $\dot{\boldsymbol{\mu}}_I$  and  $\mathbf{m}_I$ , the latter defined with respect to their molecular centre of mass  $\mathbf{R}_I$ . They can be considered as the sum of an electronic ( $^e$ ) and a nuclear ( $^n$ ) contribution,

$$\dot{\boldsymbol{\mu}}_I = \dot{\boldsymbol{\mu}}_I^e + \dot{\boldsymbol{\mu}}_I^n = -\langle \Phi_{\mathbf{R}} | \hat{\boldsymbol{\mu}} | \Phi_{\mathbf{R}} \rangle_{r_I} + \sum_{v \in I} Z_v \dot{\mathbf{R}}_v, \quad (2.62a)$$

$$\mathbf{m}_I = \mathbf{m}_I^e + \mathbf{m}_I^n = -\langle \Phi_{\mathbf{R}} | \hat{\mathbf{m}} | \Phi_{\mathbf{R}} \rangle_{r_I} + \frac{1}{2c} \sum_{v \in I} Z_v \mathbf{R}_v \times \dot{\mathbf{R}}_v, \quad (2.62b)$$

with  $\hat{\mathbf{m}} = \frac{1}{2c} \hat{\mathbf{r}} \times \hat{\mathbf{p}}$  and invoking the classical approximation for the nuclear subsystem summing over the contributions of all nuclei  $v$  belonging to molecule  $I$ . The subscript  $r_I$  suggests the requirement of the integral over the electronic degrees of freedom to be carried out within some well-defined subspace of molecule  $I$ , which would always yield the correct bulk polarisation after summing over molecular contributions  $I$ . An accurate way acknowledging this is to partition space based on the use of maximally localised Wannier functions.<sup>13,17,66</sup>

The electronic observables are calculated by means of the complete adiabatic KS states obtained *via* projected NVPT.<sup>1</sup> Like in the previous section, the magnetic dipole moments are calculated within a CSGT, applying the distributed origin gauge of MLWO.<sup>65</sup> However, both observables are further gathered in a distributed molecular origin gauge to give molecular moments as defined in Equation 2.62.<sup>17</sup>

Any change of the origin gauge gives rise to a *gauge transport term* adding to the magnetic dipole moment.<sup>6,7,28,38</sup> Generally, when the origin is shifted from  $\mathcal{O}$  to  $\mathcal{O}'$  with  $\mathcal{O}' = \mathcal{O} + \Delta\mathbf{R}$ , the magnetic dipole moment at the new reference takes the form

$$\mathbf{m}^{\mathcal{O}'} = \mathbf{m}^{\mathcal{O}} - \frac{1}{2c} (\Delta\mathbf{R} \times \boldsymbol{\mu} + \boldsymbol{\mu} \times \Delta\dot{\mathbf{R}}), \quad (2.63)$$

where  $\boldsymbol{\mu}$  denotes the permanent dipole moment of the fragment or molecule. Imposing a molecular origin gauge considers the gauge transport term with the common origin  $\mathcal{O}$  lying at each centre of mass of the molecules, respectively. The centre-of-mass motion  $\Delta\dot{\mathbf{R}}$  can be separated from normal

<sup>1</sup>Within the BO approximation, the observables would vanish due to the assumed stationarity of the nuclear reference.<sup>16,67</sup>



modes that appear in the frequency range focussed on in this work (800–1800 cm<sup>-1</sup>), making the last term in Equation 2.63 negligible. The gauge transport term represents a very important feature taking effect in supramolecular correlations as will be illustrated in the forthcoming chapters.<sup>1,17</sup>

### 2.4.3 Electric Field Perturbation

The calculation of UV-Vis and ECD spectra is based on electronic excitations that can be described as perturbation to the BO approximation (*vertical excitations*).<sup>m</sup> In order to compute excitations for large systems, TD-DFT provides an efficient scheme based on KS DFT (see Section 2.2).<sup>34,44</sup> Calculating the electronic response function (Equation 2.28) includes the virtual orbital space, which is a desired feature here as it allows for dissection of transition vectors into orbital contributions. The used TD-DFT scheme is based on perturbation with an electric field of the shape

$$\delta V^{\text{ext}}(\mathbf{r}, t) = -|\mathbf{E}_0| \exp(i\omega t) \mathbf{z}(\mathbf{r}), \quad (2.64)$$

with amplitude  $|\mathbf{E}_0|$  and field component  $\mathbf{z}(\mathbf{r})$ .<sup>34</sup>

Recently, Bannwarth and co-workers presented a simplified scheme, sTD-DFT, that also turned out to be more robust.<sup>74</sup> Therein, two electron integrals used for calculating the total electronic susceptibility *via* Equation 2.28 are approximated. sTD-DFT uses short-range damped Coulomb (and exchange) terms of (transition) charge density monopoles  $q_{pq}^{\nu}$  obtained from Löwdin population analysis,<sup>75</sup>

$$\varphi_p^*(\mathbf{r})\varphi_q(\mathbf{r})\frac{1}{|\mathbf{r}'-\mathbf{r}|}\varphi_r^*(\mathbf{r}')\varphi_s(\mathbf{r}') \approx \sum_{\nu,\mu} q_{pq}^{\nu}q_{rs}^{\mu}\gamma_{\nu\mu}, \quad (2.65)$$

where the indices  $\{p, q, r, s\}$  denote occupied and virtual orbitals alike and  $\gamma_{\nu\mu}$  is the Mataga-Nishimoto-Ohno-Klopman damped Coulomb operator.<sup>74</sup>

In the Schrödinger picture, the UV-Vis absorption intensities depend on the *oscillator strength*  $D_{f\leftarrow i}$  of the electronic transitions,

$$D_{f\leftarrow i}^{\text{L}} = \left| \langle \Phi_i | \hat{\boldsymbol{\mu}} | \Phi_f \rangle \right|^2, \quad (2.66a)$$

$$D_{f\leftarrow i}^{\text{v}} = \omega_{if}^{-2} \left| \langle \Phi_i | \hat{\boldsymbol{\mu}} | \Phi_f \rangle \right|^2, \quad (2.66b)$$

<sup>m</sup>Hence, no relaxation of the nuclear positions is considered.

## 2 BASIC THEORY

whereas the ECD intensity is proportional to the *rotational strength*,

$$R_{f \leftarrow i}^L = \Im \left( \langle \Phi_i | \hat{\boldsymbol{\mu}} | \Phi_f \rangle \cdot \langle \Phi_f | \hat{\mathbf{m}} | \Phi_i \rangle \right), \quad (2.67a)$$

$$R_{f \leftarrow i}^V = \omega_{if}^{-1} \Re \left( \langle \Phi_i | \hat{\boldsymbol{\mu}} | \Phi_f \rangle \cdot \langle \Phi_f | \hat{\mathbf{m}} | \Phi_i \rangle \right). \quad (2.67b)$$

The superscripts mark the dipole length (<sup>L</sup>) and velocity (<sup>V</sup>) formalism,<sup>n</sup> that can be used to calculate the electric transition dipole moment (ETDM),

$$\boldsymbol{\mu}_{f \leftarrow i}^L = \langle \Phi_i | \hat{\boldsymbol{\mu}} | \Phi_f \rangle = \langle \Phi_i | \hat{\mathbf{r}} | \Phi_f \rangle, \quad (2.68a)$$

$$\boldsymbol{\mu}_{f \leftarrow i}^V = \Im \left( \langle \Phi_i | \hat{\boldsymbol{\mu}} | \Phi_f \rangle \right) = \omega_{if}^{-1} \langle \Phi_i | \nabla | \Phi_f \rangle. \quad (2.68b)$$

The magnetic transition dipole moment (MTDM) is

$$\mathbf{m}_{f \leftarrow i} = \langle \Phi_i | \hat{\mathbf{m}} | \Phi_f \rangle = \frac{1}{2c} \langle \Phi_i | \hat{\mathbf{r}} \times \nabla | \Phi_f \rangle. \quad (2.69)$$

For exact wavefunctions the two representations of oscillator and rotational strength will be identical, according to the hypervirial theorem of quantum mechanics.<sup>76</sup> However, when approximations are introduced into the wavefunctions or, as it is the case for sTD-DFT, into the formalism, the hypervirial theorem is not necessarily satisfied.<sup>o</sup> Therefore, comparing the two formulations is an important reference to the quality of the approximated results; Bannwarth and co-workers were able to show that sTD-DFT delivers robust UV-Vis and ECD spectra improved with respect to the corresponding simplified Tamm-Dancoff approximation.<sup>74</sup>

---

<sup>n</sup>The equations of dipole and rotational strength are also valid for IR and VCD intensities, but are not considered in this work as the vibrational spectra are obtained in the Heisenberg formulation (see previous section).

<sup>o</sup>The UV-Vis and ECD spectra shown in this work are based on the velocity dipole formalism.

Part I

VIBRATIONAL CIRCULAR DICHROISM AND  
SUPRAMOLECULAR CHIRALITY



### 3 THE SAMPLING PROBLEM<sup>a</sup>

Vibrational circular dichroism (VCD) has enriched the toolbox of analysing and classifying chiral supramolecular organisation due to its outstanding sensitivity towards absolute and relative configurations<sup>6,77-79</sup> and its ability to resolve induced chirality in originally achiral regions.<sup>80-84</sup> VCD patterns emerging from supramolecular chirality (also referred to as *superchirality*) or related supramolecular effects, such as *chirality transfer* and *induced VCD*, have well been manifested in experimental investigations.<sup>82,83,85-90</sup> Yet, theoretical treatment of VCD in large-scale supramolecular setups is challenging as the size of molecular models is limited depending on the computational costs of the *ab initio* method.<sup>38,91-95</sup> Moreover, the success of a VCD calculation depends on proper coverage of thermal fluctuations in the molecules, because the conformational freedom strongly determines the spectrum.<sup>6</sup> In large setups, this poses substantial complications with regard to structure sampling and energy calculations,<sup>96</sup> when a plethora of very similar minimum-energy structures deliver delicate VCD signatures that hardly can be dealt with, let alone entropic effects that require considering finite temperature and dynamics. Hence, the widely employed search for representative minimum-energy molecular structures is not always an expedient task in a (highly correlated) supramolecular assembly.<sup>97</sup> To tackle the sampling issue, there exist powerful empirical models that are attractive due to their low-cost regime. Nevertheless, a crucial point often lies in the modulation and polarisation of charges, which requires an accurate description of the electronic structure. Hybrid *ab initio* / empirical approaches make use of molecular models, such as QM/MM setups.<sup>71,98-100</sup> Additionally, the non-local nature of VCD has been addressed by means of the tensor transfer<sup>101,102</sup> and the fragmentation method.<sup>103</sup>

Another approach represents Nuclear Velocity Perturbation Theory (NVPT) convoluted with a trajectory phase space obtained from *ab initio* molecular dynamics (AIMD) simulations at finite temperature.<sup>17,18</sup> AIMD calculations allow for an improved sampling of the statistical ensemble, naturally including entropic effects and coupled oscillations. Hence, thermal fluctuations, anharmonicities, and peak shapes will enter the spectra.<sup>73,104</sup> Moreover, conveniences, such as fully periodic su-

---

<sup>a</sup>This chapter is based on a published work written by the author. See reference 1.

percells and enhanced sampling methods may be profitably exploited. Upon the modelling of supramolecular aggregates, one can identify building blocks of non-local chirality and eventually gets to understand how supramolecular VCD is formed.<sup>1</sup> Even though AIMD may be ill-reputed as being expensive, one has to keep in mind the alternative, static calculations of harmonic IR/VCD spectra, requires the explicit calculation and diagonalisation of the full Hessian. AIMD demands computational effort only one order of magnitude higher than that of a Hessian calculation. In turn, a static calculation does not yield spectral anharmonicities, conformational flexibility of the molecules, and turns out mercurial when it comes to periodic boundary conditions.

To evaluate the reliability of both, theoretical and experimental VCD signals, the concept of *robustness* has been introduced and widely discussed.<sup>105-109</sup> There, a mode or peak in the VCD spectrum is considered robust if it is not prone to be altered much (or even inverted) by conformational changes, supramolecular interactions, or the level of theory of the underlying quantum mechanical calculation.<sup>106</sup> As elaborated by Nicu and Magyarfalvi, the angle between magnetic and electronic transition dipole moment (MTDM/ETDM) was found to strongly determine the robustness of the underlying mode. A strong deviation from orthogonality, that is to say, angles below  $10^\circ$  or above  $170^\circ$  will always lead to highly robust VCD peaks in the spectrum.<sup>106</sup> Close to orthogonality, the VCD peak sign readily swaps, even though the peak intensity may be high.<sup>b</sup> Within this chapter, several issues touching the robustness concept will be raised, primarily because non-local VCD bands are expected to turn out less robust amid the increased flexibility within the structural arrangements and the highly increased number of degrees of freedom. Although the determination of VCD robustness from AIMD simulations will not be discoursed within this work, the interested reader may notice that all calculated VCD spectra shown are equipped with the standard error range, which represents nothing less but a measure of peak robustness.

The aim of this chapter is the thorough elaboration of both an effective and illustrative comprehension of VCD in the condensed phase. Theoretical concepts of solvent effects, non-local VCD, and chirality induction are scrutinised by means of AIMD simulations with the objective of looking at thermodynamically realistic models. In the first part, the extension of the gauge-invariant VCD spectrum to the crystalline phase is presented together with the first fully periodic quantum mechanical calculations of solid-state VCD in crystals of L-alanine. The second part contains a scrutiny of (S)-lactic acid cluster formation, solvation, and resulting supramolecular VCD signatures, which raises doubts about the common explanatory models of VCD generation in aqueous solution.

---

<sup>b</sup>The value of a scalar product is determined by both, vector lengths and the angle between them.

## 4 L-ALANINE: ENHANCEMENT OF VCD IN MOLECULAR CRYSTALS<sup>a</sup>

Chiral supramolecular packing is a genuine property of a surprisingly large ratio of abundant crystals of biological as well as non-biological molecules; they are labelled as *chiral crystals*.<sup>110</sup> Frequently encountered space groups, such as  $P2_12_12_1$ , form chiral crystals regardless of whether the molecule itself is chiral or not. Dryzun and co-workers base this observation on the simple fact that if the asymmetric unit were achiral, it would form a space group of higher symmetry.<sup>110</sup> If there is supramolecular chirality present, coherent molecular vibrations in the crystal (phonons) will be able to pick up the chirality and lead to non-zero VCD signals, even if the molecule would no longer exhibit such signals when subjected to a disordered arrangement (*e.g.*, in solution).<sup>111</sup> Since the very beginning of VCD experimentation there have been reports of chirality imposition from the supramolecular arrangement. This is also because the resulting VCD signals appear quite enhanced and were thus easy to track with the spectrometers of that time having lower sensitivity.<sup>112–116</sup>

Early descriptions of VCD were accompanied by the establishment of the powerful coupled oscillator (CO) model to empirically account for distinct spectral patterns.<sup>112–117</sup> The CO model has been able to relate VCD amplification in proteins to supramolecular chirality.<sup>88,116,118–120</sup> Yet, it is given mixed assessment in view of the complex nature of supramolecular chirality.<sup>112,114,119</sup> Nevertheless, it is understood that enhanced VCD results from collective dynamics, if two preconditions are complied with: strong supramolecular connection and chiral long-range spatial order.<sup>1,121–123</sup> Lately, VCD entered the field of amyloid fibrils that are related to neurodegenerative disorders such as Alzheimer's disease.<sup>124</sup> It was shown by Nafie and co-workers that these proteins exhibit strongly amplified VCD signals in the amide region,<sup>118,125</sup> whose shape and intensity furthermore are highly sensitive to the precise fibril structure.<sup>119,123,126</sup> Interestingly, similar signals could be found in corresponding peptide microcrystals, which led to the conclusion that fibril braiding is not required for the formation of giant VCD.<sup>127,128</sup> This is inasmuch meaningful as very often chiral superstructures are advocated as

---

<sup>a</sup>This chapter is based on a published work written by the author. See reference 1.

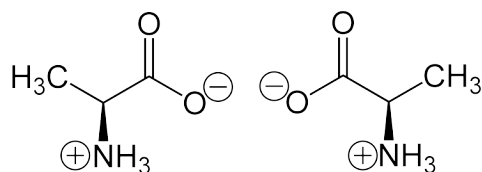


FIGURE 4.1 Lewis formulae of (*S*)- and (*R*)-alanine, also referred to as L-(+)- and D-(−)-alanine, respectively, in their zwitterionic forms.

a cause of VCD enhancement,<sup>87</sup> although it is known that coupled oscillation would not span more than two or three crystal layers.<sup>123,129</sup> It seems conceivable that fibrillar chirality moulds local chirality or *vice versa*, but that the actual vibrational coupling happens at the molecular level. It is therefore worthwhile to foster thorough understanding of VCD emerging in chiral crystals.<sup>110,114,130,131</sup>

This section presents the example of L-alanine (Figure 4.1), whose chiral crystal structure allows for supramolecular correlation of atomic motion leading to non-locally enhanced VCD. The calculations are based on realistic supercells undergoing AIMD simulations, where a fully periodic quantum description of the electronic structure of the crystal is ensured by means of periodic boundaries.



## 4.1 Gauge Invariance in the Crystalline Phase

In a crystal, when calculating gauge-invariant VCD, the issue of finding a well-defined origin under periodic boundary conditions (*cf.* Section 2.4) re-emerges. Before, in solution, well-definition of the common origin was maintained by assuming fast decorrelation of molecular motion with intermolecular distance.<sup>17,132,133</sup> As a consequence, the chiral centres appeared *pseudo-isolated* as long as large simulation cells were used. In ordered phase, however, the appearance of phonons is expected, which is entailed with large correlation lengths; the pseudo isolation ansatz seems to be inapplicable. Although the inclusion of itinerant currents in the calculation of the total cell magnetisation has been suggested to address this issue,<sup>17</sup> *modern theory of magnetisation* does not deliver phonons or long-range range correlations in a crystal, but only ensures the inclusion of surface magnetisation in the total account. Fortunately, there is no need to refrain from the solution-state approach, but it has to be ensured that supercells large enough to contain the important correlations of molecular motion in space are employed.<sup>b</sup> In the range of mid-IR VCD spectroscopy, these correlations comprise mainly optical phonons that can already be included by defining a supercell with double or triple unit cell extents, depending on the actual system.

In an adaption of the formulation of gauge-invariant VCD spectra in condensed phase,<sup>17</sup> another level of gauge sampling is added, probing the rest frame at *every* molecule in the supercell, repeatedly making use of the nearest image convention at each origin. The molecular magnetic moment of molecule  $I$  with rest frame at molecule  $K$  is

$$\begin{aligned}
 \mathbf{m}_I|_K &= \mathbf{m}_I - \frac{1}{2c} ((\mathbf{r}_K - \mathbf{r}_I) \times \dot{\boldsymbol{\mu}}_I + i \boldsymbol{\mu}_I \times (\dot{\mathbf{r}}_K - \dot{\mathbf{r}}_I)) \\
 &= \mathbf{m}_I - \frac{1}{2c} ((\mathbf{r}_K - \mathbf{r}_I) \times \dot{\boldsymbol{\mu}}_I) \\
 &= \mathbf{m}_I + \frac{1}{2c} (\mathbf{r}_{IK} \times \dot{\boldsymbol{\mu}}_I),
 \end{aligned} \tag{4.1}$$

assuming  $\dot{\mathbf{r}}_K - \dot{\mathbf{r}}_I \approx 0$ , and  $\mathbf{m}_I = \mathbf{m}_I|_I$ . The gauge-invariant VCD spectrum is obtained by averaging over all origins  $|_K$  ( $N_{\text{mol}}$  denotes the number of molecules in the simulation cell) and the

---

<sup>b</sup>In any case, the cell dimensions should at least be 2.5 to 3 times the expected correlation depth to avoid artefacts down to periodic image correlation while collecting the distributed molecular origins.

complete TCF reads

$$\begin{aligned}
\langle \dot{\boldsymbol{\mu}} \cdot \mathbf{m}(t) \rangle &= \frac{1}{N_{\text{mol}}} \sum_K \left\langle \left( \sum_J \dot{\boldsymbol{\mu}}_J \right) \cdot \left( \sum_I \mathbf{m}_I|_K(t) \right) \right\rangle \\
&= \frac{1}{N_{\text{mol}}} \sum_{I,J,K} \left\langle \dot{\boldsymbol{\mu}}_J \cdot \left( \mathbf{m}_I(t) + \frac{1}{2c} (\mathbf{r}_{IK}(t) \times \dot{\boldsymbol{\mu}}_I(t)) \right) \right\rangle \quad (4.2) \\
&= \frac{1}{N_{\text{mol}}} \sum_{I,J,K} \left\langle \dot{\boldsymbol{\mu}}_J \cdot \mathbf{m}_I(t) + \frac{1}{2c} \dot{\boldsymbol{\mu}}_J \cdot (\mathbf{r}_{IK}(t) \times \dot{\boldsymbol{\mu}}_I(t)) \right\rangle,
\end{aligned}$$

with *direct term*,  $\dot{\boldsymbol{\mu}}_J \cdot \mathbf{m}_I(t)$ , and *gauge transport term*,  $\frac{1}{2c} \dot{\boldsymbol{\mu}}_J \cdot (\mathbf{r}_{IK}(t) \times \dot{\boldsymbol{\mu}}_I(t))$  that can be rewritten to scale linear with the lever,  $\frac{1}{2c} \mathbf{r}_{IK}(t) \cdot (\dot{\boldsymbol{\mu}}_I(t) \times \dot{\boldsymbol{\mu}}_J)$ .<sup>c</sup>

$\mathbf{r}_{IK}(t) = \mathbf{r}_I(t) - \mathbf{r}_K(t)$  is the shift of molecular centres of mass. In crystalline phase under periodic boundary conditions (PBC), a nearest image convention is used, wherein  $|\mathbf{r}_{IK} - a\mathbf{L}|$  becomes minimal ( $\mathbf{L}$  is the lattice vector,  $a = -1, 0, 1$ ). Supramolecular couplings are encoded in those terms where  $I \neq J$ ; in the process, it is interesting to compare  $\sum_{I \neq J}$  with  $\sum_{I=J}$  as the latter term includes intramolecular correlations only. Under PBC, wrapping of complete molecules is ensured by having their respective centre of mass lying within the supercell.<sup>d</sup> The first term in the last line of Equation 4.2 contains direct correlations to the intrinsic magnetic moments of the molecules, while the second represents the radially dependent gauge transport carrying valuable information about the origin of VCD intensities in terms of intermolecular correlation.<sup>68</sup> In the coupled oscillator (CO) model, this term represents the coupling of dipole moments.<sup>112,114</sup>

To facilitate multidimensional spatial analysis of IR and VCD intensities, a generalised spatially resolved TCF is applied, suitable for periodic supercells.<sup>17,132</sup> The resulting radial absorptivity in  $n$ -dimensional spherical coordinates reads

$$\alpha(\omega, r) \propto \int_{-\infty}^{\infty} dt e^{-i\omega t} \Omega_n r^{n-1} C_{\dot{\boldsymbol{\mu}}\dot{\boldsymbol{\mu}}}^{\text{rad}}(t, r) \quad (4.3a)$$

$$\Delta\alpha(\omega, r) \propto \int_{-\infty}^{\infty} dt e^{-i\omega t} \Omega_n r^{n-1} C_{\dot{\boldsymbol{\mu}}\mathbf{m}}^{\text{rad}}(t, r), \quad (4.3b)$$

<sup>c</sup>For analogies to the CO model, see Abbate and co-workers.<sup>112</sup>

<sup>d</sup>The entire adaption is very similar to the formulation of local chirality and scaled moments in the neat liquid situation (*i.e.*, with more than one centre of chirality) in the work of Scherrer and co-workers.<sup>17</sup> In the present case, no scaling of moments is performed. In the limit of zero scaling, the normalisation constant becomes  $(\frac{1}{N_{\text{mols}}})^{-1/2}$  and it enters Equation 4.2 by division of the entire term with the number of molecules per supercell.

where  $\Omega_n$  is the solid angle and with the radially resolved time-correlation function,

$$C_{\mathbf{ab}}^{\text{rad}}(t, r) = \langle \mathbf{a} \cdot \mathbf{b}(t, r) \rangle. \quad (4.4)$$

The calculation of radial functions  $\alpha(w, r)/\Delta\alpha(w, r)$  stands in analogy to Equation 2.60, supported by a regularisation procedure. Smooth local densities  $\rho^a$  and  $\rho^b$  are defined on three-dimensional coordinates,

$$\rho^a(t, \mathbf{r}) = \sum_I^{N_{\text{mol}}} \mathbf{a}_I(t) \left(2\pi\sigma^2\right)^{-\frac{3}{2}} \exp\left[-\frac{(\mathbf{r} - \mathbf{r}_I(t))^2}{2\sigma^2}\right]. \quad (4.5)$$

Following the scheme presented by Heyden and co-workers<sup>132</sup> and Scherrer and co-workers *et al.*<sup>17</sup>, the comprehensive TCF,  $C_{\mathbf{ab}}(t)$ , undergoes dissection like

$$\begin{aligned} C_{\mathbf{ab}}(t) &= \int d^3r \rho^a(0, \mathbf{r}) \int d^3r' \rho^b(t, \mathbf{r}') \\ &= \int d^3r \int d^3\Delta r \rho^a(0, \mathbf{r}) \rho^b(t, \mathbf{r} + \Delta\mathbf{r}) \\ &= \int d^3\Delta r C_{\rho^a \rho^b}(t, \Delta\mathbf{r}) \\ &\equiv \int dr r^{n-1} \int d\Omega_n C_{\rho^a \rho^b}(t, \mathbf{r}) \\ &= \int dr \Omega_n r^{n-1} C_{\mathbf{ab}}^{\text{rad}}(t, r), \end{aligned} \quad (4.6)$$

where  $n$  is determined by integration over solid angle  $\Omega_n$ . The integration variable  $\Delta\mathbf{r}$  has been formally substituted by  $\mathbf{r}$  which henceforth represents the intermolecular distance. It yields the radially resolved TCF,

$$C_{\mathbf{ab}}^{\text{rad}}(t, r) = \langle \mathbf{a} \cdot \mathbf{b}(t, r) \rangle = 3\Theta_0^{\mathbf{ab}} = \frac{1}{\Omega_n} \int d\Omega_n C_{\rho^a \rho^b}(t, \mathbf{r}), \quad (4.7)$$

where  $\Theta_0^{\mathbf{ab}}$  relates to the work of Scherrer.<sup>17</sup> Equation 4.3 allows to choose  $n$  with respect to preferred spatial directions. Accordingly, the crystal topology of L-alanine suggests  $n \in \{1, 2\}$ .

In an anisotropic crystal topology, the inclusion of quadrupole moment contributions has to be considered.<sup>7</sup> In this work, calculated VCD is compared with experimental powder spectra (*cf.* Figure 4.2) that imply taking the rotational average and consider the thus isotropic  $\alpha/\Delta\alpha$ ; hence, quadrupole terms do not enter Equation 2.60 or 4.3.

## 4.2 The Origins of VCD Enhancement

### 4.2.1 Conformational Freedom

Figure 4.2 depicts the measured solid-state IR and VCD spectra of L-alanine that were obtained from ground crystals in a nujol oil mull.<sup>6</sup> From the experiment, it is clearly visible that the carbonyl peak at  $1600\text{ cm}^{-1}$  in the VCD spectrum appears very enhanced – a feature that is missing completely in VCD measurements of aqueous alanine solutions, although the IR spectra are comparable.<sup>6</sup> It has been reasoned by Nafie and co-workers that in solution, the conformational freedom of the molecule accounts for the almost vanishing band. In this sense, the carbonyl band is considered less robust, while the VCD would average out as the molecule takes on conformers of opposite (local) orientation. In the crystal, in turn, the rigid environment would clamp certain conformers and impede full averaging of the VCD, which subsequently would lead to the prominent VCD signal. According to this view, the crystal scaffold exerts a conformational bias on single alanine molecules, but the genesis of VCD still happens on the local scale. However, if one extends the idea of molecular conformers to the supramolecular scale (*superconformers*), supramolecular coupling between alanine molecules may combine with long-range spatial order that is found in the crystal, but not in solution. Yet, instead of averaging out in solution, supramolecular VCD would not emerge at all unless the molecules undergo some sort of aggregation (see also Chapter 5).

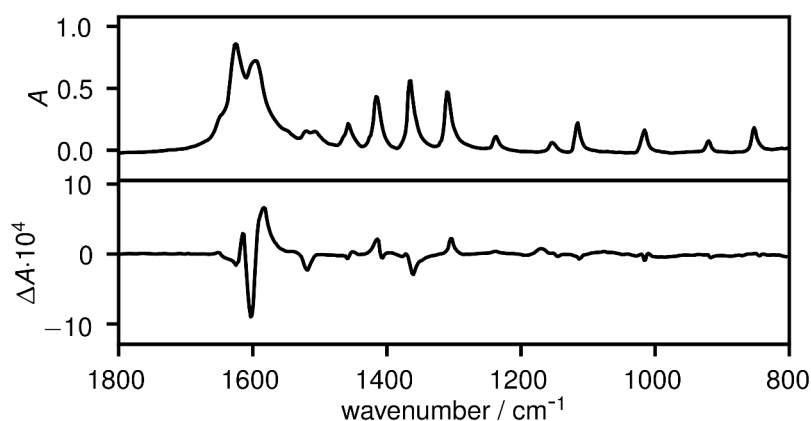


FIGURE 4.2 IR (top) and VCD (bottom) solid-state spectra of L-alanine, measured as ground crystals in a nujol mull (literature data<sup>6</sup>).

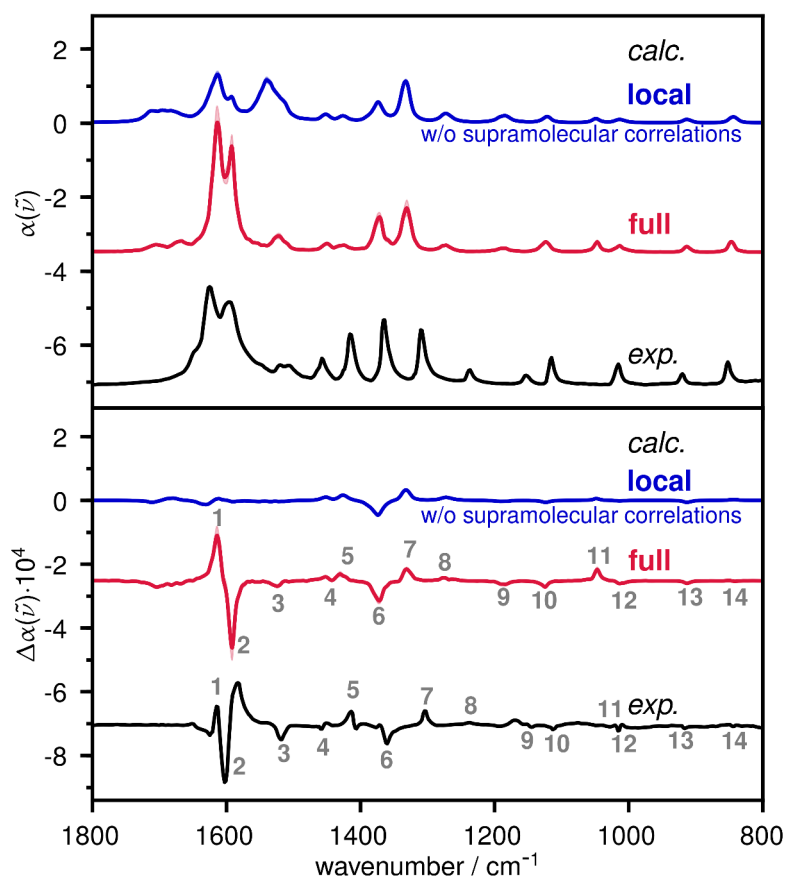


FIGURE 4.3 Calculated anharmonic solid-state IR (top) and VCD (bottom) spectra of crystalline L-alanine and experimental data, measured as a nujol oil mull (literature data<sup>6</sup>). The theoretical result is substantially influenced by the contribution of supramolecular correlations (Reprinted from reference 1. ©2018 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim).

#### 4.2.2 Local and Supramolecular Scope

Figure 4.3 compares the calculated solid-state IR and VCD spectra of an L-alanine crystal<sup>134</sup> with experimental data.<sup>6</sup> The AIMD simulations reproduce the measured data very well, including peak shapes that are mainly due to anharmonic and temperature effects.<sup>73</sup> Especially the double peak in the carbonyl region ( $1593/1614\text{ cm}^{-1}$ ) is well covered in the “full” picture (red curve in Figure 4.3), albeit it exhibits a simple bisignate shape, whereas the corresponding experimental signal appears to be more complex. The experimental spectrum indicates present scattering (all IR peaks show extended tails to low frequencies) – known as Christiansen effect.<sup>135</sup> For positive VCD peaks this settles as a negative contribution at higher frequency and a positive contribution at lower frequency and *vice versa*. To some extent, this could explain the observed differences between theory and experiment.

However, the experimental signal at position 1/2 clearly shows a superposition of at least two variants of carbonyl oscillator couplets, whose emergence may be down to morphological origins, such as edge or surface reconstructions. As the calculations consider an infinite bulk crystal, no such morphological features are included. For this reason, the calculated VCD signature appears without those artefacts.

The morphological differences between the theoretical model and the actual experimental conditions is reflected by another deviation between calculated and measured spectra. Two modes at about  $1450\text{ cm}^{-1}$  (peaks 4 and 5) exhibit an underestimated oscillator strength in the computed spectrum. Although these signals do not contribute much to VCD, they represent a quite distinct feature of the crystal IR absorption. Looking at the corresponding power spectrum gives information on how well these modes are reproduced by AIMD.<sup>e</sup> Regardless of the modes' oscillator or rotational strength, a power spectrum contains the all vibrational modes of the simulation. In Figure 4.4 those of alanine appear with equal population, which indicates that they are well-reproduced in the AIMD, and correspond to a peak from the experiment. However, the two modes mentioned above apparently receive a reduced IR activity due to increased symmetry. This is very likely connected to missing phonon deflections (*i.e.*, symmetry breaking) in the comparably small supercell that has been used for the calculations. Hence, the deviation is likely to disappear upon consideration of larger simulation cells.

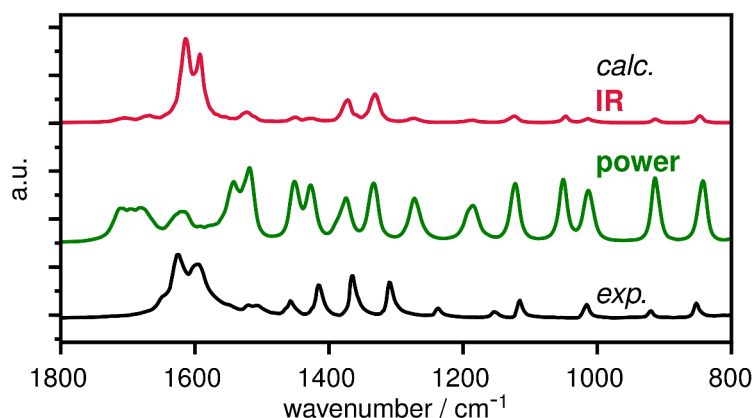


FIGURE 4.4 Calculated and experimental solid-state IR spectra as well as the calculated power spectrum of crystalline L-alanine.<sup>6</sup>

<sup>e</sup>The power spectrum can be calculated, in analogy to the IR spectrum, by means of the Fourier transformed autocorrelation function of atomic velocities, scaled with the atomic mass,<sup>35,73</sup>

$$P(\omega) \propto \sum_v M_v \int_{-\infty}^{\infty} dt e^{-i\omega t} C_{\mathbf{R}_v, \mathbf{R}_v}(t).$$

Figure 4.3 discloses that there is a difference in the computed spectra considering vibrational modes in the *local* scope – those of single molecules within the time-dependent crystal bias potential – or in the *supramolecular* scope, by means of all intermolecular correlations occurring in the supercell. Apparently, none but the latter can correctly describe the experimental data. Most interestingly, the strongly VCD-enhanced carbonyl bands 1 and 2 seem to be exclusively bound to supramolecular effects. Additionally, the IR band 3 undergoes attenuation in the full-scale picture. The following sections aim to get to the fully understand this supramolecular phenomenon by diligent dissection of the AIMD simulation cell.

## 4.3 Spatial Analysis

### 4.3.1 Crystal Topology of L-Alanine

L-Alanine crystallises in the  $P2_12_12_1$  space group;<sup>134</sup> hence it forms chiral crystals.<sup>110</sup> The zwitterionic molecules are closely held together by strong electrostatic interactions between the oppositely charged amino and carboxylate groups, with additional crossing hydrogen bonds (Figure 4.5a; see also Figure 4.1). Consequentially, the supramolecular arrangement (*i.e.*, the crystal structure) is stable at room temperature and neither rotation nor displacement of molecules is observed in the AIMD simulations (see Figure A.1). This implies that the conformational situation – in the supramolecular sense – of alanine is strongly determined by the crystal packing.

Having a closer look at the crystal's translation symmetry in *xy*-direction, one finds the amino groups of alanine arranged in a helical fashion (Figure 4.5b), which represents the inherent chirality of the unit cell as discussed by Dryzun and co-workers.<sup>110</sup> It should be noted that in alanine, the EDTM of the carbonyl-amino oscillators are collinear with the orientation of the amino group – a characteristic further examined within this chapter. Hence, the alanine case is strongly reminiscent of the “spiral staircase” model by Gulotta and co-workers, wherein the helical arrangement of coupled oscillating dipoles leads to strong VCD.<sup>116</sup> It also resembles the explanatory models of twisted crystal layers employed to explain VCD enhancement in amyloid fibrils and related compounds.<sup>118,129</sup> However in the present case, no distortions of the unit cell are considered, but supramolecular chirality is a direct result of translation symmetry. Care has to be taken when deploying the CO model, as the presence of helical chirality alone does not guarantee the emergence of supramolecular VCD, unless it

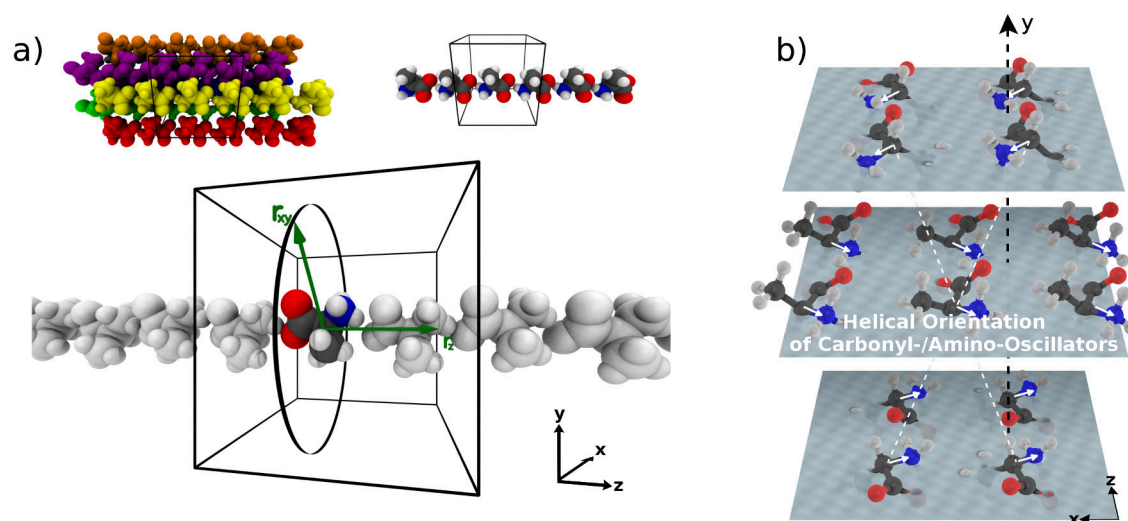


FIGURE 4.5 Crystal topology of crystalline L-alanine with, a), supramolecular chains, defining the z-vector (upper right), that are collinearly staggered (upper left). Cylindrical coordinates are declared with the radii  $r_z$  and  $r_{xy}$  defined as distance from the reference point, respectively (bottom); b), orientation of L-alanine molecules in the crystal with white arrows indicating the orientation of the carbonyl/amino oscillators (Reprinted from reference 1. ©2018 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim).

is accompanied by correlated motion. As for alanine, the question arises whether the supramolecular chirality indeed generates the characteristic features of the VCD spectrum, shown in Figure 4.3.

#### 4.3.2 Radially Resolved IR and VCD spectra

To gain insight into the spatial origins of the observed supramolecular effects, the crystal topology can be theoretically traced and the isotropic spectra calculated with spatial resolution, according to Equation 4.3.<sup>17</sup> To this end, appropriate spherical coordinates, as depicted in Figure 4.5a, are chosen. Since the alanine molecules form chains in z-direction that are collinearly packed, it appears logical to define relative distances  $r_z = |\vec{r}_z|$  and  $r_{xy} = |\vec{r}_{xy}|$  to account for this arrangement. Along these dimensions, in-space correlations are portrayed and allow to localise spectral key features depending on the chosen radial vector (see Section 4.1).<sup>f</sup>

The IR spectrum of crystalline L-alanine ( $\alpha$ ) is mainly composed of local ( $r_{xy}/r_z = 0$ ) dipole-dipole correlations of positive sign (Figure 4.6a). However, supramolecular dipole correlations interlace

<sup>f</sup> Positive and negative signals correspond to parallel and anti-parallel orientation of correlated dipole moments, respectively.



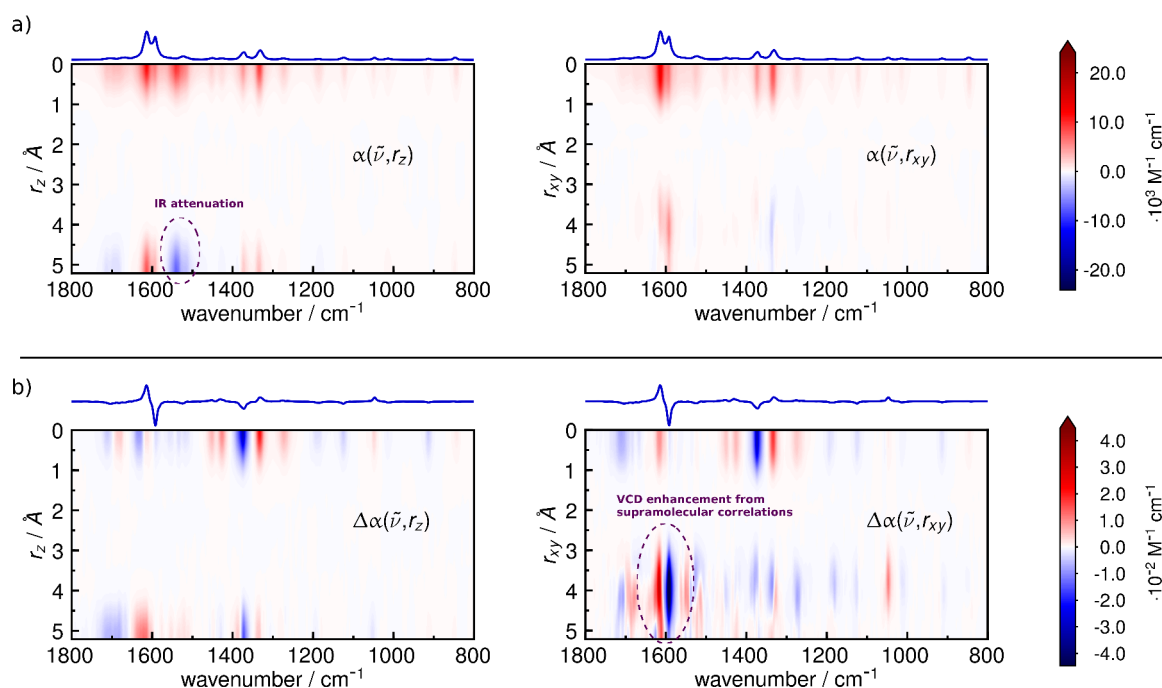


FIGURE 4.6 Radial decomposition of the isotropic IR (a) and VCD (b) spectra of crystalline L-alanine along the longitudinal axis,  $z$ , (left) and the polar axis lying in the  $xy$ -plane (right) (Reprinted from reference 1. ©2018 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim).

at intermolecular distance ( $r_z \approx 5 \text{ \AA}$ ;  $r_{xy} \approx 4 \text{ \AA}$ ). The  $z$ -dimension points in direction of the molecular chains; at  $1542 \text{ cm}^{-1}$ , strong negative correlations can be found, whereas the double peak at  $1593$  and  $1614 \text{ cm}^{-1}$  is amplified by strong positive correlations. Obviously, the  $r_z$ -dependence of IR intensity originates in zwitterionic charges being neutralised, invisible within the single-molecule frame. Also in the  $xy$ -dimension (*i.e.*, orthogonal to molecular chains), in-space correlations of dipole moments can be found, but they do not significantly alter the IR spectrum.

The spatially resolved crystal VCD spectrum ( $\Delta\alpha$ ) is shown in Figure 4.6b. Although intermolecular dipole correlations in  $z$ -dimension are clearly present – in analogy to the IR spectrum – they do not significantly contribute to the overall spectral shape. Even more interestingly, the carbonyl peaks at  $1593$  and  $1614 \text{ cm}^{-1}$ , that represent the main, amplified signals in the VCD spectrum, cannot be traced in this dimension. These, in turn, appear very clearly at the distance of the neighbouring molecules in  $xy$ -direction ( $r_{xy} \approx 4 \text{ \AA}$ ). Hence, these peaks can be put down exclusively to supramolecular correlations. Analysing the radial vibrational density of states (VDOS) of specific atom types, which corresponds to a spatially resolved interatomic power spectrum,<sup>132</sup> aids to further characterising of their nature. The vibrational coupling between carbonyl carbon atoms in Figure 4.7 clearly shows

#### 4 L-ALANINE: ENHANCEMENT OF VCD IN MOLECULAR CRYSTALS

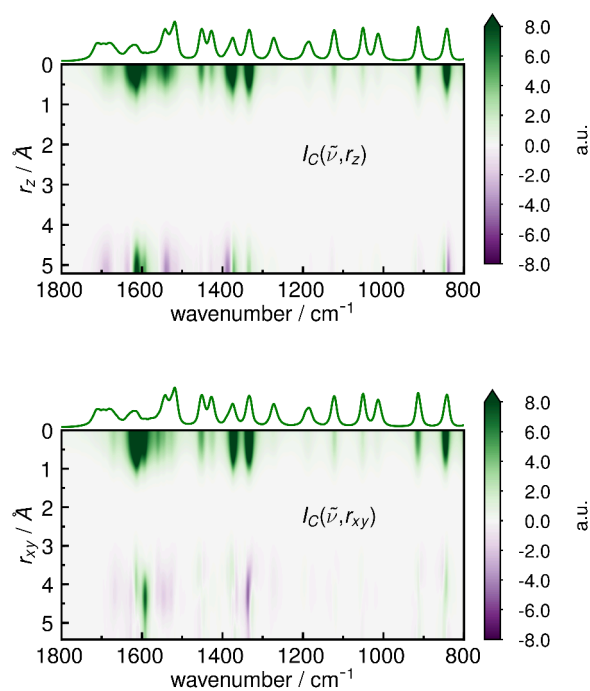


FIGURE 4.7 Radial decomposition of the vibrational density of states of the carbonyl carbon atoms of L-alanine in z- (top) and xy- (bottom) dimension. The peaks at  $r_z/r_{xy} > 0$  reveal that there is a true supramolecular correlation of atomic motion in the carbonyl region.

that at frequencies of non-local IR and VCD, also correlated motion enters the VDOS. Hence, the non-local VCD is based on collective motion of molecules in the crystal.

## 4.3.3 Non-Local Vibrations

The upper results suggest that the essence of VCD enhancement in alanine crystals lies in *coherence* of fundamental modes of the molecules. In Figure 4.8, effective modes, extracted from AIMD simulations<sup>136</sup>, and the corresponding transition dipole moments are depicted with focus on the (VCD enhanced) carbonyl region (see Table 4.1 for a complete assignment of modes and frequencies). Gradually increasing the number of molecules that enter the extraction of modes reveals, piece by piece, how the amplified VCD signal builds up. The fundamental modes of a *single* L-alanine molecule exhibit a strong ETDM in accordance with the molecule's zwitterionic state. Furthermore, in crystal organisation, the symmetric carbonyl stretching motion<sup>137</sup> mixes with the amino umbrella motion<sup>137</sup> giving two fundamental modes with linear and curved charge flow, respectively (Figure 4.8, left).<sup>138</sup> The firm electrostatic connection between molecules combines the modes to give enhanced or neutralised polarisation. According to translation symmetry in z-dimension, mode phases take effect in generating IR active and inactive combination bands of parallel ( $\sigma$ ) and anti-parallel ( $\pi$ ) phase, respectively.<sup>139</sup>

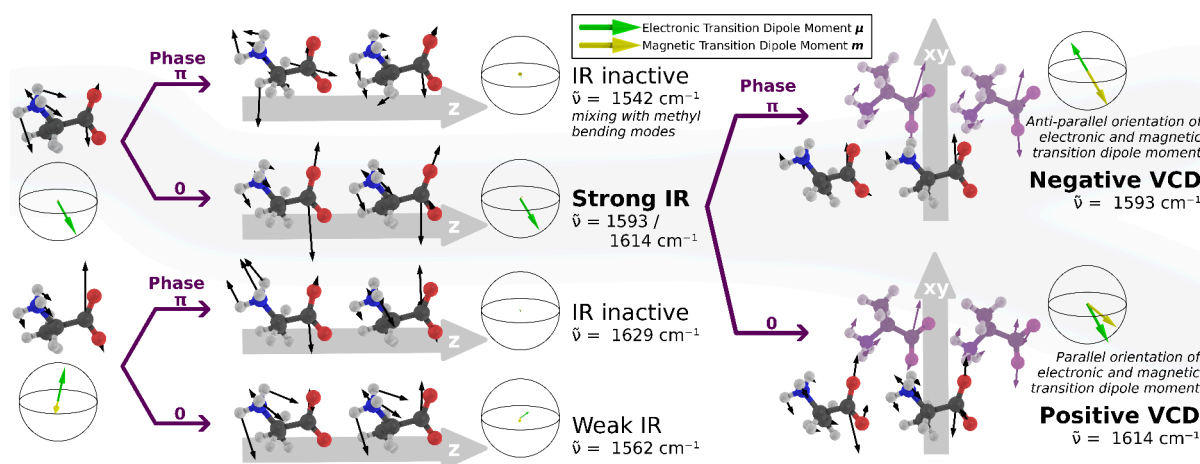


FIGURE 4.8 Calculated effective modes of L-alanine crystals in the carbonyl region with increasing supramolecular scope. The intermolecular correlation of fundamental modes gives rise to symmetric (zero phase) and anti-symmetric (phase  $\pi$ ) combinations that result in oscillators with increased or reduced transition dipole moments per molecule (globes). The number of considered molecules is augmented gradually (from one to four molecules). The IR intensity is determined by the size of the total transition dipole moment. VCD enhancement originates solely at the intermolecular distance  $r_{xy}$  of 4-5 Å (Reprinted from reference 1. ©2018 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim).

TABLE 4.1 Assignment of calculated vibrational frequencies to main features of the effective modes analysis of crystalline L-alanine. The frequency values are given with a precision of  $5 \text{ cm}^{-1}$ .

$\tilde{\nu} / \text{cm}^{-1}$	Assignment	Comment
1665/1710	$\delta_{\text{as}}(\text{NH}_3)$	
1630	$\delta_{\text{s}}(\text{NH}_3) / \nu_{\text{as}}(\text{COO})$	IR inactive
1595/1615	$\delta_{\text{s}}(\text{NH}_3) / \nu_{\text{as}}(\text{COO})$	VCD enhanced
1565	$\delta_{\text{s}}(\text{NH}_3) / \nu_{\text{as}}(\text{COO})$	
1540	$\delta_{\text{s}}(\text{NH}_3) / \nu_{\text{as}}(\text{COO}) / (\delta_{\text{as}}(\text{CH}_3))$	IR inactive
1520	$\delta_{\text{as}}(\text{CH}_3)$	
1450	$\delta_{\text{s}}(\text{CH}_3)$	underestimated dipole strength
1430	$\nu_{\text{s}}(\text{COO}) / \rho_{\text{NH}_3}(\text{C-H})$	underestimated dipole strength
1380	$\nu_{\text{s}}(\text{COO}) / \rho_{\text{CH}_3}(\text{C-H})$	
1330	$\nu_{\text{s}}(\text{COO}) / \rho_{\text{COO}}(\text{C-H})$	
1275	$\rho_{\text{s}}(\text{CH}_3/\text{NH}_3)$	
1185	$\rho(\text{NH}_3)$	
1125	$\nu(\text{C-N})$	
1045	$\rho_{\text{as}}(\text{CH}_3/\text{NH}_3)$	
1015	$\nu(\text{C-CH}_3)$	
915	$\nu(\text{C-COO})$	
845	$\nu_{\text{s}}(\text{C-COO/C-N/C-CH}_3)$	

v: stretching,  $\delta$ : bending,  $\rho$ : rocking; s: symmetric, as: anti-symmetric

The peaks at 1593 and 1614  $\text{cm}^{-1}$  appear IR enhanced by the in-phase motion, whereas the transition dipole moments of the other vibrations at 1542 and 1629  $\text{cm}^{-1}$  vanish – as expected from translational symmetry of the crystal since these modes correspond to  $\vec{k} \neq \vec{0}$  phonons. Also at 1562  $\text{cm}^{-1}$ , charge neutralisation results in a remaining weak IR signal. Notably, no VCD intensity can be deduced from this coupling (see also Figure 4.6b). It is of special importance that even temporary breaking of translation symmetry, as it happens in the AIMD by temperature effects, does not play an important part here. However, taking into account the alignment of molecular chains in xy-dimension, one gets to strong coupling of the z-chain modes (Figure 4.8, right). Now, the chiral arrangement of oscillators (Figure 4.5b) takes effect; coherence of molecular modes gives rise to strongly amplified VCD signals. Their sign is determined by parallel (o) or anti-parallel ( $\pi$ ) phase in xy (while remaining  $\vec{k} = \vec{0}$  optical phonons), leading to positive and negative VCD, respectively. To a great extent, this is mediated by the *gauge transport term* that scales linearly with distance (see Equation 4.2). In this regard, the results quantify the semi-empirical predictions of the CO model.<sup>114,116</sup> Nevertheless, there are important *direct* terms contributing to the VCD spectrum as well (Figure 4.9).<sup>112</sup> Strong supramolecular coupling breaks the degeneracy in frequency of o and  $\pi$  phase modes and ensures

their separation in the spectrum; VCD can thus probe the specific supramolecular arrangement.

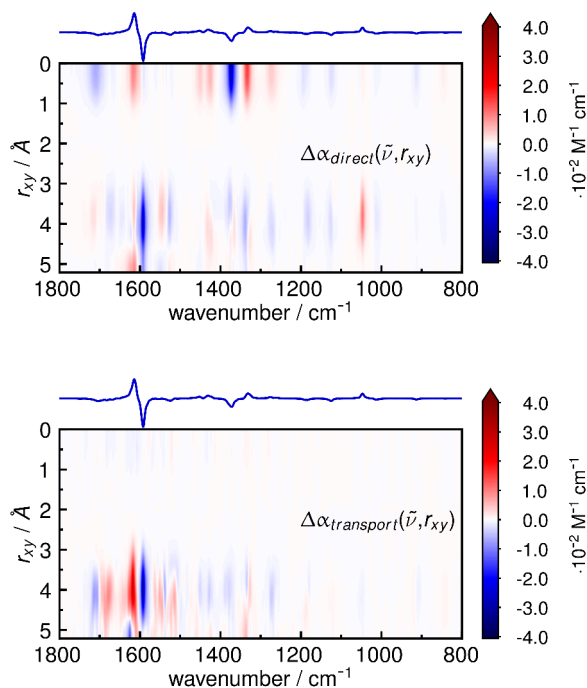


FIGURE 4.9 Radial decomposition of the direct (top) and the gauge transport (bottom) term of the isotropic VCD spectrum of crystalline L-alanine along the polar axis lying in the  $xy$ -plane.

## 4.4 Conclusion

This chapter has unravelled the non-local character of specific VCD features in molecular crystals of L-alanine. The anharmonic quantum chemical approach quantitatively describes the collective motion in the chiral supramolecular packing. Thorough dissection of the vibrational spectra is combined with an effective mode analysis to give novel insights into the mechanisms of charge polarisation in the crystal. The presence of coherent motion goes beyond the static picture of local vibrational modes embedded in a chiral environment. Consequentially, L-alanine does not exhibit VCD enhancement in aqueous solution.<sup>6</sup> Explicit consideration of the crystal topology is vital to correctly detect the key supramolecular chirality, where crystal irregularities (*e.g.*, twisting) are not required to obtain the observed VCD enhancement. While the characteristic features of the vibrational spectra can be found within the range of neighbouring molecules ( $\zeta$  Å), this concept is likely to apply equally to DNA, protein fibrils, or membranes, as long as *ab initio* models of the local packing structure can be designed.



## 5 (*S*)-LACTIC ACID: NON-LOCAL VCD AND CHIRALITY IN SOLUTION

The appearance of supramolecular VCD is not restricted to highly ordered crystalline systems, but can also be found in disordered systems, such as in solutions.<sup>140</sup> However, to allow for supramolecular correlation of motion or induced polarisation, strong supramolecular connections remain an important prerequisite. This is guaranteed in a network of hydrogen bonds between solvent and solute, but also in molecular clusters that may favourably form in solution. Although dominant contributions to the spectrum are usually located within the first solvation shell,<sup>17,132</sup> the tumultuous nature of non-covalent interactions in liquids hampers the straightforward calculation of non-local VCD. The additional scaling with intermolecular distance of the especially important gauge transport term skyrockets the computational demands to fulfil the requirement of converged (*i.e.*, ergodic) sampling. Then again, the bulk effect, that is the explicit inclusion solvent molecules, delivers a range of robust<sup>a</sup> VCD bands of *local* nature that are in fact determined by the realistic sampling of conformers and, perhaps, the inclusion of entropy in MD simulations. That sampling is less of an issue for local VCD is supported by the successful implementation and applications of hybrid QM/MM schemes by Capelli and co-workers<sup>98,100,141</sup> and can even be taken advantage of in MD-based charge flow models.<sup>69,71</sup> When it comes to non-local VCD, a full QM scheme or at least a variant containing all molecules involved in polarisation has to be considered. Scaled molecular moments as introduced by Scherrer and co-workers<sup>17</sup> alleviate the unwanted effects of signal noise from distant molecules and enforce quick decorrelation in space, which strongly reduces the computational overhead. Yet, even within the first solvation shell, correlations converge slowly.

In particular for aqueous solutions, observations of *chirality induction*<sup>b</sup> have been reported.<sup>100,142–145</sup> Xu and co-workers refer it to fundamental modes of achiral solvent molecules (*e.g.*, water) that attain VCD intensity due to the presence of a chiral solute.<sup>90,142</sup> Until now, the description of these non-local VCD phenomena happens on a heuristic basis,<sup>146</sup> although their significance for probing

---

<sup>a</sup>see Chapter 3

<sup>b</sup>Also referred to as *chirality transfer* although no literal shift, but rather a spread of chirality is reckoned.

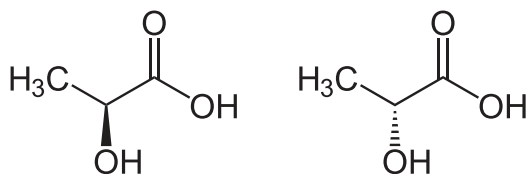


FIGURE 5.1 Lewis formulae of (*S*)- and (*R*)-lactic acid, also referred to as L-(+)- and D-(−)-lactic acid, respectively.

solvation and non-covalent interactions has well been acknowledged.<sup>82,97</sup> Identification of the key manipulators remains obscure in both, experiment and calculation. One has to keep in mind that hydrogen bond networks are subjected to lively structural fluctuations,<sup>4,147</sup> resulting in the temporal formation and dissolution of chiral sub-structures at all times, balanced to achirality. The intrusion of chiral solutes into this equilibrium leads to a non-negligible free energy bias due to the formation of stable diastereomers with specific chiral sub-structures of the hydrogen bond network. However, in the end there are usually few key solvent molecules that strongly bind to the solute and determine VCD.<sup>90,146,148</sup> Short-lived supramolecular species may be the key behind the observed *chirality induction*, yet experimental evidence for this is sparse.

The following section illustrates the example of (*S*)-lactic acid (LA) subject to a disordered environment (Figure 5.1). Beginning with assessing cluster formation in an aprotic environment, the effect of explicit solvation in the bulk phase will be the subject of AIMD simulations of realistic supercells. It will be shown, how the solute's chirality imprinted into the solvation shell can be retrieved as non-local VCD.

## 5.1 (*S*)-Lactic Acid in Aprotic Environment

### 5.1.1 Evidence for Aggregation

In Figure 5.2, the measured IR and VCD spectra of (*S*)-LA in chloroform-*d*<sub>3</sub> are shown.<sup>140</sup> Comparable to solid-state measurements of L-alanine (*cf.* Figure 4.2), also here, the carbonyl band, now in the range of 1700 to 1800 cm<sup>-1</sup>, appears very enhanced with respect to the remaining peaks. From the position of the IR peak it can be deduced that the carboxylic acid is protonated, as expected for non-protic solvation. Furthermore, its double peak nature, whose shape depends on the concentration of the LA solution,<sup>140</sup> indicates that more than one carboxylic species is present, whose



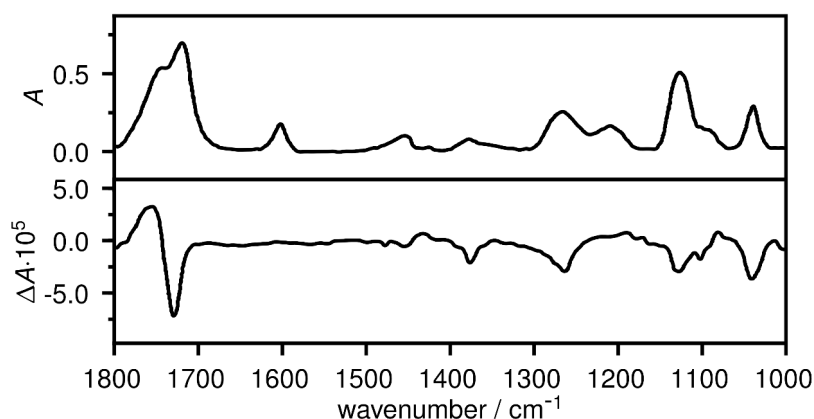


FIGURE 5.2 IR (top) and VCD (bottom) spectrum of *S*-lactic acid, measured as 0.1 M solution in chloroform-*d*<sub>3</sub>. (literature data<sup>140</sup>)

vibrational frequency is altered by the extent of hydrogen bonding. These could result from different conformers of the LA molecules, but rather suggest the presence of self-aggregation, hence the formation of molecular clusters.<sup>97,146,149</sup> The slightly enhanced VCD pattern seems to support this idea as the presence of coupled carbonyl oscillators may be the underlying reason, similar to peptides or bicamphor.<sup>97</sup>

Inspired by the textbook example of acetic acid dimers in gas phase,<sup>150</sup> the *head-on* dimer of two LA molecules seems likely and stable enough to persist in solution. Yu and co-workers pursued the idea of self-aggregation by combining VCD spectroscopy with a computational cluster analysis.<sup>97,140</sup> They proposed LA dimers to be present in solution, but concluded that the characteristic VCD pattern has to evolve from higher orders of aggregation. In the following section, the results of AIMD of LA clusters in non-interacting environment are presented.<sup>c</sup>

### 5.1.2 Monomer and Dimer

Figure 5.3 shows the optimised structures of the monomer and the head-on dimer of LA. An important degree of freedom of the molecules is the dihedral angle of the bond between the carboxylic group and the centre of chirality (molecular “core”). It is generally assumed, that a flip in dihedral an-

<sup>c</sup> The results base on AIMD simulations in a vacuum cell, which is a suitable model for aprotic environments,<sup>151</sup> even though the solvent may interact with the solute through electrostatic effects. In most cases, polarisable continuum models (PCM) are employed for describing the latter. The AIMD simulations, however, do not include PCM and the results presented underline the high relevance of entropy and dynamics over (weak) electrostatic interaction.

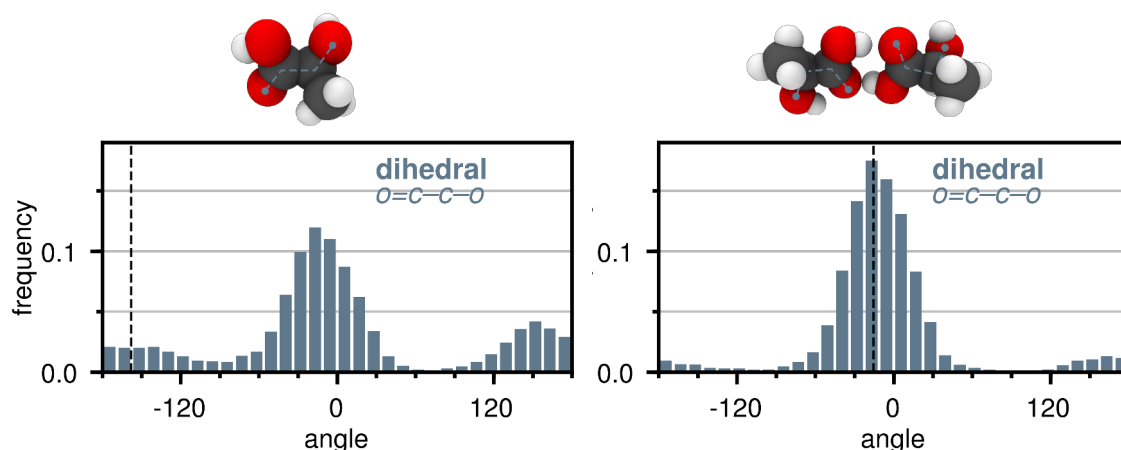


FIGURE 5.3 Top: Topologies of the (*S*)-lactic acid monomer (left) and head-on dimer (right). Bottom: Dihedral angle distribution of the molecular core (indicated in blue) as obtained from AIMD simulations. The dashed vertical lines indicate the dihedral angle of the optimised geometries that entered the normal mode analysis.

gle between two functional groups alters the VCD signature due to vibrational modes spanning both groups consequently connecting the chiral centre with the rest of the molecule. In the AIMD trajectory, the entire range of possible dihedral angles is covered, but one orientation is clearly favoured where a stabilising intramolecular hydrogen bond between the hydroxy and the carbonyl group can be formed (*i.e.*, not the conformer shown for the monomer in Figure 5.3 left). For static calculations, only one single dihedral has been chosen, respectively.

In Figure 5.4, IR and VCD spectra computed from AIMD and static calculations are compared with experimental data measured in chloroform- $d_3$ . Clearly visible for both protocols is the red-shift of the carbonyl band when going from monomeric to dimeric LA. The seemingly good fit between experimental and calculated IR spectra for monomeric LA is deceptive, since the underlying exchange-correlation functional is of GGA type (for computational details see Appendix A), which notoriously underestimates IR frequencies. This means that the actual position of the carbonyl peak should result in significantly higher frequency. In view of this, the presence of LA aggregates in chloroform- $d_3$  is therefore highly suggested. Anharmonicities enter the MD-based spectra and consequently all frequencies are red-shifted with respect to the harmonic spectra, except for the carbonyl peak that is either at equal position (monomer) or blue-shifted (dimer). After re-examination of the dihedral angle distribution in Figure 5.3, it becomes clear that two frequency shifting factors are in effect at the same time: Non-covalent interactions of intramolecular (conformers) and intermolecular (aggregation) kind. The presence of the intramolecular hydrogen bond between the hydroxy and carbonyl group slightly red-shifts the vibrational frequency of the latter. Since it is once present

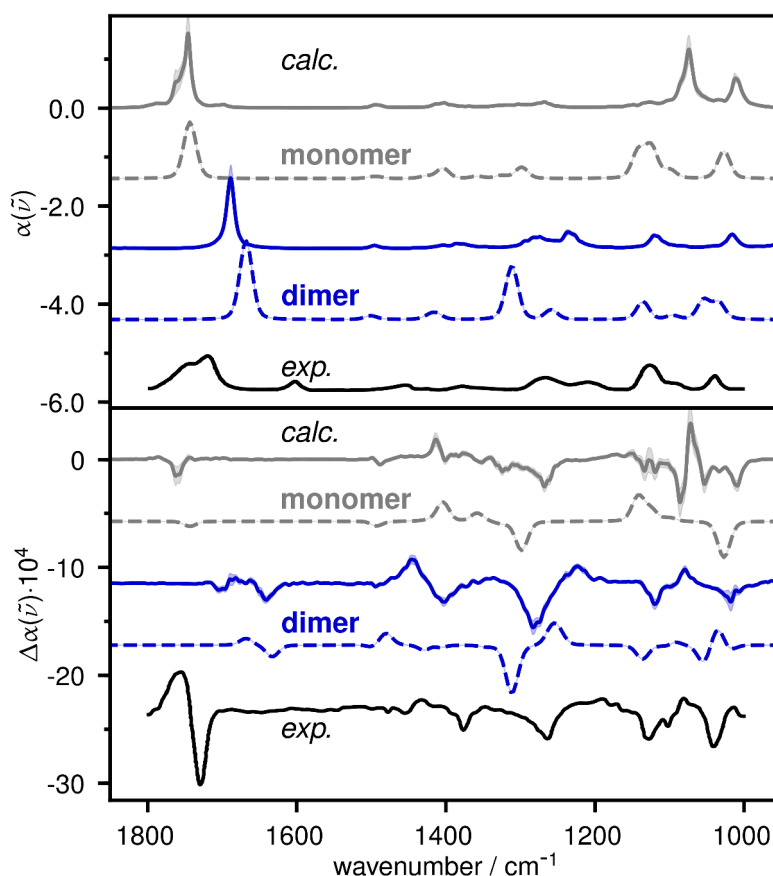


FIGURE 5.4 Calculated IR (top) and VCD (bottom) spectra of (*S*)-lactic acid species compared with experimental data, measured as 0.1 M solution in chloroform-*d*<sub>3</sub> (literature data<sup>140</sup>). A solid line represents spectra obtained from AIMD simulations (shown here with unscaled frequencies), whereas dashed lines correspond to conventional Hessian-based static calculations, artificially convoluted with a Gaussian function.

(dimer) and once absent (monomer) in the static calculations, the carbonyl band is shifted with respect to AIMD frequencies where thermal fluctuations alleviate the its strength completely – it is an artefact of minimum energy structures.

The agreement between the dimeric VCD spectrum in Figure 5.4 that originates from AIMD calculations and the experiment is captivating, especially due to the reproduction of the peak fine structure. Only the carbonyl band seems to be poorly reproduced in both AIMD and static calculations. A plain reason behind the missing VCD intensity might be LA head-on dimers not being the prevalent species in solution. This likens to the alanine case (*cf.* Chapter 4), because also there carbonyl stretching vibrations acquire rotational strength only under certain symmetric conditions. Figure 5.5 illustrates the effect of symmetry the two LA molecules experience. Considered as isolated

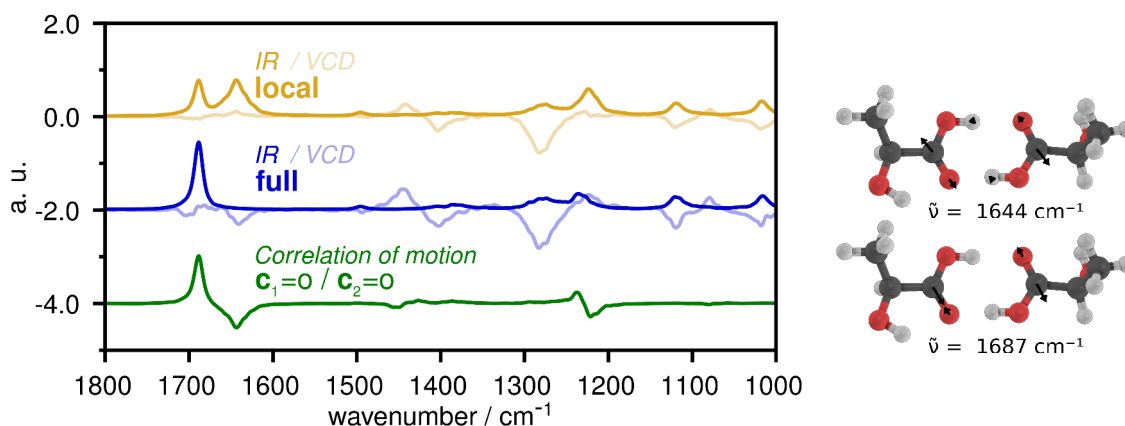


FIGURE 5.5 Left: Computed AIMD IR and VCD spectra of the monomers-in-the-dimer (yellow) and the full dimer (blue) of *S*-lactic acid. The correlation of motion of the two carboxylic carbon atoms is shown in green. Right: Two combination bands of the asymmetric carboxylic stretching vibration. Only the lower mode at higher frequency is IR active.

monomers within the dimer structure, the carboxylic group actually possesses two different localised modes emerging from the combination of the asymmetric stretch of the two molecules (see Table 5.1 for full assignment of frequencies to the vibrational modes). The correlation of motion of the carboxylic carbon atoms clarifies that these involve one constructive and one destructive combination, in accordance to the (local)  $D_{2h}$  symmetry of the carboxylic connection and the asymmetric stretch.<sup>d</sup> As a result of local symmetry, the carboxylic groups are indisposed to carry VCD unless symmetry lifting occurs.

TABLE 5.1 Assignment of calculated vibrational frequencies to main features of the effective modes analysis of dimeric (*S*)-lactic acid (*cf.* Figure 5.5).

$\tilde{\nu} / \text{cm}^{-1}$	Assignment
1644/1687	$\nu(\text{C}=\text{O})$
1494/1498	$\delta_{\text{as}}(\text{CH}_3)$
1404	$\delta_{\text{s}}(\text{CH}_3)$
1403/1449	$\rho(\text{O}-\text{H})$
1274/1321	$\rho(\text{C}-\text{H})$
1225	$\nu(\text{C}-\text{OH})$
1078/1121	$\rho(\text{CH}_3)$
1016	$\nu(\text{C}-\text{CH}_3)$

$\nu$ : stretching,  $\delta$ : bending,  $\rho$ : rocking; s: symmetric, as: anti-symmetric

<sup>d</sup>The still visible VCD traces are a result of the very large magnetic transition dipole moment (MTDM), especially in the case of the IR inactive mode at 1644  $\text{cm}^{-1}$ ; it weakly couples with the hydroxy group.

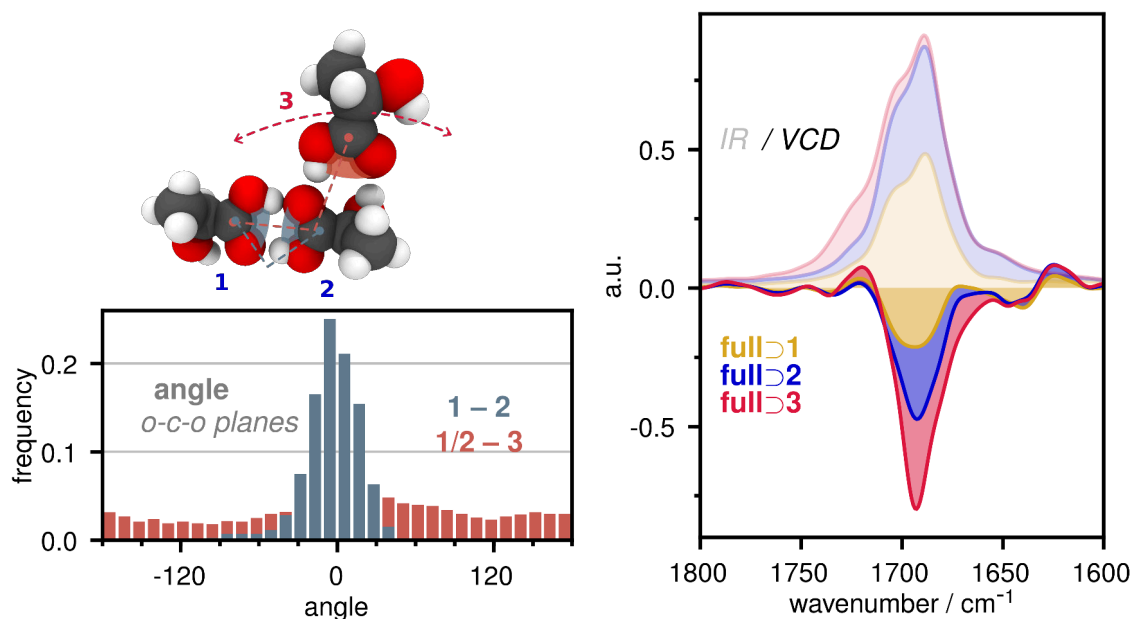


FIGURE 5.6 Left: Topology of the (S)-lactic acid trimer model “dimer + X”, together with the distribution of angles between the carboxylic planes of the molecules. Right: Molecular sub-correlations of the total IR (pale) and VCD (vivid) band in the carbonyl region. Although molecule 3 does not contribute much to the total IR spectrum, it is as equally involved in the generation of the VCD signal as the other two molecules.

### 5.1.3 Trimer and the Breaking of Symmetry

LA offers another binding site that opens up a further possibility of aggregation: the hydroxy group. Formation of a side-on *dimer* has not proved to be stable; corresponding AIMD simulations always favoured the formation of the head-on kind. Nevertheless, this does not exclude the possibility of higher orders of agglomeration. As stated in the previous section, symmetry lifting should lead to very strong VCD signals. Hence, agglomeration of the kind “dimer + X” seems to be a viable option.<sup>97</sup> AIMD calculations of a LA trimer in vacuum support this option. Apart from casual instabilities and rearrangements, the cluster maintains a specific configuration (Figure 5.6, left): Whilst a stable dimer (as before) is formed, the third molecule moves freely, always sticking to one of the hydroxy groups – and breaking symmetry. As a result, a strong VCD signal is created in the carbonyl region, visualised in Figure 5.6 (right) through the analysis of molecular sub-correlations.<sup>c</sup> It is inter-

<sup>c</sup>Molecular sub-correlations are obtained by the formula,

$$C_{ab}^{\text{full}\triangleright I}(t) = \frac{1}{2} \left( \langle \mathbf{a}_I \cdot \mathbf{b}(t) \rangle + \langle \mathbf{a} \cdot \mathbf{b}_I(t) \rangle \right),$$

where the index denotes the  $I$ -th molecule. See also Equations 2.60 and 2.61.

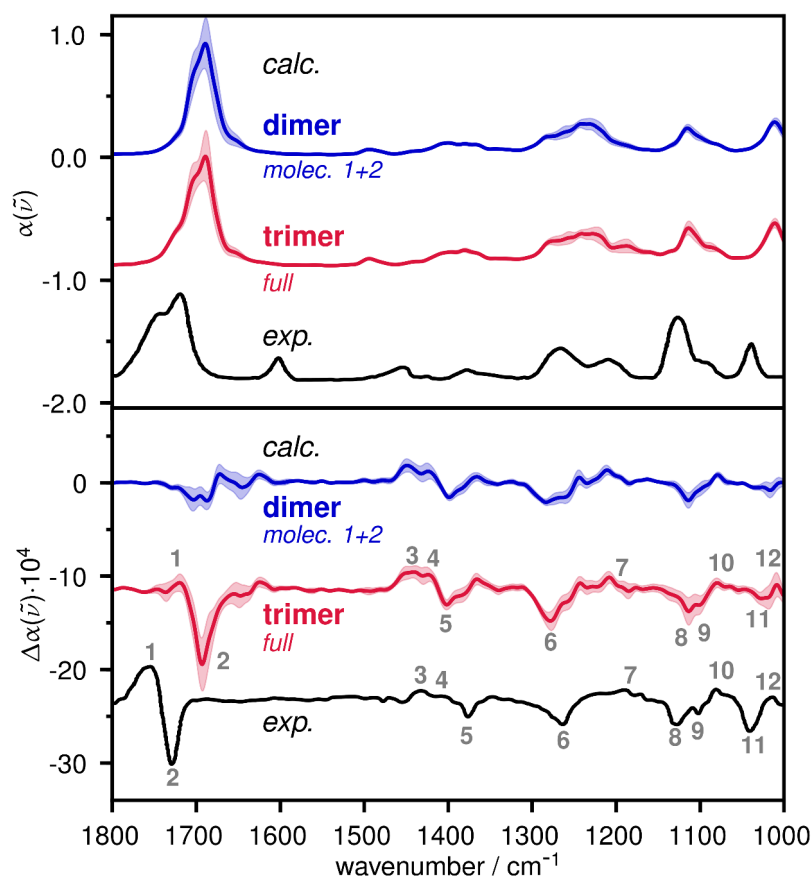


FIGURE 5.7 Calculated anharmonic IR (top) and VCD (bottom) spectra of an (*S*)-lactic acid trimer and experimental data, measured as 0.1 M solution in chloroform-*d*<sub>3</sub> (literature data<sup>140</sup>).

esting to see that, although the third molecule does not contribute much to the total IR spectrum, it is as equally involved in the generation of the VCD signal as the other two molecules. The question remains as to whether there are underlying coupled oscillations. Figure 5.6 discloses the third molecule's carbonyl oscillator exhibits a relative orientation unapt to constructively interfere with the stable dimer to yield IR intensity. It partially acts as independent oscillator, which is indicated by the IR shoulder at higher frequency. However, in terms of VCD, the relative orientation leads to strong signals due to coupled oscillation. Coupling happens at the junction of the acidic proton that provides the necessary strong supramolecular connection between the carboxylic groups.<sup>f</sup> Hence, with the third molecules being anchored to the chiral centre, it conveys supramolecular chirality.

Turning to the global spectrum in Figure 5.7, all peaks are recognisably well sampled by the AIMD

<sup>f</sup>One has to keep in mind that both, ETDM and MTDM, involved in carbonyl stretching vibrations are very large and the sensitivity of VCD in that region to any asymmetric perturbation is very high.

and the agreement with the experiment is excellent. The absence of carbonyl VCD of the dimer in Figure 5.7 supports the symmetry argument, which invokes an external perturbation to the head-on cluster to have carbonyl VCD enter the spectrum (“dimer + X”). While this non-local signal relies on coupled oscillations enabled by strong supramolecular connection, the remaining peaks in the VCD spectrum are of local nature and thus barely affected by aggregation.

The results are supported by recent matrix isolation (MI) experiments where the shape and intensity of the carbonyl band is arguably determined by temperature, but its rotational strength never reaches the magnitude of solution spectra.<sup>97</sup> The model of clusters in vacuum proves to be a valid description of LA in aprotic solvents such as chloroform.

## 5.2 Preliminary Considerations on VCD in Water

### 5.2.1 Induced VCD – Induced Chirality

Speaking of supramolecular VCD, an often made distinction between *induced chirality* and *induced VCD* (iVCD) can be found in literature.<sup>82,83,90</sup> The former is understood as an energetic bias on the conformers of an achiral molecule or atomic group within a chiral environment, whilst iVCD addresses vibrational modes that experience chiral electronic or magnetic polarisation from their embedding. In Chapter 4, collective motion represented another kind of supramolecular VCD effects. Having a closer look, induced chirality cannot be ranked among non-local VCD effects, because the ETDM and MTDM emerge from the molecules themselves; they merely are found within a chiral bias situation. Collective motion and iVCD, in turn, require not only supramolecular chirality, but also a molecular or quantum mechanical response of the environment to the motion of the perpetrator.

Interestingly, Heyden and co-workers could theoretically show that in IR spectra of water, correlated motion and electrostatic effects from the supramolecular environment can be distinguished.<sup>132</sup> Their important contribution underlined the collective character of low-frequency modes, whereas dipole correlations at higher frequencies (*e.g.*, the water bending mode at  $1650\text{ cm}^{-1}$ ) are due to electronic polarisation. The situation becomes more complex when chirality induction and iVCD appear at the same time, because, in principle, one effect does not exclude the other. For instance, one could

think of an achiral molecule as strong “donor” of polarisation (according to Heyden and co-workers) that acquires true (*i.e.*, local) chirality due to supramolecular chirality induction. Therewith, the molecule itself would have transformed into a chiral player that henceforth can induce VCD within its surrounding.

After all, induced polarisation requires the presence of polarisable charge density. Electron lone pairs provide important candidates in hydrogen-bonded networks as they are prone to sensitively respond along their electrostatic ties – as it is the case in liquid water.<sup>132</sup>

### 5.2.2 An Adaptive Moment Scaling Scheme (“Background Correction“)

To cope with the sampling problem (*cf.* Chapter 3) in disordered phase, the scaling of moments has been introduced.<sup>17</sup> The first solvation shell usually remains unaffected by moment scaling, yet superfluous dipole correlations between different solvent molecules as well as local dipole correlations may still enter the IR and VCD spectra. Thermodynamic convergence thereof requires large amounts of AIMD computation time. Similar to the experiment, having traces of solvent peaks in the spectra need not be a problem, unless solvent signals overlap with the regions of interest. It may nevertheless be desirable to filter out unnecessary contributions to IR absorption and VCD stemming from solvent molecules.<sup>§</sup> As this likens to measuring a baseline spectrum in experiment, the approach can be considered as virtual “background correction“.

In a follow-up to the work presented by Scherrer and co-workers,<sup>17</sup> a correlation term is defined, describing background signals (BG), being subtracted from the gauge-invariant formulation of the VCD spectrum<sup>h</sup> making use of moment scaling,

$$\langle \dot{\mathbf{\mu}} \cdot \mathbf{m}(t) \rangle' = \langle \dot{\mathbf{\mu}} \cdot \mathbf{m}(t) \rangle - \langle \dot{\mathbf{\mu}} \cdot \mathbf{m}(t) \rangle^{\text{BG}}. \quad (5.1)$$

The background term comprises molecular moments lying outside a region, defined by a selection matrix,  $C_{IJ}^{\text{BG}}(t)$ ,

<sup>§</sup>For instance, in the calculations the water bending mode overlaps with carbonyl stretching mode, which is worsened when using GGA exchange-correlation functionals.

<sup>h</sup>Analogously the IR spectrum.



$$\begin{aligned}
\langle \dot{\mathbf{m}} \cdot \mathbf{m}(t) \rangle' &= \sum_{I,J}^{N_{\text{mol}}} P_{IJ}(t) \left\langle \dot{\mathbf{m}}_J \cdot \mathbf{m}_I(t) + \frac{1}{2c} \dot{\mathbf{m}}_J \cdot (\mathbf{r}_{I\mathcal{O}}(t) \times \dot{\mathbf{m}}_I(t)) \right\rangle \\
&\quad - \sum_{I,J}^{N_{\text{mol}}} C_{IJ}^{\text{BG}}(t) \left\langle \dot{\mathbf{m}}_J \cdot \mathbf{m}_I(t) + \frac{1}{2c} \dot{\mathbf{m}}_J \cdot (\mathbf{r}_{I\mathcal{O}}(t) \times \dot{\mathbf{m}}_I(t)) \right\rangle \quad (5.2) \\
&= \sum_{I,J}^{N_{\text{mol}}} \left( P_{IJ}(t) - C_{IJ}^{\text{BG}}(t) \right) \left\langle \dot{\mathbf{m}}_J \cdot \mathbf{m}_I(t) + \frac{1}{2c} \dot{\mathbf{m}}_J \cdot (\mathbf{r}_{I\mathcal{O}}(t) \times \dot{\mathbf{m}}_I(t)) \right\rangle,
\end{aligned}$$

where matrix  $P_{IJ}(t) = p_I(t)p_J(t)$  is the product of scaling vectors of molecules  $I$  and  $J$  with common origin  $\mathcal{O}$ .<sup>i</sup> The selection matrix can be identified as the product of two new reciprocal scaling vectors with original scaling matrix,  $C_{IJ}^{\text{BG}}(t) = P_{IJ} \cdot (1 - \tilde{p}_I(t)) \cdot (1 - \tilde{p}_J(t))$ .

It follows

$$\langle \dot{\mathbf{m}} \cdot \mathbf{m}(t) \rangle' = \sum_{I,J}^{N_{\text{mol}}} P_{IJ}(t) \left( \tilde{p}_I(t) + \tilde{p}_J(t) - \tilde{P}_{IJ}(t) \right) \left\langle \dot{\mathbf{m}}_J \cdot \mathbf{m}_I(t) + \frac{1}{2c} \dot{\mathbf{m}}_J \cdot (\mathbf{r}_{I\mathcal{O}}(t) \times \dot{\mathbf{m}}_I(t)) \right\rangle. \quad (5.3)$$

Two scaling functions are in effect: Original scaling  $P_{IJ}(t)$  selects the active region around the common origin and enforces decorrelation within the boundaries of the supercell, whereas  $\tilde{p}_I(t)$ ,  $\tilde{p}_J(t)$ , and  $\tilde{P}_{IJ}(t)$  select a core region (*i.e.*, using a smaller cutoff) remaining unaffected by background correction.

This scaling scheme purges non-converged, hence superfluous dipole correlations within the environment but retains direct and indirect correlations with the chiral solute.<sup>j</sup> It is capable of isolating solute-solvent correlations from solvent noise, provided that the latter does not contain information relevant to the calculation of VCD. In particular in water where solvent molecules possess large dipole moments, the removal of background signal is inevitable when supramolecular correlations are to be studied.

<sup>i</sup>Moment scaling is realised by means of Fermi functions,

$$p_I(t) = \left( 1 + \exp \left( \frac{|\mathbf{r}_{I\mathcal{O}}(t)| - r_0}{D} \right) \right)^{-1},$$

with cutoff radius  $r_0$  and scaling parameter  $D$ .<sup>17,132</sup>

<sup>j</sup>No normalisation has to be done here, because only one centre of chirality is considered within the first solvation shell.<sup>17</sup>

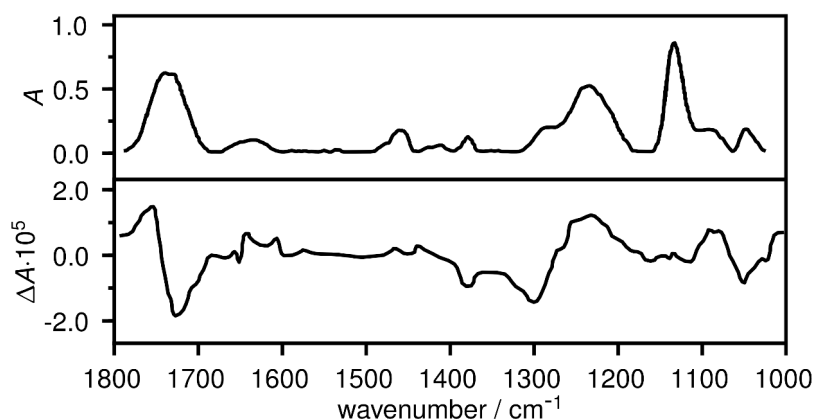


FIGURE 5.8 IR (top) and VCD (bottom) spectra of *S*-lactic acid, measured as 2.0 M solution in water. (literature data<sup>140</sup>)

### 5.3 (*S*)-Lactic Acid in Aqueous Solution

For investigating chirality induction in aqueous systems, (*S*)-lactic acid (LA) seems very suitable as it possesses two functional groups prone to form hydrogen bonds, which may effectively intrude the solvent network. In Figure 5.8, the measured IR and VCD spectra of LA in aqueous solution are shown.<sup>140</sup> It is immediately evident that the carbonyl band in the range of 1700 to 1800  $\text{cm}^{-1}$  exhibits a much weaker rotational strength with respect to the rest of the spectrum than in chloroform solution. There is a new IR peak at 1650  $\text{cm}^{-1}$ , the position where the water bending mode should appear. Also, VCD traces can be seen for this band. The authors ascribe this to an "chirality transfer" from LA to the solvent. Either way, the presence of the 1650  $\text{cm}^{-1}$  peak is a known phenomenon in vibrational solution spectra of carboxylic acids and indicates the presence of *active* water molecules, referring to molecules forming strong and comparably persistent hydrogen bonds with the acid.<sup>152</sup> If LA formed a stable head-on dimer also in aqueous solution,<sup>k</sup> the number of hydrogen binding sites open to solvent would be drastically reduced. Hence, only the hydroxy groups could be allocated as possible binding sites for active water molecules. Alternatively, water could form new clusters with the acid.

The tarnished carbonyl VCD in Figure 5.8 raises doubts about the extent of self-aggregation. In aqueous solution, the supramolecular situation of LA changes dramatically. Water molecules represent lucrative hydrogen bond partners and compete with the tendency of LA to self-aggregate and thus saturate its binding sites. In vacuum (and so in aprotic solvents), self-aggregation is the only option

<sup>k</sup>This has been suggested by Losada and co-workers.<sup>140</sup>

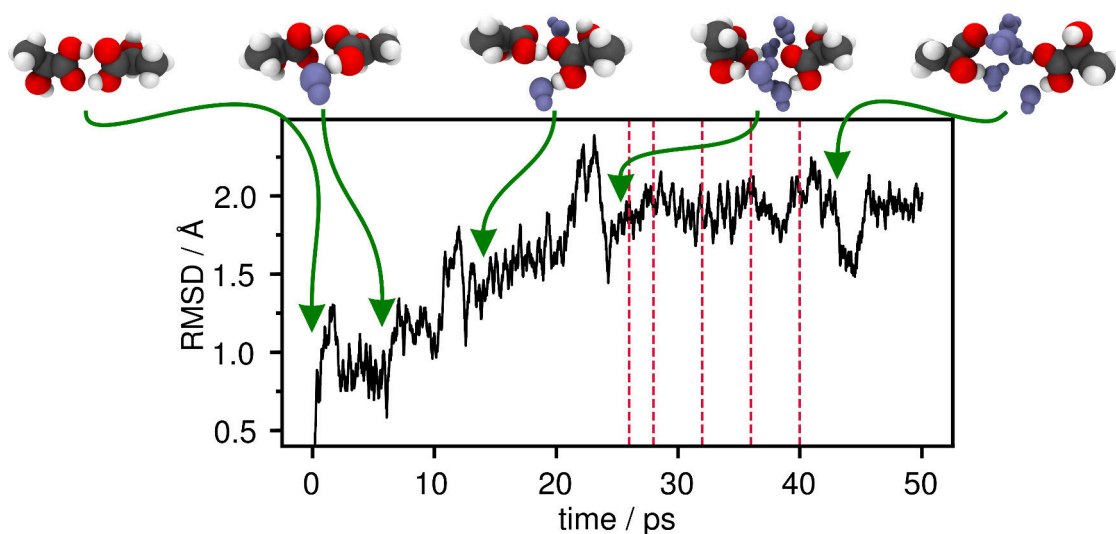


FIGURE 5.9 (*S*)-Lactic acid in aqueous solution: RMSD of the solute's heavy atoms and assigned restructuring events during the canonical AIMD. Relevant water molecules are shown in blue. The vertical lines mark furcation points of microcanonical AIMD for production of IR and VCD spectra.

to minimise free energy, yet in aqueous solution, water forms strong – perhaps even stronger – hydrogen bonds and, as there is no hydrophobic effect expected,<sup>1</sup> the maintenance of LA-LA clusters seems both energetically and entropically unfavourable.

### 5.3.1 Restructuring of the Hydrogen Bond Network

The present study continues by subjecting a head-on dimer of LA to AIMD within a fully periodic supercell of water. It is hence completely solvated in a bulk-like supramolecular environment. In Figure 5.9, the trajectory timeline of LA is presented *via* the root mean square deviation (RMSD) of heavy atoms. The continuous rise in value indicates that the initial configuration is not maintained, but unstable in favour of restructuring the hydrogen bond network. The assigned structure depictions in Figure 5.9 represent characteristic steps in the transition from a head-on dimer to completely dissolved monomers. It appears the two LA molecules act increasingly independent from each other. The initial carboxylic connection is dissolved by gradually installing water molecules between the linkages. Once separated, the LA monomers do not find back to re-form the head-on dimer (or any other cluster) within the investigated time scale.

<sup>1</sup>The hydrophobic effect describes the rise of solvent entropy upon increasing the solvent-accessible surface area, related to the *entropy of the free volume*.

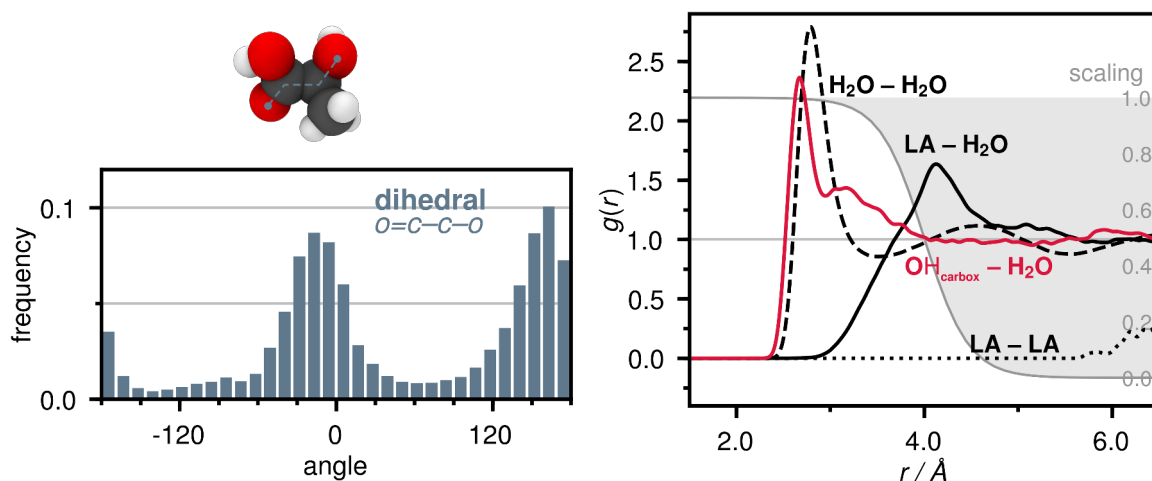


FIGURE 5.10 Left: Dihedral angle distribution of the molecular core of (*S*)-lactic acid in aqueous solution, obtained from AIMD simulations. Right: Radial distribution function (RDF) of molecular centres of mass (black) and with respect to the position of the protonated oxygen atom of the carboxylic group (red). The grey region forms a Fermi function with a 4 Å cutoff that has been employed in this work for moment scaling.

The results shown in the following were obtained from microcanonical AIMD runs sampled from the plateau region (between 25 and 40 ps) of the canonical trajectory (see Figure 5.9). Figure 5.10 (left) shows the dihedral angle distribution of the molecular core in LA. In contrast to aprotic environments, the carboxylic group populates both rotamers in approximately the same ratio. There is clear evidence that the stabilising effect of the intramolecular hydrogen bond between hydroxy and carboxylate group is completely levelled out by the presence of water. This supports the idea that the structure and energetics of LA are no longer determined by (small) local changes of conformers, but that the actual thermodynamics happen on another stage: the organisation of the solvation shell. Having a closer look at the radial distribution function (RDF)<sup>m</sup> in Figure 5.10 (right, black curves), the enduring separation of the two LA molecules becomes visible. Furthermore, organisation of the water structure remains unaffected by the presence of a solute, because the water RDF equals reported data;<sup>153</sup> the hydrogen bond network is maintained.

Looking at the distribution of water molecules around LA, a major drawback of moment scaling within the molecular origin gauge becomes palpable. Although there is a clear maximum in the RDF at about 4 Å, which indicates the presence of a (first) solvation shell, this information is misleading and in the end may lead to wrongly chosen cutoff distances. This is because the molecular origin gauge gathers the spherical distribution of molecular centres of mass around the molecular

<sup>m</sup>Performed on the distributed molecular origin gauge as the VCD calculation is carried out in this gauge.

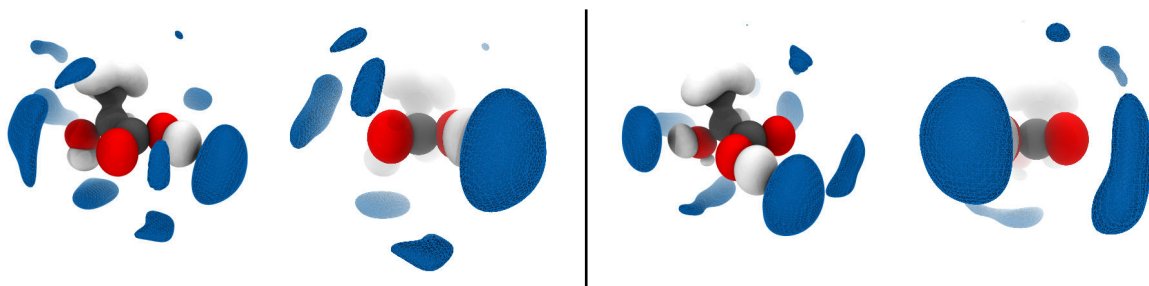


FIGURE 5.11 Spatial distribution functions (SDF) of atoms in (*S*)-lactic acid (carbon, hydrogen and oxygen atoms respectively shown in gray, white, and red) and water (only oxygen atoms, blue) obtained from AIMD simulations. The two rotamers of lactic acid (left and right) are shown in two different aspects, respectively (60% isosurface value).

origin. No isotropic information is contained, let alone the presence of specific hydrogen bonds. However, after shifting the RDF origin to the position of the protonated carboxylic oxygen atom (Figure 5.10, right, red curve), the disparate near-order of the solvation shell at the hydrogen binding site is uncovered. Interestingly, two maxima at 2.7 and 3.2 Å point to the ambivalent function of the oxygen atom as hydrogen bond donor and acceptor, respectively. Against this background, scaling of molecular moments should not be done on the basis of the molecular origin gauge, but rather atomic origins, such as the oxygen atoms of LA.<sup>n</sup>

A more detailed picture is provided by the spatial distribution function (SDF) of atoms (Figure 5.11). The solvation of the two rotamers of LA basically happens in the same way, but slight differences appear. Common to both is the clear absence of the intramolecular hydrogen bond. The hydroxy hydrogen atom instead is interacting with the water atoms in its direct surrounding. A clear difference is the establishment of a water connection between hydroxy and carboxylic group that requires a pre-orientation of the two functional groups (angle 0° according to Figure 5.10; left in Figure 5.11). An important observation in both conformers is the asymmetric hydration of the carboxylic group. Although on average the water molecules sitting at the protonated site lie in plane, the distribution at the unprotonated site is dragged out of plane. Very likely, this observation is connected with the presence of a second hydrogen binding site at the chiral centre (hydroxy group) that remotely biases the solvation shell. Hence, there is evidence for a chiral restructuring of the water network that can also be felt close to the carboxylic group.

<sup>n</sup>The calculation of VCD is still performed within the distributed molecular origin gauge. One has to keep in mind that the intermolecular distance enters the equations twice (*cf.* Equation 5.3): In the calculation of the gauge transport term, as well as in the determination of the scaling pre-factor  $p(t)$ . There is no requirement to choose the same value for both terms.

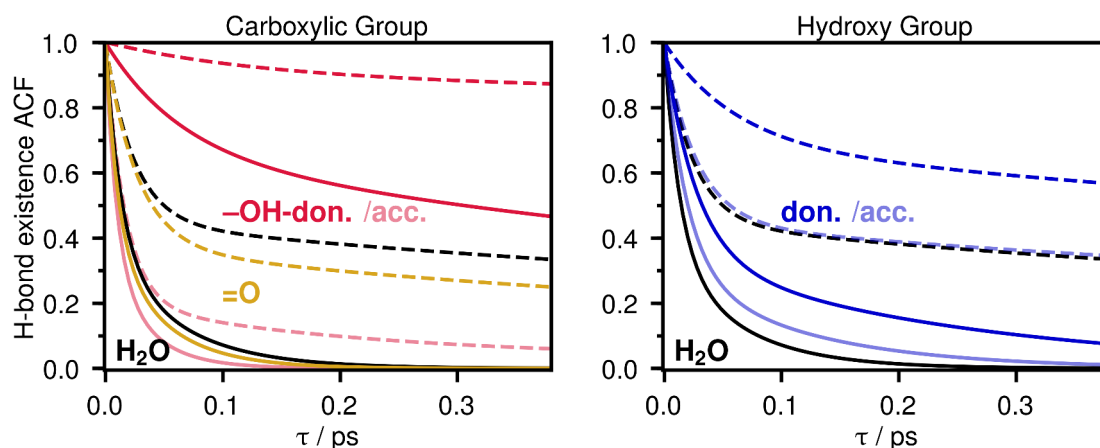


FIGURE 5.12 Lifetime analysis of hydrogen bonds of (*S*)-lactic acid in water, divided into the two sites of the carboxylic group (left) and the single site of the hydroxy group (the protonated oxygen atoms can act as both, HB donor and acceptor), based on continuous (solid lines) and intermittent (dashed lines) lifetimes. The water reference is shown in black.

The lifetime analysis of hydrogen bonds, shown in Figure 5.12, gives further insight. It is calculated as the time autocorrelation function (ACF) of the intermittent or continuous hydrogen bond population,

$$C_x(t) = \left\langle \frac{h_{ij}h_{ij}(t)}{h_{ij}^2} \right\rangle, \quad (5.4)$$

fitted to a multi-exponential function  $C_x(t) = a_1e^{-t/\tau_1} + a_2e^{-t/\tau_2} [+a_3e^{-t/\tau_3}]$ .<sup>154</sup> It sheds light into whether the fluctuation (*i.e.*, formation and breakup) of hydrogen bonds is increasing or decreasing due to the presence of LA. Regarding the carboxylic group (Figure 5.12, left), the protonated oxygen atom exhibits large differences between acting as hydrogen bond acceptor and donor. The latter is endowed with much larger lifetimes than the average water-water linkage or when acting as acceptor. This reveals a persistent water molecule residing at this site, most likely relating to the active water molecule appearing at 1650  $\text{cm}^{-1}$  in the IR spectrum (*cf.* Figure 5.8). Hydrogen bonds to the unprotonated site of the carboxylic group range among the average lifetimes of those of water. A similar situation results in the case of the hydroxy group (Figure 5.12, right), but its function as hydrogen bond donor again exhibits increased linkage lifetimes, though not as strong as in the case of the carboxylic group.

Adding LA to water clearly agitates the preceding hydrogen bond network, yet the solute is incapable of maintaining a stable head-on dimer in solution. In exchange, its acidic protons firmly bind up to two water molecules, which leads to chiral restructuring of its first solvation shell. The key question,

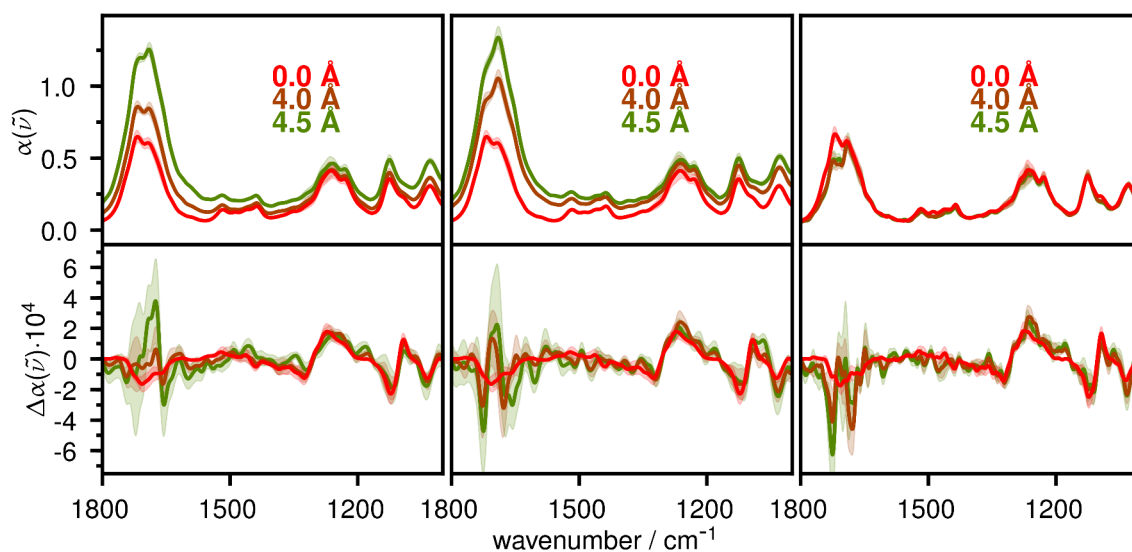


FIGURE 5.13 Computed AIMD IR (top) and VCD (bottom) spectra of (*S*)-lactic acid in aqueous solution following different moment scaling schemes at increasing cutoff value (from red to green). Left: Conventional scaling according to Scherrer.<sup>17</sup> Centre: Conventional scaling with protonated carboxylic oxygen atoms as cutoff reference. Right: Adaptive scaling according to Equation 5.3 with protonated carboxylic oxygen atoms as cutoff reference.

to what extent this invokes non-local VCD, remains.

### 5.3.2 Supramolecular VCD in Solution

In Section 5.2, the adaptive moment scaling scheme for solvated systems has been presented. Water molecules require an especially extensive sampling, because they join a large electric dipole moment with a small molecular size. Frequent reorientations and a flexible hydrogen bond network add high portions of superfluous dipolar contributions. Although they are expected to average out along the AIMD, this would need unnecessarily high amounts of computational effort. In the previous section, it was rationalised that there exist key sites of solvation; it is therefore reasonable to isolate their influence on IR and VCD by means of the adaptive scaling scheme.

IR and VCD spectra of LA in solution were calculated based on scaled moments using different cutoff radii,<sup>o</sup> whereas three variants of molecular scaling are presented (Figure 5.13).

<sup>o</sup>A cutoff radius of 0 Å stands for local contributions, *i.e.*, solely considering moments the stem from the solute.

First, conventional scaling within the distributed molecular origin gauge (left) yields converged IR and VCD bands, except for the carbonyl region where a lot of numerical noise, apparently stemming from the water bending motion, adds to the IR absorption as well. Similar observations have been made in AIMD calculations of VCD in aqueous solution of (*R*)-propylene oxide.<sup>17</sup> In the present case, the solvent seems to strongly contribute to the calculated VCD, which, even though they are not well-converged, raise the suspicion that the water supercell is chirally over-structured. This means the structural chirality originating from the chiral solutes is maintained until the borders of the simulation box. Although this need not affect the local nature of solvation of LA, it limits the expectations regarding the equilibration with longer simulation times.<sup>P</sup>

Shifting, second, the scaling reference to the protonated carboxylic oxygen atoms (*cf.* Figure 5.10) does not change the IR spectrum much, but alters shape and improves quality of the carbonyl VCD signal. The difference made here is to exclude water molecules from the calculation lying "behind" the LA molecule, that is in the region of the methyl group. It is interesting to realise that, before, those molecules had a say in the overall VCD calculation, which underlines the need of more sophisticated filtering methods than plain spherical scaling.

The adaptive scheme, third, masks out all correlations between solvent molecules. This way, the IR spectrum does not anymore change with increasing cutoff radius and also the VCD spectrum is further improved in quality. At the same time, the shaping of the carbonyl band culminates in a mainly negative rotational strength for the entire carbonyl band.

Comparing the computed spectra from adaptive scaling with the experiment (Figure 5.14), a fair agreement can be recognised. The difference between local and solvated (non-local) picture strongly resembles the previous cases: Apart from the carbonyl signal, there is no evidence of direct influence of the supramolecular environment on the VCD spectrum. These regions also agree best with experiment. In contrast, the carbonyl vibration can only limitedly attain rotational strength in the local picture. As before, the small signal stems from direct coupling with the hydroxy group. The actual feature as it is seen in experiment, however, requires the inclusion of supramolecular correlations. In more general words, carbonyl vibrations exhibit VCD in aqueous solution mainly due to the presence of water molecules and *not* as a result of self-aggregation. Yet, calculations yield two distinct carbonyl peaks in the IR and VCD spectra, whilst measurements contain a more narrow double peak.<sup>Q</sup> The computational frequency broadening of the carbonyl stretching vibration may be the result of erroneous coupling with water bending modes.<sup>r</sup> In fact, this would be an artefact emerging from using the BLYP exchange-correlation functional, which underestimates the frequency of the

<sup>P</sup>Chirality and VCD have to decay at some distant point, otherwise the experimental spectrum would look different.

<sup>Q</sup>The corresponding experimental VCD double peak has either not been resolved or it is not present at all.

<sup>r</sup>Yet, the evidence for this is low, because in the EM analysis the modes appear separated.



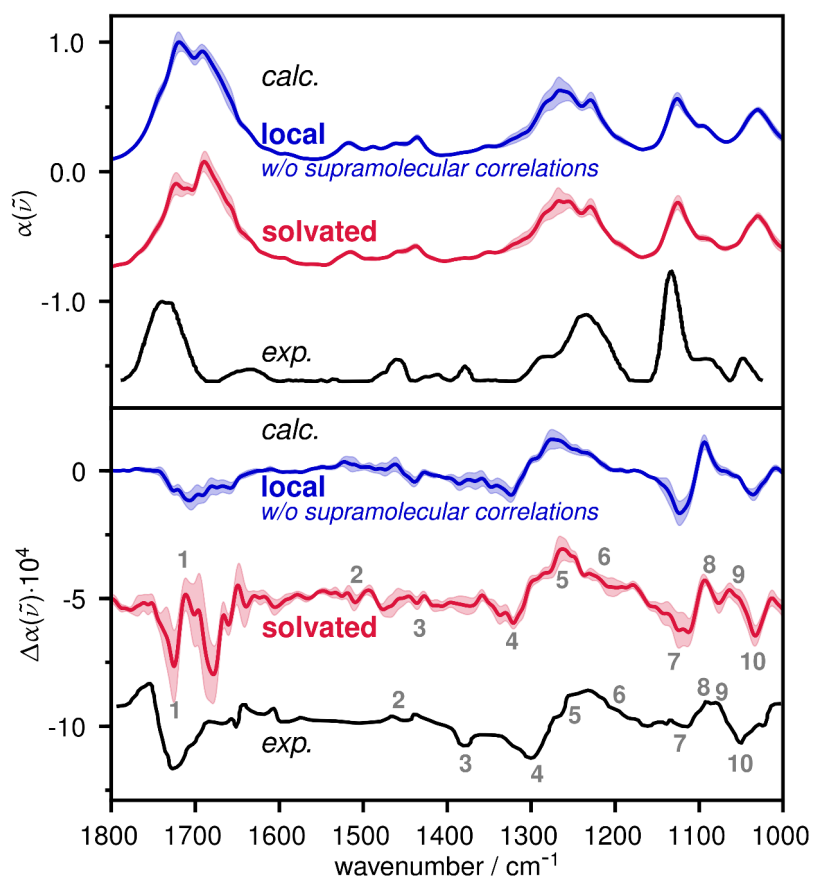


FIGURE 5.14 Calculated anharmonic IR (top) and VCD (bottom) spectra of (*S*)-lactic acid in water from AIMD, contrasted with experimental data, measured as 2.0 M solution in water (literature data<sup>140</sup>). The term "local" refers to spectral calculations only including moments stemming from the chiral molecule itself.

carbonyl stretch but not that of water bending. If the difference in separation of peaks decreased, the computed spectrum would exhibit one negative VCD band as well.

### 5.3.3 Chirality Induction and Solvation Model for Carboxylic Acids

It is worthwhile to have a look at how LA and water respectively contribute to the non-local VCD of the carbonyl region (Figure 5.15). After implementation of adaptive scaling, hardly any contributions to the IR spectrum stemming from water remain, but a clear participation in generating the VCD signal sticks out. At this point, it is not clear whether this correlation of dipole moments stems from correlated motion or electronic polarisation. Yet, as it was stated by Heyden and co-workers

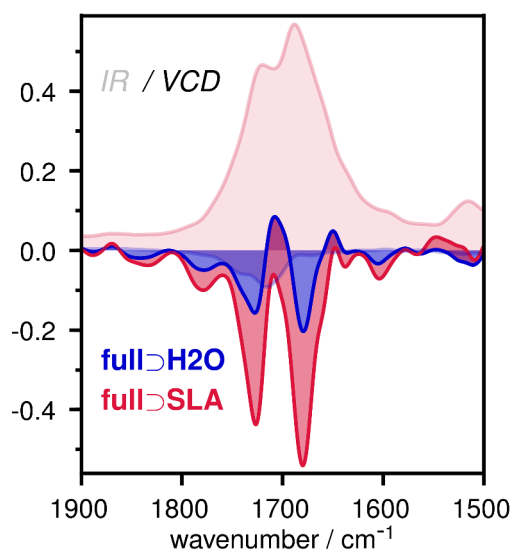


FIGURE 5.15 Molecular sub-correlations of the total IR (pale) and VCD (vivid) band in the carbonyl region ("H<sub>2</sub>O": water, "SLA": (*S*)-lactic acid). Due to adaptive scaling, the water molecules do not contribute to the IR spectrum. However, in generating the VCD signal, they are clearly involved.

(and earlier in this chapter), collective motion does not occur in the case of pure water within this range of frequency, but all correlation is of purely electrostatic origin.<sup>132</sup> This suggests that the same conclusion applies to the case of LA in water.

Figure 5.16 visualises the response of the solvation shell to a carbonyl oscillator.<sup>5</sup> As explained earlier, the carbonyl stretch entails a strong EDTM, which hardly can leave the HB network unaffected amid the omnipresence of (polarisable) electron lone pairs (Figure 5.16, left). Indeed, the local vibration invokes transition current density (TCD) of non-local nature. This exemplary analysis illustrates the polarising effect of carbonyl stretching vibrations. As the immediately attached solvent molecules "redirect" the experienced polarisation, the chiral restructuring of the solvation shell takes effect and induces VCD in LA. Figure 5.16 also illustrates intramolecular correlation with the hydroxy group, which causes weak local VCD.

In contrast to examples of the previous sections, not collective motion invokes rotational strength in the carbonyl oscillators but an electronic polarisation response of the solvent. The resulting VCD signal becomes significantly smaller than that stemming from coupled oscillators, for two reasons. First, only one causative oscillator is in effect and the electronic response does not yield transition dipole

<sup>5</sup>In this analysis, all atoms except for those of the oscillator were held fixed to mask out coupled oscillation.

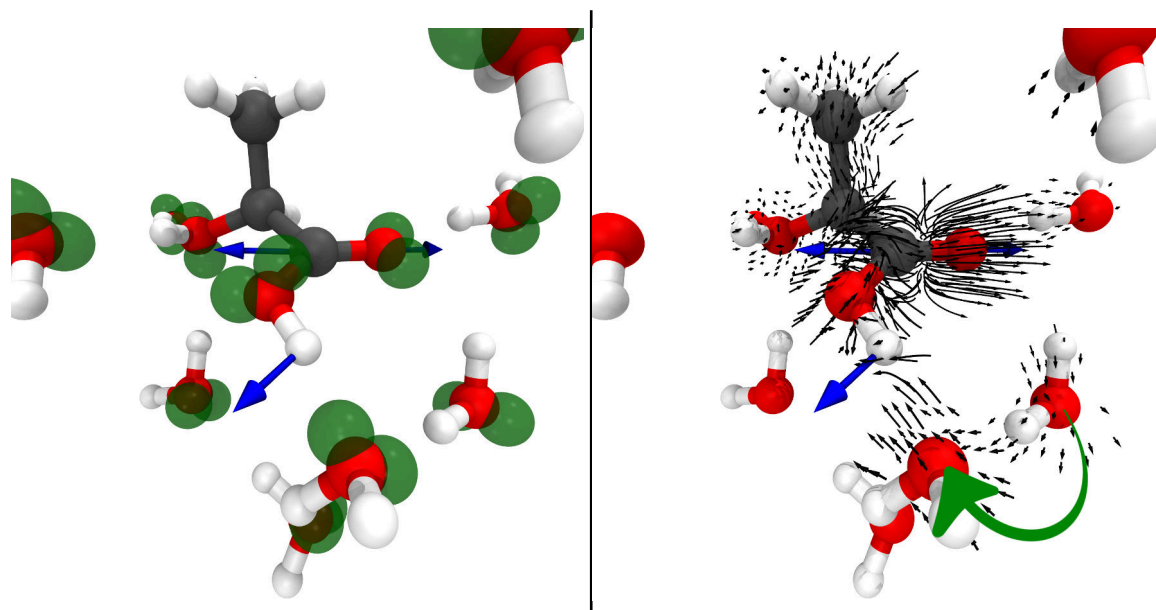


FIGURE 5.16 Snapshot from the AIMD trajectory of (*S*)-lactic acid in water, showing the carbonyl stretching mode (blue arrows). The "active water" molecule is hydrogen-bound to the protonated carboxylic oxygen atom. Left: Location of Wannier Centres (green), corresponding to electron lone pairs. Right: Transition current density (TCD) invoked by the carbonyl vibration in its surrounding. The green arrow marks the chiral redirection of current density that causes iVCD.

moments as high as additional oscillators. Second, VCD induction happens on shorter distances, hence scaling with gauge transport, which became especially important in solid-state (*cf.* Chapter 4), is less pronounced in solution. One has to admit it remains a difficult task to obtain IR and VCD spectra of satisfactory quality, having water as solvent.<sup>†</sup> But with regard to technical and biological importance of LA and beyond, water remains the most important solvent for this kind of systems.

## 5.4 Conclusion

This chapter unveiled the complicated nature of supramolecular chirality in solution, when increased mobility and actively interacting solvent molecules determine the IR and VCD spectrum. It could be shown that comparably inexpensive AIMD simulations of vacuum supercells are already of high usefulness to describe aggregation effects and supramolecular VCD amplification due to coupled oscillation. The inclusion of entropy and dynamics leads to spectra agreeing very well with the ex-

<sup>†</sup>Also in this regard, theory agrees with experiment.<sup>152</sup>

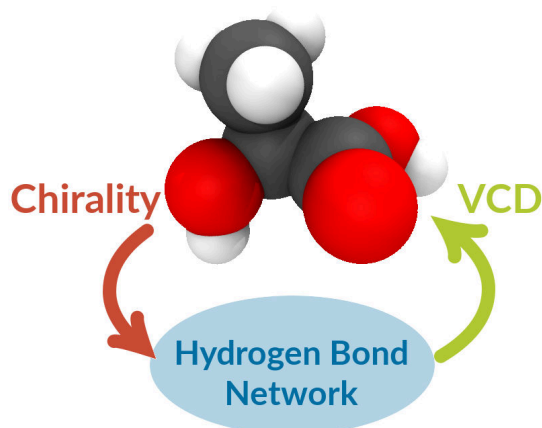


FIGURE 5.17 A novel scheme of induced VCD in solution.

periment, although no PCM has been installed. It underlines the failure of conventional frequency analysis in sampling the shallow PES of supramolecular aggregates, which is easily accessible by means of AIMD. In an aprotic environment, LA tends to form clusters of three or more monomers, which is indicated only by comparably large VCD signals in the carbonyl range, but cannot be deduced from IR absorption.

An adaptive scheme of moment scaling has been presented in order to calculate non-local VCD in solution. The supramolecular situation of LA changes completely in aqueous solution, as the solvent molecules compete with hydrogen bonds of LA aggregates and eventually lead to a complete dissolution thereof. The presence of a firmly bound ("active") water molecule can be inferred from the IR spectrum; the AIMD confirmed its location at the protonated oxygen atom of the carboxylic group. Nevertheless, identifying characteristic structural motifs and long-lived species may be helpful in effectively computing solution-state spectra.<sup>90</sup> Hydrogen bonds play an important role in generating carbonyl VCD, which leads to a novel perspective of how VCD emerges in HB networks (Figure 5.17). In terms of VCD, the main tasks of the solvent are the acquisition of chirality from the chiral centre, the restructuring of the solvation shell, and the conveyance of chirality to the dipole oscillator by means of induced VCD.

## Part II

# COLOUR TUNING OF PHYCOCYANOBILIN IN PROTEINS



## 6 THE PHYTOCHROME SUPERFAMILY

Tetrapyrroles showcase an overwhelming abundance in living species, accompanied by an impressive functional versatility.<sup>155</sup> They can grossly be classified into cyclic and open-chain forms, the former comprising the well-known porphyrins that build the precursors of hemes and chlorophylls, renowned for their important role in oxygen transport, oxidoreductase enzymes, and photosynthesis.<sup>156</sup> The second group, open-chain tetrapyrroles, also referred to as bilin chromophores<sup>a</sup>, enjoy less popularity, but they too are ubiquitous in living species. They may appear like metabolic products of their cyclic counterparts, but in most cases exhibit genuine functions in the organism, the cognisance of which is ever growing.<sup>156</sup> An important section of bile pigments is formed by light-absorbing, water-soluble biliproteins. One subgroup, phycobiliproteins, can be found in the thylakoid membranes of cyanobacteria and algae as accessory pigments that absorb different wavelengths of light and transfer the energy to photosynthetic chlorophyll *a*.<sup>157</sup> Being organised in highly elaborated light-harvesting phycobilisomes, they take advantage of wavelengths of light that lie within the “green gap” of chlorophyll, which allows their holders to populate even habitats of dire light conditions, such as in deep water.<sup>158,159</sup> Phytochromes, as another class of biliproteins, contain photoresponsive bilins acting as photosensors<sup>b</sup> that typically entail histidine kinase activity to regulate photomorphogenesis and photoperiodism.<sup>160–163</sup> In the first place, the occurrence of phytochromes was assumed to be restricted to higher plants but ever since, members of this class have also been found in cyanobacteria, bacteria, and fungi.<sup>160,163–167</sup> Recently, variants of (bacterio)phytochromes were found, genuine to cyanobacteria, that are therefore called *cyanobacteriochromes* (CBCRs).<sup>170,171</sup> They fulfil a widespread range of tasks in the cell connected to the response towards light, such as positive phototaxis.<sup>171–173</sup> Although being phylogenetically related to phytochromes, CBCRs bear unique features: While the sensory module of phytochromes needs to encompass at least a PAS–GAF–PHY tridomain (Figure 6.1), CBCRs require only a chromophore-binding GAF domain to maintain their biochemical properties.<sup>168</sup> Most intriguing, however, is the ability of CBCRs to have their absorption tuned over the entire range of the visible spectrum, includ-

---

<sup>a</sup>Named after bile pigments of mammals.

<sup>b</sup>They are furthermore sensitive to oxygen levels and heme metabolism in the cell that determine the availability of the cofactor.<sup>160</sup>

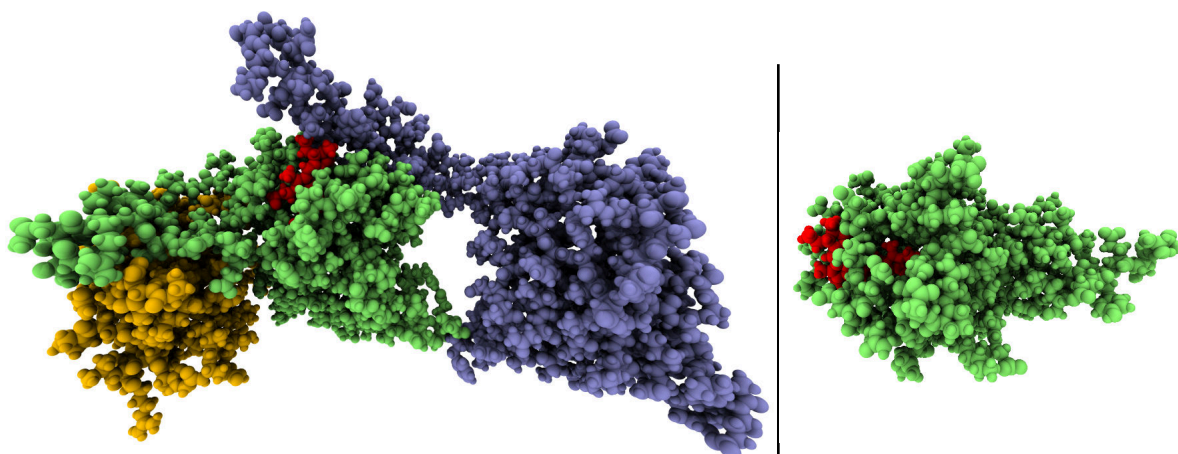


FIGURE 6.1 Left: Sensory module of CphI, a phytochrome of cyanobacterial origin. Phytochromes comprise a PAS (yellow)–GAF (green)–PHY (blue) tridomain,<sup>168</sup> wherein the latter acts as signal transmitter.<sup>166,169</sup> Right: Photoactive GAF domain (green) of AnPixJ, a cyanobacteriochrome (CBCR).<sup>170</sup> In both cases, the cofactor phycocyanobilin (PCB) is covalently bound to the GAF domain (red).

ing near-IR and near-UV.<sup>171,174–177</sup> Over the past years, a highly diverse, less differentiated toolbox of photoresponsive proteins, tailored to individual habitats and functionalities of acquiring metabolic energy, that are common to cyanobacteria has been uncovered.<sup>c</sup>

Nature tunes the properties of tetrapyrroles to their specific ends by variation of their chemical structures,<sup>155</sup> yet another more subtle level of functional determination originates in the role of the apoprotein. Very frequently too, it is observed that chemically equivalent cofactors obtain considerably different properties and tasks.<sup>155</sup> Although commonly assumed that in these cases non-covalent interactions between apoprotein and cofactor account for the modulations, it often remains obscure how they precisely take effect. Amongst others, electrostatic effects, hydrogen bonding (HB), and sterical effects are possible ways in which the protein environment can influence the cofactor.<sup>155,174,178–180</sup> The cofactor itself brings in coordination sites, electronic structure, and conformational flexibility. Here, bilin cofactors substantially deviate from their cyclic pendants as they access a strongly increased conformational phase space allowing for ring flips and intramolecular hydrogen bonds. At the same time, the electronic structure and related properties strongly differ due to the lack of aromaticity and a (partial) loss of  $\pi$ -conjugation.<sup>162,181</sup>

Phytochromes, bacteriochromes, and CBCRs share a specific set of bilin cofactors among them-

<sup>c</sup>From the evolutionary perspective it clearly is an advantage – or rather an imperative – of the cell having available various means for adaptation that can be switched to by re-regulation of gene expression under selection pressure.



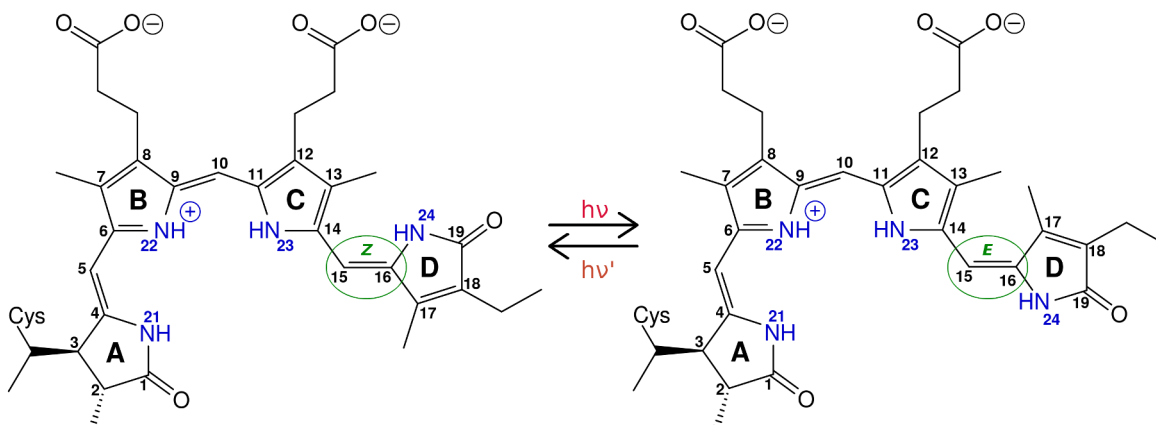


FIGURE 6.2 Lewis formula of phycocyanobilin (PCB) in its *ZZZssa* (left) and *ZZEssa* (right) configurations. Reversible photoisomerisation of the C<sub>15</sub>-C<sub>16</sub> double bond triggers the photoswitch.

selves.<sup>182</sup> Plant phytochromes exclusively bind *phytochromobilin* (PΦB), fungal phytochromes and bacteriochromes *biliverdin* (BV). However, phytochromes in cyanobacteria bind either biliverdin or *phycocyanobilin* (PCB, Figure 6.2). In turn, CBCRs can have phycocyanobilin or *phycoviolobilin* (PVB) as chromophore.<sup>175</sup> Regarding the diversity of CBCRs, it is uncertain whether there also exist variants containing phytochromobilin, but it seems that, in this regard, plants are monophyletic.<sup>182</sup> All members of the photoreceptor superfamily are characterised by the chromophore being covalently attached<sup>d</sup> to the protein by means of thioether bridges, linking to a highly conserved cysteine residue in the protein *via* pyrrole ring A (*cf.* Figure 6.1),<sup>160</sup> except for PVB that binds to two cysteine residues.<sup>175,183</sup> Often (but not always) the chromophore exhibits a *ZZZssa* configuration, which denotes the configuration of the double and single bonds of the three methylene bridges (atoms C<sub>5</sub>, C<sub>10</sub>, and C<sub>15</sub>; *cf.* Figure 6.2). With the bilin rings arranged more or less in a co-planar fashion, the conjugated system extends from ring A until ring D. In practice, however, ring A and ring D are observed to be rotated out of plane.<sup>155</sup>

The photoreceptor proteins operate as *photoswitches*.<sup>161</sup> On the basis of a reversible photoconversion mechanism they can switch between two different absorption forms. Exciting the chromophore at its absorption maximum results in photochemical *Z/E* isomerisation of the C<sub>15</sub>-C<sub>16</sub> double bond. The flipping of the D ring is associated with conformational changes in both cofactor and protein, but also in the interaction pattern between them, which triggers a biochemical response.<sup>160,184-189</sup> The two states, *Z* and *E*, are denoted as *dark state* and *photoproduct*, respectively. However, the photoproduct is likewise back-transformed into the dark state by photochemical *E/Z* isomerisation

<sup>d</sup>The cofactor is therefore denoted as prosthetic group.

after excitation at its own absorption maximum. Depending on the colour ( $x$ ) of absorbed light, the states are labelled as “P<sub>x</sub>”. Phytochromes and bacteriochromes switch between red-absorbing (P<sub>r</sub>) dark states and far-red-absorbing (P<sub>fr</sub>) photoproducts,<sup>161,164</sup> yet, as introduced above, the colour diversity in CBCRs (and phycobiliproteins) is huge; moreover they are capable to tune the absorption properties of dark state and photoproduct, independently. Typical examples of CBCR absorption colour pairs are hence red/green, green/red, blue/green, red/orange, green/orange, violet/orange, blue/yellow, or blue/teal.<sup>171,172,175–177,190,191</sup>

The colour spectrum of CBCRs is exceptional, regarding the very limited number of (chemically different) chromophores they bind. Regarding the mechanisms behind *colour tuning*, four general options have been considered.<sup>170,172,175,176,178,192–196</sup> Rockwell and co-workers suggested a protonation-dependent (proteochromic) mechanism to be the origin of colour change.<sup>172,178,193</sup> Escobar and co-workers brought in a solvatochromic model that would determine colour by different amounts of solvent molecules present near the chromophore for dark state and photoproduct,<sup>195</sup> which was taken up by Song and co-workers who related colour tuning with the presence of charged protein residues that shift the partial charge pattern within the chromophore (in analogy to the case of rhodopsin<sup>179</sup>).<sup>197</sup> Dual-cysteine mechanisms have been confirmed in a large group of CBCRs, wherein PCB is chemically transformed into PVB by linkage to a second cysteine residue;<sup>175,183,196</sup> in PVB, conjugation is interrupted by a sp<sup>3</sup>-centre at the ring A/B or B/C link, which accounts for a substantial blue-shift in absorption. The effect of the  $\pi$ -conjugation length on the bilin colour were discussed all along.<sup>155,170,192</sup> Hence the last option is distortion of the chromophore by steric interaction with the protein, twisting of the outer rings (A and D), and shortening of conjugation (“trapped-twist” model).<sup>155,170,176,194</sup> Peng and co-workers experimentally related the dihedral angle of ring D of PCB in allophycocyanin B (a phycobilisome) with the absorption maxima ranging from 590 to 670 nm and thus proved the bathochromic shift originating in an enhanced level of co-planarity that results in an increased conjugation length.<sup>159</sup> Recently, Wiebeler and co-workers showed the hypsochromic shift in the green-absorbing photoproduct of Slr1393g3 (a CBCR-carrying PCB) to emerge from a decrease of conjugation, again originating from the out-of-plane rotation of ring D.<sup>174</sup> They also ruled out an solvatochromic effect on the spectrum, but underlined the important role of the apoprotein in adjusting the chromophore in the intended conformational state.

The superfamily of photosensor biliproteins are the “visual pigments” of their holders<sup>155</sup> used for innumerable biochemical purposes related to light. Their function as photoswitches allows for a precise measurement of light intensities of any wavelength, be it to evaluate rewarding positions to set up a photosystem, distinguish day from night, or navigate motion just to name a few.<sup>162</sup> The

treasure of photosensors provided by nature with which organisms adapt to their environment bears high utility for biochemical applications in bioimaging and optogenetics, but also in therapy and catalysis.<sup>172,181,191,198–209</sup>

## 6.1 A Case Study: P<sub>r</sub> States of CphiΔ<sub>2</sub> and AnPixJg<sub>2</sub>

This chapter focusses on two photosensor proteins that can both be found in cyanobacteria:

CphiΔ<sub>2</sub>,<sup>e</sup> a phytochrome with characteristic PAS-GAF-PHY tridomain (Figure 6.1, left), represents a well-studied member of the phytochrome superfamily;<sup>54,166,169,175,186–188,210–229</sup>

AnPixJg<sub>2</sub>,<sup>f</sup> a cyanobacteriochrome (Figure 6.1, right), has been discovered later and still lags behind in being explored.<sup>2,170,171,178,183,193,195,197,230,231</sup> The proteins both contain PCB as cofactor (Figure 6.2) and exhibit a red-absorbing dark state (P<sub>r</sub>). However, photoconversion yields a far-red absorbing photoproduct (P<sub>fr</sub>) in the case of CphiΔ<sub>2</sub>, meanwhile AnPixJg<sub>2</sub> produces a green-absorbing state (P<sub>g</sub>) – spectral shifts in opposite directions by more than 100 nm. The photoproduct formation does neither imply chemical changes to the chromophore, nor a changed protonation state.<sup>178,195,231</sup> It rather seems they resemble the objects studied by Peng and Wiebeler (see above), which both concluded the D-ring twist and thus extended or reduced conjugation length to determine the absorption properties.<sup>159,174</sup> Agreed on in principle, similarly coloured pairs should rest on similar spectral tuning mechanisms; hence all phytochromes operate the same way and results obtained for green/red Slr1393g<sub>3</sub><sup>174</sup> are expected to be transferable to AnPixJg<sub>2</sub>.

This in particular matters, because to date no crystal structures of the P<sub>fr</sub> and P<sub>g</sub> state of CphiΔ<sub>2</sub> and AnPixJg<sub>2</sub>, respectively, that could build the basis of theoretical structure models, exists. With the two P<sub>r</sub> states being well characterised by UV-Vis, Raman, and NMR experiments,<sup>g</sup> they seem to be perfect candidates to subject colour tuning to profound investigation. Unfortunately, relating *in vivo* or *in vitro* data to structures stemming from X-ray diffraction may lead to equivocal results. Even though crystal data provides invaluable information and forms the basis of any theoretical struc-

<sup>e</sup> CphiΔ<sub>2</sub> stands for “cyanobacterial phytochrome 1”, *i.e.*, the first phytochrome found in cyanobacteria, more precisely *Synechocystis* sp. PCC 6803 (slr0473).<sup>166,210,211</sup> It functions as light-regulated histidine protein kinase.<sup>166,169</sup> Δ<sub>2</sub> denotes the sensory module of the protein comprising 514 N-terminal residues (Δ<sub>1</sub> comprises 492 N-terminal residues).<sup>212</sup>

<sup>f</sup> AnPixJg<sub>2</sub> denominates the second GAF domain (out of four) of a phototaxis regulator found in heterocyst-forming cyanobacteria *Anabaena* (Nostoc) sp. PCC 7120 (all1069).<sup>170,183</sup> *Pix*, originally *pis* (PhototaxIS), is a class of regulatory genes required for positive phototactic movement (PositIve phototaXis), whereas *J* is a serial letter.<sup>173</sup>

<sup>g</sup>*cf.* references<sup>2,169,180,185,186,188,189,197,215,216,219,221,222,226,227,229,231–235</sup>

ture model, the static picture does not reflect the real situation of the protein.<sup>236,237</sup> There can be substantial deviations from solution-state structure, especially in terms of solvation and regarding highly flexible side-chain or backbone segments.<sup>238</sup> In turn, three-dimensional structures obtained by NMR experiments in solution can suffer from a lack of explanation when differences to X-ray data become significant.<sup>239</sup>

This chapter follows up preceding experimental (and theoretical) work on Cph1 $\Delta$ 2 and AnPixJg2 in order to provide remedy for the conflict of static and solvated pictures. With the help of FFMD and AIMD simulations at finite temperature, realistic models augment existing concepts with dynamics and entropy, which affects non-covalent interactions and the accessible conformational space the most (Figure 6.3). Structural fluctuations can profitably be correlated with calculated spectroscopic patterns and thus increase the explanatory power of experimental spectroscopy. In the first part, long-scale FFMD simulations of Cph1 $\Delta$ 2 and AnPixJg2 in solution are analysed and observations, such as structural heterogeneity, are discussed. The second part presents the results of NMR and UV-Vis calculations based on AIMD simulations within a QM/MM framework; the relevance of choosing a sampling scheme is discussed alongside. Eventually, the obtained results are used to spectroscopically characterise the cofactor in its protein environment and conclusions will be drawn regarding the colour tuning mechanism of PCB in both proteins.

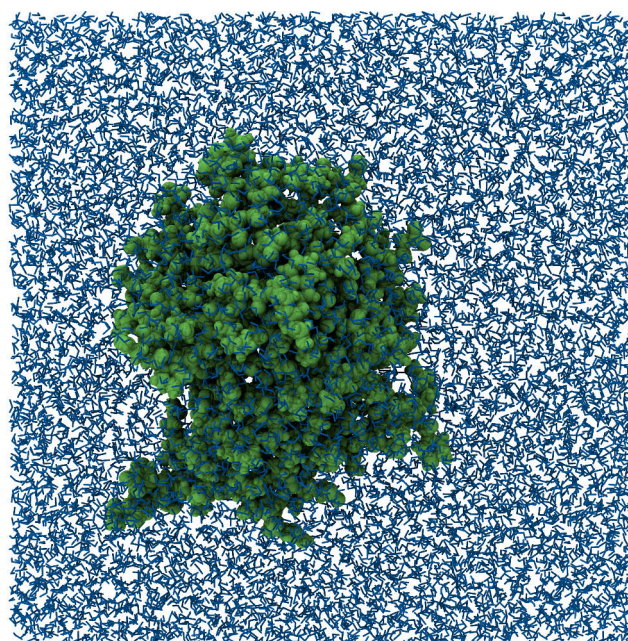


FIGURE 6.3 Depiction of a fully periodic simulation supercell of AnPixJg2 in water subject to FFMD and QM/MM AIMD simulations.

## 7 ANPIXJ AND CPHI: PROTEIN STRUCTURE AND DYNAMICS

### 7.1 AnPixJg2

#### 7.1.1 Evidence for Microheterogeneity

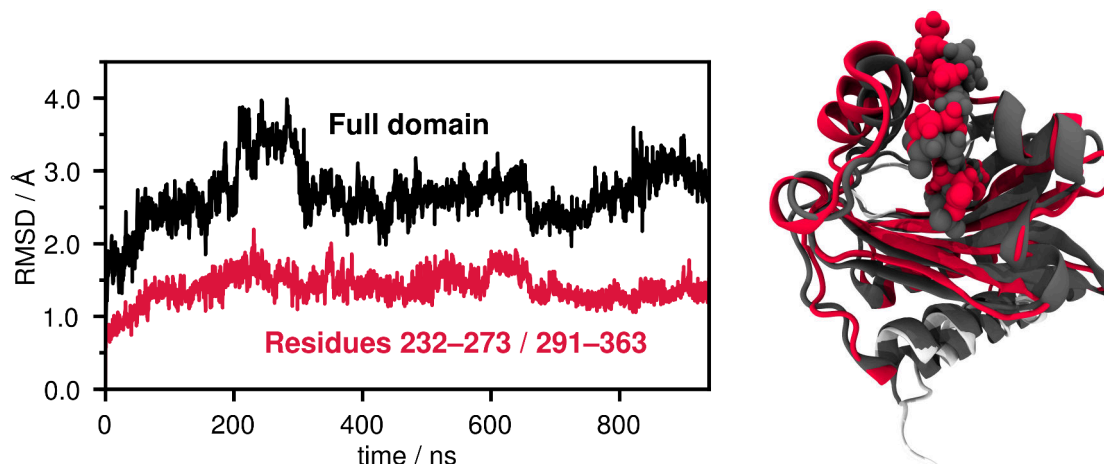


FIGURE 7.1 Left: RMSD of the AnPixJg2 backbone relative to the crystal structure from X-ray diffraction: Full domain (black) and without considering the N-terminal  $\alpha$ -helix and the unstructured loop region between residues 274 and 290 (red).<sup>2,183</sup> Right: Backbone after 1  $\mu$ s of FFMD simulation, aligned to the crystal structure (PCB shown as spheres).

Figure 7.1 shows the RMSD timeline of backbone atoms of AnPixJg2 relative to the starting structure from X-ray diffraction.<sup>4</sup> Excluding the N-terminal  $\alpha$ -helix and the unstructured loop region between residues 274 and 290 from the analysis reduces deviations drastically. The increased mobility of excluded regions is not surprising as they are especially exposed to solvent in the single GAF

<sup>4</sup>Parts of the FFMD study was performed by Laura Katharina Scarbath-Evers as a Master's project.<sup>2</sup>

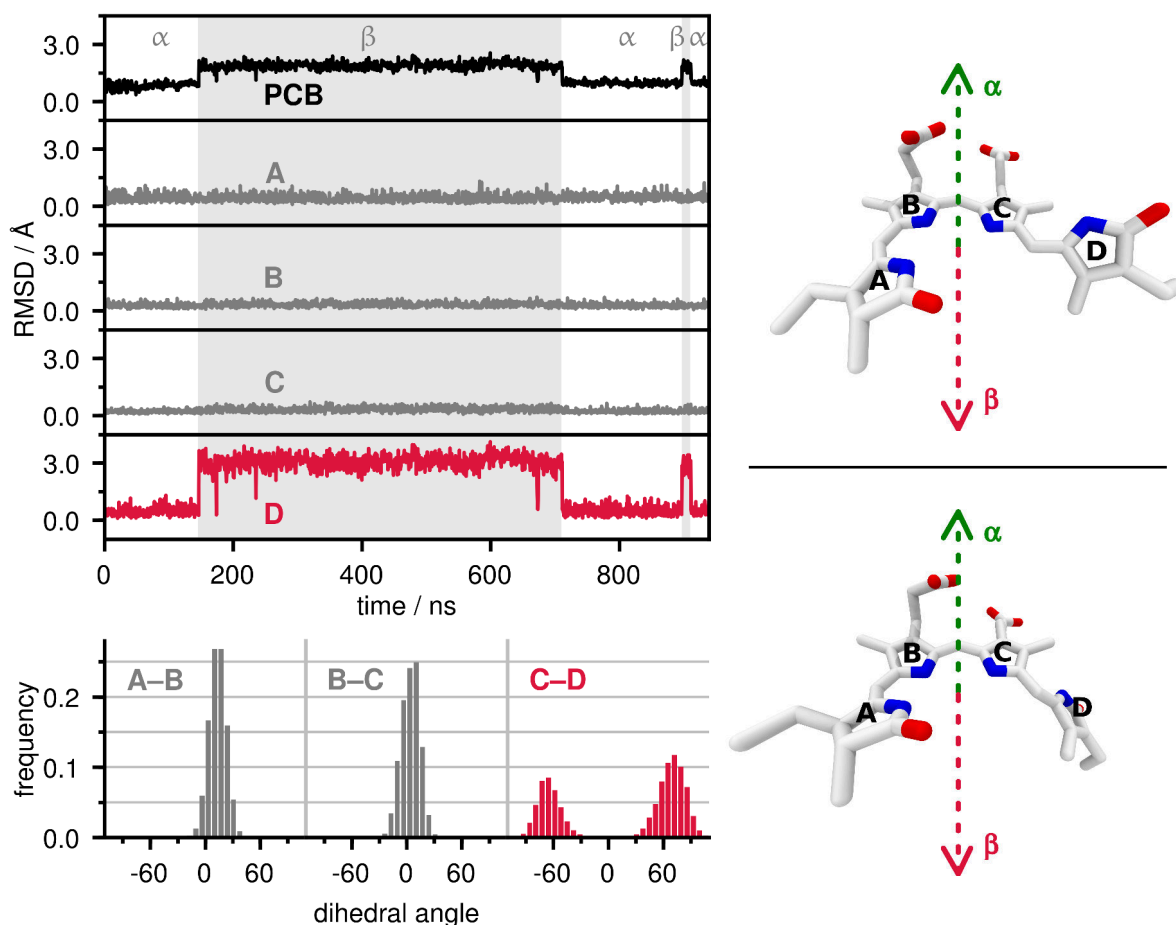


FIGURE 7.2 Top left: RMSD of PCB heavy atoms together with the individual deviations of pyrrole rings A, B, C, and D relative to the crystal structure from X-ray diffraction. The white and grey areas mark the presence of a  $\alpha$  or  $\beta$  orientation of ring D (RMSD shown in red). Bottom left: Dihedral distribution of the three methylene bridges connecting the pyrrole rings. Right:  $\alpha$  (top) and  $\beta$  (bottom) sub-state of PCB.<sup>2,223</sup>

domain, whereas in crystal the N-terminal  $\alpha$ -helix is bound to another, non-photoactive GAF domain (AnPixJg3).<sup>183</sup> Over the course of 1  $\mu$ s of FFMD simulation, the domain structure remains stable (Figure 7.1, right). The chromophore moves slightly out of the binding pocket together with adjacent  $\alpha$ -helix whereto the chromophore is linked. Also, the diameter of the entire domain, and so its volume, are slightly increased. It should be noted that in their crystal-derived solid-state structure, the protein residues are densely packed resulting in overexpression of hydrophobic protein-protein interactions while solvent effects are reduced.<sup>b</sup> It will be illustrated in the next section how immersing AnPixJg2 in a water supercell affects restructuring of the chromophore-binding pocket.

<sup>b</sup>This is an immediate consequence of the *salting out* process: the protein is “dehydrated”, otherwise it would not crystallise.

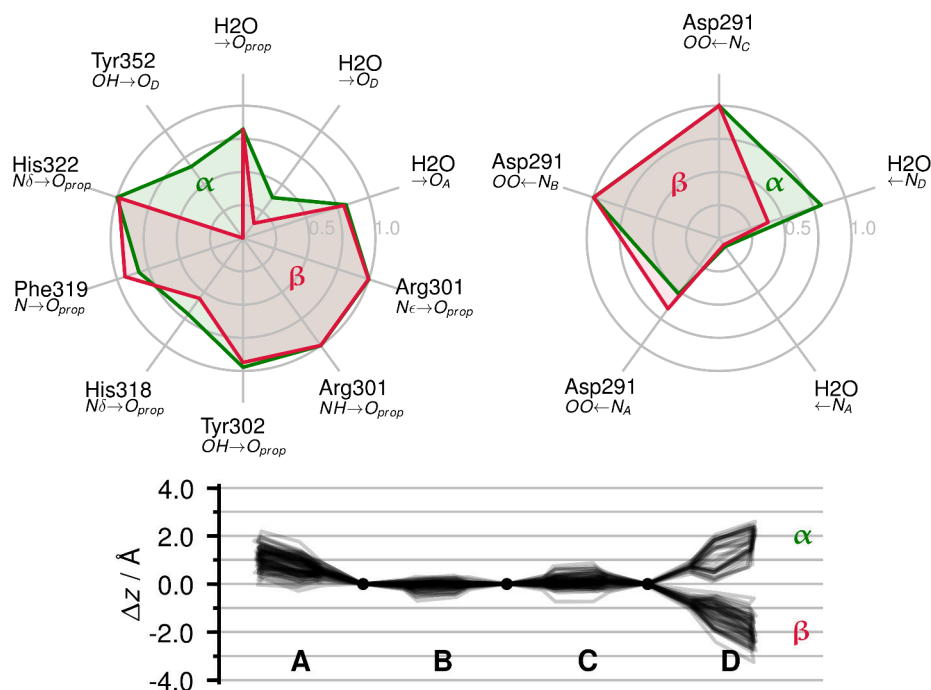


FIGURE 7.3 Top: Radar plots of the occurrence of hydrogen bond contacts between PCB and apoprotein or solvent, where PCB serves as acceptor (left) and donor (right). Bottom: Skeletal deviation plot<sup>155,240</sup> showing the z-displacements of PCB heavy atoms from the mean plane.

Turning to the cofactor PCB, a structural heterogeneity with the presence of two sub-states can be identified, which cannot be inferred from the crystal structure. Figure 7.2 illustrates the mobility of individual pyrrole rings. While rings A, B, and C exhibit low RMSD values over the course of the MD simulation, ring D shows characteristic jumps from below 1 to 3 Å and *vice versa*. These well-separated sub-states of PCB can be characterised on the basis of analysing dihedral angles between the pyrrole rings. Therein, both, A-B and B-C, dihedral distributions contain one sharp maximum indicating the three rings staying mainly co-planar with solely ring A being slightly rotated out of plane by 15 degrees. On the other hand, the dihedral angle between ring C and D shows a bimodal distribution with maxima at positive and negative angles of 70 degrees, respectively. Rockwell and co-workers suggested the two faces of PCB pointing either towards the conserved histidine or aspartate residue be labelled as α and β face.<sup>223</sup> Hence the P<sub>r</sub> state of AnPixJg2 possesses α and β sub-states *via* ring D.<sup>2</sup>

The conformational *heterogeneity* of ring D involves change to the interaction pattern with the apoprotein, which is shown in Figure 7.3 as hydrogen bond occurrences. Both sub-states are firmly anchored to the carboxylic group of Asp291, mainly by rings B and C, which is consistent with the

outer rings being rotated out of plane. Furthermore, they share a manifold of hydrogen bonds and salt bridges to the protein *via* the propionate side chains; especially links to Arg301, Tyr302, and the highly conserved His322, but also to the backbone (via Phe19) represent very strong connections that prevail throughout the simulation without having to compete with water. Nevertheless, the solvent too demonstrates its presence by solvating the charged side chains of PCB. Remarkably, water also binds strongly to the carbonyl oxygen atom of ring A, while in the crystal structure this site is occupied by the indole proton of Trp289. Yet, in the MD study no such hydrogen bond is registered. This already suggests that substantial changes in the chromophore binding pocket have occurred (see below section). An obvious difference between the two sub-states lies in the connection pattern regarding ring D: The  $\alpha$  sub-state features significant hydrogen bonding to Tyr352 *via* the carbonyl group of ring D, but shares it with a solvent molecule. In turn, the  $\beta$  sub-state misses these features, which is evident in the case of the protein link, though not in terms of solvation. Carbonyl group and NH group alike appear undersolvated, a deficiency either of choosing an inappropriate cutoff in hydrogen bond analysis or of the FFMD simulation itself; inspection of the MD trajectory reveals that the latter is true. Because of the  $\beta$  sub-state being absent in the crystal, proper solvation thereof requires a simulated time of about 500 ns (*i.e.*, half the simulation time). The partial absence of hydrogen bonding partners accounts for the low values in hydrogen bond occurrence.<sup>c</sup> Figure 7.3 further depicts the skeletal projection of PCB originally introduced by Kratky for hydroporphinoid ligands<sup>240</sup> (see also the review of Senge and co-workers<sup>155</sup>), which has slightly been adapted here for linear tetrapyrroles. This appealing abstraction intuitively illustrates the local features of conformational freedom of the cofactor; ring D clearly changes between  $\alpha$  and  $\beta$  conformation, yet ring A also appears mobile. It should not be ignored the mobility of ring D is higher in the  $\beta$  sub-state due to the missing hydrogen bond connection to Tyr352.

Conformational heterogeneity of bilin cofactors is not an unknown phenomenon as it was suggested by a number of spectroscopic investigations<sup>2,54,172,178,185,188,213,214,217,235,241–243</sup> of phytochromes and CBCRs. It bears the question as to how different sub-states affect photoconversion or dye colour. Quantum yields of fluorescence in particular can possibly be increased by reducing the extent of heterogeneity through point mutations or temperature.

---

<sup>c</sup>For further analysis, the fully solvated binding pocket has been considered.



### 7.1.2 Restructuring of the Binding Pocket

In the previous section, significant re-organisation of the chromophore's embedding has been signalled. Figure 7.4 presents three different snapshots of the binding pocket of AnPixJg2. The uppermost depiction shows the crystal structure having PCB as  $\alpha$  conformer. Therein one finds meticulously arranged protein-cofactor interactions and a blatant absence of water molecules. The important role of Trp289 in AnPixJg2 has been discussed since its discovery.<sup>178,195</sup> In the crystal structure it clearly forms a hydrogen bond to ring A of PCB and maintains a  $\pi$ - $\pi$  interaction with ring D. However, looking at the solvated  $\alpha$  sub-state (Figure 7.4, middle), residue and cofactor are dispersed with no interaction left at all. Instead, a water molecule has taken over the position at ring A and also occasionally interacts with the aspartate anchor. In fact, Trp289 is highly mobile showing regular ring flips. Therefrom, the “gatekeeper” role proposed by Escobar and co-workers seems very unlikely.<sup>2</sup> They suggested the bulky residue shields the binding pocket from solvent in the  $P_r$  state and opens up upon photoconversion to the  $P_g$  state (“hydration model” of solvatochromic colour tuning<sup>195</sup>). Other interactions remain widely intact, especially those engaging the highly conserved Asp291 and His322 residues, anchors to the pyrrole rings and the propionate side chain of ring B. In the  $\beta$  sub-state too (Figure 7.4, bottom) no changes affecting these interactions are observed, in line with Figure 7.3, but Trp289 now is completely unavailable for  $\pi$ - $\pi$  interaction with ring D. Furthermore the lost contact to Tyr352 can be seen, the latter now being hydrogen bonded to water, whereas ring D does not encounter a different protein residue on the  $\beta$  face but binds to solvent molecules (or faces void; see above discussion). Regarding Figure 7.4, the water content in the binding pocket has doubtlessly increased.

### 7.1.3 Hydrophobic Interaction and Solvation

Facing increased mobility of Trp289 in the simulation, the discussion about its putatively unique role in the binding pocket should be supported with a diligent examination of this peculiar supra-molecular interaction involving PCB and Trp289.

First, dissolution of the connection begs the question as to whether the employed force field in the MD study (CHARMM22) describes the non-covalent interactions properly. Table 7.1 summarises results of static calculations of the interaction energy between Trp289 and PCB, based on the crystal

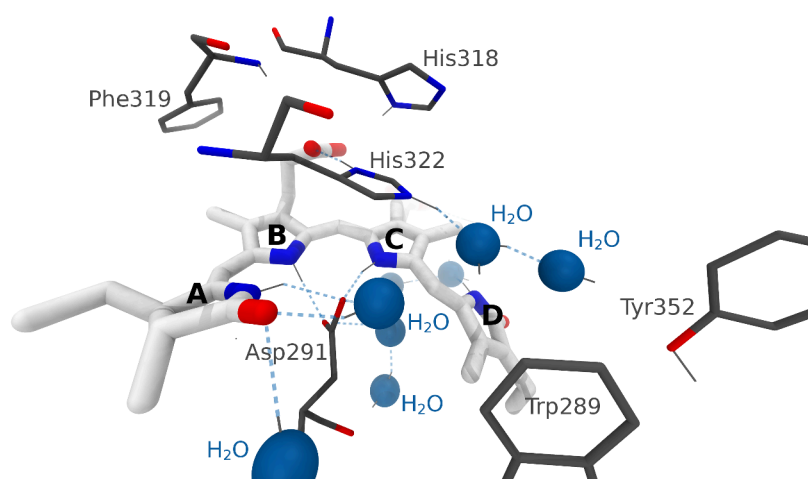
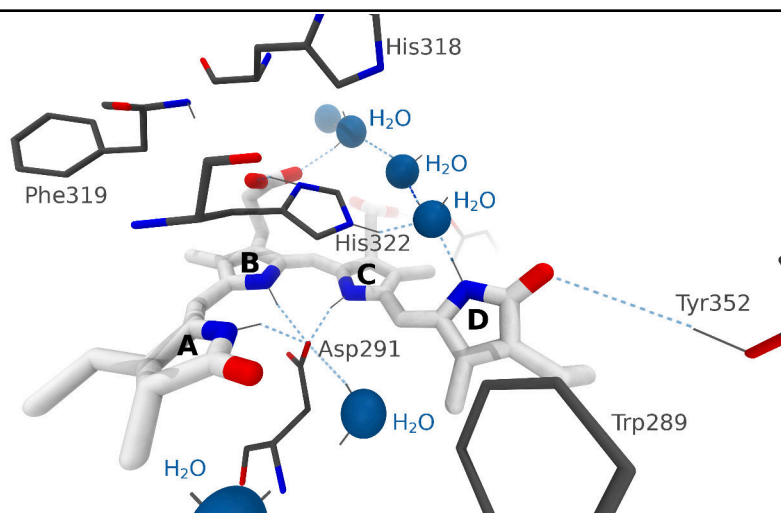
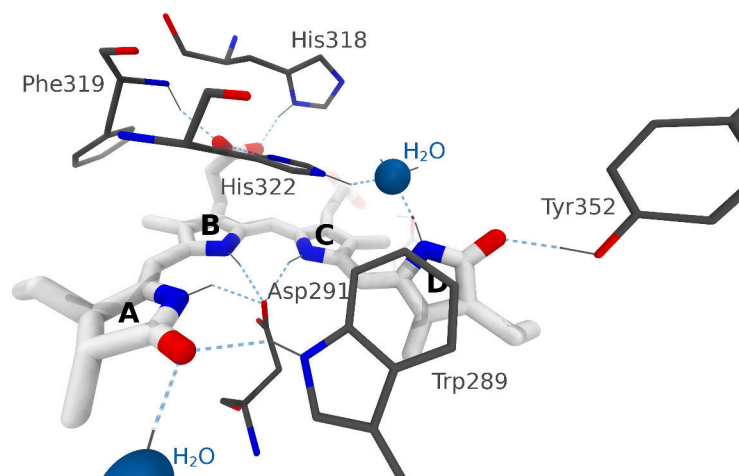


FIGURE 7.4 Chromophore-binding pocket of AnPixJg2 in the crystal (top) and after long-scale FFMD simulations, showcasing  $\alpha$  (middle) and  $\beta$  sub-states (bottom). Non-polar hydrogen atoms have been omitted for clarity.

TABLE 7.1 Interaction energies in  $\text{kJ mol}^{-1}$  of Trp289 and PCB, considering either the full chromophore, or ring D alone ( $\pi$ - $\pi$  interaction), calculated at different levels of theory. The quantum calculations were carried out with the def2-TZVP basis set (without correction of BSSE).<sup>244</sup> D<sub>3</sub> denotes dispersion correction by Grimme and co-workers.<sup>245</sup>

Calculation Method	Trp289-PCB <sub>A-D</sub>	Trp289-PCB <sub>D</sub>
BLYP	-5.32	+16.77
RHF	-3.51	+15.83
B <sub>3</sub> LYP	-12.16	+11.03
BLYP + D <sub>3</sub>	-64.57	-30.90
B <sub>3</sub> LYP + D <sub>3</sub>	-62.78	-29.53
MP <sub>2</sub>	-76.45	-41.72
CHARMM22	-51.61	-32.94

structure.<sup>d</sup> Therein, the protein residue is considered to interact either with the full chromophore or ring D only. CHARMM22 is compared to quantum mechanical calculations at different levels of theory. Expectedly for non-covalent interactions, plain DFT methods and HF substantially differ from calculations including dispersion interaction, either as empirical correction (D<sub>3</sub><sup>245</sup>) or by means of post Hartree Fock methods (MP<sub>2</sub>). Considering  $\pi$ - $\pi$  interaction with ring D alone even leads to repulsive terms that are reversed by dispersion energy. CHARMM22 describes the interaction very well for both models. This becomes even more true when considering, firstly, the overbinding effect of MP<sub>2</sub> calculations<sup>246</sup> and, more importantly, the correction of the basis set superposition error (BSSE) being omitted in this study for simplicity. The BSSE strongly affects hydrogen bonds (*i.e.*, the interaction involving the full chromophore), leading to an overestimation of interaction energies in all quantum calculations.<sup>247</sup>

Another study analyses the hypothesis that within the simulations, solvation and equilibration of the crystal structure, though being performed very gently (for computational details see Appendix A), leads to the artificial separation of Trp289 and PCB. Figure 7.5 shows the results of an adapted FFMD run where the hydrogen bond involving the carbonyl group of ring A has been constrained with a strong harmonic potential.<sup>e</sup> By this, the interaction is enforced while the chromophore-binding pocket is solvated over the course of more than 250 ns of MD simulation. The link remains intact, of course, as long as the constraint is enabled. However, the release of Trp289 (transition from grey to white in Figure 7.5) leads to immediate separation from PCB. The protein residue notably backs away from ring A in a two step process (the transition from 6 to 9 Å marks a ring flip of Trp289). Thus, the crystalline interaction is not stable in solution even if the protein solvation is carried out

<sup>d</sup>The quantum calculations of this analysis have been carried out by Michael Türk in a student project.

<sup>e</sup>Also referred to as restrain.

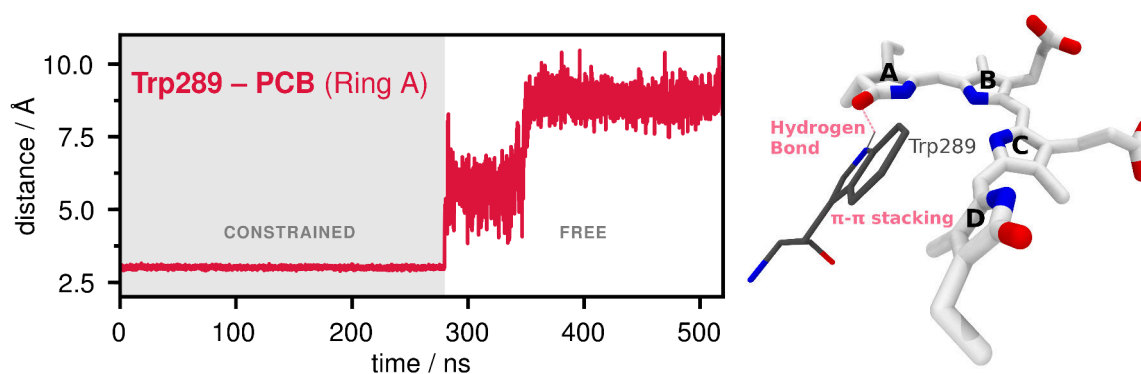


FIGURE 7.5 Left: Evolution of the Trp289–ring A hydrogen bond in a MD simulation under constraint (grey area) and after release (white area). Right: Supramolecular situation of Trp289 and PCB in the crystal structure: hydrogen bond to ring A and  $\pi$ - $\pi$  interaction with ring D.

separately from constraint release. Once Trp289 and PCB have separated, it is thermodynamically unlikely that they will find back to each other to re-form their connection. Although  $\pi$ - $\pi$  interactions can attain significant influence in a biochemical system, in the present case, other factors dominate. The enthalpic contribution of this bond with about  $30 \text{ kJ mol}^{-1}$  (*cf.* Table 7.1) is too low to keep the residue in place given the effects of entropy and ring puckering. This becomes more important as there is strong competition of water that also attains attractive dispersion interactions with  $\pi$  systems. The exclusion of small, flexible solvent molecules, however, is entropically unfavoured.<sup>f</sup> The hydrogen bond between PCB and Trp289 can easily be replaced by water as well, possibly resulting in an even more favourable energy balance. But even if formation of a PCB-Trp289 assembly were exergonic, there is no attractive force reaching deep into space (such as in electrostatic interactions). As a member of London dispersion forces,  $\pi$ - $\pi$  attraction is bound to a decrease with separation like  $r^{-6}$ .<sup>g</sup> As soon as the residues are separated, they do not “see” each other anymore, which is reflected in MD simulations.

The increased water content in the binding pocket has been identified in the previous section. Figure 7.6 presents detailed views of water chains that the two sub-states form over the course of simulation. The intrusion of water takes place from the rear side of the chromophore (*i.e.*, the side of the propionate chains) – another contradiction to the “gatekeeper” function of Trp289<sup>195</sup> – establishing delicate water chains above and below the cofactor. Intriguingly, there is a preferred position where a *persistent water molecule* is located at any time. Although being related to pyrrole water molecules

<sup>f</sup> See also Section 5.3.

<sup>g</sup>For instance, a (temporary) shift in distance from 4.5 to 5 Å would already halve the interaction energy.

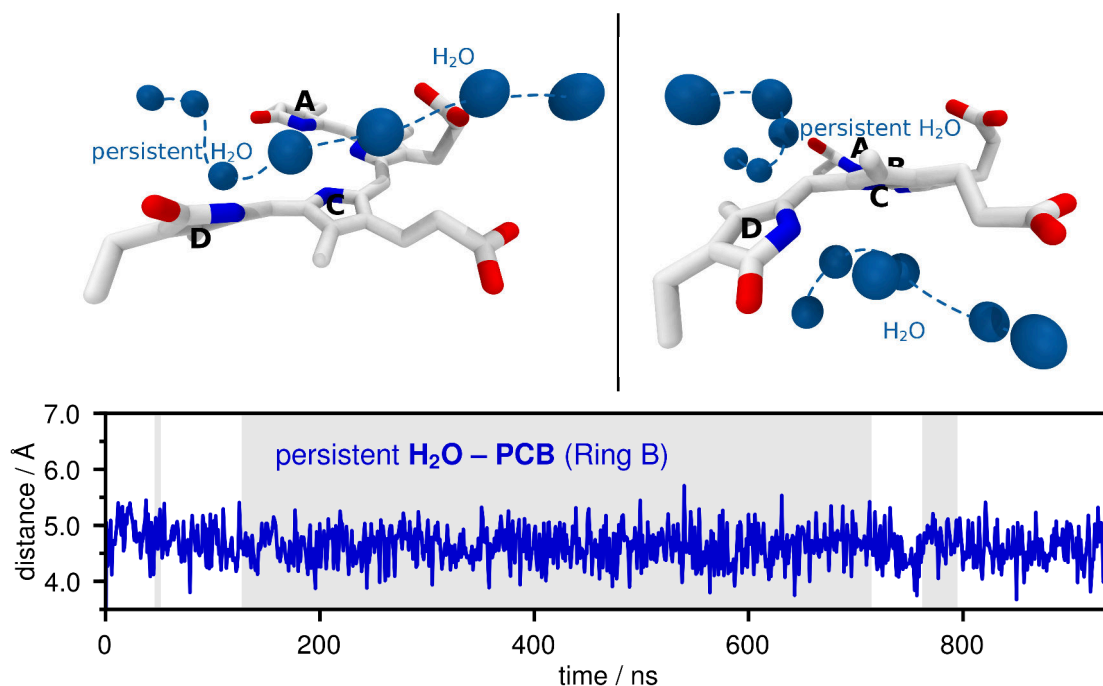


FIGURE 7.6 Top: Formation of water chains and the establishment of a persistent water molecule in  $\alpha$  (left) and  $\beta$  (right) sub-state, respectively. Bottom: Distance timeline of the persistent water molecule to ring B of PCB. The change in background shade marks events when it is substituted by a new water molecule.

(*e.g.*, in CphI $\Delta$ 2), its distance to the pyrrole rings A, B, and C is increased (about 5 Å; *cf.* Figure 7.6, bottom). Yet, this does not impede the solvent molecule from sitting stable in front of ring A. Substitution of the persistent molecule by a new molecule happens at times (shaded areas in Figure 7.6, bottom), but the site, initially occupied by Trp289, is never vacant. This interesting result paints a picture quite disparate to what has been discussed on the basis of the crystal structure of AnPixJg<sub>2</sub>. At the present point, it is too early to speculate on the role of these water chains and the persistent water, but more fascinating is the fact that solvation of the binding pocket does not stop the photoconversion mechanism to operate.<sup>231</sup> Nevertheless, it is unclear whether and how solvation affects the signalling mechanism of AnPixJ as a whole. If all non-covalent apoprotein-cofactor interactions were replaced by water, it would be difficult for the chromophore to communicate its configuration to the protein. As has been discussed before, most likely, the propionate side chains play a key role in signal transmission as they represent the only parts of the apoprotein forming strong connections to the cofactor, even in a solvating environment (*cf.* Figure 7.3).

The FFMD study of AnPixJg<sub>2</sub> hints at the imperative to let go of the perception that hydrophobic protein-protein interactions prevail under solution conditions. Moreover, it becomes clear that the

transition from crystal structure to a fully solvated model requires substantial amounts of simulation time. It pays off by providing insight into state heterogeneity and its implications for non-covalent interactions. The hydration model appears rather unrealistic a scenario of colour tuning as in the simulation unidirectional solvation of the chromophore-binding pocket happens already in the  $P_r$  state.<sup>2</sup> In view of established water chains and persistent solvent positions, it is hard to imagine this process being reversible.<sup>195</sup> In contrast, solvation of the binding pocket is an exergonic process, its reversion endergonic. The role in colour tuning taken by Tyr<sub>352</sub> and residues bound to the propionates can only be surmised as it remains obscure how the interaction pattern would look like in the  $P_g$  state.

## 7.2 CphI $\Delta_2$

The  $P_r$  state of AnPixJg<sub>2</sub> resembles much that of CphI $\Delta_2$  (comparing GAF domains only).<sup>183</sup> However, PCB inserts reversely with ring D pointing out of the domain. The multidomain nature of CphI $\Delta_2$  is characterised by an important feature, common to all phytochromes:<sup>169</sup> a tongue-like protrusion built of highly conserved residues that intimately connects to the GAF domain and shields the chromophore binding pocket from the solvent (*cf.* Figure 6.1). There can hardly be any doubt that this extension plays a key role in signal transmission, even more so as it forms a salt bridge to Asp<sub>207</sub> (anchor to pyrrole rings of PCB) is readily detached in the  $P_{fr}$  state.<sup>160,187,248</sup>

Within this study CphI $\Delta_2$  functions as reference protein, whose structure and signalling mechanism is well-studied. In Figure 7.7, the RMSD timeline from the FFMD trajectory is shown. Due to the multidomain nature of the phytochrome, deviation from crystal structure appears very large; over the course of the MD simulation relative orientation of the domains changes slightly. More important, however, is the integrity of individual domains, in particular of the GAF domain, which is confirmed by Figure 7.7.

In contrast to AnPixJg<sub>2</sub>, CphI $\Delta_2$  does not exhibit  $\alpha/\beta$  heterogeneity in its  $P_r$  state, but only appears in the  $\alpha$  sub-state (Figure 7.8). Nevertheless, NMR and Raman studies confirmed the existence of different protonation states of adjacent histidine residues, mainly His<sub>260</sub> and His<sub>290</sub>.<sup>188,214</sup> This work is based on the variant, where His<sub>260</sub> and His<sub>290</sub> are singly protonated at their N $\epsilon$  and N $\delta$  position, respectively, labelled in the literature as “Pr-II”.<sup>188</sup> Figure 7.7 presents the hydrogen bond contacts of PCB in this sub-state. As the chromophore-binding pocket is protected from solvent by

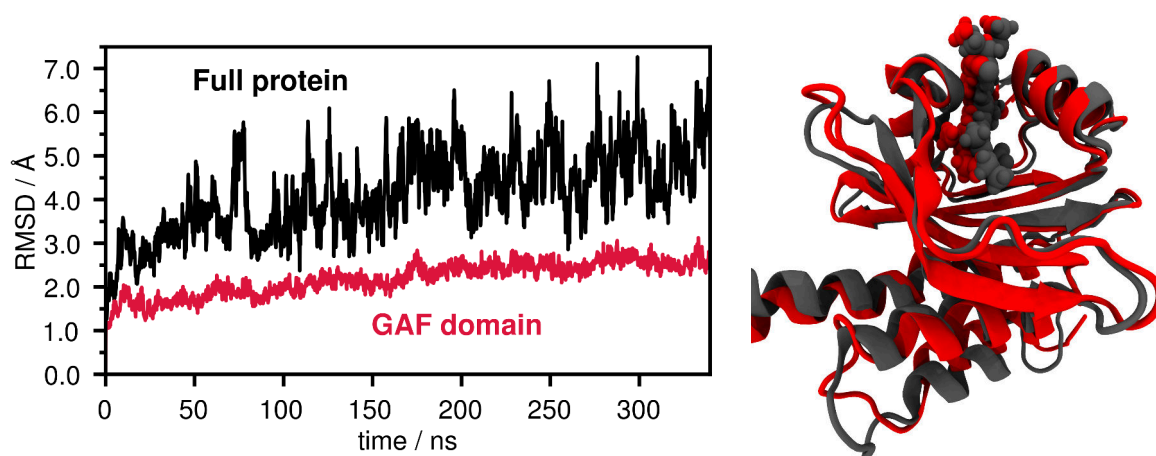


FIGURE 7.7 Left: RMSD of the Cph1Δ<sub>2</sub> backbone relative to the crystal structure from X-ray diffraction: Full protein (black) and only considering the GAF domain (red).<sup>169</sup> Right: Backbone after 400 ns of FFMD simulation, aligned to the crystal structure (PCB shown as spheres).

the PHY domain protrusion, PCB is well embedded showing strong hydrogen bonds to its neighbouring protein residues. Interestingly, the rings are less strongly connected to the central aspartate, except for ring C. Instead, a persistent water molecule (“pyrrole water”), that could also be found in the crystal, takes over the role of hydrogen bonding. Another cause of loose contact is Asp207 providing only one binding site (the backbone oxygen atom), but in AnPixJg<sub>2</sub> both oxygen atoms of the negatively charged carboxylate of Asp291 hold together the pyrrole rings.

Figure 7.9 portrays the binding pocket of PCB in Cph1Δ<sub>2</sub>. It is worthwhile comparing it with AnPixJg<sub>2</sub> (Figure 7.4) and taking a look at the highly conserved feature cysteine-aspartate-histidine (cysteine is not shown in the snapshots). An important feature of the reversed aspartate orientation is the formation of a salt bridge to the PHY domain. This connection is likely to be detached upon photoconversion to the P<sub>fr</sub> state when the nitrogen atom of ring D connects to aspartate instead, which enables drawing conclusions regarding the way of signal transfer from PCB to the apoprotein.<sup>186,220,221</sup>

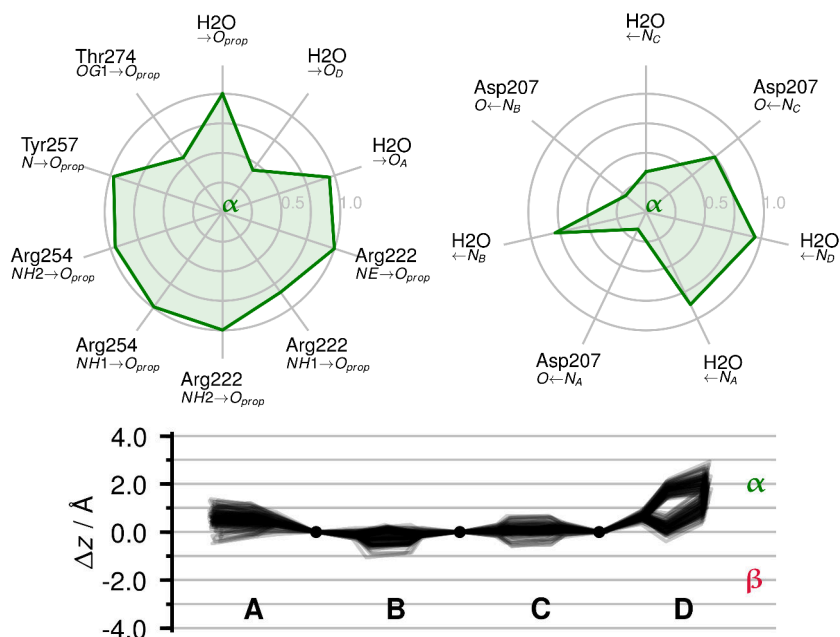


FIGURE 7.8 Top: Radar plots of the occurrence of hydrogen bond contacts between PCB and apoprotein or solvent, where PCB serves as acceptor (left) and donor (right). Bottom: Skeletal deviation plot<sup>155,240</sup> showing the z-displacements of PCB heavy atoms from the mean plane.

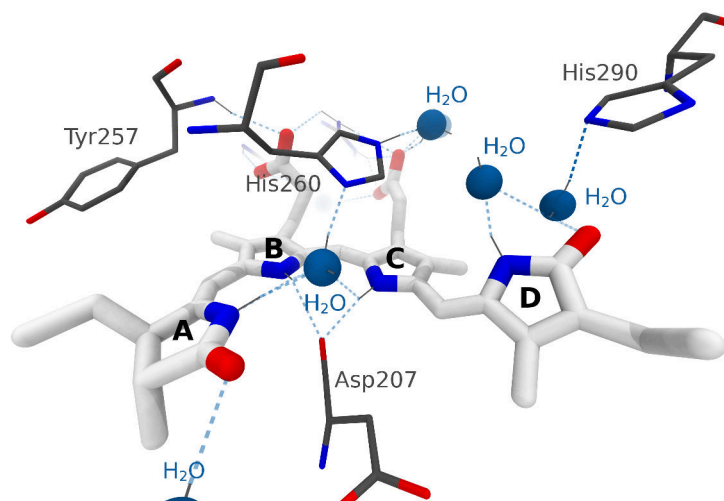


FIGURE 7.9 Chromophore-binding pocket of Cphi $\Delta$ 2 (“Pr-II”<sup>188</sup>) after 400 ns of FFMD simulation. Non-polar hydrogen atoms have been omitted for clarity. The PHY domain (not shown) forms a salt bridge with the carboxylate group of Asp207 sticking out at the lower edge of the image.



## 8 PCB: SPECTROSCOPIC CHARACTERISATION

### 8.1 NMR Spectroscopy

#### 8.1.1 Sampling Statistical Ensembles

Working with MD trajectories requires choosing a statistical ensemble and the level of theory, but also deciding on how to sample representative snapshots from the ensembles for subsequent calculations of linear response (*i.e.*, NMR shieldings and absorption spectra).<sup>18</sup> In the present study, apart from choosing either canonical (NVT) or isothermal-isobaric (NPT) conditions, different levels of theory lead to disparate timescales spanned by the trajectories (see Section 2.3). While FFMD simulations range in orders of magnitude of nanoseconds, AIMD simulations typically span not more than 30 picoseconds. At the same time, the *ab initio* linear response calculations have to base on *ab initio* geometries, ruling out direct structure sampling from FFMD trajectories.<sup>a</sup> Moreover, spectroscopically important degrees of freedom, such as molecular vibrations, may be poorly described by force fields, yet they reach ergodicity already within a few dozens of picoseconds. With regard to structural events in proteins it would nevertheless be desirable to span timescales of nanoseconds, which reach ergodicity thereof, rather than sampling short periods of picoseconds. To take account of both timescales, two different sampling protocols are employed in this work to acquire spectra from trajectories (Figure 8.1).

*Protocol 1* (P1) directly samples the isothermal-isobaric trajectory from FFMD simulations, but augments structures of MM quality by short AIMD runs in the QM/MM scheme, finished with a simulated annealing<sup>35</sup> to relax deflections of bonds, angles, and dihedrals without removing characteristic structural motives. This protocol takes advantage of long timescales within the FFMD run without

---

<sup>a</sup>Moreover, FFMD simulations are usually performed using the SHAKE algorithm, which constrains all bonds involving hydrogen.

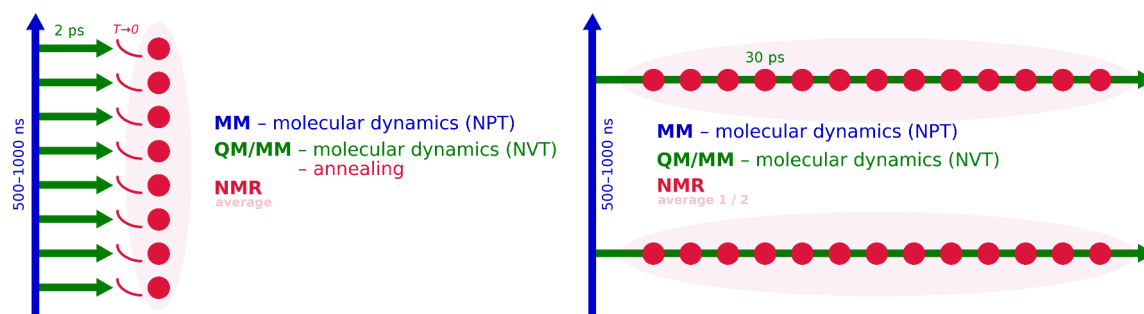


FIGURE 8.1 Protocols of spectrum acquisition from MD simulations. Left: Direct sampling of the NPT ensemble, augmented by short canonical AIMD equilibration runs and annealing (P1). Right: Sampling along canonical AIMD trajectories having started from FFMD snapshots (P2).

loosing accuracy in calculated spectroscopic signatures (see also Figure 8.2, right); yet it purges thermal fluctuations and artificially dips entropy.

*Protocol 2* (P2) resembles the “conventional” sampling scheme that is also used for computation of VCD spectra in Part I. Snapshots subject to linear response calculations are thereafter sampled from the AIMD trajectory without post-processing, thus including fully thermal fluctuations and entropy effects. While the full information picture gives P2 an important relevance, it must rely on merely a few long AIMD runs due to their respective computational overhead: In this work, the FFMD trajectory has been forked for P2 only twice, while P1 is based on up to 40 FFMD snapshots. On the other hand, P2 retrieves about 100 or more images per AIMD trajectory at regular intervals that serve for spectroscopic detail.

### 8.1.2 Size of the QM Part and Convergence

For QM/MM calculations, the size of the QM part has to be chosen carefully depending on the type of spectroscopic property (see Section 2.3). Figure 8.2 shows the relative (normed) root mean square error (RMSE) of calculated NMR shieldings of chromophore atoms. In the left graph, the RMSE plotted against QM part size points to specific residues of importance. Naturally, all (charged) residues involved in hydrogen binding to the chromophore need to be considered in the QM part, including water molecules. Apart from this, no meaningful influence on NMR shieldings can be found.<sup>b</sup> Notably, <sup>13</sup>C shieldings are barely affected by the environment, which is certainly related to their mainly quaternary nature. The effect of annealing has already been mentioned. The right graph

<sup>b</sup>Eventually, a QM part of between 260 and 280 atoms has been used in this work, depending on the included number water molecules.

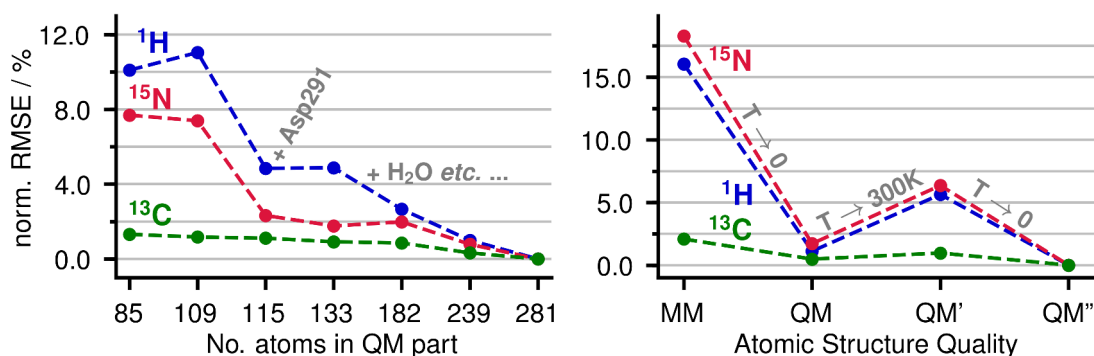


FIGURE 8.2 Normed RMSE of calculated shieldings with increasing size of the QM part (left) and after annealing and equilibration of FFMD snapshots (right) (P<sub>2</sub>).

in Figure 8.2 underlines the importance of using QM rather than MM based molecular structures; the relative error of calculated shieldings drops from over 15 % to less than 2 % after direct annealing (*i.e.*, without further equilibration) of the FFMD snapshot. Interestingly, thermal fluctuations let the error grow again above 5 %. Performing final annealing therefore is a prerequisite in order to obtain reasonable precision when averages are calculated *via* P<sub>1</sub>, where only few linear response calculations enter.

In this regard, P<sub>2</sub> differs due to frequent sampling of the AIMD trajectory. Figure 8.3 illustrates the relative RMSE depending on the number of snapshots used for average calculation. The sensitivity of calculated shieldings towards thermal fluctuations and supramolecular interactions is expressed by the spread, also shown in the figure. The atom types fluctuate differently with <sup>1</sup>H being most susceptible. This corresponds to their chemical exposure to the environment together with the lack of core electrons.

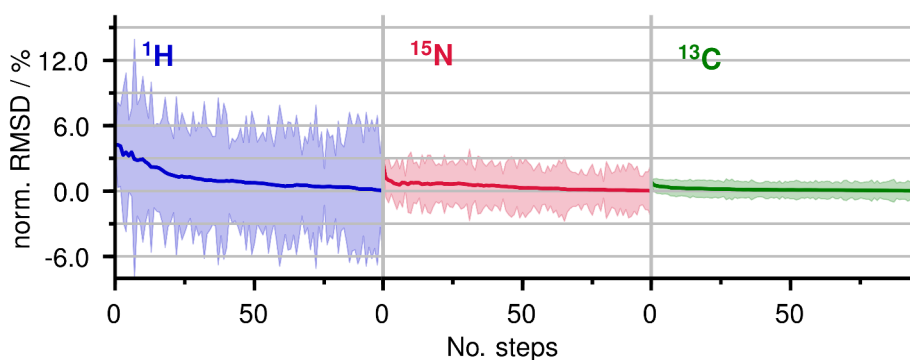


FIGURE 8.3 Normalised RMSE of calculated shieldings by means of AIMD sampling plotted against the number of snapshots, indicating convergence of the linear response calculation. The filled area marks the sensitivity of calculated shieldings towards thermal fluctuations and supramolecular interaction as RMSE.

## 8.1.3 Charge and Conjugation Pattern of PCB

Charge localisation and conjugation pattern of PCB in the protein environment has always been a matter of broad discussions.<sup>197,215</sup> However, elevated temperatures and the effect of ring pucker- ing were seldomly considered. Information on conjugation can straightforwardly be extracted from AIMD trajectories by looking at the average bond distances as shown in Figure 8.4. It clearly can be differentiated between, on the one side, isolated single ( $>150$  pm) and double ( $<140$  pm) bonds in rings A and D, and the conjugated system of rings B and C on the other. Still, the latter shows a subtle Peierls instability, wherefrom a bond alternation pattern could be crafted (Figure 8.4, right). Nevertheless, electronic delocalisation between rings B and C appears very clearly.

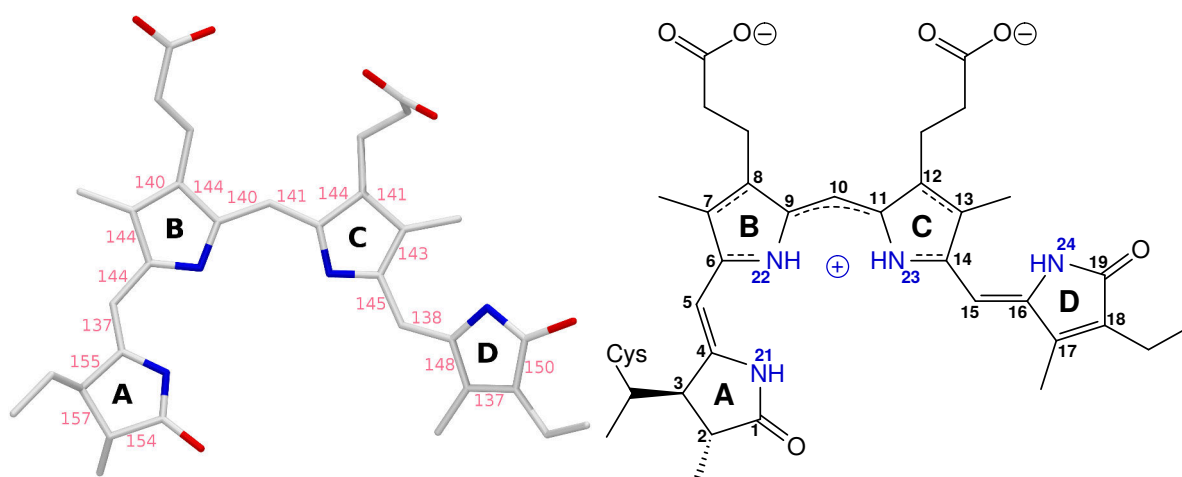


FIGURE 8.4 Left: Average C–C bond lengths in pm, extracted from the AIMD trajectory of PCB in AnPixJg2 (the standard error is 0.4 pm). Right: Proposed lewis representation of the conjugation pattern in PCB, based on the AIMD analysis (*cf.* Figure 6.2).

As of yet, localising the positive charge at either ring B or ring C were among the favoured considerations. Recently, NMR experiments came to the conclusion that in Cph1Δ<sub>2</sub>, ring B hosts the charge, because it exhibits a <sup>15</sup>N upshift.<sup>215</sup> Yet, regarding the delocalisation level pictured above, it seems unlikely a “localised” charge can be traced to any atom of the chromophore. Table 8.1 lists the average partial charges from Mulliken population analysis,<sup>249</sup> obtained as averages over AIMD trajectories. The results support the hypothesis of a delocalised positive charge, even though ring C gets assigned a number slightly increased with respect to ring B (0.45 vs. 0.24), while ring D does not carry a positive charge.<sup>c</sup> Hence, the symmetry of rings B and C is reflected also in the charge pattern, which implies the differences in <sup>15</sup>N chemical shifts can hardly serve as indicator thereof. Interestingly, structural

<sup>c</sup>But it does so in the P<sub>fr</sub> state according to the proposed photoconversion mechanism.<sup>185</sup>

TABLE 8.1 Averaged Mulliken charges<sup>249</sup> in *e* of PCB substructures, calculated along the AIMD trajectory of AnPixJg2.

Moiety	Partial Charge
Ring A + Me <sub>AB</sub>	+0.24
Ring B/C + Me <sub>BC</sub>	+0.58
Ring D + Me <sub>CD</sub>	-0.03
Ring B	+0.24
Ring C	+0.45
Me <sub>BC</sub>	-0.12
Me <sub>sum</sub>	-0.55

symmetry is indeed reflected in measured <sup>13</sup>C chemical shifts, which will be discussed in the following section.

#### 8.1.4 Information Conveyed by <sup>13</sup>C Chemical Shifts

It has been reported that <sup>13</sup>C chemical shifts of PCB in CphIΔ2 showcase a remarkable symmetry within the conjugated system between atoms 6 and 14 (*cf.* Figure 8.4).<sup>188,197,215,226</sup> In AnPixJg2 the pattern looks very much alike, but not as ideal as in the phytochrome. Figure 8.5 collects experimental and calculated <sup>13</sup>C chemical shifts of PCB in both proteins. In addition to symmetry between C6 and C14, a characteristic alternation pattern, spanning chemical shifts between 110 and 150 ppm, stands out. This underlines that the carbon atoms, though being part of a conjugated system, are chemically inequivalent. In fact, the electronic system is structured by the presence of electronegative atoms in the rings. The two proteins' similarity regarding <sup>13</sup>C chemical shifts is remarkable. Comparing with calculated chemical shifts, it becomes apparent that precision limits of the prediction do not fulfil requirements to account for subtle effects on carbon shifts. After all, the assignment of carbon atoms can be achieved by calculations but a decision of whether a measured spectrum belongs to AnPixJg2 or CphIΔ2 is unreachable on the basis of calculated ensemble averages. Yet, one has to keep in mind that the calculated chemical shifts correspond to the underlying structural variables; for instance, overestimation of bond lengths may result in deviation of calculated shieldings, albeit the linear response calculation itself were accurate. Consequently, deterioration of precision does not lie in the nuclear shieldings but in the structural models of FFMD and AIMD alike. This underlines that looking at absolute values does not grasp all information lying within the theoretical model. By means of statistical analysis of variable distributions, the connection between structural parameters and calculated chemical shifts can be looked at. An existing correlation would be more robust against

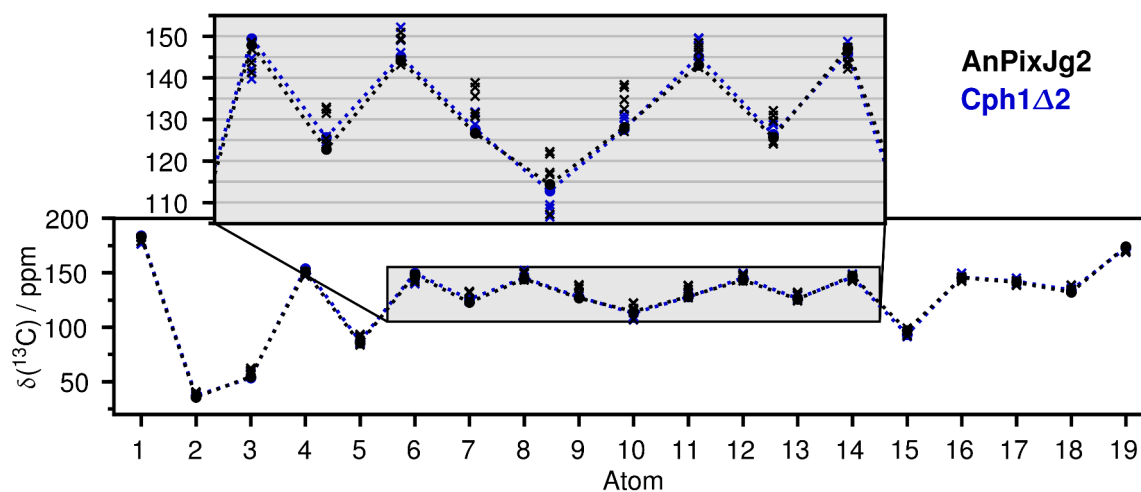


FIGURE 8.5  $^{13}\text{C}$  chemical shifts of PCB in AnPixJg2 (black) and CphI $\Delta$ 2 (blue).<sup>197</sup> The crosses mark calculated values from MD trajectories after protocols P<sub>1</sub> and P<sub>2</sub> alike.

over- or underestimation of variables and make a comparison to the experiment reliable.

Recently, Song and co-workers reported a peculiar difference pattern of chemical shifts upon photoconversion in AnPixJg2 and CphI $\Delta$ 2.<sup>197</sup> Therein, the carbon atoms belonging to rings C and D exhibit shift changes by more than 10 ppm when transforming from the P<sub>r</sub> state into the respective photoproduct. The authors attribute this observation to a “large-scale rearrangement of the critical interactions [that] cannot be explained only by the configurational changes of the chromophore.”<sup>197</sup> Taken literally, they suggest removal of hydrogen bonds, emersion of a new charged side chain or any other drastic modification to the chromophore. However as shown in Figure 8.2 (left),  $^{13}\text{C}$  chemical shifts undergo only small changes upon inclusion of protein environment into the QM part. On the other hand, Figure 8.3 reveals that thermal fluctuations alone result in the  $^{13}\text{C}$  shifts varying by more than 0.5 % (*i.e.*, 5–9 ppm). Also, chemically very similar carbon atoms (*e.g.*, atoms 6, 9, 11, and 14) differ by more than 20 ppm.

The experiment clearly shows the symmetric pattern of chemical shifts between rings B and C being perturbed. In analogy to the afore-mentioned inferring of structural symmetry from NMR patterns, the observed variation of chemical shifts indicates structural changes that let rings B and C become unequal. At the same time, the overall shape of the pattern, such as the alternation of chemical shifts, remains intact. This limits the degree of allowed chemical changes to the chromophore (*e.g.*, ruling out deconjugation).

Another observation of Song and co-workers is the opposite trend of the difference pattern going from the P<sub>r</sub> states to either P<sub>fr</sub> or P<sub>g</sub>. This reversibility needs to be reflected in the causing process,

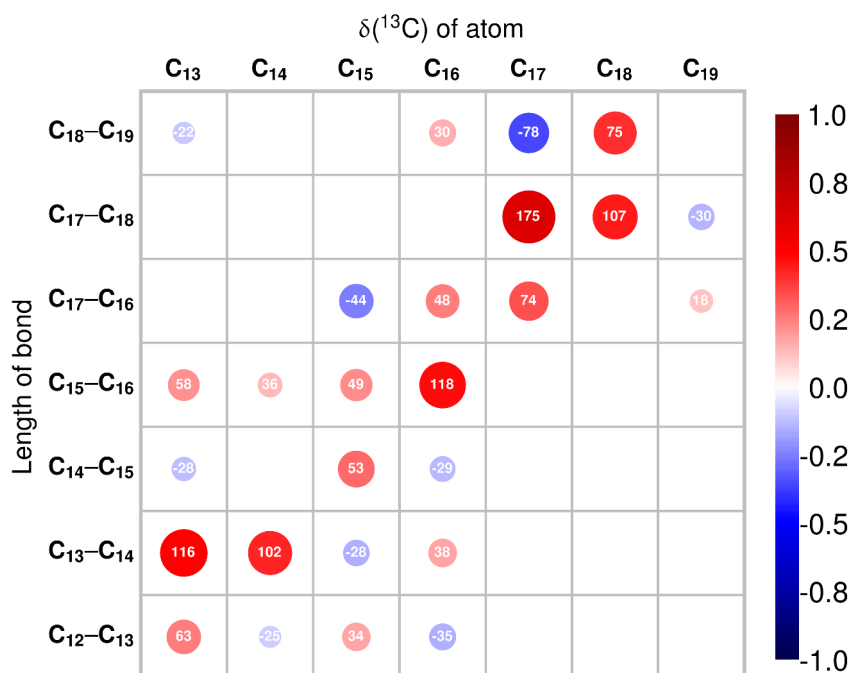


FIGURE 8.6 Correlation matrix  $\mathbf{M}$  (see Equation 8.1) representing the sensitivity of  $^{13}\text{C}$  chemical shifts against C–C bond lengths in rings C and D. Colour and size of the circles correspond to the Pearson correlation coefficient (see colour bar on the right), while the white numbers denote the slope of linear regression (in  $\text{ppm } \text{\AA}^{-1}$ ). Excluded from view are entries with significance level lower than 95 % (p-value).

which suggests even more that the latter can be found only in the cofactor itself. It would therefore be insightful to draw the connection between chemical shifts and the distribution of structural parameters.

Figure 8.6 presents the correlation matrix  $\mathbf{M}$  of  $^{13}\text{C}$  shifts with C–C bond lengths in rings C and D, based on the assumption of linear behaviour following the equation,

$$\Delta\delta_i^{\text{C}} \propto M_{ij} \cdot \Delta l_j, \quad (8.1)$$

with  $\Delta\delta_i^{\text{C}}$  being the change in  $^{13}\text{C}$  chemical shift,  $\Delta l_j$  the change in bond length, and  $M_{ij}$  denoting the matrix elements of  $\mathbf{M}$ . It contains various significant correlations proving the statistical connection of the two dimensions. Intriguingly, Figure 8.6 shows several entries that depend on one variable only while one would expect multivariate behaviour, because of the multivalent nature of the carbon atoms. For instance, atoms 13 and 14 alike correlate to the distance between them but not or only weakly to the bonds connecting them with atoms 12 or 15. Atom 16 shows a similar behaviour. Especially interesting are entries of atoms correlating with bonds that do not include themselves,

TABLE 8.2 Hypothetical changes in C–C bond lengths based on measured difference pattern of  $^{13}\text{C}$  chemical shifts (in ppm) going from the  $P_{fr}$  state of  $\text{CphI}\Delta_2$  to the  $P_g$  state of  $\text{AnPixJg}_2$ .<sup>197</sup> The trends denoted by +/- are inferred by inverse application of Equation 8.1 to Figure 8.6.

Atom	$\Delta\delta^{\text{C}}(P_g-P_{fr})^{197}$	C13–C14	C14–C15	C15–C16	C16–C17	C17–C18	C18–C19
C13	-11.1	-		-			
C14	-12.3	-					
C15	+7.7		(+)	(+)	(-)		
C16	-14.6			-	(-)		
C17	+10.4				(+)	+	(-)
C18	+0.8					+*	-*
	$\Delta l_{\text{bond}}$	-	(+)	-	(-)	+	-

\*Small difference for C18 implies bond lengths C17–C18 and C18–C19 to change reciprocally.

as seen for entries C15/C16–C17 and C17/C18–C19. It can be inferred from the asymmetric correlation pattern, that **M** reflects the electronic connection of atoms, governed by conjugation and electron-withdrawing atoms and substituents.

**M** can be used to interpret the carbon chemical shift differences described by Song and co-workers. By inversion of Equation 8.1, implications for bond lengths can be retrieved. Table 8.2 gives an overview of implied structural changes in the PCB moiety spanning atoms 13 to 19. Putting all together, three items that are significantly related to the measured chemical shift pattern there stand out. Going from the  $P_{fr}$  state of  $\text{CphI}\Delta_2$  to the  $P_g$  state of  $\text{AnPixJg}_2$ , correlation analysis suggests bonds C13–C14 and C15–C16 to be substantially shortened, whereas bond C17–C18 is increased in length. This is a forthright indication to a change in extent of  $\pi$ -conjugation between rings C and D, which is obvious for (double) bond C15–C16, whose shortening corresponds to reduced delocalisation through this bond. But also regarding bond C13–14, delineating possible mesomeric structures of PCB points to increased conjugation between rings B and C, while it is reduced between rings C and D. The role of bond C17–C18, though being clearly lengthened, in turn, does not seem to be connected to conjugation on the first glance. Yet, it possibly indicates increased conjugation *within* ring D after the  $\pi$  system has been cut off rings B and C.<sup>d</sup> Although these results remain vague on the basis of still improvable correlation coefficients in Figure 8.6, the variation of bond lengths fulfils the extended and invertible nature of the measured difference pattern. It has been shown above that carbon atoms are insensitive against (non-drastring) supramolecular modifications. It seems there-

<sup>d</sup>Another reason can be presumed in coordination of ring D to the charged sidechain of Asp207 in the  $P_{fr}$  state. It partially quenches the positive mesomeric (+M) effect of the carbonyl group acting on both, bond C17–C18 and the pyrrole nitrogen atom.



fore reasonable to suppose that the observed differences in carbon chemical shifts indicate increased conjugation in the  $P_{fr}$  state, while the same is reduced in the  $P_g$  state.

### 8.1.5 $^{15}\text{N}$ and $^1\text{H}$ Chemical Shifts: Sensors of Hydrogen Bonding

While  $^{13}\text{C}$  chemical shifts of the  $P_r$  states of AnPixJg2 and CphI $\Delta$ 2 are characterised by a symmetric pattern in rings B and C,  $^{15}\text{N}$  and  $^1\text{H}$  chemical shifts do not show such mirrored behaviour, but differ by more than 10 and 2 ppm, respectively.<sup>215</sup> This mysterious dichotomy has posed the question as to how the latter are actually determined. Since  $^{13}\text{C}$  chemical shifts and conjugation underpin the positive charge of PCB being highly delocalised between rings B and C, the obvious difference between pyrrol nitrogen atoms, suggesting the contrary, surprises.<sup>e</sup> Matysik and co-workers suspected that, in the case of CphI $\Delta$ 2, the conjugated system along the C–C bonds is “invisible” to the pyrrole nitrogen atoms,<sup>197,215</sup> but would not dismiss the idea of shifts indicating a distribution of charge.<sup>f</sup> Elgabarty and co-workers hinted at the explicit susceptibility of (in their work)  $^1\text{H}$  chemical shifts towards hydrogen bond parameters.<sup>236,237</sup> Looking back, Figure 8.2 is reminiscent of how atoms exposed to environment react strongly to changes therein. Hence, it is worthwhile to study the relation of relevant chemical shifts to structural parameters, that is, how covalent and non-covalent interactions account for the observations.

Figure 8.7 shows characteristic correlation plots of  $^1\text{H}$  and  $^{15}\text{N}$  chemical shifts with the extent of supramolecular connection of all four pyrrole rings, represented by covalent N–H bonds on one side and non-covalent hydrogen bonds to the environment (apoprotein or water) on the other. There is clear evidence for a correlation between  $^1\text{H}$  shifts and the hydrogen bond length, with ring B seeming to fall out of this trend. However, all four ring protons are visibly correlated to the N–H bond length. This fosters the idea of ring B being only weakly bound to aspartate in both proteins. It is interesting to see that the resemblance of NMR parameters between AnPixJg2 and CphI $\Delta$ 2 also appears in Figure 8.7, although the absolute values deviate. Differences in measured chemical shifts are therefore down to a varied distribution of the underlying structural variable. Most prominently, ring C has a shorter (*i.e.*, stronger) hydrogen bond to aspartate in AnPixJg2, which corresponds to

<sup>e</sup>It has been understood that the pyrrole ring with lower chemical shift for  $^{15}\text{N}$  carries the positive charge, such as ring B in the case of CphI $\Delta$ 2 (consequently, in AnPixJg2 it would reside at ring C).<sup>215</sup>

<sup>f</sup>They refer to a work of Bachovchin and co-workers. However, that work treats hydrogen bonding of histidine residues, whose (highly localised) charge is stabilised by a negatively charged boronic species,<sup>250</sup> whereas PCB operates with diffuse charges.

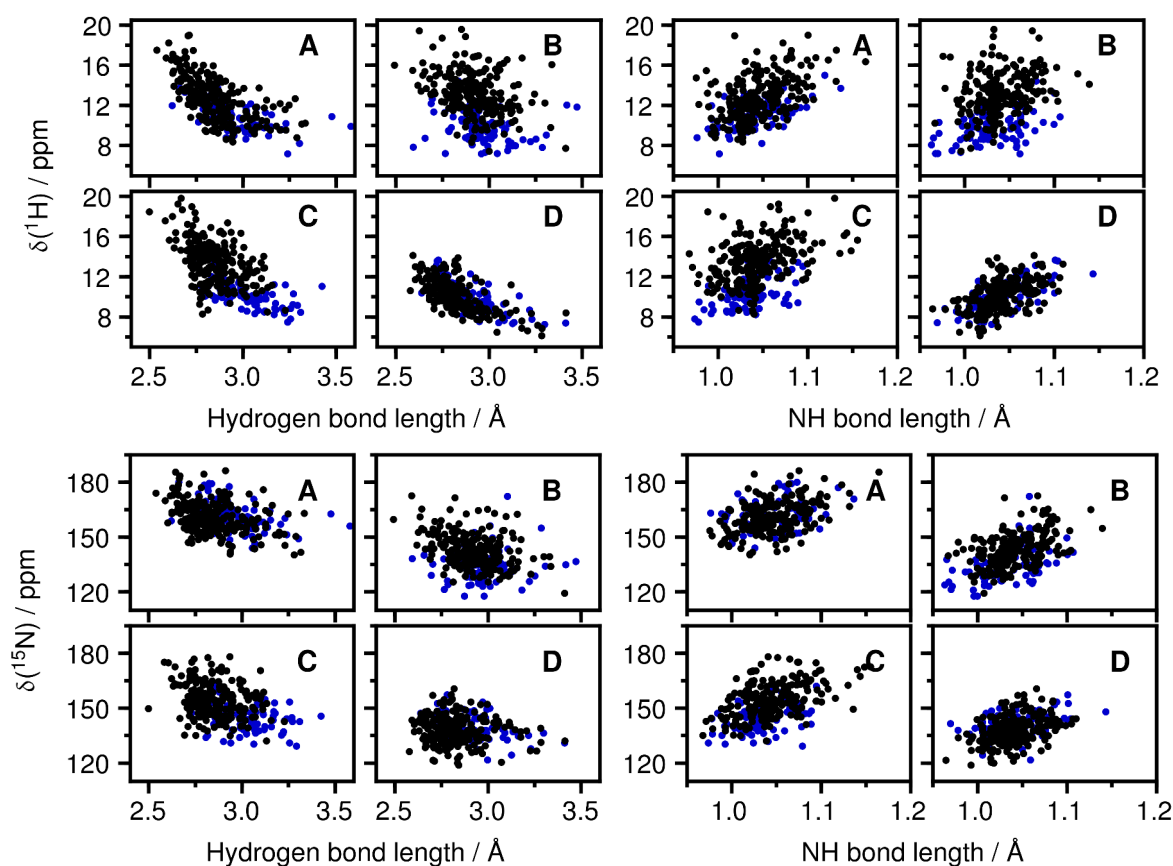


FIGURE 8.7 Correlation between calculated chemical shifts of both,  $^1\text{H}$  (upper panels) and  $^{15}\text{N}$  (lower panels), with the length of hydrogen bonds (left) and N-H bonds (right). The results for AnPixJg2 are shown in black, those for CphI $\Delta$ 2 in blue.

residue Asp291 presenting its negatively charged carboxylate.<sup>§</sup> Ring B too exhibits a slight stretch of its N-H bond in AnPixJg2, supposedly for the same reason, even though it does not react strongly to the hydrogen bond itself.  $^{15}\text{N}$  shifts are mainly determined by the length of N-H bond, that way equally responding to electrostatics in the binding pocket. Apart from this, the individual differences in chemical structure are reflected by a parallel shift of the correlation pattern; in this way, rings A and D are shifted downfield and upfield, respectively. They build a frame for rings B and C that are almost identical, reflecting the symmetry of the  $^{13}\text{C}$  nuclei. More insight into the chemical nature of individual pyrrole rings is provided by correlation of  $^{15}\text{N}$  with  $^1\text{H}$  chemical shifts; the fate of the two atom species is naturally entangled. Indeed, in Figure 8.8 all rings appear quite distinct and correlated. The important message within these correlations is that  $^{15}\text{N}$  and  $^1\text{H}$  chemical shifts cannot be regulated independently, but require, in one way or another, chemical changes to the molecule.

<sup>§</sup>In CphI $\Delta$ 2, Asp207 presents its neutral backbone carbonyl group; cf. Figures 7.4 and 7.9.

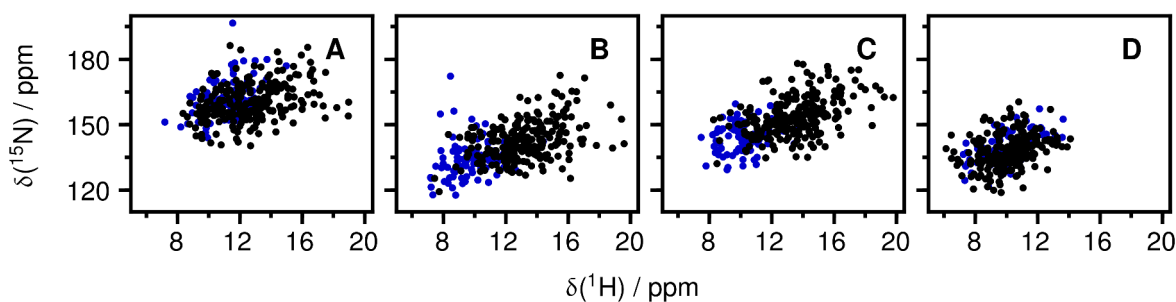


FIGURE 8.8 Correlation of calculated  $^{15}\text{N}$  with  $^1\text{H}$  chemical shifts. The results for AnPixJg2 are shown in black, those for CphI $\Delta$ 2 in blue.

Once more the correlation plots in Figure 8.7 can be translated to a correlation matrix, in analogy to Equation 8.1, which is shown in Figure 8.9. Additionally, it accounts for correlations to the conjugated system represented by the set of C–C bonds. Remarkably, there is hardly any connection of either  $^1\text{H}$  or  $^{15}\text{N}$  chemical shifts to any C–C bond of the chromophore's core. Hence, just as it was presumed from experiments, there is clear evidence that both chemical shifts of pyrrole nitrogen and hydrogen atoms do not respond to the conjugated system. The correlation matrix instructively visualises the determination of chemical shifts by hydrogen bonding (the odd nature of ring B is recognised as well). Consequently in the case of  $^1\text{H}$  and  $^{15}\text{N}$ , no conclusions can be drawn regarding the symmetry of rings B and C, as it was done for  $^{13}\text{C}$  shifts. Instead, the chemical shift pattern is determined *exogenously*. To continue this thought, pyrrole nitrogen atoms appear unlikely to be electronically involved in colour tuning as changes in conjugation would not affect much their electronic state. In turn, it is questionable whether coordination of a charged residue (*e.g.*, Asp207 in CphI $\Delta$ 2) would considerably change the colour.<sup>197</sup>

As a matter of fact, observed correlations do not prove causality. The origins of changes to variable distributions, which could conceivably involve charged residues or a steric restraint, cannot be diverted from the statistical analysis. For that, inspections of the underlying simulated chemical structure is required; nevertheless, it is worthwhile to collate implications. It has been emphasised that the margins of chemical tuning to PCB are limited after certain scenarios could unambiguously be excluded. Therefore, there is every indication that the twist of ring D determines the colour of PCB and that this evidence too is inscribed within the chemical shifts.

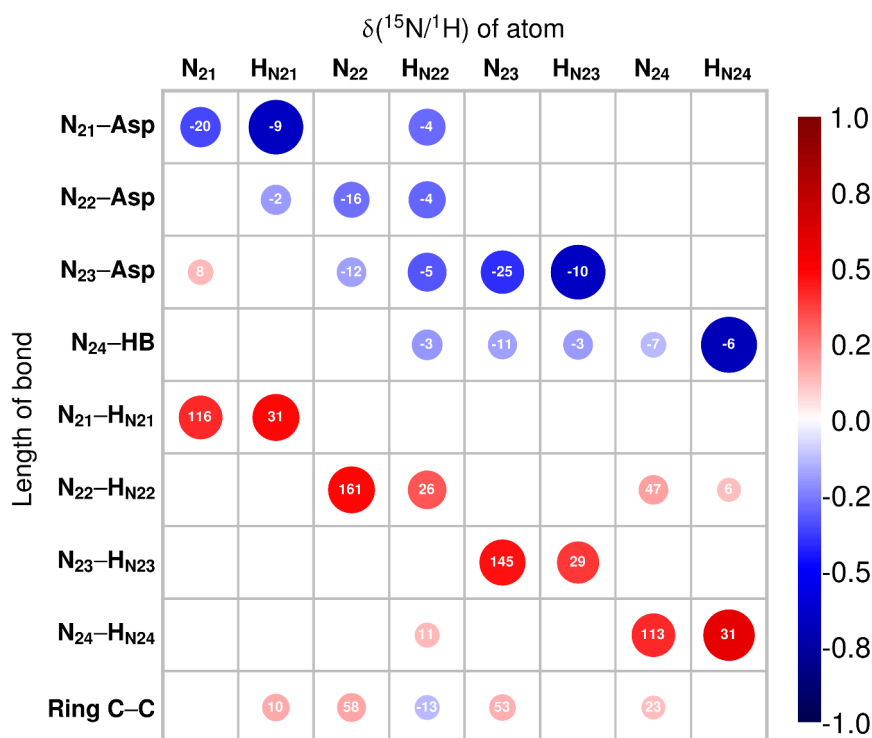


FIGURE 8.9 Correlation matrix in analogy to Equation 8.1, showing the sensitivity of  $^{15}\text{N}$  and  $^1\text{H}$  chemical shifts towards the length of hydrogen and N-H bonds. Additionally, maximal correlation found with any C-C bond of the pyrrole rings is shown (last row). Colour and size of the circles correspond to the Pearson correlation coefficient (see colour bar on the right), while the white numbers denote the slope of linear regression (in  $\text{ppm \AA}^{-1}$ ). Excluded from view are entries with significance level lower than 95 % (p-value).

### 8.1.6 Sampling Protocol for Experimental Conditions

The decision on whether to choose  $P_1$  or  $P_2$  for the computation of NMR shieldings depends on the posed scientific question.  $P_2$  (direct AIMD sampling at elevated temperature) turns out best to analyse correlations as it provides variable distributions broad enough to apply statistical tools (previous section).  $P_1$ , in turn, strongly narrows the variable spread leaving it unapt to generate a statistical dataset – at least when it comes to molecular vibrations. Yet, computation of NMR spectra relies on averaging over extracted snapshots from the MD trajectory.

Figure 8.10 presents calculated 2D  $^1\text{H}\text{-}^{15}\text{N}$  NMR spectra of PCB in AnPixJg2, computed after protocol  $P_1$  or  $P_2$ . They are compared with data from *cross-polarisation magic angle spinning* (CP/MAS)

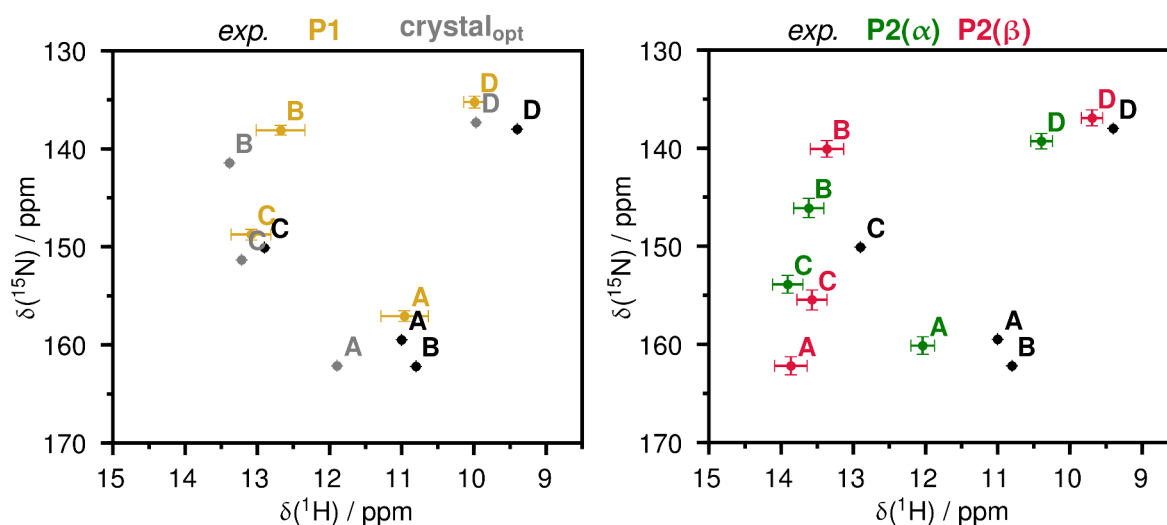


FIGURE 8.10  $2D$   $^1H$ - $^{15}N$  NMR spectra of AnPixJg2 obtained from QM/MM MD following protocol P<sub>1</sub> (left) and P<sub>2</sub> (right); see also Figure 8.1. The computed spectra are compared to experimental data from CP/MAS measurements at  $-40^\circ C$ . The grey signals correspond to direct calculations from the crystal structure after geometry optimisation. The error bars mark the standard error of ensemble average.

measurements, recorded at  $-40^\circ C$ .<sup>h</sup> For P<sub>2</sub>, all  $^1H$  chemical shifts are issued with a clear downfield shift compared to P<sub>1</sub>. One has to keep in mind the AIMD simulations were performed at 300 K (for computational details see Appendix A). At this temperature, molecular bonds substantially deviate from their local potential minimum. In the case of N-H bonds, the shape of the potential is highly anharmonic due to hydrogen bonding. The effect thereof is an asymmetric stretch of the bonds, which leads to an increase in  $^1H$  chemical shifts. Another issue of P<sub>2</sub>, which delivers two independent AIMD spectra for  $\alpha$  and  $\beta$ , stems from short timescales leading to *locked states* of the supramolecular setting. For instance, P<sub>2</sub>( $\alpha$ ) in Figure 8.10 appears highly shifted with respect to P<sub>2</sub>( $\beta$ ) or the experiment. Obviously, the AIMD sampled a (supramolecular) structure not representative for the experiment. From these findings, it is highly advisable to deploy P<sub>1</sub> when spectra from thermodynamic ensembles are to be calculated that are highly influenced by supramolecular interactions with moderate frequency. Furthermore, simulated annealing (or geometry optimisation) corresponds to a temperature of 0 K, approximating conditions of the CP/MAS NMR experiments best.

A disagreement between computation and experiment affects ring B, whose chemical shift is expected higher for  $^{15}N$  ( $>+20$  ppm) and lower for  $^1H$  ( $<-1.5$  ppm) in the experiment. By virtue of the quantum chemical connection, depicted in Figure 8.8, only chemical changes (*i.e.*, changing the structural model) could account for size and kind of this large deviation, given the assignment of

<sup>h</sup>The data was provided by Chen Song from the group of Jörg Matysik.

TABLE 8.3  $^{15}\text{N}$  chemical shifts of PCB CphI $\Delta$ 2 in ppm. The results calculated in this work are compared to dynamic nuclear polarisation solid-state MAS measurements and QM/MM calculations based on geometry optimisation of the crystal structure.

Ring	Experiment <sup>215</sup>	Reference <sup>215</sup>	This work (P1)	This work (P2)
A	159.3	159.5	157.1	161.2
B	146.3	141.5	138.1	143.1
C	156.9	153.8	148.8	154.7
D	132.6	130.0	135.3	138.1

peaks from experiment was carried out correctly.

In the case of CphI $\Delta$ 2, minor discrepancies for both protocols characterise the result of  $^{15}\text{N}$  chemical shifts, shown in Table 8.3. The corresponding measurements were performed at 100 K, but P2 seems to better meet the results.<sup>215</sup> Mroginski and co-workers extensively studied the effect of protonation of His260 and His290 in CphI $\Delta$ 2.<sup>214</sup> Hence, different structure models may have to be evaluated for both, AnPixJg2 and CphI $\Delta$ 2. Addressing this issue, however, lies beyond the scope of this work and will be part of the ongoing study.

## 8.2 UV-Vis and ECD Spectroscopy

### 8.2.1 Absorption Spectra from AIMD

Electronic excitation happens within timescales short enough to capture deflections from molecular vibrations and rotations (*vertical absorption*), in contrast to free induction decay in NMR that exhibits comparably large relaxation times ( $T_1$ ) and yields temperature-dependent coalesced signals of structural averages.<sup>217,218,241,251</sup> One important cause of line broadening in UV-Vis spectra is therefore the roto-vibrational finestructure.<sup>i</sup> Yet, single-point calculations return line spectra corresponding to electronic excitations of only the current molecular geometry. If this geometry has been optimised beforehand, the resulting UV-Vis (line) spectrum merely reflects the absorption properties of the molecule in its PES minimum – even if several conformers are considered (*cf.* sampling for VCD in Chapter 3). Hence, as static calculations do not account for thermodynamic line broadening, a frequent workaround is to convolute spectral lines with Gaussian functions, mimicking the normal distribution of the molecular degrees of freedom.

AIMD simulations do not face this problem, but can give continuous spectra without additional broadening (see Section 2.3).<sup>j</sup> Regarding the sampling protocols presented in Subsection 8.1.1, however, an important adaption has to be made. As a result from the disparate timescales of NMR and electronic excitation, the ensemble spectrum is not an average of calculated roots but a superposition thereof.<sup>k</sup> Individual spectral lines of each AIMD snapshot are merely convoluted with a narrow Lorentzian function for regularisation and to account for short lifetimes of excited states.<sup>6</sup> Following the new protocol, the finestructure (*i.e.*, peak shapes) of the obtained UV-Vis (or ECD) spectrum is determined by the probability distribution of the underlying statistical ensemble.

### 8.2.2 Robustness of Absorption Properties of PCB

Figure 8.11 presents UV-Vis spectra calculated for PCB in AnPixJg2 and Cph1Δ2 on the basis of AIMD simulations. Therein, different sizes of the quantum mechanical part (QMP) were used –

<sup>i</sup>The finestructure is not resolved in solution spectra, but can be retrieved, *e.g.*, in gas phase.

<sup>j</sup>Yet, AIMD can not cover the exact vibronic finestructure as long as the nuclear degrees of freedom are treated classically, let alone working under the Born-Oppenheimer approximation (*cf.* Section 2.1).

<sup>k</sup>This relates to the interaction picture of quantum mechanics.<sup>39</sup>

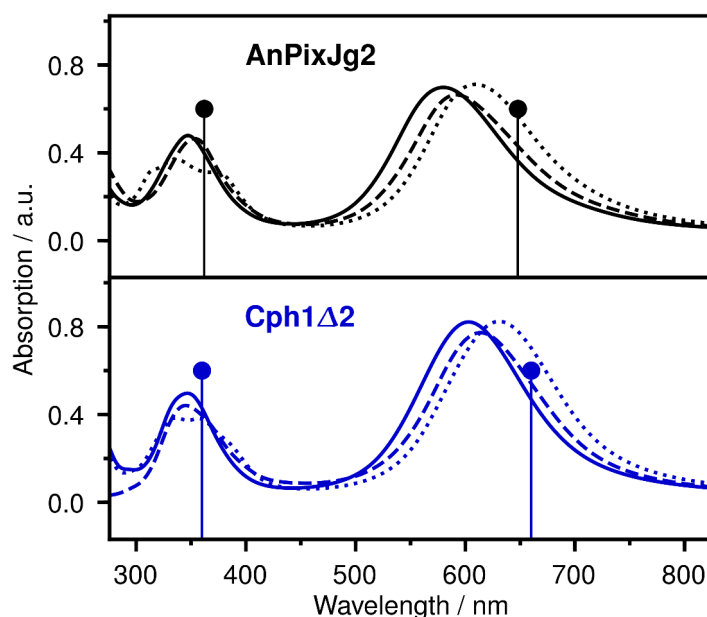


FIGURE 8.11 Calculated UV-Vis spectra of PCB in AnPixJg2( $\alpha$ ) (top) and Cph1 $\Delta$ 2 (bottom) for different sizes of the QM part (QMP), including either the full cofactor and neighbouring protein residues (dashed lines) or a reduced chromophore core consisting of the  $\pi$  system with adjacent atomic groups (solid lines). The dotted lines represent calculations without consideration of the protein. The pins mark experimental values.<sup>169,170</sup>

similar to the calculation of nuclear shieldings in Section 8.1. But in contrast to NMR, here, the influence of the supramolecular environment is insignificant, even towards the propionate sidechains of PCB. The inclusion of the full chromophore and neighbouring protein residues does only slightly red-shift the result compared to the absorption of a reduced chromophore core (for QMP definitions and further computational details see Appendix A). In both proteins, the calculated results are slightly blue-shifted with respect to the experiment, the absorption maximum of lower energy ( $S_0 \rightarrow S_1$ ) with about 0.15 eV not as much as the one of high energy ( $S_0 \rightarrow S_2$ ) with about 0.22 eV.<sup>1</sup> These results are very similar to the calculations in CBCR Slr1393g3 by Wiebeler and co-workers<sup>174</sup> and attest the employed method (sTD-DFT/CAM-B<sub>3</sub>LYP) to be a viable and cost-effective approach for PCB in proteins.<sup>74,252–254</sup>

A main question regarding colour tuning in phytochromes and CBCRs is how the protein influences the cofactor to gain the desired properties. The solvatochromic model presumes an electrostatic effect,<sup>195</sup> but the previous chapter has already indicated this an unlikely scenario. The protein effect

<sup>1</sup>The difference in energy is not proportional to the difference in wavelength; with  $E = h\frac{c}{\lambda}$ , lines equidistant in energy appear closer at low wavelengths.



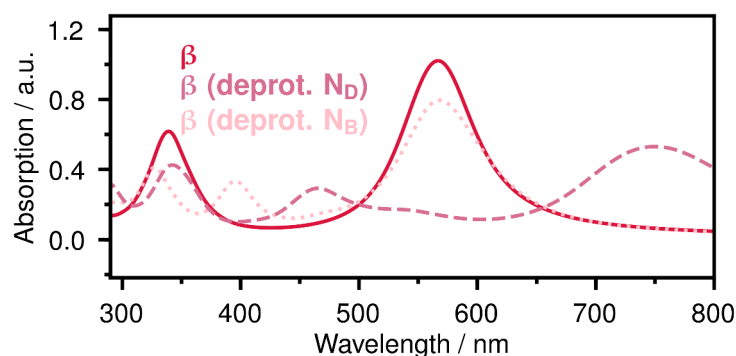


FIGURE 8.12 Calculated UV-Vis spectra of PCB in AnPixJg<sub>2</sub> ( $\beta$ ) for deprotonation of an “allowed” (ring B) and a “forbidden” (ring D) position (protocol P<sub>I</sub>). The positions of all remaining atoms were kept unchanged.

could also be of steric nature. Indeed, Matute and co-workers reported distorted structures from crystal structures showing better agreement with experiment than re-optimised geometries.<sup>225,255</sup> Figure 8.11 shows that also in the case of AnPixJg<sub>2</sub> and CphI $\Delta$ <sub>2</sub>, complete removal of the protein from calculation, together with the PCB sidechains, only slightly improves the position of the lowest energy absorption band. This result highly suggests the protein exerts, above all, steric influence on the cofactor.<sup>174</sup>

Although it has been shown the protonation states do not influence the dye spectrum,<sup>231</sup> it is worthwhile to theoretically look at the effect of deprotonation of PCB on absorption properties. Chemically it can be distinguished between “allowed” and “forbidden” sites of deprotonation. In the NMR study, ring B exhibited a loose proton compared to the other pyrroles (*cf.* Section 8.1). It represents an “allowed” (*i.e.*, acidic) site, because deprotonation would not perturb conjugation. Ring D, in turn, marks a “forbidden” position as deprotonation would leave a negative charge at that site. Compensation with the positive charge of rings B and C would essentially destroy the C<sub>15</sub>–C<sub>16</sub> double bond, leading to “enforced conjugation” with ring C. By all means, it would represent a very artificial, unstable situation. The classification of rings B and D is reflected in Figure 8.12: Deprotonation of N<sub>B</sub> only affects the absorption properties in second order (*i.e.*, the S<sub>0</sub>→S<sub>2</sub> transition), while depriving ring D of its acidic proton drastically red-shifts the lowest energy absorption band almost into the near-infrared region. As this state represents “enforced conjugation”, it can be considered as a pre-emption of the P<sub>fr</sub> state where conjugation between rings D and C putatively happens on a natural basis. Interestingly, this result suggests the absorption properties of PCB can be tuned to far-red regions even for a naturally red-absorbing state by deprotonation only.

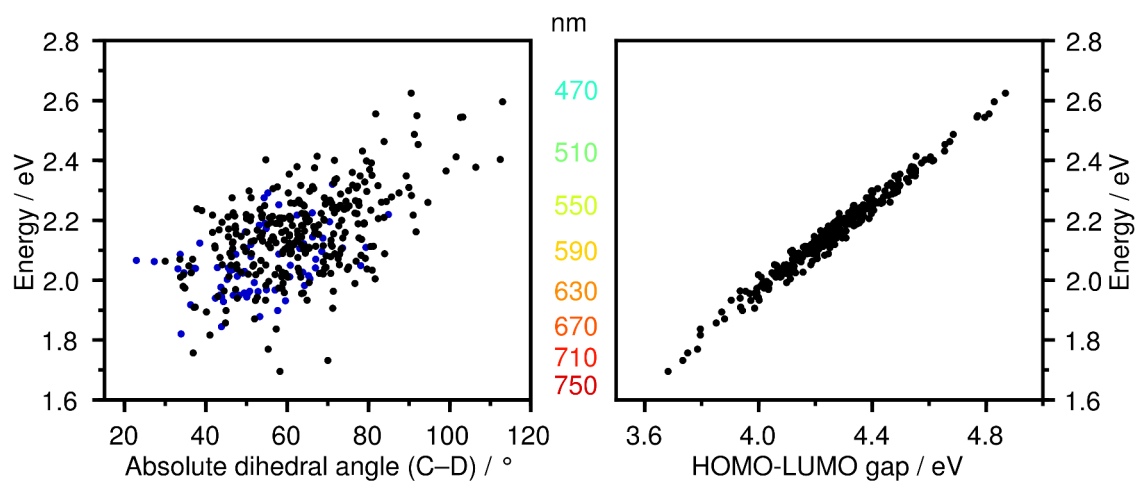


FIGURE 8.13 Left: Correlation of the  $S_0 \rightarrow S_1$  absorption energy with the out-of-plane rotation of ring D with respect to ring C in AnPixJg2 (black) and CphI $\Delta$ 2 (blue). Right: Correlation of the absorption energy with the HOMO-LUMO gap.

### 8.2.3 Colour Tuning is Mediated by Ring D

Wiebeler and co-workers as well as Peng and co-workers related colour tuning of PCB to the rotation of ring D, which leads to either a decrease or an increase of co-planarity with respect to the conjugated system formed by rings B and C.<sup>159,174</sup> The effective length of conjugation determines the energy gap of the first electronic transition. In view of the NMR study in Section 8.1, wherein characteristic  $P_g/P_{fr}$  difference patterns could be related to conjugation length, a similar mechanism of colour tuning can be expected for AnPixJg2 and CphI $\Delta$ 2.

Figure 8.13 presents the correlation of  $S_0 \rightarrow S_1$  absorption energy with rotation of ring D, based on AIMD trajectories. Although both proteins are simulated in the  $P_r$  state only, the dependence of the absorption energy on the ring D twist is clearly visible: Dihedral angles between ring C and D in the range of 20 to 40 degrees lead to far-red absorption, whereas angles above 80 degrees result in green absorption. As both proteins contain PCB with an average dihedral angle of about 60 degrees in the dark state, the latter absorbs red light and is accordingly labelled as  $P_r$  state. The right plot in Figure 8.13 affirms that the electronic absorption is a clean HOMO-LUMO transition, showing high correlation to the frontier orbital gap.

Turning towards the shape of frontier orbitals, illustrated in Figure 8.14, it can be seen that the entire conjugated system of PCB as well as ring D are involved in the  $S_0 \rightarrow S_1$  transition.<sup>225,255</sup> Yet, the pyr-

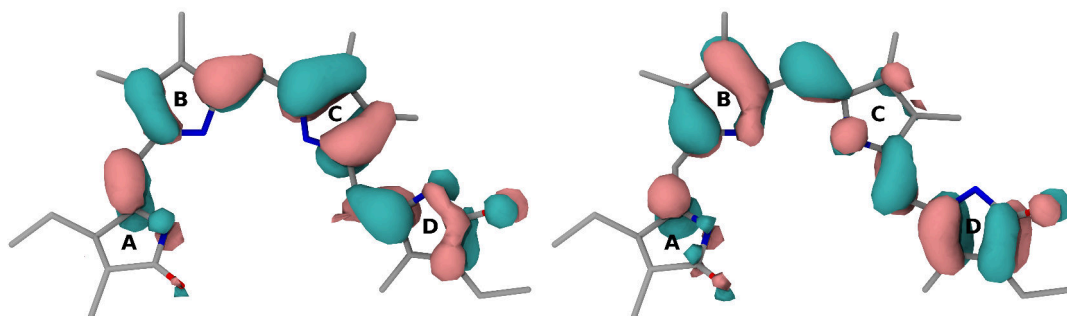


FIGURE 8.14 Frontier molecular orbitals involved in the  $S_0 \rightarrow S_1$  transition: HOMO (left) and LUMO (right).

role nitrogen atoms are only slightly included in both orbitals, which is in line with the deprotonation study shown above. The protonated site of pyrrole is orthogonal to the conjugated system and de- or re-protonation does not take influence on the mesomeric effect of the nitrogen atom. At the same time, Figure 8.14 emphasises that the effective conjugation length clearly wins over electrostatic effects. The conjugated system, and so the frontier orbitals, are well shielded from the (supramolecular) environment.<sup>m</sup> In this regard, PCB differs from retinal, where a negatively charged protein residue directly influences the frontier orbitals through the iminium moiety of the cofactor carrying a (localised) positive charge.<sup>179</sup> Ring D strongly participates in the  $S_0 \rightarrow S_1$  transition, as it carries non-zero orbital coefficients for both the HOMO and the LUMO. In this way, Figure 8.14 underlines the relation of ring twist and absorption frequency by means of effective conjugation length, adjusted by ring D. Ring A, in turn, cannot participate in conjugation, wherefore the ring twist does not influence the absorption energy here.<sup>174</sup>

#### 8.2.4 Differentiation of $P_r$ Sub-States in AnPixJg2

Optical spectroscopy provides insights not only into the colour tuning mechanism, but also gives experimental evidence regarding structural heterogeneity.  $\alpha$  and  $\beta$  conformers have been found as putative  $P_r$  sub-states of AnPixJg2, yet they cannot be identified by means of NMR spectroscopy. Although UV-Vis does not resolve the two states either (Figure 8.15), it has been reported by Rockwell and co-workers that the  $\alpha$  and  $\beta$  conformers of tetrapyrroles lead to characteristic patterns in an

<sup>m</sup>Problematic in the context of *open-chain* systems such as PCB is moreover the consultation of the Gouterman model,<sup>256</sup> originally proposed for *cyclic* tetrapyrroles.<sup>197</sup> The model does not apply, because it assumes symmetry and degeneracy of HOMO and LUMO, respectively – prerequisites that do not exist in the open-chain form of tetrapyrroles.

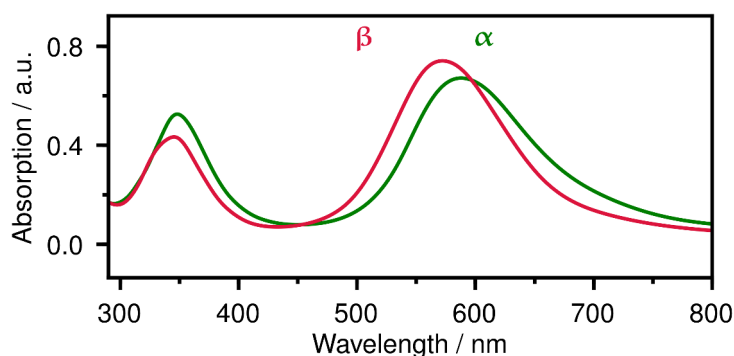


FIGURE 8.15 Calculated UV-Vis spectra of PCB in AnPixJg2 in its  $\alpha$  and  $\beta$  sub-state, respectively.

ECD spectrum.<sup>223</sup> The relation of D ring orientation and circular dichroism is not surprising, as the twisted conjugated system clearly outlines a left- or right-handed helix. Indeed, calculated rotational strengths for ECD of PCB exhibit a mirror-like behaviour of  $\alpha$  and  $\beta$  sub-state, shown in Figure 8.16. The individual spectra contain the characteristic change of sign going from high to low energy absorption, which likens them to many other examples in the phytochrome superfamily.<sup>172,220</sup>

In contrast, the ECD spectrum of AnPixJg2 ( $P_r$ ) measured *in vivo* looks quite different.<sup>n</sup> Instead of showcasing two peaks of opposite sign, it contains a positive signal for both electronic transitions (Figure 8.16, bottom). With this, the spectrum cannot unambiguously be assigned to either the  $\alpha$  or  $\beta$  sub-state, as it resembles more the former at higher wavelengths, but appears like the latter at low wavelengths.<sup>o</sup>

The occurrence of the *in vivo* spectrum can be comprehended by having a closer look at how the ECD signal is connected to ring D rotation. It is important to note that the  $\alpha$  and  $\beta$  sub-states do not behave like enantiomers. Thus, their individual ECD spectra need not combine to give the baseline. In fact, the two absorption maxima act differently due to different extents of chirality of the involved orbitals. As a consequence, the quality of mirror-like behaviour of the ECD signals differs, as shown in Figure 8.17. While the high energy absorption is truly mirrored, this is not the case for the low energy absorption. Thus, coexistence of  $\alpha$  and  $\beta$  sub-states does not lead to complete cancellation of ECD, but gives a surplus with respect to the low energy absorption band. As the experimental spectrum contains two positive ECD bands, it indicates that AnPixJg2 possesses a heterogeneous  $P_r$  state with  $\alpha$  and  $\beta$  sub-states, whereas a slight imbalance exists in favour of the  $\alpha$  sub-state.

<sup>n</sup>The experimental spectrum was recorded by Susanne Altmayer in the groups of Jörg Matysik and Annette Beck-Sickingher.

<sup>o</sup>The negative band at 300 nm and below stems from tryptophan residues of the apoprotein.

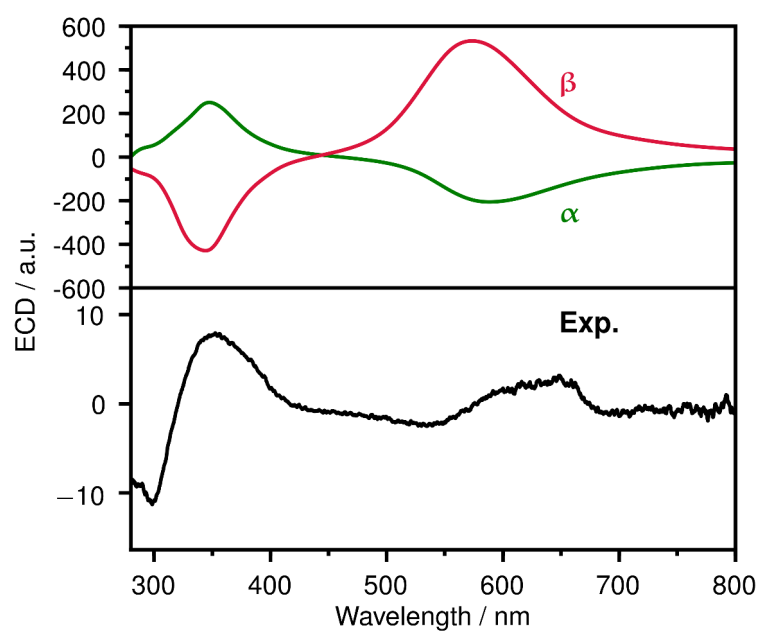


FIGURE 8.16 Top: Calculated ECD spectrum of PCB in AnPixJg2 in its  $\alpha$  and  $\beta$  sub-state, respectively. Bottom: Measured ECD spectrum of the  $P_r$  state of AnPixJg2 *in vivo*.

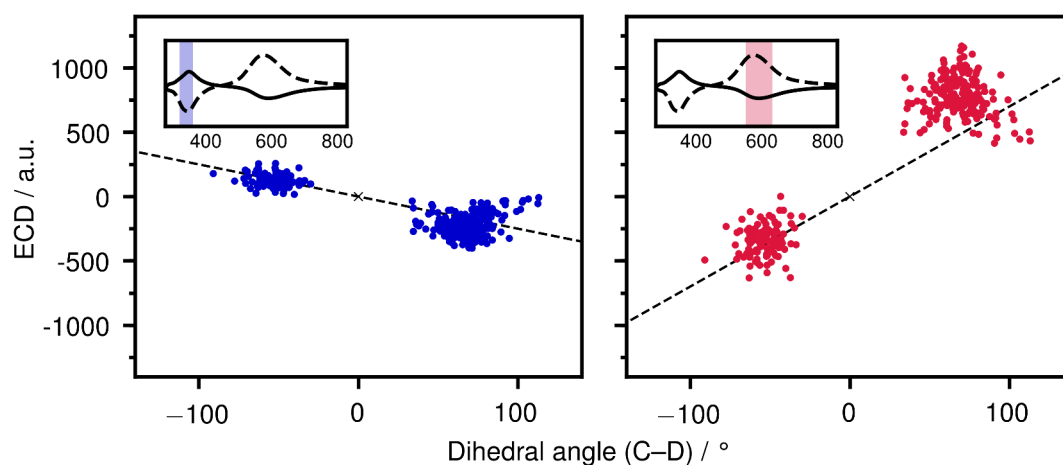


FIGURE 8.17 Correlation of ECD rotational strengths of PCB in AnPixJg2 with the dihedral angle between rings C and D. The results for two electronic transitions are shown as indicated by the insets.

### 8.3 Conclusion

In this chapter, a thorough exploration of structural as well as spectroscopic characteristics of PCB in AnPixJg2 and Cph1 $\Delta$ 2 has been presented. Based on large scale FFMD simulations, it has been shown that in AnPixJg2 the chromophore-binding pocket undergoes substantial reorganisation related to solvation. This work underlines that hydrophobic interactions cannot compete with the presence of water and should not be overrated in terms of explaining the protein's functionality. Furthermore, a new rotamer of highly mobile ring D of PCB in AnPixJg2 was found, which is not present in the crystal structure. Labelled as  $\alpha$  and  $\beta$ , the two sub-states can be differentiated by means of ECD spectroscopy. Structural fluctuations of PCB in both proteins were recorded by means of QM/MM MD simulations. With this, a statistical dataset was generated, building the basis for deeper interpretation of experimental NMR and UV-Vis patterns. It became clear that correlating spectroscopic signatures with structural patterns bears high informative value. Although only the P<sub>r</sub> states of AnPixJg2 and Cph1 $\Delta$ 2 were analysed, conclusions regarding the colour tuning mechanism could be drawn. The characteristic chemical shift pattern observed for ring C and D upon photoconversion indicates a changed conjugation length within the photoproducts. While P<sub>fr</sub> exhibits increased conjugation, the opposite was inferred for P<sub>g</sub>. UV-Vis calculations supported the hypothesis that colour tuning of PCB is performed by means of ring D. Depending on how strong it is participating in conjugation, it renders the absorption maximum with red or blue-shift.

## BIBLIOGRAPHY





## REFERENCES

- [1] Jähnigen, S.; Scherrer, A.; Vuilleumier, R.; Sebastiani, D. Chiral Crystal Packing Induces Enhancement of Vibrational Circular Dichroism. *Angew. Chem. Int. Ed.* 2018, *57*, 13344–13348.
- [2] Scarbath-Evers, L. K.; Jähnigen, S.; Elgabarty, H.; Song, C.; Narikawa, R.; Matysik, J.; Sebastiani, D. Structural Heterogeneity in a Parent Ground-State Structure of AnPixJg<sub>2</sub> Revealed by Theory and Spectroscopy. *Phys. Chem. Chem. Phys.* 2017, *19*, 13882–13894.
- [3] Lehn, J.-M. *Supramolecular Chemistry: Concepts and Perspectives*; VCH Verlagsgesellschaft mbH, Weinheim, 1995.
- [4] Temelso, B.; Shields, G. C. The Role of Anharmonicity in Hydrogen-Bonded Systems: The Case of Water Clusters. *J. Chem. Theory Comput.* 2011, *7*, 2804–2817.
- [5] Pasteur, L. Mémoire Sur La Relation Qui Peut Exister Entre La Forme Cristalline et La Composition Chimique, et Sur La Cause de La Polarisation Rotatoire. *Comptes Rendus L'Académie Sci.* 1848, *26*, 535–538.
- [6] Nafie, L. A. *Vibrational Optical Activity: Principles and Applications*; John Wiley & Sons Ltd., 2011.
- [7] Barron, L. D. *Molecular Light Scattering and Optical Activity*; Cambridge University Press: Cambridge, 2004.
- [8] Holzwarth, G.; Hsu, E. C.; Mosher, H. S.; Faulkner, T. R.; Moscovitz, A. Infrared Circular Dichroism of Carbon-Hydrogen and Carbon-Deuterium Stretching Modes. Observations. *J. Am. Chem. Soc.* 1974, *96*, 251–252.
- [9] Purcell, E. M.; Torrey, H. C.; Pound, R. V. Resonance Absorption by Nuclear Magnetic Moments in a Solid. *Phys. Rev.* 1946, *69*, 37–38.

## REFERENCES

- [10] Rabi, I. I.; Zacharias, J. R.; Millman, S.; Kusch, P. A New Method of Measuring Nuclear Magnetic Moment. *Phys. Rev.* 1938, *53*, 318.
- [11] Callaghan, P. T. *Translational Dynamics and Magnetic Resonance: Principles of Pulsed Gradient Spin Echo NMR*; Oxford University Press, 2011.
- [12] Hohenberg, P.; Kohn, W. The Inhomogeneous Electron Gas. *Phys. Rev.* 1964, *136*, B864.
- [13] Martin, R. M. *Electronic Structure: Basic Theory and Practical Methods*; Cambridge University Press, 2004.
- [14] Sebastiani, D.; Parrinello, M. A New Ab-Initio Approach for NMR Chemical Shifts in Periodic Systems. *J. Phys. Chem. A* 2001, *105*, 1951–1958.
- [15] Putrino, A.; Sebastiani, D.; Parrinello, M. Generalized Variational Density Functional Perturbation Theory. *J. Chem. Phys.* 2000, *113*, 7102–7109.
- [16] Scherrer, A.; Agostini, F.; Sebastiani, D.; Gross, E. K. U.; Vuilleumier, R. Nuclear Velocity Perturbation Theory for Vibrational Circular Dichroism: An Approach Based on the Exact Factorization of the Electron-Nuclear Wave Function. *J. Chem. Phys.* 2015, *143*, 074106.
- [17] Scherrer, A.; Vuilleumier, R.; Sebastiani, D. Vibrational Circular Dichroism from Ab Initio Molecular Dynamics and Nuclear Velocity Perturbation Theory in the Liquid Phase. *J. Chem. Phys.* 2016, *145*, 084101.
- [18] Tuckerman, M. E. *Statistical Mechanics: Theory and Molecular Simulation*; Oxford University Press Inc., New York, 2010.
- [19] Ryckaert, J.-P.; Ciccotti, G.; Berendsen, H. J. Numerical Integration of the Cartesian Equations of Motion of a System with Constraints: Molecular Dynamics of n-Alkanes. *J. Comput. Phys.* 1977, *23*, 327–341.
- [20] VandeVondele, J.; Krack, M.; Mohamed, F.; Parrinello, M.; Chassaing, T.; Hutter, J. QUICKSTEP: Fast and Accurate Density Functional Calculations Using a Mixed Gaussian and Plane Waves Approach. *Comput. Phys. Commun.* 2005, *167*, 103–128.
- [21] Weber, V.; Iannuzzi, M.; Giani, S.; Hutter, J.; Declerck, R.; Waroquier, M. Magnetic Linear

- Response Properties Calculations with the Gaussian and Augmented-Plane-Wave Method. *J. Chem. Phys.* 2009, *131*, 014106.
- [22] Schrödinger, E. An Undulatory Theory of the Mechanics of Atoms and Molecules. *Phys. Rev.* 1926, *28*, 1049–1070.
- [23] Szabo, A.; Ostlund, N. S. *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory*; Dover Publications Inc., 1996.
- [24] Abedi, A.; Maitra, N. T.; Gross, E. K. U. Correlated Electron-Nuclear Dynamics: Exact Factorization of the Molecular Wavefunction. *J. Chem. Phys.* 2012, *137*, 22A530.
- [25] Abedi, A.; Maitra, N. T.; Gross, E. K. U. Exact Factorization of the Time-Dependent Electron-Nuclear Wave Function. *Phys. Rev. Lett.* 2010, *105*, 123002.
- [26] Alonso, J. L.; Clemente-Gallardo, J.; Echenique-Robba, P.; Jover-Galtier, J. A. Comment on Correlated Electron-Nuclear Dynamics: Exact Factorization of the Molecular Wavefunction. *J. Chem. Phys.* 2013, *139*, 087101.
- [27] Abedi, A.; Maitra, N. T.; Gross, E. K. U. Response to Comment on Correlated Electron-Nuclear Dynamics: Exact Factorization of the Molecular Wavefunction. *J. Chem. Phys.* 2013, *139*, 087102.
- [28] Jackson, J. D. *Classical Electrodynamics*, 3rd ed.; John Wiley & Sons Ltd, 1999.
- [29] Agostini, F.; Tavernelli, I.; Ciccotti, G. Nuclear Quantum Effects in Electronic (Non)Adiabatic Dynamics. *Eur. Phys. J. B Condens. Matter Phys.* 2018, *91*, 139.
- [30] Schild, A.; Agostini, F.; Gross, E. K. U. Electronic Flux Density beyond the Born-Oppenheimer Approximation. *J. Phys. Chem. A* 2016, *120*, 3316–3325.
- [31] Eich, F. G.; Agostini, F. The Adiabatic Limit of the Exact Factorization of the Electron-Nuclear Wave Function. *J. Chem. Phys.* 2016, *145*, 054110.
- [32] Born, M.; Oppenheimer, R. Zur Quantentheorie Der Molekeln. *Ann. Phys.* 1927, *389*, 457–484.

## REFERENCES

- [33] Born, M.; Huang, K. *Dynamical Theory of Crystal Lattices*; Oxford University Press, 1954.
- [34] Reinhold, J. *Quantentheorie Der Moleküle*, 3rd ed.; B. G. Teubner Verlag, 2006.
- [35] Marx, D.; Hutter, J. *Ab Initio Molecular Dynamics: Basic Theory and Advanced Methods*, 1st ed.; Cambridge University Press, 2009.
- [36] Nafie, L. A. Adiabatic Molecular Properties beyond the Born–Oppenheimer Approximation. Complete Adiabatic Wave Functions and Vibrationally Induced Electronic Current Density. *J. Chem. Phys.* 1983, 79, 4950–4957.
- [37] Nafie, L. A.; Freedman, T. B. Vibronic Coupling Theory of Infrared Vibrational Transitions. *J. Chem. Phys.* 1983, 78, 7108–7116.
- [38] Scherrer, A.; Vuilleumier, R.; Sebastiani, D. Nuclear Velocity Perturbation Theory of Vibrational Circular Dichroism. *J. Chem. Theory Comput.* 2013, 9, 5305–5312.
- [39] McQuarrie, D. A. *Quantum Chemistry*, 2nd ed.; University Science Books, 2007.
- [40] Kohn, W.; Sham, L. J. Self-Consistent Equations Including Exchange and Correlation Effects. *Phys Rev* 1965, 140, A1133.
- [41] Koch, W.; Holthausen, M. C. *A Chemists Guide to Density Functional Theory*, 2nd ed.; Wiley-VCH Verlag GmbH, 2001.
- [42] Jacob, C. R.; Reiher, M. Spin in Density-Functional Theory. *Int. J. Quantum Chem.* 2012, 112, 3661–3684.
- [43] Runge, E.; Gross, E. K. U. Density-Functional Theory for Time-Dependent Systems. *Phys. Rev. Lett.* 1984, 52, 997–1000.
- [44] Burke, K.; Werschnik, J.; Gross, E. K. U. Time-Dependent Density Functional Theory: Past, Present, and Future. *J. Chem. Phys.* 2005, 123, 062206.
- [45] Casida, M. E. In *Recent Developments and Applications of Modern Density Functional Theory*; Seminario, J., Ed.; Theoretical and Computational Chemistry; Elsevier, 1996; Vol. 4; pp 391–439.

- [46] Sternheimer, R. M. Electronic Polarizabilities of Ions from the Hartree-Fock Wave Functions. *Phys. Rev.* 1954, *96*, 951–968.
- [47] Gonze, X. Perturbation Expansion of Variational Principles at Arbitrary Order. *Phys. Rev. A* 1995, *52*, 1086–1095.
- [48] Baroni, S.; Giannozzi, P.; Testa, A. Green's-Function Approach to Linear Response in Solids. *Phys. Rev. Lett.* 1987, *58*, 1861–1864.
- [49] Gonze, X. Adiabatic Density-Functional Perturbation Theory. *Phys. Rev. A* 1995, *52*, 1096–1114.
- [50] Gonze, X. First-Principles Responses of Solids to Atomic Displacements and Homogeneous Electric Fields: Implementation of a Conjugate-Gradient Algorithm. *Phys. Rev. B* 1997, *55*, 10337–10354.
- [51] Swope, W. C.; Andersen, H. C.; Berens, P. H.; Wilson, K. R. A Computer Simulation Method for the Calculation of Equilibrium Constants for the Formation of Physical Clusters of Molecules: Application to Small Water Clusters. *J. Chem. Phys.* 1982, *76*, 637–649.
- [52] Berendsen, H. J. C. *Simulating the Physical World*; Cambridge University Press, 2007.
- [53] Vanommeslaeghe, K.; Hatcher, E.; Acharya, C.; Kundu, S.; Zhong, S.; Shim, J.; Darian, E.; Guvench, O.; Lopes, P.; Vorobyov, I.; Mackerell Jr., A. D. CHARMM General Force Field: A Force Field for Drug-like Molecules Compatible with the CHARMM All-Atom Additive Biological Force Fields. *J. Comput. Chem.* 2010, *31*, 671–690.
- [54] Kaminski, S.; Mroginiski, M. A. Molecular Dynamics of Phycocyanobilin Binding Bacteriophytochromes: A Detailed Study of Structural and Dynamic Properties. *J. Phys. Chem. B* 2010, *114*, 16677–16686.
- [55] Vanommeslaeghe, K.; Raman, E. P.; MacKerell, A. D. Automation of the CHARMM General Force Field (CGenFF) II: Assignment of Bonded Parameters and Partial Atomic Charges. *J. Chem. Inf. Model.* 2012, *52*, 3155–3168.
- [56] MacKerell, A. D.; Bashford, D.; Bellott, M.; Dunbrack, R. L.; Evanseck, J. D.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; Joseph-McCarthy, D.; Kuchnir, L.; Kuczera, K.; Lau, F.

## REFERENCES

- T. K.; Mattos, C.; Michnick, S.; Ngo, T.; Nguyen, D. T.; Prodhom, B.; Reiher, W. E.; Roux, B.; Schlenkrich, M.; Smith, J. C.; Stote, R.; Straub, J.; Watanabe, M.; Wiórkiewicz-Kuczera, J.; Yin, D.; Karplus, M. All-Atom Empirical Potential for Molecular Modeling and Dynamics Studies of Proteins. *J. Phys. Chem. B* 1998, *102*, 3586–3616.
- [57] Warshel, A.; Levitt, M. Theoretical Studies of Enzymic Reactions: Dielectric, Electrostatic and Steric Stabilization of the Carbonium Ion in the Reaction of Lysozyme. *J. Mol. Biol.* 1976, *103*, 227–249.
- [58] Sherwood, P. Hybrid Quantum Mechanics/Molecular Mechanics Approaches. *NIC Ser.* 2000, *3*, 285–305.
- [59] Carloni, P.; Rothlisberger, U.; Parrinello, M. The Role and Perspective of Ab Initio Molecular Dynamics in the Study of Biological Systems. *Acc. Chem. Res.* 2002, *35*, 455–464.
- [60] Cao, L.; Ryde, U. On the Difference between Additive and Subtractive QM/MM Calculations. *Front. Chem.* 2018, *6*, 89–89.
- [61] Ihrig, A. C.; Schiffmann, C.; Sebastiani, D. Specific Quantum Mechanical/Molecular Mechanical Capping-Potentials for Biomolecular Functional Groups. *J. Chem. Phys.* 2011, *135*, 214107.
- [62] Laino, T.; Mohamed, F.; Laio, A.; Parrinello, M. An Efficient Real Space Multigrid QM/MM Electrostatic Coupling. *J. Chem. Theory Comput.* 2005, *1*, 1176–1184.
- [63] Laino, T.; Mohamed, F.; Laio, A.; Parrinello, M. An Efficient Linear-Scaling Electrostatic Coupling for Treating Periodic Boundary Conditions in QM/MM Simulations. *J. Chem. Theory Comput.* 2006, *2*, 1370–1378.
- [64] Keith, T. A.; Bader, R. F. Calculation of Magnetic Response Properties Using a Continuous Set of Gauge Transformations. *Chem. Phys. Lett.* 1993, *210*, 223–231.
- [65] Silvestrelli, P. L.; Marzari, N.; Vanderbilt, D.; Parrinello, M. Maximally-Localized Wannier Functions for Disordered Systems: Application to Amorphous Silicon. *Solid State Commun.* 1998, *107*, 7–11.

- [66] Marzari, N.; Vanderbilt, D. Maximally Localized Generalized Wannier Functions for Composite Energy Bands. *Phys. Rev. B* 1997, *56*, 12847–12865.
- [67] Nafie, L. A. Velocity-Gauge Formalism in the Theory of Vibrational Circular Dichroism and Infrared Absorption. *J. Chem. Phys.* 1992, *96*, 5687–5702.
- [68] Abbate, S.; Longhi, G.; Kwon, K.; Moscovitz, A. The Use of Cross-Correlation Functions in the Analysis of Circular Dichroism Spectra. *J. Chem. Phys.* 1998, *108*, 50–62.
- [69] Choi, J.-H.; Cho, M. Direct Calculations of Mid- and near-IR Absorption and Circular Dichroism Spectra of Chiral Molecules Using QM/MM Molecular Dynamics Simulation Method. *J. Chem. Theory Comput.* 2011, *7*, 4097–4103.
- [70] Horníček, J.; Kaprálová, P.; Bouř, P. Simulations of Vibrational Spectra from Classical Trajectories: Calibration with Ab Initio Force Fields. *J. Chem. Phys.* 2007, *127*, –.
- [71] Yang, S.; Cho, M. Direct Calculations of Vibrational Absorption and Circular Dichroism Spectra of Alanine Dipeptide Analog in Water: Quantum Mechanical/Molecular Mechanical Molecular Dynamics Simulations. *J. Chem. Phys.* 2009, *131*, 135102.
- [72] Thomas, M.; Kirchner, B. Classical Magnetic Dipole Moments for the Simulation of Vibrational Circular Dichroism by Ab Initio Molecular Dynamics. *J. Phys. Chem. Lett.* 2016, *7*, 509–513.
- [73] Thomas, M.; Brehm, M.; Fligg, R.; Vohringer, P.; Kirchner, B. Computing Vibrational Spectra from Ab Initio Molecular Dynamics. *Phys. Chem. Chem. Phys.* 2013, *15*, 6608–6622.
- [74] Bannwarth, C.; Grimme, S. A Simplified Time-Dependent Density Functional Theory Approach for Electronic Ultraviolet and Circular Dichroism Spectra of Very Large Molecules. *Comput. Theor. Chem.* 2014, *1040-1041*, 45–53.
- [75] Löwdin, P.-O. On the Non-Orthogonality Problem Connected with the Use of Atomic Wave Functions in the Theory of Molecules and Crystals. *J. Chem. Phys.* 1950, *18*, 365–375.
- [76] Hirschfelder, J. O. Classical and Quantum Mechanical Hypervirial Theorems. *J. Chem. Phys.* 1960, *33*, 1462–1466.

## REFERENCES

- [77] Pollok, C. H.; Riesebeck, T.; Merten, C. Photoisomerization of a Chiral Imine Molecular Switch Followed by Matrix-Isolation VCD Spectroscopy. *Angew. Chem. Int. Ed.* 2017, *56*, 1925–1928.
- [78] Domingos, S. R.; Sanders, H. J.; Hartl, F.; Buma, W. J.; Woutersen, S. Switchable Amplification of Vibrational Circular Dichroism as a Probe of Local Chiral Structure. *Angew. Chem. Int. Ed.* 2014, *53*, 14042–14045.
- [79] Wesolowski, S. S.; Pivonka, D. E. A Rapid Alternative to X-Ray Crystallography for Chiral Determination: Case Studies of Vibrational Circular Dichroism (VCD) to Advance Drug Discovery Projects. *Bioorg. Med. Chem. Lett.* 2013, *23*, 4019–4025.
- [80] Zhu, H.; Yi, J.; Li, M.-Y.; Xiao, J.; Zhang, L.; Yang, C.-W.; Kaindl, R. A.; Li, L.-J.; Wang, Y.; Zhang, X. Observation of Chiral Phonons. *Science* 2018, *359*, 579–582.
- [81] Sato, H.; Tamura, K.; Takimoto, K.; Yamagishi, A. Solid State Vibrational Circular Dichroism towards Molecular Recognition: Chiral Metal Complexes Intercalated in a Clay Mineral. *Phys. Chem. Chem. Phys.* 2018, *20*, 3141–3147.
- [82] Merten, C. Vibrational Optical Activity as Probe for Intermolecular Interactions. *Phys. Chem. Chem. Phys.* 2017, *19*, 18803–18812.
- [83] Choi, H.; Cho, K. J.; Seo, H.; Ahn, J.; Liu, J.; Lee, S. S.; Kim, H.; Feng, C.; Jung, J. H. Transfer and Dynamic Inversion of Coassembled Supramolecular Chirality through 2D-Sheet to Rolled-up Tubular Structure. *J. Am. Chem. Soc.* 2017, *139*, 17711–17714.
- [84] Merten, C.; Pollok, C. H.; Liao, S.; List, B. Stereochemical Communication within a Chiral Ion Pair Catalyst. *Angew. Chem. Int. Ed.* 2015, *54*, 8841–8845.
- [85] Avilés Moreno, J. R.; López González, J. J.; Partal Ureña, F.; Vera, F.; Ros, M. B.; Sierra, T. Study of the Photoinduced Supramolecular Chirality in Columnar Liquid Crystals by Infrared and VCD Spectroscopies. *J. Phys. Chem. B* 2012, *116*, 5090–5096.
- [86] Li, M.-C.; Wang, H.-F.; Chiang, C.-H.; Lee, Y.-D.; Ho, R.-M. Lamellar-Twisting-Induced Circular Dichroism of Chromophore Moieties in Banded Spherulites with Evolution of Homochirality. *Angew. Chem. Int. Ed.* 2014, *53*, 4450–4455.



- [87] Rizzo, P.; Guerra, G. Intense Chiral Optical Phenomena in Racemic Polymers by Cocrystallization with Chiral Guest Molecules: A Brief Overview. *Chirality* 2016, 28, 29–38.
- [88] Wen, T.; Wang, H.-F.; Li, M.-C.; Ho, R.-M. Homochiral Evolution in Self-Assembled Chiral Polymers and Block Copolymers. *Acc. Chem. Res.* 2017, 50, 1011–1021.
- [89] Sadlej, J.; Dobrowolski, J. C.; Rode, J. E. VCD Spectroscopy as a Novel Probe for Chirality Transfer in Molecular Interactions. *Chem. Soc. Rev.* 2010, 39, 1478–1488.
- [90] Perera, A. S.; Cheramy, J.; Merten, C.; Thomas, J.; Xu, Y. IR, Raman, and Vibrational Optical Activity Spectra of Methyl Glycidate in Chloroform and Water: The Clusters-in-a-Liquid Solvation Model. *ChemPhysChem* 2018, 19, 2234–2242.
- [91] Srebro-Hooper, M.; Autschbach, J. Calculating Natural Optical Activity of Molecules from First Principles. *Annu. Rev. Phys. Chem.* 2017, 68, 399–420.
- [92] Bloino, J.; Barone, V. A Second-Order Perturbation Theory Route to Vibrational Averages and Transition Properties of Molecules: General Formulation and Application to Infrared and Vibrational Circular Dichroism Spectroscopies. *J. Chem. Phys.* 2012, 136, 124108.
- [93] Cheeseman, J.; Frisch, M.; Devlin, F.; Stephens, P. Ab Initio Calculation of Atomic Axial Tensors and Vibrational Rotational Strengths Using Density Functional Theory. *Chem. Phys. Lett.* 1996, 252, 211–220.
- [94] Nicu, V. P.; Neugebauer, J.; Wolff, S. K.; Baerends, E. J. A Vibrational Circular Dichroism Implementation within a Slater-Type-Orbital Based Density Functional Framework and Its Application to Hexa- and Hepta-Helicenes. *Theor Chem Account* 2008, 119, 245–263.
- [95] Reiter, K.; Kühn, M.; Weigend, F. Vibrational Circular Dichroism Spectra for Large Molecules and Molecules with Heavy Elements. *The Journal of Chemical Physics* 2017, 146, 054102.
- [96] Marty, R.; Frauenrath, H.; Helbing, J. Aggregates from Perylene Bisimide Oligopeptides as a Test Case for Giant Vibrational Circular Dichroism. *J. Phys. Chem. B* 2014, 118, 11152–11160.
- [97] Perera, A. S.; Cheramy, J.; Poopari, M. R.; Xu, Y. Aggregation of Lactic Acid in Cold Rare-

## REFERENCES

- Gas Matrices and the Link to Solution: A Matrix Isolation-Vibrational Circular Dichroism Study. *Phys. Chem. Chem. Phys.* 2019, 21, 3574–3584.
- [98] Giovannini, T.; Olszowska, M.; Cappelli, C. Effective Fully Polarizable QM/MM Approach to Model Vibrational Circular Dichroism Spectra of Systems in Aqueous Solution. *J. Chem. Theory Comput.* 2016, 12, 5483–5492.
- [99] Bünnemann, K.; Merten, C. Solvation of N,C-Protected Valine: Interactions with DMSO and a Chiral Solvating Agent. *J. Phys. Chem. B* 2016, 120, 9434–9442.
- [100] Giovannini, T.; Del Frate, G.; Lafiosca, P.; Cappelli, C. Effective Computational Route towards Vibrational Optical Activity Spectra of Chiral Molecules in Aqueous Solution. *Phys. Chem. Chem. Phys.* 2018, 20, 9181–9197.
- [101] Bouř, P.; Sopková, J.; Bednářová, L.; Malo, P.; Keiderling, T. A. Transfer of Molecular Property Tensors in Cartesian Coordinates: A New Algorithm for Simulation of Vibrational Spectra. *J. Comput. Chem.* 1997, 18, 646–659.
- [102] Kessler, J.; Andrushchenko, V.; Kapitan, J.; Bouř, P. Insight into Vibrational Circular Dichroism of Proteins by Density Functional Modeling. *Phys. Chem. Chem. Phys.* 2018, 20, 4926–4935.
- [103] Choi, J.-H.; Kim, J.-S.; Cho, M. Amide I Vibrational Circular Dichroism of Polypeptides: Generalized Fragmentation Approximation Method. *J. Chem. Phys.* 2005, 122, 174903.
- [104] Le Barbu-Debus, K.; Scherrer, A.; Bouchet, A.; Sebastiani, D.; Vuilleumier, R.; Zehnacker, A. Effect of Puckering Motion and Hydrogen Bond Formation on the Vibrational Circular Dichroism Spectrum of a Flexible Molecule: The Case of (S)-1-Indanol. *Phys. Chem. Chem. Phys.* 2018, 20, 14635–14646.
- [105] Magyarfalvi, G.; Tarczay, G.; Vass, E. Vibrational Circular Dichroism. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* 2011, 1, 403–425.
- [106] Nicu, V. P. Revisiting an Old Concept: The Coupled Oscillator Model for VCD. Part 2: Implications of the Generalised Coupled Oscillator Mechanism for the VCD Robustness Concept. *Phys. Chem. Chem. Phys.* 2016, 18, 21213–21225.

- [107] Nicu, V. P.; Baerends, E. J. On the Origin Dependence of the Angle Made by the Electric and Magnetic Vibrational Transition Dipole Moment Vectors. *Phys. Chem. Chem. Phys.* 2011, *13*, 16126–16129.
- [108] Gussem, E. D.; Bultinck, P.; Feledziak, M.; Marchand-Brynaert, J.; Stevens, C. V.; Herrebout, W. Vibrational Circular Dichroism versus Optical Rotation Dispersion and Electronic Circular Dichroism for Diastereomers: The Stereochemistry of 3-(1'-Hydroxyethyl)-1-(3'-Phenylpropanoyl)-Azetidin-2-One. *Phys. Chem. Chem. Phys.* 2012, *14*, 8562–8571.
- [109] Nicu, V. P.; Baerends, E. J. Robust Normal Modes in Vibrational Circular Dichroism Spectra. *Phys. Chem. Chem. Phys.* 2009, *11*, 6107–6118.
- [110] Dryzun, C.; Avnir, D. On the Abundance of Chiral Crystals. *Chem. Commun.* 2012, *48*, 5874–5876.
- [111] Pérez-Mellor, A.; Zehnacker, A. Vibrational Circular Dichroism of a 2,5-Diketopiperazine (DKP) Peptide: Evidence for Dimer Formation in Cyclo LL or LD Diphenylalanine in the Solid State. *Chirality* 2017, *29*, 89–96.
- [112] Abbate, S.; Bruhn, T.; Pescitelli, G.; Longhi, G. Vibrational Optical Activity of BODIPY Dimers: The Role of Magnetic–Electric Coupling in Vibrational Excitons. *J. Phys. Chem. A* 2017, *121*, 394–400.
- [113] Nicu, V. P. Revisiting an Old Concept: The Coupled Oscillator Model for VCD. Part 1: The Generalised Coupled Oscillator Mechanism and Its Intrinsic Connection to the Strength of VCD Signals. *Phys. Chem. Chem. Phys.* 2016, *18*, 21202–21212.
- [114] Covington, C. L.; Nicu, V. P.; Polavarapu, P. L. Determination of the Absolute Configurations Using Exciton Chirality Method for Vibrational Circular Dichroism: Right Answers for the Wrong Reasons? *J. Phys. Chem. A* 2015, *119*, 10589–10601.
- [115] Taniguchi, T.; Monde, K. Exciton Chirality Method in Vibrational Circular Dichroism. *J. Am. Chem. Soc.* 2012, *134*, 3695–3698.
- [116] Gulotta, M.; Goss, D. J.; Diem, M. Ir Vibrational CD in Model Deoxyoligonucleotides: Observation of the B → Z Phase Transition and Extended Coupled Oscillator Intensity Calculations. *Biopolymers* 1989, *28*, 2047–2058.

## REFERENCES

- [117] Holzwarth, G.; Chabay, I. Optical Activity of Vibrational Transitions: A Coupled Oscillator Model. *J. Chem. Phys.* 1972, *57*, 1632–1635.
- [118] Measey, T. J.; Schweitzer-Stenner, R. Vibrational Circular Dichroism as a Probe of Fibrillogenesis: The Origin of the Anomalous Intensity Enhancement of Amyloid-like Fibrils. *J. Am. Chem. Soc.* 2011, *133*, 1066–1076.
- [119] Kurouski, D.; Lu, X.; Popova, L.; Wan, W.; Shanmugasundaram, M.; Stubbs, G.; Dukor, R. K.; Lednev, I. K.; Nafie, L. A. Is Supramolecular Filament Chirality the Underlying Cause of Major Morphology Differences in Amyloid Fibrils? *J. Am. Chem. Soc.* 2014, *136*, 2302–2312.
- [120] Keiderling, T. A.; Stephens, P. J. Vibrational Circular Dichroism of Dimethyl Tartrate. A Coupled Oscillator. *J. Am. Chem. Soc.* 1977, *99*, 8061–8062.
- [121] Kubelka, J.; Kim, J.; Bouř, P.; Keiderling, T. A. Contribution of Transition Dipole Coupling to Amide Coupling in IR Spectra of Peptide Secondary Structures. *Vib. Spectrosc.* 2006, *42*, 63–73.
- [122] Kubelka, J.; Keiderling, T. A. Differentiation of  $\beta$ -Sheet-Forming Structures: Ab Initio-Based Simulations of IR Absorption and Vibrational CD for Model Peptide and Protein  $\beta$ -Sheets. *J. Am. Chem. Soc.* 2001, *123*, 12048–12058.
- [123] Pazderková, M.; Pazderka, T.; Shanmugasundaram, M.; Dukor, R. K.; Lednev, I. K.; Nafie, L. A. Origin of Enhanced VCD in Amyloid Fibril Spectra: Effect of Deuteriation and pH. *Chirality* 2017, *29*, 469–475.
- [124] Ellis, R. J.; Pinheiro, T. J. T. Danger – Misfolding Proteins. *Nature* 2002, *416*, 483–484.
- [125] Ma, S.; Cao, X.; Mak, M.; Sadik, A.; Walkner, C.; Freedman, T. B.; Lednev, I. K.; Dukor, R. K.; Nafie, L. A. Vibrational Circular Dichroism Shows Unusual Sensitivity to Protein Fibril Formation and Development in Solution. *J. Am. Chem. Soc.* 2007, *129*, 12364–12365.
- [126] Iyer, A.; Roeters, S. J.; Kogan, V.; Woutersen, S.; Claessens, M. M. A. E.; Subramaniam, V. C-Terminal Truncated  $\alpha$ -Synuclein Fibrils Contain Strongly Twisted  $\beta$ -Sheets. *J. Am. Chem. Soc.* 2017, *139*, 15392–15400.

- [127] Kurouski, D.; Handen, J. D.; Dukor, R. K.; Nafie, L. A.; Lednev, I. K. Supramolecular Chirality in Peptide Microcrystals. *Chem. Commun.* 2015, 51, 89–92.
- [128] Torii, H.; Sato, H. Intermediate Length-Scale Chirality Related to the Vibrational Circular Dichroism Intensity Enhancement upon Fibril Formation in a Gelation Process. *Phys. Chem. Chem. Phys.* 2018, 20, 14992–14996.
- [129] Welch, W. R. W.; Kubelka, J.; Keiderling, T. A. Infrared, Vibrational Circular Dichroism, and Raman Spectral Simulations for  $\beta$ -Sheet Structures with Various Isotopic Labels, Interstrand, and Stacking Arrangements Using Density Functional Theory. *J. Phys. Chem. B* 2013, 117, 10343–10358.
- [130] Shtukenberg, A. G.; Punin, Y. O.; Gujral, A.; Kahr, B. Growth Actuated Bending and Twisting of Single Crystals. *Angew. Chem. Int. Ed.* 2014, 53, 672–699.
- [131] Gonzalez, J. J. L.; Urena, F. P.; Moreno, J. R. A.; Mata, I.; Molins, E.; Claramunt, R. M.; Lopez, C.; Alkorta, I.; Elguero, J. The Chiral Structure of 1H-Indazoles in the Solid State: A Crystallographic, Vibrational Circular Dichroism and Computational Study. *New J. Chem.* 2012, 36, 749–758.
- [132] Heyden, M.; Sun, J.; Funkner, S.; Mathias, G.; Forbert, H.; Havenith, M.; Marx, D. Dissecting the THz Spectrum of Liquid Water from First Principles via Correlations in Time and Space. *Proc. Natl. Acad. Sci.* 2010, 107, 12068–12073.
- [133] Heyden, M.; Sun, J.; Forbert, H.; Mathias, G.; Havenith, M.; Marx, D. Understanding the Origins of Dipolar Couplings and Correlated Motion in the Vibrational Spectrum of Water. *J. Phys. Chem. Lett.* 2012, 3, 2135–2140.
- [134] Wilson, C. C.; Myles, D.; Ghosh, M.; Johnson, L. N.; Wang, W. Neutron Diffraction Investigations of L- and D-Alanine at Different Temperatures: The Search for Structural Evidence for Parity Violation. *New J. Chem.* 2005, 29, 1318–1322.
- [135] Nafie, L. A.; Dukor, R. K. In *Chiral Analysis*; Busch, K. W., Busch, M. A., Eds.; Elsevier: Amsterdam, 2006; pp 505–544.
- [136] Martinez, M.; Gaigeot, M.-P.; Borgis, D.; Vuilleumier, R. Extracting Effective Normal Modes from Equilibrium Dynamics at Finite Temperature. *J. Chem. Phys.* 2006, 125, 144106.

## REFERENCES

- [137] Chowdhry, B. Z.; Dines, T. J.; Jabeen, S.; Withnall, R. Vibrational Spectra of  $\alpha$ -Amino Acids in the Zwitterionic State in Aqueous Solution and the Solid State: DFT Calculations and the Influence of Hydrogen Bonding. *J. Phys. Chem. A* 2008, *112*, 10333–10347.
- [138] Freedman, T. B.; Chernovitz, A. C.; Zuk, W. M.; Paterlini, M. G.; Nafe, L. A. Vibrational Circular Dichroism in the Methine Bending Modes of Amino Acids and Dipeptides. *J. Am. Chem. Soc.* 1988, *110*, 6970–6974.
- [139] Tasumi, M.; Shimanouchi, T. Crystal Vibrations and Intermolecular Forces of Polymethylene Crystals. *J. Chem. Phys.* 1965, *43*, 1245–1258.
- [140] Losada, M.; Tran, H.; Xu, Y. Lactic Acid in Solution: Investigations of Lactic Acid Self-Aggregation and Hydrogen Bonding Interactions with Water and Methanol Using Vibrational Absorption and Vibrational Circular Dichroism Spectroscopies. *J. Chem. Phys.* 2008, *128*, 014508.
- [141] Cappelli, C.; Bloino, J.; Lipparini, F.; Barone, V. Toward Ab Initio Anharmonic Vibrational Circular Dichroism Spectra in the Condensed Phase. *J. Phys. Chem. Lett.* 2012, *3*, 1766–1773.
- [142] Losada, M.; Xu, Y. Chirality Transfer through Hydrogen-Bonding: Experimental and Ab Initio Analyses of Vibrational Circular Dichroism Spectra of Methyl Lactate in Water. *Phys. Chem. Chem. Phys.* 2007, *9*, 3127–3135.
- [143] Losada, M.; Nguyen, P.; Xu, Y. Solvation of Propylene Oxide in Water: Vibrational Circular Dichroism, Optical Rotation, and Computer Simulation Studies. *J. Phys. Chem. A* 2008, *112*, 5621–5627.
- [144] Yang, G.; Xu, Y. Probing Chiral Solute-Water Hydrogen Bonding Networks by Chirality Transfer Effects: A Vibrational Circular Dichroism Study of Glycidol in Water. *J. Chem. Phys.* 2009, *130*, 164506.
- [145] Merten, C.; Xu, Y. Chirality Transfer in a Methyl Lactate-Ammonia Complex Observed by Matrix-Isolation Vibrational Circular Dichroism Spectroscopy. *Angew. Chem. Int. Ed.* 2013, *52*, 2073–2076.
- [146] Bünnemann, K.; Pollok, C. H.; Merten, C. Explicit Solvation of Carboxylic Acids for Vi-

- brational Circular Dichroism Studies: Limiting the Computational Efforts without Losing Accuracy. *J. Phys. Chem. B* 2018, *122*, 8056–8064.
- [147] Auer, B. M.; Skinner, J. L. IR and Raman Spectra of Liquid Water: Theory and Interpretation. *J. Chem. Phys.* 2008, *128*, 224511.
- [148] Bünnemann, K.; Merten, C. Solvation of a Chiral Carboxylic Acid: Effects of Hydrogen Bonding on the IR and VCD Spectra of  $\alpha$ -Methoxyphenylacetic Acid. *Phys. Chem. Chem. Phys.* 2017, *19*, 18948–18956.
- [149] Sadlej, J.; Dobrowolski, J. C.; Rode, J. E.; Jamróz, M. H. DFT Study of Vibrational Circular Dichroism Spectra of D-Lactic Acid–Water Complexes. *Phys. Chem. Chem. Phys.* 2006, *8*, 101–113.
- [150] Zhang, M.; Chen, L.; Yang, H.; Ma, J. Theoretical Study of Acetic Acid Association Based on Hydrogen Bonding Mechanism. *J. Phys. Chem. A* 2017, *121*, 4560–4568.
- [151] Kubelka, J.; Huang, R.; Keiderling, T. A. Solvent Effects on IR and VCD Spectra of Helical Peptides: DFT-Based Static Spectral Simulations with Explicit Water. *J. Phys. Chem. B* 2005, *109*, 8231–8243.
- [152] Max, J.-J.; Chapados, C. Infrared Spectroscopy of Aqueous Carboxylic Acids: Comparison between Different Acids and Their Salts. *J. Phys. Chem. A* 2004, *108*, 3324–3337.
- [153] Clark, G. N.; Cappa, C. D.; Smith, J. D.; Saykally, R. J.; Head-Gordon, T. The Structure of Ambient Water. *Mol. Phys.* 2010, *108*, 1415–1433.
- [154] Gowers, R. J.; Carbone, P. A Multiscale Approach to Model Hydrogen Bonding: The Case of Polyamide. *J. Chem. Phys.* 2015, *142*, 224907.
- [155] Senge, M. O.; MacGowan, S. A.; O'Brien, J. M. Conformational Control of Cofactors in Nature - the Influence of Protein-Induced Macrocyclic Distortion on the Biological Function of Tetrapyrroles. *Chem. Commun.* 2015, *51*, 17031–17063.
- [156] Kadish, K., Smith, K., Guillard, R., Eds. *The Porphyrin Handbook*; San Diego: Academic, 2000.

## REFERENCES

- [157] Sarovar, M.; Ishizaki, A.; Fleming, G. R.; Whaley, K. B. Quantum Entanglement in Photosynthetic Light-Harvesting Complexes. *Nat. Phys.* 2010, 6, 462.
- [158] Grossman, A. R.; Schaefer, M. R.; Chiang, G. G.; Collier, J. L. The Phycobilisome, a Light-Harvesting Complex Responsive to Environmental Conditions. *Microbiol. Rev.* 1993, 57, 725–749.
- [159] Peng, P.-P.; Dong, L.-L.; Sun, Y.-F.; Zeng, X.-L.; Ding, W.-L.; Scheer, H.; Yang, X.; Zhao, K.-H. The Structure of Allophycocyanin B from *Synechocystis* PCC 6803 Reveals the Structural Basis for the Extreme Redshift of the Terminal Emitter in Phycobilisomes. *Acta Crystallogr. Sect. D* 2014, 70, 2558–2569.
- [160] Rockwell, N. C.; Su, Y.-S.; Lagarias, J. C. Phytochrome Structure and Signaling Mechanisms. *Annu. Rev. Plant Biol.* 2006, 57, 837–58.
- [161] Butler, W. L.; Norris, K. H.; Siegelman, H. W.; Hendricks, S. B. Detection, Assay, and Preliminary Purification of the Pigment Controlling Photoresponsive Development of Plants. *Proc. Natl. Acad. Sci.* 1959, 45, 1703–1708.
- [162] Sage, L. C. *Pigment of the Imagination: A History of Phytochrome Research*; San Diego: Academic, 1992.
- [163] Tu, S.-L.; Lagarias, J. C. *Handbook of Photosensory Receptors*; Wiley-VCH Verlag GmbH, 2005; pp 121–149.
- [164] Huq, E.; Quail, P. H. *Handbook of Photosensory Receptors*; Wiley-VCH Verlag GmbH, 2005; pp 151–170.
- [165] Vierstra, R. D.; Karniol, B. *Handbook of Photosensory Receptors*; Wiley-VCH Verlag GmbH, 2005; pp 171–195.
- [166] Yeh, K.-C.; Wu, S.-H.; Murphy, J. T.; Lagarias, J. C. A Cyanobacterial Phytochrome Two-Component Light Sensory System. *Science* 1997, 277, 1505–1508.
- [167] Schmidt, A.; Sauthof, L.; Szczepek, M.; Lopez, M. F.; Escobar, F. V.; Qureshi, B. M.; Michael, N.; Buhrke, D.; Stevens, T.; Kwiatkowski, D.; von Stetten, D.; Mroginski, M. A.;



- Krauß, N.; Lamparter, T.; Hildebrandt, P.; Scheerer, P. Structural Snapshot of a Bacterial Phytochrome in Its Functional Intermediate State. *Nat. Commun.* 2018, 9, 4912.
- [168] Karniol, B.; Wagner, J. R.; Walker, J. M.; Vierstra, R. D. Phylogenetic Analysis of the Phytochrome Superfamily Reveals Distinct Microbial Subfamilies of Photoreceptors. *Biochem. J.* 2005, 392, 103–116.
- [169] Essen, L.-O.; Mailliet, J.; Hughes, J. The Structure of a Complete Phytochrome Sensory Module in the Pr Ground State. *Proc. Natl. Acad. Sci.* 2008, 105, 14709–14714.
- [170] Narikawa, R.; Fukushima, Y.; Ishizuka, T.; Itoh, S.; Ikeuchi, M. A Novel Photoactive GAF Domain of Cyanobacteriochrome AnPixJ That Shows Reversible Green/Red Photoconversion. *J. Mol. Biol.* 2008, 380, 844–855.
- [171] Ikeuchi, M.; Ishizuka, T. Cyanobacteriochromes: A New Superfamily of Tetrapyrrole-Binding Photoreceptors in Cyanobacteria. *Photochem. Photobiol. Sci.* 2008, 7, 1159–1167.
- [172] Rockwell, N. C.; Martin, S. S.; Lagarias, J. C. Red/Green Cyanobacteriochromes: Sensors of Color and Power. *Biochemistry* 2012, 51, 9667–9677.
- [173] Suzuki, F.; Fujita, H.; Ikeuchi, M.; Yoshihara, S.; Geng, X. X. Novel Putative Photoreceptor and Regulatory Genes Required for the Positive Phototactic Movement of the Unicellular Motile Cyanobacterium *Synechocystis* Sp. PCC 6803. *Plant Cell Physiol.* 2000, 41, 1299–1304.
- [174] Wiebeler, C.; Rao, A. G.; Gärtner, W.; Schapiro, I. The Effective Conjugation Length Is Responsible for the Red/Green Spectral Tuning in the Cyanobacteriochrome Slr1393g3. *Angew. Chem. Int. Ed.* 2019, 58, 1934–1938.
- [175] Rockwell, N. C.; Martin, S. S.; Feoktistova, K.; Lagarias, J. C. Diverse Two-Cysteine Photocycles in Phytochromes and Cyanobacteriochromes. *Proc. Natl. Acad. Sci.* 2011, 108, 11854–11859.
- [176] Enomoto, G.; Hirose, Y.; Narikawa, R.; Ikeuchi, M. Thiol-Based Photocycle of the Blue and Teal Light-Sensing Cyanobacteriochrome Tlr1999. *Biochemistry* 2012, 51, 3050–3058.
- [177] Rockwell, N. C.; Martin, S. S.; Gulevich, A. G.; Lagarias, J. C. Phycoviolobilin Formation

## REFERENCES

- and Spectral Tuning in the DXCF Cyanobacteriochrome Subfamily. *Biochemistry* 2012, 51, 1449–1463.
- [178] Rockwell, N. C.; Martin, S. S.; Gulevich, A. G.; Lagarias, J. C. Conserved Phenylalanine Residues Are Required for Blue-Shifting of Cyanobacteriochrome Photoproducts. *Biochemistry* 2014, 53, 3118–3130.
- [179] Sekharan, S.; Sugihara, M.; Buss, V. Origin of Spectral Tuning in Rhodopsin—It Is Not the Binding Pocket. *Angew. Chem. Int. Ed.* 2007, 46, 269–271.
- [180] Rohmer, T.; Lang, C.; Gärtner, W.; Hughes, J.; Matysik, J. Role of the Protein Cavity in Phytochrome Chromoprotein Assembly and Double-Bond Isomerization: A Comparison with Model Compounds. *Photochem. Photobiol.* 2010, 86, 856–861.
- [181] Scheer, H.; Yang, X.; Zhao, K.-H. Biliproteins and Their Applications in Bioimaging. *Procedia Chem.* 2015, 14, 176–185.
- [182] Lamparter, T. Evolution of Cyanobacterial and Plant Phytochromes. *FEBS Lett.* 2004, 573, 1–5.
- [183] Narikawa, R.; Ishizuka, T.; Muraki, N.; Shiba, T.; Kurisu, G.; Ikeuchi, M. Structures of Cyanobacteriochromes from Phototaxis Regulators AnPixJ and TePixJ Reveal General and Specific Photoconversion Mechanism. *Proc. Natl. Acad. Sci.* 2013, 110, 918–923.
- [184] Rüdiger, W.; Thümmler, F.; Cmiel, E.; Schneider, S. Chromophore Structure of the Physiologically Active Form (P(Fr)) of Phytochrome. *Proc. Natl. Acad. Sci.* 1983, 80, 6244–6248.
- [185] Song, C.; Rohmer, T.; Tiersch, M.; Zaanen, J.; Hughes, J.; Matysik, J. Solid-State NMR Spectroscopy to Probe Photoactivation in Canonical Phytochromes. *Photochem. Photobiol.* 2013, 89, 259–273.
- [186] Rohmer, T.; Lang, C.; Bongards, C.; Gupta, K. B. S. S.; Neugebauer, J.; Hughes, J.; Gärtner, W.; Matysik, J. Phytochrome as Molecular Machine: Revealing Chromophore Action during the Pfr – Pr Photoconversion by Magic-Angle Spinning NMR Spectroscopy. *J. Am. Chem. Soc.* 2010, 132, 4431–4437.
- [187] Heyes, D. J.; Khara, B.; Sakuma, M.; Hardman, S. J. O.; O’Cualain, R.; Rigby, S. E. J.; Scrut-

- ton, N. S. Ultrafast Red Light Activation of Synechocystis Phytochrome Cph1 Triggers Major Structural Change to Form the Pfr Signalling-Competent State. *PLoS One* 2012, 7, e52418.
- [188] Song, C.; Psakis, G.; Lang, C.; Mailliet, J.; Gärtner, W.; Hughes, J.; Matysik, J. Two Ground State Isoforms and a Chromophore D-Ring Photoflip Triggering Extensive Intramolecular Changes in a Canonical Phytochrome. *Proc. Natl. Acad. Sci.* 2011, 108, 3842–3847.
- [189] Song, C.; Psakis, G.; Kopycki, J.; Lang, C.; Matysik, J.; Hughes, J. The D-Ring, Not the a-Ring, Rotates in Synechococcus OS-B Phytochrome. *J. Biol. Chem.* 2014, 289, 2552–2562.
- [190] Hirose, Y.; Rockwell, N. C.; Nishiyama, K.; Narikawa, R.; Ukaji, Y.; Inomata, K.; Lagarias, J. C.; Ikeuchi, M. Green/Red Cyanobacteriochromes Regulate Complementary Chromatic Acclimation via a Protochromic Photocycle. *Proc. Natl. Acad. Sci.* 2013, 110, 4974–4979.
- [191] Wu, X.-J.; Yang, H.; Sheng, Y.; Zhu, Y.-L.; Li, P.-P. Fluorescence Properties of a Novel Cyanobacteriochrome GAF Domain from Spirulina That Exhibits Moderate Dark Reversion. *Int. J. Mol. Sci.* 2018, 19, 2253.
- [192] Zhao, K.-H.; Haessner, R.; Cmiel, E.; Scheer, H. Type I Reversible Photochemistry of Phycocerythrocyanin Involves Z/E-Isomerization of  $\alpha$ -84 Phycoviolobilin Chromophore. *Biochim. Biophys. Acta BBA - Bioenerg.* 1995, 1228, 235–243.
- [193] Fukushima, Y.; Iwaki, M.; Narikawa, R.; Ikeuchi, M.; Tomita, Y.; Itoh, S. Photoconversion Mechanism of a Green/Red Photosensory Cyanobacteriochrome AnPixJ: Time-Resolved Optical Spectroscopy and FTIR Analysis of the AnPixJ-GAF2 Domain. *Biochemistry* 2011, 50, 6328–6339.
- [194] Rockwell, N. C.; Martin, S. S.; Lagarias, J. C. Mechanistic Insight into the Photosensory Versatility of DXCF Cyanobacteriochromes. *Biochemistry* 2012, 51, 3576–3585.
- [195] Velazquez Escobar, F.; Utesch, T.; Narikawa, R.; Ikeuchi, M.; Mroginski, M. A.; Gärtner, W.; Hildebrandt, P. Photoconversion Mechanism of the Second GAF Domain of Cyanobacteriochrome AnPixJ and the Cofactor Structure of Its Green-Absorbing State. *Biochemistry* 2013, 52, 4871–4880.
- [196] Lim, S.; Rockwell, N. C.; Martin, S. S.; Dallas, J. L.; Lagarias, J. C.; Ames, J. B. Photocon-

## REFERENCES

- version Changes Bilin Chromophore Conjugation and Protein Secondary Structure in the Violet/Orange Cyanobacteriochrome NpF2163g3. *Photochem. Photobiol. Sci.* 2014, 13, 951–962.
- [197] Song, C.; Narikawa, R.; Ikeuchi, M.; Gärtner, W.; Matysik, J. Color Tuning in Red/Green Cyanobacteriochrome AnPixJ: Photoisomerization at C<sub>15</sub> Causes an Excited-State Destabilization. *J. Phys. Chem. B* 2015, 119, 9688–9695.
- [198] Schäfer, E.; Nagy, F. *Handbook of Photosensory Receptors*; Wiley-VCH Verlag GmbH, 2005; pp 197–210.
- [199] Velazquez Escobar, F.; Hildebrandt, T.; Utesch, T.; Schmitt, F. J.; Seuffert, I.; Michael, N.; Schulz, C.; Mroginski, M. A.; Friedrich, T.; Hildebrandt, P. Structural Parameters Controlling the Fluorescence Properties of Phytochromes. *Biochemistry* 2014, 53, 20–29.
- [200] Rockwell, N. C.; Ohlendorf, R.; Möglich, A. Cyanobacteriochromes in Full Color and Three Dimensions. *Proc. Natl. Acad. Sci.* 2013, 110, 806–807.
- [201] Rumyantsev, K. A.; Shcherbakova, D. M.; Zakharova, N. I.; Verkhusha, V. V.; Tur-overov, K. K. Design of Near-Infrared Single-Domain Fluorescent Protein GAF-FP Based on Bacterial Phytochrome. *Cell Tissue Biol.* 2017, 11, 16–26.
- [202] Shu, X.; Royant, A.; Lin, M. Z.; Aguilera, T. A.; Lev-Ram, V.; Steinbach, P. A.; Tsien, R. Y. Mammalian Expression of Infrared Fluorescent Proteins Engineered from a Bacterial Phytochrome. *Science* 2009, 324, 804–807.
- [203] Zhang, J.; Wu, X.-J.; Wang, Z.-B.; Chen, Y.; Wang, X.; Zhou, M.; Scheer, H.; Zhao, K.-H. Fused-Gene Approach to Photoswitchable and Fluorescent Biliproteins. *Angew. Chem. Int. Ed.* 2010, 49, 5456–5458.
- [204] Simon, J.; Losi, A.; Zhao, K.-H.; Gärtner, W. FRET in a Synthetic Flavin- and Bilin-Binding Protein. *Photochem. Photobiol.* 2017, 93, 1057–1062.
- [205] Piatkevich, K. D.; Subach, F. V.; Verkhusha, V. V. Engineering of Bacterial Phytochromes for Near-Infrared Imaging, Sensing, and Light-Control in Mammals. *Chem. Soc. Rev.* 2013, 42, 3441–3452.

- [206] Hou, Y.-N.; Ding, W.-L.; Jiang, S.-P.; Miao, D.; Tan, Z.-Z.; Hu, J.-L.; Scheer, H.; Zhao, K.-H. Bright Near-Infrared Fluorescence Bio-Labeling with a Biliprotein Triad. *Biochim Biophys Acta Bioenerg - Mol Cell Res* 2019, 1866, 277–284.
- [207] Filonov, G. S.; Piatkevich, K. D.; Ting, L.-M.; Zhang, J.; Kim, K.; Verkhusha, V. V. Bright and Stable Near-Infrared Fluorescent Protein for in Vivo Imaging. *Nat. Biotechnol.* 2011, 29, 757.
- [208] Ryu, M.-H.; Gomelsky, M. Near-Infrared Light Responsive Synthetic c-Di-GMP Module for Optogenetic Applications. *ACS Synth. Biol.* 2014, 3, 802–810.
- [209] Shimizu-Sato, S.; Huq, E.; Tepperman, J. M.; Quail, P. H. A Light-Switchable Gene Promoter System. *Nat. Biotechnol.* 2002, 20, 1041.
- [210] Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.; Hirose, M.; Sugiura, M.; Sasamoto, S.; Kimura, T.; Hosouchi, T.; Matsuno, A.; Muraki, A.; Nakazaki, N.; Naruo, K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda, M.; Tabata, S. Sequence Analysis of the Genome of the Unicellular Cyanobacterium *Synechocystis* Sp. Strain PCC6803. II. Sequence Determination of the Entire Genome and Assignment of Potential Protein-Coding Regions. *DNA Res.* 1996, 3, 109–36.
- [211] Hughes, J.; Lamparter, T.; Mittmann, F.; Hartmann, E.; Gärtner, W.; Wilde, A.; Börner, T. A Prokaryotic Phytochrome. *Nature* 1997, 386, 663–663.
- [212] Hahn, J.; Strauss, H. M.; Landgraf, F. T.; Giménez, H. F.; Lochnit, G.; Schmieder, P.; Hughes, J. Probing Protein–Chromophore Interactions in Cph1 Phytochrome by Mutagenesis. *FEBS J.* 2006, 273, 1415–1429.
- [213] Takeda, K.; Terazima, M. Photoinduced Orientation Change of the Dimer Structure of the Pr-i State of Cph1Δ2. *Biochemistry* 2018, 57, 5058–5071.
- [214] Velazquez Escobar, F.; Lang, C.; Takiden, A.; Schneider, C.; Balke, J.; Hughes, J.; Alexiev, U.; Hildebrandt, P.; Mroginiski, M. A. Protonation-Dependent Structural Heterogeneity in the Chromophore Binding Site of Cyanobacterial Phytochrome Cph1. *J. Phys. Chem. B* 2017, 121, 47–57.

## REFERENCES

- [215] Stöppler, D.; Song, C.; van Rossum, B.-J.; Geiger, M.-A.; Lang, C.; Mroginski, M.-A.; Jagtap, A. P.; Sigurdsson, S. T.; Matysik, J.; Hughes, J.; Oschkinat, H. Dynamic Nuclear Polarization Provides New Insights into Chromophore Structure in Phytochrome Photoreceptors. *Angew. Chem. Int. Ed.* 2016, *55*, 16017–16020.
- [216] Yang, Y.; Heyne, K.; Mathies, R. A.; Dasgupta, J. Non-Bonded Interactions Drive the Sub-Picosecond Bilin Photoisomerization in the Pfr State of Phytochrome Cph1. *ChemPhysChem* 2016, *17*, 369–374.
- [217] Kim, P. W.; Rockwell, N. C.; Martin, S. S.; Lagarias, J. C.; Larsen, D. S. Heterogeneous Photodynamics of the Pfr State in the Cyanobacterial Phytochrome Cph1. *Biochemistry* 2014, *53*, 4601–4611.
- [218] Kim, P. W.; Pan, J.; Rockwell, N. C.; Chang, C.-W.; Taylor, K. C.; Clark Lagarias, J.; Larsen, D. S. Ultrafast E to Z Photoisomerization Dynamics of the Cph1 Phytochrome. *Chem. Phys. Lett.* 2012, *549*, 86–92.
- [219] Spillane, K. M.; Dasgupta, J.; Mathies, R. A. Conformational Homogeneity and Excited-State Isomerization Dynamics of the Bilin Chromophore in Phytochrome Cph1 from Resonance Raman Intensities. *Biophys. J.* 2012, *102*, 709–717.
- [220] Mailliet, J.; Psakis, G.; Feilke, K.; Sineshchekov, V.; Essen, L.-O.; Hughes, J. Spectroscopy and a High-Resolution Crystal Structure of Tyr263 Mutants of Cyanobacterial Phytochrome Cph1. *J. Mol. Biol.* 2011, *413*, 115–127.
- [221] Song, C.; Psakis, G.; Lang, C.; Mailliet, J.; Zaanen, J.; Gärtner, W.; Hughes, J.; Matysik, J. On the Collective Nature of Phytochrome Photoactivation. *Biochemistry* 2011, *50*, 10987–10989.
- [222] Mroginski, M. A.; von Stetten, D.; Escobar, F. V.; Strauss, H. M.; Kaminski, S.; Scheerer, P.; Günther, M.; Murgida, D. H.; Schmieder, P.; Bongards, C.; Gärtner, W.; Mailliet, J.; Hughes, J.; Essen, L.-O.; Hildebrandt, P. Chromophore Structure of Cyanobacterial Phytochrome Cph1 in the Pr State: Reconciling Structural and Spectroscopic Data by QM/MM Calculations. *Biophys. J.* 2009, *96*, 4153–4163.
- [223] Rockwell, N. C.; Shang, L.; Martin, S. S.; Lagarias, J. C. Distinct Classes of Red/Far-Red Photochemistry within the Phytochrome Superfamily. *Proc. Natl. Acad. Sci.* 2009, *106*, 6123–6127.

- [224] Hahn, J.; Strauss, H. M.; Schmieder, P. Heteronuclear NMR Investigation on the Structure and Dynamics of the Chromophore Binding Pocket of the Cyanobacterial Phytochrome Cph1. *J. Am. Chem. Soc.* 2008, *130*, 11170–11178.
- [225] Matute, R. A.; Contreras, R.; Pérez-Hernández, G.; González, L. The Chromophore Structure of the Cyanobacterial Phytochrome Cph1 as Predicted by Time-Dependent Density Functional Theory. *J. Phys. Chem. B* 2008, *112*, 16253–6.
- [226] Rohmer, T.; Lang, C.; Hughes, J.; Essen, L.-O.; Gärtner, W.; Matysik, J. Light-Induced Chromophore Activity and Signal Transduction in Phytochromes Observed by  $^{13}\text{C}$  and  $^{15}\text{N}$  Magic-Angle Spinning NMR. *Proc. Natl. Acad. Sci.* 2008, *105*, 15229–15234.
- [227] Murgida, D. H.; von Stetten, D.; Hildebrandt, P.; Schwinté, P.; Siebert, F.; Sharda, S.; Gärtner, W.; Mroginiski, M. A. The Chromophore Structures of the Pr States in Plant and Bacterial Phytochromes. *Biophys. J.* 2007, *93*, 2410–2417.
- [228] Rockwell, N. C.; Lagarias, J. C. The Structure of Phytochrome: A Picture Is Worth a Thousand Spectra. *Plant Cell* 2006, *18*, 4–14.
- [229] Rohmer, T.; Strauss, H.; Hughes, J.; de Groot, H.; Gärtner, W.; Schmieder, P.; Matysik, J.  $^{15}\text{N}$  MAS NMR Studies of Cph1 Phytochrome: Chromophore Dynamics and Intramolecular Signal Transduction. *J. Phys. Chem. B* 2006, *110*, 20580–20585.
- [230] Xu, X.-L.; Gutt, A.; Mechelke, J.; Raffelberg, S.; Tang, K.; Miao, D.; Valle, L.; Borsarelli, C. D.; Zhao, K.-H.; Gärtner, W. Combined Mutagenesis and Kinetics Characterization of the Bilin-Binding GAF Domain of the Protein Slr1393 from the Cyanobacterium *Synechocystis* PCc6803. *Chembiochem* 2014, *15*, 1190–1199.
- [231] Song, C.; Velazquez Escobar, F.; Xu, X.-L.; Narikawa, R.; Ikeuchi, M.; Siebert, F.; Gärtner, W.; Matysik, J.; Hildebrandt, P. A Red/Green Cyanobacteriochrome Sustains Its Color despite a Change in the Bilin Chromophores Protonation State. *Biochemistry* 2015, *54*, 5839–5848.
- [232] Song, C.; Lang, C.; Mailliet, J.; Hughes, J.; Gärtner, W.; Matysik, J. Exploring Chromophore-Binding Pocket: High-Resolution Solid-State  $^1\text{H}$ – $^{13}\text{C}$  Interfacial Correlation NMR Spectra with Windowed PMLG Scheme. *Appl. Magn. Reson.* 2012, *42*, 79–88.

## REFERENCES

- [233] Rohmer, T.; Matysik, J.; Mark, F. Solvation and Crystal Effects in Bilirubin Studied by NMR Spectroscopy and Density Functional Theory. *J. Phys. Chem. A* 2011, 115, 11696–11714.
- [234] Ulijasz, A. T.; Cornilescu, G.; Cornilescu, C. C.; Zhang, J.; Rivera, M.; Markley, J. L.; Vierstra, R. D. Structural Basis for the Photoconversion of a Phytochrome to the Activated Far-Red Light-Absorbing Form. *Nature* 2010, 463, 250–254.
- [235] Song, C.; Mroginski, M. A.; Lang, C.; Kopycki, J.; Gärtner, W.; Matysik, J.; Hughes, J. 3D Structures of Plant Phytochrome A as Pr and Pfr from Solid-State NMR: Implications for Molecular Function. *Front. Plant Sci.* 2018, 9, 498.
- [236] Elgabarty, H.; Schmieder, P.; Sebastiani, D. Unraveling the Existence of Dynamic Water Channels in Light-Harvesting Proteins: Alpha-C-Phycocyanobilin in Vitro. *Chem. Sci.* 2013, 4, 755–763.
- [237] Watermann, T.; Elgabarty, H.; Sebastiani, D. Phycocyanobilin in Solution - a Solvent Triggered Molecular Switch. *Phys. Chem. Chem. Phys.* 2014, 16, 6146–6152.
- [238] Billeter, M.; Vendrell, J.; Wider, G.; Avilés, F. X.; Coll, M.; Guasch, A.; Huber, R.; Wüthrich, K. Comparison of the NMR Solution Structure with the X-Ray Crystal Structure of the Activation Domain from Procarboxypeptidase B. *J. Biomol. NMR* 1992, 2, 1–10.
- [239] Snyder, D. A.; Chen, Y.; Denissova, N. G.; Acton, T.; Aramini, J. M.; Ciano, M.; Karlin, R.; Liu, J.; Manor, P.; Rajan, P. A.; Rossi, P.; Swapna, G. V. T.; Xiao, R.; Rost, B.; Hunt, J.; Montelione, G. T. Comparisons of NMR Spectral Quality and Success in Crystallization Demonstrate That NMR and X-Ray Crystallography Are Complementary Methods for Small Protein Structure Determination. *J. Am. Chem. Soc.* 2005, 127, 16505–16511.
- [240] Kratky, C.; Waditschatka, R.; Angst, C.; Johansen, J. E.; Plaquevent, J. C.; Schreiber, J.; Eschenmoser, A. Die Sattelkonformation Der Hydroporphinoiden Nickel(II)-Komplexe: Struktur, Ursprung Und Stereochemische Konsequenzen. *Helv. Chim. Acta* 1985, 68, 1312–1337.
- [241] Kim, P. W.; Freer, L. H.; Rockwell, N. C.; Martin, S. S.; Lagarias, J. C.; Larsen, D. S. Femtosecond Photodynamics of the Red/Green Cyanobacteriochrome NpR6012g4 from *Nostoc punctiforme*. I. Forward Dynamics. *Biochemistry* 2012, 51, 608–618.



- [242] Gottlieb, S. M.; Kim, P. W.; Rockwell, N. C.; Hirose, Y.; Ikeuchi, M.; Lagarias, J. C.; Larsen, D. S. Primary Photodynamics of the Green/Red-Absorbing Photoswitching Regulator of the Chromatic Adaptation E Domain from *Fremyella Diplosiphon*. *Biochemistry* 2013, 52, 8198–8208.
- [243] von Stetten, D.; Günther, M.; Scheerer, P.; Murgida, D. H.; Mroginski, M. A.; Krauß, N.; Lamparter, T.; Zhang, J.; Anstrom, D. M.; Vierstra, R. D.; Forest, K. T.; Hildebrandt, P. Chromophore Heterogeneity and Photoconversion in Phytochrome Crystals and Solution Studied by Resonance Raman Spectroscopy. *Angew. Chem. Int. Ed.* 2008, 47, 4753–4755.
- [244] Weigend, F.; Ahlrichs, R. Balanced Basis Sets of Split Valence, Triple Zeta Valence and Quadruple Zeta Valence Quality for H to Rn: Design and Assessment of Accuracy. *Phys. Chem. Chem. Phys.* 2005, 7, 3297–3305.
- [245] Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. A Consistent and Accurate Ab Initio Parametrization of Density Functional Dispersion Correction (DFT-D) for the 94 Elements H-Pu. *J. Chem. Phys.* 2010, 132, 154104.
- [246] Tkatchenko, A.; DiStasio, R. A.; Head-Gordon, M.; Scheffler, M. Dispersion-Corrected Møller–Plesset Second-Order Perturbation Theory. *J. Chem. Phys.* 2009, 131, 094106.
- [247] Boys, S.; Bernardi, F. The Calculation of Small Molecular Interactions by the Differences of Separate Total Energies. Some Procedures with Reduced Errors. *Mol. Phys.* 1970, 19, 553–566.
- [248] Noack, S.; Michael, N.; Rosen, R.; Lamparter, T. Protein Conformational Changes of *Agrobacterium* Phytochrome Agp1 during Chromophore Assembly and Photoconversion. *Biochemistry* 2007, 46, 4164–4176.
- [249] Mulliken, R. S. Electronic Population Analysis on LCAO–MO Molecular Wave Functions. I. *J. Chem. Phys.* 1955, 23, 1833–1840.
- [250] Bachovchin, W. W.; Wong, W. Y. L.; Farr-Jones, S.; Shenvi, A. B.; Kettner, C. A. Nitrogen-15 NMR Spectroscopy of the Catalytic-Triad Histidine of a Serine Protease in Peptide Boronic Acid Inhibitor Complexes. *Biochemistry* 1988, 27, 7689–7697.
- [251] Altoè, P.; Climent, T.; De Fusco, G. C.; Stenta, M.; Bottoni, A.; Serrano-Andrés, L.;

## REFERENCES

- Merchán, M.; Orlandi, G.; Garavelli, M. Deciphering Intrinsic Deactivation/Isomerization Routes in a Phytochrome Chromophore Model. *J. Phys. Chem. B* 2009, *113*, 15067–15073.
- [252] Falklöf, O.; Durbeej, B. Red-Light Absorption and Fluorescence of Phytochrome Chromophores: A Comparative Theoretical Study. *Chem. Phys.* 2013, *425*, 19–28.
- [253] Falklöf, O.; Durbeej, B. Modeling of Phytochrome Absorption Spectra. *J. Comput. Chem.* 2013, *34*, 1363–1374.
- [254] Yanai, T.; Tew, D. P.; Handy, N. C. A New Hybrid Exchange–Correlation Functional Using the Coulomb-Attenuating Method (CAM-B3LYP). *Chem. Phys. Lett.* 2004, *393*, 51–57.
- [255] Matute, R. A.; Contreras, R.; González, L. Time-Dependent DFT on Phytochrome Chromophores: A Way to the Right Conformer. *J. Phys. Chem. Lett.* 2010, *1*, 796–801.
- [256] Gouterman, M.; Wagnière, G. H.; Snyder, L. C. Spectra of Porphyrins: Part II. Four Orbital Model. *J. Mol. Spectrosc.* 1963, *11*, 108–127.
- [257] Hanwell, M. D.; Curtis, D. E.; Lonie, D. C.; Vandermeersch, T.; Zurek, E.; Hutchison, G. R. Avogadro: An Advanced Semantic Chemical Editor, Visualization, and Analysis Platform. *J. Cheminformatics* 2012, *4*, 17.
- [258] Martínez, L.; Andrade, R.; Birgin, E. G.; Martínez, J. M. PACKMOL: A Package for Building Initial Configurations for Molecular Dynamics Simulations. *J. Comput. Chem.* 2009, *30*, 2157–2164.
- [259] Martínez, J. M.; Martínez, L. Packing Optimization for Automated Generation of Complex System’s Initial Configurations for Molecular Dynamics and Docking. *J. Comput. Chem.* 2003, *24*, 819–825.
- [260] The CP2K Developers Group. 2001; [www.cp2k.org](http://www.cp2k.org).
- [261] Becke, A. D. Density-Functional Exchange-Energy Approximation with Correct Asymptotic Behavior. *Phys. Rev. A* 1988, *38*, 3098–3100.
- [262] Lee, C.; Yang, W.; Parr, R. G. Development of the Colle-Salvetti Correlation-Energy Formula into a Functional of the Electron Density. *Phys. Rev. B* 1988, *37*, 785–789.

- [263] VandeVondele, J.; Hutter, J. Gaussian Basis Sets for Accurate Calculations on Molecular Systems in Gas and Condensed Phases. *J. Chem. Phys.* 2007, *127*, 114105.
- [264] Goedecker, S.; Teter, M.; Hutter, J. Separable Dual-Space Gaussian Pseudopotentials. *Phys. Rev. B* 1996, *54*, 1703–1710.
- [265] Hartwigsen, C.; Goedecker, S.; Hutter, J. Relativistic Separable Dual-Space Gaussian Pseudopotentials from H to Rn. *Phys. Rev. B* 1998, *58*, 3641–3662.
- [266] Krack, M. Pseudopotentials for H to Kr Optimized for Gradient-Corrected Exchange-Correlation Functionals. *Theor. Chem. Acc.* 2005, *114*, 145–152.
- [267] Bussi, G.; Donadio, D.; Parrinello, M. Canonical Sampling through Velocity Rescaling. *J. Chem. Phys.* 2007, *126*, 014101.
- [268] CPMD 4.1.0, Copyright 2000–2019. Jointly by IBM Corp. and by Max Planck Institute, Stuttgart. 1990; [www.cpmd.org](http://www.cpmd.org).
- [269] Troullier, N.; Martins, J. L. Efficient Pseudopotentials for Plane-Wave Calculations. *Phys. Rev. B* 1991, *43*, 1993–2006.
- [270] Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Petersson, G. A.; Nakatsuji, H.; Li, X.; Caricato, M.; Marenich, A.; Bloino, J.; Janesko, B. G.; Gomperts, R.; Mennucci, B.; Hratchian, H. P.; Ortiz, J. V.; Izmaylov, A. F.; Sonnenberg, J. L.; Williams-Young, D.; Ding, F.; Lipparini, F.; Egidi, F.; Goings, J.; Peng, B.; Petrone, A.; Henderson, T.; Ranasinghe, D.; Zakrzewski, V. G.; Gao, J.; Rega, N.; Zheng, G.; Liang, W.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Throssell, K.; Montgomery, J. A., Jr.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Keith, T.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Millam, J. M.; Klene, M.; Adamo, C.; Cammi, R.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Farkas, O.; Foresman, J. B.; Fox, D. J. Gaussian 09, Revision d.01. 2016,
- [271] Humphrey, W.; Dalke, A.; Schulten, K. VMD – Visual Molecular Dynamics. *J Mol Graph* 1996, *14*, 33–38.

## REFERENCES

- [272] Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of Simple Potential Functions for Simulating Liquid Water. *J. Chem. Phys.* 1983, 79, 926–935.
- [273] Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kalé, L.; Schulten, K. Scalable Molecular Dynamics with NAMD. *J. Comput. Chem.* 2005, 26, 1781–1802.
- [274] Feller, S. E.; Zhang, Y.; Pastor, R. W.; Brooks, B. R. Constant Pressure Molecular Dynamics Simulation: The Langevin Piston Method. *J. Chem. Phys.* 1995, 103, 4613–4621.
- [275] Nosé, S. A Unified Formulation of the Constant Temperature Molecular Dynamics Methods. *J. Chem. Phys.* 1984, 81, 511–519.
- [276] Schiffmann, C.; Sebastiani, D. Artificial Bee Colony Optimization of Capping Potentials for Hybrid Quantum Mechanical/Molecular Mechanical Calculations. *J. Chem. Theory Comput.* 2011, 7, 1307–1315.
- [277] Jensen, F. Polarization Consistent Basis Sets: Principles. *J. Chem. Phys.* 2001, 115, 9113–9125.
- [278] Mulder, F. A. A.; Filatov, M. NMR Chemical Shift Data and Ab Initio Shielding Calculations: Emerging Tools for Protein Structure Determination. *Chem. Soc. Rev.* 2010, 39, 578–590.
- [279] Barone, G.; Gomez-Paloma, L.; Duca, D.; Silvestri, A.; Riccio, R.; Bifulco, G. Structure Validation of Natural Products by Quantum-Mechanical GIAO Calculations of <sup>13</sup>C NMR Chemical Shifts. *Chem. – Eur. J.* 2002, 8, 3233–3239.
- [280] Neese, F. The ORCA Program System. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* 2012, 2, 73–78.
- [281] Neese, F. Software Update: The ORCA Program System, Version 4.0. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* 2018, 8, e1327.
- [282] Field, M. J. The pDynamo Program for Molecular Simulations Using Hybrid Quantum Chemical and Molecular Mechanical Potentials. *J. Chem. Theory Comput.* 2008, 4, 1151–1161.
- [283] Gowers, R. J.; Linke, M.; Barnoud, J.; Reddy, T. J. E.; Melo, M. N.; Seyler, S. L.; Domański, J.;

- Dotson, D. L.; Buchoux, S.; Kenney, I. M.; Beckstein, O. MDAnalysis: A Python Package for the Rapid Analysis of Molecular Dynamics Simulations. Proceedings of the 15th Python in Science Conference. 2016; pp 98–105.
- [284] Michaud-Agrawal, N.; Denning, E. J.; Woolf, T. B.; Beckstein, O. MDAnalysis: A Toolkit for the Analysis of Molecular Dynamics Simulations. *J. Comput. Chem.* 2011, 32, 2319–2327.
- [285] Oliphant, T. E. Python for Scientific Computing. *Comput. Sci. Eng.* 2007, 9, 10–20.
- [286] Millman, K. J.; Aivazis, M. Python for Scientists and Engineers. *Comput. Sci. Eng.* 2011, 13, 9–12.
- [287] Perez, F.; Granger, B. E. IPython: A System for Interactive Scientific Computing. *Comput. Sci. Eng.* 2007, 9, 21–29.
- [288] Python Software Foundation. Python Language Reference, Version 3.6. [www.python.org](http://www.python.org).
- [289] Oliphant, T. A Guide to NumPy. 2006; [www.numpy.org](http://www.numpy.org).
- [290] van der Walt, S.; Colbert, S. C.; Varoquaux, G. The NumPy Array: A Structure for Efficient Numerical Computation. *Comput. Sci. Eng.* 2011, 13, 22–30.
- [291] Jones, E.; Oliphant, T.; Peterson, P. SciPy: Open Source Scientific Tools for Python. 2001; [www.scipy.org](http://www.scipy.org).
- [292] Hunter, J. D. Matplotlib: A 2D Graphics Environment. *Comput. Sci. Eng.* 2007, 9, 90–95.
- [293] Stone, J. Tachyon. An Efficient Library for Parallel Ray Tracing and Animation. 1995; [jedi.ks.uiuc.edu/~johns/raytracer](http://jedi.ks.uiuc.edu/~johns/raytracer).
- [294] The PyMOL Molecular Graphics System, Version 1.7, Schrödinger, LLC. 2014; [pymol.org](http://pymol.org).
- [295] GIMP 2.8. [gimp.org](http://gimp.org).



## LIST OF ABBREVIATIONS

acc.	Acceptor
ACF	Autocorrelation function
AIMD	<i>Ab initio</i> molecular dynamics
AnPixJg <sub>2</sub>	<i>see footnote f on page 77</i>
B <sub>3</sub> LYP	Becke+Hartree-Fock/Lee-Yang-Parr (exchange/correlation)
BLYP	Becke/Lee-Yang-Parr (exchange/correlation)
BO	Born-Oppenheimer
BOMD	Born-Oppenheimer molecular dynamics
BSSE	Basis set superposition error
CA	Complete adiabatic
calc.	Calculated
CAM	Coulomb-attenuating method
CBCR	Cyanobacteriochrome
CC	Coupled cluster
CHARMM	Chemistry at Harvard macromolecular mechanics
CIS(D)	Configuration interaction, including single and (perturbative) double excitations
CO	Coupled oscillator
COM	Centre of mass
CP	Cross polarisation
Cphi $\Delta_2$	<i>see footnote e on page 77</i>
CPMD	Car-Parrinello molecular dynamics
CSGT	Continuous set of gauge transformations
DFPT	Density functional perturbation theory
DFT	Density functional theory
DMOG	Distributed molecular origin gauge
don.	Donor
ECD	Electronic circular dichroism
EM	Effective mode

## LIST OF ABBREVIATIONS

ETDM	Electronic transition dipole moment
FT	Fourier transform/ed
exp.	Experimental
FFMD	Force field molecular dynamics
FWHM	Full width half maximum
GAF	cGMP phosphodiesterase/adenylyl cyclase/FhlA
GAPW	Gaussian-augmented plane waves
GGA	Generalised gradient approximation
GPW	Gaussian and plane waves
GTH	Goedecker-Teter-Hutter (pseudopotential)
HB	Hydrogen bond
HF	Hartree-Fock
HOMO	Highest occupied molecular orbital
IR	Infrared
iVCD	Induced vibrational circular dichroism
KS	Kohn-Sham
LA	Lactic acid
LDA	Local density approximation
LUMO	Lowest unoccupied molecular orbital
MAS	Magic angle spinning
MD	Molecular dynamics
MFPT	Magnetic field perturbation theory
MI	Matrix isolation
MLWO	Maximally localised Wannier orbital
MM	Molecular mechanics
MP <sub>2</sub>	Møller-Plesset (perturbation theory of second order)
MTDM	Magnetic transition dipole moment
NM	Normal mode
NMR	Nuclear magnetic resonance
NVE	Microcanonical (ensemble)
NVPT	Nuclear velocity perturbation theory
NVT	Canonical (ensemble)
P <sub>1</sub> /P <sub>2</sub>	Protocol 1/2
PAS	period circadian/aryl hydrocarbon receptor nuclear translocator/single-minded
PBC	Periodic boundary conditions



## LIST OF ABBREVIATIONS

PCB	Phycocyanobilin
PCM	Polarisable continuum models
PES	Potential energy surface
PHY	Phytochrome specific
PLA	Polylactic acid
PP	Pseudopotential
QM	Quantum mechanics
QM/MM	(Hybrid) Quantum mechanics/molecular mechanics
QMP	QM part
RDF	Radial distribution function
RMSD	Root mean square deviation
RMSE	Root mean square error
SDF	Spatial distribution function
SLA	( <i>S</i> )-lactic acid
sTD-DFT	Simplified time-dependent density functional theory
TCD	Transition current density
TCF	Time-correlation function
TDA	Tamm-Dancoff density functional approach
TD-DFT	Time-dependent density functional theory
TDPEs	Time-dependent potential energy surface
TDSE	Time-dependent Schrödinger equation
UV	Ultraviolet
VCD	Vibrational circular dichroism
VDOS	Vibrational density of states



## LIST OF FIGURES

4.1	Lewis formulae of ( <i>S</i> )- and ( <i>R</i> )-alanine, also referred to as L-(+)- and D-(−)-alanine, respectively, in their zwitterionic forms. . . . .	34
4.2	IR (top) and VCD (bottom) solid-state spectra of L-alanine, measured as ground crystals in a nujol mull (literature data <sup>6</sup> ). . . . .	38
4.3	Calculated anharmonic solid-state IR (top) and VCD (bottom) spectra of crystalline L-alanine and experimental data, measured as a nujol oil mull (literature data <sup>6</sup> ). The theoretical result is substantially influenced by the contribution of supramolecular correlations (Reprinted from reference 1. ©2018 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim). . . . .	39
4.4	Calculated and experimental solid-state IR spectra as well as the calculated power spectrum of crystalline L-alanine. <sup>6</sup> . . . . .	40
4.5	Crystal topology of crystalline L-alanine with, a), supramolecular chains, defining the z-vector (upper right), that are collinearly staggered (upper left). Cylindrical coordinates are declared with the radii $r_z$ and $r_{xy}$ defined as distance from the reference point, respectively (bottom); b), orientation of L-alanine molecules in the crystal with white arrows indicating the orientation of the carbonyl/amino oscillators (Reprinted from reference 1. ©2018 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim). . . . .	42
4.6	Radial decomposition of the isotropic IR (a) and VCD (b) spectra of crystalline L-alanine along the longitudinal axis, z, (left) and the polar axis lying in the xy-plane (right) (Reprinted from reference 1. ©2018 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim). . . . .	43
4.7	Radial decomposition of the vibrational density of states of the carbonyl carbon atoms of L-alanine in z- (top) and xy- (bottom) dimension. The peaks at $r_z/r_{xy} > 0$ reveal that there is a true supramolecular correlation of atomic motion in the carbonyl region. . . . .	44

## LIST OF FIGURES

4.8	Calculated effective modes of L-alanine crystals in the carbonyl region with increasing supramolecular scope. The intermolecular correlation of fundamental modes gives rise to symmetric (zero phase) and anti-symmetric (phase $\pi$ ) combinations that result in oscillators with increased or reduced transition dipole moments per molecule (globes). The number of considered molecules is augmented gradually (from one to four molecules). The IR intensity is determined by the size of the total transition dipole moment. VCD enhancement originates solely at the intermolecular distance $r_{xy}$ of 4-5 Å (Reprinted from reference 1. ©2018 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim). . . . .	45
4.9	Radial decomposition of the direct (top) and the gauge transport (bottom) term of the isotropic VCD spectrum of crystalline L-alanine along the polar axis lying in the $xy$ -plane. . . . .	47
5.1	Lewis formulae of ( <i>S</i> )- and ( <i>R</i> )-lactic acid, also referred to as L-(+)- and D-(−)-lactic acid, respectively. . . . .	50
5.2	IR (top) and VCD (bottom) spectrum of <i>S</i> -lactic acid, measured as 0.1 M solution in chloroform- $d_3$ . (literature data <sup>140</sup> ) . . . . .	51
5.3	Top: Topologies of the ( <i>S</i> )-lactic acid monomer (left) and head-on dimer (right). Bottom: Dihedral angle distribution of the molecular core (indicated in blue) as obtained from AIMD simulations. The dashed vertical lines indicate the dihedral angle of the optimised geometries that entered the normal mode analysis. . . . .	52
5.4	Calculated IR (top) and VCD (bottom) spectra of ( <i>S</i> )-lactic acid species compared with experimental data, measured as 0.1 M solution in chloroform- $d_3$ (literature data <sup>140</sup> ). A solid line represents spectra obtained from AIMD simulations (shown here with unscaled frequencies), whereas dashed lines correspond to conventional Hessian-based static calculations, artificially convoluted with a Gaussian function. . . . .	53
5.5	Left: Computed AIMD IR and VCD spectra of the monomers-in-the-dimer (yellow) and the full dimer (blue) of <i>S</i> -lactic acid. The correlation of motion of the two carboxylic carbon atoms is shown in green. Right: Two combination bands of the asymmetric carboxylic stretching vibration. Only the lower mode at higher frequency is IR active. . . . .	54
5.6	Left: Topology of the ( <i>S</i> )-lactic acid trimer model “dimer + X”, together with the distribution of angles between the carboxylic planes of the molecules. Right: Molecular sub-correlations of the total IR (pale) and VCD (vivid) band in the carbonyl region. Although molecule 3 does not contribute much to the total IR spectrum, it is as equally involved in the generation of the VCD signal as the other two molecules. . . . .	55

5.7	Calculated anharmonic IR (top) and VCD (bottom) spectra of an ( <i>S</i> )-lactic acid trimer and experimental data, measured as 0.1 M solution in chloroform- <i>d</i> <sub>3</sub> (literature data <sup>140</sup> ). . . . .	56
5.8	IR (top) and VCD (bottom) spectra of <i>S</i> -lactic acid, measured as 2.0 M solution in water. (literature data <sup>140</sup> ) . . . . .	60
5.9	( <i>S</i> )-Lactic acid in aqueous solution: RMSD of the solute's heavy atoms and assigned restructuring events during the canonical AIMD. Relevant water molecules are shown in blue. The vertical lines mark furcation points of microcanonical AIMD for production of IR and VCD spectra. . . . .	61
5.10	Left: Dihedral angle distribution of the molecular core of ( <i>S</i> )-lactic acid in aqueous solution, obtained from AIMD simulations. Right: Radial distribution function (RDF) of molecular centres of mass (black) and with respect to the position of the protonated oxygen atom of the carboxylic group (red). The grey region forms a Fermi function with a 4 Å cutoff that has been employed in this work for moment scaling. . . . .	62
5.11	Spatial distribution functions (SDF) of atoms in ( <i>S</i> )-lactic acid (carbon, hydrogen and oxygen atoms respectively shown in gray, white, and red) and water (only oxygen atoms, blue) obtained from AIMD simulations. The two rotamers of lactic acid (left and right) are shown in two different aspects, respectively (60% isosurface value). . . . .	63
5.12	Lifetime analysis of hydrogen bonds of ( <i>S</i> )-lactic acid in water, divided into the two sites of the carboxylic group (left) and the single site of the hydroxy group (the protonated oxygen atoms can act as both, HB donor and acceptor), based on continuous (solid lines) and intermittent (dashed lines) lifetimes. The water reference is shown in black. . . . .	64
5.13	Computed AIMD IR (top) and VCD (bottom) spectra of ( <i>S</i> )-lactic acid in aqueous solution following different moment scaling schemes at increasing cutoff value (from red to green). Left: Conventional scaling according to Scherrer. <sup>17</sup> Centre: Conventional scaling with protonated carboxylic oxygen atoms as cutoff reference. Right: Adaptive scaling according to Equation 5.3 with protonated carboxylic oxygen atoms as cutoff reference. . . . .	65
5.14	Calculated anharmonic IR (top) and VCD (bottom) spectra of ( <i>S</i> )-lactic acid in water from AIMD, contrasted with experimental data, measured as 2.0 M solution in water (literature data <sup>140</sup> ). The term "local" refers to spectral calculations only including moments stemming from the chiral molecule itself. . . . .	67

## LIST OF FIGURES

- 5.15 Molecular sub-correlations of the total IR (pale) and VCD (vivid) band in the carbonyl region ("H<sub>2</sub>O": water, "SLA": (*S*)-lactic acid). Due to adaptive scaling, the water molecules do not contribute to the IR spectrum. However, in generating the VCD signal, they are clearly involved. . . . . 68
- 5.16 Snapshot from the AIMD trajectory of (*S*)-lactic acid in water, showing the carbonyl stretching mode (blue arrows). The "active water" molecule is hydrogen-bonded to the protonated carboxylic oxygen atom. Left: Location of Wannier Centres (green), corresponding to electron lone pairs. Right: Transition current density (TCD) invoked by the carbonyl vibration in its surrounding. The green arrow marks the chiral redirection of current density that causes iVCD. . . . . 69
- 5.17 A novel scheme of induced VCD in solution. . . . . 70
- 6.1 Left: Sensory module of CphI, a phytochrome of cyanobacterial origin. Phytochromes comprise a PAS (yellow)–GAF (green)–PHY (blue) tridomain,<sup>168</sup> wherein the latter acts as signal transmitter.<sup>166,169</sup> Right: Photoactive GAF domain (green) of AnPixJ, a cyanobacteriochrome (CBCR).<sup>170</sup> In both cases, the cofactor phycocyanobilin (PCB) is covalently bound to the GAF domain (red). . . . . 74
- 6.2 Lewis formula of phycocyanobilin (PCB) in its *ZZZssa* (left) and *ZZEssa* (right) configurations. Reversible photoisomerisation of the C<sub>15</sub>-C<sub>16</sub> double bond triggers the photoswitch. . . . . 75
- 6.3 Depiction of a fully periodic simulation supercell of AnPixJg2 in water subject to FFMD and QM/MM AIMD simulations. . . . . 78
- 7.1 Left: RMSD of the AnPixJg2 backbone relative to the crystal structure from X-ray diffraction: Full domain (black) and without considering the N-terminal  $\alpha$ -helix and the unstructured loop region between residues 274 and 290 (red).<sup>2,183</sup> Right: Backbone after 1  $\mu$ s of FFMD simulation, aligned to the crystal structure (PCB shown as spheres). . . . . 79
- 7.2 Top left: RMSD of PCB heavy atoms together with the individual deviations of pyrrole rings A, B, C, and D relative to the crystal structure from X-ray diffraction. The white and grey areas mark the presence of  $\alpha$  or  $\beta$  orientation of ring D (RMSD shown in red). Bottom left: Dihedral distribution of the three methylene bridges connecting the pyrrole rings. Right:  $\alpha$  (top) and  $\beta$  (bottom) sub-state of PCB.<sup>2,223</sup> . 80

7.3	Top: Radar plots of the occurrence of hydrogen bond contacts between PCB and apoprotein or solvent, where PCB serves as acceptor (left) and donor (right). Bottom: Skeletal deviation plot <sup>155,240</sup> showing the z-displacements of PCB heavy atoms from the mean plane. . . . .	81
7.4	Chromophore-binding pocket of AnPixJg2 in the crystal (top) and after long-scale FFMD simulations, showcasing $\alpha$ (middle) and $\beta$ sub-states (bottom). Non-polar hydrogen atoms have been omitted for clarity. . . . .	84
7.5	Left: Evolution of the Trp289–ring A hydrogen bond in a MD simulation under constraint (grey area) and after release (white area). Right: Supramolecular situation of Trp289 and PCB in the crystal structure: hydrogen bond to ring A and $\pi$ - $\pi$ interaction with ring D. . . . .	86
7.6	Top: Formation of water chains and the establishment of a persistent water molecule in $\alpha$ (left) and $\beta$ (right) sub-state, respectively. Bottom: Distance timeline of the persistent water molecule to ring B of PCB. The change in background shade marks events when it is substituted by a new water molecule. . . . .	87
7.7	Left: RMSD of the CphI $\Delta$ 2 backbone relative to the crystal structure from X-ray diffraction: Full protein (black) and only considering the GAF domain (red). <sup>169</sup> Right: Backbone after 400 ns of FFMD simulation, aligned to the crystal structure (PCB shown as spheres). . . . .	89
7.8	Top: Radar plots of the occurrence of hydrogen bond contacts between PCB and apoprotein or solvent, where PCB serves as acceptor (left) and donor (right). Bottom: Skeletal deviation plot <sup>155,240</sup> showing the z-displacements of PCB heavy atoms from the mean plane. . . . .	90
7.9	Chromophore-binding pocket of CphI $\Delta$ 2 (“Pr-II” <sup>188</sup> ) after 400 ns of FFMD simulation. Non-polar hydrogen atoms have been omitted for clarity. The PHY domain (not shown) forms a salt bridge with the carboxylate group of Asp207 sticking out at the lower edge of the image. . . . .	90
8.1	Protocols of spectrum acquisition from MD simulations. Left: Direct sampling of the NPT ensemble, augmented by short canonical AIMD equilibration runs and annealing (P1). Right: Sampling along canonical AIMD trajectories having started from FFMD snapshots (P2). . . . .	92
8.2	Normed RMSE of calculated shieldings with increasing size of the QM part (left) and after annealing and equilibration of FFMD snapshots (right) (P2). . . . .	93

## LIST OF FIGURES

- 8.3 Normalised RMSE of calculated shieldings by means of AIMD sampling plotted against the number of snapshots, indicating convergence of the linear response calculation. The filled area marks the sensitivity of calculated shieldings towards thermal fluctuations and supramolecular interaction as RMSD. . . . . 93
- 8.4 Left: Average C–C bond lengths in pm, extracted from the AIMD trajectory of PCB in AnPixJg<sub>2</sub> (the standard error is 0.4 pm). Right: Proposed lewis representation of the conjugation pattern in PCB, based on the AIMD analysis (*cf.* Figure 6.2). . . 94
- 8.5 <sup>13</sup>C chemical shifts of PCB in AnPixJg<sub>2</sub> (black) and CphIΔ<sub>2</sub> (blue).<sup>197</sup> The crosses mark calculated values from MD trajectories after protocols P<sub>1</sub> and P<sub>2</sub> alike. . . . . 96
- 8.6 Correlation matrix **M** (see Equation 8.1) representing the sensitivity of <sup>13</sup>C chemical shifts against C–C bond lengths in rings C and D. Colour and size of the circles correspond to the Pearson correlation coefficient (see colour bar on the right), while the white numbers denote the slope of linear regression (in ppm Å<sup>-1</sup>). Excluded from view are entries with significance level lower than 95 % (p-value). . . . . 97
- 8.7 Correlation between calculated chemical shifts of both, <sup>1</sup>H (upper panels) and <sup>15</sup>N (lower panels), with the length of hydrogen bonds (left) and N-H bonds (right). The results for AnPixJg<sub>2</sub> are shown in black, those for CphIΔ<sub>2</sub> in blue. . . . . 100
- 8.8 Correlation of calculated <sup>15</sup>N with <sup>1</sup>H chemical shifts. The results for AnPixJg<sub>2</sub> are shown in black, those for CphIΔ<sub>2</sub> in blue. . . . . 101
- 8.9 Correlation matrix in analogy to Equation 8.1, showing the sensitivity of <sup>15</sup>N and <sup>1</sup>H chemical shifts towards the length of hydrogen and N-H bonds. Additionally, maximal correlation found with any C–C bond of the pyrrole rings is shown (last row). Colour and size of the circles correspond to the Pearson correlation coefficient (see colour bar on the right), while the white numbers denote the slope of linear regression (in ppm Å<sup>-1</sup>). Excluded from view are entries with significance level lower than 95 % (p-value). . . . . 102
- 8.10 2D <sup>1</sup>H-<sup>15</sup>N NMR spectra of AnPixJg<sub>2</sub> obtained from QM/MM MD following protocol P<sub>1</sub> (left) and P<sub>2</sub> (right); see also Figure 8.1. The computed spectra are compared to experimental data from CP/MAS measurements at -40 °C. The grey signals correspond to direct calculations from the crystal structure after geometry optimisation. The error bars mark the standard error of ensemble average. . . . . 103



8.11	Calculated UV-Vis spectra of PCB in AnPixJg2( $\alpha$ ) (top) and CphI $\Delta$ 2 (bottom) for different sizes of the QM part (QMP), including either the full cofactor and neighbouring protein residues (dashed lines) or a reduced chromophore core consisting of the $\pi$ system with adjacent atomic groups (solid lines). The dotted lines represent calculations without consideration of the protein. The pins mark experimental values. <sup>169,170</sup> . . . . .	106
8.12	Calculated UV-Vis spectra of PCB in AnPixJg2 ( $\beta$ ) for deprotonation of an “allowed” (ring B) and a “forbidden” (ring D) position (protocol P1). The positions of all remaining atoms were kept unchanged. . . . .	107
8.13	Left: Correlation of the So $\rightarrow$ Si absorption energy with the out-of-plane rotation of ring D with respect to ring C in AnPixJg2 (black) and CphI $\Delta$ 2 (blue). Right: Correlation of the absorption energy with the HOMO-LUMO gap. . . . .	108
8.14	Frontier molecular orbitals involved in the So $\rightarrow$ Si transition: HOMO (left) and LUMO (right). . . . .	109
8.15	Calculated UV-Vis spectra of PCB in AnPixJg2 in its $\alpha$ and $\beta$ sub-state, respectively. . . . .	110
8.16	Top: Calculated ECD spectrum of PCB in AnPixJg2 in its $\alpha$ and $\beta$ sub-state, respectively. Bottom: Measured ECD spectrum of the P <sub>r</sub> state of AnPixJg2 <i>in vivo</i> . . . . .	111
8.17	Correlation of ECD rotational strengths of PCB in AnPixJg2 with the dihedral angle between rings C and D. The results for two electronic transitions are shown as indicated by the insets. . . . .	111
A.1	Root mean square deviation (RMSD) of heavy atoms of L-alanine along the canonical trajectory with respect to the crystal structure. . . . .	xiii
A.2	Calculated <sup>13</sup> C, <sup>1</sup> H, and <sup>15</sup> N (from top) chemical shifts of PCB plotted against experimental data (black: AnPixJg2; blue: CphI $\Delta$ 2), after referencing to experiment by linear regression. <sup>278,279</sup> The empty circles mark atoms from the protein backbone chosen as internal standard (see text). . . . .	xvii



## LIST OF TABLES

4.1	Assignment of calculated vibrational frequencies to main features of the effective modes analysis of crystalline L-alanine. The frequency values are given with a precision of $5 \text{ cm}^{-1}$ . . . . .	46
5.1	Assignment of calculated vibrational frequencies to main features of the effective modes analysis of dimeric ( <i>S</i> )-lactic acid ( <i>cf.</i> Figure 5.5). . . . .	54
7.1	Interaction energies in $\text{kJ mol}^{-1}$ of Trp289 and PCB, considering either the full chromophore, or ring D alone ( $\pi$ - $\pi$ interaction), calculated at different levels of theory. The quantum calculations were carried out with the def2-TZVP basis set (without correction of BSSE). <sup>244</sup> D <sub>3</sub> denotes dispersion correction by Grimme and co-workers. <sup>245</sup> . . . . .	85
8.1	Averaged Mulliken charges <sup>249</sup> in <i>e</i> of PCB substructures, calculated along the AIMD trajectory of AnPixJg2. . . . .	95
8.2	Hypothetical changes in C–C bond lengths based on measured difference pattern of <sup>13</sup> C chemical shifts (in ppm) going from the P <sub>fr</sub> state of CphIΔ <sub>2</sub> to the P <sub>g</sub> state of AnPixJg2. <sup>197</sup> The trends denoted by +/- are inferred by inverse application of Equation 8.1 to Figure 8.6. . . . .	98
8.3	<sup>15</sup> N chemical shifts of PCB CphIΔ <sub>2</sub> in ppm. The results calculated in this work are compared to dynamic nuclear polarisation solid-state MAS measurements and QM/MM calculations based on geometry optimisation of the crystal structure. . . . .	104
A.1	Computational details of <i>ab initio</i> molecular dynamics calculations for the VCD study. . . . .	xiii
A.2	Computational details of force field and <i>ab initio</i> molecular dynamics calculations for the protein study. . . . .	xv
A.3	Scaling parameters for calculated NMR nuclear shieldings using linear regression to experimental chemical shifts (Equation A.1). . . . .	xvi

## LIST OF TABLES

A.4	AnPixJg2: Information on protein residues included into the QM part, and corresponding capping atoms. In UV-Vis calculations, hydrogen capping has been used for protein CA atoms. . . . .	xviii
A.5	Cphi $\Delta$ 2: Information on protein residues included into the QM part, and corresponding capping atoms. In UV-Vis calculations, hydrogen capping has been used for protein CA atoms. . . . .	xix

# APPENDIX



# A COMPUTATIONAL DETAILS

## A.1 Vibrational Circular Dichroism and Supramolecular Chirality

### A.1.1 Preparations

A perfect infinite crystal of L-alanine was considered based on the crystal structure from neutron diffraction experiment (CCDC 278464).<sup>134</sup> L-alanine crystallises in the  $P2_12_12_1$  space group with cell parameters 6.036(3), 12.342(5), and 5.788(3) Å. A supercell was constructed consisting of  $2 \times 1 \times 2$  unit cells containing 16 molecules in total.

For the study of (*S*)-lactic acid (LA) in vacuum, monomer, dimer, and trimer structures were generated with the AVOGADRO software.<sup>257</sup> For aqueous solution, a the head-on dimer structure (see main text) from the vacuum study was solvated with 123 water molecules in a supercell with parameters 20.0, 14.0, 14.0 Å using the PACKMOL package.<sup>258,259</sup> This corresponds to a density of 1.014 g cm<sup>-3</sup> and a LA concentration of 0.85 mol L<sup>-1</sup>. All vacuum structures were geometry optimised beforehand with CP2K at the same level of theory used in AIMD simulations.<sup>20</sup>

### A.1.2 *Ab Initio* Molecular Dynamics Simulations

Born-Oppenheimer AIMD simulations were carried out with the CP2K<sup>260</sup> program package *via* the Quickstep module,<sup>20</sup> using the BLYP functional for exchange-correlation,<sup>261,262</sup> the DZVP-MOLOPT-GTH (alanine) and DZVP-MOLOPT-SR-GTH (lactic acid) basis set,<sup>263</sup> and GTH pseudo-potentials<sup>264-266</sup> with a 0.5 fs time step, a density cutoff of 400 Ry, and Grimme's dispersion correction (D3).<sup>245</sup> Fully periodic boundary conditions were imposed. After equilibration at 340 K using a massive thermostat (CSVR<sup>267</sup>), a canonical (NVT) production run was performed.

## A COMPUTATIONAL DETAILS

Statistically independent microcanonical (NVE) trajectories were branched off this production run (for more details see Table A.1).

### A.1.3 NVPT Calculations

Every 4 fs along the NVE trajectories, a projected NVPT calculation of electronic response properties has been carried out (see Section 2.4). The method is available in the latest release of the CPMD<sup>38,268</sup> electronic structure package. The calculations have been performed using density functional perturbation theory<sup>15</sup> with Troullier-Martins<sup>269</sup> pseudo-potentials and the BLYP functional, employing a plane wave cutoff of 70 Ry.

### A.1.4 Calculation of IR and VCD Spectra

All TCF calculations were performed according to Equation 2.60. Scaled moments and correction of spectral background were calculated according to Equation 5.3, where cutoff radii  $r_0$  and  $\tilde{r}_0$  of 6.0 and 4.0 Å, respectively, as well as a scaling parameter  $D$  of 0.25 Å were employed. Canonical average was obtained by averaging over the microcanonical trajectories. The shown iD spectra have been smoothed in the frequency domain *via* convolution with a Gaussian filter with  $\sigma$ -values of 2 cm<sup>-1</sup> (alanine; LA: vacuum, monomer/dimer) and 5 cm<sup>-1</sup> (LA: vacuum, trimer; aqueous solution). The calculated frequencies were scaled with a factor of 1.035 to overcome inaccuracies in the GGA exchange-correlation functional which notoriously underestimates bond strengths leading to lower frequency values. The issue is likely to disappear within a higher level of theory (*e.g.*, including exact exchange).

For alanine, spatially-resolved absorptivities according to Equation 4.3 were calculated on a cubic grid with a spacing of 0.464, 0.475, and 0.445 Å, respectively, and a regularisation parameter  $\sigma$  of 0.4 Å. All spectra have been smoothed by means of a Gaussian filter ( $\sigma = 2$  cm<sup>-1</sup>).

The effective mode analysis has been performed with the MOLSIM package.<sup>136</sup>



TABLE A.1 Computational details of *ab initio* molecular dynamics calculations for the VCD study.

Parameter	L-Alanine	(S)-Lactic acid
Space Group	P <sub>2</sub> <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>	–
ABC/ $\alpha\beta\gamma$ (Å/deg)	12.072, 12.342, 11.576 / 90.0, 90.0, 90.0	20.0, 14.0, 14.0 / 90.0, 90.0, 90.0*
Temperature (NVT)	340 K	340 K
Thermostat (NVT)	CSVR	CSVR
Functional	BLYP-D <sub>3</sub>	BLYP-D <sub>3</sub>
Basis Set (CP2K)	DZVP-MOLOPT-GTH	DZVP-MOLOPT-SR-GTH
Density Cutoff (CP2K)	400 Ry	400 Ry
PW Cutoff (CPMD)	70 Ry	70 Ry
No. of NVEs	6	9/14/5/5 (M/D/T/W)**
$t_{\text{NVT}}$	>20 ps	>20 ps
$t_{\text{NVE}}$	20 ps	20 ps/30 ps*

\*In aqueous solution. \*\*M=Monomer, D=Dimer, T=Trimer, W=Aqueous solution

### A.1.5 Static Calculations

IR and VCD calculations for the LA monomer and dimer based on double harmonic approximation were performed with Gaussian 09.<sup>270</sup> The geometries were optimised using the BLYP functional for exchange-correlation,<sup>261,262</sup> def2-TZVP basis set, and Grimme's dispersion correction (D<sub>3</sub>).<sup>245</sup> The obtained linespectrum was convoluted with a Lorentzian of FWHM of 5 cm<sup>-1</sup>.

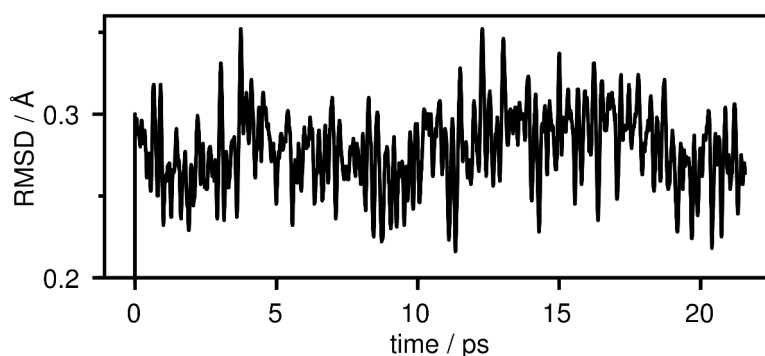


FIGURE A.1 Root mean square deviation (RMSD) of heavy atoms of L-alanine along the canonical trajectory with respect to the crystal structure.

### A.2 Colour Tuning of Phycocyanobilin in Proteins

#### A.2.1 Preparations

Starting geometries for P<sub>r</sub> states of AnPixJg2 and Cph1Δ2 were obtained as X-ray structures from the Protein Data Bank (PDB; AnPixJg2: 3W2Z,<sup>170</sup> Cph1Δ2: 2VEA<sup>169</sup>). The simulation cell was set up using VMD,<sup>271</sup> solvating the protein in a box of TIP3P water.<sup>272</sup> Possible protonation states of protein residues were set to correspond pH 7.0. The protonation state of histidine was determined by visual inspection of the chemical environment, which is the usual procedure (AnPixJg2: His293-E, His322-P, rest: D; Cph1Δ2: His260-E, His290-D, rest: D).

#### A.2.2 Force Field Molecular Dynamics Simulations

FFMD simulations were carried with the NAMD package<sup>273</sup> under periodic boundary conditions (PBC) using the CHARMM22 force field<sup>53,56</sup> in the isothermal-isobaric (NPT) ensemble and the combined Nosé–Hoover Langevin piston method with a period of 200 fs.<sup>274,275</sup> For PCB, force field parameters obtained by Mroginiski and co-workers were used.<sup>54</sup> For van der Waals interactions a cut-off of 10 Å was employed. Bond lengths between heavy atoms and hydrogen atoms were held fixed using the SHAKE algorithm<sup>19</sup> with a time step of 2 fs. The system was equilibrated, first, by optimisation of the water cell, followed by heat up and equilibration keeping the protein positions fixed. Then, the protein structure was optimised with fixed water positions. Final equilibration was carried out after heating up the entire simulation cell to the desired temperature, followed by the production run (see Table A.2 for more details).

#### A.2.3 QM/MM Molecular Dynamics Simulations

Snapshots from FFMD simulations were transferred into a QM/MM setup. AIMD simulations were performed in CP2K.<sup>260</sup> The partitioning of the protein into QM and MM part, as well as employed capping atoms, is listed in Table A.4 and Table A.5. QM/MM bond interfaces were handled using an optimized capping potential introduced between for the bond between  $\alpha$  and  $\beta$  carbon atoms.<sup>61,276</sup> The MM part was described by the CHARMM22 force field (see previous section).

TABLE A.2 Computational details of force field and *ab initio* molecular dynamics calculations for the protein study.

Parameter	FFMD	AIMD	NMR
Ensemble	NPT	NVT	NVT
No. of atoms	37651 / 104960**	$\approx 100_{\text{QM}}$	$\approx 280_{\text{QM}}$
Temperature	300 K	300 K	0 / 300 K*
Pressure	1.01325 bar	–	–
Barostat	Langevin	–	–
Thermostat	Nosé–Hoover	Nosé–Hoover	–
ABC <sub>QM</sub> (Å)	–	30.0, 30.0, 30.0	35.0, 35.0, 35.0
Force Field / Functional	CHARMM22	BLYP-D <sub>3</sub>	BLYP-D <sub>3</sub>
Basis Set	–	TZVP-GTH	pcS <sub>2</sub> /3
Density Cutoff	–	320 Ry	400 Ry
No. of Snapshots	–	–	10–25 / 200*
$t_{\text{NVT}} / t_{\text{NVT}}$	$\approx 1 \mu\text{s}$	>20 ps	>20 ps
Other	SHAKE		

\*P<sub>1</sub> / P<sub>2</sub> (see Subsection 8.1.1). \*\*AnPixJg2 / CphiΔ<sub>2</sub> + H<sub>2</sub>O

The size of the QM box was set to 30.0, 30.0, 30.0 Å. The BLYP functional<sup>261,262</sup> was employed together with Grimme’s dispersion correction (D<sub>3</sub>),<sup>245</sup> using the GPW scheme<sup>20</sup>, GTH pseudo-potentials,<sup>264–266</sup> a density cutoff of 320 Ry, and the TZVP-GTH basis set. The QM/MM interaction term was evaluated within the GEEP scheme presented by Laino and co-workers.<sup>62,63</sup> Depending on the chosen protocol (see Subsection 8.1.1) equilibration runs of different length (5–10 ps) were carried out under massive thermostating using Nosé–Hoover chains of length 5 at 330 K and a coupling constant of 10 fs.<sup>18,275</sup> For subsequent production runs (P<sub>2</sub>), QM and MM subsystems were coupled to separate thermostat units with a coupling constant of 500 fs. Annealing (P<sub>1</sub>) was performed with a factor of 0.995.

#### A.2.4 Calculation of Nuclear Shieldings

NMR calculations were performed in the same setup as AIMD simulations, using MFPT (see Section 2.4). However, the size of the QM/MM part was chosen such that RMSE of chemical shifts were below 1 % (see Figure 8.2 and Table A.4/Table A.5). Water molecules within 2.6 Å to PCB were included in the QM part as well. Snapshots were extracted from the AIMD trajectory according to the sampling protocol (see Subsection 8.1.1). All atoms in the QM region were treated at an all-electron level using the Gaussian-augmented plane waves (GAPW) method of CP2K<sup>21</sup> with density cutoff

## A COMPUTATIONAL DETAILS

TABLE A.3 Scaling parameters for calculated NMR nuclear shieldings using linear regression to experimental chemical shifts (Equation A.1).

Nucleus	$a$ (a.u. ppm <sup>-1</sup> )	$b$ (a.u.)
<sup>1</sup> H	-1.1506	30.923
<sup>13</sup> C	-1.0599	165.56
<sup>15</sup> N	-1.1278	225.06

of 400 Ry and pcS-3 basis set (pcS-2 on oxygen and sulphur atoms).<sup>277</sup> The size of the QM box was set to 35.0, 35.0, 35.0 Å. Gauge origin was treated using the IGAIM method of CP2K,<sup>21</sup> which is an adaption of the CSGT method based on *Atoms in Molecules*.<sup>64</sup> For comparison to experiment theoretical shieldings were scaled using the slope  $a$  and intercept  $b$  obtained by linear regression analysis of experimental shifts,<sup>278,279</sup>

$$\delta_{\text{CAL}} = (b - \sigma_{\text{EXP}}) / a. \quad (\text{A.1})$$

The resulting parameters for each nuclear kind is shown in Table A.3. To improve the statistical data basis of <sup>1</sup>H and <sup>15</sup>N shieldings, internal standards stemming from chromophore and protein backbone were chosen (Methyl group for <sup>1</sup>H, amide group of cysteine for <sup>15</sup>N). Figure A.2 plots calculated against experimental chemical shifts. Convergence of computed NMR chemical shieldings with simulation time (after P2) is shown in Figure 8.3.

### A.2.5 UV-Vis and ECD Calculations

UV-Vis and ECD spectra were calculated with ORCA<sup>280,281</sup> using simplified TD-DFT (Section 2.4) the CAM-B3LYP functional<sup>254</sup>, and a def2-TZVP basis set.<sup>244</sup> Structures were based on the same AIMD snapshots used for NMR calculations. Partitioning into QM and MM subsystem was carried out using PDYNAMO;<sup>282</sup> whereas MM charges entered the calculations as point charges and hydrogen capping has been used for bonds between QM and MM part. Two different QM part sizes were studied; the small variant comprised 72 core atoms of PCB only (including capping atoms) after the propionate side chains were reduced to methyl groups. For the full chromophore, protein residues were included according to Table A.4 and Table A.5 and capping was always realised at the position of the CA atoms of protein residues. The shown spectra are based on the velocity dipole formalism of the oscillator strength (see Equation 2.66).

## A.2 COLOUR TUNING OF PHYCOCYANOBILIN IN PROTEINS

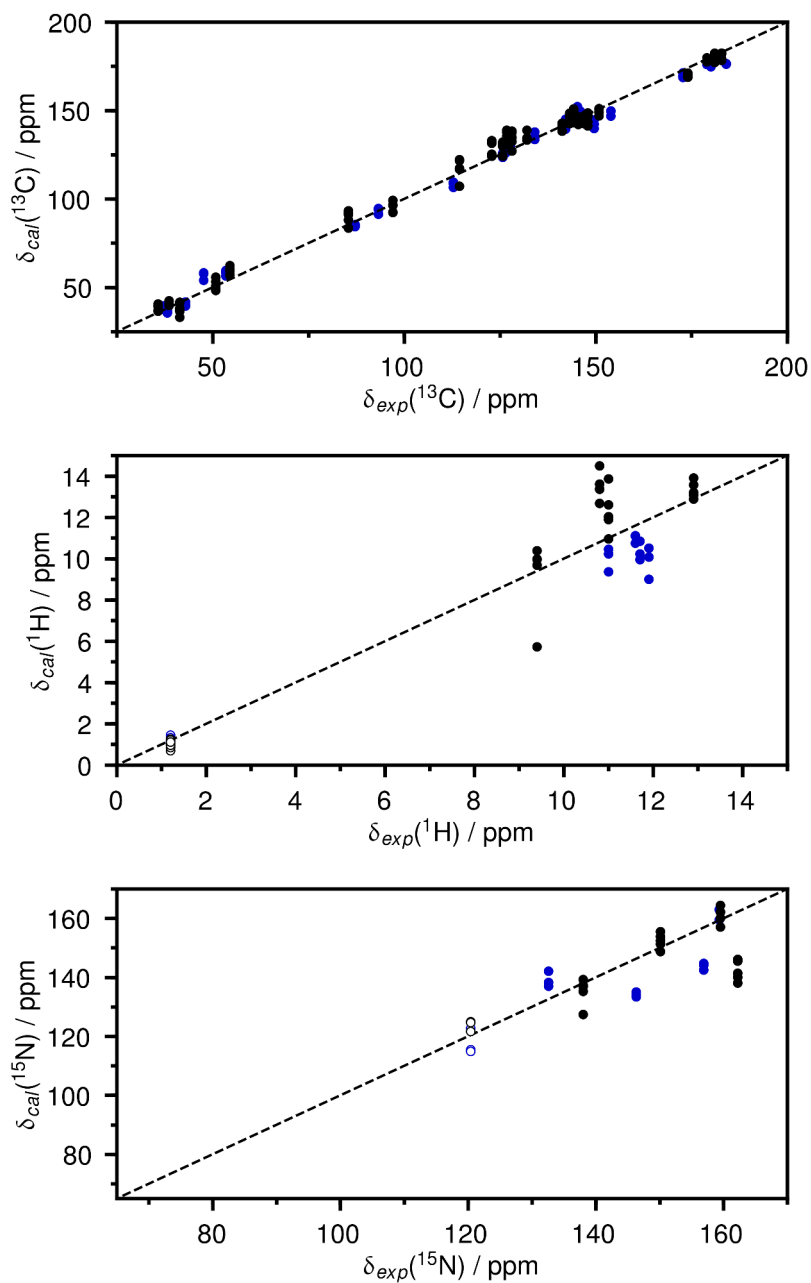


FIGURE A.2 Calculated  ${}^{13}\text{C}$ ,  ${}^1\text{H}$ , and  ${}^{15}\text{N}$  (from top) chemical shifts of PCB plotted against experimental data (black: AnPixJg2; blue: Cph1Δ2), after referencing to experiment by linear regression.<sup>278,279</sup> The empty circles mark atoms from the protein backbone chosen as internal standard (see text).

## A COMPUTATIONAL DETAILS

TABLE A.4 AnPixJg2: Information on protein residues included into the QM part, and corresponding capping atoms. In UV-Vis calculations, hydrogen capping has been used for protein CA atoms.

Residue	AIMD	NMR	Capping Atom	UV-Vis (full / red <sup>**</sup> )
PCB	x	x	–	x / x
Trp289	x	x	CA	x / –
Asp291	x	x	CA	x / –
His293	–	x	CA	–
Arg301	–	x	CA	–
Tyr302	–	x	CA	–
Phe308	–	x	CA	–
His318	–	x	CA	–
Phe319	–	x	CA	–
Ser320	x	x	CA	–
Cys321	x	x	–	x / –
His322	x	x	C	x / –
Tyr334	–	x	CA	–
Water*	–	x	–	–

\*Water molecules within a distance of 2.6 Å to PCB. \*\*The reduced QM part for UV-Vis calculations comprise 72 core atoms of PCB only (including capping atoms), whereas the propionate chains have been reduced to methyl groups.

## A.2 COLOUR TUNING OF PHYCOCYANOBILIN IN PROTEINS

TABLE A.5 Cph1 $\Delta$ 2: Information on protein residues included into the QM part, and corresponding capping atoms. In UV-Vis calculations, hydrogen capping has been used for protein CA atoms.

Residue	AIMD	NMR	Capping Atom	UV-Vis (full / red <sup>**</sup> )
PCB	x	x	–	x / x
Asp207	x	x	CB / CA <sup>***</sup>	x / –
Ile208	x	x	C / C, CB <sup>***</sup>	x / –
Arg222	–	x	CB	– / –
Arg254	–	x	CB	– / –
Ala256	–	x	CA	– / –
Tyr257	–	x	–	– / –
His258	x	x	CA / CB <sup>***</sup>	– / –
Cys259	x	x	–	x / –
His260	x	x	C	x / –
Tyr263	–	x	CA	– / –
His290	–	x	CA	– / –
Water*	x	x	–	– / –

\*Water molecules within a distance of 2.6 Å to PCB. \*\*The reduced QM part for UV-Vis calculations comprise 72 core atoms of PCB only (including capping atoms), whereas the propionate chains have been reduced to methyl groups.

\*\*\*AIMD /NMR

## A COMPUTATIONAL DETAILS

### A.3 Other

Python-based MDANALYSIS<sup>283,284</sup> was used for handling protein topologies, as well as for the calculation of protein RMSDs. Geometric analysis, QM/MM setups, and calculations of RDF, SDF, VDOS, as well as any post-processing for theoretical spectrum acquisition was carried out using own implementations in PYTHON 3.6<sup>285–288</sup> and numerical libraries NUMPY and SCIPY.<sup>289–291</sup> The geometric criterion for hydrogen bonds was a distance of 3.5 Å (3.0 Å in the case of hydrogen bond lifetimes in Part II) and an angle of 130 degrees. Plots were generated with MATPLOTLIB.<sup>292</sup> Molecular visualisations were created with VMD<sup>271</sup> using the Tachyon Ray Tracer<sup>293</sup> and PYMOL.<sup>294</sup> All figures were post-processed with GIMP.<sup>295</sup> This thesis was written in L<sup>A</sup>T<sub>E</sub>X.



# SASCHA JÄHNIGEN, M. SC.

Geburtsdatum 1. März 1989  
Geburtsort Karl-Marx-Stadt, jetzt Chemnitz

## WERDEGANG

10/2011 – 11/2013 Universität Leipzig  
Masterstudium Chemie; Abschluss: *Master of Science* (mark: 1.1)

07/2012 – 10/2012 Universität Granada, Spanien  
Vertiefungspraktikum (Erasmus)

03/2012 – 07/2012 Complutense-Universität Madrid, Spanien  
Master in Organischer Chemie (Erasmus)

10/2008 – 09/2011 Universität Leipzig  
Bachelorstudium Chemie, Abschluss: *Bachelor of Science* (mark: 1.3)

10/2008 – 11/2013 Mitglied der Studienstiftung des Deutschen Volkes e.V.

09/2007 – 05/2008 Zivildienst in der mobilen Behindertenhilfe, Chemnitz

09/1999 – 06/2007 Gymnasium Einsiedel, Chemnitz, Abschluss: *Abitur* (mark: 1.0)

## PUBLIKATIONSLISTE

S. Jähnigen, A. Scherrer, R. Vuilleumier, and D. Sebastiani; Chiral Crystal Packing Induces Enhancement of Vibrational Circular Dichroism, *Angew. Chem., Int. Ed.*, 2018, 57, 13344–13348.

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 spectroscopy and theory, *Phys. Chem. Chem. Phys.*, 2017, 19, 13882–13894.

M. von Domaros, S. Jähnigen, J. Friedrich, and B. Kirchner; Quantum cluster equilibrium model of (*N*)-methylformamide-water binary mixtures, *J. Chem. Phys.*, 2016, 144, 064305.

A. Schmied, A. Straube, T. Grell, S. Jähnigen, and E. Hey-Hawkins; Heterobimetallic complexes with highly flexible 1,1'-bis(phospholanoalkyl)ferrocene ligands, *Dalton Trans.*, 2015, 44, 18760–18768.