

Assessment and Application of Quantum Mechanical Methods for the Calculation of Large and Condensed Systems

Dissertation
zur
Erlangung des Doktorgrades (Dr. rer. nat.)
der
Mathematisch-Naturwissenschaftlichen Fakultät
der
Rheinischen Friedrich-Wilhelms-Universität Bonn

vorgelegt von
Sarah Löffelsender, geb. Schmitz
aus
Siegburg

Bonn, May 2025

Angefertigt mit Genehmigung der Mathematisch-Naturwissenschaftlichen Fakultät der Rheinischen
Friedrich-Wilhelms-Universität Bonn

Gutachter / Betreuer:	Prof. Dr. Stefan Grimme
Gutachter:	Prof. Dr. Marc de Wergifosse
Tag der Promotion:	03.11.2025
Erscheinungsjahr:	2025

Publication List

Parts of this thesis have been published in peer-reviewed journals.

1. Sarah Schmitz, Jakob Seibert, Katja Ostermeir, Andreas Hansen, Andreas H. Göller, and Stefan Grimme, *Quantum Chemical Calculation of Molecular and Periodic Peptide and Protein Structures*, The Journal of Physical Chemistry B, **124**.18 (2020) 3636, 10.1021/acs.jpcb.0c00549
2. Sarah Löffelsender, Marilù G. Maraldi, and Marc de Wergifosse, *All-atom quantum mechanical (AQM) methodologies for one- and two-photon absorption of realistic systems*, ChemPhysChem, **26**.11 (2025) e202401121, 10.1002/cphc.202401121
3. Jan Riebe, Benedikt Bädorf, Sarah Löffelsender, Matias Ezequiel Gutierrez Suburu, María Belén Rivas Aiello, Cristian Alejandro Strassert, Stefan Grimme, Jochen Niemeyer, *Molecular folding governs switchable singlet oxygen photoproduction in porphyrin-decorated bistable rotaxanes*, Communications Chemistry, **7**.1 (2024) 171, 10.1038/s42004-024-01247-7
4. Sarah Löffelsender, Peter Schwerdtfeger, Stefan Grimme, and Jan-Michael Mewes, *It's Complicated: On Relativistic Effects and Periodic Trends in the Melting and Boiling Points of the Group 11 Coinage Metals*, Journal of the American Chemical Society **144**.1 (2022) 485, 10.1021/jacs.1c10881

Furthermore, the following studies were performed, but are not part of the thesis.

1. Jakob Seibert, Jana Pisarek, Sarah Schmitz, Christoph Bannwarth, and Stefan Grimme *Extension of the element parameter set for ultra-fast excitation spectra calculation (sTDA-xTB)*, Molecular Physics **117**.(9-12) (2019) 1104, 10.1080/00268976.2018.1510141
2. Sarah Löffelsender, Pierre Beaujean, and Marc de Wergifosse *Simplified quantum chemistry methods to evaluate non-linear optical properties of large systems*, WIREs Computational Molecular Science **14**.1 (2024) e1695, 10.1002/wcms.1695

Abstract

Quantum mechanical (QM) calculations of large and condensed matter require the availability of fast and reliable methods. This thesis assesses several QM methodologies with challenging systems pushing the boundaries of these approaches. First, the semi-empirical quantum mechanical methods GFN n -xTB ($n = 1, 2$) are tested for the geometry optimization of large (metallo-)proteins up to 5000 atoms using an all-atom single structure quantum mechanical (ASQM) ansatz. Thus, two benchmark sets are compiled for comparison against experimental data. In general, both extended tight binding models prove to efficiently yield accurate structures of various (metallo-)proteins with and without prosthetic groups.

Besides the ASQM workflow, two all-atom dynamic structure QM approaches are assessed for the calculation of one- and two-photon absorption (1PA, and 2PA) spectra of realistic systems. Additionally, the recently introduced dt-sTD-DFT-xTB method is evaluated, which has the potential to significantly reduce computational requirements. The ASQM workflow is tested on the two challenging proteins bR and iLOV. Analysis reveals the importance of including the chromophore's environment explicitly to properly describe chromophore-protein interactions, and transitions inside tryptophan units, impacting the spectra considerably. To furthermore resolve side-features (*e.g.*, broadening), ADQM schemes are evaluated for the computation of 1- and 2PA spectra of iLOV's chromophore, the flavin mononucleotide (FMN) in aqueous solution, either relying on Boltzmann-averaged spectra with implicit solvation (ADQM-B), or using uncorrelated snapshots from a MD simulation (ADQM-MD) of explicitly solvated FMN. While the ADQM-B workflow provides little improvement over the ASQM scheme, ADQM-MD reveals striking agreement with experimental data for FMN in aqueous solution, underlining the importance of considering all atoms in an explicit dynamic manner. Next, the ADQM-B scheme is applied to reveal structural features of two bistable rotaxanes, complementing experimental efforts to explain their photochemical properties. Based on a multilevel scheme, Gibbs energies for several conformations are calculated, revealing that strong folding in the molecular structure is responsible for the unexpected photochemical behavior.

Finally, the recently proposed TI-MD- λ DFT approach is assessed to explore relativistic effects on phase transition points of the coinage metals Cu, Ag, and Au. While the boiling points (BPs) show increasing relativistic effects with increasing nuclear charge, the melting points (MPs) defy this correlation, as the MP of Au in the spin-orbit relativistic (SOR) and the non-relativistic (NR) limits are very similar. An in-depth study considering thermodynamical quantities reveals a strong stabilization of the SOR Au liquid due to relativistic effects, rendering NR Au similar to SOR Ag, confirming a half-a-century old hypothesis by P. Pyykkö.

Contents

I	Introduction and Theoretical Background	1
1	Introduction	3
2	Theoretical Background	9
2.1	Electronic Ground State Methods	9
2.1.1	Definition of the Electronic Hamiltonian and Wave Function	9
2.1.2	Wave Function Theory	11
2.1.3	Kohn-Sham Density Functional Theory	13
2.1.4	Composite Methods	15
2.1.5	Relativistic Effects	16
2.1.6	Effective Core Potentials	16
2.1.7	Semi-Empirical Quantum Mechanical Methods	17
2.1.8	Periodic Boundary Conditions	20
2.2	Excited State Calculation	21
2.2.1	Time-Dependent Density Functional Theory	21
2.2.2	Simplified Time-Dependent Density Functional Theory	22
2.2.3	One-Photon Absorption	25
2.2.4	Two-Photon Absorption	25
2.3	Gibbs Energy Calculations to obtain Melting and Boiling Points	27
2.4	Remarks on Geometry Preparation of Proteins	29
II	Studies on Semi-Empirical Quantum Mechanical Methods	31
3	Quantum Chemical Calculation of Molecular and Periodic Peptide and Protein Structures	33
4	All-Atom Quantum Mechanical (AQM) Methodologies for One- and Two-Photon Absorption of Realistic Systems	37
5	Molecular Folding Governs Switchable Singlet Oxygen Photoproduction in Porphyrin-Decorated Bistable Rotaxanes	41

III	Quantum Mechanical Studies on Relativistic Effects	45
6	It's Complicated: On Relativistic Effects and Periodic Trends in the Melting and Boiling Points of the Group 11 Coinage Metals	47
IV	Conclusions and Outlooks	51
	Bibliography	57
A	Quantum Chemical Calculation of Molecular and Periodic Peptide and Protein Structures	79
B	All-Atom Quantum Mechanical (AQM) Methodologies for One- and Two-Photon Absorption of Realistic Systems	91
C	Molecular Folding Governs Switchable Singlet Oxygen Photoproduction in Porphyrin-Decorated Bistable Rotaxanes	113
D	It's Complicated: On Relativistic Effects and Periodic Trends in the Melting and Boiling Points of the Group 11 Coinage Metals	125

Part I

Introduction and Theoretical Background

Introduction

Over the past decades, quantum mechanical (QM) simulations have played an increasingly important role in chemical investigations. Thanks to developments of QM methodologies and available high-performance computers, theoretical calculations have become a valuable tool for the prediction and explanation of experimental findings, saving resources and research time. It is established that measurable physical and chemical properties depend on the three-dimensional (3D) structure of a system, including its direct environment.¹ Several indirect experimental methods for the elucidation of a molecular structure were developed, *e.g.*, nuclear magnetic resonance (NMR), mass spectrometry, but the correct interpretation and evaluation of structure investigations yielding a full three-dimensional geometry of realistic (bio-)chemical systems with several hundreds or thousands of atoms remains demanding. Note that throughout this thesis, the term "realistic system" is employed to emphasize the importance of going beyond the single molecule or small cluster approach by explicitly including most of the environmental and (if possible) dynamic structural effects to obtain a model system approximating reality. For large, condensed matter, crystal X-ray diffraction experiments are nowadays established, leading to atomistic 3D structures, but losing important information about hydrogen atoms and consequently the total charge of the given system. Both information are crucial for theoretical research, and consequently, experimental crystal structures need to undergo several preparation steps before calculations can be conducted. Figure 1.1 depicts a currently established workflow for the preparation of proteins (see Section 2.4 for more remarks on protein preparation). Using the experimental X-ray structure as input, the user manually needs to select possible configurations of the side chains. After removing excess configurations, the hydrogen atoms are added, often by employing predetermined pK_a values of the single amino acids (AA) in water. However, pK_a values of AA in water and in protein environments can differ significantly.² Both steps are decisive for accurate investigations, but are prone to errors. In a next step, the charge (and the total spin if a metal center is involved) is determined and solvent molecules are taken into account. Such changes usually lead to tensions inside the system, and the resulting structure must be relaxed.

Geometry optimizations necessitate the use of computationally feasible routines, with a multitude of different methods being available nowadays.^{3,4} Very accurate, but computationally demanding methods like coupled cluster (CC) are only applicable to small molecules or cutouts (below 100 atoms, see Chapter 4). Density functional theory (DFT) is the workhorse of modern quantum chemistry, and is applicable to medium-sized systems, see Chapters 5 and 6 for examples. However, QM methods are generally computationally rather demanding, jeopardizing their application for biological

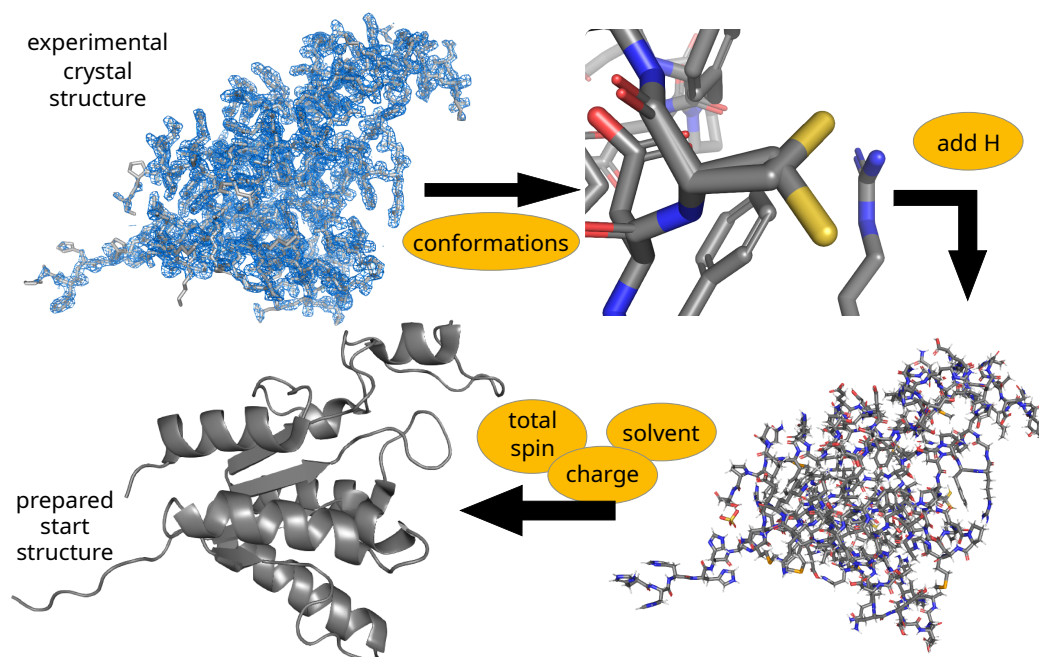


Figure 1.1: General workflow to prepare an experimental X-ray protein structure for further theoretical investigations. See text for a description of the single steps.

relevant proteins.^{5,6} Therefore, more approximate methods were developed, including quantum mechanic/molecular mechanic (QM/MM)^{7–9} approaches, ONIOM schemes,^{10–14} local fragmentation models,^{15–17} or composite methods.^{18–22} Although leading to reasonable structures with respect to experiment,^{23,24} structure optimizations may still require exhaustive computational times, *e.g.*, several months for the example shown in in Chapter 3 for the HF-3c method, and remain the bottleneck in quantum chemistry regarding structure investigations of large systems. Reducing the computational costs to a minimum while still considering an atomistic level is achieved by force fields.²⁵ Examples are the AMBER*^{26,27} and the OPLS2005²⁸ protein force fields, although many other exist.^{29–32} While being computationally efficient, specialized force fields are mainly applicable to problems similar to the fitting set and only provide limited usability for deviating situations. General force fields are applicable to most elements across the periodic table, but older variants are widely not trustworthy.³³ Note that newer versions achieved significantly improved results.³⁰ In general, polarization effects are usually missing,^{34–36} and prosthetic groups or metal atoms are often described poorly or not at all. Furthermore, chemical reactions, including proton transfers, are in many cases not possible. Force fields describe atom interactions classically, without considering electrons explicitly. In consequence, interactions of quantum nature, including hydrogen-bonds (H-bonds) or dispersion interactions responsible for many intra- and intermolecular interactions in large (bio-)systems, needs to be added using additional corrections.

A bridge between highly parametrized FF and QM are semi-empirical quantum mechanical (SQM) methods,^{37–41} which can treat all atoms in a quantum mechanical way, even when the system size exceeds 1000 atoms. The first SQM methods were based on Hartree-Fock (HF) theory,^{42–44} relying on three main simplifications which are mostly used until today: 1. diminution of basis functions by

using, *e.g.*, a minimal basis set, 2. exclusion of explicit core electrons (by implicit consideration in the parameterization) solely focusing on the chemically more active valence electrons, and 3. neglect of some atomic orbital (AO) two-electron integrals. Varying levels of simplifications and different parameterizations have lead to various flavors of HF-based SQM approaches, *e.g.*, ZINDO,⁴⁵ AM1,⁴⁶ PM3,⁴¹ or PM6,⁴⁷ to name only a few. Later on, SQM methods based on density functional theory (DFT) emerged as well, yielding the density functional tight binding (DFTB) approaches.^{40,48–51} Generally, SQM methods capture dispersion interactions inadequately, but corrections like, the empirical dispersion corrections schemes DFT-D x ($2 \leq x \leq 4$)^{52–55} can be added *a posteriori* just as for HF and DFT methods. Specific interactions such as hydrogen or halogen bonds are described poorly with SQM models due to usually small and thus inflexible basis sets. Therefore, classical corrections are commonly applied. In many cases, parameters in SQM approaches are formulated atom-pair-wise,⁵⁶ limiting their use to only the small part of possible interactions contained in the fitting set. Consequently, a high number of parameters are required, and therefore usually only few elements are parameterized, limiting the general applicability of those methods.

In 2017, a new semi-empirical quantum mechanical method, dubbed GFN-xTB (in the following GFN1-xTB)⁵⁷ was introduced, followed in 2019 by its successor GFN2-xTB.³⁷ Both approaches are parameterized in an element-specific fashion, reducing the number of required parameters and making them applicable out-of-the-box to almost the entire periodic table. Due to the inclusion of anisotropic electrostatic interactions, GFN2-xTB avoids any pair-specific or classical force-field-type parameters, which were still used in GFN1-xTB. For uncommon intramolecular interactions, the method is therefore superior to pair-wise parameterized models as it allows for the computation of binding situations beyond the fitting set (see Section 2.1.7, and refs. [37, 56, 57] for detailed information). GFN2-xTB was used as an "all-atom QM" (AQM) methodology to compute accurate (non-covalent) systems and properties for multiple chemical situations, including large organic⁵⁸ and transition metal complexes,⁵⁹ transition states,⁶⁰ redox potentials,⁶¹ and drug design.⁶² Note that the AQM terminology was first introduced in the context of interactions of light with large systems,⁶³ and that it is extended in this thesis for structural investigations of computationally demanding systems. Large, realistic protein structures were not included during the fitting procedure of GFN n -xTB ($n = 1, 2$), although geometry optimizations of proteins are possible.⁶⁴ Consequently, the quality of GFN n -xTB ($n = 1, 2$) optimized protein structures remained undetermined until 2020. One aim of this thesis was therefore to assess GFN n -xTB ($n = 1, 2$) to reproduce experimental structures of proteins with and without prosthetic groups, pushing the boundaries of this method by including metalloproteins up to 5000 atoms.⁶⁵

Knowing the strengths and weaknesses of GFN2-xTB lays the foundation for further property calculations, such as one- or two-photon absorption (1PA or 2PA, respectively) spectra. 2PA is a non-linear optical (NLO) property that was predicted in 1931 by Maria Gertrude Göppert-Mayer⁶⁶, and first observed 30 years later⁶⁷ after the discovery of laser.⁶⁸ A significant benefit of 2PA over the commonly employed 1PA is its reduced phototoxicity, as both incident photons possess an energy lower than the excited state transition energy required for 1PA. 2PA exhibits a higher degree of tissue penetration (up to 1 nm) compared to 1PA (50-80 μm in normal confocal laser scanning microscopy) and is a non-invasive imaging technique. Many medical relevant fields would benefit from designed systems with enhanced 2PA cross-sections (σ^{2PA}), such as cancer,^{69–71} kidney research,^{72–74} and neuroscience.⁷⁵ Possible applications exists in living organisms, which was demonstrated either in fixed^{76–78} or even in freely moving animals,^{79,80} but also in other research fields, such as in data storage research,^{81,82} and information processing.⁸³ Investigation of 1- and 2PA necessitates an accurate (ground state) structure.^{84–87} GFN2-xTB overcomes the computational bottleneck of

geometry optimization of large systems, but is not designed to yield a ground state wave function for further theoretical exploration of such spectroscopic properties. Including all atoms in ground and excited state calculations is computationally demanding when it comes to realistic systems, prohibiting the use of default QM methodologies, as TD-HF, or TD-DFT.^{88,89} Therefore, different simplified models for the calculation of excited state properties were developed, such as finite field,⁹⁰ sum-over-states,^{91,92} or the Tamm-Dancoff approximation.^{93,94} Another approach coined simplified Tamm-Dancoff approximation (sTDA),⁹⁵ was introduced in 2013 by Grimme *et al.*⁹⁶ sTDA originally enables for the calculation of 1PA and electronic circular dichroism (ECD) spectra in the UV-region, including very large condensed systems or a multitude of configurations. The approach relies on three simplifications, including the truncation of the single excitation space (see Section 2.2.2 and [96] for further information). One year later, Bannwarth *et al.*⁹⁷ extended this approach to the simplified time-dependent density functional theory (sTD-DFT). To gain a reliable ground state wave function, Grimme *et al.*^{98,99} developed a method for the ultra-fast calculation of simplified ground state wave functions of hybrid functional character, dubbed sTD-DFT-xTB, parameterized for parts of the periodic table. Note that GFN1- and GFN2-xTB were developed in 2017 and 2019, and are not designed to conduct further spectroscopic investigations. In 2018, sTDA-xTB and sTD-DFT-xTB¹⁰⁰ were extended to almost all elements across the periodic table, illustrating the accuracy for medium to large complex metal systems. Further response properties^{89,101} were also implemented by de Wergifosse *et al.* to the std2 package,⁹⁵ including polarizability α and first hyperpolarizability β ,¹⁰² excited state absorption (ESA),¹⁰³ spin-flip scheme,¹⁰⁴ optical rotation,¹⁰⁵ and two-photon absorption.¹⁰⁶

In 2021, Beaujean *et al.*²⁴ introduced the dual threshold (dt) sTD-DFT, further reducing computational requirements by separating the truncated singly-excited configuration space into chromophore and surroundings. They used an all-atom single structure QM (ASQM) approach to investigate the influence of explicit surroundings on the first hyperpolarizability β of the large fluorescent proteins iLOV and bacteriorhodopsin (bR). Neglecting dynamic structural effects, the dt-sTD-DFT scheme showed qualitatively similar results compared to the more expensive single-threshold sTD-DFT approach, and remarkable agreement with respect to experiment. The influence of dynamic structural effects on the first hyperpolarizability of small to medium sized tryptophan-rich peptides was investigated, too.¹⁰⁷ Performing sTD-DFT calculations for a full, Boltzmann-weighted conformer-rotamer ensemble (CRE), the implicit all-atom dynamic structure QM workflow (ADQM-B) compares well to experimental investigations. The influence of explicit dynamic structural effects, *e.g.*, incorporated through a molecular dynamics (MD) simulation (ADQM-MD), were additionally explored, albeit neglecting explicit surroundings.

However, 2PA spectra for realistic systems using AQM schemes need further benchmarking. Calculating them, though, is an inherently difficult task, as it is formally accessible as a third-order molecular response property.^{108,109} Additionally, excited states of conjugated systems are commonly affected by the bond length distribution as well as bond angles of the structure.⁸⁴⁻⁸⁷ Last but not least, the close vicinity of a chromophore modifies spectra considerably, making a realistic representation of the surroundings essential for accurate theoretical calculations.^{1,63,110} In the literature, many theoretical 2PA investigations therefore rely on QM/MM schemes,¹¹¹⁻¹²² introduced already in 1976 by Warshel *et al.*,⁷ to elucidate environmental effects. The part of the system considered responsible for the target property (usually the chromophore) is computed with a higher level QM method, while the description of the environment (a protein or surrounding solvent molecules) relies on computationally more feasible classical MM routines. However, interactions with quantum nature, *i.e.*, long range, or directional non-covalent interactions are particularly difficult to describe at the intersection between

QM and MM parts.¹²³ Note that several protocols to overcome these shortcomings exist, although this issue still remains relevant.²³ With the xTB family of methods at hand, different AQM methodologies for various excited state properties including the environment and dynamic structural effects explicitly are feasible. Nevertheless, combination of multiple SQM approaches in the computation process of excited state features needs thorough investigations to identify the most practical workflow without compromising accuracy. Consequently, ASQM, ADQM-B, and ADQM-MD workflows are assessed for the computation of 1- and 2PA spectra of realistic systems (see Chapter 4).

Pushing the computational boundaries further, the ADQM-B scheme can also aid in the investigation of possible structures on the potential energy surface (PES) of functional molecular switches. Those systems feature different stable states, which can be selectively addressed to change their molecular properties by an external stimulus.^{124–129} Among others, molecular muscles and elevators,^{130–133} nanovalves,^{134,135} nanorobots,¹³⁶ switchable catalysts,^{137–141} and materials with switchable photophysical properties^{142–144} should be mentioned here. A relevant class of supramolecular complexes in this context are rotaxanes, which are mechanically interlocked systems with a macrocycle encircling a thread with multiple binding stations.^{128,129,145–147} The modification of one binding station results in a significant motion of the macrocycle to the other station, and vice versa. Consequently, the stations are shielded/revealed, relevant in the domain of catalysis^{140,141}, or the surroundings of attached groups exhibit substantial changes. Accordingly, photophysical properties of chromophores attached to the thread can be tuned, enabling *e.g.*, switchable fluorescence.¹⁴⁸ The tunable functions are intrinsically linked to the molecular conformation in each state, and predicting the dominant three-dimensional geometry can be a formidable challenge during the design process, as folded conformations may occur.¹³⁷

One popular motif for a binding station is an ammonium/amine group, which is preferred when protonated. During deprotonation, the second binding station, *e.g.*, a triazolium, is favored. Using this combination of binding stations results in co-conformations being predominant by > 99 %, making this rotaxane effectively binary.¹⁴⁹ Depending on the relative affinity, and guided through external stimulus, the macrocycle translocates from one to the other binding station. One possible application area for this rotaxanes, relevant in the medical field,^{150–152} or in organic synthesis,^{153–155} is the controllable singlet dioxygen (¹O₂) photoproduction (for simplicity named singlet oxygen production, or ¹O₂ production in the following). To reduce toxicity, and increase reactivity, the rotaxanes in the *off*-state should suppress ¹O₂ production completely, while the *on*-state should yield the desired agent. In a simplified picture, a black-hole quencher (BHQ) should suppress the ¹O₂ production of the photosensitizer (PS) when they are in close proximity to one another, while a large distance between BHQ and PS leads to the *on*-state. The synthesis of such a switchable rotaxane¹⁵⁶ resulted indeed in a separate *on*- and *off*-state, but vice versa to the original design. Additionally, the ¹O₂ quantum yield was considerably lower in the rotaxanes compared to the free PS, and the thread length has been identified as a significant factor in the ¹O₂ production. In the context of such highly flexible systems, it is essential to account for molecular folding, which is where theoretical ADQM-B simulations can be applied to screen the large PES using the efficient GFN2-xTB method (see Chapter 5). For more reliable data, the CRE should further be refined with higher level DFT calculations, yielding final Gibbs energies with a multilevel scheme, enabling the profound analysis of the conformational space.

Another part of this thesis focuses on the exploration of relativistic effects on melting and boiling points (MPs and BPs, respectively) for atomistic systems. Until today there exists no generally accepted microscopic explanation how phase transitions, *i.e.*, melting of ice, occur. To reveal the mechanisms, quantum mechanical calculations could assist, but are inherently difficult, as inter- and

intramolecular interactions between all atoms or molecules involved need to be described properly.¹⁵⁷ Only heating up (or cooling down) a simulation box using MD or Monte-Carlo approaches describe the macroscopic picture insufficiently, as possible defects, different solid-state structures, super-heating or -cooling, different reaction paths, large volume changes, etc. influence phase transitions significantly. Especially for the boiling point, large volume changes occur,¹⁵⁸ increasing the *per se* challenging task of computing phase transitions. Additionally, the simulation of covalently bonded or metallic solids requires an accurate description of electron correlation, a computational task that is particularly demanding for bulk systems.^{159–161} Despite the prevalent application of DFT in calculating phase transitions, the performance of density functional approximations (DFAs) for bulk properties is element-specific,^{162–164} limiting their comprehensive utilization.

Instead of directly simulating phase transition processes, Mewes *et al.*^{165–167} proposed an indirect method called TI-MD- λ DFT, yielding phase transition points by the intersection of Gibbs energy curves. Gibbs energies at specific temperatures are obtained by combination of MD simulations at a DFT level of theory with thermodynamic integration (TI), and thermodynamic perturbation theory (TPT). Moreover, an empirical scaling of the Hamiltonian is applied to mitigate above mentioned DFA-related systematic errors. Those Gibbs energies are subsequently extrapolated for each phase, with their intersections giving the respective phase transition points. It has been shown for closed-shell Group 12 elements Zn, Cd, and Hg that this approach is able to correctly predict MPs and BPs, and furthermore, it could clarify the influence of relativistic effects on the anomalie of the MP of Hg.¹⁶⁶ A multitude of relativistic effects associated with Au have been examined, including the phenomenon of its characteristic yellow color,¹⁶⁸ the extraordinary high electron affinity (~ 2.4),¹⁶⁹ the occurrence of the oxidation state +5,¹⁷⁰ or aurophilic interactions.¹⁷¹ However, the impact of relativistic effects on the MPs and BPs of the coinage metals Cu, Ag, and Au requires further investigation. Pushing the boundaries of the TI-MD- λ DFT approach, this workflow is therefore applied to the electronically complicate case of open-shell Group 11 elements (cf. Chapter 6) by computing their non-relativistic, scalar relativistic and spin-orbit relativistic phase transition points.

In the following, Chapter 2 introduces theoretical models used in this thesis, concentrating on electronic structure ground and excited state theory, as well as the calculation of Gibbs energy curves for atomistic systems. Then, the main research results are presented in Parts II and III while the full publications are printed in the appendices (A to D). In Part II, studies on large systems are conducted using semi-empirical methods. First, the SQM method GFN2-xTB is assessed for the description of proteins, using an all-atom single structure approach in Chapter 3. Second, the influence of the environment as well as dynamic structural effects on 1- and 2PA spectra is explored in Chapter 4. In Chapter 5, investigations of the conformational space of large organic molecules are shown. In the next Part III, the influence of relativistic effects on phase transitions is evaluated in Chapter 6. Finally, a summary of the achievements of this thesis, and a perspective regarding future studies are given in Part IV.

Theoretical Background

This chapter provides an overview over different theoretical concepts applied in this thesis. First, different electronic ground state methods (see Section 2.1) are introduced. Subsequently, the calculation of excited states for 1- and 2PA spectra is presented Section 2.2. Then, Gibbs energy calculations for the computation of melting and boiling points is shown in Section 2.3. The primary sources for this chapter are references [172, 173], which also provide a more detailed description for interested readers. Moreover, references [37, 56, 63, 89, 96, 97, 165, 166] were utilized for recently developed and validated methodologies.

2.1 Electronic Ground State Methods

This section gives a brief overview over commonly applied electronic ground state methods. Additionally, relativistic effects, effective core potentials and periodic boundary conditions are briefly examined. Furthermore, the semi-empirical tight binding method GFN2-xTB is introduced. If not stated otherwise, atomic units are used in all equations.

2.1.1 Definition of the Electronic Hamiltonian and Wave Function

Modern ground state methods calculate the total energy E of a system by applying the operator $\hat{H}$ (termed Hamiltonian) to the wave function Ψ , *i.e.*, connecting the system's structure to its energy.

$$\hat{H}\Psi = E\Psi \quad (2.1)$$

In this so-called stationary Schrödinger equation,¹⁷⁴ the Hamiltonian is not dependent on time (see Section 2.1.5 for the time-dependent variant). First, details about the time-independent Hamiltonian $\hat{H}$ are given, before the wave function Ψ is discussed. $\hat{H}$ consists of the kinetic $\hat{T}$ and the potential $\hat{V}$ energy of electrons e and nuclei n .

$$\begin{aligned} \hat{H} &= \hat{T} + \hat{V} \\ &= \hat{T}_e + \hat{T}_n + \hat{V}_{ee} + \hat{V}_{en} + \hat{V}_{nn} \end{aligned} \quad (2.2)$$

The potential energy $\hat{V}$ represents the Coulomb interaction between two particles (electron-electron: $\hat{V}_{ee}$, nucleus-nucleus: $\hat{V}_{nn}$ or electron-nucleus: $\hat{V}_{en}$), while $\hat{T}$ represents the kinetic energy which can be split into the kinetics of electrons and nuclei, respectively ($\hat{T}_e$, and $\hat{T}_n$). For all elements, the nucleus is much heavier than their respective electrons, *e.g.*, the mass is 1800 times higher in the smallest possible nucleus H, and already 430 000 times higher in uranium. Consequently, Born and Oppenheimer¹⁷⁵ introduced two approximations, based on the assumption that the coupling between the motion of electrons and nuclei can be disregarded. From this follows: i) the kinetic energy of the nuclei T_n can be neglected, as the velocity of the electrons is much higher than that of the nuclei, and ii) the Coulomb potential V_{nn} between two nuclei is a constant, and thus needs to be calculated only once. For the majority of chemical interesting cases, the Born-Oppenheimer approximation leads to negligible errors ($\sim 10^{-4}$ au for H_2 and smaller for heavier atoms)^{172,176}, and is therefore used almost universally. Throughout this thesis, it is applied, too.

Consequently, the Hamiltonian is separated into nuclei-nuclei interactions and electron-dependent terms (see eq. 2.3).

$$\hat{H} = \hat{H}_{el} + V_{nn} \quad (2.3)$$

The electronic Hamiltonian $\hat{H}_{el}$ consists of interactions including electrons, *i.e.*, $\hat{T}_e$, $\hat{V}_{ee}$ and $\hat{V}_{en}$.

$$\hat{T}_e = -\frac{1}{2} \sum_i^{N_{el}} \nabla_i^2 \quad (2.4)$$

$$\hat{V}_{ee} = \sum_i^{N_{el}} \sum_{j>i}^{N_{el}} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} \quad (2.5)$$

$$\hat{V}_{en} = \sum_i^{N_{el}} \sum_A^{N_{at}} \frac{Z_A}{|\mathbf{r}_i - \mathbf{R}_A|}. \quad (2.6)$$

N_{el} and N_{at} are the total number of electrons and nuclei, respectively, Z_A is the nuclear charge of nucleus A, and $\mathbf{r}_i$, $\mathbf{r}_j$, and $\mathbf{R}_A$ are the coordinates of electron i and j , and of nucleus A, respectively. ∇_i^2 , the Laplace operator, denotes the second derivative in all three directions, *i.e.*, $\nabla_i^2 = \frac{\partial^2}{\partial x_i^2} + \frac{\partial^2}{\partial y_i^2} + \frac{\partial^2}{\partial z_i^2}$.

It might be useful to separate the one ($\hat{h}_i = \hat{T}_e + \hat{V}_{en}$) and two-electron ($\hat{V}_{ee}$) integrals, with the latter being the bottleneck in many calculations. In total, the electronic Schrödinger equation then reads:

$$\hat{H}_e \Psi_e = \left(\sum_i^{N_{el}} \hat{h}_i + \sum_i^{N_{el}} \sum_{j>i}^{N_{el}} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} \right) \Psi = E_e \Psi_e. \quad (2.7)$$

All in all, the desired quantity E_{tot} (the total energy), is the (constant) Coulomb energy V_{nn} and the expectation value of the electronic Hamiltonian (the total electronic energy E_e), here written in the Dirac notation.

$$E_{tot} = E_e + V_{nn} \quad (2.8)$$

$$= \langle \Psi_e | \hat{H}_e | \Psi_e \rangle + \sum_A^{N_{at}} \sum_{B>A}^{N_{at}} \frac{Z_A Z_B}{|\mathbf{R}_A - \mathbf{R}_B|} \quad (2.9)$$

Note that only electronic terms are considered further, therefore the index e is omitted from now on.

After defining the Hamiltonian, the (electronic) wave function Ψ is now introduced briefly. To put it in a nutshell, Ψ provides a mathematical connection between the spatial positions and spins of the electrons of a given system to its (electronic) energy. In quantum mechanics, an electron is represented by one-particle wave function $\phi(x)$, also called orbital or spin wave function. The latter has four-dimension, three spacial components x, y and z , and a spin component $\sigma_i \in \alpha, \beta$, as the electron as fermion is known to have a spin of $\frac{1}{2}$.

$$\phi_i(x) = \psi_i(r)\sigma_i \quad (2.10)$$

For the hydrogen atom, comprising a positive and a negative charge, the Schrödinger equation can be solved exactly, yielding the well-known hydrogenic orbitals. Orbitals of elements with more than one electron adopt, in principle, the similar form, and are commonly approximated with hydrogenic orbitals. When computing integrals, the mathematical expressions of hydrogenic orbitals in multi-electron systems are extremely demanding, therefore simpler formulations are mainly used in quantum mechanical calculations: Slater-type and Gaussian-type orbitals (STO, and GTO, respectively). The former are accurate solutions of the Schrödinger equation for hydrogen-like systems,¹⁷⁷ but are computationally challenging for more than one- or two-center two-electron integrals. Therefore, the less demanding GTOs are commonly used. However, a single GTO is insufficient to model the orbital accurately, due to its inability to reproduce the nuclear cusp and its overly rapid decay at larger distances from the nucleus. Therefore multiple GTOs are contracted (also called linear combination), to simulate the shape of one STO. It was found that at least three GTOs are required to resemble one STO.

The wave function needs to be antisymmetric, and the Pauli principle^{177,178} is a direct consequence, stating that two electrons need to differ in at least one quantum number. Therefore, the wave function is commonly represented as (normalized) Slater Determinant (SD), consisting of N orthonormal one-electron wave-functions (or MOs), which are occupied by N electrons.

$$\Psi_0 \approx \Phi^{SD}(1, 2, \dots, N) = \frac{1}{\sqrt{N!}} \begin{vmatrix} \phi_1(x_1) & \phi_1(x_2) & \dots & \phi_N(x_1) \\ \phi_1(x_2) & \phi_2(x_2) & \dots & \phi_N(x_2) \\ \vdots & \vdots & \ddots & \vdots \\ \phi_1(x_N) & \phi_2(x_N) & \dots & \phi_N(x_N) \end{vmatrix} \quad (2.11)$$

In many electronic ground state methods, only one SD is taken, as *e.g.*, in the Hartree-Fock (HF) (see Sec. 2.1.2), or in density functional theory (DFT, see Sec. 2.1.3), which is a good approximation for many systems.

2.1.2 Wave Function Theory

Hartree-Fock One established approach to translate eq. 2.1 to practically solvable equations is the Hartree-Fock (HF) theory.¹⁷⁹ Although it lacks crucial electron correlations, it is the base for many other, more sophisticated methods like coupled cluster (CC)^{172,180–183} or configuration interaction (CI)^{172,173,180}. As they are merely touched in this thesis, the reader is referred to references [172, 173] for an introduction of these methods. HF approximate the exact total energy of eq. 2.1, applying the variational principle, *i.e.*, the resulting HF energy E^{HF} must be larger or equal to the exact energy E .

Note that equality would only hold for an exact wave function and considering all interactions. The HF equivalent of the Schrödinger eq. 2.1 for the computation of the electronic energy is written in eq. 2.12.

$$\hat{f}_i \phi_i = \epsilon_i \phi_i \quad (2.12)$$

The Hamiltonian $\hat{H}$ of eq. 2.1 is replaced by the Fock operator $\hat{f}$ (eq. 2.13), the wave function Ψ is approximated by molecular spin orbitals ϕ (eigenfunctions of $\hat{f}$), and the total electronic energy E_e is estimated by the sum over all orbital energies ϵ_i (eigenvalues of $\hat{f}$). $\hat{f}$ takes the form:

$$\hat{f}_i = \hat{h}_i + \sum_j^{N_{el}} (\hat{J}_{ij} - \hat{K}_{ij}), \quad (2.13)$$

where $\hat{h}$ is the one-electron operator accounting for $\hat{T}_e$ and $\hat{V}_{en}$, and the interaction between two electrons $\hat{V}_{ee}$ is estimated by the difference of the Coulomb J_{ij} and the exchange term K_{ij} . The exchange energy K does not have a classical counterpart, but is commonly described as the lowering of the energy due to the hypothetical exchange of two indistinguishable particles (*e.g.*, two electrons with the same spin). In HF, the Coulomb energy J is the interaction of an electron with the mean field of all other electrons, *i.e.*, the approximated repulsion between all electrons of a system. This mean-field approximation is the main ingredient in HF and is introduced to avoid the many-body problem. Consequently, no explicit electron correlation effects are accounted for in HF. In passing, note that HF is free of the self-interaction error (SIE), as the Coulomb-type self-interaction J_{ii} is equal to the exchange-type self-interaction K_{ii} . $\hat{h}$, $\hat{J}$ and $\hat{K}$ take the form:

$$\hat{h}|\phi_i\rangle = -\frac{1}{2}\nabla_i^2 - \sum_A^{N_{at}} \frac{Z_A}{|\mathbf{r}_i - \mathbf{R}_A|} |\phi_i\rangle \quad (2.14)$$

$$\hat{J}_j|\phi_i\rangle = \langle\phi_j|\frac{1}{r_{ij}}|\phi_j\rangle|\phi_i\rangle = \left(\int \phi_j^* \frac{1}{r_{ij}} \phi_j dx_j\right) |\phi_i\rangle \quad (2.15)$$

$$\hat{K}_j|\phi_i\rangle = \langle\phi_j|\frac{1}{r_{ij}}|\phi_i\rangle|\phi_j\rangle = \left(\int \phi_j^* \frac{1}{r_{ij}} \phi_i dx_j\right) |\phi_j\rangle \quad (2.16)$$

In HF (and many other QC methods), atomic spin orbitals (AOs) χ_i are nucleus-centered. Linear combination of all N_{AO} atomic orbitals (LCAO) yields the molecular orbitals (MOs) ϕ of the system and filled with the electrons following the "Aufbau" principle.

$$\phi_i = \sum_{\mu}^{N_{AO}} C_{\mu i} \chi_{\mu} \quad (2.17)$$

The shape of the MOs depends on the underlying system, therefore the coefficients $C_{\mu i}$ are optimized during the electronic ground state calculation such that the resulting wave function yields a minimum total energy (see eq. 2.1). Translating eq. 2.12 into a linear matrix equation yields the Roothaan-Hall formula 2.18.^{184,185}

$$\mathbf{FC} = \mathbf{SC}\epsilon \quad (2.18)$$

$\mathbf{F}$ is the Fock matrix (eq. 2.19), $\mathbf{C}$ consists of the coefficients of eq. 2.17, and $\mathbf{S}$ is the overlap matrix,

written in eq. 2.20. The latter is inserted to ensure orthonormal orbitals.

$$F_{\mu\nu} = \langle \chi_\mu | \hat{f} | \chi_\nu \rangle \quad (2.19)$$

$$S_{\mu\nu} = \langle \chi_\mu | \chi_\nu \rangle \quad (2.20)$$

The Fock matrix depends on the coefficient $C_{\mu i}$ through the density matrix P , therefore eq. 2.18 needs to be solved iteratively. This procedure is called self-consistent field (SCF) approach and depends on linear variation of the MO coefficients $C_{\mu i}$. After reaching a predefined threshold, the final Hartree-Fock energy E^{HF} is obtained with the converged, self-consistent coefficients as shown in eq. 2.21.

$$E^{HF} = \sum_i^{N_{el}} \langle \phi_i | \hat{h}_i | \phi_i \rangle + \sum_i^{N_{el}} \sum_{j>i}^{N_{el}} (\langle \phi_i | \hat{f}_{ij} | \phi_i \rangle - \langle \phi_i | \hat{K}_{ij} | \phi_i \rangle) \quad (2.21)$$

Formally, HF scales with $O(N^4)$ (N equals number of basis function), although modern codes decrease the actual computational effort.

Correlation As outlined before, HF is a mean-field theory. Consequently, the potential energy between two explicit electrons V_{ee} is approximated by the interaction of one electron with the average field of all other electrons. Due to the Pauli principle, electrons with the same spin interact within HF (Fermi correlation), but electrons with opposite spin are uncorrelated (Coulomb correlation). Therefore, the HF solution is always an approximation to the exact energy, even when considering a complete basis set (CBS). The missing correlation energy E_{corr} (eq. 2.22) is normally about 1% of the total energy.

$$E_{corr} = E^{exact} - E^{HF} \quad (2.22)$$

Although this seems negligible, for many properties, relative energies, and sometimes even covalent bonds, it plays a crucial role. Additionally, London dispersion is poorly described within HF, as it is a long-range interaction. As mentioned above, CI and CC methods improve upon the HF ground state by taking into account multiple electron configurations. Up to now, the most notable approach is the CCSD(T) ansatz ("gold standard" in computational chemistry),^{180,183} which truncates the cluster operator used in coupled cluster theory after double excitations, and computes the triple excitations perturbatively. Although much more accurate than HF, CCSD(T) scales formally with $O(N^7)$, making it only feasible for relatively small molecules.

2.1.3 Kohn-Sham Density Functional Theory

Nowadays, the workhorse of quantum chemistry is Kohn-Sham Density Functional Theory (KS-DFT). It is based on the Hohenberg-Kohn theorem,¹⁸⁶ which is split into two statements: i) the electronic energy of the ground state of a system is determined by its electron density $\rho(\mathbf{r})$. ii) Variation of the electron density leads to higher energies than that of the ground state, *i.e.*, the ground state electronic energy can be approximated using variations of a test electron density. Although proven in 1964,¹⁸⁶ the authors did not give any recipe for constructing such a functional, and its form remains unknown until today. This "orbital-free" functional is only dependent on three variables, but leads to low accuracy comparable to HF, and is merely used today.^{187,188} One of the issues is the unknown kinetic energy

$T_e[\rho]$, for which no exact expression exists. Therefore, Kohn and Sham¹⁸⁹ (re-)introduced the concept of orbitals, by using the assumption of non-interacting electrons. This divides $T_e[\rho]$ into the kinetic energy $T_S[\rho]$ (eq. 2.23, computed using a Slater determinant in the "non-interacting limit"), and a small remaining term, which is absorbed in the exchange-correlation energy $E_{XC}[\rho]$ (eq. 2.24). This exchange-correlation term is the only approximate term in KS-DFT (Kohn-Sham DFT). Note that orbitals in KS-DFT generally apply the LCAO approximation introduced in Section 2.1.2. Further known ingredients of KS-DFT are the Coulomb interaction between electrons and nuclei E_{ne} (eq. 2.25) and between electrons and electrons in J (eq. 2.26).

$$T_S = \sum_i^{N_{el}} \left\langle \phi_i \left| -\frac{1}{2} \nabla^2 \right| \phi_i \right\rangle \quad (2.23)$$

$$E_{XC}[\rho] = E_X[\rho] + E_C[\rho] \quad (2.24)$$

$$E_{ne}[\rho] = - \sum_A^{N_{at}} \int \frac{Z_A \rho(\mathbf{r})}{|\mathbf{r}_A - \mathbf{r}|} d\mathbf{r} \quad (2.25)$$

$$J[\rho] = \frac{1}{2} \int \int \frac{\rho(\mathbf{r})\rho(\mathbf{r}')}{|\mathbf{r}_A - \mathbf{r}|} d\mathbf{r} d\mathbf{r}' \quad (2.26)$$

All in all, the total energy in the KS-DFT approach reads:

$$E[\rho] = T_S + E_{ne}[\rho] + J[\rho] + E_{XC}[\rho] \quad (2.27)$$

There are two major drawbacks in KS-DFT, which current developments try to improve: i) The exact form of the exchange-correlation functional is unknown. ii) The Coulomb and exchange terms for identical electrons, J_{ii} and K_{ii} , do not cancel, as it was the case for HF. Consequently, an electron interacts with itself in KS-DFT, an unphysical behavior, which is called self-interaction error (SIE). The influence of E_{XC} on the total energy is small compared to the influence of $T_e[\rho]$ in orbital-free DFT, nevertheless it is the critical ingredient in density functional approximations (DFAs) regarding their accuracy, and even can lead to (un)bound states, where the opposite should be the case.^{190,191} Over the past decades, developments concentrated on the improvement of exchange-correlation functionals, resulting in a plethora of different DFAs.⁴ Each has its own strengths and weaknesses - there is no universal functional, and the DFA needs to be selected carefully considering from the beginning the system of interest and the target property. Note that throughout this thesis, DFT always refers to KS-DFT.

DFAs can be ranked according to their description of the density using the "Jacob's ladder" of DFT.¹⁹² The simplest DFA, which is on the first and lowest rung, is based on the uniform electron gas. Here, the density varies only slowly, leading to the local spin density approximation (LSDA). It describes some metallic systems well, but is merely used to compute molecular systems due to its overall poor performance. The generalized gradient approximation (GGA) improves the density description by taking into account the first derivative of the density $\Delta\rho$, placing it on the second rung of Jacob's ladder. Consequently, the electron density can adapt better to heterogeneous systems. Therefore, it is often applied for fast QC calculations, with a scaling of formally $O(N^3)$. Prominent GGA functionals are PBE^{193,194}, BP86^{195,196}, and PBEsol^{197,198}. On the third rung, higher derivatives of the density $\nabla^2\rho$ or the kinetic energy density τ . These *meta*-GGA functionals generally improve the

accuracy compared to GGAs although formally scaling similar as GGAs ($O(N^3)$.) Famous examples are TPSS¹⁹⁹, and r^2 SCAN²⁰⁰.

Hybrid functionals, on the fourth rung, incorporate exact HF exchange using KS orbitals (eq. 2.28).

$$E_{XC}^{hybrid} = (1 - a_x)E_X^{(m)GGA} + a_x E_X^{HF} + E_C^{(m)GGA} \quad (2.28)$$

a_x equals the amount of exact exchange from the HF approach that is mixed in, and is either determined based on the adiabatic connection formalism, or empirically. Interpolating the non-correlated HF approach and the correlated DFT relying on the adiabatic connection often leads to an improved description of systems that are prone to the SIE, although this comes to a higher computational cost ($O(N^4)$). Examples are PBE0,²⁰¹ or the B3LYP^{202,203} functional. Double hybrid (DH) functionals occupy the last (fifth) rung and include virtual orbitals by incorporating many-body perturbation theory expressions, usually applying the Møller-Plesset approximation obtained at the second order (MP2). DHs increasing the formal scaling to $O(N^5)$, but improving the description of electron correlation.¹⁷² Prominent examples are the B2PLYP²⁰⁴ and the PWPB95²⁰⁵ functional.

The introduced DFT methods normally suffer from the overlocalization problem, which means that delocalized systems are favored over localized systems. Therefore, range-separated (also called long-range corrected) functionals were developed, aiming at an improvement of the overlocalization problem. In this class of functionals, the exchange energy is divided into a short- and a long-range part, usually represented by a density exchange functional and the exact HF expression, respectively. Partition is usually done by an error function $erf(\omega r_{12})$, which depends on the distance r_{12} of two electrons, and the parameter ω , which is usually determined by fitting to experimental data (typically 0.30 – 0.50/bohr). The ω B97X²⁰⁶ functional is to be mentioned here, which is relied on (with a dispersion correction) in this thesis, too. In practice, range-separated DFAs are particularly valuable in difficult cases such as for excited state calculations, for cases of loosely bound electrons, and therefore for charge-transfer and Rydberg states, or for structure predictions of zwitterionic structures.

2.1.4 Composite Methods

DFT methods are in general more accurate than HF, but still suffer from shortcomings (*vide supra*). Especially the combination of DFAs with small basis sets results in efficient but often inaccurate calculations. Therefore, composite methods like HF-3c and DFT-3c were developed, applying three corrections to the underlying ground state method.^{18–22,207,208} All of them are computationally more efficient than their underlying ground state method in the basis set limit, as they apply relatively small (modified) basis sets. It is known that small basis sets leads to the basis set superposition error (BSSE), as all atoms try to maximize their basis sets (*i.e.*, their description) by "borrowing" basis functions of nearby atoms. Therefore, most 3c methods use the geometrical counterpoise (gCP) correction^{209,210} to decrease the influence of the BSSE. Additionally, short-ranged basis set (SRB) errors are improved, which mainly influence covalent bond lengths of many (*meta*)GGAs.

Last but not least, all composite approaches use a London dispersion correction (D3⁵⁵ or D4²¹¹) to enhance the description of non-covalent interactions (NCI). London dispersion effects are part of attractive Van-der-Waals (VdW) interactions in the long-range region of a system, and they are crucial in non-covalently bound systems like the benzene-dimer.^{190,191} As mean-field methods like HF and (semi-)local KS-DFT are not able to describe such long-range correlation effects, Grimme and coworkers developed corrections, termed DFT-Dn ($n = 1, 2, 3, 4$),^{52–55,211,212} which are added a

posteriori to a calculation. All of them come at negligible computational cost compared to the ground state approaches, but improve their description of NCIs significantly.

2.1.5 Relativistic Effects

Until now, all electronic ground state methods are based on a non-relativistic Hamiltonian. For most calculations concerning organic chemistry, this approximation yields reasonable results, but if heavy atoms are included in a calculation, relativistic effects influence the physical and chemical behavior significantly. The famous color of gold is commonly known, which appears only due to relativistic effects.¹⁶⁸ Other examples are the low melting point of mercury,²¹³ the extraordinary high electron affinity of gold,¹⁶⁹ or the HALA (heavy atom on the light atom) effect in NMR.²¹⁴ In Chapter 6, the influence of relativistic effects on the melting and boiling points of Au is investigated.²¹⁵

Their base are core-electrons which move at a non-negligible fraction of the speed of light c (usually $> 1/3 \cdot c$),¹⁷² leading to a substantial increase of the atom's effective mass. The $1s$ orbital contracts by the same factor as the effective mass increases, affecting all other s -orbitals to preserve orthogonality. This normally leads to shorter bond lengths. Furthermore, the size of d - and f -orbitals increase and thus, they become more diffuse.

As this is a considerable change in the electronic structure of an heavy atom, relativistic effects need to be considered in their simulations for the correct description of their chemical behavior. Therefore, the non-relativistic Schrödinger equation is exchanged by the relativistic Dirac equation 2.29.

$$\hat{H}_{Dirac} \Psi = E \Psi \quad (2.29)$$

$$\hat{H}_{Dirac} = c \boldsymbol{\alpha} \cdot \mathbf{p} + \beta m c^2 + \mathbf{V} \quad (2.30)$$

α and β are 4×4 matrices, $\mathbf{p}$ is the momentum operator, m resembles the relativistic electron mass, and $\mathbf{V}$ is an electrostatic potential. Note that in the non-relativistic limit, *i.e.*, light particles with small velocities, the Dirac Hamiltonian simplifies to the Schrödinger Hamiltonian.

$\hat{H}_{Dirac}$ contains different relativistic terms, next to default non-relativistic kinetic and potential energy operators. Two of them, the *mass-velocity* and the *Darwin* correction increase the computational complexity only slightly, and yield together the *scalar-relativistic* correction. The *mass-velocity* correction couples velocity and mass of an electron, *i.e.*, increasing speed correlates to an increase of the effective mass. The *Darwin* correction describes high frequency oscillations around an electron's mean position, called *Zitterbewegung*. The third relativistic term describes *spin-orbit relativistic* (SOR) interactions. In short, they occur due to the interaction of the electron's spin with its self-induced magnetic field due to its movement. Including SOR interactions significantly increases computational demands.

2.1.6 Effective Core Potentials

Elements in the lower part of the periodic table require increasing computational effort as the number of (core) electrons increases. Additionally, relativistic effects become more important (many core electrons). However, chemistry mostly takes place at the valence orbitals, and reducing computational requirements is often done using *Effective Core Potentials* (ECPs), also called *Pseudo-Potentials* (PPs). In a nutshell, core electrons are treated with a pre-computed potential (ECP), while only valence electrons (or electrons in high-lying orbitals) are considered explicitly in the quantum

chemical calculation. Pseudo-potential are able to capture relativistic effects by design (mostly scalar interactions) without running a full relativistic calculation. It is furthermore possible to include spin-orbit effects in ECPs (increasing the computational time), or to exclude all relativistic effects to yield non-relativistic ECPs, as done in Chapter 6. Therefore, the influence of relativity on the property of interest can easily be investigated with ECPs.

2.1.7 Semi-Empirical Quantum Mechanical Methods

Popular DFAs can nowadays compute properties for a couple of hundred atoms, depending on the task. Geometry optimization, molecular dynamic (MD) simulations, or the generation of conformer/rotamer ensembles (CREs), however, are often computationally too demanding for *ab-initio* methods, and even efficient composite DFT methods reach their limit of what is computationally feasible. In this context, semi-empirical quantum mechanical (SQM) methods arose, based either on HF, or on DFT. For HF-based approaches, NDDO^{216,217}, ZINDO⁴⁵, and MNDO²¹⁸ are to be mentioned here. The DFT-based density functional tight binding (DFTB)^{37,49,50} is briefly introduced in the following.

Density Functional Tight Binding Different parameterizations exist, but all DFTB methods have in common: i) the use of a minimal basis set, ii) neglect of three- and four-center integrals, and iii) treatment of valence electrons only, while core electrons (and their interactions) are implicitly accounted for in the parameterization. In short, DFT evaluates the electron density of a given structure, while DFTB examines differences of the electron density with respect to a precomputed reference state of neutral, non-interacting atoms.

$$\rho = \rho_0 + \Delta\rho \text{ with } \rho_0 = \sum_A \rho_0^A, \quad (2.31)$$

Therefore, energy differences are described with density fluctuations $\delta\rho$, expanded in a power series.

$$E[\rho] = E^{(0)}[\rho_0] + E^{(1)}[\rho_0, \delta\rho] + E^{(2)}[\rho_0, (\delta\rho)^2] + E^{(3)}[\rho_0, (\delta\rho)^3] + \dots \quad (2.32)$$

Truncation after different orders lead to different DFTB flavors, *i.e.*, truncating after 1st order is called DFTB1,⁴⁸ truncating after the 3rd order is called DFTB3.⁴⁰ From the 2nd order onwards, self-consistent charge (SCC) is reached.⁵¹ Dispersion is not accounted for, but can be added empirically just as for WFT/DFT. Furthermore, specific interactions as *e.g.*, hydrogen-bonds H-X or halogen bonds R-X can be corrected, as they are normally described poorly within the DFTB approach.

GFN2-xTB One recently proposed DFTB3 variant is the so called GFN2-xTB (*geometry, frequency, non-covalent interactions - extended tight binding*) method,³⁷ which was benchmarked and afterwards applied to gain results in this thesis. Just as its precursor GFN1-xTB⁵⁷ (see ref. [56] for comparison of all GFN methods), GFN2-xTB is a special purpose TB method, to quickly compute geometries, frequencies and non-covalent interactions for systems with up to thousands of atoms. One main feature in GFN2-xTB is the exclusion of any pair specific parameters, which are either global or element-specific. GFN2-xTB is parameterized for all elements up to radon, and consequently widely applicable to many chemical challenges. It employs a minimal valence basis set of atom centered, contracted Gaussian functions, and additionally adds polarization functions for some main group elements (see Table 2.1 for details). Core electrons are treated implicitly via the parameterization.

One novelty in GFN2-xTB is the calculation of anisotropic electrostatic interactions and exchange-correlation effects up to second order in the multipole expansion.³⁷ Additionally, dispersion interactions are included self-consistently, using the modified density-dependent D4 dispersion model.^{55,211} Due to inclusion of anisotropic electrostatics, no force-field (FF) type corrections for hydrogen or halogen bonds are required, unlike for the predecessor GFN1-xTB⁵⁷, and most other DFTB variants in literature.

In GFN2-xTB, the total energy is the sum over repulsion and dispersion interactions, the extended Hückel-type energy, (an)isotropic electrostatic (AES and IES) and exchange-correlation (AXC and IXC) effects, and an entropic contribution G .

$$E_{GFN2-xTB} = E_{rep} + E_{disp} + E_{EHT} + E_{IES+IXC} + E_{AES} + E_{AXC} + G_{Fermi}, \quad (2.33)$$

E_{rep} arises formally at zeroth order of eq. 2.32, and is calculated in a classical manner, using atom pairwise potentials.

$$E_{rep} = \sum_{AB} \frac{Y_A^{eff} Y_B^{eff}}{R_{AB}} e^{-(a_A a_B)^{1/2} (R_{AB})^{k_f}} \quad (2.34)$$

Here, Y_A^{eff} , Y_B^{eff} , a_A and a_B are element-specific parameter (Y_{eff} influence the repulsion interaction strength), and k_f is a global parameter set to $\frac{3}{2}$, except if both atoms are first row atoms (H or He), where it equals 1.

Dispersion effects are included formally up to second order in the Taylor series, *i.e.*, $E_{disp} = E_{disp}^{(0)} + E_{disp}^{(1)} + E_{disp}^{(2)}$. GFN2-xTB employs a modified, self-consistent D4 London dispersion

Table 2.1: GFN2-xTB Slater-Type atomic orbital basis functions employed for the different elements across the periodic table.³⁷

element	basis function
H	ns
He	ns, (n+1)p
Ne	nsp, (n+1)d
Group 1, Be-F, Zn, Cd, Hg-Po	nsp
Group 2, 13-18	nspd
TM, lanthanoides	nd, (n+1)sp

variant,²¹¹ which is shown in eq. 2.35.

$$\begin{aligned}
 E_{disp} = & - \sum_{n=6,8} s_n \sum_{A>B} \frac{C_n^{AB}(q_A, CN_{cov}^A, q_B, CN_{cov}^B)}{R_{AB}^n} f_n^{damp,BJ}(R_{AB}) \\
 & - s_9 \sum_{A>B>C} \frac{[3 \cdot \cos(\theta_{ABC})\cos(\theta_{BCA})\cos(\theta_{CAB}) + 1] C_9^{ABC}(CN_{cov}^A, CN_{cov}^B, CN_{cov}^C)}{(R_{AB}R_{AC}R_{BC})^3} \\
 & \times f_9^{damp,0}(R_{AB}, R_{AC}, R_{BC})
 \end{aligned} \tag{2.35}$$

R_{AB} is the distance between atoms A and B , θ_{ABC} represents the angle between atoms A , B , and C , f_{damp} are (Becke-Johnson- (BJ) and zero-(0)) damping functions (defined in references [212] and [54]), and C_n^{AB} and C_9^{ABC} are dispersion coefficients that depend on the atomic (partial) charges (C_6 and C_8) and coordination numbers (CNs). The first term in eq. 2.35 is a two body term, which is charge and CN-dependent via C_n^{AB} . In consequence, the two-body dispersion influences the electronic energy and is optimized self-consistently. Dominant many-body dispersion effects are included using the second three-body term, called Axilrod-Teller-Muto (ATM).^{219,220} It is charge independent and therefore does not influence the electronic energy.

The next term in eq. 2.33, E_{EHT} , is mainly responsible for covalent bonds and contains element-specific and global parameters. Within the LCAO ansatz, E_{EHT} reads:

$$E_{EHT} = \sum_i n_i \langle \Psi_i | \hat{H}_0 | \Psi_i \rangle \tag{2.36}$$

$$\equiv \sum_{\kappa} \sum_{\lambda} P_{\kappa\lambda} H_{\lambda\kappa} \tag{2.37}$$

$$\text{with } P_{\kappa\lambda} = P_{\kappa\lambda}^0 + \delta P_{\kappa\lambda} \tag{2.38}$$

$P_{\kappa\lambda}$ is a density matrix element and $H_{\lambda\kappa}$ resembles matrix elements that depends on the CN. Through parameters, $\hat{H}$ accounts to some extent for neglected on-site $\delta\rho^{(2)}$ effects. Additionally, short-range many-body effects are captured. Due to the CN-dependency, hypervalency is possible for large CNs through a lowering of d-level orbitals.

For charged or polar systems, the actual density differ from the reference density, *i.e.*, $\rho \neq \rho^{ref}$, and consequently net charges exist. In this approach, they are treated with shell-wise partitioned Mulliken charges, leading to equation 2.39.

$$E_{IES+IXC} = \frac{1}{2} \sum_{A,B} \sum_{l \in A} \sum_{l' \in B} q_{A,l} q_{B,l'} \gamma_{AB,ll'} + \frac{1}{3} \sum_A \sum_{l \in A} \Gamma_{A,l} q_{A,l}^3 \tag{2.39}$$

$q_{A,l}$ is an isotropic monopole charge of the l shell of atom A , $\gamma_{AB,ll'}$, called Mataga-Nishimoto-Ohno-Klopman (MNOK)^{221–223} formula, damps the Coulomb interaction $q \times q$. The MNOK term is distance dependent, and behaves differently for each element, as it contains element-specific parameters. It is included to enable long-range Coulomb interactions and decrease their magnitude in the short-range. $\Gamma_{A,l}$ is a shell-wise parameter, depending on element-specific and global parameters. It mainly repairs for large terms in the second XC term. All in all, the isotropic ES and XC contribution

mainly stabilize charged atomic states. Additionally, it helps to partially remedy shortcomings from the minimal valence basis set.

The AES term (eq. 2.40) describes electrostatic interactions between non-uniform charge distributions, *i.e.*, between monopoles q and dipoles μ , monopoles and quadrupoles Θ , and dipoles and dipoles. Consequently, it improves the description of the anisotropic density around an atom.

$$E_{AES} = E_{q\mu} + E_{q\Theta} + E_{\mu\mu} \quad (2.40)$$

μ and Θ are computed from electric dipole and quadrupole moment integrals between respective AOs. In short, the anisotropic electrostatic contributions are used for a better description of NCIs of less coordinated atoms (e.g. hydrogen or halogen atoms), and therefore, no further (FF-type) corrections are needed in GFN2-xTB.

An E_{AXC} is proposed for the first time in the DFTB context within the GFN2-xTB method. It captures changes due to polarization and mitigates to some extent shortcomings from the minimal valence basis set.

$$E_{AXC} = \sum_A (f_{XC}^{\mu_A} |\mu_A|^2 + f_{XC}^{\Theta_A} ||\Theta_A||^2) \quad (2.41)$$

μ and Θ corresponds to the cumulative atomic dipole moment and the traceless quadrupole moment, respectively, and f_{XC} are fitted, element-specific parameters to account for anisotropic density distributions.

The last term G_{Fermi} is introduced to handle e.g. static correlation or open-shell cases via a finite temperature ansatz (also referred to as Fermi-smearing)²²⁴:

$$G_{Fermi} = k_B T_{el} \sum_{\sigma} \sum_i^{\alpha, \beta, N_{MO}} \left[n_{i\sigma} \ln(n_{i\sigma}) + (1 - n_{i\sigma}) \ln(1 - n_{i\sigma}) \right]. \quad (2.42)$$

k_B refers to Boltzmann's constant, T_{el} is the electronic temperature (300 K by default), and $n_{i\sigma}$ describes the (fractional) occupation number of the molecular spin orbital.

GFN2-xTB is nowadays used in many quantum chemical analysis, for geometry optimization of large molecules, to generate and investigate CREs, for the detection of transition states, or to compute chemical properties.^{58–62,225} Furthermore, it is shown in Chapter 3 that GFN2-xTB is also able to quickly optimize (metallo-)proteins⁶⁵ up to 5000 atoms, yielding structures in the same error range as experimental uncertainties.

2.1.8 Periodic Boundary Conditions

Until now, all consideration belonged to molecular systems. For the calculation of metals or explicitly solvated systems, this approach reaches its limitations, as the system size increases drastically. Furthermore, the outer atoms or molecules tend to boil off into space, which is usually an undesired behavior. Therefore, periodic boundary conditions (PBC) are often introduced (see Fig. 2.1). Here, the system of interest is placed inside an appropriate box. The latter is then duplicated in all three dimensions such that the central box (and in principle all following boxes) is surrounded in every direction. In all boxes, the arrangements of atoms are mirrors of the central box, *i.e.*, if atom 1 in the central box moves to the left, an analogue movement in all other boxes with the same velocity and the same distance occurs. When one atom leaves the central box during a calculation, its mirror image will

enter the central box on the other side from the respective mirror box. Although PBCs are a simple tool to simulate periodic systems, attention must be paid to the size of the cell and to the truncation of long-range interactions. Electrostatic interactions, *e.g.*, are long-ranged, which means that an atom in the central box can in principle interact with its mirror image, which is physically unsound. Therefore, those interactions need to be truncated without losing physical sound interactions with other atoms. Consequently, the cell should be large enough to describe all necessary inter- (and intra-)molecular interactions properly.

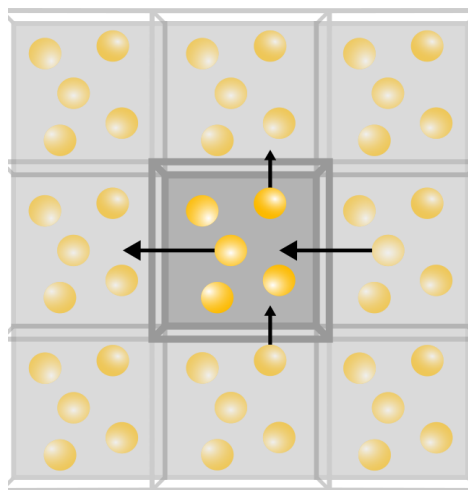


Figure 2.1: Sketch of multiplied cells used in calculations employing periodic boundary condition.

2.2 Excited State Calculation

Some properties of interest lay beyond the ground state picture and need to include a time-dependent external potential. This is for example the case, when electromagnetic radiation in a specific wavelength excites an electron from a bound to an unbound state. In the following, occupied molecular orbitals (MOs) are denoted as $i, j, k, l \dots$, unoccupied (virtual) ones as $a, b, c, d \dots$, MOs of any kind are written as $p, q, r, s \dots$, and atomic orbitals are given in Greek letters $\alpha, \beta, \gamma, \delta \dots$. Additionally, two-electron four-center integrals are written in the Mulliken notation:

$$(pq|rs) = \int dr_1 dr_2 \Psi_p(r_1) \Psi_q(r_1) \frac{1}{r_{12}} \Psi_r(r_2) \Psi_s(r_2). \quad (2.43)$$

The following Section gives a brief overview over time-dependent density functional theory and its simplified variant. Additionally, the computation of one- and two-photon absorption is highlighted.

2.2.1 Time-Dependent Density Functional Theory

The Hohenberg-Kohn theorems (*vide supra*) were proven for the time-independent ground state. In 1984, Runge and Gross²²⁶ similarly connect the electron density to an external potential for the time-dependent case,²²⁶ leading to the Runge-Gross theorems for time-dependent density functional

theory (TD-DFT). The non-Hermitian eigenvalue problem to be solved, called Casida's equations,²²⁷ is shown in eq. 2.44, assuming real orbitals.

$$\begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B} & \mathbf{A} \end{bmatrix} - \omega \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{pmatrix} \mathbf{X}_\zeta(\omega) \\ \mathbf{Y}_\zeta(\omega) \end{pmatrix} = \begin{pmatrix} \mu_\zeta \\ \mu_\zeta \end{pmatrix}, \quad (2.44)$$

$\mathbf{A}$ and $\mathbf{B}$ represent orbital rotation Hessian super-matrices, the unknown excitation and de-excitation amplitudes are depicted as $\mathbf{X}$ and $\mathbf{Y}$ eigenfunctions, ω is the eigenvalue vector with the dimension of the number of roots (or states) and describes the excitation energies, and μ_ζ is the perturbation.

For the calculation of an excited state, the perturbation is set to 0, and eq. 2.44 is transformed to a linear Hermitian problem, depicted in eq. 2.45

$$\begin{aligned} (\mathbf{A} - \mathbf{B})^{\frac{1}{2}} (\mathbf{A} + \mathbf{B}) (\mathbf{A} - \mathbf{B})^{\frac{1}{2}} \mathbf{Z} &= \omega^2 \mathbf{Z} \\ \text{with } \mathbf{Z} &= (\mathbf{A} - \mathbf{B})^{\frac{1}{2}} (\mathbf{X} + \mathbf{Y}). \end{aligned} \quad (2.45)$$

Assuming a spin-restricted global hybrid exchange correlation functional, the $\mathbf{A}$ and $\mathbf{B}$ super-matrices read:

$$A_{ia,jb} = \delta_{ij}\delta_{ab}(\epsilon_a - \epsilon_i) + 2(ia|jb) - a_x(ij|ab) + (1 - a_x)(ia|f_{XC}|jb) \quad (2.46)$$

$$B_{ia,jb} = 2(ia|bj) - a_x(ib|aj) + (1 - a_x)(ia|f_{XC}|bj) \quad (2.47)$$

Orbital energies of occupied and virtual orbitals are denoted as ϵ_i and ϵ_a (resulting from a DFT ground state calculation), a_x is the amount of non-local Fock-exchange of the underlying hybrid functional, and the two-electron four-center integrals are written in the Mulliken notation (see eq. 2.43). Eqs. 2.46 and 2.47 contain two different type of two-electron four-center integrals: the Coloumb-type integrals $(ij|ab)$, and the exchange-type integrals $(ia|jb)$, $(ia|bj)$, and $(ib|aj)$. In the adiabatic approximation,^{227,228} which work well for low-lying valence states, the derivative of the TD-XC functional, also known as TD-XC kernel f_{XC} , is derived from its time-independent counterpart, and is denoted as:

$$f_{XC}(r_1, r_2) = \frac{\delta^2 E_{XC}[\rho]}{\delta\rho(r_1)\delta\rho(r_2)}. \quad (2.48)$$

2.2.2 Simplified Time-Dependent Density Functional Theory

TD-DFT can nowadays be applied to 200-400 atoms,^{89,229,230} depending on the basis set and atoms of interest. Consequently, it is the workhorse for *ab-initio* excited state property calculations. However, realistic systems easily exceed 1000 atoms, jeopardizing the use of default TD-DFT for investigations of their excited state properties. To still enable such investigations, Grimme and co-workers⁹⁷ proposed a simplified variant coined sTD-DFT in 2014. It applies the same simplifications as the previously developed simplified Tamm-Dancoff approximation (sTDA)⁹⁶, introducing three approximations to TD-DFT: i) neglect of the XC kernel f_{XC} , ii) truncation of the singly-excited ($i \rightarrow a$) configuration space, and iii) approximation of the two-electron integrals by short-ranged damped Coulomb interactions of point charges (see eq. 2.49).

$$(pq|rs) \approx \sum_A^N \sum_B^N q_{pq}^A q_{rs}^B \Gamma_{AB} \quad (2.49)$$

q_{pq}^A are (transition) charge densities from a Löwdin population analysis²³¹ and Γ_{AB} is the Mataga-Nishimoto-Ohno-Klopman (MNOK)^{221–223} formula to damp short-range interactions of the point charges q . If the approximated integral is of Coulomb-type, Γ_{AB}^J reads:

$$\Gamma_{AB}^J = \left(\frac{1}{(R_{AB})^{y_J} + (a_x \eta)^{-y_J}} \right)^{\frac{1}{y_J}}$$

with $\eta = \frac{\eta(A) + \eta(B)}{2}$.

(2.50)

R_{AB} is the distance between atoms A and B , a_x is the amount of Fock-exchange, and η resembles the mean chemical hardness of two atoms A and B . For the latter, tabulated values are used, which were derived for all elements of the periodic table consistently.²³² y_J is a linear function and depend on a_x :

$$y_J = 0.20 + 1.83a_x$$
(2.51)

Exchange-type integrals decay differently than Coulomb-type integrals, and therefore are damped as:

$$\Gamma_{AB}^K = \left(\frac{1}{(R_{AB})^{y_K} + \eta^{-y_K}} \right)^{\frac{1}{y_K}},$$
(2.52)

with y_K :

$$y_K = 1.42 + 0.48a_x$$
(2.53)

Altogether, the sTD-DFT method contains 4 global parameters, which were fitted against reference excitation energies⁹⁶ to yield accurate UV/Vis and CD spectra.

Eq. 2.49 is based on the zero differential overlap (ZDO)²³³ approximation, assuming Löwdin-orthogonalized basis functions λ and Löwdin-orthogonalized LCAO coefficients C^{low} :

$$(ia|jb) \approx \sum_{\alpha\beta} C_{i\alpha}^{low*} C_{a\alpha}^{low} C_{j\beta}^{low*} C_{b\beta}^{low} (\lambda_\alpha \lambda_\alpha | \lambda_\beta \lambda_\beta).$$
(2.54)

Therefore, (transition) density charges q can be obtained for each atom A using the respective coefficients:

$$q_{ia}^A = \sum_{\alpha \in A}^N C_{ai}^{low*} C_{a\alpha}^{low}.$$
(2.55)

Considering the above approximations, the super-matrix elements $\mathbf{A}'$ and $\mathbf{B}'$ in the sTD-DFT formalism are denoted in eqs. 2.56 and 2.57.

$$\mathbf{A}'_{ia,jb} = \delta_{ij} \delta_{ab} (\epsilon_a - \epsilon_i) + \sum_{A,B}^N (2q_{ia}^A q_{jb}^B \Gamma_{AB}^K - q_{ij}^A q_{ab}^B \Gamma_{AB}^J)$$
(2.56)

$$\mathbf{B}'_{ia,jb} = \sum_{A,B}^N (2q_{ia}^A q_{bj}^B \Gamma_{AB}^K - a_x q_{ib}^A q_{aj}^B \Gamma_{AB}^K)$$
(2.57)

Originally, the singly-excited configuration space is truncated using only one threshold E_{thresh} , which will be called single threshold (st). Literature suggests^{234–237} that considering carefully chosen MOs instead of the full variant significantly decreases computational requirements, with only a minor loss of accuracy. In sTD(A)-DFT, the MO space (*i.e.*, adequate selection of occupied and unoccupied orbitals) is truncated evenly around the highest occupied and lowest unoccupied molecular orbitals (HOMO, and LUMO, respectively), considering the MO energy ϵ and the amount of Fock exchange a_x .

$$\begin{aligned}\epsilon &> \epsilon_{HOMO} - 2(1 + 0.8a_x)E_{thr} \\ \epsilon &< \epsilon_{LUMO} + 2(1 + 0.8a_x)E_{thr}\end{aligned}\quad (2.58)$$

Then, the primary configuration state functions (P-CSFs) are chosen using their diagonal Hamiltonian element A . If the excitation energy of occupied i to unoccupied a MO is smaller or equal to the given threshold, the singly-excited configuration is considered further.

$$A_{ia,ia} \leq E_{thr} \quad (2.59)$$

Secondary CSFs (S-CSFs), which have a non-negligible impact, are chosen for remaining configurations with energies higher than $10^{-4} E_h$, using a second-order perturbation contribution.

$$E_{jb}^{(2)} = \sum_{ia}^{P-CSF} \frac{|A_{ia,jb}|^2}{A_{ia,ia} - A_{jb,jb}} > 10^{-4} E_h \quad (2.60)$$

All in all, the sum of P-CSFs and S-CSFs is the total amount of CSFs taken into account for construction of the $\mathbf{A}'$ and $\mathbf{B}'$ super-matrices. Löffelsender *et al.*⁶³ showed for FMN (51 atoms) that this procedure decreases the CSF space from 12096 to about 350-1400, depending on the used thresholds.

Cases where the excited state property mainly stems from a small part of the system, say a chromophore in a fluorescent protein (FP), lead to the development of the dual threshold (dt) method, using two energy thresholds E_{high} and E_{low} for the truncation of the CSF space. In consequence, the responsible part (chromophore) is treated in the high layer with a higher threshold (*i.e.*, higher accuracy) while the environment, although considered with a lower threshold, can still be described explicitly. After truncation of the MO space following eq. 2.58, the density ζ_i computed from the MO coefficients C is divided into the *high* and *low* part:

$$\begin{aligned}\zeta_i &= \sum_{\alpha} C_{\alpha i}^2, \quad \text{with } \alpha \in \text{high-layer} \\ \zeta_i &> 0.1 \rightarrow E_{high} \\ \zeta_i &\leq 0.1 \rightarrow E_{low}\end{aligned}$$

P-CSFs are computed as in the st-sTD-DFT approach (see eq. 2.59), applying the respective energy thresholds. The MO space for S-CSF of the high-layer was determined already in the first step, while S-CSFs for the low layer follow a slightly different approach:

$$E_{low} < A_{jb,jb} < 2(1 + 0.8a_x)E_{low} \quad (2.61)$$

Both S-CSFs are evaluated with the second-order perturbation shown before (eq. 2.60) to explore if they need further investigation. Finally, the total amount of CSFs is again calculated as the sum over all P- and S-CSFs, *i.e.*,

$$N_{CSFs} = N_{P-CSFs_{high}} + N_{P-CSFs_{low}} + N_{S-CSFs}. \quad (2.62)$$

2.2.3 One-Photon Absorption

In an 1PA process, a system is excited from the ground to an excited state by absorption of a photon with a frequency ω . Figure 2.2 left side illustrates an exemplary one-photon absorption and emission process. Note that the energy of the photon during emission can be lower than the energy during absorption, depicted with a dashed arrow in Figure 2.2. To be visible in the one-photon absorption (1PA) spectrum, an excited state ν needs to have a non-zero oscillator strength f . The latter depends on the electric dipole moment $\vec{\mu}$ from ground (0) to excited (ν) state as well as on the excitation energy ω .

$$f_{0\nu} = \frac{2}{3} \omega_{\nu 0} \vec{\mu}_{0\nu} \vec{\mu}_{\nu 0} \quad (2.63)$$

$\vec{\mu}$ uses the the eigenfunctions $\mathbf{X}$ and $\mathbf{Y}$ of eq. 2.44 solved without perturbation in the sTD-DFT context. In the length formalism L , and assuming a spin-restricted case, the transition dipole moment for exciting an electron from an occupied to an unoccupied orbital ψ_i to ψ_a is shown in the following equation:

$$\vec{\mu}_{0\nu}^L = \sqrt{2} \sum_{ia}^{N_{CSF}} \langle \psi_i | \vec{r} | \psi_a \rangle (X_{ia,\nu} + Y_{ia,\nu}). \quad (2.64)$$

The resulting oscillator strengths yield a stick spectrum, which is converted to a broadened spectrum (to compare to experiment spectra) using Gaussian broadening functions (see Section 2.2.4).

2.2.4 Two-Photon Absorption

Two-photon absorption (2PA, sometimes referred to as TPA) is a non-linear optical process of third order. A system absorbs two photons of the same or different wavelengths simultaneously over a virtual state to yield an excited state. Figure 2.2 shows an exemplary two-photon absorption and emission process. Calculation of 2PA spectra in the sTD-DFT formalism was first introduced in ref. [106] in 2022.

To obtain 2PA strengths, the (s)TD-DFT formula (eq. 2.44) needs to be solved two times, once with and once without a perturbation. This leads to virtual (η) and excited (n) states, respectively, needed for the computation of the two-photon transition dipole moments $M_{\zeta\eta}^{0 \rightarrow n}$. Although formally of third order, they can be extracted from the linear response vectors $\mathbf{X}_n$, $\mathbf{X}_\eta(-\omega_n/2)$, $\mathbf{Y}_n$ and $\mathbf{Y}_\eta(-\omega_n/2)$.

$$M_{\zeta\eta}^{0 \rightarrow n} = -A' + B' \quad (2.65)$$

A' and B' are denoted in the following:

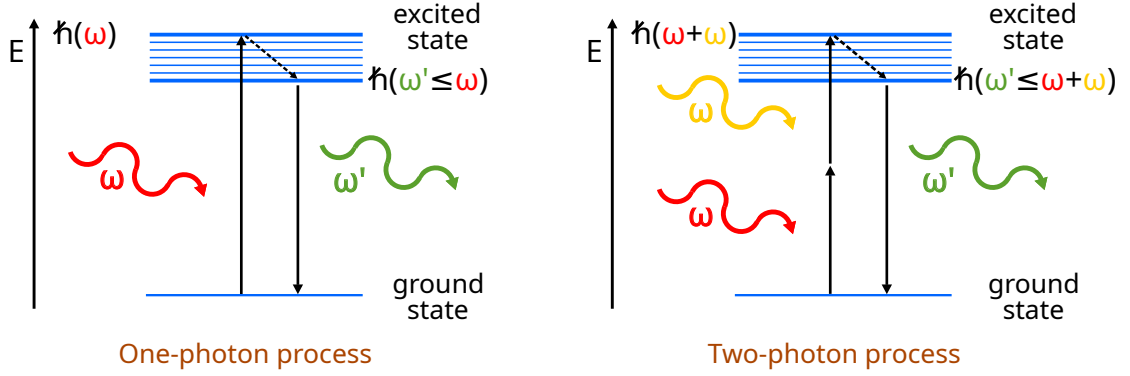


Figure 2.2: Sketch of an one (left) and two-photon (right) absorption and emission process. The dashed arrow illustrates a possible de-excitation before photon emission occurs.

$$A' = \sum_{perm.\xi,\zeta,\eta} \left\{ \sum_{aij} X_{n,ai} [\mu_{\zeta,ij}(1 - \delta_{\zeta\eta})] Y_{\eta,aj}(-\omega_n/2) \right\} \quad (2.66)$$

and

$$B' = \sum_{perm.\xi,\zeta,\eta} \left\{ \sum_{iab} X_{n,ai} [\mu_{\zeta,ab}(1 - \delta_{\zeta\eta})] Y_{\eta,bi}(-\omega_n/2) \right\} \quad (2.67)$$

μ is the perturbation, and calculated as $\mu_{\zeta,pq} = \langle \Phi_p | \hat{\mu}_{\zeta} | \Phi_q \rangle$, the Kronecker delta $\delta_{\zeta\eta}$ ensures the neglect of transition between two excited states, whilst $perm.\xi, \zeta, \eta$ denotes permutations between cartesian coordinates and their respective frequencies.

The transition strength matrix elements S are the product of 2P transition dipole moments:

$$S_{\zeta\eta,\xi\nu} = M_{\zeta\eta}^{0 \rightarrow n} M_{\xi\nu}^{0 \rightarrow n}. \quad (2.68)$$

Then, the rotationally averaged microscopic 2PA strengths $\langle \delta^{2PA} \rangle$ are computed via eq. 2.69.

$$\langle \delta^{2PA} \rangle = \frac{F}{30} \sum_{\zeta,\eta} S_{\zeta\zeta,\eta\eta} + \frac{G}{30} \sum_{\zeta,\eta} S_{\zeta\eta,\zeta\eta} + \frac{H}{30} \sum_{\zeta,\eta} S_{\zeta\eta,\eta\zeta} \quad (2.69)$$

The parameters F , G , and H depend on the experimental setup to compare with, and all equal 2 for parallel linearly polarized light. For a proper comparison to experimental data, the microscopic 2PA strengths are converted to macroscopic (measurable) 2PA cross-section σ^{2PA} (in Goepfert-Mayer (GM) units, $1 \text{ GM} = 10^{-50} \text{ cm}^4 \text{ s Photon}^{-1} \text{ molecule}^{-1}$) via:

$$\sigma^{2PA} = \frac{N\pi^3 \alpha a_0^5 (2\omega)^2}{c} \langle \delta^{2PA} \rangle S(2\omega, \omega_n, \Gamma) \quad (2.70)$$

Depending on the experimental setup, $N = 1$ or 2 for the single or double beam experiment, respectively, taking into account the photon dissipation rate of the laser beams.²³⁸ α denotes the fine structure constant, a_0 is Bohr's atomic radius, c resembles again the speed of light, and $S(2\omega, \omega_n, \Gamma)$ is a broadening function to convert the computed sticks to a line spectrum. For condensed phases, S is

normally a Gaussian line shaping function, depending on the damping factor Γ and the frequencies ω :

$$S(2\omega, \omega_n, \Gamma) = G(2\omega, \omega_n, \Gamma) = \frac{\sqrt{\ln 2}}{\sqrt{\pi} \Gamma} \exp \left[-\ln 2 \left(\frac{2\omega - \omega_n}{\Gamma} \right)^2 \right] \quad (2.71)$$

2.3 Gibbs Energy Calculations to obtain Melting and Boiling Points

An indirect attempt for the calculation of phase transition points, called TI-MD- λ DFT approach, is described in detail in refs. [165, 166]. It pins down melting and boiling points (MP and BP) of elements by calculating the intersection of the Gibbs energy curves (see Fig. 2.3) of the respective phases (solid and liquid for MP, liquid and gas for BP) at their respective equilibrium volumes and at normal pressure. Combining thermodynamic integration (TI), MD simulations and thermodynamic perturbation theory (TPT), all at a DFT level of theory, it overcomes DFA shortcomings using a scaling factor λ which is based on experimental cohesive energies E_{coh} . Altogether, this TI-MD- λ DFT yields outstanding results for BPs^{165,166,215} and pins down the MPs with deviations up to 10%.^{166,215}

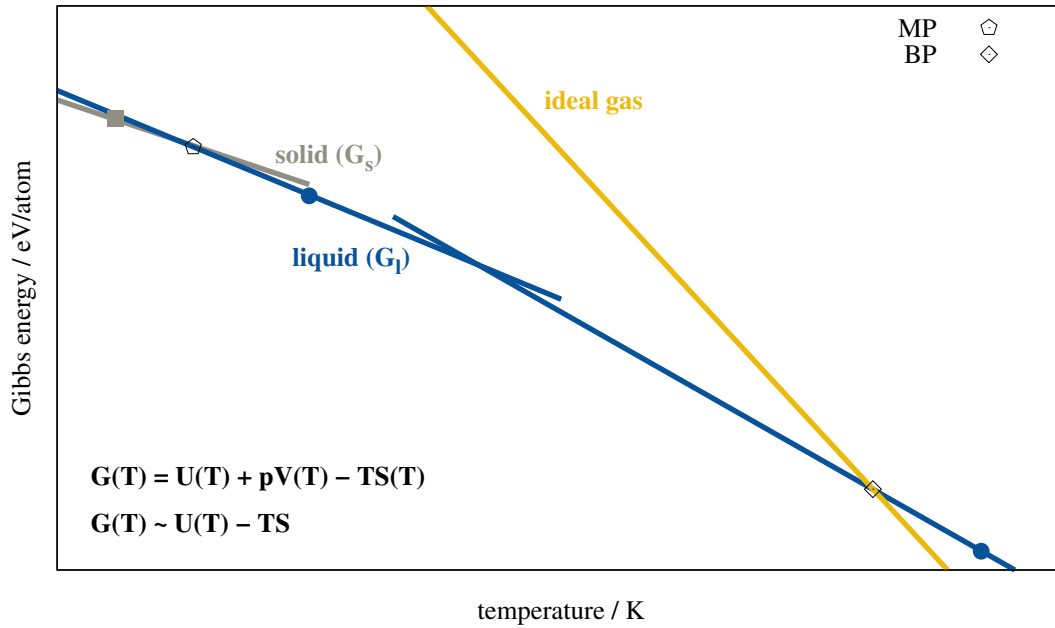


Figure 2.3: Sketch of the Gibbs energy intersections of solid, liquid, and gas phase. Solid, colored dots represent the calculated Gibbs energy at a specific temperature, lines are (linear) approximations to Gibbs energy curves.

A phase transition occurs, when the Gibbs energy curves of both phases intersect, *i.e.*, when they are equal ($G_{solid} = G_{liquid}$ for MP, and $G_{liquid} = G_{gas}$ for BP). In general, the Gibbs energy is composed of the internal energy U , the entropy S and the volume work pV , all dependent on the temperature T :

$$G(T) = U(T) + pV(T) - TS(T) = F(T) + pV(T). \quad (2.72)$$

The difference of U and TS is referred to as Helmholtz free energy F . In case all simulations (solid and liquid) are conducted at equilibrium volume, the volume work $pV \rightarrow 0$, and thus can be neglected. This simplifies eq. 2.72 to:

$$G(T) \approx F(T) = U(T) - TS(T). \quad (2.73)$$

While the internal energy U is accessible in QM calculations, computation of accurate entropies remains a challenge.²³⁹ Therefore, it is difficult to compute the Gibbs energy of a specific state, and this step is circumvented by integrating from a calculable reference to a fully-interacting DFT one, called thermodynamic integration. This step is different for solid and liquid phases, so in the following, each phase is discussed separately.

Solid The reference for the solid phase is the ideal (harmonic) crystal. The Gibbs energy of the solid consists of the total electronic energy $E_{\text{eq}, 0\text{K}}$, vibrational (phonon) contributions $G_{\text{vib}, 0\text{K}}$ at 0 K, anharmonic contributions G_{anharm} , and corrections calculated with TPT G_{TPT} .

$$G_{\text{sol}} = E_{\text{eq}, 0\text{K}} + G_{\text{vib}, 0\text{K}} + G_{\text{anharm}} + G_{\text{TPT}}. \quad (2.74)$$

$E_{\text{eq}, 0\text{K}}$ as well as $G_{\text{vib}, 0\text{K}}$ are computed at the predetermined equilibrium volume $V_{\text{eq}, \text{sol}}$. The final vibrational contributions are calculated in the harmonic approximation, for example using the program Phonopy.²⁴⁰

Afterwards, TI is used to investigate the anharmonic contributions, by integrating from an harmonic (ideal) crystal to the fully interacting DFT crystal.

$$G_1 - G_0 = \int_0^1 d\lambda \langle U_1(\mathbf{R}) - U_0(\mathbf{R}) \rangle_\lambda \quad (2.75)$$

The Gauss-Legendre quadrature with three quadrature points is used to approximate eq. 2.75, leading to three additional MD simulations with mixed harmonic and DFT forces.

$$\int_{-1}^1 f(x) dx \approx \sum_{i=1}^n w_i f(x_i) \quad (2.76)$$

Until now, the settings for the MD simulations can be selected according to computational feasibility, *e.g.*, SOR effects etc. can be neglected. The incorporation of refined computational settings, as well as other (higher-rung) DFT functionals or phenomena, such as SO coupling, is facilitated by TPT. In short, TPT compares a "lower" level of theory (1) to a "higher" level (2), thereby revealing the influence of the investigated topic. Thus, it is possible to achieve more sophisticated numerical settings without increasing the computational demands of the MD simulation. In the case of the TI-MD- λ DFT approach, TPT is applied to several independent configurations of the MD simulation, and the resulting differences are averaged for a more comprehensive estimate. Consequently, the influence on the final Gibbs energy of refined numerical settings, a computational demanding DFT functional, and SOR effects is investigated thoroughly.

$$G_2 - G_1 = -\frac{1}{\beta} \ln \langle e^{-\beta[U_2(R) - U_1(R)]} \rangle_1 \quad (2.77)$$

Instead of solving the full equation, we use the second-order cumulant expansion. Here, already the

second term gives contributions well below 1 meV/atom, which is sufficiently accurate.

$$G_2 - G_1 \approx \langle \Delta U \rangle_1 - \frac{\beta}{2} \langle (\Delta U - \langle \Delta U \rangle)^2 \rangle_1 \quad (2.78)$$

Liquid The reference for the liquid is an ideal gas-like non-interacting system at the predetermined equilibrium volume $V_{\text{eq, liq}}$ of the liquid supercell. Its Gibbs energy is simple to compute:

$$G^{id} = F^{id} = -k_B T \ln(Z(\Theta, T, V, N)) + pV, \text{ with} \quad (2.79)$$

$$Z(\Theta, T, V, N) = \frac{(\Theta V)^N}{\Lambda^{3N} N!} \text{ and } \Lambda = h \sqrt{\frac{\beta}{2\pi m}}$$

k_B denotes Boltzmann constant, Θ resembles the electronic degeneracy, N is the number of atoms in the investigated supercell, Λ is the de Brogli wavelength, and m the mass of the atom. For a given simulation temperature T and a predetermined equilibrium volume V , this term is constant.

The Gibbs energy of the interacting DFT liquid can be extracted from TI, applying the interpolation of the internal energy of a non-interacting liquid to a fully-interacting DFT liquid. In contrast to the solid, where the ideal crystal is easy to simulate, the computation of the non-interacting liquid is unfeasible. Briefly, the non-interacting atoms do not experience any forces, no attractive but also not repulsive one, so they can in principle fuse during a MD simulation. Therefore, λ in eq. 2.75 is substitute by a function $\lambda(x)$, which was developed in 2018 by Dorner *et al.*²⁴¹

$$\lambda(x) = \left(\frac{x+1}{2} \right)^{\frac{1}{1-\kappa}} \quad (2.80)$$

Consequently, eq. 2.75 now reads:

$$G_1 - G_0 = \frac{1}{2(1-\kappa)} \int_{-1}^1 dx \langle U_1^{\text{pot}}(\mathbf{R}) \lambda(x)^\kappa \rangle_x \quad (2.81)$$

In contrast to Dorner *et al.*²⁴¹, the TI-MD- λ DFT approach conducts simulations at a lower level of theory, and uses TPT as before to numerically converge the final liquid Gibbs energy. Additionally, it enables the consideration of SOR effects. Eq. 2.81 is approximated with the Gauss-Lobatto quadrature, for which the values $n = 7$ and $\kappa = 0.7$ have been chosen in the present work (cf. Chapter 6).

Gas The Gibbs energy of the gas phase is approximated as an ideal gas. It is calculated as shown before (eq. 2.79) for a temperature T near the BP, at the equilibrium volume and at ambient pressure. The natural logarithm ($\ln(Z)$) is solved by the Stirling approximation, considering an arbitrary number of particles. Note that the volume work is not negligible for the gas phase and is considered explicitly.

2.4 Remarks on Geometry Preparation of Proteins

Preparing the geometry of medium-sized (<500 atoms) molecules is non-trivial,²⁴² and generation of large molecules (>500 atoms) is even more demanding. Most of the time, quantum chemical calculations start with the energetically lowest-lying conformer, but pinning down its three dimensional

structure is a challenge. For proteins, the three dimensional structure can be extracted from NMR or X-Ray experiments, as stated in Chapter 1. Protein NMR-structures are commonly measured in solution, leading to an ensemble of different structures, and it remains elusive, which structure to take for further QM calculations. X-Ray structures provide quite detailed positions of heavy atoms, although different side chain conformations may appear, and must be considered manually. However, information about the spatial arrangement of the hydrogen atoms are not resolved. It is trivial to protonate amino-acids using their experimentally known pK_a -values, measured in water for the single amino-acid of interest. However, there are (at least) five amino-acids with varying protonation states (arginine, lysine, aspartic acid, glutamic acid, and the very special case of histidine). It was shown experimentally² that the pK_a -values may differ due to the environment, leading to the inherently difficult task to determine the correct protonation state of a protein. Although this might sound negligible for a 5000 atom protein, H-bonds or electrostatic interactions due to a charged side chain can lead to wrong binding situation.²⁴³ Therefore, yielding a correct 3D-structure of a protein, including H-atoms, remains in many cases a matter of chance, but may decide over success or failure of subsequent QM calculations. Due to the unknown protonation state, the total charge (and the spin, if metal atoms are part of the structure) of a protein remains elusive, too. Last but not least, crystallization agents as well as solvent molecules need to be accounted for, as their presence might impact the protein structure. Collectively, these unknown variables aggravates the already difficult protein calculations further, and the resulting data should be interpreted with appropriate caution.

Part II

Studies on Semi-Empirical Quantum Mechanical Methods

Quantum Chemical Calculation of Molecular and Periodic Peptide and Protein Structures

Sarah Schmitz¹, Jakob Seibert¹, Katja Ostermeir¹, Andreas Hansen¹, Andreas H. Göller², and Stefan Grimme¹

Received: January 20, 2020

First published: April 10, 2020

Reprinted with permission[‡] from:

S. Schmitz, J. Seibert, K. Ostermeir, A. Hansen, A. H. Göller, and S. Grimme, *Quantum Chemical Calculation of Molecular and Periodic Peptide and Protein Structures*, The Journal of Physical Chemistry B, **124(18)** (2020) 3636-3646, DOI: 10.1021/acs.jpcb.0c00549

Copyright © 2020 American Chemical Society

Own manuscript contributions:

- Data curating in parts
- Performing most calculations
- Evaluation and interpretation of all results
- Visualization of all results
- Writing and revising most parts of the manuscript

¹ Mulliken Center for Theoretical Chemistry, Rheinische Friedrich-Wilhelms-Universität Bonn, Berlingstr. 4, 53115 Bonn, Germany

² Computational Molecular Design, Pharmaceuticals, R&D, Bayer AG, 42096 Wuppertal, Germany

[‡] Permission requests to reuse material from this chapter should be directed to American Chemical Society

Understanding structure-property relationships for proteins is a challenging task until today. Computational chemistry can aid, but necessitates an accurate molecular structure. For proteins, this implies error-prone preparation of an experimental structure, as illustrated in Figure 1.1, and further computational refining. As outlined above, pure DFT or WFT methods are computationally too demanding for AQM calculations on realistic proteins. Various approximations exist,^{8–23,244–248} but have their own limitations. Reducing computational requirements leads to semi-empirical QM (SQM)^{40,41,47–51} or classical approaches,^{26–32} which can routinely be used for geometry optimizations of systems with up to ten thousand atoms, *i.e.*, biological relevant proteins. In this Chapter, the SQM methods GFN n -xTB ($n = 1, 2$) are examined thoroughly for (metallo-)protein structures up to 5000 atoms, applying an ASQM approach. Extending the study of Martinez *et al.*,²⁴⁹ two comprehensive benchmark sets were compiled for the evaluation of geometry optimizations of GFN n -xTB ($n = 1, 2$) considering protein structures, and compared to experimental crystal X-ray structures. If possible, comparison to competitor methods are also conducted, namely the two special-purpose force fields AMBER*^{26,27} and OPLS2005²⁸, the composite method HF-3c¹⁹, and the SQM methods PM6-D3H4X^{39,250} and PM7²⁵¹. The first test set consists of 70 peptides and proteins, and contains only organic elements (H, C, N, O) and sulfur, ranging from small peptides and fibrils to large proteins (70 – 1657 atoms, 897 atoms on average), with total charges ranging from –2 to +2. From the first set, a small subset of 10 oligopeptides (90 – 109 atoms) is created to evaluate packing effects, *i.e.*, the difference between molecular and periodic treatment. The second test set concentrates on proteins containing various metals (Cu, Fe, Ca, Zn, Hg, Co, Ni, Cd), and metal-containing prosthetic groups, such as heme. The chosen 20 metalloproteins are larger in size (1034 – 4804 atoms) and have total charges from –15 to +7. Note that GFN2-xTB is focused in this study, and therefore, results for the precursor GFN1-xTB are not discussed in detail.

For the evaluation of packing effects, the neutralized peptides of the subset were optimized using HF-3c with implicit solvation as well as with periodic boundary conditions (see Figure 2.1) to mimic crystal behavior. HF-3c is known to overbind molecular crystals,²⁵² therefore a scaled variant (sHF-3c²⁵²) is applied, too. Note that at the time of the study, PBC calculations with GFN n -xTB were not implemented. The C_α root mean square deviation (C_α -RMSD) is used to assess both the difference between theory and experiment, and the packing effect. Molecular equilibrium sHF-3c and HF-3c structures yield similar deviations with respect to experimental results (C_α -RMSD = 0.56 and 0.60 Å, respectively), while sHF-3c structures optimized under PBC perform slightly better (0.43 Å). Note that all workflows show deviations close to the apparent experimental uncertainty of 0.5 Å.^{253,254} The packing effect, evaluated by comparing sHF-3c calculations with and without PBC amounts on average to 0.70 Å, again similar in size than the apparent experimental uncertainty. Therefore, comparing geometrically optimized molecular structures to experimental crystal structures is valid, although the packing effect should be further investigated using larger systems and preferably crystal protein structures.

Next, structures of the first test set are evaluated using GFN n -xTB ($n = 1, 2$), AMBER*, OPLS2005 and HF-3c. Assessing the computational timings, it becomes visible that HF-3c is barely applicable to the larger test systems, as it needs on average 1.5 h per optimization step. The longest optimization consequently needed 4 months to complete, which makes this approach not feasible for screening purposes. GFN2-xTB took on average 17 s/step, and the force fields are fastest (much less than 1 s/step), due to their much simpler theory.

Visual inspection of all structures reveals no unfolding during optimization and furthermore shows preservation of the secondary structure features for all tested methods. Only regions with random coil

deviate more from experimental structures. To assess the quality of all methods, the backbone angles Φ , Ψ , and ω , the first side chain angle χ_1 and two RMSDs, namely C_α -RMSD and ha-RMSD (RMSD between all non-hydrogen atoms) are calculated with respect to the corresponding experimental structures. Analysis reveals that GFN2-xTB is of similar quality as the specialized OPLS2005 force field, and superior to the other tested models. Afterwards, bond length distributions, which are of high importance for spectroscopic investigations, are evaluated for CC, CO, CN, and CS bonds. The force fields, as well as HF-3c yield narrow bond length distributions that deviate significantly from experimental measurements. GFN2-xTB yields much broader distributions, even for the difficult CS bond, and therefore performs best for this descriptor. Note that both tested PM methodologies show technical problems for 50 systems of the first test set. As GFN2-xTB generally leads to more accurate results while being computationally less demanding, PM6-D3H4X and PM7 are not further discussed here for the remaining proteins. Putting it all together, GFN2-xTB is a promising candidate to yield fast and accurate molecular equilibrium structures of proteins without prosthetic groups at a quantum mechanical level of theory.

In the second test set, metalloproteins are investigated, each containing either single metal atoms or prosthetic groups with metals like heme-groups. As common force fields are not able to treat any other than common prosthetic groups, such as heme, and HF-3c is known to yield poor results for metals, only GFN n -xTB is assessed here. Considering the above introduced geometrical descriptors, GFN2-xTB yields similar results for the second as for the first test set. The secondary structure is well preserved, and random coil regions are stabilized if near a metal center. Additionally, the first coordination shell around metal atoms is evaluated empirically by means of the Wiberg/Mayer bond order.^{255,256} GFN2-xTB mostly preserves experimental coordination spheres, and only in few cases, a ligand change is observed, or the coordination number is increased. Therefore, we conclude that GFN2-xTB shows generally good performance also for challenging structures such as metalloproteins.

In conclusion, the comparison of molecular structures against periodic crystal structures was validated, by structure optimizations of a small subset with an implicit solvation model and with periodic boundary conditions. For the first test set, GFN2-xTB yields similar geometrical descriptors than the special purpose force field OPLS2005, and better results than all other tested methods. Regarding the bond length distributions, it yields best estimates compared to experimental structures. Furthermore, results for the second test set show that GFN2-xTB describes proteins containing common and uncommon metal groups with a similar quality as without those groups. Coordination spheres are mostly preserved during the optimization process. All in all, GFN2-xTB provides reasonable structures for realistic (metallo-)proteins, rendering it a promising candidate for further (spectroscopic) investigations.

All-Atom Quantum Mechanical (AQM) Methodologies for One- and Two-Photon Absorption of Realistic Systems

Sarah Löffelsender¹, Marilù G. Maraldi², and Marc de Wergifosse²

Received: December 16, 2024

First published: April 08, 2025

Reprinted with permission from[‡]:

S. Löffelsender, M. G. Maraldi, and M. de Wergifosse, ChemPhysChem, **26.11** (2025) e202401121,

DOI: 10.1002/cphc.202401121

Copyright © 2025 Wiley-VCH GmbH

Own manuscript contributions:

- Data curating
- Preparation of calculations
- Performing all calculations
- Evaluation and interpretation of the results
- Visualization of most results
- Writing and revising the manuscript

¹ Mulliken Center for Theoretical Chemistry, Clausius Institute for Physical and Theoretical Chemistry, Rheinische Friedrich-Wilhelms-Universität Bonn, Berlingstr. 4, 53115 Bonn, Germany

² Institute of Condensed Matter and Nanosciences, Université Catholique de Louvain, B-1348 Louvain-la-Neuve, Belgium

[‡] Permission requests to reuse material from this chapter should be directed to Wiley-VCH GmbH

The direct environment of a molecule influences its geometry, energy, dynamics and reactivity, and thus ground and excited state properties.¹ As shown in Chapter 3, the SQM method GFN2-xTB yields reliable structures for large and challenging proteins, and the encouraging results in bond length distributions render it a potential candidate for further spectroscopic investigations. Thus, this study investigated possible AQM workflows for the computation of 1- and 2PA spectra for realistic systems, considering the environment explicitly at a QM level of theory, and either excluding (all-atom single structure, ASQM) or including dynamic structural effects (all-atom dynamic structure, ADQM). The ASQM method was assessed by means of the challenging fluorescent proteins bacteriorhodopsin (bR, 3835 atoms) and iLOV (1849 atoms), optimized at the ONIOM ω B97X-D^{206,257}/6-31G*^{258–263}:GFN2-xTB³⁷/GBSA(water)^{264–266} level of theory.²⁴ The two ADQM workflows ADQM-B and ADQM-MD were tested for the flavin mononucleotide (FMN, chromophore of iLOV), using GFN2-xTB³⁷ for geometry generation. Additionally, implicit and explicit dynamic structural effects were evaluated. The ADQM-B approach with its indirect dynamic structural effects is based on a conformer-rotamer ensemble (CRE) of FMN in gas phase or implicit solvation (water), using the programs CREST^{242,267} and CENSO^{268,269} to obtain a final Gibbs energy ranked CRE (r^2 SCAN-3c^{55,200,207} electronic energy + G_{mRRHO} ^{270–272}(GFN2-xTB)// r^2 SCAN-3c for the ensemble in gas phase as well as r^2 SCAN-3c electronic energy + δG_{solv} (COSMO-RS[H₂O])^{273–275} + G_{mRRHO} (GFN2[GBSA]-bhess) // r^2 SCAN-3c[DCOSMO-RS]^{276,277} in implicit water). For all relevant structures, absorption spectra were computed and Boltzmann-weighted to yield a final averaged spectrum. For the ADQM-MD scheme, FMN was simulated in i) gas phase, ii) implicit water, and iii) explicit water. For the latter, water molecules were added to the lowest-energy conformer of FMN using spheres of increasingly large radii of 7 Å, 9 Å, 10 Å, 11 Å, and 12 Å. Then, MD simulations were conducted at the GFN2-xTB level of theory. Uncorrelated snapshots from each MD simulation laid the foundation for further absorption spectra calculation, and were equally weighted to yield averaged spectra. 1- and 2PA spectra as well as natural transition orbitals (NTOs) were computed with the simplified (dt-)sTD-DFT-xTB^{24,95–97,99,106} method, and compared to respective experimental 1- and 2PA spectra.^{278–286} To complement theoretical investigations, 1- and 2PA spectra of the lowest-energy conformers of the chromophores with respect to r^2 SCAN-3c energies were explored in vacuum and in implicit solvation (FMN only) using the high level approach RI-CC2²⁸⁷/aug-cc-pVDZ²⁸⁸.

Based on the ASQM scheme, comparison of experimental and theoretical 1- and 2PA spectra for bR show reasonable agreement, as the main features are captured by sTD-DFT-xTB. Additionally, the recently introduced²⁴ dt-sTD-DFT-xTB methodology performs well while reducing the computational costs by decreasing the explicitly considered configurational space. Inspection of natural transition orbitals reveals that all atoms need to be included in the excited state calculation, as tryptophan residues have non-negligible impact on the high energy part of the spectra, often absorbing photons without chromophore contribution. This was already found for the photoactive yellow protein PYP by de Wergifosse *et al.*²⁸⁹ For bR, NTO analysis furthermore reveals a known^{290–292} charge transfer (CT) from a nearby tyrosine to the chromophore retinal, enhancing and red-shifting σ^{2PA} of this excitation compared to the chromophore in gas phase, while the oscillator strength decreases significantly. The same analysis based on iLOV confirms the ASQM scheme as well, as many excitations involve surrounding residues or are residue-only. Experimental 1- and 2PA spectra are well reproduced, although the second peak in both spectra is blue-shifted. The literature-known sensitivity of photon absorption processes to the planarity of the isoalloxazine ring^{84,85,87} may impact the calculated spectra, as the optimized ring is slightly bent compared to experimental structure. For both chromophores (retinal and FMN), including the surroundings in the calculations has a far more significant impact

than just increasing the level of theory from sTD-DFT-xTB to RI-CC2. Nevertheless, side features as broadening or shoulders in the spectra are not resolved correctly, which is most likely due to the underlying single structure approach.

Therefore, ADQM schemes are probed using flavin mononucleotide (FMN), including dynamic structural effects either implicitly (ADQM-B) or explicitly (ADQM-MD). Resulting 1- and 2PA spectra were compared to experimental ones, revealing that the ADQM-B scheme hardly improves over single structure spectra. When using a large solvation shell of at least 11 Å, however, ADQM-MD spectra provide striking agreement with experimental 1- and 2PA spectra, catching the difficult inhomogeneous broadening of the second absorption band. NTO analysis reveals that the inhomogeneous broadening appears due to a charge transfer including parts of the chromophore (especially the interaction of the moving HPO_4^- with the isoalloxazine ring) and surrounding water molecules, which can only be captured when accounting explicitly for the surroundings together with explicit dynamics. The same inhomogeneous broadening of the second peak is also visible in the 1PA spectrum of iLOV. As the ASQM scheme failed in correctly describing this feature, the ADQM-MD scheme could aid and should be investigated in a future study. All in all, the results clearly emphasize the importance of explicitly including the environment in the calculation of 1- and 2PA spectra by applying explicit molecular dynamics.

Molecular Folding Governs Switchable Singlet Oxygen Photoproduction in Porphyrin-Decorated Bistable Rotaxanes

Jan Riebe¹, Benedikt Bädorf², Sarah Löffelsender², Matias Ezequiel Gutierrez Suburu³, María Belén Rivas Aiello³, Cristian Alejandro Strassert³, Stefan Grimme², Jochen Niemeyer¹

Received: March 15, 2024

First published: August 07, 2024

Reprinted with permission from:

J. Riebe, B. Bädorf, S. Löffelsender, M. E. Gutierrez Suburu, M. B. Rivas Aiello, C. A. Strassert, S. Grimme, and J. Niemeyer, *Communications Chemistry* **7** (2024), DOI: 10.1038/s42004-024-01247-7
Copyright © 2024 The Authors. Published under the license CC BY 4.0

Own manuscript contributions:

- Geometry preparations
- Preparation and performing parts of calculations
- Supervision for all calculations
- Interpretation of results
- Writing and revising parts of the manuscript

¹ Faculty of Chemistry (Organic Chemistry) and Center for Nanointegration Duisburg-Essen (CENIDE), University of Duisburg-Essen, Universitätsstrasse 7, 45141 Essen, Germany

² Mulliken Center for Theoretical Chemistry, Clausius Institute for Physical and Theoretical Chemistry, Rheinische Friedrich-Wilhelms-Universität Bonn, Beringstr. 4, 53115 Bonn, Germany

³ Institut für Anorganische und Analytische Chemie, CeNTech, CiMIC, SoN, Universität Münster, Heisenbergstr. 11, 48149 Münster, Germany

As shown in Chapter 4, the ADQM-B workflow shows little improvement over single structure methods regarding 1- and 2PA spectra. Nevertheless, it provides valuable insights into possible structures of the potential energy surface (PES) of highly flexible systems. The experimental part of this work (performed by the collaborators) first focuses on the design of two bistable rotaxanes with a photosensitizer (Zn(II) tetraphenylporphyrin, ZnTPP), a black-hole-quencher (BHQ-2), and two linkers of different lengths (C_7 or C_1) between both binding stations (compounds **1a/b**, see Appendix C, Fig. 1d). ZnTPP has a strong absorption band at around 420 nm and a high singlet oxygen production ($\Phi_{\Delta} = 0.72 \pm 0.12$), while BHQ-2, featuring a dialkylaniline, a central dimethoxybenzene, and a terminal nitrobenzene linked by two diazo groups, has a broad absorption range of 350 – 650 nm and should thus be able to suppress 1O_2 production when in close proximity to ZnTPP. To suppress the singlet oxygen production completely, the design envisages that the quencher is positioned within close proximity to the photosensitizer in the *off*-state (**1a/b**⁺), while it is at a long distance in the *on*-state (**H-1a/b**²⁺). Subsequent NMR-investigations confirm the pH-induced shift of the macrocycle between the amine/ammonium-station and the triazolium-station.

Measurements of the 1O_2 production demonstrate that the rotaxane **1a**, featuring the long C_7 linker, indeed exhibits two reversibly switchable states, with the *on*-state yielding about 500 % increased singlet oxygen production compared to the *off*-state ($\Phi_{\Delta} = 0.03 \pm 0.003$ for **H-1a**²⁺ vs. $\Phi_{\Delta} = 0.15 \pm 0.03$ for **1a**⁺). However, the direction of switching is reversed regarding the original design, and the production is significantly reduced compared to ZnTPP alone. In contrast, rotaxane **1b**, featuring the short C_1 linker, shows no significant 1O_2 quantum yield in both protonation states ($\Phi_{\Delta} < 0.01 \pm 0.001$ for **H-1b**²⁺, $\Phi_{\Delta} = 0.01 \pm 0.001$ for **1b**⁺). Control rotaxanes without the quencher show similar results as the ZnTPP alone ($\Phi_{\Delta} = 0.48 - 0.60$), demonstrating that i) the quencher is indeed affecting the singlet oxygen production, ii) the protonation state is not influencing the 1O_2 quantum yield, and iii) the conformational change in the structures of **1a** must play a crucial role.

Conducted fluorescence lifetime measurements (by other collaborators) could not explain the unique behavior of **1a**⁺ in the **1a/b** series. Therefore, theoretical investigations on the conformational space of all four rotaxanes are conducted, focusing on possible intramolecular interactions of the ZnTPP unit. To this end, the potential energy surface (PES) is scanned for all **1a/b** variants, using the program CREST^{242,267} in combination with the GFNn-xTB ($n = 1, 2$) methods^{37,56,57}, and the force field variant GFN-FF,³⁰ employing the ALPB(THF) implicit solvation model.²⁹³ To additionally increase the level of theory, representative conformations are evaluated at the PBEh-3c level of theory²⁰ to yield multilevel Gibbs energies (PBEh-3c electronic energy + $\delta G_{solv}(\text{COSMO-RS}[\text{THF}])^{273-275}$ + $G_{mRRHO}(\text{GFN2}[\text{ALPB}]\text{-hess})^{270-272}$ // GFN2-xTB [ALPB:THF]. The investigated conformations include energetically low-lying structures as well as higher-lying ones, focusing on a variety of Zn atom interactions. Based on the Zn atom binding motif in its axial (octahedral) position and the folding of the whole system, all conformers are organized in three categories: *closed*, *half-open*, and *open*.

Evaluation shows distinct differences in the folding behavior of **1a**⁺ and the other variants. The protonated **H-1a**²⁺ form is predominantly characterized by a strongly folded binding motif (*closed* category), exhibiting the lowest Gibbs energy among all categories, despite unfavorable thermal and solvation contributions. The long C_7 linker enables a triple π stack, probably stabilizing the high charge (+2) of this folded conformation via strong electrostatic interactions. Additionally, the strong folding can explain unexpected shifts observed in the NMR spectra. Both **1b** variants show limited flexibility, resulting in mostly *closed* conformations with similar binding motifs, which are thermodynamically favored. Therefore, **H-1a**²⁺ and both **1b** protonation states show similar singlet oxygen quantum yields. In contrast, **1a**⁺ proofed to be more flexible with many different Zn atom

binding motifs. Several *half-open* conformers were found, showing similar Gibbs energies as *closed* conformations. This unique behavior allows for better diffusion of $^3\text{O}_2$ to the photosensitizer.

Putting it all together, the BHQ-moiety is in closer proximity to the photosensitizer in **H-1a**²⁺ and both **1b** variants. On the one hand, this facilitates energy transfer from the ZnTPP to the quencher, suppressing the singlet oxygen production. On the other hand, diffusion of $^3\text{O}_2$ is inhibited as the binding capabilities of the Zn atom is limited through the close distance between ZnTPP and BHQ. The more flexible **1a**⁺ shows multiple *close* and *half-open* conformations, with the BHQ-moiety closer or in further distance from the Zn atom. The *closed* conformers for the **1a**⁺ variant explain the lower singlet oxygen production compared to ZnTPP alone ($\Phi_{\Delta} = 0.15 \pm 0.03$ vs. $\Phi_{\Delta} = 0.72 \pm 0.12$), as quenching is active and diffusion of $^3\text{O}_2$ is hindered. However, in *half-open* conformers, BHQ and ZnTPP are in further distance, deteriorating the quenching, and improving the diffusion of $^3\text{O}_2$ to the Zn atom. This explains the higher, but not equivalent, $^1\text{O}_2$ quantum yield compared to the other variants and ZnTPP alone, and moreover, the reversed switching behavior of the **1a** variants. All in all, this investigation reveals the importance of considering molecular folding for the design of flexible systems through QM calculations.

Part III

Quantum Mechanical Studies on Relativistic Effects

It's Complicated: On Relativistic Effects and Periodic Trends in the Melting and Boiling Points of the Group 11 Coinage Metals

Sarah Löffelsender¹, Peter Schwerdtfeger², Stefan Grimme¹, and Jan-Michael Mewes¹

Received: October 14, 2021

First published: December 29, 2021

Reprinted with permission from[‡]:

S. Löffelsender, P. Schwerdtfeger, S. Grimme, J.-M. Mewes, *Journal of the American Chemical Society*, **144**(1) (2022), 485-494,

DOI: 10.1021/jacs.1c10881

Copyright © 2021 American Chemical Society

Own manuscript contributions:

- Preparation and performance of all calculations excluding non-relativistic potential generation
- Evaluation and interpretation of the results
- Visualization of all results
- Writing the initial draft of the manuscript
- Revising of the manuscript

¹ Mulliken Center for Theoretical Chemistry, Rheinische Friedrich-Wilhelms-Universität Bonn, Beringstr. 4, 53115 Bonn, Germany

² Centre for Theoretical Chemistry and Physics, The New Zealand Institute for Advanced Study, Massey University, Auckland Campus, 0632 Auckland, New Zealand

[‡] Permission requests to reuse material from this chapter should be directed to American Chemical Society

Although even the smallest children are already aware of phase transition processes (*e.g.*, melting of ice), their microscopic mechanics are still under debate. Directly simulating phase transitions is challenging, as technical details such as superheating or supercooling affects the melting and boiling points (MP, and BP) considerably. Therefore, Mewes *et al.*^{165–167} developed an indirect approach to compute phase transition points relying on intersection of Gibbs energy curvatures. Internal energies are determined using MD simulations at a DFT level of theory, and subsequent thermodynamic integrations (TI) lead to Gibbs energies at a specific temperature. Increased numerical accuracy as well as inclusion of spin-orbit relativistic (SOR) effects are received using thermodynamic perturbation theory (TPT). Consequently, Gibbs energies are extrapolated linearly or quadratically, yielding the final phase transition points. For a meaningful discussion, DFT-related systematic errors for each element, *i.e.*, over- or underbinding, need to be considered. Therefore, the phase transition points are scaled, using an empirical scaling factor λ based on the ratio of experimental and theoretical cohesive energies E_{coh} . The latter is additionally employed to determine the density functional approximation (DFA) with the best cost-accuracy ratio for further MD simulations. As shown for closed-shell Group 12 elements Zn, Cd, and Hg, this so-called TI-MD- λ DFT approach yields accurate MPs and BPs.¹⁶⁶ In this work, the approach is challenged by investigations on relativistic effects on phase transition points of the open-shell coinage metals Cu, Ag, and Au.

First, E_{coh} and equilibrium volumes at 0 K (V_{eq}^0) for all Group 11 elements are assessed and compared to experimental results. The GGAs PBE^{193,194}, PBE-D3(BJ)^{55,212} and PBEsol^{197,198}, and the meta-GGA SCAN²⁹⁴ are tested. SCAN and PBEsol reproduce E_{coh} and V_{eq}^0 most consistently for all coinage metals. Considering computational efforts, the more approximate PBEsol functional is chosen for further calculations. For the SOR case, the lattice for all three metals is known experimentally (*fcc*), but for the non-relativistic (NR) limit, it needs to be determined first. Therefore, E_{coh} for the face-centered cubic (*fcc*), body-centered cubic (*bcc*), and hexagonal close-packed (*hcp*) lattices at 0 K are calculated with the PBEsol functional. For SOR and NR, the *fcc* lattice is yielded as most stable packing, in line with (SOR) experimental results. Note that for Au, *fcc*, and *hcp* are energetically closer in the NR case compared to the SOR case.

After defining DFA and lattice, MPs and BPs for the SOR cases are computed. BPs agree excellently with experimental data, with deviations below 1 %. This is in line with previous investigations.^{165,166} The gas phase Gibbs energy is computed as ideal gas (compare eq. 2.79), *i.e.*, no simulation is needed. Therefore, simulations of the liquid phases apparently yield accurate Gibbs energies, and are furthermore improved using the λ -scaling.

On the contrary, MPs deviate between < 2 % (Ag) and about 10 % (Cu, and Au). To exclude the linear extrapolation of the solid and liquid Gibbs energies as source of error, simulations at different temperatures are conducted. New simulation temperatures are chosen such that they are close to the old estimated MP. Consequently, Gibbs energy curves are approximated with a quadratic extrapolation. This ansatz still leads to significant deviations from the experiment, but fits well to previously reported computations of the MP of Cu with PBE and LDA.²⁹⁵ Additionally, the system size is increased, and harder PAWs (projector augmented waves) are investigated, without significant improvement of the results. Furthermore, attempts were made to increase the level of theory by TPT to the meta-GGA SCAN, which should in principle describe the complex electronic structure of the coinage metals better, but again without any significant enhancement. It was shown before that DFT methods generally suffer from poor descriptions of the polarization of the lower lying *d*-shells,^{296,297} and the poor performance of DFT compared to coupled cluster for the deep-lying 6s orbital is also known.²⁹⁸

Moreover, the intersection of the Gibbs energy curves is steep in the BP case, but in the MP region, the curves run almost parallel. Therefore, an error of 1 meV/atom in the Gibbs energy results in a roughly 10 times higher deviation in the MP than in the BP (1 K vs. 10 K for BP and MP, respectively). Comparing the computed electronic entropies to experimental ones reveals that the liquid phase is simulated rather well, which reflects in the BPs. However, the solid entropies deviate more, rendering the solid simulations more challenging than the liquid ones. Nevertheless, comparing SOR and NR phase transitions in a next step should benefit from some sort of error cancellation, and relativistic shifts are consequently more reliable.

To assess relativistic effects on MPs and BPs, SOR and scalar relativistic (SR) effects are successively switched off. It becomes visible that the majority of relativistic effects stem from SR effects. Turning off spin-orbit coupling insignificantly affects MPs or BPs of the coinage metals, just as reported for the Group 12 metals. Comparing non-relativistic (NR) and SOR calculations, BPs are influenced considerably, decreasing the BP about -123 K in case of Cu, over -225 K for Ag, and -831 K for Au, just as expected due to the nuclear charge. Interestingly, results for MPs show a contrasting situation. While the MP of Ag is similarly decreased (-210 K) as its BP when turning off SOR effects, the MP of Cu and Au are slightly increased within the non-relativistic limit ($+40$ K, and $+30$ K, for Cu, and Au, respectively). Moreover, the MP of Cu is affected slightly more than the MP of Au, although both lie within the error bars of ± 30 K. Surprisingly, the MP of NR Au is therefore similar to the one of SOR Au, defying the typically stronger influence of relativity on heavier elements. This was traced back to a strong stabilization of the SOR liquid compared to the NR case, while the solid does not profit as much from including an relativistic treatment. In contrast, Ag and Cu solids benefit more from SOR effects. For NR Au, E_{coh} for *fcc* and *hcp* lattices are energetically distinctly closer than in the SOR case, which is presumably the reason for the lower stability of the NR liquid. All in all, the results render NR Au similar to SOR Ag, consistent with a half-century-old hypothesis by Pyykkö, proposing the difference between Ag and Au essentially stems from relativity.²⁹⁹

Comparing relativistic effects of Group 11 and Group 12¹⁶⁶ elements reveals that except for the MP of Ag, Group 12 elements experience larger relative relativistic effects than the coinage metals. While phase transitions of the Group 12 metals all behave similarly to their lightest member Zn when turning off relativity, this does not hold true for the coinage metals. Additionally, MPs and BPs in Group 12 are similarly affected by relativistic effects, while they are not within Group 11, as outlined above. Thus, relativistic effects alter phase transitions of Group 12 elements more strongly than those of Group 11 elements.

Part IV

Conclusions and Outlooks

The importance of quantum mechanical methods in the explanation or prediction of experimental results or in the design of new systems with enhanced properties can rarely be overrated. To this end, an accurate three-dimensional structure of the system of interest is mandatory. Nonetheless, the environment can also exert a substantial influence on the accurate determination of properties, and its inclusion as an explicit factor in certain theoretical investigations merits consideration. Although high-performance computational resources are available, many *ab-initio* QM algorithms reach their limit when calculating large and condensed systems with several hundreds or thousands of atoms in an all-atom QM way. As a consequence, approximate QM methods that are commonly able to include the environment of a system, need to be evaluated. One recently developed semi-empirical QM approach, GFN2-xTB, is based on density functional tight binding, and is parameterized for almost the whole periodic table. Consequently, quantum mechanical calculations of all atoms for a variety of chemical challenges are possible, including structure optimizations of proteins with prosthetic groups as well as molecular dynamic simulations of explicitly solvated structures.

In a first step, GFN2-xTB was therefore assessed for ASQM geometry optimizations of proteins, pushing the boundaries by including systems up to 5000 atoms (see Chapter 3). To first validate the ansatz of comparing experimental crystal structures to molecular geometries used in the following, the packing effect is estimated for 10 small oligopeptides. It was shown that the packing effect is of similar size as the apparent experimental uncertainty of 0.5 Å.^{253,254} Therefore, comparing molecular optimizations to experimental crystal structures is valid.

The first test set, containing 70 protein structures without prosthetic groups, is optimized with GFN2-xTB and final equilibrium structures are compared to experiment. Additionally, the QM approach HF-3c, as well as two special purpose force fields (OPLS2005 and AMBER*), and two semi-empirical methods (PM6-D3H4X, and PM7) were chosen. GFN2-xTB yields similar results as OPLS2005 regarding geometric descriptors like the C_α -RMSD, while all other methods perform worse. Note that the semi-empirical PMx methods were excluded from the studies due to technical problems during the calculations. Additionally, bond length distributions were assessed, which are crucial for spectroscopic investigations. The FFs as well as HF-3c yield too narrow bond length distribution, while GFN2-xTB reproduces experimental findings best, even catching the large bond length range of the important C-S bond. Therefore, GFN2-xTB enables fast QM optimizations of organic proteins. The second test set contains large proteins with prosthetic groups, and is only tested with GFN2-xTB. Note that both force fields lack parameters for any than usual prosthetic groups including many metal atoms, and HF-3c is known to perform poorly for metal atoms, therefore they are excluded. For metalloproteins, GFN2-xTB yields similar results as for organic proteins regarding geometric descriptors. Furthermore, the coordination sphere around metal centers is described well. All in all, GFN2-xTB proofed to be an accurate and fast alternative for the computation of protein structures with the advantage of being applicable out-of-the-box to unusual binding situations, for which common FFs lack parameters.

Due to the strong performance of GFN2-xTB regarding bond length distributions of equilibrium protein structures, further spectroscopic investigations of realistic systems are possible. In Chapter 4, the influence of the environment on 1- and 2PA spectra of the fluorescent proteins bR (chromophore: retinal) and iLOV (chromophore: FMN) is assessed, using an ASQM approach. Comparison of theoretical to experimental 1- and 2PA spectra show reasonable agreement, as the main features are captured. The additionally tested dt-sTD-DFT-xTB approach, which further reduces the computational demands, gives similar results as the sTD-DFT-xTB method. Inspection of natural transition orbitals reveal that all atoms need to be included in the excited state calculation, as tryptophan and tyrosine

residues have non-negligible impact on the high energy parts of the spectra. Next to influencing the excited states of the chromophore itself, residues in any part of the protein also show absorbance without chromophore contribution, underlining the importance of the ASQM approach. Note that the choice between high-level RI-CC2 and low-level (dt-)sTD-DFT-xTB methodologies barely affect the quality of the simulated spectra. Nevertheless, side features as broadening of peaks or shoulders are not resolved correctly, which is most likely due to the neglect of dynamic structural effects.

Therefore, ADQM schemes are probed using flavin mononucleotide (FMN), including dynamic structural effects either implicitly (ADQM-B) or explicitly (ADQM-MD). Resulting 1- and 2PA spectra at the dt-sTD-DFT-xTB level of theory were compared to experiment, revealing that the ADQM-B scheme hardly improves over single structure spectra in this case. ADQM-MD results, nevertheless, provide striking agreement to experimental 1- and 2PA spectra, even catching the inhomogeneous broadening of the second absorption band. As shown by NTO analyses, the latter is due to a charge transfer including parts of the chromophore and surrounding water molecules, which can only be captured when accounting explicitly for the surroundings in an explicit dynamic manner. All in all, the results strongly advice to include surrounding atoms as well as explicit dynamic structural effects in the computation of 1- and 2PA spectra of realistic systems to gain quantitative agreement with experimental results.

Although the ADQM-B scheme shows little improvement compared to single structures in the computation of 1- and 2PA spectra, it provides valuable insights into structural features of the PES. Therefore, it was applied to investigate the conformational space of two bistable rotaxanes, whose photochemical behavior could not be clarified experimentally (see Chapter 5). The two investigated rotaxanes, featuring ZnTPP as photosensitizer, a black-hole-quencher and two linkers of different lengths were designed for a switchable singlet oxygen production by addition of acid or base. Experimentally it was found that switching is only possible if the linker is sufficiently long. Nevertheless, the $^1\text{O}_2$ production is lower than in the free ZnTPP. Additionally, the *on/off* states are reversed with respect to the original design.

Theoretical calculations reveal that strongly folded conformations of both **1b** variants and **H-1a**²⁺ dominate, as indicated by their low Gibbs energies. The investigated *half-open* and *open* conformations show higher Gibbs energies, despite favorable solvation contributions. Thus, the PS and the BHQ unit exhibit many intramolecular interactions, enabling a fast energy transfer from ZnTPP to the quencher. Additionally, the diffusion of $^3\text{O}_2$ to the reactive Zn is hindered. Therefore, the three systems show almost no $^1\text{O}_2$ production and thus represent *off* states. In contrast, **1a**⁺ features higher conformational versatility, displaying many different intramolecular Zn interactions. The Zn atom interacts with the quencher as well as with other parts of the system, rendering it more flexible than its counterpart. Furthermore, conformations with free sides in the octahedral position of the Zn atom are similarly stable as *closed* conformations with intramolecular interactions containing the Zn atom. Consequently, diffusion of $^3\text{O}_2$ towards Zn is facilitated while energy transfer from Zn to the quencher is impaired. As a result, **1a**⁺ represents the *on* state, although exhibiting a lower $^1\text{O}_2$ production than the free ZnTPP. All in all, this investigation reveals the importance of molecular folding for highly charged species, leading to a deeper understanding of the influence of the design of functional switches on their operations.

In the last part, relativistic effects on the melting and boiling points of the coinage metals Cu, Ag, and Au are investigated extensively. To this end, the recently introduced Gibbs energy-based TI-MD- λ DFT approach is applied for the computation of Gibbs energy curves of solid and liquid phases in the spin-orbit relativistic (SOR), scalar relativistic (SR) and non-relativistic (NR) limit.

Comparing the SOR BPs to experimental data reveals an excellent agreement with deviations below 1 %. On the contrary, the SOR MPs show higher differences, with absolute deviations of 2-10 %, in line with previous theoretical studies of MPs¹⁶⁶ and especially the MP of Cu.²⁹⁵ On the one hand, MPs are in general computationally more challenging, as solid and liquid Gibbs energy curves run almost parallel close to their intersection, while the liquid-gas intersection is more steep. This leads to a considerably larger error in the calculation of MPs than in BPs, as an error of 1 meV/atom in the calculated Gibbs energy results in about 10 K deviation in the MP, but only about 1 K deviation in the BP. On the other hand, exploration of the liquid and solid entropies reveals that the solid entropies deviate considerably from experimental data, especially for Cu. Note that liquid entropies are in line with experimental results, which is reflected in the good agreement of the SOR BPs. The deviating calculated MPs were therefore traced back to the more complex structures of the solid of the open-shell Group 11 elements.

In the next step, the influence of relativistic effects on MPs and BPs is explored. As expected by their nuclear charge, the BP of Cu exhibits the smallest SOR-NR shift (−123 K), followed by Ag (−225 K), and finally Au (−831 K). However, the MPs defy this relation, as the MP of Au shows a similarly small relativistic shift of 30 K than Cu (40 K), while it is of opposite sign and significantly larger for Ag (−210 K). Exploration of several thermodynamic quantities shows that the liquid phase of Au benefits stronger from including relativistic effects in the calculations than the solid phase, contrary to Ag and Cu, where the solid phase profits more. Altogether, this renders NR Au quite similar to SOR Ag, confirming a half-a-century old hypothesis of Pyykkö.²⁹⁹

To conclude, this thesis provides various insights into the application of quantum mechanical methods of realistic systems, although several challenges still await consideration. Commonly, molecular structures are compared to experimental crystal structure, and this ansatz was validated for small peptides in Chapter 3. Nevertheless, this issue should be investigated further for larger peptides and preferably large crystal proteins applying GFN*n*-xTB (*n* = 1, 2) to obtain a deeper understanding of packing effects, which can nowadays be achieved with current implementations of the xTB family into PBC containing packages.³⁰⁰ Additionally, Zheng *et al.*⁶⁴ investigated GFN2-xTB for the quantum refinement of experimental crystal structures for a small set of 13 proteins. This topic should be assessed thoroughly for a larger benchmark set including uncommon binding motives, cofactors or heavy metal atoms. In particular, a quantum mechanical refinement of proteins crystallized with the aid of uncommon crystalization agents could benefit from an improvement that goes beyond the usual FF treatment. Moreover, GFN*n*-xTB (*n* = 1, 2) could be used in the investigation of ligand binding processes in proteins, starting from the *apo*-protein and the free ligand. Another impressive challenge would be to include dynamic structural effects in the calculation of (fluorescent) proteins. As stated in Chapter 4, the 1PA spectra for iLOV could improve when dynamic structural effects are considered. In general, the influence of the environment including explicit dynamics would be desirable for other excited state properties such as the first and second hyperpolarizability or excited state absorption. Implementations are available in the std2 program, and the study presented in Chapter 4 provides a technical foundation and tested workflows. Next to excited states, ground state properties as well as structural investigations normally benefit from explicitly considering surroundings including solvation, if, and only if the solvation sphere is large enough. It would be a considerable task to include explicit solvation molecules in the generation of CREs for large and flexible systems as the rotaxanes of Chapter 5. All methods have been demonstrated to yield accurate results for large and complex systems in a fraction of the time usually required by QM methods. Consequently, the feasibility of high-throughput or screening applications for medium-sized systems becomes apparent, facilitating

the identification of optimal candidates from a set of hundreds to thousands of systems.

The recent TI-MD- λ FDT approach was demonstrated to yield good predictions for phase transition points of atomistic systems, and the extension to molecular systems is the next logical step. The TI approach integrates from a non-interacting reference to a fully-interacting DFT liquid, which is practically computed with MD simulations of partly-interacting atoms. The calculation of the Gibbs energy of the liquid phase necessitates the incorporation of damping intermolecular interactions during those MD-simulation, while it is crucial to fully account for intramolecular interactions to prevent the separation of the atoms constituting the molecular entity. For this, a proper implementation is crucial, which consequently separates intra- and intermolecular interactions in the simulation of the liquid phase.

Overall, this work has demonstrated the importance of accurate and efficient methods for the calculation of large, condensed systems. A particular emphasis was placed on the explicit description of the environment, encompassing both the exclusion and inclusion of explicit dynamic effects. Consequently, functional workflows were established that can be employed in the examination of other systems and their properties. It is thus advantageous to continue conducting research in these promising disciplines employing the investigated methodologies.

Bibliography

- [1] T. J. Zuehlsdorff and C. M. Isborn, *Modeling absorption spectra of molecules in solution*, Int. J. Quantum Chem. **119** (2019) e25719.
- [2] M. J. Harms et al., *The pKa Values of Acidic and Basic Residues Buried at the Same Internal Location in a Protein Are Governed by Different Factors*, J. Mol. Biol. **389** (2009) 34.
- [3] M. Bursch, J.-M. Mewes, A. Hansen, and S. Grimme, *Best-Practice DFT Protocols for Basic Molecular Computational Chemistry***, Angew. Chem., Int. Ed. **61** (2022) e202205735.
- [4] L. Goerigk et al., *A look at the density functional theory zoo with the advanced GMTKN55 database for general main group thermochemistry, kinetics and noncovalent interactions*, Phys. Chem. Chem. Phys. **19** (48 2017) 32184.
- [5] N. March and C. C. Matthai, *The application of quantum chemistry and condensed matter theory in studying amino-acids, protein folding and anticancer drug technology*, Theor. Chem. Acc. **125** (2010) 193.
- [6] A. Shurki and A. Warshel, *Structure/Function Correlations of Proteins using MM, QM/MM, and Related Approaches: Methods, Concepts, Pitfalls, and Current Progress*, Adv. Protein Chem. **66** (2003) 249.
- [7] A. Warshel and M. Levitt, *Theoretical studies of enzymic reactions: dielectric, electrostatic and steric stabilization of the carbonium ion in the reaction of lysozyme*, J. Mol. Biol. **103** (1976) 227.
- [8] H. M. Senn and W. Thiel, "QM/MM Methods for Biological Systems," *Atomistic Approaches in Modern Biology: From Quantum Chemistry to Molecular Simulations*, ed. by M. Reiher, Berlin, Heidelberg: Springer Berlin Heidelberg, 2007 173.
- [9] J. Fanfrlík, J. Brynda, J. Řezáč, P. Hobza, and M. Lepšík, *Interpretation of Protein/Ligand Crystal Structure using QM/MM Calculations: Case of HIV-1 Protease/Metallacarborane Complex*, J. Phys. Chem. B. **112** (2008) 15094.
- [10] P. Canfield, M. G. Dahlbom, N. S. Hush, and J. R. Reimers, *Density-functional geometry optimization of the 150 000-atom photosystem-I trimer*, J. Chem. Phys. **124** (2006) 024301.
- [11] M. Svensson et al., *ONIOM: a multilayered integrated MO+ MM method for geometry optimizations and single point energy predictions. A test for Diels- Alder reactions and Pt(P(*t*-Bu)₃)²⁺ + H₂ oxidative addition*, J. Phys. Chem. **100** (1996) 19357.

- [12] F. Maseras and K. Morokuma, *IMOMM: A new integrated ab initio + molecular mechanics geometry optimization scheme of equilibrium structures and transition states*, J. Comput. Chem. **16** (1995) 1170.
- [13] S. Humbel, S. Sieber, and K. Morokuma, *The IMOMO method: Integration of different levels of molecular orbital approximations for geometry optimization of large systems: Test for n-butane conformation and S_N2 reaction: $RCl+Cl^-$* , J. Chem. Phys. **105** (1996) 1959.
- [14] S. Dapprich, I. Komáromi, K. Byun, K. Morokuma, and M. J. Frisch, *A new ONIOM implementation in Gaussian98. Part I. The calculation of energies, gradients, vibrational frequencies and electric field derivatives*, J. Mol. Struct. THEOCHEM **461-462** (1999) 1.
- [15] D. W. Zhang and J. Z. H. Zhang, *Molecular fractionation with conjugate caps for full quantum mechanical calculation of protein–molecule interaction energy*, J. Chem. Phys. **119** (2003) 3599.
- [16] M. Wada and M. Sakurai, *A quantum chemical method for rapid optimization of protein structures*, J. Comput. Chem. **26** (2005) 160.
- [17] K. M. J. Merz, *Using Quantum Mechanical Approaches to Study Biological Systems*, Acc. Chem. Res. **47** (2014) 2804.
- [18] E. Caldeweyher and J. G. Brandenburg, *Simplified DFT methods for consistent structures and energies of large systems*, J. Condens. Matter Phys. **30** (2018) 213001.
- [19] R. Sure and S. Grimme, *Corrected small basis set Hartree-Fock method for large systems*, J. Comput. Chem. **34** (2013) 1672.
- [20] S. Grimme, J. G. Brandenburg, C. Bannwarth, and A. Hansen, *Consistent structures and interactions by density functional theory with small atomic orbital basis sets*, J. Chem. Phys. **143** (2015) 054107.
- [21] M. Cutini et al., *Assessment of Different Quantum Mechanical Methods for the Prediction of Structure and Cohesive Energy of Molecular Crystals*, J. Chem. Theory Comput. **12** (2016) 3340.
- [22] J. G. Brandenburg, C. Bannwarth, A. Hansen, and S. Grimme, *B97-3c: A revised low-cost variant of the B97-D density functional method*, J. Chem. Phys. **148** (2018) 064104.
- [23] H. M. Senn and W. Thiel, *QM/MM methods for biomolecular systems*, Angew. Chem., Int. Ed. **48** (2009) 1198.
- [24] P. Beaujean, B. Champagne, S. Grimme, and M. de Wergifosse, *All-Atom Quantum Mechanical Calculation of the Second-Harmonic Generation of Fluorescent Proteins*, J. Phys. Chem. Lett. **12** (2021) 9684.
- [25] T. Lazaridis and M. Karplus, *Effective Energy Functions for Protein Structure Prediction*, Curr. Opin. Struct. Biol. **10** (2000) 139.
- [26] D. McDonald and W. Still, *AMBER torsional parameters for the peptide backbone*, Tetrahedron Lett. **33** (1992) 7743.

-
- [27] D. M. Ferguson and P. A. Kollman, *Can the Lennard–Jones 6-12 function replace the 10-12 form in molecular mechanics calculations?* J. Comput. Chem. **12** (1991) 620.
- [28] J. L. Banks et al., *Integrated Modeling Program, Applied Chemical Theory (IMPACT)*, J. Comput. Chem. **26** (2005) 1752.
- [29] A. K. Rappe, C. J. Casewit, K. S. Colwell, W. A. I. Goddard, and W. M. Skiff, *UFF, a full periodic table force field for molecular mechanics and molecular dynamics simulations*, J. Am. Chem. Soc. **114** (1992) 10024.
- [30] S. Spicher and S. Grimme, *Robust Atomistic Modeling of Materials, Organometallic, and Biochemical Systems*, Angew. Chem., Int. Ed. **59** (2020) 15665.
- [31] A. Warshel and M. Levitt, *Theoretical studies of enzymic reactions: Dielectric, electrostatic and steric stabilization of the carbonium ion in the reaction of lysozyme*, J. Mol. Biol. **103** (1976) 227.
- [32] A. Warshel, P. K. Sharma, M. Kato, and W. W. Parson, *Modeling electrostatic effects in proteins*, Biochim. Biophys. Acta - Proteins Proteom. **1764** (2006) 1647.
- [33] R. K. Mishra et al., *cemff: A force field database for cementitious materials including validations, applications and opportunities*, Cem. Concr. Res. **102** (2017) 68.
- [34] L. Rulíšek and Z. Havlas, *Theoretical Studies of Metal Ion Selectivity. † 2. DFT Calculations of Complexation Energies of Selected Transition Metal Ions (Co^{2+} , Ni^{2+} , Cu^{2+} , Zn^{2+} , Cd^{2+} , and Hg^{2+}) in Metal-Binding Sites of Metalloproteins*, J. Phys. Chem. A **106** (2002) 3855.
- [35] S. Sinnecker and F. Neese, *QM/MM calculations with DFT for taking into account protein effects on the EPR and optical spectra of metalloproteins. Plastocyanin as a case study*, J. Comput. Chem. **27** (2006) 1463.
- [36] A. Schmiedekamp and V. Nanda, *Metal-activated histidine carbon donor hydrogen bonds contribute to metalloprotein folding and function*, J. Inorg. Biochem. **103** (2009) 1054.
- [37] C. Bannwarth, S. Ehlert, and S. Grimme, *GFN2-xTB—An Accurate and Broadly Parametrized Self-Consistent Tight-Binding Quantum Chemical Method with Multipole Electrostatics and Density-Dependent Dispersion Contributions*, J. Chem. Theory Comput. **15** (2019) 1652.
- [38] D. R. Kanis, M. A. Ratner, and T. J. Marks, *Design and construction of molecular assemblies with large second-order optical nonlinearities. Quantum chemical aspects*, Chem. Rev. **94** (1994) 195.
- [39] J. J. P. Stewart, *Optimization of Parameters for Semiempirical Methods V: Modification of NDDO Approximations and Application to 70 Elements*, J. Mol. Model. **13** (2007) 1173.
- [40] M. Gaus, Q. Cui, and M. Elstner, *DFTB3: Extension of the Self-Consistent-Charge Density-Functional Tight-Binding Method (SCC-DFTB)*, J. Chem. Theory Comput. **7** (2011) 931.
- [41] J. J. P. Stewart, *Optimization of Parameters for Semiempirical Methods I. Method*, J. Comput. Chem. **10** (1989) 209.

- [42] T. Husch, A. C. Vaucher, and M. Reiher, *Semiempirical Molecular Orbital Models Based on the Neglect of Diatomic Differential Overlap Approximation*, Int. J. Quantum Chem. **118** (2018) e25799.
- [43] E. Hückel, *Quantentheoretische Beiträge zum Benzolproblem: I. Die Elektronenkonfiguration des Benzols und verwandter Verbindungen*, Z. für Phys. **70** (1931) 204.
- [44] E. Hückel, *Quantentheoretische Beiträge zum Benzolproblem: II. Quantentheorie der induzierten Polaritäten*, Z. für Phys. **72** (1931) 310.
- [45] J. Ridley and M. Zerner, *An Intermediate Neglect of Differential Overlap Technique for Spectroscopy: Pyrrole and the Azines*, Theor. chim. acta **32** (1973) 111.
- [46] M. J. S. Dewar, E. G. Zoebisch, E. F. Healy, and J. J. P. Stewart, *Development and Use of Quantum Mechanical Molecular Models. 76. AM1: A New General Purpose Quantum Mechanical Molecular Model*, J. Am. Chem. Soc. **107** (1985) 3902.
- [47] J. J. P. Stewart, *Optimization of Parameters for Semiempirical Methods V: Modification of NDDO Approximations and Application to 70 Elements*, J. Mol. Model. **13** (2007) 1173.
- [48] G. Seifert and H. Eschrig, *LCAO- $X\alpha$ Calculations of Transition Metal Clusters*, Phys. Status Solidi **127** (1985) 573.
- [49] G. Seifert, *Tight-Binding Density Functional Theory: An Approximate Kohn-Sham DFT Scheme*, J. Phys. Chem. A **111** (2007) 5609.
- [50] M. Elstner and G. Seifert, *Density Functional Tight Binding*, Philos. Trans. R. Soc. A **372** (2014) 20120483.
- [51] M. Elstner et al., *Self-Consistent-Charge Density-Functional Tight-Binding Method for Simulations of Complex Materials Properties*, Phys. Rev. B **58** (1998) 7260.
- [52] S. Grimme, *Accurate description of van der Waals complexes by density functional theory including empirical corrections*, J. Comput. Chem. **25** (2004) 1463.
- [53] S. Grimme, *Semiempirical GGA-type density functional constructed with a long-range dispersion correction*, J. Comput. Chem. **27** (2006) 1787.
- [54] S. Grimme, J. Antony, S. Ehrlich, and H. Krieg, *A Consistent and Accurate Ab Initio Parametrization of Density Functional Dispersion Correction (DFT-D) for the 94 Elements H-Pu*, J. Chem. Phys. **132** (2010) 154104.
- [55] E. Caldeweyher, C. Bannwarth, and S. Grimme, *Extension of the D3 Dispersion Coefficient Model*, J. Chem. Phys. **147** (2017) 034112.
- [56] C. Bannwarth et al., *Extended tight-binding quantum chemistry methods*, WIREs Comput. Mol. Sci. **11** (2021) e1493.
- [57] S. Grimme, C. Bannwarth, and P. Shushkov, *A Robust and Accurate Tight-Binding Quantum Chemical Method for Structures, Vibrational Frequencies, and Noncovalent Interactions of Large Molecular Systems Parametrized for All spd-Block Elements ($Z = 1-86$)*, J. Chem. Theory Comput. **13** (5 2017) 1989.

-
- [58] F. Bohle and S. Grimme, *Efficient structural and energetic screening of fullerene encapsulation in a large supramolecular double decker macrocycle*, J. Serb. Chem. Soc. **84** (2019) 837.
- [59] M. Bursch, H. Neugebauer, and S. Grimme, *Structure optimisation of large transition-metal complexes with extended tight-binding methods*, Angew. Chem. Int. Ed. **58** (2019) 11078.
- [60] S. Dohm, M. Bursch, A. Hansen, and S. Grimme, *Semiautomated transition state localization for organometallic complexes with semiempirical quantum chemical methods*, J. Chem. Theory Comput. **16** (2020) 2002.
- [61] H. Neugebauer, F. Bohle, M. Bursch, A. Hansen, and S. Grimme, *Benchmark study of electrochemical redox potentials calculated with semiempirical and DFT methods*, J. Phys. Chem. A **124** (2020) 7166.
- [62] P. Raghavan et al., *Incorporating Synthetic Accessibility in Drug Design: Predicting Reaction Yields of Suzuki Cross-Couplings by Leveraging AbbVie's 15-Year Parallel Library Data Set*, J. Am. Chem. Soc. **146** (2024) 15070.
- [63] S. Löffelsender, M. G. Maraldi, and M. de Wergifosse, *All-Atom Quantum Mechanical Methodologies for One- and Two-Photon Absorption of Realistic Systems*, ChemPhysChem **26** (2025) e202401121.
- [64] M. Zheng et al., *Including crystallographic symmetry in quantum-based Refinement: Q|R#2*, Acta Crystallogr. D **76** (2020) 41.
- [65] S. Schmitz et al., *Quantum Chemical Calculation of Molecular and Periodic Peptide and Protein Structures*, J. Phys. Chem. B. **124** (2020) 3636.
- [66] M. Göppert-Mayer, *Über Elementarakte mit zwei Quantensprüngen*, Ann. Phys. (Berl.) **401** (1931) 273.
- [67] W. Kaiser and C. G. B. Garrett, *Two-Photon Excitation in $\text{CaF}_2:\text{Eu}^{2+}$* , Phys. Rev. Lett. **7** (6 1961) 229.
- [68] T. Maiman, *Stimulated emission of radiation in ruby*, Nat. **187** (1960) 493.
- [69] S. Alexander, G. E. Koehl, M. Hirschberg, E. K. Geissler, and P. Friedl, *Dynamic imaging of cancer growth and invasion: a modified skin-fold chamber model*, Histochem. Cell Biol. **130** (2008) 1147.
- [70] J. Paoli, M. Smedh, and M. B. Ericson, "Multiphoton laser scanning microscopy—a novel diagnostic method for superficial skin cancers," *Seminars in Cutaneous Medicine and Surgery*, vol. 28, WB Saunders, 2009 190.
- [71] K. Tanaka et al., *In vivo optical imaging of cancer metastasis using multiphoton microscopy: a short review*, Am. J. Transl. Res. **6** (2014) 179.
- [72] K. W. Dunn et al., *Functional studies of the kidney of living animals using multicolor two-photon microscopy*, Am. J. Physiol. Cell Physiol. **283** (2002) C905.

- [73] A. Diaspro, G. Chirico, and M. Collini,
Two-photon fluorescence excitation and related techniques in biological microscopy,
Q. Rev. Biophys. **38** (2005) 97.
- [74] S. Ashworth, R. Sandoval, G. Tanner, and B. Molitoris,
Two-photon microscopy: visualization of kidney dynamics, Kidney Int. **72** (2007) 416.
- [75] K. Svoboda and R. Yasuda,
Principles of two-photon excitation microscopy and its applications to neuroscience,
Neuron **50** (2006) 823.
- [76] R. K. Benninger and D. W. Piston,
Two-photon excitation microscopy for the study of living cells and tissues,
Curr. Protoc. Cell Biol. **59** (2013) 4.
- [77] C. Huang et al., *Long-term optical brain imaging in live adult fruit flies*,
Nat. Commun. **9** (2018) 872.
- [78] J. Demas et al., *High-speed, cortex-wide volumetric recording of neuroactivity at cellular resolution using light beads microscopy*, Nat. Methods **18** (2021) 1103.
- [79] F. Helmchen, M. S. Fee, D. W. Tank, and W. Denk, *A miniature head-mounted two-photon microscope: high-resolution brain imaging in freely moving animals*, Neuron **31** (2001) 903.
- [80] W. Zong et al., *Large-scale two-photon calcium imaging in freely moving mice*,
Cell **185** (2022) 1240.
- [81] T. Fischer and N. A. Hampp, *Two-Photon Absorption of Bacteriorhodopsin: Formation of a Red-Shifted Thermally Stable Photoproduct F620*, Biophys. J. **89** (2005) 1175.
- [82] B. Yao et al.,
Polarization holographic high-density optical data storage in bacteriorhodopsin film,
Appl. Opt. **44** (2005) 7344.
- [83] D. Oesterhelt, C. Bräuchle, and N. Hampp,
Bacteriorhodopsin: a biological material for information processing,
Q. Rev. Biophys. **24** (1991) 425.
- [84] C. Neiss, P. Saalfrank, M. Parac, and S. Grimme,
Quantum Chemical Calculation of Excited States of Flavin-Related Molecules,
J. Phys. Chem. A **107** (2003) 140.
- [85] K. Schwinn, N. Ferré, and M. Huix-Rotllant, *UV-visible absorption spectrum of FAD and its reduced forms embedded in a cryptochrome protein*,
Phys. Chem. Chem. Phys. **22** (22 2020) 12447.
- [86] S. Salzmann, M. R. Silva-Junior, W. Thiel, and C. M. Marian,
Influence of the LOV Domain on Low-Lying Excited States of Flavin: A Combined Quantum-Mechanics/Molecular-Mechanics Investigation,
J. Phys. Chem. B **113** (2009) 15610.
- [87] M. P. Kabir, P. Ghosh, and S. Gozem,
Electronic Structure Methods for Simulating Flavin's Spectroscopy and Photophysics: Comparison of Multi-reference, TD-DFT, and Single-Reference Wave Function Methods,
J. Phys. Chem. B **128** (2024) 7545.

-
- [88] T. Andruniów and M. Olivucci, *QM/MM studies of light-responsive biological systems*, vol. 31, Challenges and Advances in Computational Chemistry and Physics, Springer, 2021, ISBN: 978-3-030-57721-6.
- [89] S. Löffelsender, P. Beaujean, and M. de Wergifosse, *Simplified quantum chemistry methods to evaluate non-linear optical properties of large systems*, WIREs Comput. Mol. Sci. **14** (2024) e1695.
- [90] H. D. Cohen and C. C. J. Roothaan, *Electric Dipole Polarizability of Atoms by the Hartree—Fock Method. I. Theory for Closed-Shell Systems*, J. Chem. Phys. **43** (1965) S34.
- [91] J. F. Ward, *Calculation of Nonlinear Optical Susceptibilities Using Diagrammatic Perturbation Theory*, Rev. Mod. Phys. **37** (1965) 1.
- [92] B. Orr and J. Ward, *Perturbation Theory of the Non-Linear Optical Polarization of an Isolated System*, Mol. Phys. **20** (1971) 513.
- [93] A. L. Fetter and J. D. Walecka, *Quantum theory of many-particle systems*, New York: McGraw-Hill, 1971.
- [94] S. Hirata and M. Head-Gordon, *Time-dependent density functional theory within the Tamm–Dancoff approximation*, Chem. Phys. Lett. **314** (1999) 291.
- [95] *stda, a program for computing excited states and response functions via simplified TD-DFT methods (sTDA, sTD-DFT, and SF-sTD-DFT)*, <https://github.com/grimme-lab/stda>, 2023.
- [96] S. Grimme, *A simplified Tamm-Dancoff density functional approach for the electronic excitation spectra of very large molecules*, J. Chem. Phys. **138** (2013).
- [97] C. Bannwarth and S. Grimme, *A simplified time-dependent density functional theory approach for electronic ultraviolet and circular dichroism spectra of very large molecules*, Comput. Theor. Chem. **1040-1041** (2014) 45.
- [98] S. Grimme and C. Bannwarth, *Ultra-fast computation of electronic spectra for large systems by tight-binding based simplified Tamm-Dancoff approximation (sTDA-xTB)*, J. Chem. Phys. **145** (2016).
- [99] C. Bannwarth and S. Grimme, *sTDA-xTB for ground state calculations*, <https://github.com/grimme-lab/xtb4stda>, 2016.
- [100] J. Seibert, J. Pisarek, S. Schmitz, C. Bannwarth, and S. Grimme, *Extension of the element parameter set for ultra-fast excitation spectra calculation (sTDA-xTB)*, Mol. Phys. **117** (2019) 1104.
- [101] M. de Wergifosse and S. Grimme, *Perspective on Simplified Quantum Chemistry Methods for Excited States and Response Properties*, J. Phys. Chem. A **125** (2021) 3841.
- [102] M. de Wergifosse and S. Grimme, *Nonlinear-response properties in a simplified time-dependent density functional theory (sTD-DFT) framework: evaluation of the first hyperpolarizability*, J. Chem. Phys. **149** (2018) 024108.

- [103] M. de Wergifosse and S. Grimme,
Nonlinear-response properties in a simplified time-dependent density functional theory (sTD-DFT) framework: evaluation of excited-state absorption spectra,
J. Chem. Phys. **150** (2019) 094112.
- [104] M. de Wergifosse, C. Bannwarth, and S. Grimme,
A simplified spin-flip time-dependent density functional theory (SF-sTD-DFT) approach for the electronic excitation spectra of very large diradicals, J. Phys. Chem. A **123** (2019) 5815.
- [105] M. de Wergifosse, J. Seibert, and S. Grimme,
Simplified time-dependent density functional theory (sTD-DFT) for molecular optical rotation,
J. Chem. Phys. **153** (2020) 084116.
- [106] M. de Wergifosse, P. Beaujean, and S. Grimme, *Ultrafast Evaluation of Two-Photon Absorption with Simplified Time-Dependent Density Functional Theory*,
J. Phys. Chem. A **126** (2022) 7534.
- [107] J. Seibert, B. Champagne, S. Grimme, and M. de Wergifosse, *Dynamic Structural Effects on the Second-Harmonic Generation of Tryptophane-Rich Peptides and Gramicidin A*,
J. Phys. Chem. B **124** (2020) 2568.
- [108] J. Olsen and P. Jørgensen,
Linear and nonlinear response functions for an exact state and for an MCSCF state,
J. Chem. Phys. **82** (1985) 3235.
- [109] T. Helgaker et al.,
Recent advances in wave function-based methods of molecular-property calculations,
Chem. Rev. **112** (2012) 543.
- [110] T. Giovannini and C. Cappelli,
Continuum vs. atomistic approaches to computational spectroscopy of solvated systems,
Chem. Commun. **59** (38 2023) 5644.
- [111] J. Olesiak-Banska et al., *Revealing Spectral Features in Two-Photon Absorption Spectrum of Hoechst 33342: A Combined Experimental and Quantum-Chemical Study*,
J. Phys. Chem. B **117** (2013) 12013.
- [112] K. Matczyszyn et al., *One- and Two-Photon Absorption of a Spiropyran–Merocyanine System: Experimental and Theoretical Studies*, J. Phys. Chem. B **119** (2015) 1515.
- [113] N. A. Murugan et al., *Hybrid density functional theory/molecular mechanics calculations of two-photon absorption of dimethylamino nitro stilbene in solution*,
Phys. Chem. Chem. Phys. **13** (27 2011) 12506.
- [114] M. Paterson, J. Kongsted, O. Christiansen, K. Mikkelsen, and C. Nielsen,
Two-photon Absorption Cross sections: an Investigation of Solvent Effects. Theoretical Studies on Formaldehyde and Water, J. Chem. Phys. **125** (2006) 184501.
- [115] T. N. Ramos, L. R. Franco, D. L. Silva, and S. Canuto,
Calculation of the one- and two-photon absorption spectra of water-soluble stilbene derivatives using a multiscale QM/MM approach, J. Chem. Phys. **159** (2023) 024309.

-
- [116] D. L. Silva et al., *The Role of Molecular Conformation and Polarizable Embedding for One- and Two-Photon Absorption of Disperse Orange 3 in Solution*, J. Phys. Chem. B **116** (2012) 8169.
- [117] M. Wielgus et al., *Two-Photon Solvatochromism II: Experimental and Theoretical Study of Solvent Effects on the Two-Photon Absorption Spectrum of Reichardt's Dye*, ChemPhysChem **14** (2013) 3731.
- [118] M. Rossano-Tapia, J. M. H. Olsen, and A. Brown, *Two-Photon Absorption Cross-Sections in Fluorescent Proteins Containing Non-canonical Chromophores Using Polarizable QM/MM*, Front. Mol. Biosci. **7** (2020).
- [119] A. H. Steindal, J. M. H. Olsen, K. Ruud, L. Frediani, and J. Kongsted, *A combined quantum mechanics/molecular mechanics study of the one- and two-photon absorption in the green fluorescent protein*, Phys. Chem. Chem. Phys. **14** (16 2012) 5440.
- [120] K. Sneskov et al., *Computational screening of one- and two-photon spectrally tuned channelrhodopsin mutants*, Phys. Chem. Chem. Phys. **15** (20 2013) 7567.
- [121] A. A. A. Attia, P. Reinholdt, and J. Kongsted, *Modeling One- and Two-Photon Excitation of 4'-(Hydroxymethyl)-4,5',8-trimethylpsoralen in Complex with DNA: Solving Electron Spill-Out Problems in Polarizable QM/MM Calculations*, Adv. Theor. Simul. **4** (2021) 2000294.
- [122] N. A. Murugan and R. Zaleśny, *Multiscale Modeling of Two-Photon Probes for Parkinson's Diagnostics Based on Monoamine Oxidase B Biomarker*, J. Chem. Inf. Model. **60** (2020) 3854.
- [123] T. Giovannini, F. Egidi, and C. Cappelli, *Molecular spectroscopy of aqueous solutions: a theoretical perspective*, Chem. Soc. Rev. **49** (16 2020) 5664.
- [124] W. R. Browne and B. L. Feringa, "Chiroptical Molecular Switches," *Molecular Switches*, vol. 1, John Wiley & Sons, Ltd, 2011, chap. 5 121.
- [125] S. Erbas-Cakmak, D. A. Leigh, C. T. McTernan, and A. L. Nussbaumer, *Artificial molecular machines*, Chem. Rev. **115** (2015) 10081.
- [126] L. Zhang, V. Marcos, and D. A. Leigh, *Molecular machines with bio-inspired mechanisms*, Proc. Natl. Acad. Sci. **115** (2018) 9397.
- [127] H. Wang, H. K. Bisoyi, X. Zhang, F. Hassan, and Q. Li, *Visible light-driven molecular switches and motors: recent developments and applications*, Chem. Eur. J. **28** (2022) e202103906.
- [128] E. A. Neal and S. M. Goldup, *Chemical consequences of mechanical bonding in catenanes and rotaxanes: isomerism, modification, catalysis and molecular machines for synthesis*, Chem. Commun. **50** (2014) 5128.
- [129] P. Wu, B. Dharmadhikari, P. Patra, and X. Xiong, *Rotaxane nanomachines in future molecular electronics*, Nanoscale Adv. **4** (17 2022) 3418.
- [130] C. J. Bruns and J. F. Stoddart, *Rotaxane-based molecular muscles*, Acc. Chem. Res. **47** (2014) 2186.

- [131] A. Goujon, E. Moulin, G. Fuks, and N. Giuseppone, *[c2]Daisy Chain Rotaxanes as Molecular Muscles*, *CCS Chem.* **1** (2019) 83.
- [132] J. Kohn, S. Spicher, M. Bursch, and S. Grimme, *Quickstart guide to model structures and interactions of artificial molecular muscles with efficient computational methods*, *Chem. Commun.* **58** (2 2022) 258.
- [133] G. Du, E. Moulin, N. Jouault, E. Buhler, and N. Giuseppone, *Muscle-like Supramolecular Polymers: Integrated Motion from Thousands of Molecular Machines*, *Angew. Chem. Int. Ed.* **51** (2012) 12504.
- [134] M. Karimi, H. Mirshekari, M. Aliakbari, P. Sahandi-Zangabad, and M. R. Hamblin, *Smart mesoporous silica nanoparticles for controlled-release drug delivery*, *Nanotechnol. Rev.* **5** (2016) 195.
- [135] W. Chen, C. A. Glackin, M. A. Horwitz, and J. I. Zink, *Nanomachines and Other Caps on Mesoporous Silica Nanoparticles for Drug Delivery*, *Acc. Chem. Res.* **52** (2019) 1531.
- [136] M. T. Nistor and A. G. Rusu, "Chapter 3 - Nanorobots With Applications in Medicine," *Polymeric Nanomaterials in Nanotherapeutics*, ed. by C. Vasile, Micro and Nano Technologies, Elsevier, 2019 123, ISBN: 978-0-12-813932-5.
- [137] D. A. Leigh, V. Marcos, and M. R. Wilson, *Rotaxane Catalysts*, *ACS Catal.* **4** (2014) 4490.
- [138] V. Blanco, D. A. Leigh, and V. Marcos, *Artificial switchable catalysts*, *Chem. Soc. Rev.* **44** (15 2015) 5341.
- [139] A. Martinez-Cuezva, A. Saura-Sanmartin, M. Alajarin, and J. Berna, *Mechanically Interlocked Catalysts for Asymmetric Synthesis*, *ACS Catal.* **10** (2020) 7719.
- [140] V. Blanco, D. A. Leigh, U. Lewandowska, B. Lewandowski, and V. Marcos, *Exploring the Activation Modes of a Rotaxane-Based Switchable Organocatalyst*, *J. Am. Chem. Soc.* **136** (2014) 15775.
- [141] K. Eichstaedt et al., *Switching between Anion-Binding Catalysis and Aminocatalysis with a Rotaxane Dual-Function Catalyst*, *J. Am. Chem. Soc.* **139** (2017) 9376.
- [142] C.-S. Kwan, A. S. C. Chan, and K. C.-F. Leung, *A Fluorescent and Switchable Rotaxane Dual Organocatalyst*, *Org. Lett.* **18** (2016) 976.
- [143] X. Ma et al., *A room temperature phosphorescence encoding [2]rotaxane molecular shuttle*, *Chem. Sci.* **7** (7 2016) 4582.
- [144] A. Ghosh, I. Paul, M. Adlung, C. Wickleder, and M. Schmittl, *Oscillating Emission of [2]Rotaxane Driven by Chemical Fuel*, *Org. Lett.* **20** (2018) 1046.
- [145] S. F. M. van Dongen, S. Cantekin, J. A. A. W. Elemans, A. E. Rowan, and R. J. M. Nolte, *Functional interlocked systems*, *Chem. Soc. Rev.* **43** (1 2014) 99.
- [146] J. E. M. Lewis, M. Galli, and S. M. Goldup, *Properties and emerging applications of mechanically interlocked ligands*, *Chem. Commun.* **53** (2 2017) 298.

-
- [147] D. Sluysmans and J. F. Stoddart, *The burgeoning of mechanically interlocked molecules in chemistry*, Trends Chem. **1** (2019) 185.
- [148] Z.-Q. Cao et al., *A Perylene-Bridged Switchable [3]Rotaxane Molecular Shuttle with a Fluorescence Output*, Asian J. Org. Chem. **4** (2015) 212.
- [149] G. Ragazzon, A. Credi, and B. Colasson, *Thermodynamic Insights on a Bistable Acid–Base Switchable Molecular Shuttle with Strongly Shifted Co-conformational Equilibria*, Chem. Eur. J. **23** (2017) 2149.
- [150] D. García-Fresnadillo, *Singlet Oxygen Photosensitizing Materials for Point-of-Use Water Disinfection with Solar Reactors*, ChemPhotoChem **2** (2018) 512.
- [151] J. H. Correia, J. A. Rodrigues, S. Pimenta, T. Dong, and Z. Yang, *Photodynamic Therapy Review: Principles, Photosensitizers, Applications, and Future Directions*, Pharm. **13** (2021).
- [152] P. S. Maharjan and H. K. Bhattarai, *Singlet Oxygen, Photodynamic Therapy, and Mechanisms of Cancer Cell Death*, J. Oncol. **2022** (2022) 7211485.
- [153] D. Kalaitzakis, M. Sofiadis, V. Tsopanakis, T. Montagnon, and G. Vassilikogiannakis, *Merging singlet-oxygen induced furan oxidations with organocatalysis: synthesis of enantiopure cyclopentanones and hydrindanes*, Org. Biomol. Chem. **18** (15 2020) 2817.
- [154] A. A. Ghogare and A. Greer, *Using Singlet Oxygen to Synthesize Natural Products and Drugs*, Chemical Reviews **116** (2016) 9994.
- [155] P. Bayer, R. Pérez-Ruiz, and A. Jacobi von Wangelin, *Stereoselective Photooxidations by the Schenck Ene Reaction*, ChemPhotoChem **2** (2018) 559.
- [156] J. Riebe et al., *Molecular folding governs switchable singlet oxygen photoproduction in porphyrin-decorated bistable rotaxanes*, Commun. Chem. **7** (2024) 171.
- [157] J. M. Haile, I. Johnston, A. J. Mallinckrodt, and S. McKay, *Molecular dynamics simulation: elementary methods*, Comput. Phys. **7** (1993) 625.
- [158] M. P. Allen and D. J. Tildesley, *Computer simulation of liquids*, Oxford university press, Oxford; 640 pgs., 2017.
- [159] P. Fulde, *Electron correlations in molecules and solids*, vol. 100, Springer Science & Business Media, Berlin; 480 pgs., 1995.
- [160] J. J. Shepherd, G. Booth, A. Grüneis, and A. Alavi, *Full configuration interaction perspective on the homogeneous electron gas*, Phys. Rev. B **85** (8 2012) 081103.
- [161] G. H. Booth, A. Grüneis, G. Kresse, and A. Alavi, *Towards an exact description of electronic wavefunctions in real solids*, Nat. **493** (2013) 365.
- [162] A. J. Cohen, P. Mori-Sánchez, and W. Yang, *Insights into Current Limitations of Density Functional Theory*, Sci. **321** (2008) 792.
- [163] R. O. Jones, *Density functional theory: Its origins, rise to prominence, and future*, Rev. Mod. Phys. **87** (3 2015) 897.

- [164] M. G. Medvedev, I. S. Bushmarinov, J. Sun, J. P. Perdew, and K. A. Lyssenko, *Density functional theory is straying from the path toward the exact functional*, Sci. **355** (2017) 49.
- [165] J.-M. Mewes and O. R. Smits, *Accurate elemental boiling points from first principles*, Phys. Chem. Chem. Phys. **22** (41 2020) 24041.
- [166] J.-M. Mewes and P. Schwerdtfeger, *Exclusively Relativistic: Periodic Trends in the Melting and Boiling Points of Group 12*, Angew. Chem., Int. Ed. **60** (2021) 7703.
- [167] F. Dorner, Z. Sukurma, C. Dellago, and G. Kresse, *Melting Si: Beyond Density Functional Theory*, Phys. Rev. Lett. **121** (19 2018) 195701.
- [168] K. Glantschnig and C. Ambrosch-Draxl, *Relativistic effects on the linear optical properties of Au, Pt, Pb and W*, New J. Phys. **12** (2010) 103048.
- [169] P. Schwerdtfeger, *Relativistic and electron-correlation contributions in atomic and molecular properties: benchmark calculations on Au and Au₂*, Chem. Phys. Lett. **183** (1991) 457.
- [170] P. Schwerdtfeger, P. D. W. Boyd, S. Brienne, and A. K. Burrell, *Relativistic effects in gold chemistry. 4. Gold(III) and gold(V) compounds*, Inorg. Chem. **31** (1992) 3411.
- [171] P. Pykkö, *Strong Closed-Shell Interactions in Inorganic Chemistry*, Chem. Rev. **97** (1997) 597.
- [172] F. Jensen, *Introduction to Computational Chemistry. Third Edition*. John Wiley & Sons, Ltd., 2017, ISBN: 978-1-118-82599-0.
- [173] A. Szabo and N. S. Ostlund, *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory*. Dover Publications, Inc., 1996, ISBN: 978-0-486-69186-2.
- [174] E. Schrödinger, *An Undulatory Theory of the Mechanics of Atoms and Molecules*, Phys. Rev. **28** (6 1926) 1049.
- [175] M. Born and R. Oppenheimer, *Zur Quantentheorie der Molekeln*, Ann. Phys. **389** (1927) 457.
- [176] S. Hirata, E. B. Miller, Y.-y. Ohnishi, and K. Yagi, *On the Validity of the Born- Oppenheimer Separation and the Accuracy of Diagonal Corrections in Anharmonic Molecular Vibrations*, J. Phys. Chem. A **113** (2009) 12461.
- [177] J. C. Slater, *Atomic Shielding Constants*, Phys. Rev. **36** (1 1930) 57.
- [178] W. Pauli, *Über den Zusammenhang des Abschlusses der Elektronengruppen im Atom mit der Komplexstruktur der Spektren*, Z. Phys. **31** (1925) 765.
- [179] D. R. Hartree and W. Hartree, *Self-consistent field, with exchange, for beryllium*, Proc. R. Soc. Lond. A Math. Phys. Sci. **150** (1935) 9.
- [180] T. Helgaker, P. Jorgensen, and J. Olsen, *Molecular electronic-structure theory*, John Wiley & Sons, Ltd., 2013.
- [181] R. J. Bartlett, *Many-body perturbation theory and coupled cluster theory for electron correlation in molecules*, Annu. Rev. Phys. Chem. **32** (1981) 359.

-
- [182] R. J. Bartlett and M. Musiał, *Coupled-cluster theory in quantum chemistry*, Rev. Mod. Phys. **79** (2007) 291.
- [183] R. J. Bartlett, “Chapter 42 - How and why coupled-cluster theory became the pre-eminent method in an ab initio quantum chemistry,” *Theory and Applications of Computational Chemistry*, ed. by C. E. Dykstra, G. Frenking, K. S. Kim, and G. E. Scuseria, Amsterdam: Elsevier, 2005 1191.
- [184] C. C. J. Roothaan, *New developments in molecular orbital theory*, Rev. Mod. Phys. **23** (1951) 69.
- [185] G. G. Hall, *The molecular orbital theory of chemical valency VIII. A method of calculating ionization potentials*, Proc. R. Soc. Lond. A Math. Phys. Sci. **205** (1951) 541.
- [186] P. Hohenberg and W. Kohn, *Inhomogeneous Electron Gas*, Phys. Rev. **136** (1964) B864.
- [187] S. Laricchia, L. A. Constantin, E. Fabiano, and F. Della Sala, *Laplacian-Level Kinetic Energy Approximations Based on the Fourth-Order Gradient Expansion: Global Assessment and Application to the Subsystem Formulation of Density Functional Theory*, J. Chem. Theory Comput. **10** (2014) 164.
- [188] D. García-Aldea and J. E. Alvarellos, *Kinetic energy density study of some representative semilocal kinetic energy functionals*, J. Chem. Phys. **127** (2007) 144109.
- [189] W. Kohn and L. J. Sham, *Self-Consistent Equations Including Exchange and Correlation Effects*, Phys. Rev. **140** (1965) A1133.
- [190] E. J. Meijer and M. Sprik, *A density-functional study of the intermolecular interactions of benzene*, J. Chem. Phys. **105** (1996) 8684.
- [191] S. Grimme, A. Hansen, J. G. Brandenburg, and C. Bannwarth, *Dispersion-Corrected Mean-Field Electronic Structure Methods*, Chem. Rev. **116** (2016) 5105.
- [192] J. P. Perdew and K. Schmidt, *Jacob’s ladder of density functional approximations for the exchange-correlation energy*, AIP Conf. Proc. **577** (2001) 1.
- [193] J. P. Perdew, K. Burke, and M. Ernzerhof, *Generalized Gradient Approximation Made Simple*, Phys. Rev. Lett. **77** (18 1996) 3865.
- [194] J. P. Perdew, K. Burke, and M. Ernzerhof, *Generalized Gradient Approximation Made Simple [Phys. Rev. Lett. 77, 3865 (1996)]*, Phys. Rev. Lett. **78** (7 1997) 1396.
- [195] J. P. Perdew, *Density-functional approximation for the correlation energy of the inhomogeneous electron gas*, Phys. Rev. B **33** (12 1986) 8822.
- [196] A. D. Becke, *Density-functional exchange-energy approximation with correct asymptotic behavior*, Phys. Rev. A **38** (6 1988) 3098.

- [197] J. P. Perdew et al.,
Restoring the Density-Gradient Expansion for Exchange in Solids and Surfaces,
Phys. Rev. Lett. **100** (13 2008) 136406.
- [198] J. P. Perdew et al., *Erratum: Restoring the Density-Gradient Expansion for Exchange in Solids and Surfaces [Phys. Rev. Lett. 100, 136406 (2008)]*, Phys. Rev. Lett. **102** (3 2009) 039902.
- [199] V. N. Staroverov, G. E. Scuseria, J. Tao, and J. P. Perdew, *Comparative assessment of a new nonempirical density functional: Molecules and hydrogen-bonded complexes*,
J. Chem. Phys. **119** (2003) 12129.
- [200] J. W. Furness, A. D. Kaplan, J. Ning, J. P. Perdew, and J. Sun,
Accurate and Numerically Efficient r2SCAN Meta-Generalized Gradient Approximation,
J. Phys. Chem. Lett. **11** (2020) 8208.
- [201] C. Adamo and V. Barone,
Toward reliable density functional methods without adjustable parameters: The PBE0 model,
J. Chem. Phys. **110** (1999) 6158.
- [202] A. D. Becke, *Density-functional thermochemistry. III. The role of exact exchange*,
J. Chem. Phys. **98** (1993) 5648.
- [203] P. J. Stephens, F. J. Devlin, C. F. Chabalowski, and M. J. Frisch,
Ab Initio Calculation of Vibrational Absorption and Circular Dichroism Spectra Using Density Functional Force Fields, J. Phys. Chem. **98** (1994) 11623.
- [204] S. Grimme,
Semiempirical hybrid density functional with perturbative second-order correlation,
J. Chem. Phys. **124** (2006) 034108.
- [205] L. Goerigk and S. Grimme, *Efficient and Accurate Double-Hybrid-Meta-GGA Density Functionals—Evaluation with the Extended GMTKN30 Database for General Main Group Thermochemistry, Kinetics, and Noncovalent Interactions*,
J. Chem. Theory Comput. **7** (2011) 291.
- [206] J.-D. Chai and M. Head-Gordon,
Systematic optimization of long-range corrected hybrid density functionals,
J. Chem. Phys. **128** (2008) 084106.
- [207] S. Grimme, A. Hansen, S. Ehlert, and J.-M. Mewes,
r2SCAN-3c: A “Swiss army knife” composite electronic-structure method,
J. Chem. Phys. **154** (2021) 064103.
- [208] M. Müller, A. Hansen, and S. Grimme, *ω B97X-3c: A composite range-separated hybrid DFT method with a molecule-optimized polarized valence double- ζ basis set*,
J. Chem. Phys. **158** (2023) 014103.
- [209] H. Kruse and S. Grimme,
A geometrical correction for the inter- and intra-molecular basis set superposition error in Hartree-Fock and density functional theory calculations for large systems,
J. Chem. Phys. **136** (2012) 154101.

-
- [210] J. G. Brandenburg et al., *Geometrical Correction for the Inter- and Intramolecular Basis Set Superposition Error in Periodic Density Functional Theory Calculations*, J. Phys. Chem. A **117** (2013) 9282.
- [211] E. Caldeweyher et al., *A generally applicable atomic-charge dependent London dispersion correction*, J. Chem. Phys. **150** (2019) 154122.
- [212] S. Grimme, S. Ehrlich, and L. Goerigk, *Effect of the damping function in dispersion corrected density functional theory*, J. Comput. Chem. **32** (2011) 1456.
- [213] F. Calvo, E. Pahl, M. Wormit, and P. Schwerdtfeger, *Evidence for Low-Temperature Melting of Mercury owing to Relativity*, Angew. Chem., Int. Ed. **52** (2013) 7583.
- [214] J. B. Kleine Büning, S. Grimme, and M. Bursch, *Machine learning-based correction for spin-orbit coupling effects in NMR chemical shift calculations*, Phys. Chem. Chem. Phys. **26** (6 2024) 4870.
- [215] S. Löffelsender, P. Schwerdtfeger, S. Grimme, and J.-M. Mewes, *It's Complicated: On Relativistic Effects and Periodic Trends in the Melting and Boiling Points of the Group 11 Coinage Metals*, J. Am. Chem. Soc. **144** (2022) 485.
- [216] A. S. Christensen, T. Kubař, Q. Cui, and M. Elstner, *Semiempirical Quantum Mechanical Methods for Noncovalent Interactions for Chemical and Biochemical Applications*, Chem. Rev. **116** (2016) 5301.
- [217] T. Husch, A. C. Vaucher, and M. Reiher, *Semiempirical Molecular Orbital Models Based on the Neglect of Diatomic Differential Overlap Approximation*, Int. J. Quantum Chem. **118** (2018) e25799.
- [218] M. J. S. Dewar and W. Thiel, *Ground States of Molecules. 38. The MNDO Method. Approximations and Parameters*, J. Am. Chem. Soc. **99** (1977) 4899.
- [219] B. M. Axilrod and E. Teller, *Interaction of the van der Waals Type Between Three Atoms*, J. Chem. Phys. **11** (1943) 299.
- [220] Y. Muto, *Force between nonpolar molecules*, Proc. Phys.-Math. Soc. Jpn. **17** (1943) 629.
- [221] K. Nishimoto and N. Mataga, *Electronic Structure and Spectra of Some Nitrogen Heterocycles*, Z. Phys. Chem. **12** (1957) 335.
- [222] K. Ohno, *Some remarks on the Pariser-Parr-Pople method*, Theor. Chim. Acta **2** (1964) 219.
- [223] G. Klopman, *A Semiempirical Treatment of molecular Structures. II. Molecular Terms and Application to diatomic Molecules*, J. Am. Chem. Soc. **86** (1964) 4550.
- [224] N. D. Mermin, *Thermal Properties of the Inhomogeneous Electron Gas*, Phys. Rev. **137** (5A 1965) A1441.
- [225] E. R. Sekera et al., *Utilization of bis-MPA Dendrimers for the Calibration of Ion Mobility Collision Cross Section Calculations*, J. Am. Soc. Mass Spectrom. **35** (2024) 1101.

- [226] E. Runge and E. K. U. Gross, *Density-Functional Theory for Time-Dependent Systems*, Phys. Rev. Lett. **52** (12 1984) 997.
- [227] M. E. Casida, "Time-dependent density functional response theory for molecules," *Recent Advances In Density Functional Methods: (Part I)*, World Scientific, 1995 155.
- [228] E. Gross, J. Dobson, and M. Petersilka, *Density Functional Theory II*, Springer Series (1996).
- [229] J. M. Herbert, "Chapter 3 - Density-functional theory for electronic excited states," *Theoretical and Computational Photochemistry*, ed. by C. García-Iriepa and M. Marazzi, Elsevier, 2023 69, ISBN: 978-0-323-91738-4.
- [230] C. Adamo and D. Jacquemin, *The calculations of excited-state properties with Time-Dependent Density Functional Theory*, Chem. Soc. Rev. **42** (3 2013) 845.
- [231] P.-O. Löwdin, *On the Non-Orthogonality Problem Connected with the Use of Atomic Wave Functions in the Theory of Molecules and Crystals*, J. Chem. Phys. **18** (1950) 365.
- [232] D. C. Ghosh and N. Islam, *Semiempirical evaluation of the global hardness of the atoms of 103 elements of the periodic table using the most probable radii as their size descriptors*, Int. J. Quantum Chem. **110** (2010) 1206.
- [233] R. G. Parr, *A Method for Estimating Electronic Repulsion Integrals Over LCAO MO'S in Complex Unsaturated Molecules*, J. Chem. Phys. **20** (1952) 1499.
- [234] R. J. Buenker and S. D. Peyerimhoff, *Individualized configuration selection in CI calculations with subsequent energy extrapolation*, Theor. Chim. Acta **35** (1974) 33.
- [235] R. J. Buenker and S. D. Peyerimhoff, *Energy extrapolation in CI calculations*, Theor. Chim. Acta **39** (1975) 217.
- [236] S. Grimme and M. Waletzke, *A combination of Kohn–Sham density functional theory and multi-reference configuration interaction methods*, J. Chem. Phys. **111** (1999) 5645.
- [237] F. Neese, *A spectroscopy oriented configuration interaction procedure*, J. Chem. Phys. **119** (2003) 9428.
- [238] M. T. Beerepoot, D. H. Friese, N. H. List, J. Kongsted, and K. Ruud, *Benchmarking two-photon absorption cross sections: performance of CC2 and CAM-B3LYP*, Phys. Chem. Chem. Phys. **17** (2015) 19306.
- [239] P. Pracht and S. Grimme, *Calculation of absolute molecular entropies and heat capacities made simple*, Chem. Sci. **12** (19 2021) 6551.
- [240] A. Togo and I. Tanaka, *First principles phonon calculations in materials science*, Scr. Mater. **108** (2015) 1.
- [241] F. Dorner, Z. Sukurma, C. Dellago, and G. Kresse, *Melting Si: Beyond Density Functional Theory*, Phys. Rev. Lett. **121** (19 2018) 195701.
- [242] P. Pracht, F. Bohle, and S. Grimme, *Automated exploration of the low-energy chemical space with fast quantum chemical methods*, Phys. Chem. Chem. Phys. **22** (14 2020) 7169.

-
- [243] S. Maki-Yonekura, K. Kawakami, K. Takaba, T. Hamaguchi, and K. Yonekura, *Measurement of charges and chemical bonding in a cryo-EM structure*, Commun. Chem. **6** (2023) 98.
- [244] L. Cao and U. Ryde, *On the difference between additive and subtractive QM/MM calculations*, Front. Chem. **6** (2018) 89.
- [245] H. Lin and D. G. Truhlar, *QM/MM: what have we learned, where are we, and where do we go from here?* Theor. Chem. Acc. **117** (2007) 185.
- [246] M. J. Ramos and P. A. Fernandes, *Computational enzymatic catalysis*, Acc. Chem. Res. **41** (2008) 689.
- [247] D. Balcells and F. Maseras, *Computational approaches to asymmetric synthesis*, New J. Chem. **31** (2007) 333.
- [248] S. F. Sousa et al., *Application of quantum mechanics/molecular mechanics methods in the study of enzymatic reaction mechanisms*, WIREs Comput. Mol. Sci. **7** (2017) e1281.
- [249] H. J. Kulik, N. Luehr, I. S. Ufimtsev, and T. J. Martinez, *Ab Initio Quantum Chemistry for Protein Structures*, J. Phys. Chem. B **116** (2012) 12501.
- [250] J. Řezáč, K. E. Riley, and P. Hobza, *Benchmark Calculations of Noncovalent Interactions of Halogenated Molecules*, J. Chem. Theory Comput. **8** (2012) 4285.
- [251] J. J. Stewart, *Optimization of parameters for semiempirical methods VI: more modifications to the NDDO approximations and re-optimization of parameters*, J. Mol. Model. **19** (2013) 1.
- [252] E. Caldeweyher and J. Gerit Brandenburg, *Simplified DFT methods for consistent structures and energies of large systems*, J. Phys. Condens. Matter. **30** (2018).
- [253] E. Eyal, S. Gerzon, V. Potapov, M. Edelman, and V. Sobolev, *The Limit of Accuracy of Protein Modeling: Influence of Crystal Packing on Protein Structure*, J. Mol. Biol. **351** (2005) 431.
- [254] T. Flores, C. Orengo, D. Moss, and J. Thornton, *Comparison of conformational characteristics in structurally similar protein pairs*, Protein Sci. **2** (1993) 1811.
- [255] K. Wiberg, *Application of the pople-santry-segal CNDO method to the cyclopropylcarbinyl and cyclobutyl cation and to bicyclobutane*, Tetrahedron **24** (1968) 1083.
- [256] I. Mayer, *On bond orders and valences in the Ab initio quantum chemical theory*, Int. J. Quantum Chem. **29** (1986) 73.
- [257] J.-D. Chai and M. Head-Gordon, *Long-range corrected hybrid density functionals with damped atom–atom dispersion corrections*, Phys. Chem. Chem. Phys. **10** (44 2008) 6615.
- [258] R. Ditchfield, W. J. Hehre, and J. A. Pople, *Self-Consistent Molecular-Orbital Methods. IX. An Extended Gaussian-Type Basis for Molecular-Orbital Studies of Organic Molecules*, J. Chem. Phys. **54** (1971) 724.

- [259] M. M. Francel et al., *Self-consistent molecular orbital methods. XXIII. A polarization-type basis set for second-row elements*, J. Chem. Phys. **77** (1982) 3654.
- [260] M. S. Gordon, J. S. Binkley, J. A. Pople, W. J. Pietro, and W. J. Hehre, *Self-consistent molecular-orbital methods. 22. Small split-valence basis sets for second-row elements*, J. Am. Chem. Soc. **104** (1982) 2797.
- [261] P. C. Hariharan and J. A. Pople, *The influence of polarization functions on molecular orbital hydrogenation energies*, Theor. Chim. Acta **28** (1973) 213.
- [262] W. J. Hehre, R. Ditchfield, and J. A. Pople, *Self-Consistent Molecular Orbital Methods. XII. Further Extensions of Gaussian-Type Basis Sets for Use in Molecular Orbital Studies of Organic Molecules*, J. Chem. Phys. **56** (1972) 2257.
- [263] B. P. Pritchard, D. Altarawy, B. Didier, T. D. Gibbsom, and T. L. Windus, *A New Basis Set Exchange: An Open, Up-to-date Resource for the Molecular Sciences Community*, J. Chem. Inf. Model. **59** (2019) 4814.
- [264] A. Onufriev, D. Bashford, and D. A. Case, *Exploring protein native states and large-scale conformational changes with a modified generalized born model*, Proteins: Struct., Funct., Bioinf. **55** (2004) 383.
- [265] G. Sigalov, A. Fenley, and A. Onufriev, *Analytical electrostatics for biomolecules: Beyond the generalized Born approximation*, J. Chem. Phys. **124** (2006) 124902.
- [266] A. W. Lange and J. M. Herbert, *Improving Generalized Born Models by Exploiting Connections to Polarizable Continuum Models. I. An Improved Effective Coulomb Operator*, J. Chem. Theory Comput. **8** (2012) 1999.
- [267] P. Pracht et al., *CREST: Conformer-Rotamer Ensemble Sampling Tool - A program for the exploration of low-energy molecular chemical space*, <https://github.com/crest-lab/crest>, 2020.
- [268] S. Grimme et al., *Efficient Quantum Chemical Calculation of Structure Ensembles and Free Energies for Nonrigid Molecules*, J. of Phys. Chem. A **125** (2021) 4039.
- [269] Universität Bonn, MCTC, Bonn, Germany, and F. Bohle, *CENSO - Commandline Energetic Sorting of Conformer-Rotamer Ensembles, Version 1.2.0*, <https://github.com/grimme-lab/censo/releases/>, 2022.
- [270] S. Grimme, *Supramolecular Binding Thermodynamics by Dispersion-Corrected Density Functional Theory*, Chem. Eur. J. **18** (2012) 9955.
- [271] S. Spicher and S. Grimme, *Efficient Computation of Free Energy Contributions for Association Reactions of Large Molecules*, J. Phys. Chem. Lett. **11** (2020) 6606.
- [272] S. Spicher and S. Grimme, *Single-Point Hessian Calculations for Improved Vibrational Frequencies and Rigid-Rotor-Harmonic-Oscillator Thermodynamics*, J. Chem. Theory Comput. **17** (2021) 1701.
- [273] A. Klamt, *Conductor-like Screening Model for Real Solvents: A New Approach to the Quantitative Calculation of Solvation Phenomena*, J. Phys. Chem. **99** (1995) 2224.

-
- [274] A. Klamt, V. Jonas, T. Bürger, and J. C. W. Lohrenz, *Refinement and Parametrization of COSMO-RS*, J. Phys. Chem. A **102** (1998) 5074.
- [275] A. Klamt, *The COSMO and COSMO-RS solvation models*, WIREs Comput. Mol. Sci. **1** (5 2011) 699.
- [276] S. Sinnecker, A. Rajendran, A. Klamt, M. Diedenhofen, and F. Neese, *Calculation of Solvent Shifts on Electronic g-Tensors with the Conductor-Like Screening Model (COSMO) and Its Self-Consistent Generalization to Real Solvents (Direct COSMO-RS)*, J. Phys. Chem. A **110** (2006) 2235.
- [277] M. Renz, M. Kess, M. Diedenhofen, A. Klamt, and M. Kaupp, *Reliable Quantum Chemical Prediction of the Localized/Delocalized Character of Organic Mixed-Valence Radical Anions. From Continuum Solvent Models to Direct-COSMO-RS*, J. Chem. Theory Comput. **8** (2012) 4189.
- [278] Y. De Coene et al., *Fluorescence-free spectral dispersion of the molecular first hyperpolarizability of bacteriorhodopsin*, J. Phys. Chem. C **121** (2017) 6909.
- [279] E. Winder, M. H. Griep, D. Lueking, and C. Friedrich, *Chemical and biological sensing utilizing fused bacteriorhodopsin protein hybrids*, Proc. 26th Army Sci. Conf. **14** (2008).
- [280] R. R. Birge and C.-F. Zhang, *Two-photon double resonance spectroscopy of bacteriorhodopsin. Assignment of the electronic and dipolar properties of the low-lying $1 A^*$ - g-like and $1 B^{*+}$ u-like π , π^* states*, J. Chem. Phys. **92** (1990) 7178.
- [281] X. Ran et al., *Probing the electron transfer between iLOV protein and Ag nanoparticles*, Mol. **25** (2020) 2544.
- [282] J. Torra et al., *Tailing miniSOG: structural bases of the complex photophysics of a flavin-binding singlet oxygen photosensitizing protein*, Sci. Rep. **9** (2019) 2428.
- [283] R. J. Homans et al., *Two photon spectroscopy and microscopy of the fluorescent flavoprotein, iLOV*, Phys. Chem. Chem. Phys. **20** (2018) 16949.
- [284] L. M. Antill, S.-y. Takizawa, S. Murata, and J. R. Woodward, *Photoinduced flavin-tryptophan electron transfer across vesicle membranes generates magnetic field sensitive radical pairs*, Mol. Phys. **117** (2019) 2594.
- [285] N. H. List et al., *Effect of chromophore encapsulation on linear and nonlinear optical properties: the case of “miniSOG”, a protein-encased flavin*, Phys. Chem. Chem. Phys. **16** (2014) 9950.
- [286] S. S. Díaz, H. Al-Zubaidi, A. C. Ross-Obare, and S. O. Obare, *Chemical reduction of chlorpyrifos driven by flavin mononucleotide functionalized titanium (IV) dioxide*, Phys. Sci. Rev. **5** (2020) 20200007.
- [287] C. Hättig and F. Weigend, *CC2 excitation energy calculations on large molecules using the resolution of the identity approximation*, J. Chem. Phys. **113** (2000) 5154.
- [288] J. Dunning Thom H., *Gaussian basis sets for use in correlated molecular calculations. I. The atoms boron through neon and hydrogen*, J. Chem. Phys. **90** (1989) 1007.

- [289] M. de Wergifosse,
Computing Excited States of Very Large Systems with Range-Separated Hybrid Functionals and the eXact Integral Simplified Time-Dependent Density Functional Theory (XsTD-DFT), J. Phys. Chem. Lett. **15** (2024) 12628.
- [290] K. Hayashi, M. Tanigawara, and J. Kishikawa,
Measurements of the driving forces of bio-motors using the fluctuation theorem, Biophys. **8** (2012) 67.
- [291] K. Yanai, K. Ishimura, A. Nakayama, and J.-y. Hasegawa,
First-Order Interacting Space Approach to Excited-State Molecular Interaction: Solvatochromic Shift of p-Coumaric Acid and Retinal Schiff Base., J. Chem. Theory Comput. **14** **7** (2018) 3643.
- [292] J. K. Lanyi, *Bacteriorhodopsin*, Annu. Rev. Physiol. **66** (2004) 665.
- [293] S. Ehlert, M. Stahn, S. Spicher, and S. Grimme,
Robust and Efficient Implicit Solvation Model for Fast Semiempirical Methods, J. Chem. Theory Comput. **17** (7 2021) 4250.
- [294] J. Sun, A. Ruzsinszky, and J. P. Perdew,
Strongly Constrained and Appropriately Normed Semilocal Density Functional, Phys. Rev. Lett. **115** (3 2015) 036402.
- [295] L.-F. Zhu, B. Grabowski, and J. Neugebauer,
Efficient approach to compute melting properties fully from ab initio with application to Cu, Phys. Rev. B **96** (22 2017) 224202.
- [296] C. Thierfelder, P. Schwerdtfeger, and T. Saue,
 ^{63}Cu and ^{197}Au nuclear quadrupole moments from four-component relativistic density-functional calculations using correct long-range exchange, Phys. Rev. A **76** (3 2007) 034502.
- [297] C. J. Cramer and D. G. Truhlar,
Density functional theory for transition metals and transition metal chemistry, Phys. Chem. Chem. Phys. **11** (2009) 10757.
- [298] D. A. Götz, R. Schäfer, and P. Schwerdtfeger, *The performance of density functional and wavefunction-based methods for 2D and 3D structures of Au10*, J. Comput. Chem. **34** (2013) 1975.
- [299] J. Desclaux and P. Pykkö, *Dirac-Fock one-centre calculations. The molecules CuH, AgH and AuH including p-type symmetry functions*, Chem. Phys. Lett. **39** (1976) 300.
- [300] A. Katbashev et al., *Overview on Building Blocks and Applications of Efficient and Robust Extended Tight Binding*, J. Phys. Chem. A **129** (2025) 2667.

Appendix

Quantum Chemical Calculation of Molecular and Periodic Peptide and Protein Structures

Sarah Schmitz¹, Jakob Seibert¹, Katja Ostermeir¹, Andreas Hansen¹, Andreas H. Göller², and Stefan Grimme¹

Received: January 20, 2020

First published: April 10, 2020

Reprinted with permission[‡] from:

S. Schmitz, J. Seibert, K. Ostermeir, A. Hansen, A. H. Göller, and S. Grimme, *Quantum Chemical Calculation of Molecular and Periodic Peptide and Protein Structures*, The Journal of Physical Chemistry B, **124(18)** (2020) 3636-3646, DOI: 10.1021/acs.jpcb.0c00549

Copyright © 2020 American Chemical Society

Own manuscript contributions:

- Data curating in parts
- Performing most calculations
- Evaluation and interpretation of all results
- Visualization of all results
- Writing and revising most parts of the manuscript

¹ Mulliken Center for Theoretical Chemistry, Rheinische Friedrich-Wilhelms-Universität Bonn, Berlingstr. 4, 53115 Bonn, Germany

² Computational Molecular Design, Pharmaceuticals, R&D, Bayer AG, 42096 Wuppertal, Germany

[‡] Permission requests to reuse material from this chapter should be directed to American Chemical Society

Quantum Chemical Calculation of Molecular and Periodic Peptide and Protein Structures

Sarah Schmitz, Jakob Seibert, Katja Ostermeir, Andreas Hansen,* Andreas H. Göller, and Stefan Grimme*

Cite This: *J. Phys. Chem. B* 2020, 124, 3636–3646

Read Online

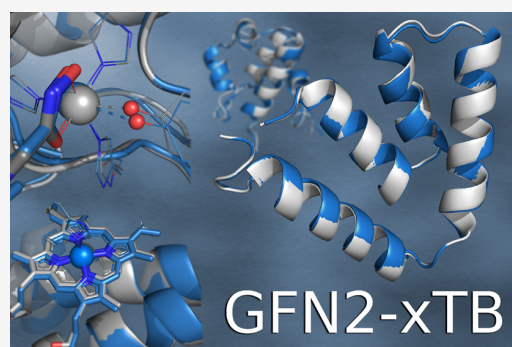
ACCESS |

Metrics & More

Article Recommendations

Supporting Information

ABSTRACT: Special-purpose classical force fields (FFs) provide good accuracy at very low computational cost, but their application is limited to systems for which potential energy functions are available. This excludes most metal-containing proteins or those containing cofactors. In contrast, the GFN2-xTB semiempirical quantum chemical method is parametrized for almost the entire periodic table. The accuracy of GFN2-xTB is assessed for protein structures with respect to experimental X-ray data. Furthermore, the results are compared with those of two special-purpose FFs, HF-3c, PM6-D3H4X, and PM7. The test sets include proteins without any prosthetic groups as well as metalloproteins. Crystal packing effects are examined for a set of smaller proteins to validate the molecular approach. For the proteins without prosthetic groups, the special purpose FF OPLS-2005 yields the smallest overall RMSD to the X-ray data but GFN2-xTB provides similarly good structures with even better bond-length distributions. For the metalloproteins with up to 5000 atoms, a good overall structural agreement is obtained with GFN2-xTB. The full geometry optimizations of protein structures with on average 1000 atoms in wall-times below 1 day establishes the GFN2-xTB method as a versatile tool for the computational treatment of various biomolecules with a good accuracy/computational cost ratio.



INTRODUCTION

In order to investigate peptides and proteins computationally, an accurate molecular structure needs to be determined at first. Calculating the structure of biologically relevant proteins consisting of several thousand atoms is computationally challenging,^{1,2} and even Hartree–Fock (HF) or density functional theory (DFT) methods cannot be applied with reasonably large atomic orbital basis sets. Hence, large protein structures are typically investigated with more approximate approaches, such as local fragmentation methods,^{3–5} repetitive application of multilevel ONIOM schemes,⁶ or QM-MM approaches.^{7–10} Smaller peptides and proteins, however, were also investigated with first-principles QM methods employing small atomic orbital (AO) basis sets,^{11–13} as well as semi-empirical methods,¹⁴ such as DFTB,^{15–17} PM3, PM6, or PM7.^{18–21}

In general, for calculating the atomistic structure of large biomolecules, QM methods compete with much faster, specially devised classical force fields (FFs).²² However, in most FFs, electronic effects such as polarization^{23–25} are usually missing and chemical reactions or proton transfers, which may play an important role in protein studies, are often not possible with FFs. Additionally, empirical parameters for metal containing systems, which account for approximately 30% of the known proteins, are difficult to obtain, and except for very common prosthetic

groups like heme groups, these parameters are typically not included in standard FFs. In contrast to FFs, QM methods capture all these effects naturally and routinely, making the latter more generally applicable, at least for single energy point calculations on subsystems of the respective peptides or proteins due to their several orders of magnitude higher computational cost.^{26,27}

Promising new candidates for an efficient QM treatment of proteins are the recently developed semiempirical, tight-binding based GFN-xTB²⁸ and GFN2-xTB²⁹ methods.³⁰ Both are parametrized for all elements up to radon, include a generalized Born implicit solvation model (GBSA), a very efficient geometry optimizer, and the well established D3 or D4 London dispersion correction.^{31–34} GFN-xTB (in the following dubbed GFN1-xTB) has been applied successfully in the efficient calculation of electronic spectra of biomolecules including a molecular dynamic (MD) treatment³⁵ as well as for optimizing the geometries of large lanthanoid metal complexes³⁶ and for

Received: January 20, 2020

Revised: April 9, 2020

Published: April 10, 2020



protonation states of various organic molecules.³⁷ Furthermore, GFN2-xTB performed remarkably well in a recent study on structure optimization of large transition metal complexes.³⁴ Since GFN2-xTB is the improved successor of GFN1-xTB, we focus here on GFN2-xTB and show the GFN1-xTB results mainly in the [Supporting Information](#). Additionally, we compare the quality of GFN2-xTB structures with corresponding HF-3c¹³ results. HF-3c is a Hartree–Fock method with a small (mostly minimal) Gaussian AO basis set that includes the D3 dispersion correction, a geometrical correction for the basis set superposition error (termed gCP scheme³⁸), and a correction for short-ranged basis set incompleteness effects. It has been already shown that HF-3c performs reasonably well for the computation of protein and related structures^{13,26,39,40} but at a significantly higher computational cost compared with the xTB methods. In a recent study by Zheng et al.⁴¹ GFN2-xTB was successfully applied in the context of protein optimization, but their investigation was limited to a few proteins.

Structure optimizations based on the two standard FFs OPLS2005⁴² and AMBER*,^{43,44} which are in common use for calculating protein structures, are assessed for comparison, too. We confess the limited practical relevance and artificial character of optimizing proteins yielding equilibrium structures at 0 K and moreover of the uncertainties in the comparison with experimental structures obtained at finite temperature (typically 100 K). However, since rigorously including temperature as well as zero-point-vibrational effects is computationally much too expensive for the herein considered set of structures, we see no alternative for benchmarking on equilibrium structures. Nevertheless, we are convinced that our approach still yields valuable insights and implications for computational studies of biomolecules.

Following and extending the study of Martinez et al.,¹² we have compiled a set of 70 peptide and protein structures as well as a set of 20 larger metalloproteins to assess the above-mentioned methods in their ability to reproduce the respective X-ray reference structures. Since the three-dimensional structure of a protein in the condensed phase results from interactions of atoms within the protein and with their environment (solvent or neighboring molecule), it is mandatory to go beyond an isolated molecule treatment in the calculations. Ideally, calculations under periodic boundary conditions (PBC) for the entire protein crystal should be conducted in order to allow a fair and consistent comparison of theory and experimental X-ray structures. Unfortunately, due to unavailable software implementations and the large size of the considered proteins and their unit cells, such treatments are rarely possible. A common approximation is to optimize protein structures with a solvation model (i.e., in solution) but still comparing them with experimental X-ray structures. In this way, important effects in the crystal such as electrostatic screening are included but other interactions crucial for the crystal packing such as intermolecular Pauli exchange–repulsion, dispersion, and charge-transfer effects are missing to a large extent. These crystal packing effects have been explored by periodic DFT calculations for the small crambin protein crystal,⁴⁵ but little is known about their magnitude for proteins more generally. Hence, we decided to shed some light on this question and conducted periodic calculations with the scaled HF-3c method (dubbed sHF-3c)⁴⁶ in comparison with the respective molecular calculations for a subset of 10 small peptides, for which a full geometry optimization was computationally feasible. The C_α RMSD between the periodic and molecular optimizations is used as a

theoretical measure for the packing effect which is discussed in section [Periodic Optimizations](#).

In the next sections, the computational details and the test sets are described before the theoretical methods are assessed and conclusions are drawn.

■ TEST SETS AND COMPUTATIONAL DETAILS

Two main test sets were compiled and employed to assess the methods mentioned above. The first set is dubbed Peptide and Protein test set, and the second one is named Metalloprotein test set.

First Test Set. The first test set consists of 70 molecular peptide and protein structures determined by X-ray crystallography taken from the Protein Data Bank.⁴⁷ It comprises 22 protein fibrils as well as peptides such as hormones and inhibitors and larger proteins with up to 1657 atoms. The average number of atoms is 897. In contrast to the study of Martinez et al.,¹² we do not include structures obtained from NMR measurements, since their accuracy is mostly not sufficient for our purpose. The peptides and proteins of the first test set do not contain any metals. Their molecular weight is below 12 kDa, and their total charges range from -2 to $+2$. We exclude peptides and proteins with unnatural amino acids, noncrystallized heavy atoms, or those that show more than 30% sequence identity or expression tags as well as aggregates. Structures were prepared using the Maestro program of the Schrödinger software package⁴⁸ and sorted by the number of atoms, resulting in the aforementioned set of 70 structures that meet the desired selection criteria (see the [Supporting Information, Table S1](#)). Periodic calculations, performed for 10 protein fibrils from this first test set, are specified in section [Periodic Optimizations](#) of this chapter.

Second Test Set. The second test set consists of 20 metalloproteins. They have been selected to contain various metals (Cu, Fe, Ca, Zn, Hg, Co, Ni, Cd) and cover a wide range of biological functions, e.g., monooxygenase or electron transport proteins. Large proteins up to a molecular weight of 40 kDa (1034–4804 atoms) and total charges from -15 to $+7$ were included. Most of the proteins include several subunits and/or are bound to ligands such as heme groups providing a real challenge for all computational methods. Further information can be found in the [Supporting Information, Table S3](#).

Structure Preparation. The structure preparation step for the systems in the benchmark has been conducted in the following general manner. In the case of multiple molecules in the unit cell, the first molecule is chosen as starting point for the calculation. Dangling bonds are saturated with hydrogens, and disulfide bonds between cysteines were assigned appropriately.⁴⁸ Some structures contain explicit water molecules, with their oxygen positions resolved in the X-ray structure. These crystal water molecules are removed except for those within 5 Å around the metals, since these are considered to be important for the first coordination sphere. Protonation states of aspartic acids, glutamic acids, lysines, arginines, and histidines are calculated for a pH of 7 by employing Schrödinger's PROPKA pK_a prediction tool⁴⁹ which produces the corresponding total molecular charge. In the case of the metalloproteins, possible metal oxidation states (see [Table 1](#)) are calculated with Schrödinger's EPIK program.^{50,51} The QM calculations are conducted with the proposed charge states and final charges are chosen, for which the optimized structure most closely shows integer orbital occupations and a maximized HOMO–LUMO

Table 1. Description of the Metalloprotein Test Set^a

PDB entry	metal	protein	proposed EPIK charges	final charge
1DXG	Fe	desulforedoxin	−10, −11, −12	−11
1B7V	Fe	cytochrome	−5, −6	−6
3D2N	Zn	mbn1 tandem zinc finger 1 and 2 domain		4
1AG6	Cu	plastocyanin	−8, −9, −10	−9
4CUE	Ca	transcription		−6
4XPX	Fe	hemerythrin	1, 2, 3	1
5FTZ	Cu	monooxygenase	6, 7, 8	6
4FWX	Fe	myoglobin	0, 1	0
1NX2	Ca	calpain		1
1IAC	Hg	zinc endopeptidase		−12
1AST	Zn	metalloproteinase (astacin)		−12
1IAA	Cu	zinc endopeptidase	−11, −12, −13	−13
1IAB	Co	zinc endopeptidase	−11, −12	−12
1IAE	Ni	zinc endopeptidase	−11, −12	−12
1QJJ	Zn	astacin + inhibitor	−11, −12	−11
1EPT	Ca	hydrolase (serine protease)		7
1CON	Cd, Ca	concanavalin		1
1A7V	Fe	cytochrome <i>c</i>	5, 6, 7	5
1BZM	Zn	carbonic anhydrase	−1, 0	0
1YME	Zn	carboxypeptidase		2

^aThe proposed charges by the Schrödinger EPIK program in the case of multiple oxidation states and final charges used are shown.

gap. For the electron transporter 1AG6 an exception was made. A total charge of −10 fulfills the above-described requirements but refers to an unusual oxidation state of copper. The state proposed by the experimentalists is Cu(II),⁵² leading to a final charge of −9. Furthermore, a fractional occupation density analysis (FOD)^{53,54} was carried out to exclude metalloproteins with artificial static correlation due to erratic structure preparation via the EPIK software (see Figure S3 in the Supporting Information for an example).

Molecular Optimizations. As the optimized structures depend on both the type of geometry optimizer used and the implicit solvent model, it is desirable to use as similar computational settings as possible for the assessment of the different approaches. However, for technical reasons, FF and QM calculations cannot be performed under exactly the same computational settings, but we tried to minimize the differences as much as possible. First, the same prepared start structure is used for full geometry optimizations. For all optimizations convergence thresholds for energy change and residual force of $E_{\text{conv}} = 5 \times 10^{-5} E_h$, $G_{\text{conv}} = 4 \times 10^{-3} E_h/\text{Bohr}$ are employed, which correspond to the optimization criteria *loose* in GFN*n*-xTB ($n = 1, 2$). For GFN*n*-xTB we employ the default implicit solvent model (GBSA(water)). The optimizer in the xtb code is based on approximate normal coordinates (ANC) from Lindh et al.'s model Hessian⁵⁵ and is also used in HF-3c calculations. The latter is performed in combination with the DCOSMO-RS(water)⁵⁶ solvent model employing the Turbomole program package version 7.2.^{57,58}

The FF optimizations were conducted with both the OPLS2005 FF⁴² and the AMBER* FF^{43,44} employing the Schrödinger software package with extended cutoffs for noncovalent interactions (van der Waals, 8.0 Å; electrostatics, 20 Å; H-bond, 4 Å) and the built-in implicit solvent model with parameters for aqueous solution. To resemble the GFN*n*-xTB

optimization procedure as closely as possible, we applied the similar LBFGS (limited-memory Broyden–Fletcher–Goldfarb–Shanno) algorithm for the FFs optimization and allowed for a number of steps uniquely determined by the xtb convergence threshold (unlimited number of steps).

For the optimizations with PM6-D3H4X^{59,60} and PM7,⁶¹ the MOPAC⁶² software is used for the single point calculations in combination with the efficient optimizer implemented in xtb. This optimizer is also used for the GFN*n*-xTB and HF-3c optimizations and employs the same optimization settings for all methods. The COSMO solvation model⁶³ ($\epsilon = 80.4$) is applied during MOPAC single point calculations. Due to technical problems, we restricted the comparison between the PM methods and GFN2-xTB to a subset of 20 smaller peptides for which the PM calculations were computationally feasible (see the Supporting Information for details; MOPAC's COSMO implementation restricts the system size).

For metal containing systems Hartree–Fock often suffers from convergence problems, and hence, the metalloproteins are not treated with HF-3c. Furthermore, the assessed FFs do not contain parameters for most metals, and therefore they are not able to treat metalloproteins.

Though a DFT/MM model with harmonic constraints can yield accurate results for metal sites of crystal structures, the aim of this work is to show that GFN2-xTB is able to compute such systems without any special preparation. Furthermore, comparing different methods (DFT/MM with constraints against GFN2-xTB without constraints) in a benchmark study, which is the focus of our paper, would bias the results.

In consequence, the metalloprotein test set is only optimized with GFN*n*-xTB as described above.

Periodic Optimizations. The original HF-3c method is known to overbind molecular crystals.⁴⁶ To account for this, the dipole–quadrupole dispersion term of HF-3c (s_8) is scaled by a factor of 0.6143, leading to sHF-3c as described by Caldeweyher and Brandenburg.⁴⁶ It is applied to a small subset of 10 peptides for periodic sHF-3c optimizations as implemented in the CRYSTAL17⁶⁴ package with a modified convergence criterion for the energy change of $10^{-6} E_h$ based on the BFGS algorithm.^{65–68} Charged structures were neutralized by (de-)protonating charged amino acid side chains (cf. Table 2). Cell parameters and symmetry groups were chosen as indicated in the PDB files. Further information on the subset can be found in the Supporting Information, Table S6.

Table 2. Neutralized Peptides and Their Mutation Sites for the Periodic Calculations

title	initial charge	mutation
3SGS	−2	GLU5 → GLH5, ASP2 → ASH2
3FPO	1	HIP1 → HIE1
3Q2X	1	LYS2 → LYN2
3NVG	1	NH3MET1 → NH2MET1
2Y29	1	LYS1 → LYN1

Geometric Descriptors and Coordination Numbers.

The three-dimensional structure of a protein can be characterized by the Φ and Ψ torsion angles of the peptide backbone, as well as by the first side chain torsion angle χ_1 and the planarity angle ω of the peptide bond (cf. Supporting Information, Figure S1). The quality of the optimized structures is evaluated by visual inspection as well as by the deviations of the calculated and experimentally determined angles as $\Delta\alpha =$

$\alpha_{\text{exp}} - \alpha_{\text{calc}}$, where α represents one of the angles specified above. In addition, the C_α RMSD (root-mean-square deviation) and the RMSD of all atoms except hydrogen (heavy atom RMSD) between experimental X-ray and optimized structure are compared. Both values were calculated based on the atom best-fit Kabsch algorithm.⁶⁹

Furthermore, a bond length distribution analysis for CC, CN, CO, and CS bonds of the optimized to the experimental data is performed with the ASE⁷⁰ software suite. Additionally, computational timings for the optimizations are given as the time for a complete optimization divided by the number of cycles needed for this optimization. The arithmetic mean is determined in section Results and Discussion.

The coordination numbers of the metals are empirically determined using Wiberg–Mayer bond orders^{71,72} of ≥ 0.2 as a threshold, that is, atoms with bonds larger than this value are counted as neighbored. The mean absolute error (MAE) is used to investigate the difference between metal and ligand bond distances in the experimental and the GFN2-xTB optimized structure.

RESULTS AND DISCUSSION

In the first subsection, periodic optimizations of the small subset of 10 oligopeptides are presented and discussed. Afterward, the results of the molecular HF-3c(DCOSMO-RS), GFN2-xTB-(GBSA), and FF(water) optimizations are shown. Finally, the metalloprotein test set is considered. Further information (name, number of atoms and residues, charge, percentage of α -helix, β -sheet and unstructured loop, protein class, and X-ray resolution) about the test is presented in the Supporting Information (Tables S1 and S3).

Periodic Optimizations. The molecular structure of a protein in solution and as a solid (crystal) differs, and if a molecular approach for theoretical studies is chosen, this deviation should be investigated.^{73–76} For this purpose, we conducted sHF-3c optimizations under PBC and molecular sHF-3c and unscaled HF-3c(DCOSMO-RS(water)) optimizations for a subset of 10 small peptides. The C_α RMSDs with respect to the corresponding experimental X-ray structure are shown in Figure 1 and Table S7 of the Supporting Information. Both molecular optimizations (sHF-3c vs HF-3c) yield on average a similar C_α RMSD (0.56 and 0.60 Å, respectively). With an average C_α RMSD of 0.43 Å the sHF-3c optimizations under PBC perform slightly better than the molecular ones except for the two peptides 2Y29 and 3SGS. For both a hydrogen transfer was observed during the optimization, which could be the reason for the larger errors of the periodic case. The similar RMSDs of sHF-3c and HF-3c imply that the changed s_8 coefficient in the dispersion treatment has a smaller influence on the structure compared to the effect of the PBC.

Experimentally determined protein structures by X-ray diffraction have an intrinsic uncertainty, and many attempts were made to quantify it.^{74,77} In a comprehensive study, Eyal et al.⁷⁴ investigated the structural deviation of proteins with identical sequence crystallized in the same space group but by different authors. The C_α RMSD between these structures amounts on average to about 0.5 Å. Flores et al.⁷⁷ found a similar average C_α RMSD of 0.51 Å but for a much smaller test set. The RMSDs of the sHF-3c and HF-3c optimized structures with respect to the respective X-ray structures are either within or close to this apparent experimental uncertainty. Note that these differences are not to be confused with experimental errors.

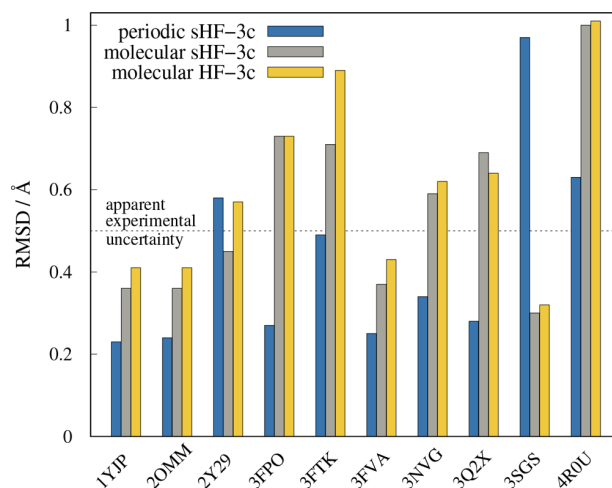


Figure 1. C_α RMSDs (in Å) for the subset of 10 peptides, employing periodic boundary conditions (sHF-3c) and within an molecular approach (sHF-3c and HF-3c), in each case with respect to the respective X-ray structures. Black dashed line at $y = 0.5$ Å depicts the found apparent experimental uncertainty.

Table 3 shows the C_α RMSDs for all peptides optimized employing periodic boundary conditions (sHF-3c, PBC) with

Table 3. C_α RMSD (in Å) between Structures Optimized with sHF-3c Employing Periodic Boundary Conditions (PBC) and sHF-3c (Molecular) for the Subset of 10 Protein Fibrils

PDB entry	C_α RMSD
1YJP	0.45
3NVG	0.56
3FPO	0.78
2OMM	0.40
3FVA	0.40
3Q2X	0.81
4R0U	1.26
3FTK	0.57
2Y29	0.57
3SGS	1.21
average	0.70

respect to molecular optimizations (sHF-3c, molecular). This value is used to estimate the packing effect, i.e., the structural difference between solid (crystal) and molecular structures. On average the packing effect calculated here amounts to 0.70 Å, which is of similar size as the apparent experimental uncertainty of about 0.5 Å. Hence, we conclude that a comparison of structures taken from molecular optimizations to experimental X-ray structures is legitimate for benchmarking purposes. An in-depth investigation of this topic is outside the scope of this study but should be conducted as soon as it becomes computationally feasible (e.g., when a periodic implementation of GFN2-xTB is available).

Molecular Optimizations. After validating the molecular approach, molecular optimizations of the first test set are performed with HF-3c, GFN n -xTB, OPLS2005, and AMBER*. We first discuss some computer timings to demonstrate the viability of a full QM approach to protein structures in practical applications.

Timings. For a better comparison of the timings, all optimizations were conducted on the same CPU cores (Intel

Xeon E3-1270 v5, 3.6 GHz, 8 M cache). HF-3c is one of the few wave function based QM methods that can be applied at all to optimize proteins of the here considered size. However, the usability of this method is severely limited. The most time-consuming optimization in the first test set took 4 months on four CPU cores. The average time per optimization step (energy and gradient) amounts roughly to 1.5 h, making this approach computationally unfeasible for large scale studies on proteins of this size (with on average 897 atoms). In contrast, the selected special purpose FFs need much less than a second for one optimization step. GFN2-xTB on average needs about 17 s/step, which is sufficiently efficient for protein optimizations with up to 5000 atoms. Due to the intrinsically worse scaling behavior of computation time with system size (cubically for QM vs quadratically for FFs), however, methods like GFN2-xTB are difficult to apply without further approximations for systems with 10 000 or more atoms.

Geometrical Descriptors. Figure 2 shows average deviations of the four considered dihedral angles as well as the C_α and the heavy atoms RMSD with respect to the experimental X-ray structure.

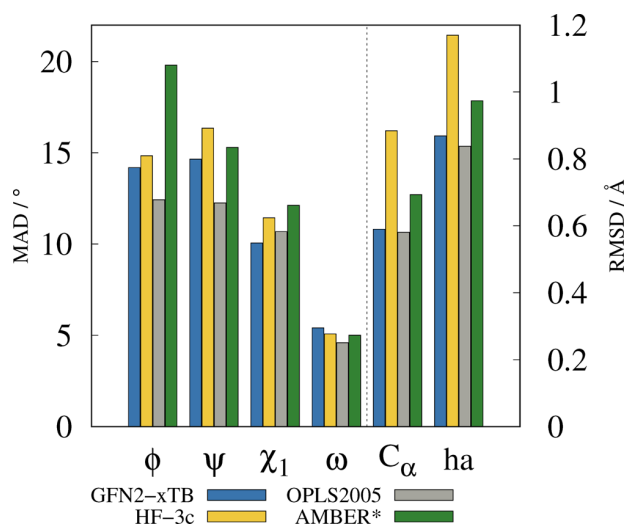


Figure 2. Average deviations of four types of dihedral angles (in degree) for the first test set with respect to the crystal structure as well as average C_α and heavy atoms RMSD (in Å).

The deviations of the angles, i.e., $\Delta\Phi$, $\Delta\Psi$, $\Delta\chi_1$, and $\Delta\omega$ are soft descriptors regarding local displacements of the protein backbone. They can be used to check for the right folding of the protein. If the protein partially unfolds or rearranges strongly in the optimization, the angle deviations increase, revealing these changes. The deviations for the angles are depicted in Figure 2. GFN2-xTB is not discussed as it performs similarly to GFN1-xTB; for details see the Supporting Information. In general, GFN2-xTB performs better than both tested PM methods for the subset of 20 peptides and proteins for most of the used geometrical descriptors. More information regarding the comparison of GFN2-xTB with PM6-D3H4X and PM7 can be found in the Supporting Information.

OPLS2005 outperforms all the other tested methods except for the first side chain angle χ_1 , where GFN2-xTB yields the best result with an average deviation of 10.1°. The planarity of the peptide moiety ω is slightly worse with GFN2-xTB when compared to HF-3c and both FFs. AMBER* and HF-3c yield

the worst results for Φ , Ψ , and χ_1 . Visual inspection of the overall structures (cf. Figure 3) shows that the secondary structure of the proteins is preserved during the GFN2-xTB optimization.

The RMSDs show a similar trend as the deviations for the angles (cf. Figure 2 and Table S2). The heavy atoms RMSD is about 0.3 Å systematically larger than the C_α RMSD for all tested methods as it includes also the floppy side chains and therefore not discussed further here. GFN2-xTB performs as well as the special purpose force field OPLS2005 directly followed by AMBER*. HF-3c produces the largest RMSDs with respect to the crystal structure. The C_α RMSDs of all investigated methods are close to the mentioned estimated apparent experimental uncertainty of 0.5 Å,^{74,77} confirming a good overall performance of the theoretical methods.

For a more detailed look into the quality of GFN2-xTB optimized structures, overlays of the proteins with very good and the worst C_α RMSD with respect to the experimental crystal structure are shown in Figure 3. For the systems 4QXX and 1YJP, the side chains are reasonably well reproduced by GFN2-xTB. This was already indicated by the small deviation of the first side angle χ_1 obtained with the latter. In general, it can be seen that the secondary structures of the proteins and protein fibrils remain intact during optimization. This is confirmed by the small deviations of the angles and RMSDs. The structures with small C_α RMSD cover all secondary structure features such as α -helices and β -sheets. Even for the worst cases with the highest C_α RMSD, the main structural motifs are preserved. The overlays presented in Figure 3 depict that the largest structural deviations are observed in random coil regions. Considering the peptides 2OVO and 3B7H (C_α RMSDs of 0.87 Å and 1.14 Å, respectively), the GFN2-xTB method reproduces the stiff and ordered secondary features well, but the floppy random coil sequences deviate clearly from the X-ray structure. In contrast, 3E4H (C_α RMSD of 0.37 Å) consisting of more than 75% unstructured loop is reproduced very well with a C_α RMSD of only 0.37 Å. 3E4H, however, has six cysteines forming three sulfur–sulfur bridges stabilizing this unstructured region, leading to a good agreement between experimental crystal and optimized structure.

All resolved water molecules in the crystal structure were removed during structure preparation. However, it remains elusive whether these water molecules are important for stabilization of the unstructured loop region, and hence, further investigation of this topic as well as a possible stabilization through explicit solvation should be performed in the future.

In conclusion, OPLS2005 overall yields the smallest deviations from experimental structures regarding the first test set, whereas AMBER* and HF-3c perform the worst. Without any special adaption to proteins, GFN2-xTB provides, however, similar good results as the special-purpose FF OPLS2005 which is very encouraging.

Selected Bond Lengths Distribution. A good description of the bond lengths is important for calculating accurately various spectroscopic properties. Therefore, the bond length distributions for all CC, CN, CO, and CS bonds are depicted and compared to experimental values in Figure 4.

Analysis of Figure 4 reveals that the FFs yield too narrow distributions with maxima close to the experimental value. As both FFs are fitted to reproduce experimental values in proteins, this behavior is expected. Furthermore, the bonds are approximated by harmonic potentials neglecting many-body effects, leading to more rigid bonds with little variability and accordingly sharper peaks. The HF-3c and GFN2-xTB QM

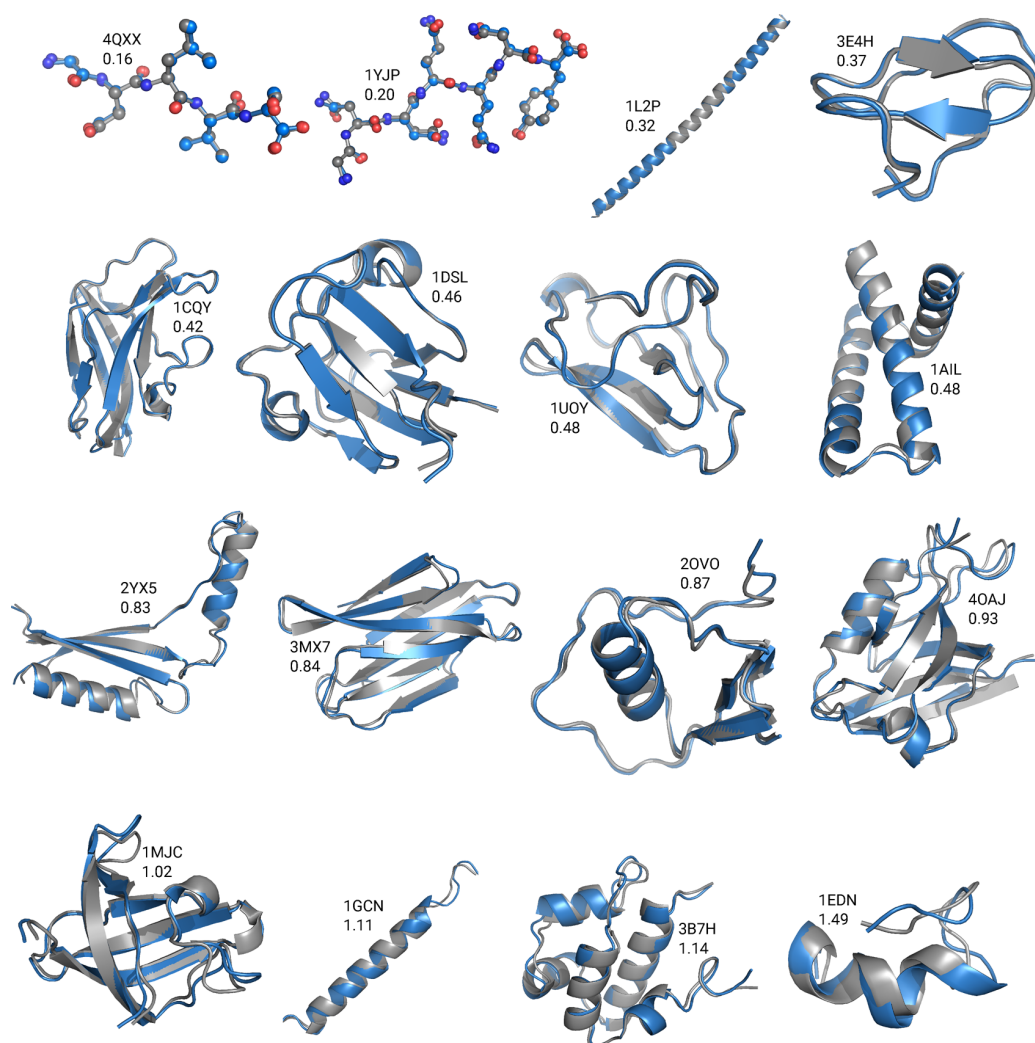


Figure 3. Overlay of experimental crystal structures (blue) and GFN2-xTB optimized structures (gray). C_{α} RMSDs (in Å) for the displayed overlays are depicted below the respective protein names.

methods are more flexible and thus yield broader distribution in better agreement with the experiment.

Overall, the different binding motifs are well described with all tested methods. Concerning C–C and C–N bonds, GFN2-xTB and HF-3c both yield too narrow bond length distributions similarly to the tested FFs. For C–O and C–S bonds, however, the distribution is much broader for GFN2-xTB and HF-3c, comparable with the experimental one, while the FFs still give significantly too narrow bond length distributions. Nevertheless, some systematic deviations are noted. HF-3c predicts, for example, the C=N and the C–S bonds too long and the C=O bond too short. The FFs deviate the most for the C=C and C–S bonds. GFN2-xTB deviates in most of the cases only slightly from the experiment. Furthermore, both FFs and HF-3c calculate the C–S single bond as being rather rigid, whereas the experimental broader distribution is captured by GFN2-xTB. This binding motif is important in the C–S–S–Cs bridges, stabilizing the tertiary structure of a protein.

Metalloproteins. Next, the GFN2-xTB results for the metalloproteins are presented; for GFN1-xTB results, see the Supporting Information (Table S5).

Geometrical Descriptors. Metal atoms in proteins are often coordinated to oxygen, nitrogen, or sulfur atoms of the respective amino acids and form cluster structures. In addition, they are frequently coordinated by water molecules. The accurate structure of metalloproteins is essential for understanding chemical behavior and biological function, and theoretical methods are needed to support structure elucidation. Since the optimization of metalloproteins is more challenging, compared with normal proteins and peptides without any prosthetic groups, the selection criteria for the second test set have been adjusted as described in the computational details.

Table 4 and Figure S3 of the Supporting Information summarize the average deviations of GFN2-xTB for all investigated dihedral angles and RMSDs of the 20 metalloproteins. The RMSDs (first row of Table 4) are somewhat larger compared to the first test set (second row of Table 4), whereas the deviations for the angles are slightly smaller.

Figure 5 depicts overlays of the best (top) and worst (bottom) calculated protein structures with respect to the C_{α} RMSD. Highly structured domains remain very well preserved during optimization, while the unstructured regions deviate considerably more from the respective crystal structures similar to what

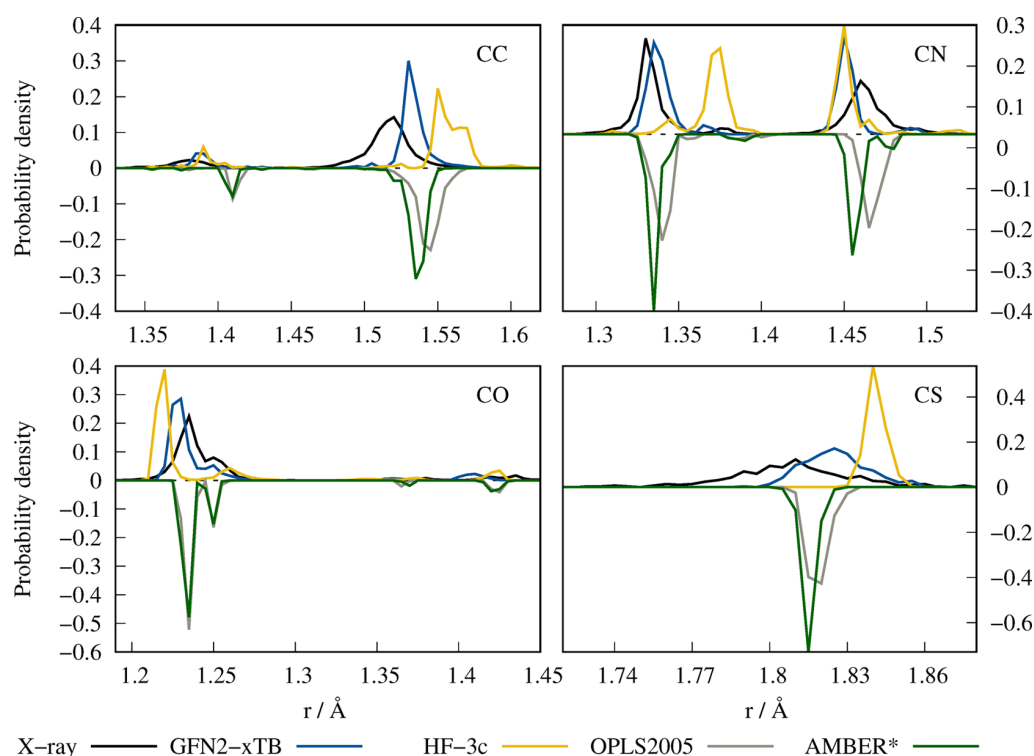


Figure 4. Probability density for CC, CN, CO, and CS bond lengths, respectively. For clarity reasons, FF results are inverted.

Table 4. Average Deviations of GFN2-xTB Optimizations of Four Types of Dihedral Angles (in Degree) of Both Test Sets with Respect to the Crystal Structure as Well as Average C_α and Heavy Atoms RMSD (in Å)

	2nd test set	1st test set
$\Delta\Phi$	12.9	14.2
$\Delta\Psi$	12.2	14.7
$\Delta\chi_1$	9.5	10.2
$\Delta\omega$	5.3	5.4
C_α	0.70	0.59
heavy atoms	0.94	0.87

was noted for the first test set. The example of 1B7V (top left) shows that the heme group stabilizes the unstructured loop such that it still matches the crystal structure reasonably well.

Coordination Sphere around Metal Centers. For an analysis of the first coordination sphere around the metal centers, we use two descriptors: the Wiberg/Mayer bond order and the MAE of the distances between the metal and its coordinated ligands (compared to experiment). On the basis of these descriptors, we classify the results into four cases. An example for each will be discussed in the following (cf. Figure 6). In the Supporting Information, details about all investigated proteins are provided (cf. Table S4).

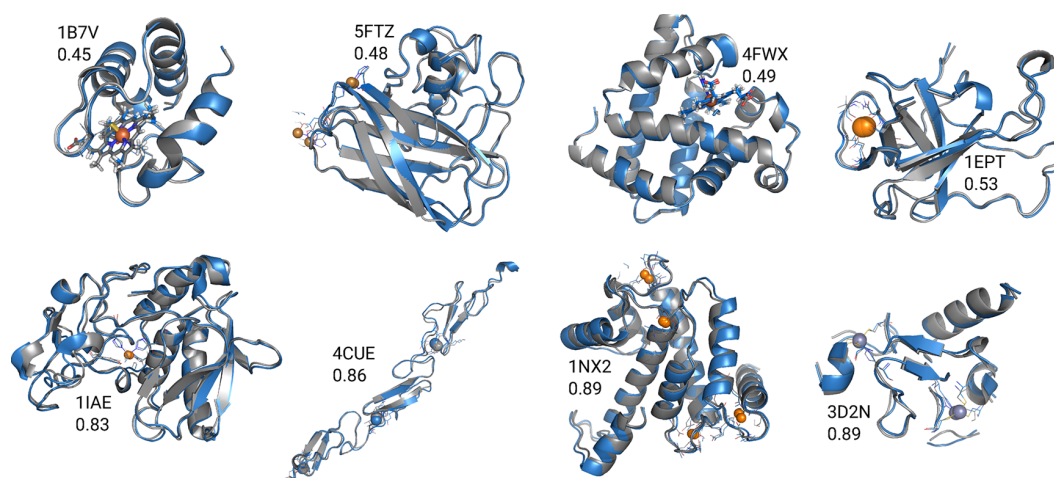


Figure 5. Overlay of experimental crystal structures (blue) and GFN2-xTB optimized structures (gray). C_α RMSDs (in Å) for the displayed overlays are depicted below the respective protein names.

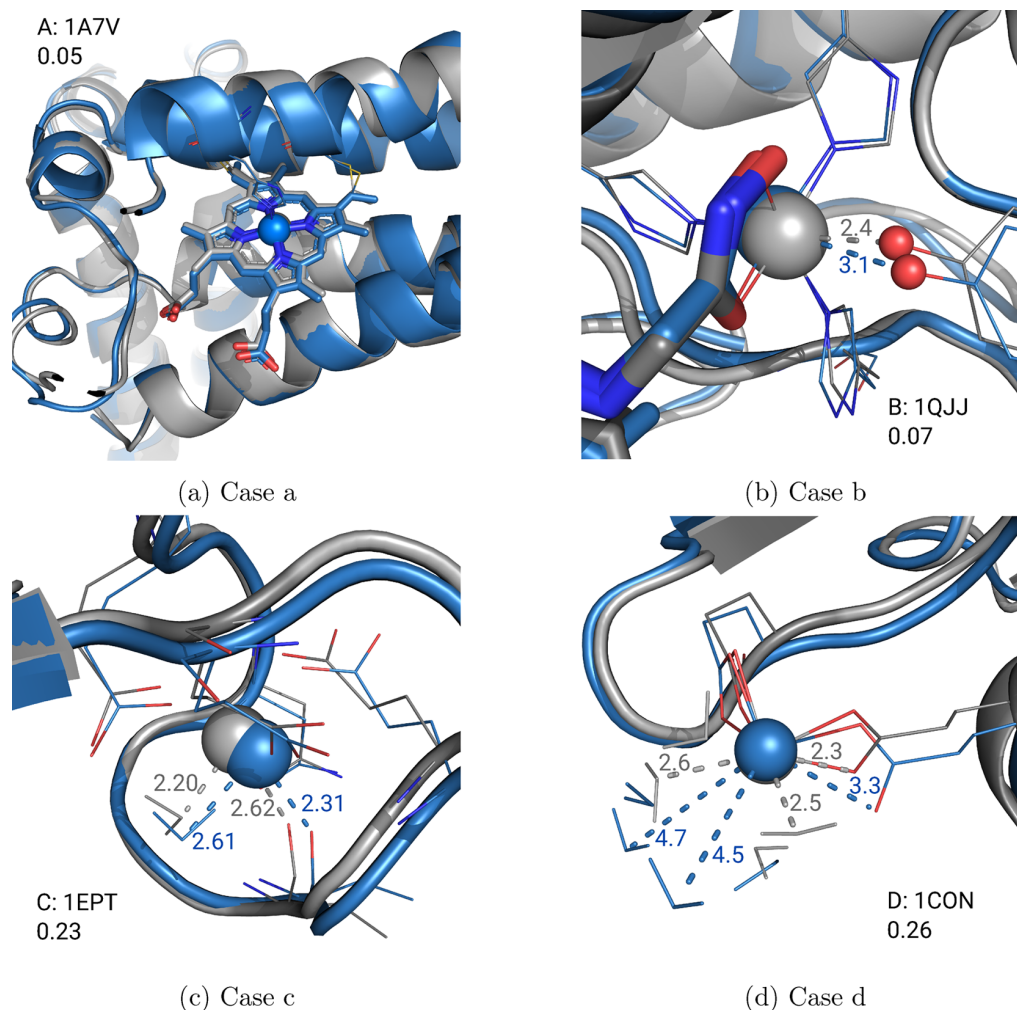


Figure 6. Exemplary structures for the four cases discussed below of changes in metal–ligand sphere during optimization. Experimental structure and experimental bond lengths (in Å) are in blue. GFN2-xTB optimized structure and bond lengths (in Å) are in gray. Only the metal with its ligands (considered in the Wiberg/Mayer bond order) are shown. The MAEs (in Å) are depicted below the respective protein names.

The first example (case a, Figure 6a) is the electron transporter 1A7V. During optimization, the coordination number of the iron center is conserved. The small MAE of only 0.05 Å indicates an excellent agreement between GFN2-xTB and the X-ray structure. The hydrolase 1QJJ is an example for case b (Figure 6b). Here, the MAE is similarly small as in the first example (0.07 Å) but the coordination number of the metal increases by one during optimization. Figure 6b depicts the approach of one ligand (atom number 1435, TYR) to the metal (zinc) by 0.7 Å in the optimization. This increases the coordination number without changing the average distances between metal and ligands too much. The example for case c (Figure 6c) is the calcium-containing hydrolase 1EPT. While the coordination number is unaffected during optimization, the MAE is as large as 0.23 Å. This indicates a ligand change, which is also visible in Figure 6c. The metal and a water molecule (atom number 2033) come close, while the backbone oxygen of a valine (ligand 528) dissociates. Case d (Figure 6d) is represented by the lectin 1CON. During optimization, the coordination number increases by 3 and the MAE amounts to 0.26 Å. The entire ligand sphere is changed in this case. In summary, the four cases can be described as follows:

- (a) $\Delta\text{CN} = 0$, MAE small $\rightarrow$ “no” change observed 13 times.
- (b) $\Delta\text{CN} \neq 0$, MAE small $\rightarrow$ (de-)approach of already coordinated ligands observed six times.
- (c) $\Delta\text{CN} = 0$, MAE large $\rightarrow$ ligand change observed once.
- (d) $\Delta\text{CN} \neq 0$, MAE large $\rightarrow$ (de-)coordination of ligands observed four times.

Furthermore, eight mixed cases are observed. For example during optimization of the transcription protein 4CUE, the backbone oxygen of glycine (ligand 160) approaches the metal by 3 Å (case d). In addition, one oxygen of glutamic acid (ligand 31) decoordinates while the other oxygen of the same glutamic acid (ligand 32) coordinates to the metal (case c).

The above classification is somewhat arbitrary, and every protein structure optimization has to be analyzed individually. Still, on the basis of the presented criteria, GFN2-xTB provides accurate structures for metalloproteins with up to about 5000 atoms. This good performance not only is true for the secondary structure of the metalloprotein but also holds for the respective metal center and their ligand sphere. Note that the average time for fully optimizing the metalloproteins amounts to 26 h per system on four CPU cores (on average 4 min per step). The

largest protein in the test set consists of 4804 atoms and took about 3.2 days for the full optimization.

CONCLUSION

We have evaluated the newly developed semiempirical GFN2-xTB method for equilibrium structure determination of organic proteins and metalloproteins. Experimental crystal structures were taken as reference. Various geometrical descriptors and RMSDs were used to assess the quality of GFN2-xTB optimized structures also in comparison to competitor methods. Due to the limitations of the available FF parametrization, we split the test set into two parts. The first test set comprises 70 peptides and proteins that only include the elements H, C, N, O, and S. The second test set contains 20 metalloproteins covering a wide range of biochemical functions.

We investigated the difference between the molecular structure of molecular protein structure and as a solid by optimizing 10 small peptides with sHF-3c with and without applying periodic boundary conditions. The C_α RMSDs of all optimized structures are very close to the estimated apparent experimental uncertainty of 0.5 Å,⁷⁴ validating the molecular approach used for all subsequent calculations.

The proteins of the first test set were optimized using GFN n -xTB ($n = 1,2$), HF-3c, as well as the two special purpose FFs OPLS2005 and AMBER*. The optimized structures were compared to the experimental crystal structures employing several geometrical descriptors. Regarding the deviations of the angles from the experimental values, GFN2-xTB yields similar results as the specialized force field OPLS2005. On average, all other tested methods (HF-3c, AMBER*, and GFN1-xTB) deviate more from the crystal structure than GFN2-xTB for this property. The same conclusions were drawn for the C_α and heavy atoms RMSD. This is in line with the conclusions drawn by Zheng et al.⁴¹ that GFN2-xTB is a promising method for the structural optimization of proteins. In general, secondary structure motifs are well conserved with all methods. Furthermore, we demonstrated that GFN2-xTB could reproduce the experimental bond lengths distributions. These distributions as provided by the FFs are, however, too narrow probably due to the neglect of QM (many-body) effects. Furthermore, GFN2-xTB is significantly faster than HF-3c so that large-scale studies of proteins are possible. In summary, we conclude that the general purpose method GFN2-xTB yields similar results as the specialized force field OPLS2005 and even better results than HF-3c and AMBER* for equilibrium structures of proteins without any prosthetic groups.

For proteins including very common prosthetic groups or pure polypeptides, most biochemical FFs give fast and accurate results. However, investigations for other systems, for which no parametrization is available, are challenging. Therefore, metalloproteins with common and uncommon metal centers are included in the second test set and investigated. Because HF-3c is known to perform poorly for many metal-containing molecules and for the FFs the necessary empirical parameters are missing, proteins of the metalloprotein test set were optimized only with the GFN n -xTB ($n = 1,2$) methods.

The deviations of the GFN n -xTB optimized structures from the experimental crystal structures were similarly small as for the first test set. Furthermore, we analyzed the coordination geometry of the metal centers with respect to the corresponding X-ray structure. Also for the local structure of the metal coordination, GFN2-xTB mostly provides a good agreement

with the reference structures and presented its potential for these challenging systems.

In conclusion, GFN2-xTB is not intended to replace specialized FFs for the investigation of biochemical systems but rather to offer a robust and efficient black-box SQM alternative for the structural optimization of systems that cannot be treated with FFs. Furthermore, as already investigated by Zheng et al.⁴¹ for a small set of proteins, GFN2-xTB is a promising method for the quantum refinement of protein structures, which should be further investigated. Even the largest (metallo-)proteins with up to 5000 atoms were fully optimized with GFN2-xTB in a few days on a small work-station or laptop computer. Due to its low computational cost for a QM method, entirely new possibilities in computational protein research are accessible with GFN2-xTB.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/acs.jpcb.0c00549>.

Cartesian coordinates (xyz) of the experimental structures (start structures) of the first test set (ZIP)

Cartesian coordinates (xyz) of the GFN2-xTB optimized structures of the first test set (ZIP)

Cartesian coordinates (xyz) of the GFN2-xTB optimized structures of the second test set (ZIP)

Cartesian coordinates (xyz) of the experimental structures (start structures) of the second test set (ZIP)

Additional information about the test sets, the periodic optimizations, results for GFN1-xTB, and the coordination sphere around metal centers (PDF)

AUTHOR INFORMATION

Corresponding Authors

Andreas Hansen – Mulliken Center for Theoretical Chemistry, Institut für Physikalische und Theoretische Chemie der Universität Bonn, D-53115 Bonn, Germany; orcid.org/0000-0003-1659-8206; Phone: +49 (0)228 7360431; Email: hansen@thch.uni-bonn.de

Stefan Grimme – Mulliken Center for Theoretical Chemistry, Institut für Physikalische und Theoretische Chemie der Universität Bonn, D-53115 Bonn, Germany; orcid.org/0000-0002-5844-4371; Phone: +49 (0)228 732351; Email: grimme@thch.uni-bonn.de

Authors

Sarah Schmitz – Mulliken Center for Theoretical Chemistry, Institut für Physikalische und Theoretische Chemie der Universität Bonn, D-53115 Bonn, Germany

Jakob Seibert – Mulliken Center for Theoretical Chemistry, Institut für Physikalische und Theoretische Chemie der Universität Bonn, D-53115 Bonn, Germany; orcid.org/0000-0002-3163-8627

Katja Ostermeir – Mulliken Center for Theoretical Chemistry, Institut für Physikalische und Theoretische Chemie der Universität Bonn, D-53115 Bonn, Germany

Andreas H. Göller – Computational Molecular Design, Pharmaceuticals, R&D, Bayer AG, D-42096 Wuppertal, Germany; orcid.org/0000-0003-4343-4063

Complete contact information is available at: <https://pubs.acs.org/10.1021/acs.jpcb.0c00549>

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

The authors thank Jan Michael Mewes, Sebastian Dohm, Jan-Gerit Brandenburg, Stephan Ehrlich, and Gregor Hagelücken for support and fruitful discussions. This work was supported by the Bayer AG and the DFG in the framework of the Gottfried-Wilhelm-Leibniz prize.

REFERENCES

- (1) March, N.; Matthai, C. C. The Application of Quantum Chemistry and Condensed Matter Theory in studying Amino-Acids, Protein Folding and Anticancer Drug Technology. *Theor. Chem. Acc.* **2010**, *120*, 193–201.
- (2) Shurki, A.; Warshel, A. Structure/Function Correlations of Proteins using MM, QM/MM, and Related Approaches: Methods, Concepts, Pitfalls, and Current Progress. *Adv. Protein Chem.* **2003**, *66*, 249–313.
- (3) Zhang, D. W.; Zhang, J. Molecular Fractionation with Conjugate Caps for full Quantum Mechanical Calculation of Protein–Molecule Interaction Energy. *J. Chem. Phys.* **2003**, *119*, 3599–3605.
- (4) Wada, M.; Sakurai, M. A. Quantum Chemical Method for Rapid Optimization of Protein Structures. *J. Comput. Chem.* **2005**, *26*, 160–168.
- (5) Merz, K. M. Using Quantum Mechanical Approaches to study Biological Systems. *Acc. Chem. Res.* **2014**, *47*, 2804–2811 (PMID: 25099338).
- (6) Canfield, P.; Dahlbom, M. G.; Hush, N. S.; Reimers, J. R. Density-Functional Geometry Optimization of the 150000-atom photosystem-I trimer. *J. Chem. Phys.* **2006**, *124*, 024301.
- (7) Senn, H. M.; Thiel, W. *Atomistic Approaches in Modern Biology*; Springer, 2006; pp 173–290.
- (8) Fanfrlík, J.; Brynda, J.; Řezáč, J.; Hobza, P.; Lepšík, M. Interpretation of Protein/Ligand Crystal Structure using QM/MM Calculations: Case of HIV-1 Protease/Metallacarborane Complex. *J. Phys. Chem. B* **2008**, *112*, 15094–15102 (PMID: 18975888).
- (9) Sproviero, E. M.; Newcomer, M. B.; Gascón, J. A.; Batista, E. R.; Brudvig, G. W.; Batista, V. S. The MoD-QM/MM methodology for Structural Refinement of Photosystem II and other biological Macromolecules. *Photosynth. Res.* **2009**, *102*, 455–470.
- (10) Steinbrecher, T.; Elstner, M. QM and QM/MM Simulations of Proteins. *Methods Mol. Biol.* **2013**, *924*, 91–124.
- (11) Van Alsenoy, C.; Yu, C.-H.; Peeters, A.; Martin, J. M. L.; Schäfer, L. Ab Initio Geometry Determinations of Proteins. 1. Crambin. *J. Phys. Chem. A* **1998**, *102*, 2246–2251.
- (12) Kulik, H. J.; Luehr, N.; Ufimtsev, I. S.; Martinez, T. J. Ab Initio Quantum Chemistry for Protein Structures. *J. Phys. Chem. B* **2012**, *116*, 12501–12509 (PMID: 22974088).
- (13) Sure, R.; Grimme, S. Corrected small Basis Set Hartree-Fock method for large Systems. *J. Comput. Chem.* **2013**, *34*, 1672–1685.
- (14) Christensen, A. S.; Kubař, T.; Cui, Q.; Elstner, M. Semiempirical Quantum Mechanical Methods for Noncovalent Interactions for Chemical and Biochemical Applications. *Chem. Rev.* **2016**, *116*, 5301–5337 (PMID: 27074247).
- (15) Elstner, M.; Frauenheim, T.; Kaxiras, E.; Seifert, G.; Suhai, S. A self-consistent Charge Density-Functional based Tight-Binding Scheme for large Biomolecules. *Phys. Status Solidi B* **2000**, *217*, 357–376.
- (16) Elstner, M.; Cui, Q.; Muni, P.; Kaxiras, E.; Frauenheim, T.; Karplus, M. Modeling Zinc in Biomolecules with the Self Consistent Charge-Density Functional Tight Binding (SCC-DFTB) Method: Applications to Structural and Energetic Analysis. *J. Comput. Chem.* **2003**, *24*, 565–581.
- (17) Elstner, M. The SCC-DFTB method and its Application to Biological Systems. *Theor. Chem. Acc.* **2006**, *116*, 316–325.
- (18) Stewart, J. J. Calculation of the Geometry of a Small Protein using Semiempirical Methods. *J. Mol. Struct.: THEOCHEM* **1997**, *401*, 195–205 (The Future of Semiempirical Methods).
- (19) Dobeš, P.; Řezáč, J.; Fanfrlík, J.; Otyepka, M.; Hobza, P. Semiempirical Quantum Mechanical Method PM6-DH2X describes the Geometry and Energetics of CK2-Inhibitor Complexes involving Halogen Bonds Well, while the Empirical Potential fails. *J. Phys. Chem. B* **2011**, *115*, 8581–8589 (PMID: 21648479).
- (20) Stewart, J. J. P. Application of the PM6 Method to Modeling Proteins. *J. Mol. Model.* **2009**, *15*, 765–805.
- (21) Martin, B. P.; Brandon, C. J.; Stewart, J. J. P.; Braun-Sand, S. B. Accuracy Issues involved in Modeling in vivo Protein Structures using PM7. *Proteins: Struct., Funct., Genet.* **2015**, *83*, 1427–1435.
- (22) Lazaridis, T.; Karplus, M. Effective Energy Functions for Protein Structure Prediction. *Curr. Opin. Struct. Biol.* **2000**, *10*, 139–145.
- (23) Rulišek, L.; Havlas, Z. Theoretical Studies of Metal Ion Selectivity. † 2. DFT Calculations of Complexation Energies of Selected Transition Metal Ions (Co²⁺, Ni²⁺, Cu²⁺, Zn²⁺, Cd²⁺, and Hg²⁺) in Metal-Binding Sites of Metalloproteins. *J. Phys. Chem. A* **2002**, *106*, 3855–3866.
- (24) Sinnecker, S.; Neese, F. QM/MM calculations with DFT for taking into account Protein Effects on the EPR and optical Spectra of Metalloproteins. Plastocyanin as a case study. *J. Comput. Chem.* **2006**, *27*, 1463–1475.
- (25) Schmiedekamp, A.; Nanda, V. Metal-activated Histidine Carbon Donor Hydrogen Bonds contribute to Metalloprotein Folding and Function. *J. Inorg. Biochem.* **2009**, *103*, 1054–1060.
- (26) Ehrlich, S.; Göller, A. H.; Grimme, S. Towards full Quantum-Mechanics-based Protein-Ligand Binding Affinities. *ChemPhysChem* **2017**, *18*, 898–905.
- (27) Ryde, U.; Söderhjelm, P. Ligand-Binding Affinity Estimates Supported by Quantum-Mechanical Methods. *Chem. Rev.* **2016**, *116*, 5520–5566.
- (28) Grimme, S.; Bannwarth, C.; Shushkov, P. A Robust and Accurate Tight-Binding Quantum Chemical Method for Structures, Vibrational Frequencies, and Noncovalent Interactions of Large Molecular Systems Parametrized for All spd-Block Elements (Z = 1 – 86). *J. Chem. Theory Comput.* **2017**, *13*, 1989–2009.
- (29) Bannwarth, C.; Ehlert, S.; Grimme, S. GFN2-xTB—An Accurate and Broadly Parametrized Self-Consistent Tight-Binding Quantum Chemical Method with Multipole Electrostatics and Density-Dependent Dispersion Contributions. *J. Chem. Theory Comput.* **2019**, *15*, 1652–1671 (PMID: 30741547).
- (30) *Semiempirical extended tight-binding program package xtb*. <https://github.com/grimme-lab/xtb> (accessed 2020-04-08).
- (31) 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.
- (32) Grimme, S.; Ehrlich, S.; Goerigk, L. Effect of the Damping Function in Dispersion Corrected Density Functional Theory. *J. Comput. Chem.* **2011**, *32*, 1456–1465.
- (33) Caldeweyher, E.; Ehlert, S.; Hansen, A.; Neugebauer, H.; Spicher, S.; Bannwarth, C.; Grimme, S. A generally applicable Atomic-Charge dependent London Dispersion Correction. *J. Chem. Phys.* **2019**, *150*, 154122.
- (34) Bursch, M.; Caldeweyher, E.; Hansen, A.; Neugebauer, H.; Ehlert, S.; Grimme, S. Understanding and Quantifying London Dispersion Effects in Organometallic Complexes. *Acc. Chem. Res.* **2019**, *52*, 258–266.
- (35) Seibert, J.; Bannwarth, C.; Grimme, S. Biomolecular Structure Information from high-speed Quantum Mechanical Electronic Spectra Calculation. *J. Am. Chem. Soc.* **2017**, *139*, 11682–11685.
- (36) Bursch, M.; Hansen, A.; Grimme, S. Fast and Reasonable Geometry Optimization of Lanthanoid Complexes with an Extended Tight Binding Quantum Chemical Method. *Inorg. Chem.* **2017**, *56*, 12485–12491.

- (37) Pracht, P.; Bauer, C. A.; Grimme, S. Automated and Efficient Quantum Chemical Determination and Energetic Ranking of Molecular Protonation Sites. *J. Comput. Chem.* **2017**, *38*, 2618–2631.
- (38) Kruse, H.; Grimme, S. A Geometrical Correction for the Inter- and Intra-molecular Basis Set Superposition Error in Hartree-Fock and Density Functional Theory Calculations for Large Systems. *J. Chem. Phys.* **2012**, *136*, 154101.
- (39) Goerigk, L.; Reimers, J. R. Efficient Methods for the Quantum Chemical Treatment of Protein Structures: The Effects of London-Dispersion and Basis-Set Incompleteness on Peptide and Water-Cluster Geometries. *J. Chem. Theory Comput.* **2013**, *9*, 3240–3251 (PMID: 26583999).
- (40) Dietschreit, J. C. B.; Peters, L. D. M.; Kussmann, J.; Ochsenfeld, C. Identifying Free Energy Hot-Spots in Molecular Transformations. *J. Phys. Chem. A* **2019**, *123*, 2163–2170.
- (41) Zheng, M.; Biczysko, M.; Xu, Y.; Moriarty, N. W.; Kruse, H.; Urzhumtsev, A.; Waller, M. P.; Afonine, P. V. Including crystallographic Symmetry in quantum-based Refinement: Q|R#2. *Acta Crystallogr., Sect. D* **2020**, *76*, 41–50.
- (42) Banks, J. L.; Beard, H. S.; Cao, Y.; Cho, A. E.; Damm, W.; Farid, R.; Felts, A. K.; Halgren, T. A.; Mainz, D. T.; Maple, J. R.; et al. Integrated Modeling Program, Applied Chemical Theory (IMPACT). *J. Comput. Chem.* **2005**, *26*, 1752–1780.
- (43) McDonald, D.; Still, W. AMBER Torsional Parameters for the Peptide Backbone. *Tetrahedron Lett.* **1992**, *33*, 7743–7746 (The International Journal for the Rapid Publication of Preliminary).
- (44) Ferguson, D. M.; Kollman, P. A. Can the Lennard-Jones 6-12 function replace the 10-12 form in Molecular Mechanics Calculations? *J. Comput. Chem.* **1991**, *12*, 620–626.
- (45) Delle Piane, M.; Corno, M.; Orlando, R.; Dovesi, R.; Ugliengo, P. Elucidating the Fundamental Forces in Protein Crystal Formation: the Case of Crambin. *Chem. Sci.* **2016**, *7*, 1496–1507.
- (46) Caldeweyher, E.; Gerit Brandenburg, J. Simplified DFT Methods for Consistent Structures and Energies of Large Systems. *J. Phys.: Condens. Matter* **2018**, *30*, 213001.
- (47) Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. The Protein Data Bank. *Nucleic Acids Res.* **2000**, *28*, 235–242.
- (48) Madhavi Sastry, G.; Adzhigirey, M.; Day, T.; Annabhimoju, R.; Sherman, W. Protein and Ligand Preparation: Parameters, Protocols, and Influence on Virtual Screening Enrichments. *J. Comput.-Aided Mol. Des.* **2013**, *27*, 221–234.
- (49) Bas, D. C.; Rogers, D. M.; Jensen, J. H. Very fast Prediction and Rationalization of pK_a Values for Protein-Ligand Complexes. *Proteins: Struct., Funct., Genet.* **2008**, *73*, 765–783.
- (50) Shelley, J. C.; Cholleti, A.; Frye, L. L.; Greenwood, J. R.; Timlin, M. R.; Uchimaya, M. Epik: a Software Program for pK_a Prediction and Protonation State Generation for Drug-like Molecules. *J. Comput.-Aided Mol. Des.* **2007**, *21*, 681–691.
- (51) Greenwood, J. R.; Calkins, D.; Sullivan, A. P.; Shelley, J. C. Towards the Comprehensive, Rapid, and Accurate Prediction of the Favorable Tautomeric States of Drug-like Molecules in Aqueous Solution. *J. Comput.-Aided Mol. Des.* **2010**, *24*, 591–604.
- (52) Xue, Y.; Oekvist, M.; Hansson, O.; Young, S. Crystal Structure of Spinach Plastocyanin at 1.7 Å Resolution. *Protein Sci.* **1998**, *7*, 2099–2105.
- (53) Grimme, S.; Hansen, A. A practicable Real-Space Measure and Visualization of Static Electron-Correlation Effects. *Angew. Chem., Int. Ed.* **2015**, *54*, 12308–12313.
- (54) Bauer, C. A.; Hansen, A.; Grimme, S. The Fractional Occupation Number Weighted Density as a Versatile Analysis Tool for Molecules with a Complicated Electronic Structure. *Chem. - Eur. J.* **2017**, *23*, 6150–6164.
- (55) Lindh, R.; Bernhardsson, A.; Karlström, G.; Malmqvist, P.-Å. On the use of a Hessian model function in molecular Geometry Optimizations. *Chem. Phys. Lett.* **1995**, *241*, 423–428.
- (56) Klamt, A.; Diedenhofen, M. Calculation of Solvation Free Energies with DCOSMO-RS. *J. Phys. Chem. A* **2015**, *119*, 5439–5445.
- (57) Furche, F.; Ahlrichs, R.; Hättig, C.; Klopper, W.; Sierka, M.; Weigend, F. Turbomole. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2014**, *4*, 91–100.
- (58) TURBOMOLE V7.2 2017, a development of University of Karlsruhe and Forschungszentrum Karlsruhe GmbH, 1989–2007, TURBOMOLE GmbH, since 2007. <http://www.turbomole.com> (accessed April 8, 2020).
- (59) Rezáč, J.; Hobza, P. Advanced Corrections of Hydrogen Bonding and Dispersion for Semiempirical Quantum Mechanical Methods. *J. Chem. Theory Comput.* **2012**, *8*, 141–151 (PMID: 26592877).
- (60) Rezáč, J.; Hobza, P. A Halogen-bonding Correction for the Semiempirical PM6Method. *Chem. Phys. Lett.* **2011**, *506*, 286–289.
- (61) Stewart, J. J. Optimization of Parameters for semiempirical Methods VI: More modifications to the NDDO approximations and re-optimization of Parameters. *J. Mol. Model.* **2013**, *19*, 1–32.
- (62) James, J.; Stewart, P. MOPAC2016, version 19.179L. Stewart Computational Chemistry, <http://OpenMOPAC.net> (accessed April 8, 2020).
- (63) Klamt, A.; Schüürmann, G. COSMO: A new approach to dielectric Screening in Solvents with explicit expressions for the screening Energy and its Gradient. *J. Chem. Soc., Perkin Trans. 2* **1993**, *2*, 799–805.
- (64) Dovesi, R.; Erba, A.; Orlando, R.; Zicovich-Wilson, C. M.; Civalieri, B.; Maschio, L.; Rérat, M.; Casassa, S.; Baima, J.; Salustro, S.; et al. Quantum-mechanical Condensed Matter Simulations with CRYSTAL. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2018**, *8*, No. e1360.
- (65) Doll, K.; Saunders, V. R.; Harrison, N. M. Analytical Hartree-Fock Gradients for Periodic Systems. *Int. J. Quantum Chem.* **2001**, *82*, 1–13.
- (66) Civalieri, B.; D'Arco, P.; Orlando, R.; Saunders, V.; Dovesi, R. Hartree-Fock Geometry Optimisation of Periodic Systems with the CRYSTAL code. *Chem. Phys. Lett.* **2001**, *348*, 131–138.
- (67) Cutini, M.; Civalieri, B.; Corno, M.; Orlando, R.; Brandenburg, J. G.; Maschio, L.; Ugliengo, P. Assessment of Different Quantum Mechanical Methods for the Prediction of Structure and Cohesive Energy of Molecular Crystals. *J. Chem. Theory Comput.* **2016**, *12*, 3340–3352.
- (68) Brandenburg, J. G.; Grimme, S. *Prediction and Calculation of Crystal Structures*; Springer, 2013; pp 1–23.
- (69) Coutias, E. A.; Seok, C.; Dill, K. A. Using Quaternions to calculate RMSD. *J. Comput. Chem.* **2004**, *25*, 1849–1857.
- (70) Larsen, A. H.; Mortensen, J. J.; Blomqvist, J.; Castelli, I. E.; Christensen, R.; Dulak, M.; Friis, J.; Groves, M. N.; Hammer, B.; Hargus, C.; et al. The Atomic Simulation Environment—a Python library for Working with Atoms. *J. Phys.: Condens. Matter* **2017**, *29*, 273002.
- (71) Wiberg, K. Application of the pople-santry-segal CNDO Method to the Cyclopropylcarbinyl and Cyclobutyl Cation and to Bicyclobutane. *Tetrahedron* **1968**, *24*, 1083–1096.
- (72) Mayer, I. On Bond Orders and Valences in the Ab Initio Quantum Chemical Theory. *Int. J. Quantum Chem.* **1986**, *29*, 73–84.
- (73) Jacobson, M. P.; Friesner, R. A.; Xiang, Z.; Honig, B. On the role of the Crystal Environment in determining Protein Side-Chain Conformations. *J. Mol. Biol.* **2002**, *320*, 597–608.
- (74) Edelman, M.; Potapov, V.; Sobolev, V.; Gerzon, S.; Eyal, E. The Limit of Accuracy of Protein Modeling: Influence of Crystal Packing on Protein Structure. *J. Mol. Biol.* **2005**, *351*, 431–442.
- (75) Hinsen, K. Structural Flexibility in Proteins: Impact of the Crystal Environment. *Bioinformatics* **2008**, *24*, 521–528.
- (76) Rapp, C. S.; Pollack, R. M. Crystal Packing Effects on Protein Loops. *Proteins: Struct., Funct., Genet.* **2005**, *60*, 103–109.
- (77) Flores, T.; Orengo, C.; Moss, D.; Thornton, J. Comparison of Conformational Characteristics in Structurally Similar Protein Pairs. *Protein Sci.* **1993**, *2*, 1811–1826.

All-Atom Quantum Mechanical (AQM) Methodologies for One- and Two-Photon Absorption of Realistic Systems

Sarah Löffelsender¹, Marilù G. Maraldi², and Marc de Wergifosse²

Received: December 16, 2024

First published: April 08, 2025

Reprinted with permission from[‡]:

S. Löffelsender, M. G. Maraldi, and M. de Wergifosse, ChemPhysChem 2401121, DOI: 10.1002/cphc.202401121

Copyright © 2025 Wiley-VCH GmbH

Own manuscript contributions:

- Data curating
- Preparation of calculations
- Performing all calculations
- Evaluation and interpretation of the results
- Visualization of most results
- Writing and revising the manuscript

¹ Mulliken Center for Theoretical Chemistry, Clausius Institute for Physical and Theoretical Chemistry, Rheinische Friedrich-Wilhelms-Universität Bonn, Berlingstr. 4, 53115 Bonn, Germany

² Institute of Condensed Matter and Nanosciences, Université Catholique de Louvain, B-1348 Louvain-la-Neuve, Belgium

[‡] Permission requests to reuse material from this chapter should be directed to Wiley-VCH GmbH

All-Atom Quantum Mechanical Methodologies for One- and Two-Photon Absorption of Realistic Systems

Sarah Löffelsender, Marilù G. Maraldi, and Marc de Wergifosse*

All-atom quantum mechanics (AQM) methodologies are assessed to evaluate one- and two-photon absorption (1PA and 2PA) of realistic systems. All-atom single structure QM (ASQM) and dynamic structure QM (ADQM) methodologies are discussed. These workflows are possible thanks to developments in simplified quantum chemistry methods and in particular with both sTD-DFT-xTB and dt-sTD-DFT-xTB schemes. The ASQM scheme is tested to compute the 1- and 2PA of two proteins: bacteriorhodopsin and iLOV. Results show that the ASQM methodology is able to describe higher-energy transitions involving π -conjugated

amino acids such as tryptophan or tyrosine. Then, two variants of the ADQM workflow are evaluated to reproduce the 1- and 2PA of the flavin mononucleotide in aqueous solution, involving either Boltzmann ensemble of conformers in implicit solvent (ADQM-Boltz.) or snapshots of molecular dynamics of explicitly solvated systems (ADQM-MD). Spectra computed with the ADQM-MD approach provide striking comparisons with respect to experiment, while the ADQM-Boltz. approach provides little change with respect to the ASQM workflow.

1. Introduction

The direct environment of a molecule modifies its ground and excited state properties including geometry, energy, dynamics, and reactivity.^[1] For example, the theoretical description of solvatochromism should involve a realistic representation of surrounding molecules around the solute to faithfully reproduce one- and two-photon absorption (1PA and 2PA) experimental spectra.^[1,2] For intermolecular interactions such as hydrogen bonds, π -stacking, or halogen bonds, an explicit treatment of the environment should naturally be included into the methodology.^[2] The computational modeling of highly flexible systems, for example, molecules in solution, necessitates the integration of dynamic structural effects.^[1,2] Such treatment can be computationally prohibitive for large molecules or molecular assemblies.^[3] **Figure 1** shows two currently established workflows^[1,2] to compute excited states of molecules in solution that include dynamic structural effects. In the first workflow (left side), the impact of the environment is modeled implicitly using a solvent model applied to i) an isolated molecule or ii) a cluster composed of the chromophore and few explicit solvent molecules that includes microsolvation interactions while keeping a moderate computational cost. Dynamic structural

effects are captured implicitly by sampling the conformational space, and Boltzmann averaged spectra are computed. The second workflow (right side of Figure 1) accounts explicitly for direct surroundings of the chromophore. Molecular dynamics (MD) simulations are performed to compute solute-solvent (chromophore-environment) configurations and account for temperature effects.^[1,2,4,5] These simulations are usually conducted using classical, Monte-Carlo, or quantum mechanical/molecular mechanical (QM/MM) MD simulations.^[2,6] For each snapshot, excited states are typically computed using a QM/MM framework (reducing the number of explicitly treated atoms drastically). An averaged spectra incorporating each snapshot is then obtained. Note that to account for polarization effects and to converge excited state properties,^[7] some authors suggested to include >250 atoms in the QM region^[8] or at least two solvation shells.^[1] This places a significant computational demand on the excited state calculation. In both workflows, two main steps can be distinguished: i) sampling geometries and ii) characterizing excited states with the environment accounted for either implicitly or explicitly.

In 1976, Warshel and Levitt^[9] introduced the QM/MM method to treat large molecular systems in a consistent way. For the development of such multiscale models, they won the Nobel Prize in Chemistry in 2013 with Karplus. In the QM/MM approach, the system considered responsible for the target property is computed quantum mechanically while surroundings are treated classically. These schemes enable the treatment of explicitly solvated systems as well as chromophores embedded in more complex matrices such as proteins. Nowadays, a plethora of schemes exists for the MM part,^[2,6,10–22] involving mechanical embedding (ME), electrostatic embedding (EE), or polarizable embedding (PE). They rely on a classical description of the solute-solvent (or chromophore-environment) electrostatic interactions, either neglecting (ME, EE) or accounting for (PE) important mutual polarizations between the solute and the environment. Nevertheless, solute-solvent (or chromophore-environment)

S. Löffelsender

Mulliken Center for Theoretical Chemistry
–Clausius Institut für Physikalische und Theoretische Chemie
University of Bonn
Beringstr. 4, 53115 Bonn, Germany

M. G. Maraldi, M. de Wergifosse
Theoretical Chemistry Group, MolecularChemistry
–Materials and Catalysis Division (MOST)
–Institute of Condensed Matter and Nanosciences
Université Catholique de Louvain
B-1348 Louvain-la-Neuve, Belgium
E-mail: marc.dewergifosse@uclouvain.be



Supporting information for this article is available on the WWW under <https://doi.org/10.1002/cphc.202401121>

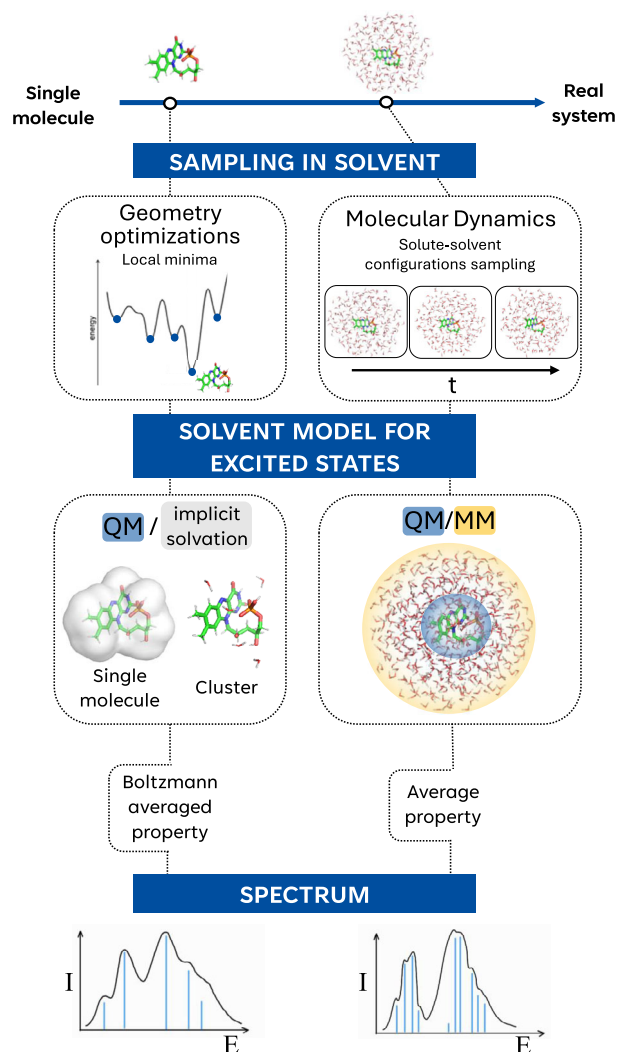


Figure 1. Two current workflows to compute excited state spectra in solution or complex environments that account for dynamic structural effects.

interactions have a quantum nature, which are particularly difficult to describe with QM/MM models.^[6] Higher-level approaches also exist such as the QM/effective fragment potential (EFP)^[23–27] or quantum embedding schemes.^[28–30]

2PA is a nonlinear optical (NLO) effect that allows to use lower energy light sources with respect to 1PA. Both incident photons have a lower energy than the excited state transition energy (half of it, if degenerate). It exhibits an improved spatial resolution and an increased penetration depth (up to 1 nm, compared to 50–80 μm in normal confocal laser scanning microscopy) as well as reduced photobleaching. These characteristics motivated the use of 2PA as a noninvasive imaging technique for numerous medical fields, including cancer^[31–33] or kidney researches,^[34–36] neuroscience,^[37] and even to probe neurological functions of (fixed or freely moving) living animals.^[38–42] 2PA is also used in data storage devices^[43,44] and information processing.^[45] These areas of research could benefit from the design of new systems with enhanced 2PA cross sections ($\sigma^{2\text{PA}}$).

Calculating 2PA spectra is a challenging task as it formally involves to evaluate a third-order molecular response

property.^[46,47] Computationally efficient expressions to evaluate $\sigma^{2\text{PA}}$ can be derived from residue analysis of the cubic response function.^[46,48] For the degenerate case, 2PA strengths can also be extracted from the residue of the quadratic response function.^[46] Since the seminal work of Olsen and Jørgensen^[46] in 1985, hierarchy of methods were implemented along the years including couple cluster (CC) and density functional theory (DFT) methods. Nanda and Krylov^[49] implemented one of the most formally accurate scheme available with the equation of motion coupled cluster singles and doubles (EOM-CCSD) method. Hattig et al.^[48,50,51] implemented CCs, CCSD, CC3, CC2 and RI-CC2 methods to evaluate 2PA. Cost-efficient (RI)-CC2 schemes are often used as a reference to evaluate 2PA strengths.^[4,52–59] CC methods are still computationally demanding and are only available to treat rather small systems (<100 atoms). Alternatively, time-dependent DFT methods^[60–62] were implemented for medium-sized systems.

QM/MM schemes were often used to evaluate 2PA of chromophores in solution or complex matrices. Model systems of explicitly solvated chromophores were employed to characterize the effect of the solvent on 2PA spectral features,^[63–67] and two-photon solvatochromism,^[68] as well as to assess dynamic structural effects.^[69] Some of them provided computational results directly comparable to experiments.^[65,66,68] For example, Olesiak-Banska et al.^[65] studied the 2PA of the Hoechst chromophore in water solution showing the emergence of a low-energy band in the 2PA spectra due to dimerization of the dye in solution. The QM/MM family of methods was also used to evaluate the 2PA of chromophores in protein environments.^[70–72] For channel rhodopsin mutants^[73] and green fluorescent proteins,^[72] design guidelines based on QM/MM studies were proposed to enhance their $\sigma^{2\text{PA}}$. Murugan and Zalesny^[71] used a CC2/MM approach to model the 2PA properties of Parkinson's diagnostic probes targeting monoamine oxidase B. To improve the modeling of theoretical 2PA spectra, Kongsted and coworkers^[74] showed for DNA intercalators the importance to include non-electrostatic repulsion into the embedding scheme, reducing the electron density leakage from the QM to the MM region. Olsen et al.^[75] showed that very accurate explicitly polarizable embedding potentials for proteins can be efficiently designed using system fragmentation strategies for QM/MM molecular response calculations. Excited state properties can also be highly sensitive to chromophore geometries.^[2,76–79] For example, Palczewska et al.^[79] examined the impact of distorting bovine rhodopsin on 2PA using QM/MM (QM, chromophore; MM, residue charges within a 5 Å distance). They demonstrated that the deformation of the chromophore as well as surrounding atoms and their charges have significant impacts on the 2PA spectrum, complementing the study of Valsson et al.^[8] on 1PA.

As already stressed above, the nature of interactions between the chromophore and its direct environment has a quantum nature. De facto, such interactions are particularly difficult to be described with QM/MM models especially if important long-range interactions or direct H-bonds between layers occur.^[6] A significant limitation associated with such schemes is to interface both QM and MM regions, although several protocols exist.^[13] A viable alternative to QM/MM schemes for the

computation of excited states is to directly treat most of the system in a QM fashion while remaining parts are accounted for by a solvent model. This is the so-called “brute force approach” as termed by Giovannini et al.^[2] for which we prefer the terminology “all-atom QM (AQM) methodology.” To compute excited states and response properties of large systems, an AQM methodology containing explicitly up to several thousands of atoms is now possible thanks to recent developments in the field of simplified quantum chemistry methods.^[80–91] The realm of simplified methods was first introduced by Grimme^[80] in 2013 with the simplified time-dependent DFT (sTD-DFT) variant using the Tamm–Dancoff approximation (sTDA). Shortly after, this was extended to the sTD-DFT method by Bannwarth and Grimme.^[81] To compute excited states of ultra large systems (up to 5000 atoms), an approach called sTD-DFT-xTB combines an extended tight binding (xTB) ground state^[82] with the sTD-DFT method. The computation of response properties with the sTD-DFT scheme was introduced by de Wergifosse et al. including the polarizability,^[83] optical rotation,^[86] excited state absorption,^[84] first hyperpolarizability,^[83] and the ultra-fast evaluation of 2PA.^[90] Among applications, Seibert et al.^[87] assessed dynamic structural effects on the first hyperpolarizability of very flexible tryptophan-rich peptides as well as the gramicidin A; Beaujean et al.^[88] proposed an AQM methodology to compute the first hyperpolarizability of two fluorescent proteins (FPs) iLOV and the bacteriorhodopsin (bR), showing striking agreement with respect to experiment. The dual-threshold sTD-DFT (dt-sTD-DFT) was introduced to reduce the computational cost of nonlinear response calculations by truncating in a different manner the space of singly excited configurations for the chromophore and the remaining of the protein.^[88] Last year, one of us with the help of the eXact integral sTD-DFT (XsTD-DFT)^[91,92] using the CAM-B3LYP range-separated hybrid functional showed that to reproduce experimental 1PA and circular dichroism spectra of the photoactive yellow protein (PYP), an AQM methodology is absolutely necessary because one of the main peak is due to a local $\pi \rightarrow \pi^*$ transition on a tryptophan and thus not involving the chromophore.^[93]

The bacteriorhodopsin (3835 atoms) is a transmembrane protein of *Halobacterium salinarum*.^[94–97] It is a light-driven proton pump that facilitates the transfer of protons from the external to the internal cellular environment. Prior studies have demonstrated that the chromophore, a retinal, interacts considerably with nearby residues (Trp86 and Tyr185) upon excitation,^[98–100] resulting in a remarkable high molar extinction coefficient.^[101] Moreover, reports indicate an exceptionally high 2PA cross section,^[102] with possible applications in data storage.^[43,44] bR undergoes photobleaching upon 2P excitation, due to the accessibility of photoproducts beyond the classical photocycle.^[103,104] Note that the retinal chromophore from bR is a different molecule with respect to the well-studied retinal vitamin A from the human eye.

iLOV is an improved variant of the light, oxygen, or voltage domain of the plant blue light receptor.^[105] This small FP (1849 atoms) originally engineered to monitor virus infection and spread owes its photophysical properties to its noncovalently bound chromophore flavin mononucleotide (FMN). Theoretical and experimental studies^[4,76–78,106–113] on FMN established that its surroundings influence significantly its excited state

properties. **Figure 2** presents experimental 1PA spectra of FMN in different environments (water, methanol, and two similar proteins iLOV and miniSOG), showing the sensitivity of FMN to its environment. At a first glance, trends among different types of surroundings can be rationalized in terms of the polarity of the environment. For 1PA spectra, flavins were found to be sensitive to the planarity of the isoalloxazine ring.^[76,111,113] Already subtle distortions, that is, the well-known “butterfly” bending of the ring system, causes substantial changes in the spectra,^[76,110,111] for example, shifting of the excitation energy. This was attributed to the mixing of $n \rightarrow \pi^*$ and $\pi \rightarrow \pi^*$ transitions.^[76,110,113] In the low-energy region (between 450 and 350 nm), FMN has dark singlet and triplet $n \rightarrow \pi^*$ states especially sensitive to the environment.^[77,110,111,113] Upon significant distortions due to surrounding atoms, those forbidden transitions are then allowed, modifying the spectra. Depending on the environment, the structure of flavin isoalloxazine rings can be planar^[76] or distorted.^[110] Kar et al.^[78] suggested the existence of a wide distribution of angles in flavins encapsulated in protein environments. In a theoretical 1PA and 2PA investigation, List et al.^[4] used a QM/MM scheme with PE and accounted for explicit dynamic structural effects to examine the behavior of FMN in aqueous solution and in the miniSOG protein. Comparisons with respect to experimental 1PA and 2PA spectra indicate that spectral shapes are rather well reproduced, while peak energy positions are blue-shifted.

The objective of this study is to establish possible AQM workflows to compute 1- and 2PA of molecules in solution and complex environments (see **Figure 3**). Note that in this contribution, the terminology “realistic systems” is used to illustrate that not only calculations on single molecules or small systems are performed but also including explicitly large parts of the environment as well as dynamic structural effects (when possible) to provide model systems closer to reality. The general idea is to rely only on QM levels of theory and using a similar approach as some of us did for the evaluation of the first hyperpolarizability of FPs.^[88] **Figure 3** gives a general overview of the methodology we will apply in this work. A first protocol is proposed for FPs

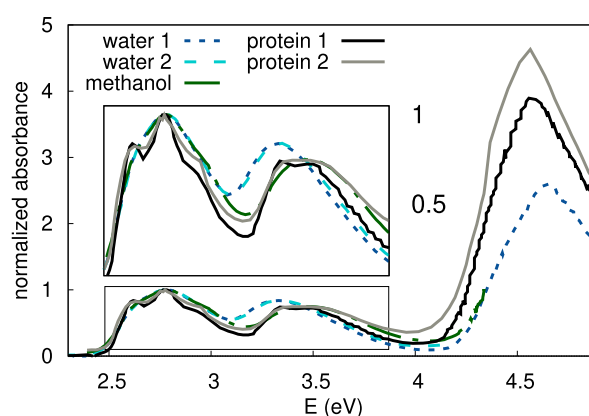


Figure 2. Experimental 1PA spectra of the FMN in different surroundings (water 1,^[106] water 2,^[4] methanol,^[107] protein 1 (iLOV),^[108] and protein 2 (miniSOG)^[109]). All spectra were normalized to one for the lowest-energy band around 2.79 nm.

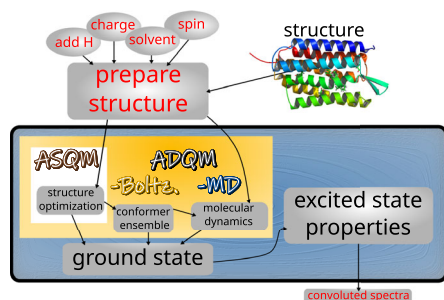


Figure 3. General workflow to obtain excited state spectra. Red font marks critical user-defined steps. In this work, each step inside the black framed box considers all atoms quantum mechanically.

where we use a single structure approach, neglecting dynamic structural effects: an all-atom single structure QM (ASQM) methodology. For this, we used the same systems as in Beaujean et al.^[88] that is, bR (≈ 3850 atoms) and iLOV (≈ 2000 atoms) (see **Figure 4** left and middle) optimized at the ONIOM ω B97X-D/6-31G*:GFN2-xTB/GBSA(water) level of theory. The sTD-DFT-xTB method^[82] is selected to compute excited states as well as 1PA and 2PA spectra directly considering entire structures. For both steps, the generalized born/surface area (GBSA) model^[114–116] is used. The challenging systems to assess the ASQM approach (bR and iLOV) were selected because experimental 1PA and 2PA spectra are available in the literature.^[102,108,109,117–119] With these systems, we are at the edge of what is computationally feasible nowadays for such scheme.

A second protocol is proposed for more structurally dynamic systems such as chromophores in solution. This protocol is named all-atom dynamic structure QM (ADQM). In that case, two approaches were tested to account for dynamic structural effects adapting usually admitted workflows (Figure 1) to AQM procedures. First, solvent effects are treated implicitly and the conformational space is screened at the tight-binding GFN2-xTB level using the conformer/rotamer sampling tool CREST.^[120] Afterward, the GFN2-xTB ensemble is refined with the r^2 SCAN-3c^[121–123] DFT method using CENSO.^[124,125] Boltzmann averaged 1- and 2PA spectra are determined using the sTD-DFT-xTB/GBSA method, accounting implicitly for dynamic structural effects. This variant of the ADQM protocol is named ADQM-Boltz. Second,

solvent effects are incorporated by a sphere of explicit solvent molecules around the chromophore. Dynamic structural effects are considered by running mixed classical/QM GFN2-xTB/GBSA MD simulations. This approach computes forces analytically at the QM GFN2-xTB level of theory to solve Newton's equation of motions. Thus, only trajectories are computed classically.^[126] For each uncorrelated snapshot selected from these MD simulations, 1- and 2PA spectra are computed at the sTD-DFT-xTB/GBSA level of theory and averaged. This variant of the ADQM protocol is called ADQM-MD. Note that in our MD simulations, nuclei move classically which might impact high-frequency motions and narrowing spectral absorption bands. Strategies to modify spectral width or asymmetry of states exists,^[127–129] that is, using the nuclear-ensemble approximation that uses a stochastic ensemble of normal coordinates generated by a Wigner distributions. This approach seems limited to rather small systems. Lukeš et al.^[130] showed that increasing the MD temperature could be used as a parameter to provide better peak broadenings equivalent to those produce by a Wigner sampling. A feature that will be investigated in future studies as a refinement of the present ADQM protocol. FMN (Figure 4 right) was selected for this first exploratory study to assess the proposed AQM protocol that includes dynamical effects of the solvent and the chromophore. Results are compared to available experimental 1PA and 2PA spectra for FMN in aqueous solution.^[4,105–108]

Note that while AQM protocols are developed to better capture the complexity of the experiment by providing the most realistic system we can model nowadays, comparing results from such calculations to experiment is still a tremendous task. Our calculations which neglect vibronic effects provide excitation energies that might need to be shifted and strengths that need to be convoluted by arbitrary functions (either Lorentzian or Gaussian) with an arbitrary width and converted into macroscopic cross sections. Converting 2PA strengths to macroscopic cross sections depend on experimental setups as stressed by Beerepoot et al.^[54] and this information is not always available in the literature. Experimental error bars are not always provided neither. On the top of this, 2PA cross sections are dependent on the square of the excitation energy. Thus, what we compare to experiment is not always completely well-defined and clear.

This study was also the opportunity to test the dt-sTD-DFT-xTB method to compute excited states, which has never been

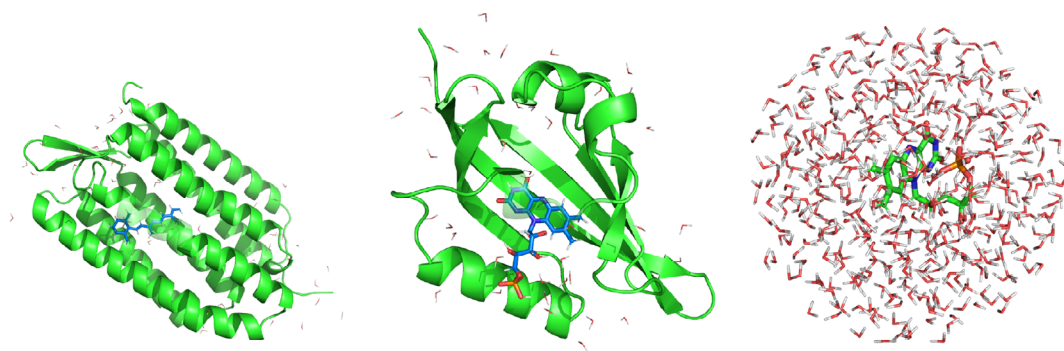


Figure 4. Chemical sketches of bacteriorhodopsin (left panel), iLOV (middle panel), and the flavin mononucleotide surrounded by a shell of 12 Å of explicit water molecules (right panel).

done until now. This scheme seems particularly relevant to reduce the computational cost when a large number of excited states calculations need to be performed, for example, when considering hundreds snapshots from MD runs. We will demonstrate its advantage and illustrate how it can reduce the computational effort while maintaining accuracy in comparison to regular sTD-DFT-xTB calculations.

This study is organized as follows: we first give a concise overview of the theory behind the dt-sTD-DFT as well as how to compute 1- and 2PA at the sTD-DFT level. Then, the computational details employed throughout this work are given (Section 3). In Section 4, results are discussed. Finally, the conclusions and outlooks are given in Section 5.

2. Theory

The dt-sTD-DFT method was developed to reduce the computational cost of sTD-DFT calculations to evaluate the first hyperpolarizability of FPs.^[88,91] Here, we extend the reach of this scheme to determine excited states and to evaluate 2PA cross sections. This section briefly recalls the theory behind the sTD-DFT method to compute excited states^[80,81] and 2PA cross sections.^[90] Details are also given about the dt-sTD-DFT scheme.^[88,91] In this study, we will use $i, j, \dots$ for occupied, $a, b, \dots$ for virtual and $p, q, \dots$ for molecular orbitals (MO) of any kind.

The sTD-DFT method is based on Casida's equations:

$$\begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B} & \mathbf{A} \end{pmatrix} \begin{pmatrix} \mathbf{X} \\ \mathbf{Y} \end{pmatrix} = \begin{pmatrix} \omega & 0 \\ 0 & -\omega \end{pmatrix} \begin{pmatrix} \mathbf{X} \\ \mathbf{Y} \end{pmatrix} \quad (1)$$

where $\mathbf{A}$ and $\mathbf{B}$ are orbital rotation Hessian matrices, $\mathbf{X}$ and $\mathbf{Y}$ are the excitation and de-excitation vectors, and ω represents the eigenvalue diagonal matrix with the dimension of the number of roots (or states) and describes the excitation energies. In 2013, Grimme^[80] proposed three simplifications to construct $\mathbf{A}$ and $\mathbf{B}$ supermatrices: i) the exchange-correlation kernel f_{xc} in $\mathbf{A}$ and $\mathbf{B}$ supermatrices is neglected to avoid time-intensive numerical integration; ii) the two-electron integrals are approximated by short-ranged damped Coulomb interactions of Löwdin transition charges; and iii) the single excitation space is truncated, leading to a significant speed-up in time with only a minor loss of accuracy.^[80,81] Matrix elements of $\mathbf{A}'$ and $\mathbf{B}'$ in sTD-DFT read the following.

$$\mathbf{A}'_{ia,jb} = \delta_{ij}\delta_{ab}(\varepsilon_a - \varepsilon_i) + \sum_{A,B}^N (2q_{ia}^A q_{jb}^B \Gamma_{AB}^K - q_{ij}^A q_{ab}^B \Gamma_{AB}^J) \quad (2)$$

$$\mathbf{B}'_{ia,jb} = \sum_{A,B}^N (2q_{ia}^A q_{bj}^B \Gamma_{AB}^K - a_x q_{ib}^A q_{aj}^B \Gamma_{AB}^K) \quad (3)$$

where two-electron MO integrals written in the chemists' notations $(pq|rs) = \iint \Psi_p^*(\mathbf{r}_1) \Psi_q(\mathbf{r}_1) \frac{1}{r_{12}} \Psi_r^*(\mathbf{r}_2) \Psi_s(\mathbf{r}_2) d\mathbf{r}_1 d\mathbf{r}_2$ are approximated as follows

$$(pq|rs) \approx \sum_A^N \sum_B^N q_{pq}^A q_{rs}^B \Gamma_{AB} \quad (4)$$

Transition charges q_{pq}^A are obtained from a Löwdin population analysis,^[131] and the Γ_{AB} function damps their interaction at short distance using the Mataga–Nishimoto–Ohno–Klopman formula.^[132–134] For a Coulomb-type integral $(ij|ab)'$, the damping formula takes the form

$$\Gamma_{AB}^J = \left(\frac{1}{(R_{AB})^{y_J} + (a_x \eta)^{-y_J}} \right)^{\frac{1}{y_J}} \quad (5)$$

with $\eta = \frac{\eta(A) + \eta(B)}{2}$

where R_{AB} denotes the distance between atom A and B , y_J is a linear function depending on the amount of Fock exchange a_x , and η is the average chemical hardness of two atoms. Exchange type integrals $(ib|aj)'$ are damped using a slightly different equation.

$$\Gamma_{AB}^K = \left(\frac{1}{(R_{AB})^{y_K} + \eta^{-y_K}} \right)^{\frac{1}{y_K}} \quad (6)$$

where y_K is again a linear function that depends on the amount of Fock exchange. Both Γ_{AB} functions were determined as $y_J = 0.20 + 1.83a_x$ and $y_K = 1.42 + 0.48a_x$, respectively.^[80,81]

To truncate the configurations interaction space (related to the size of $\mathbf{A}'$ and $\mathbf{B}'$), one single energy threshold E_{thr} is used as parameter. E_{thr} represents the maximum excitation energy to be computed to represent the 1PA spectral range. It was shown that a careful truncation of the CI space decreases computational costs in time and memory effectively without losing much accuracy.^[135–138] As a preliminary step, the MO active space is truncated as follows.

$$\varepsilon_{\text{min}} = \varepsilon_{\text{LUMO}} - 2(1 + 0.8a_x)E_{\text{thr}} \quad (7)$$

$$\varepsilon_{\text{max}} = \varepsilon_{\text{HOMO}} + 2(1 + 0.8a_x)E_{\text{thr}}$$

where $\varepsilon_{\text{HOMO}}$ and $\varepsilon_{\text{LUMO}}$ are energies of the highest occupied (HO) and lowest unoccupied (LU) molecular orbitals. Primary configuration state functions (P-CSFs), describing the single excitations from occupied MO i to unoccupied MO a , are selected when

$$A'_{ia,ia} \leq E_{\text{thr}} \quad (8)$$

For remaining CSFs with $A'_{jb,jb} > E_{\text{thr}}$, using a perturbative approach, secondary CSFs (S-CSFs) are only considered if they are coupled significantly to the P-CSFs space

$$E_{jb}^{(2)} = \sum_{ia}^{\text{P-CSFs}} \frac{|A'_{ia,jb}|^2}{A'_{jb,jb} - A'_{ia,ia}} > 10^{-4} E_h \quad (9)$$

Consequently, the total number of CSFs is the sum of P- and S-CSFs selected. For FMN at the sTD-DFT-xTB level, the space of 12 096 CSFs is drastically reduced to 342 CSFs using a E_{thr} value of 7 eV and 1413 CSFs for $E_{\text{thr}} = 10$ eV.

When considering larger systems with thousands of atoms, computational costs increase drastically. The dt-sTD-DFT scheme gives a solution to truncate even further the space of CSFs for large systems with a central chromophore.^[88,91] The dual threshold approach divides the system into two subsystems, which are treated by two different energy thresholds. The motivation is

simple: Considering a chromophore inside its explicit environment, for example, solvent, the response of the chromophore (few atoms) is much more significant than the response of the surroundings (many atoms). Therefore, CSFs involving the chromophore will impact much more the excited state manifold. Thus, a larger energy threshold (E_{high}) should be employed for the chromophore, while the surroundings are covered with a lower energy threshold (E_{low}), reducing computational costs. Note that E_{high} should at least represent the spectral range.

In the dt-sTD-DFT procedure, the truncation of the MO space is done as before (see Equation (7)) considering E_{high} as the energy threshold. Then, the occupied MO space is split into two parts, the high level (e.g., the chromophore) and the low level ones (e.g., the solvent). If an occupied MO has a density ζ higher than 10% on atoms from the high level part, it is included into the high-level occupied MO space. The remaining occupied MOs are treated within the low-level occupied MO space. For a MO i , we have the following.

$$\zeta_i = \sum_{\alpha \in \text{high level part}} C_{\alpha i}^2 \quad (10)$$

$$\zeta_i > 0.1 \rightarrow E_{\text{high}} \quad (11)$$

$$\zeta_i \leq 0.1 \rightarrow E_{\text{low}} \quad (12)$$

$C_{\alpha i}$ are LCAO coefficients and ζ_i is the electron density of the MO i on the high level part of the system.

P-CSFs are selected using Equation (8) with E_{high} and E_{low} , for single excitations, respectively, involving the high level and low level occupied MO spaces. This leads to the P-CSFs-H space if $A_{i_{\text{high}}a, i_{\text{high}}a} \leq E_{\text{high}}$ and to P-CSFs-L space if $A_{i_{\text{low}}a, i_{\text{low}}a} \leq E_{\text{low}}$. To prune the remaining space, only $j_{\text{low}} \rightarrow b$ CSFs with

$$E_{\text{low}} < A'_{j_{\text{low}}b, j_{\text{low}}b} < 2(1 + 0.8a_x)E_{\text{low}} \quad (13)$$

are considered for the low level part, while all remaining $j_{\text{high}} \rightarrow b$ CSFs are selected for the high level part. S-CSFs are then selected if

$$E_{jb}^{(2)} = \sum_{ia}^{P\text{-CSFs}-(H+L)} \frac{|A'_{iajb}|^2}{A'_{iaja} - A'_{jbjb}} \quad (14)$$

Finally, the total amount of CSFs is the sum of all primary CSFs from the high- and low-level parts as well as the S-CSFs, that is, #CSFs = #P-CSFs - H + #P-CSFs - L + #S-CSFs.

To compute a 1PA spectrum at the (dt)-sTD-DFT level, the transition dipole moment $\vec{\mu}_{0v}^L$ between the ground state and the excited state v is evaluated using X' and Y' . Considering the restricted case in the length formalism, the transition dipole moment reads

$$\vec{\mu}_{0v}^L = \sqrt{2} \sum_{ia} \vec{\mu}_{ia} (X_{ia}^{iv} + Y_{ia}^{iv}) \quad (15)$$

where $\vec{\mu}_{ia} = \langle \psi_i | \vec{r} | \psi_a \rangle$. The associated oscillator strength is obtained as follows.

$$f_{0v} = \frac{2}{3} \omega_{v0} \vec{\mu}_{0v} \cdot \vec{\mu}_{v0} \quad (16)$$

Oscillator strengths are then convoluted with Gaussian functions with a sole damping factor Γ at half width at $1/e$ maximum for all transitions to obtain the 1PA spectrum.

A 2PA spectrum is computed with the (dt)-sTD-DFT method by first evaluating two-photon (2P-) transition dipole moments $M_{\zeta\eta}^{0 \rightarrow n}$ using the following expression.^[90]

$$M_{\zeta\eta}^{0 \rightarrow n} = -A + B \quad (17)$$

where A and B read

$$A = \sum_{\text{perm}, \zeta, \zeta', \eta} \left\{ \sum_{aij} X'_{n,ai} [\mu_{\zeta,ij} (1 - \delta_{\zeta\eta})] Y'_{\eta,ai} (-\omega_n/2) \right\} \quad (18)$$

and

$$B = \sum_{\text{perm}, \zeta, \zeta', \eta} \left\{ \sum_{iab} X'_{n,ai} [\mu_{\zeta,ab} (1 - \delta_{\zeta\eta})] Y'_{\eta,bi} (-\omega_n/2) \right\} \quad (19)$$

$X'_{\eta}(-\omega_n/2)$ and $Y'_{\eta}(-\omega_n/2)$ are linear response vectors obtained at the frequency $-\omega_n/2$ from extra linear response calculations for each transition, and perm. ζ, ζ', η stands for permutations between cartesian coordinates and their respective frequencies. From transition dipole moments, the rotationally averaged 2PA transition strength $\delta^{2PA[139,140]}$ is evaluated as follows.

$$\langle \delta^{2PA} \rangle = \frac{F}{30} \sum_{\zeta, \eta} S_{\zeta\zeta, \eta\eta} + \frac{G}{30} \sum_{\zeta, \eta} S_{\zeta\eta, \zeta\eta} + \frac{H}{30} \sum_{\zeta, \eta} S_{\zeta\eta, \eta\zeta} \quad (20)$$

with $S_{\zeta\eta, \zeta v} = M_{\zeta\eta}^{0 \rightarrow n} M_{\zeta v}^{0 \rightarrow n}$

where F, G , and H are parameters which depend on the experimental setup. In the case of parallel linearly polarized incident light, F, G , and H equal 2. All 2PA spectra provided in this contribution consider this case. To compute 2PA spectra, macroscopic 2PA cross-section σ^{2PA} are computed as follows.

$$\sigma^{2PA} = \frac{N \alpha^3 a_0^5 (2\omega)^2}{c} \langle \delta^{2PA} \rangle G(2\omega, \omega_n, \Gamma) \quad (21)$$

where the parameter N depends on the experimental setup to compare with (single or double beam experiment, $N = 1$ or 2), α depicts the fine structure constant, a_0 denotes the Bohr radius, c is the speed of light, and G represents a line-shape function. Gaussians are usually used for Γ to model 2PA spectra in solution.^[54]

$$G(2\omega, \omega_n, \Gamma) = \frac{\sqrt{\ln 2}}{\sqrt{\pi} \Gamma} \exp \left[-\ln 2 \left(\frac{2\omega - \omega_n}{\Gamma} \right)^2 \right] \quad (22)$$

where Γ is the half-width at half-maximum (HWHM). Note that while an identical Γ factor is usually used for all transitions,^[49,52,90,141–144] it was shown that 2P-transitions have often different broadenings,^[145,146] originating from different interactions. In the following, broadening factors are given separately for each system and are chosen to best reproduce experiment. Note that for average ADQM-MD spectra, smaller damping factors are used to capture spectral features from out of equilibrium structures that

are naturally broadening absorption bands. All along the article, 2PA spectra are presented as function of 2ω .

3. Computational Details

To assess the ASQM protocol, structures of bR (bacteriorhodospin) and iLOV were taken from Beaujean et al.^[88] They were optimized at the ONIOM ω B97X-D^[147,148]/6-31G*^[149–154]:GFN2-xTB^[155]/GBSA(water)^[114–116] level of theory and include few explicit water molecules (57 in bR and 46 in iLOV) around the protein. The ω B97X-D range-separated hybrid exchange-correlation functional was used at the DFT level for the ONIOM high layer including the chromophore and the surrounding amino acids and water molecules in a 4 Å radius. The remaining of the protein including external water molecules was accounted for in the ONIOM low layer with the GFN2-xTB method.^[155] Excited states as well as δ^{2PA} were computed at the (dt)-sTD-DFT-xTB level of theory.^[82] Note that the implicit GBSA(water) model was used to further account for solvent effects for ground state xTB calculations. For each sTD-DFT-xTB 2PA calculation, δ^{2PA} were computed for the 20 first excited states and converted to σ^{2PA} using Equation (21). Note that computing σ^{2PA} for the 20 first excited states is enough to match the experimental 2PA spectral range for all systems considered in this study.

The ADQM protocol includes dynamic structural effects considering its two variants: ADQM-Boltz. and ADQM-MD. In the ADQM-Boltz. methodology for water solution, the deprotonated FMN conformer/rotamer ensemble (CREs) was determined at the GFN2-xTB level of theory using CREST 2.12.^[120] Solvent effects were accounted for using the GBSA implicit solvent model. The CRE for FMN in vacuum was computed in the same manner, excluding any solvent model. Both CREs were refined at the DFT level: first, by using the B97-D3^[156,157]/def2-SV(P)^[158] scheme for a cheap pre-screening, and second the composite r^2 SCAN-3c^[121–123] method to reoptimize geometries. Implicit solvent effects were included using the DCOSMO-RS(water)^[159,160] model. Final Gibbs energy rankings are obtained using the r^2 SCAN-3c electronic energy + G_{mRRHO} ^[161–163](GFN2-xTB)// r^2 SCAN-3c for the ensemble in gas phase as well as the r^2 SCAN-3c electronic energy + COSMO-RS[H₂O]^[164,165] + G_{mRRHO} (GFN2[GBSA]-bhess)// r^2 SCAN-3c[DCOSMO-RS]^[159,160] in implicit water. This procedure is performed with the command line energetic sorting algorithm CENSO 1.2.0,^[124,125] using Turbomole V7.6,^[166–168] the xtb program package V6.6.1,^[125] and the COSMOTHERM 16 program package (G_{sol} option, BP_TZVP_C30_1601 parametrization, $T = 298.15$ K, $p = 1$ atm).^[169]

Alternatively, dynamic structural effects were accounted for by taking snapshots from MD simulations, giving the ADQM-MD protocol. Accounting explicitly for solvent effects, water molecules were added to FMN lowest-energy conformer using spheres of increasingly large radii of 7, 9, 10, 11, and 12 Å with the program Chimera^[170] ("shell" solvation and TIP3PBOX solvation model). Each structure were preoptimized at the GFN2-xTB/GBSA(water) level to remove possible tensions due to the solvation process, giving better starting points for MD simulations. 1 ns long GFN2-xTB/GBSA(water) MDs were run with a time step of 4 fs to determine an NVT ensemble at 298.15 K. Note that the SHAKE algorithm was

used to constrain all bonds, and we applied four times the mass of hydrogen to all hydrogen atoms. Two polynomial wall potentials were set as follows: 1) One spherical potential to constrain the whole system and prevent water molecules to "escape," mimicking the "glass" of a real world experiment. Spheres of radius values of 11.6, (22.0), 13.2 (25.0), 14.7 (27.7), 15.3 (29.0), and 16.6 (31.45) Å (Bohr) were chosen for the potential when considering 7, 9, 10, 11, and 12 Å of water shells, respectively. Note that in the following, we always refer to FMN solvated in X Å water shell as " X Å." 2) A second wall potential was set around a central atom of FMN with a sphere radius of 2.6 Å (5.0 Bohr) to keep the referential of the molecule centered and avoid drifting apart.

A longer MD simulation (3 ns) was also conducted for the 10 Å water shell to assess the impact of the MD time length on spectra. To select uncorrelated snapshots from MD simulations, all trajectories were evaluated using the autocorrelation function (ACF) of the program TRAVIS.^[171,172] Additionally, the block averaging method was used to get the smallest time step between two uncorrelated configurations. This procedure led to 118 (7 Å), 98 (9 Å), 237 (10 Å), 134 (11 Å), and 166 (12 Å) snapshots, respectively. To complement discussions about FMN in gas phase, a 11 ns long GFN2-xTB MD simulation on FMN lowest-energy conformer was run using the same conditions as described above. 1521 snapshots were taken.

For each structure generated with the second protocol, xTB/GBSA(water) ground states for sTD-DFT-xTB calculations were computed. Excitation energies, oscillator strengths, and δ^{2PA} were determined at the (dt)-sTD-DFT-xTB level of theory. For each conformer or MD snapshot, 1PA and 2PA spectra were generated by convoluting oscillator strengths or δ^{2PA} with a Gaussian line-shape. Smaller damping factors were employed for spectra of MD snapshots before taking the average to resolve structural influences (see Section 4).

Furthermore, RI-CC2^[173]/aug-cc-pVDZ^[174] calculations were performed to compute 1- and 2PA spectra for retinal and FMN lowest-energy conformer obtained with r^2 SCAN-3c in gas phase and implicit solvent (FMN only). Turbomole 7.7.1^[175] was used to compute eight RI-CC2 transition energies and cross sections using default parameters. Note that we used a value of 1000 for the maximum dimension of the reduced space of the CC linear equations (maxred). A convergence threshold of 10^{-6} is used for the numerical Laplace transformation to compute δ^{2PA} .

The xtb program package^[125] version 6.5.1 was used for all GFN2-xTB calculations. All-atom xTB ground states were calculated with the xtb4stda program.^[176] (dt)-sTD-DFT-xTB calculations were performed using a development version of the std2 package (previously branded stda).^[80,90,177] Note that all along this study, vibronic effects are not accounted for because the gradient is not implemented yet in std2.^[80,90,177]

4. Results and Discussions

The main goal of this work is to assess two AQM protocols considering the effect of the environment explicitly without and with dynamic structural effects: ASQM and ADQM, respectively. This section is divided into two parts. Using the ASQM protocol, we

first discuss 1- and 2PA of two very large systems: bR and iLOV. This first part is also the opportunity to discuss the dt-STD-DFT approach to compute excited states for such systems. Second, we investigate the influence of dynamic structural effects on 1- and 2PA spectra for FMN in aqueous solution using both ADQM-Boltz. and ADQM-MD methodologies.

4.1. An All-Atom Single Structure QM Protocol Applied to Fluorescent Proteins

4.1.1. Bacteriorhodopsin

We first assess the ASQM workflow to compute the 1PA spectrum of bR (Figure S1, Supporting Information) at the sTD-DFT-xTB/GBSA(water) level of theory in comparison with experimental 1PA spectra from de Coene et al.^[117] and Winder et al.^[118] recorded in water. Figure 5 presents this comparison considering

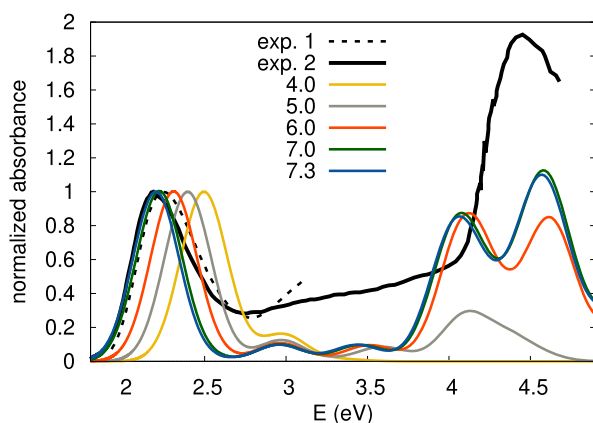


Figure 5. Experimental 1PA spectra of bR obtained by de Coene et al.^[117] (exp. 1) as well as Winder et al.^[118] (exp. 2) recorded in water in comparison with computed sTD-DFT-xTB/GBSA(water) spectra considering different energy thresholds (4.0, 5.0, 6.0, 7.0, and 7.3 eV). Convolved spectra used a damping factor $\Gamma = 0.2$ eV and were shifted by -0.4 eV, allowing the sTD-DFT-xTB ($E_{\text{thr}} = 7.3$ eV) spectrum to better match experiment.

different E_{thr} values (4.0, 5.0, 6.0, 7.0, and 7.3) eV for the sTD-DFT-xTB scheme. Note that for the ASQM procedure, only the optimized structure of bR from Beaujean et al.^[88] was used. Figure 6 shows computation times and numbers of CSFs involved for sTD-DFT-xTB calculations (left panel) as a function of E_{thr} where one can observe a direct correlation between computation times and numbers of configurations. Note that the 40 min-long excited state calculation considering $E_{\text{thr}} = 7.3$ eV with the std2 program was the largest computation we were able to run without filling completely the memory of an AMD EPYC 7742 64-Core Processor (1500.0 MHz, 1024 GB RAM) computer node. Both experiments show a first maximum around 2.25 eV (552 nm)^[117] and 2.19 eV (567 nm),^[118] respectively. Theoretical spectra were shifted by -0.4 eV to allow the sTD-DFT-xTB spectrum computed with $E_{\text{thr}} = 7.3$ eV to best match this absorption band. The experimental spectrum from Winder et al.^[118] presents two main absorption bands as also predicted by the theory. The second band energy position is actually well-reproduced by the sTD-DFT-xTB spectrum computed with the largest energy threshold when no energy shift is applied with only 0.01 eV of difference with respect to experiment. Note that at least a value of E_{thr} larger than 6 eV needs to be employed to account for this second peak providing a reasonable comparison to experiment. As expected, E_{thr} represents the spectral window. The gap between both experimental and theoretical first peak position is decreased when increasing the energy threshold going from 0.7 eV ($E_{\text{thr}} = 4$ eV) to 0.4 eV ($E_{\text{thr}} = 7.3$ eV) with respect to experiment. As both experimental spectra were normalized, we cannot discuss absolute peak intensities but only relative one. Our calculations are not able to reproduce the experimental ratio between both main absorption bands but spectral shapes remain reasonable.

Figure 7 depicts natural transition orbitals (NTOs) for the first excited state of the unoptimized retinal chromophore extracted from bR computed in vacuum with the sTD-DFT-xTB method as well as for the first and second excited states of the full protein bR at the sTD-DFT-xTB/GBSA(water) ($E_{\text{thr}} = 7.3$ eV) level of theory. The first excited state of retinal in gas phase is a pure $\pi \rightarrow \pi^*$

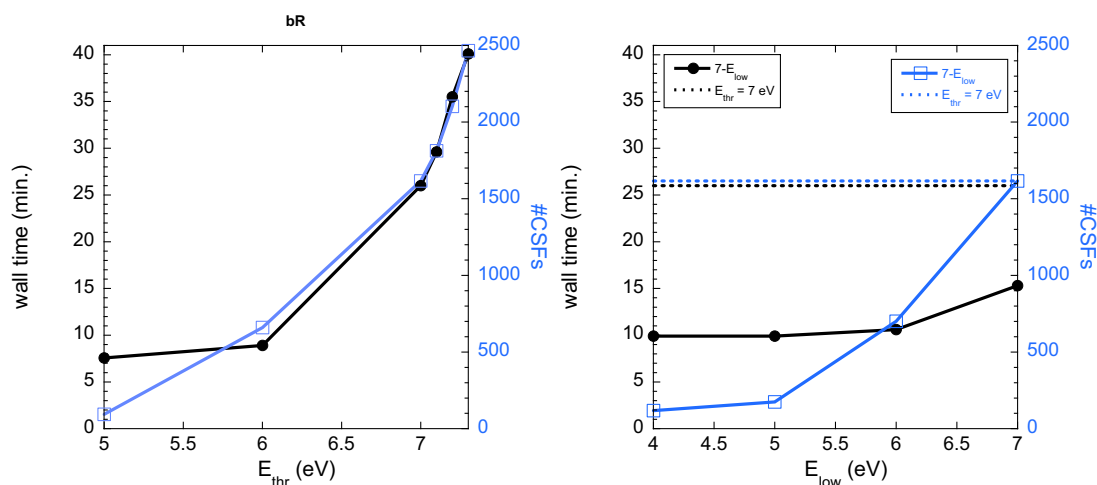


Figure 6. For bR, computation times and numbers of CSFs involved as a function of E_{thr} for sTD-DFT-xTB/GBSA(water) calculations (left panel) as well as for dt-sTD-DFT-xTB/GBSA(water) calculations (right panel) as a function of E_{low} considering $E_{\text{high}} = 7$ eV.

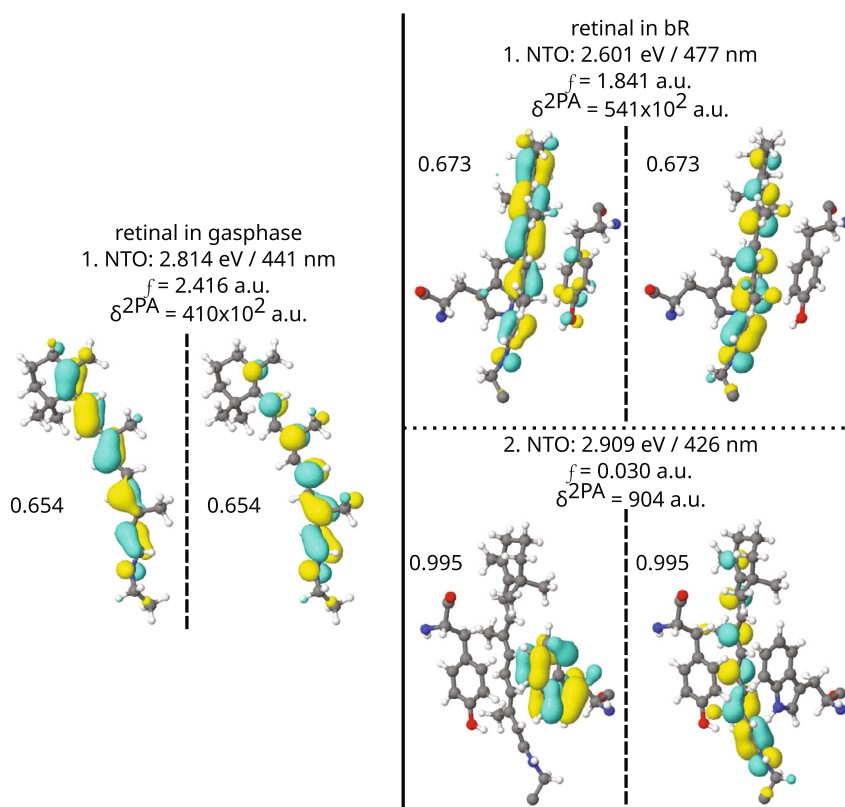


Figure 7. NTOs (left panel) computed at the sTD-DFT-xTB for the first excited state of the retinal chromophore extracted from bR as well as NTOs (right panel) computed at the sTD-DFT-xTB/GBSA(water) (both calculations with $E_{\text{thr}} = 7.3$ eV) for the first and second excited states of bR. Hole and particle NTOs are separated by dashed lines. Weights are provided for each NTO pair. The excitation energy, oscillator strength, and 2PA strength are also given for each state. Isovalue = 0.03.

excitation happening on the π -conjugated pathway of the molecule. As already mentioned in the introduction, previous studies^[98–100] showed that in the protein, the retinal interacts with a nearby tyrosine (TYR185). This is confirmed by our NTO analysis. For the first excited state, we have the same $\pi \rightarrow \pi^*$ excitation as for the retinal in gas phase but with contributions from the tyrosine in the hole NTO. This small charge transfer (CT) contribution stabilizes this electronic transition energy by 0.21 eV and the oscillator strength is reduced from 2.416 to 1.841. Interestingly, the NTOs analysis (not shown here) shows that both higher energy peaks are composed of collection of $\pi \rightarrow \pi^*$ transitions located on tryptophans (TRP80, 137, 138, and 198) from the protein. Obviously, such transitions can only be captured by accounting for the whole protein into the excited state calculation. This result advocates for the use of the ASQM methodology to compute 1PA spectra of proteins.

Figure 8 compares the sTD-DFT-xTB/GBSA(water) spectrum computed with $E_{\text{thr}} = 7.0$ eV to dt-sTD-DFT-xTB/GBSA(water) ones computed with $E_{\text{high}} = 7.0$ eV and E_{low} values of 4.0, 5.0, and 6.0 eV. To reproduce the second larger absorption band around 4 eV from the sTD-DFT-xTB(7.0 eV) calculation, we need to use at least an energy threshold of 6.0 eV for the protein surroundings because it involves $\pi \rightarrow \pi^*$ transitions located on tryptophans. Note that dt-sTD-DFT spectra were shifted by a larger energy shift of -0.6 eV to match the lowest-energy peak. Thus, the 0.2 eV difference of energy observed in Figure 8 for the second band between

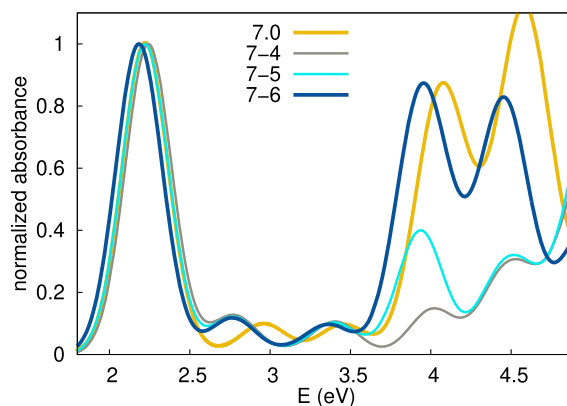


Figure 8. For bR, sTD-DFT-xTB/GBSA(water) 1PA spectrum computed with $E_{\text{thr}} = 7.0$ eV in comparison with dt-sTD-DFT-xTB/GBSA(water) ones computed with $E_{\text{high}} = 7.0$ eV and E_{low} values of 4.0, 5.0, and 6.0 eV. Convolved spectra used a damping factor $\Gamma = 0.2$ eV and were shifted by -0.4 eV for sTD-DFT-xTB and -0.6 eV for dt-sTD-DFT-xTB.

both sTD-DFT-xTB (7.0 eV) and 7-6 spectra is artificial as both excitation energies are 4.48 and 4.55 eV, respectively. Figure 6 shows computation times and numbers of CSFs involved for sTD-DFT-xTB (left panel) and dt-sTD-DFT-xTB calculations (right panel) as a function of E_{low} (see Table S2, Supporting Information for more details). Using 1351 CSFs, the 7-6 dt-sTD-DFT-xTB calculation took only 11 min, while the sTD-DFT-xTB(7.0 eV) one, using 1616 CSFs, took

26 min, reducing by more than a factor of two the computation time without affecting much the global shape of the spectrum. Note that the 7–4 spectrum is already reproducing well the sTD-DFT-xTB(7.0 eV) one for (shifted) excitation energies below 3.5 eV but the extra cost to perform 7–6 dt-sTD-DFT-xTB calculation is negligible with respect to the 7–4 one. While leading to the same configuration space of 1616 CSFs, the 7–7 calculation took less time (15 min) than the sTD-DFT-xTB(7.0 eV) one (26 min) as less CSFs were probed by the perturbative approach to determine S-CSFs due to the restricted active space defined by Equation (13).

Second, we assess the ASQM approach to evaluate the 2PA spectrum of bR with the sTD-DFT-xTB/GBSA(water) with respect to the experiment. **Figure 9** presents the experimental 2PA spectrum for bR recorded in water by Birge et al.^[102] in 1990 in comparison with computed sTD-DFT-xTB/GBSA(water) spectra considering different energy thresholds (5.0, 6.0, 7.0, 7.1, 7.2, and 7.3 eV). Note that according to data from Birge et al.^[102] (see Table 1 in ref. [102]), the experimental spectrum was not corrected for background voltage, and all 2PA cross-sections were contaminated by 1PA signals. To improve the comparison with respect to this rather old experimental 2PA spectrum, we corrected arbitrarily the baseline by shifting the minimum error bar of the lowest-energy point at 0 GM for all experimental 2PA cross sections. Note too that the 2PA spectrum was recorded for a very narrow 1 eV-wide spectral range. This 2PA spectrum shows a distinct peak at 2.27 eV with a maximum of 191 (± 38) GM, a shoulder around 2.5 eV, and the onset of a another peak around 3 eV. Computed 2PA spectra were shifted by -0.4 eV as 1PA ones. Our calculations reproduce well the global shape of the experimental 2PA spectrum even if the shoulder of the first excited state is not reproduced. It could be due to vibronic effects that are not accounted for by our calculations or to the second excited state for which the 2PA cross-section could be underestimated. Figure 7 depicts the NTO pair for the second excited state of bR computed at the sTD-DFT-xTB/GBSA(water)

($E_{\text{thr}} = 7.3$ eV) level of theory, showing a CT transition at 2.91 eV from a tryptophan (TRP86) to the retinal chromophore with $\delta^{2\text{PA}} = 904$ a.u. Note that for the first excited state, the 2PA cross section is enhanced from the gas phase retinal calculation from 410 to 541×10^2 GM for the full protein thanks to the nearby tyrosine giving the CT character to the transition. The onset of the third band is also computed. Because of the baseline problem of the experiment and that 2PA bands were not filtered from 1PA, quantitative comparisons on 2PA cross sections with respect to experiment are especially challenging. However, computed and experimental cross sections are more or less of the same order of magnitude.

Regarding the impact of the energy threshold on calculations, as we observed for 1PA spectra, the energy gap between the first experimental 2PA band and the sTD-DFT-xTB computed one decreases when increasing this threshold, while the rest of the spectrum is not changing much. **Figure 10** compares the experimental spectrum recorded by Birge et al.^[102] in water with the computed sTD-DFT-xTB/GBSA(water) spectrum considering an energy threshold of 7.0 eV as well as dt-sTD-DFT-xTB/GBSA(water) spectra with 6–4, 6–5, 7–4, 7–5, and 7–6 values of $E_{\text{high}} - E_{\text{low}}$ (see Figure S2, Supporting Information, for comparison of sTD-DFT and dt-sTD-DFT for more energy thresholds). As for 1PA, a larger energy red-shift of -0.6 eV was applied for dt-sTD-DFT-xTB 2PA spectra. The influence of chosen energy threshold on dt-sTD-DFT-xTB 2PA calculations remains negligible. All dt-sTD-DFT-xTB 2PA spectra show a first distinct peak around 2.2–2.3 eV fitting well to experiment. The main difference between sTD-DFT and dt-sTD-DFT 2PA spectra is a decreased energy difference between first and second main peaks for dt-sTD-DFT results with respect to sTD-DFT ones. The first excitation at around 2.2 eV has a somewhat smaller magnitude with the dt-sTD-DFT scheme with respect to the sTD-DFT one (7–6 eV: 120 GM vs. 7.0 eV: 142 GM). Nevertheless, the experimental spectrum is rather well reproduced by both methods, even considering small energy thresholds, for example, 6.0 eV or 6–4 eV.

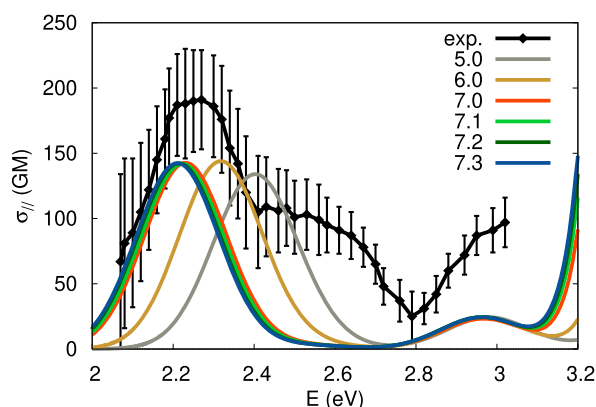


Figure 9. Experimental 2PA spectrum of bR recorded by Birge et al.^[102] in water in comparison with computed sTD-DFT-xTB/GBSA(water) spectra considering different energy thresholds (5.0, 6.0, 7.0, 7.1, 7.2, and 7.3 eV). Convolved spectra used a damping factor $\Gamma = 0.1$ eV and were shifted by -0.4 eV, allowing the sTD-DFT-xTB ($E_{\text{thr}} = 7.3$ eV) spectrum to better match experiment.

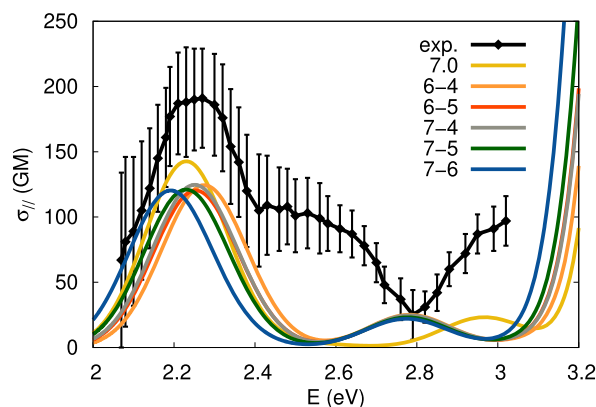


Figure 10. Experimental 2PA spectrum of bR recorded by Birge et al.^[102] in water in comparison with the computed sTD-DFT-xTB/GBSA(water) spectrum considering an energy threshold 7.0 eV as well as dt-sTD-DFT-xTB/GBSA(water) spectra with 6–4, 6–5, 7–4, 7–5, and 7–6 values of $E_{\text{high}} - E_{\text{low}}$. Convolved spectra used a damping factor $\Gamma = 0.1$ eV and were shifted by -0.4 eV for sTD-DFT-xTB and -0.6 eV for dt-sTD-DFT-xTB.

4.1.2. iLOV

For iLOV^[105] (Figure S3, Supporting Information), we further assess the performance of the ASQM workflow to evaluate 1- and 2PA of such large systems. Again, only the optimized structure of iLOV from Beaujean et al.^[88] was used. **Figure 11** presents 1PA spectra recorded in water by Ran et al.^[108] as well as Torra et al.^[109] in comparison with computed sTD-DFT-xTB/GBSA(water) spectra considering different energy thresholds (6.0, 7.0, 8.0, 9.0, and 9.6 eV). Note that no energy shifts were applied to calculated spectra and that they were normalized to match the lowest-energy band at 2.79 eV (445 nm). Both experimental 1PA spectra show three main absorption bands. The first peak presents a clearly resolved vibronic structure, similar to the one found by List et al.^[4] emerging because of the protein environment as for miniSOG (see Figure 2). The left panel of **Figure 12** shows on a logarithmic scale the computational time and number of selected CSFs as a function of E_{thr} for the sTD-DFT-xTB calculations. On an AMD EPYC 7742 64-Core Processor (1500.0 MHz,

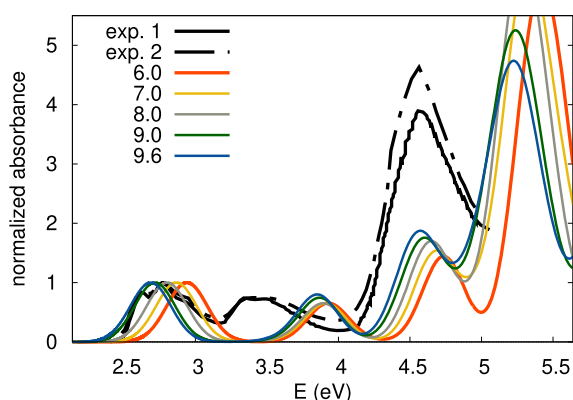


Figure 11. Experimental 1PA spectra of iLOV obtained by Ran et al.^[108] (exp. 1) as well as Torra et al.^[109] (exp. 2) recorded in water in comparison with computed sTD-DFT-xTB/GBSA(water) spectra considering different energy thresholds (6.0, 7.0, 8.0, 9.0, and 9.6 eV). Convolved spectra used a damping factor $\Gamma = 0.2$ eV and no energy shifts were applied.

1024 GB RAM) computer node, the largest sTD-DFT-xTB excited state calculation we were able to run was with $E_{\text{thr}} = 9.6$ eV considering 46 027 CSFs and took a bit less than 41 h. Using an energy threshold of 6.0 eV reduces the computation time to only 2 min. Comparing the sTD-DFT-xTB/GBSA(water) ($E_{\text{thr}} = 9.6$ eV) spectrum to experiments shows that the three main peaks are reproduced. For the first peak, the theoretical excitation energy is only 0.08 eV away from the experimental maximum and for the third peak, we have 0.07 eV of difference. Note that if only the chromophore of iLOV was accounted for, the first excited state energy would be red-shifted by 0.68 eV with respect to the calculation accounting for the whole protein. The second peak in Figure 11 is blue-shifted with respect to experiment by 0.46 eV. This blue shift is related to the underlying isoalloxazine structure and its planarity. During the ONIOM optimization,^[88] the original planar ring system of the isoalloxazine from iLOV crystal structure was slightly bended. As already mentioned in the introduction, the 1PA spectrum of FMN is very sensitive to the planarity of the isoalloxazine ring.^[76,110,111,113] To test this further, we computed the 1PA spectrum of iLOV while keeping the isoalloxazine flat. Position and intensities of the second and third excitations show significant changes, while the first signal remains unaltered (data not shown here). When increasing E_{thr} , computed 1PA spectra are red-shifted and intensities of the second and third bands are increased, better matching experiments.

Figure 13 presents NTOs for the three main excitations of the 1PA spectrum computed at the sTD-DFT-xTB/GBSA(water) ($E_{\text{thr}} = 8.0$ eV) level of theory. The first excited state at 2.80 eV presents a $\pi \rightarrow \pi^*$ transition located on the isoalloxazine ring. No CT character is observed from nearby residues explaining why the signal is almost not altered by the chromophore direct environment (see Figure 2). For the second excited state at 3.65 eV, the hole NTO is a linear combination of n orbitals located on hetero-atoms as well as σ orbitals from the isoalloxazine ring with a small contribution from a P orbital on the oxygen from a nearby asparagine. The particle NTO is the same π^* orbital as for the first excited state. This transition has a small oscillator

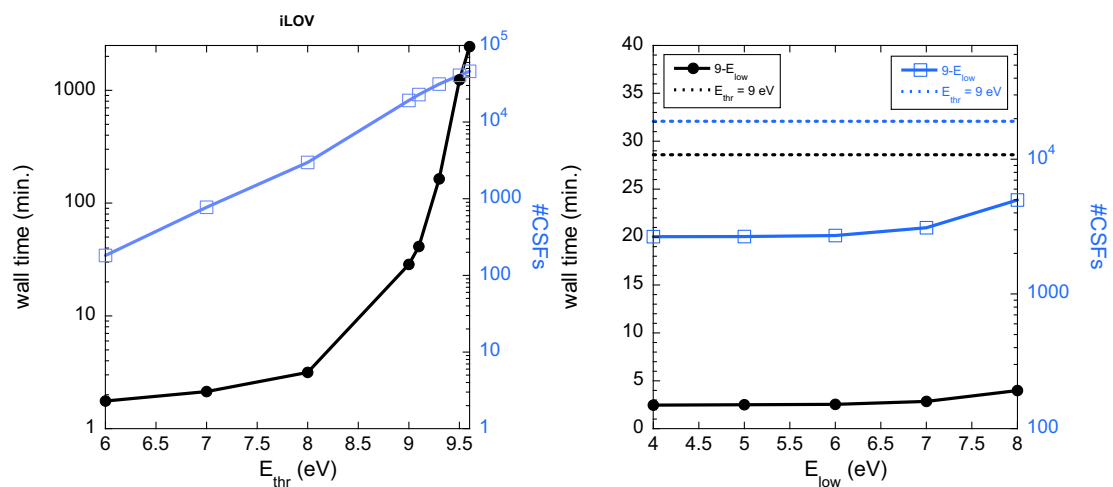


Figure 12. For iLOV, computation times and numbers of CSFs involved as a function of E_{thr} for sTD-DFT-xTB calculations (left panel) as well as for dt-sTD-DFT-xTB calculations (right panel) as a function of E_{low} considering $E_{\text{high}} = 9$ eV.

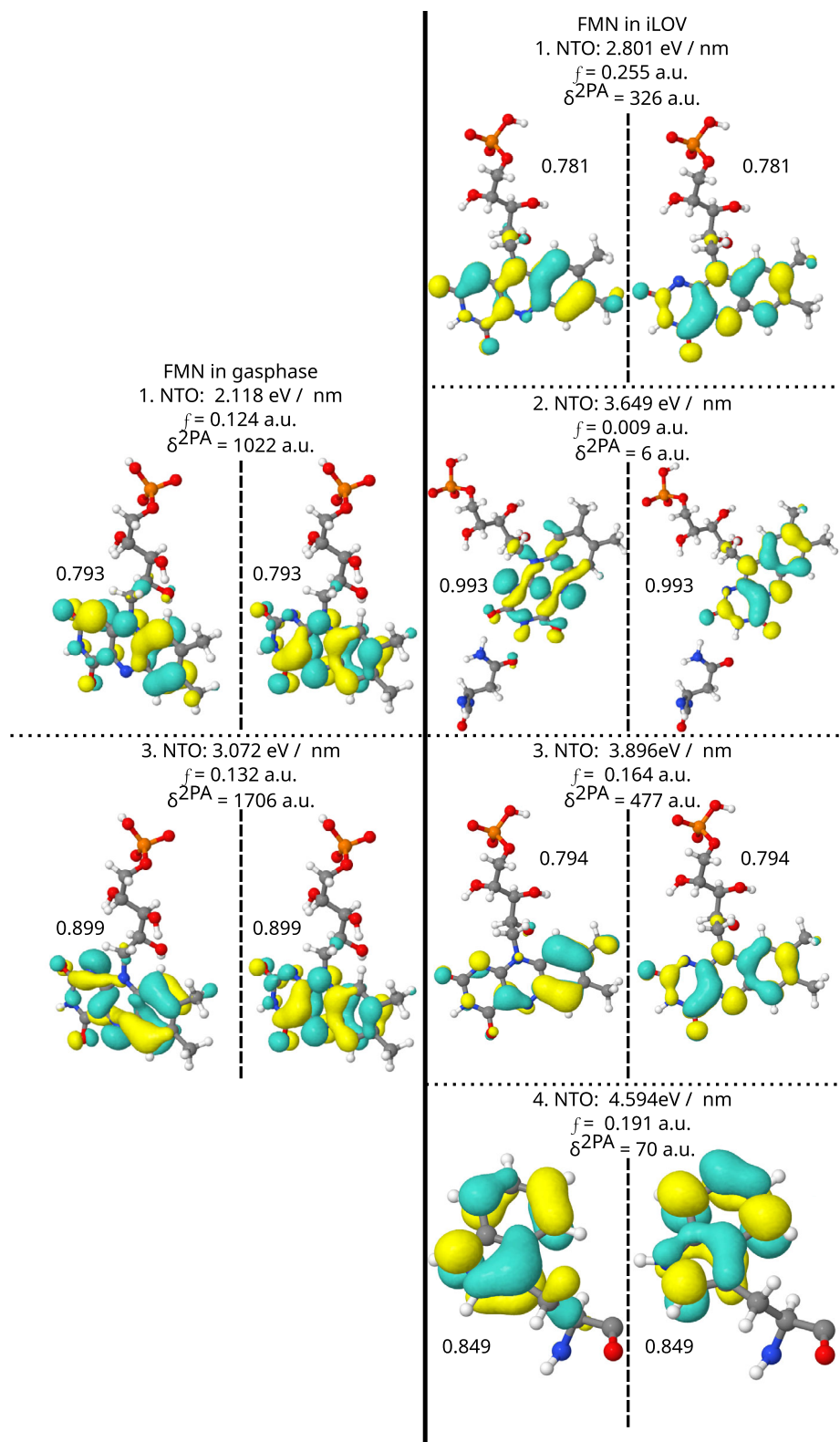


Figure 13. NTOs (left panel) computed at sTD-DFT-xTB(vacuum) for the first and third excited states of FMN chromophore extracted from iLOV as well as NTOs (right panel) computed at the sTD-DFT-xTB/GBSA(water) for the first, second, third, and fourth excited states of iLOV (both calculations with $E_{\text{thr}} = 8.0$ eV). Hole and particle NTOs are separated by dashed lines. Weights are provided for each NTO pair. The excitation energy, oscillator strength, and 2PA strength are also given for each state. Isovalue = 0.03.

strength of 0.01 and does not contribute much to the second absorption band. The third excited state is the main contribution to this second peak with an oscillator strength of 0.16. It is

another $\pi \rightarrow \pi^*$ transition located on the isoalloxazine ring. More interestingly, the fourth excited state is the result of a $\pi \rightarrow \pi^*$ excitation located on a tryptophan of the protein and

thus not involving the chromophore. As for bR and also observed for PYP by one of us,^[93] an AQM protocol is needed to compute properly absorption spectra of fluorescent proteins, especially to account for $\pi \rightarrow \pi^*$ transitions located on tryptophan residues that impact the high energy part of spectra.

Figure 14 compares experimental 1PA spectra to computed ones with both sTD-DFT-xTB/GBSA(water) ($E_{\text{thr}} = 9.0$ eV) and dt-sTD-DFT-xTB/GBSA(water) (9–4, 9–5, 9–6, 9–7, and 9–8 values of $E_{\text{high}} - E_{\text{low}}$) schemes (see Figure S4, Supporting Information, for a more detailed comparison of sTD-DFT and dt-sTD-DFT considering more energy thresholds). The right panel of Figure 12 presents computation times and numbers of CSFs as a function of E_{low} for dt-sTD-DFT-xTB calculations considering $E_{\text{high}} = 9.0$ eV (see Table S3, Supporting Information for more details). While 9–4 and 9–5 dt-sTD-DFT-xTB computed spectra failed to reproduce correctly the third absorption band with respect to the sTD-DFT-xTB spectrum, using a E_{low} value of 6.0 eV seems good enough to replicate the sTD-DFT-xTB/GBSA(water) ($E_{\text{thr}} = 9.0$ eV) spectrum. As already observed for bR, a value of $E_{\text{low}} \geq 6$ eV is necessary to properly cope with tryptophan $\pi \rightarrow \pi^*$ transitions. The configuration space is reduced from 19 165 with $E_{\text{thr}} = 9.0$ eV to 2716 CSFs with 9–6 values of $E_{\text{high}} - E_{\text{low}}$, reducing the computation time by a factor of 10. Note that for sTD-DFT-xTB calculations, it was not possible to use larger E_{thr} value than 9.6 eV on an AMD EPYC 7742 64-Core Processor (1500.0 MHz, 1024 GB RAM) computer node, but we were able to run dt-sTD-DFT-xTB calculations with $E_{\text{high}} = 10.0$ eV, increasing further the space of CSFs for the chromophore (see Figure S4, Supporting Information, lower panel). Using at least $E_{\text{low}} = 6.0$ eV, dt-sTD-DFT-xTB spectra are very similar to the sTD-DFT-xTB ($E_{\text{thr}} = 9.6$ eV) one.

Following the discussions on the 1PA of iLOV, we further assess the ASQM approach to evaluate the 2PA spectrum of iLOV with the sTD-DFT-xTB/GBSA(water). **Figure 15** presents the experimental 2PA spectrum of iLOV recorded in water by Homans et al.^[119] in comparison with computed sTD-DFT-xTB/GBSA(water) spectra considering different energy thresholds (6.0, 7.0, 8.0, 9.0, 9.3, 9.5, and 9.6 eV). The experimental 2PA

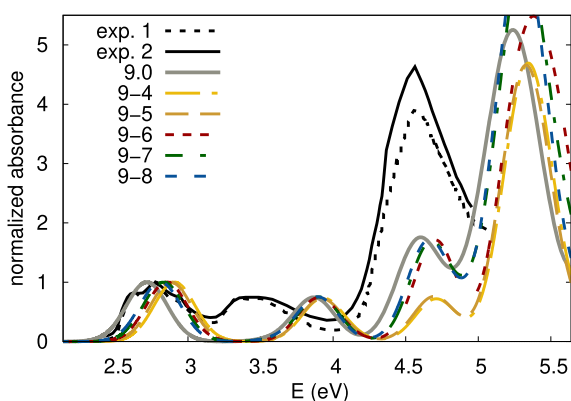


Figure 14. Experimental 1PA spectra of iLOV obtained by Ran et al.^[108] (exp. 1) as well as Torra et al.^[109] (exp. 2) recorded in water in comparison with computed sTD-DFT-xTB/GBSA(water) ($E_{\text{thr}} = 9.0$ eV) spectrum as well as dt-sTD-DFT-xTB/GBSA(water) ones with 9–4, 9–5, 9–6, 9–7, and 9–8 values of $E_{\text{high}} - E_{\text{low}}$. Convolved spectra used a damping factor $\Gamma = 0.2$ eV and no energy shifts were applied.

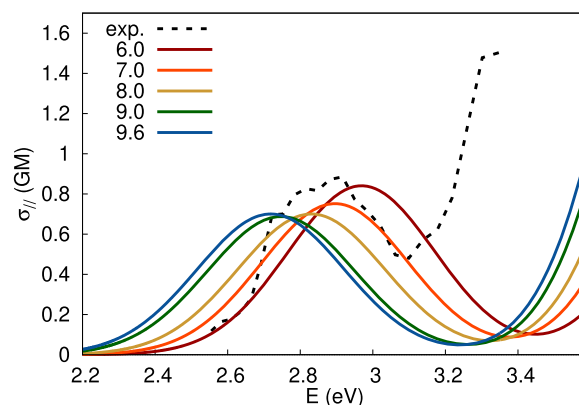


Figure 15. Experimental 2PA spectrum of iLOV recorded by Homans et al.^[119] in water in comparison with computed sTD-DFT-xTB/GBSA(water) spectra considering different energy thresholds (6.0, 7.0, 8.0, 9.0, and 9.6 eV). Convolved spectra used a damping factor $\Gamma = 0.2$ eV and no energy shifts were applied.

spectrum of iLOV^[119] is recorded for a small spectral range from 2.55 to 3.35 eV and shows a maximum at around 2.9 eV (≈ 0.9 GM). The onset of a second peak around 3.35 eV (≈ 1.5 GM) was also recorded. Computed sTD-DFT-xTB spectra are slightly shifted (± 0.1 eV) with respect to experiment but the comparison is striking, very well reproducing the 2PA cross-section of first 2PA band (with $E_{\text{thr}} = 9.6$ eV, $\sigma^{2PA} = 0.70$ GM). All sTD-DFT-xTB spectra reproduce the first experimental peak in position and intensity, while the second peak (which is not clearly resolved in the experiment) is always blue-shifted, as we also observed for computed 1PA spectra (vide supra). The variance between spectra computed with different thresholds is small; therefore, it is possible to go as low as $E_{\text{thr}} = 6.0$ eV for this system because of the small energy range considered by the experiment. For the dt-sTD-DFT-xTB spectra (see **Figure 16** and S4, Supporting Information), a similar picture emerges. Even the smallest tested 9–4 ($E_{\text{high}} - E_{\text{low}}$) thresholds lead to an excellent comparison with respect to experiment and the calculation took less than 3 min.

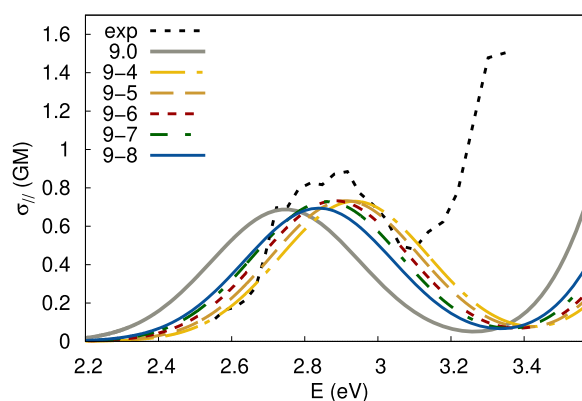


Figure 16. Experimental 2PA spectrum of iLOV recorded by Homans et al.^[119] in water in comparison with the computed sTD-DFT-xTB/GBSA(water) spectrum considering an energy threshold 9.0 eV as well as dt-sTD-DFT-xTB/GBSA(water) spectra with 9–4, 9–5, 9–6, 9–7, and 9–8 values of $E_{\text{high}} - E_{\text{low}}$. Convolved spectra used a damping factor $\Gamma = 0.2$ eV and no energy shifts were applied.

As intermediate conclusions, the ASQM methodology proposed here provided excellent comparisons for the 1- and 2PA spectra of bR and iLOV with respect to experiment. For both proteins, an ASQM workflow was necessary to describe absorption bands involving tryptophan residues, confirming what was already shown by one of us^[93] for PYP. Note that we are at edge of what is computationally feasible with quantum chemistry for such systems. However, including dynamic structural effects could improve further comparisons with respect to experiment as we will show in the next section for FMN in aqueous solution.

4.2. An All-Atom Dynamic Structure QM Protocol Applied to the Flavin Mononucleotide in Water Solution

We could still argue that for FPs, because of the network of non-covalent interactions (including H-bonds), the environment of the chromophore remains rigid enough to justify the use of an ASQM procedure. However, subtle distortions or new configurations of the chromophore can still change the 1- and 2PA spectra dramatically. For solvated dyes, an ASQM approach is certainly not sufficient. Dynamic structural effects need to be included in. In the next section, both ADQM-Boltz. and ADQM-MD approaches are assessed to compute the 1- and 2PA of FMN in aqueous solution. Before applying these ADQM schemes, 1- and 2PA spectra computed in vacuum at the sTD-DFT-xTB level are discussed and compared to reference RI-CC2 ones as well as using implicit solvent models.

Figure 2 presents experimental 1PA spectra of FMN recorded in water, methanol, iLOV, and miniSOG. Note that all spectra were normalized on the first absorption peak. As so, we cannot assess the impact of the surroundings on the molar absorption coefficient. Independent of the environment, the lowest-energy peak always arises around 2.79 eV (445 nm). In the protein surroundings, this absorption band is more vibrationally resolved. Going from water (3.31 eV or 375 nm) to methanol or proteins (3.49 eV or 355 nm) environments, the second signal is shifted by 0.18 eV (20 nm). The energy position of this second absorption band seems dependent on the polarity of the environment. Both proteins influence this peak similarly as methanol, although it was suggested that the dielectric constant of such protein environment (9.2) is way lower than the one from methanol (32.6).^[78,178] The intensity of the third peak in the UV-C region (4.59 eV or 270 nm), traced back to be influenced by the environment,^[108] decreases in water compared to FMN in protein. As we discussed in the previous section, this band is clearly impacted by the absorption of tryptophan residues from the protein, explaining this trend.

Figure 17 compares 1PA spectra of FMN computed in gas phase at RI-CC2/aug-cc-pVDZ and sTD-DFT-xTB levels of theory to experimental spectra recorded in water. It also reports the average sTD-DFT-xTB spectrum on uncorrelated snapshots from a GFN2-xTB MD that was run in vacuum for 11 ns. Note that the 1PA spectrum from Antill et al.^[106] was normalized; therefore, we extrapolated it to match the extinction coefficient of the first absorption band of the 1PA spectrum recorded in water by List et al.^[4] The first experimental absorption band around

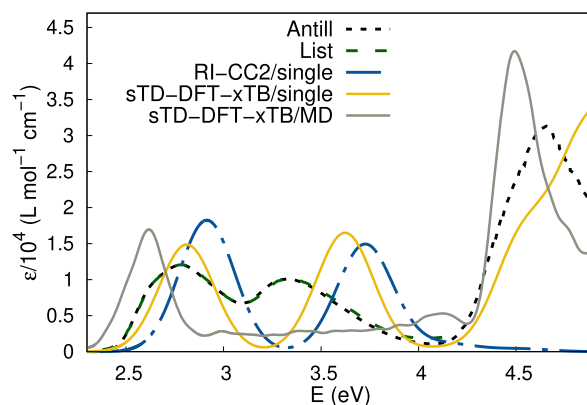


Figure 17. Experimental 1PA spectra of FMN recorded in water by Antill et al.^[106] as well as List et al.^[4] in comparison with computed RI-CC2/aug-cc-pVDZ//*r*²SCAN-3c and sTD-DFT-xTB//*r*²SCAN-3c ($E_{\text{thr}} = 16.0$ eV) spectra in vacuum as well as the average sTD-DFT-xTB spectrum on uncorrelated snapshots of a GFN2-xTB 11 ns-long MD in gas phase. Convolved spectra used a damping factor $\Gamma = 0.2$ eV, while a damping factor $\Gamma = 0.04$ eV was used for spectra computed for each MD snapshot. No energy shifts were applied.

2.79 eV is rather well-reproduced by both RI-CC2 and sTD-DFT-xTB calculations. With respect to RI-CC2, the sTD-DFT-xTB excitation energy is red-shifted by 0.1 eV, and the oscillator strength is 19% smaller. Considering the second absorption band for which experiment showed its sensitivity to the surroundings,^[78,178] as expected, large blue shifts of 0.39 and 0.28 eV with respect to experiment are observed at both RI-CC2 and sTD-DFT-xTB levels of theory, respectively. The average sTD-DFT-xTB spectrum computed from uncorrelated snapshots of a GFN2-xTB MD shows a complete compression of the second absorption band. Globally, RI-CC2 and sTD-DFT-xTB 1PA spectra in vacuum are very similar, justifying the usage of the sTD-DFT-xTB method to compute 1PA spectra for FMN.

Let us assess the ADQM-Boltz. approach to simulate the 1PA spectrum of FMN in aqueous solution. Dynamic structural effects were accounted for implicitly by Boltzmann-averaging 1PA spectra computed at the sTD-DFT-xTB/GBSA(water) for the 23 lowest-energy conformers obtained at the *r*²SCAN-3c level of theory following the CENSO^[124,125] protocol. For the sake of completeness, comparisons were made using the *r*²SCAN-3c lowest-energy conformer to compute excited states with both RI-CC2 and sTD-DFT-xTB levels using implicit solvent models (DCOSMO-RS and GBSA, respectively). **Figure 18** compares experimental 1PA spectra recorded in water^[4,106] to spectra computed at these levels of theory as well as the Boltzmann-averaged sTD-DFT-xTB/GBSA(water) spectrum obtained by following the ADQM-Boltz. workflow. It becomes very clear that either using the lowest-energy conformer only or the ensemble of conformers barely show any difference on the sTD-DFT-xTB 1PA spectra. Note that the lowest-energy conformer has only a Boltzmann weight of around 25%. The position of the first energy band at 2.79 eV is well-reproduced by both RI-CC2 and sTD-DFT-xTB methods with only 0.08 and 0.04 eV of difference, respectively. Both RI-CC2 and sTD-DFT-xTB schemes overestimate the intensity by a factor of 2.1 and 1.4. The second peak is blue-shifted by both methods

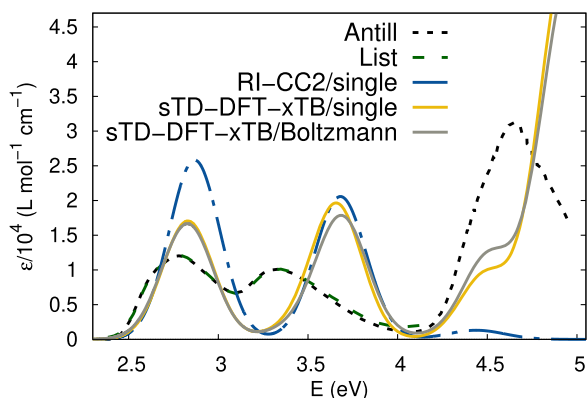


Figure 18. Experimental 1PA spectra of FMN recorded in water by Antill et al.^[106] as well as List et al.^[4] in comparison with computed RI-CC2/aug-cc-pVDZ//*r*²SCAN-3c and sTD-DFT-xTB//*r*²SCAN-3c ($E_{\text{thr}} = 16.0$ eV) spectra using implicit solvent models (DCOSMO-RS and GBSA, respectively) for water as well as the Boltzmann-averaged sTD-DFT-xTB/GBSA(water) spectrum for the conformer ensemble obtained at room temperature with the *r*²SCAN-3c method. Convolved spectra used a damping factor $\Gamma = 0.2$ eV. No energy shifts were applied.

(RI-CC2/single: 3.68 eV and sTD-DFT-xTB: 3.66 eV) compared to the experimental peak maximum at 3.31 eV. The intensity of the second peak is overestimated by both methods with respect to experiment. Nevertheless, RI-CC2 reproduces the trend between the first and second peak intensities, while sTD-DFT-xTB inverts this relation. These results are not encouraging to employ the ADQM-Boltz. workflow for this very case.

The ADQM-MD protocol is applied to compute the 1PA spectrum of FMN, explicitly accounting for water molecules and dynamic structural effects. **Figure 19** compares experimental 1PA spectra recorded in water^[4,106] to average dt-sTD-DFT/GBSA(water) spectra on uncorrelated snapshots extracted from 1 ns-long GFN2-xTB MD simulations accounting explicitly for solvent molecules using solvation spheres of increasingly large radii of 7, 9, 11 and 12 Å (see Figure S6, Supporting Information) as well

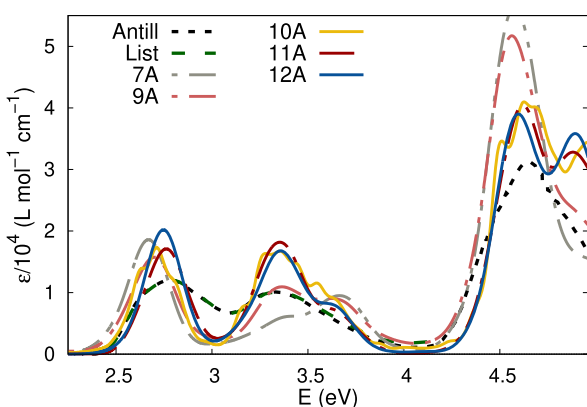


Figure 19. Experimental 1PA spectra of FMN recorded in water by Antill et al.^[106] as well as List et al.^[4] in comparison with computed average dt-sTD-DFT-xTB/GBSA(water) spectra on uncorrelated snapshots extracted from 1 ns-long GFN2-xTB MD simulations accounting explicitly for solvent molecules using solvation spheres of increasingly large radii of 7, 9, 11, and 12 Å. Note that when considering a sphere of 10 Å, a longer 3 ns-long GFN2-xTB MD simulation was run. Convolved spectra used a damping factor $\Gamma = 0.1$ eV. No energy shifts were applied.

as from a 3 ns-long GFN2-xTB MD simulation considering a solvation sphere of 10 Å. In consequence, different amounts of water molecules (124,235,307,386 and 483 for spheres of 7, 9, 10, 11, and 12 Å, respectively) were considered to assess the convergence of the ADQM-MD procedure. Note that spectra were computed with the dt-sTD-DFT/GBSA(water) method to reduce the computational cost using $E_{\text{high}} - E_{\text{low}}$ values of 11–8 eV, except for the dt-sTD-DFT-xTB(12 Å)/GBSA(water) for which we used 10–8 eV. When considering a sphere of 12 Å, the sTD-DFT-xTB calculation with $E_{\text{thr}} = 10$ eV took 46 min (on the same computer node specified above), while the dt-sTD-DFT-xTB with $E_{\text{high}} - E_{\text{low}}$ values of 10–8 eV took only 16 min without modifying much the computed spectrum, justifying this choice. The comparison with respect to experiment is more than striking, knowing that no energy shifts were applied to MD-average dt-sTD-DFT-xTB spectra. While simulated spectra are still arbitrary damped, energy position, intensities, and broadening of the three main experimental peaks are well-reproduced. Considering the dt-sTD-DFT-xTB(12 Å)/GBSA(water) MD-average spectrum, for example, the first absorption band occurs at 2.75 eV (451 nm, exp: 2.79 eV or 445 nm), and the second peak at 3.36 eV (369 nm, exp: 3.31 eV or 375 nm) with intensities of 20 194 and 16 780 L mol^{−1} cm^{−1} (exp: 12 043 and 10 122 L mol^{−1} cm^{−1}), respectively. The reproduction of the second experimental absorption band is very sensitive to the size of the solvation sphere included into the ADQM-MD procedure. At least a solvation sphere of 9 Å is needed to describe properly this second peak, but for better describing the third absorption band, a minimum 10 Å solvation sphere should be used. Only the ADQM-MD was able to account for the very specific broadening of the second absorption band. All in all, the ADQM-MD procedure that accounts explicitly for the environment and explicitly for dynamic structural effects correctly described the first three main absorption bands of the 1PA spectrum of FMN in aqueous solution.

To understand what is responsible for the broadening of the second absorption band, NTOs were calculated from two representative snapshots of the GFN2-xTB MD. **Figure 20** presents NTOs computed at dt-sTD-DFT-xTB/GBSA(water) (10 – 8 eV) for two excited states of FMN in explicit water (12 Å) that contribute to the second absorption band. We already saw for iLOV that the second absorption band is due to a $\pi \rightarrow \pi^*$ transition on the isalloxazine ring. For FMN in water, the dynamic nature of the environment directly impacts this transition. For both NTO transitions, the hole orbital is a linear combination of the π orbital from the isalloxazine ring and nearby *n* orbitals from oxygen located on water molecules as well as from the tail of FMN chromophore. Particle orbitals from both NTOs are mostly π^* orbitals located on the isalloxazine ring. In both cases, the $\pi \rightarrow \pi^*$ transition presents a CT character modulated by the interaction of the chromophore with its nearby surroundings. Thus, the dynamic interaction of the HPO_4^- group as well as water molecules with the isalloxazine ring is crucial to correctly reproduce the second absorption band and its broadening. This advocates for the use of the ADQM-MD scheme.

Let us now focus on 2PA of FMN in water, assessing both ADQM schemes for computing such spectra. As before, we firstly

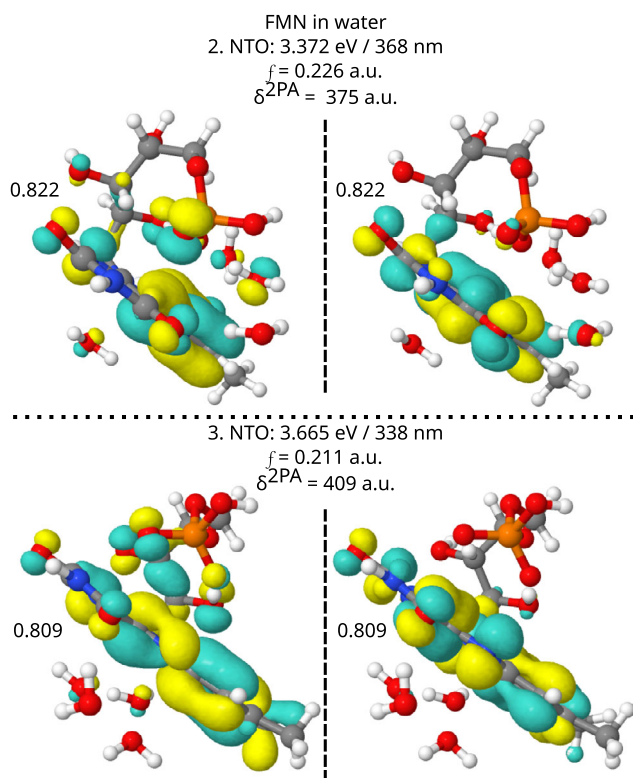


Figure 20. From two different snapshots of the GFN2-xTB MD, NTOs computed at dt-sTD-DFT-xTB/GBSA(water) (10 – 8 eV) for two excited states of FMN in explicit water (12 Å) that contribute to the second absorption band. Hole and particle NTOs are separated by dashed lines. Weights are provided for each NTO pair. The excitation energy, oscillator strength, and 2 PA strength are also given for each state. Isovalue = 0.03.

discuss FMN results obtained in vacuum. **Figure 21** compares two experimental 2PA spectra of FMN recorded in water^[4,119] to computed RI-CC2/aug-cc-pVDZ and sTD-DFT-xTB ($E_{\text{thr}} = 16.0$ eV) 2PA spectra calculated in vacuum as well as the average sTD-DFT-xTB spectrum on uncorrelated snapshots from a GFN2-xTB 11 ns-long MD in gas phase.

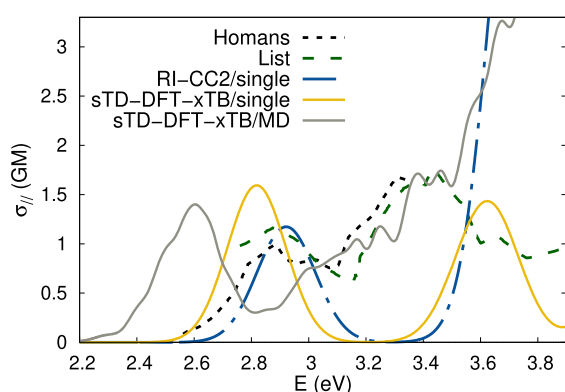


Figure 21. Experimental 2PA spectra of FMN recorded in water by Homans et al.^[119] as well as List et al.^[4] in comparison with computed RI-CC2/aug-cc-pVDZ/ r^2 SCAN-3c and sTD-DFT-xTB/ r^2 SCAN-3c ($E_{\text{thr}} = 16.0$ eV) spectra in vacuum as well as the average sTD-DFT-xTB spectrum on uncorrelated snapshots from a GFN2-xTB 11 ns-long MD in gas phase. Convolved spectra used a damping factor $\Gamma = 0.1$ eV, while a damping factor $\Gamma = 0.02$ eV was used for spectra computed for each MD snapshot. No energy shifts were applied.

MD in gas phase. Both experimental 2PA spectra were recorded for a narrow energy range. Combining both experimental data leads to an energy window larger by 0.5 eV than for iLOV. The experimental 2PA spectra (Figure 21) feature two main peaks around 2.88 and 3.43 eV. Comparing them to the RI-CC2 spectrum, we observe a good agreement for the first experimental peak energy position with only 0.04 eV of difference with respect to experiment but 8% of overestimation for the 2PA cross section. The 2PA cross section of the second absorption band is nine times larger than the experimental one (1.77 GM) while slightly underestimated at the sTD-DFT-xTB level. The average sTD-DFT-xTB spectrum computed from uncorrelated snapshots of a GFN2-xTB MD improves the comparison with respect to experiment for the second 2PA band but the onset of the band continues to climb at higher energies. Note that the sTD-DFT-xTB calculation on the optimized r^2 SCAN-3c geometry took only 4 s, while the RI-CC2 calculation needed 130 h (wall time) for the computation of both ground and excited states as well as 2PA strengths.

As 1PA and 2PA spectra are involving same transitions, we can expect that the ADQM-Boltz. shall not be particularly helpful to cope with dynamic structural effects. **Figure 22** confirms this, comparing experimental 2PA spectra recorded in water^[4,119] to spectra computed at these levels of theory as well as the Boltzmann-averaged sTD-DFT-xTB/GBSA(water) spectrum obtained by following the ADQM-Boltz. workflow.

Finally, the ADQM-MD workflow is applied to evaluate the 2PA of FMN in water. The same procedure is followed as for 1PA. **Figure 23** compares experimental 2PA spectra of FMN recorded in water by Homans et al.^[119] as well as List et al.^[4] to average dt-sTD-DFT-xTB/GBSA(water) spectra on uncorrelated snapshots extracted from 1 ns-long GFN2-xTB MD simulations accounting explicitly for solvent molecules using solvation spheres of increasingly large radii of 7, 9, 11, and 12 Å as well as from a 3 ns-long GFN2-xTB MD simulation considering a solvation sphere of 10 Å. For 2PA, the comparison with respect to

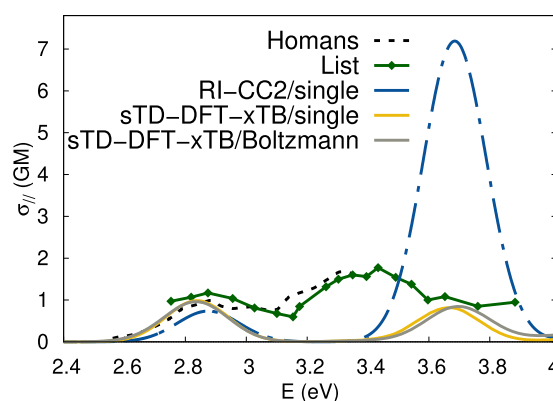


Figure 22. Experimental 2PA spectra of FMN recorded in water by Homans et al.^[119] as well as List et al.^[4] in comparison with computed RI-CC2/aug-cc-pVDZ/ r^2 SCAN-3c and sTD-DFT-xTB/ r^2 SCAN-3c ($E_{\text{thr}} = 16.0$ eV) spectra using implicit solvent models (DCOSMO-RS and GBSA, respectively) for water as well as the Boltzmann-averaged sTD-DFT-xTB/GBSA(water) spectrum for the conformer ensemble obtained at room temperature with the r^2 SCAN-3c method. Convolved spectra used a damping factor $\Gamma = 0.1$ eV. No energy shifts were applied.

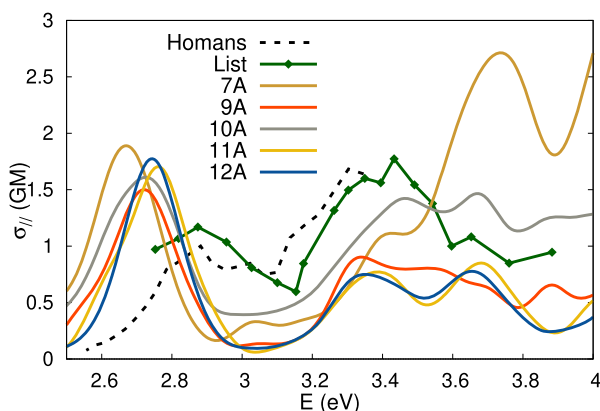


Figure 23. Experimental 2PA spectra of FMN recorded in water by Homans et al.^[119] as well as List et al.^[4] in comparison with computed average dt-sTD-DFT-xTB/GBSA(water) spectra on uncorrelated snapshots extracted from 1 ns-long GFN2-xTB MD simulations accounting explicitly for solvent molecules using solvation spheres of increasingly large radii of 7, 9, 11, and 12 Å. Note that when considering a sphere of 10 Å, a longer 3 ns-long GFN2-xTB MD simulation was run. Convolved spectra used a damping factor $\Gamma = 0.05$ eV. No energy shifts were applied.

experiment is striking especially when using at least a solvation sphere of 11 Å. As for 1PA, the nonhomogeneous broadening of the second absorption band featuring two peaks between 3.2 and 3.8 eV is well-captured by the ADQM-MD methodology due to the dynamic interaction of the isoalloxazine ring with its nearby surroundings. Figure S7, Supporting Information, provides more details about the influence of longer MD simulations for the 10 Å system. In 2014, List et al.^[4] computed the 2PA spectrum of FMN in aqueous solution at the QM/MM level (PE-CAM-B3LYP) considering only FMN at the QM level. The average spectrum was computed for 50 snapshots of a classical MD simulation followed by QM/MM (B3LYP/6-31+G**//OPLS2005) optimizations. **Figure 24** presents this spectrum in comparison with experiment

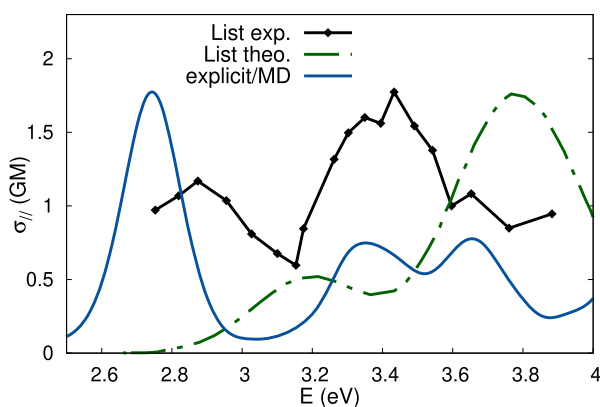


Figure 24. Experimental 2PA spectrum of FMN recorded in water by List et al.^[4] in comparison with computed average dt-sTD-DFT-xTB/GBSA(water) spectrum on uncorrelated snapshots extracted from 1 ns-long GFN2-xTB MD simulations accounting explicitly for solvent molecules using a solvation sphere of 12 Å, as well as a QM-MM average spectrum computed by List et al.^[4] at the PE-CAM-B3LYP/cc-pVDZ+ level on 50 snapshots of a classical MD simulation followed by QM/MM (B3LYP/6-31+G**//OPLS2005) optimizations (only FMN was accounted for into the QM layer). Convolved spectra used a damping factor $\Gamma = 0.05$ eV (this work) and $\Gamma = 0.35$ eV (List et al.^[4]). No energy shifts were applied.

and our average dt-sTD-DFT/GBSA(water) spectrum considering a solvation sphere of 12 Å. Note that the amount of explicit water molecules accounted for at the MM level by List et al.^[4] is comparable to our. The QM/MM spectrum^[4] is shifted by 0.4 eV with respect to experiment and does not capture the inhomogeneous broadening of the second experimental band, but reproduces the energy separation between both maxima, as well as their relative intensities. In comparison, our average dt-sTD-DFT/GBSA(water) 2PA spectrum provides peak energy positions very similar to experimental one and is able to recover the two peaks featured by the second absorption band. This result could be improved by employing other simplified schemes such as the XsTD-DFT method^[91,92] as one of us did to compute 1PA and CD spectra of PYP using CAM-B3LYP.^[93] Nevertheless, these results strongly advocates for the use of the ADQM-MD scheme to evaluate 1- and 2PA spectra of dyes in solution.

5. Conclusions

In this study, different AQM workflows were assessed to compute 1- and 2PA of realistic systems. While neglecting dynamic structural effects but still explicitly accounting for the chromophore surroundings, the performance of ASQM scheme was evaluated for two extremely challenging systems: bR (≈ 3850 atoms) and iLOV (≈ 2000 atoms), pushing the computational boundaries of this methodology. It was also the opportunity to test with success the dt-sTD-DFT scheme to reduce the computational cost of the sTD-DFT calculations. Results showed that to reproduce experimental 1- and 2PA spectra, it is important to account for the whole protein at a QM level because absorbing amino acid residues such as tryptophan or tyrosine have non-negligible impact on the higher-energy part of the spectra (>4 eV). All spectral features were not fully captured by the ASQM scheme but still comparisons with respect to experiment were rather reasonable. Future studies are planned to further reduce the gap between simulated and real systems by including dynamic structural effects using an ADQM procedure for proteins.

Furthermore, two ADQM schemes were tested to evaluate 1- and 2PA of FMN in aqueous solution. While the ADQM-Boltz. was not improving results with respect to the single structure approach, the ADQM-MD provided striking agreements with respect to experiment for both 1- and 2PA when including solvation spheres of at least 11 Å. In both spectra, the inhomogeneous broadening of the second absorption band results from a CT excitation modulated by the dynamic interaction of the HPO_4^- group as well as water molecules with the isoalloxazine ring. These interactions can only be coped with by an ADQM scheme. Note that the same feature is also present in iLOV spectra for which FMN is the chromophore and could be capture using such scheme. We clearly observed that explicitly accounting for both solvent and dynamic structural effects has a far more dramatic impact on spectra than simply choosing between both RI-CC2 or sTD-DFT-xTB methods in a single structure approach. Thus, these results clearly advocate for the use of an ADQM-MD workflow to compute 1- and 2PA of molecules in solution and to provide quantitative agreement with respect to experiment. Clearly,

sQC methods such as the (dt-)sTD-DFT-xTB schemes are crucial for AQM workflows. The establishment of AQM methodologies with the help of sQC methods opens the gate for a plethora of applications on realistic systems in reasonable computational times.

Acknowledgements

S.L. and M.de W. thank Prof. Dr. Stefan Grimme for his support all along these years. M.M. thanks the F.R.S.-F.N.R.S. for her Research Fellow position. Some of the computational resources were provided by the supercomputing facilities at the UCLouvain (CISM) and the Consortium des Équipements de Calcul Intensif en Fédération Wallonie Bruxelles (CÉCI) funded by the Fond de la Recherche Scientifique de Belgique (F.R.S.-FNRS) under convention 2.5020.11 and the Walloon Region.

Conflict of Interest

The authors declare no conflict of interest.

Author Contributions

Sarah Löffelsender: data curation (lead); formal analysis (lead); investigation (lead); methodology (equal); writing—original draft (equal); writing—review & editing (equal). **Marilù G. Maraldi:** writing—original draft (equal); writing—review & editing (supporting). **Marc de Wergifosse:** conceptualization (lead); funding acquisition (lead); methodology (equal); software (lead); supervision (lead); writing—original draft (equal); writing—review & editing (equal).

Data Availability Statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Keywords: all-atom quantum mechanics · realistic system · simplified quantum chemistry · sTD-DFT · two-photon absorption

- [1] T. J. Zuehlsdorff, C. M. Isborn, *Int. J. Quantum Chem.* **2019**, *119*, e25719.
- [2] T. Giovannini, C. Cappelli, *Chemical Communications* **2023**, *59*, 5644.
- [3] T. Andruniów, M. Olivucci, in *Volume 31 Of Challenges And Advances In Computational Chemistry And Physics* Springer, New York **2021**.
- [4] N. H. List, F. M. Pimenta, L. Holmegaard, R. L. Jensen, M. Etzerodt, T. Schwabe, J. Kongsted, P. R. Ogilby, O. Christiansen, *Phys. Chem. Chem. Phys.* **2014**, *16*, 9950.
- [5] T. Giovannini, P. Lafiosca, C. Cappelli, *J. Chem. Theory Comput.* **2017**, *13*, 4854.
- [6] T. Giovannini, F. Egidi, C. Cappelli, *Chem. Soc. Rev.* **2020**, *49*, 5664.
- [7] J. M. Milanese, M. R. Provorse, E. J. Alameda, C. M. Isborn, *J. Chem. Theory Comput.* **2017**, *13*, 2159.
- [8] O. Valsson, P. Campomanes, I. Tavernelli, U. Rothlisberger, C. Filippi, *J. Chem. Theory Comput.* **2013**, *9*, 2441.
- [9] A. Warshel, M. Levitt, *J. Mol. Biol.* **1976**, *103*, 227.
- [10] M. Svensson, S. Humbel, R. D. Froese, T. Matsubara, S. Sieber, K. Morokuma, *J. Phys. Chem.* **1996**, *100*, 19357.
- [11] F. Maseras, K. Morokuma, *J. Comput. Chem.* **1995**, *16*, 1170.
- [12] S. Humbel, S. Sieber, K. Morokuma, *J. Chem. Phys.* **1996**, *105*, 1959.
- [13] H. M. Senn, W. Thiel, *Angew. Chem. Int. Ed.* **2009**, *48*, 1198.
- [14] P. Sherwood, A. H. de Vries, M. F. Guest, G. Schreckenbach, C. R. A. Catlow, S. A. French, A. A. Sokol, S. T. Bromley, W. Thiel, A. J. Turner, S. Billeter, F. Terstegen, S. Thiel, J. Kendrick, S. C. Rogers, J. Casci, M. Watson, F. King, E. Karlsen, M. Sjøvoll, A. Fahmi, A. Schäfer, C. Lennartz, *J. Mol. Struct.: Theochem* **2003**, *632*, 1.
- [15] H. Lin, D. G. Truhlar, *Theor. Chem. Acc.* **2007**, *117*, 185.
- [16] L. Goletto, S. Gómez, J. H. Andersen, H. Koch, T. Giovannini, *Phys. Chem. Chem. Phys.* **2022**, *24*, 27866.
- [17] C. Cappelli, *Int. J. Quantum Chem.* **2016**, *116*, 1532.
- [18] C. Reichardt, T. Welton, in *Solvents and Solvent Effects in Organic Chemistry*, John Wiley & Sons, Hoboken, NJ **2011**.
- [19] L. Jensen, P. T. Van Duijnen, J. G. Snijders, *J. Chem. Phys.* **2003**, *119*, 12998.
- [20] C. Curutchet, A. Munoz-Losa, S. Monti, J. Kongsted, G. D. Scholes, B. Mennucci, *J. Chem. Theory Comput.* **2009**, *5*, 1838.
- [21] P. Reinholdt, J. Kongsted, J. M. H. Olsen, *J. Phys. Chem. Lett.* **2017**, *8*, 5949.
- [22] D. Loco, É. Polack, S. Caprasecca, L. Lagardere, F. Lipparini, J.-P. Piquemal, B. Mennucci, *J. Chem. Theory Comput.* **2016**, *12*, 3654.
- [23] M. S. Gordon, L. Slipchenko, H. Li, J. H. Jensen, in *Annual Reports in Computational Chemistry*, (Ed: D. Spellmeyer, R. Wheeler) Vol3, Elsevier, Amsterdam **2007**, pp. 177–193.
- [24] A. DeFusco, N. Minezawa, L. V. Slipchenko, F. Zahariev, M. S. Gordon, *J. Phys. Chem. Lett.* **2011**, *2*, 2184.
- [25] L. V. Slipchenko, *J. Phys. Chem. A* **2024**, *128*, 656.
- [26] C. I. Viquez Rojas, L. V. Slipchenko, *J. Chem. Theory Comput.* **2020**, *16*, 6408.
- [27] D. Kosenkov, L. V. Slipchenko, *J. Phys. Chem. A* **2011**, *115*, 392.
- [28] A. Severo Pereira Gomes, C. R. Jacob, *Annu. Rep. Progr. Chem. Sect. C: Phys. Chem.* **2012**, *108*, 222.
- [29] E. D. Larsson, P. Reinholdt, E. D. Hedegård, J. Kongsted, *J. Phys. Chem. B* **2023**, *127*, 9905.
- [30] M. Jansen, P. Reinholdt, E. D. Hedegård, C. König, *J. Phys. Chem. A* **2023**, *127*, 5689.
- [31] S. Alexander, G. E. Koehl, M. Hirschberg, E. K. Geissler, P. Friedl, *Histochem. Cell Biol.* **2008**, *130*, 1147.
- [32] J. Paoli, M. Smedh, M. B. Ericson, in *Seminars in Cutaneous Medicine and Surgery*, Vol. 28, WB Saunders, Philadelphia, PA **2009**, pp. 190–195.
- [33] K. Tanaka, Y. Toiyama, Y. Okugawa, M. Okigami, Y. Inoue, K. Uchida, T. Araki, Y. Mohri, A. Mizoguchi, M. Kusonoki, *Am. J. Transl. Res.* **2014**, *6*, 179.
- [34] K. W. Dunn, R. M. Sandoval, K. J. Kelly, P. C. Dagher, G. A. Tanner, S. J. Atkinson, R. L. Bacallao, B. A. Molitoris, *Am. J. Physiol.-Cell Physiol.* **2002**, *283*, C905.
- [35] A. Diaspro, G. Chirico, M. Collini, *Q. Rev. Biophys.* **2005**, *38*, 97.
- [36] S. Ashworth, R. Sandoval, G. Tanner, B. Molitoris, *Kidney Int.* **2007**, *72*, 416.
- [37] K. Svoboda, R. Yasuda, *Neuron* **2006**, *50*, 823.
- [38] R. K. Benninger, D. W. Piston, *Curr. Protoc. Cell Biol.* **2013**, *59*, 4.
- [39] C. Huang, J. R. Maxey, S. Sinha, J. Savall, Y. Gong, M. J. Schnitzer, *Nat. Commun.* **2018**, *9*, 872.
- [40] J. Demas, J. Manley, F. Tejera, K. Barber, H. Kim, F. M. Traub, B. Chen, A. Vaziri, *Nat. Methods* **2021**, *18*, 1103.
- [41] F. Helmchen, M. S. Fee, D. W. Tank, W. Denk, *Neuron* **2001**, *31*, 903.
- [42] W. Zong, H. A. Obenaus, E. R. Skytøen, H. Eneqvist, N. L. de Jong, R. Vale, M. R. Jorge, M.-B. Moser, E. I. Moser, *Cell* **2022**, *185*, 1240.
- [43] T. Fischer, N. A. Hampp, *Biophys. J.* **2005**, *89*, 1175.
- [44] B. Yao, Z. Ren, N. Menke, Y. Wang, Y. Zheng, M. Lei, G. Chen, N. Hampp, *Appl. Opt.* **2005**, *44*, 7344.
- [45] D. Oesterhelt, C. Bräuchle, N. Hampp, *Q. Rev. Biophys.* **1991**, *24*, 425.
- [46] J. Olsen, P. Jørgensen, *J. Chem. Phys.* **1985**, *82*, 3235.
- [47] T. Helgaker, S. Coriani, P. Jørgensen, K. Kristensen, J. Olsen, K. Ruud, *Chem. Rev.* **2012**, *112*, 543.
- [48] M. J. Paterson, O. Christiansen, F. Pawłowski, P. Jørgensen, C. Hättig, T. Helgaker, P. Salek, *J. Chem. Phys.* **2006**, *124*, 054322.
- [49] K. D. Nanda, A. I. Krylov, *J. Chem. Phys.* **2015**, *142*.
- [50] C. Hättig, O. Christiansen, P. Jørgensen, *J. Chem. Phys.* **1998**, *108*, 8355.
- [51] D. H. Fries, C. Hättig, K. Ruud, *Phys. Chem. Chem. Phys.* **2012**, *14*, 1175.
- [52] M. A. Salem, A. Brown, *J. Chem. Theory Comput.* **2014**, *10*, 3260.
- [53] M. T. P. Beerepoot, D. H. Fries, K. Ruud, *Phys. Chem. Chem. Phys.* **2014**, *16*, 5958.
- [54] M. T. Beerepoot, D. H. Fries, N. H. List, J. Kongsted, K. Ruud, *Phys. Chem. Chem. Phys.* **2015**, *17*, 19306.
- [55] M. T. P. Beerepoot, M. M. Alam, J. Bednarska, W. Bartkowiak, K. Ruud, R. Zalesny, *J. Chem. Theory Comput.* **2018**, *14*, 3677.
- [56] I. A. Elayan, A. Brown, *J. Phys. Chem. A* **2024**, *128*, 7511.

- [57] M. Cholu, M. M. Alam, M. T. P. Beerepoot, S. P. Sitkiewicz, E. Matito, K. Ruud, R. Zalesny, *J. Chem. Theory Comput.* **2022**, *18*, 1046.
- [58] K. Ahmadzadeh, X. Li, Z. Rinkevicius, P. Norman, R. Zalesny, *J. Phys. Chem. Lett.* **2024**, *15*, 969.
- [59] C. Naim, R. Zalesny, D. Jacquemin, *J. Chem. Theory Comput.* **2024**, *20*, 9093.
- [60] P. Salek, O. Vahtras, J. Guo, Y. Luo, T. Helgaker, H. Ågren, *Chem. Phys. Lett.* **2003**, *374*, 446.
- [61] F. Zahariev, M. S. Gordon, *J. Chem. Phys.* **2014**, *140*, 18A523.
- [62] S. M. Parker, D. Rappoport, F. Furche, *J. Chem. Theory Comput.* **2018**, *14*, 807.
- [63] M. Paterson, J. Kongsted, O. Christiansen, K. Mikkelsen, C. Nielsen, *J. Chem. Phys.* **2006**, *125*, 184501.
- [64] D. L. Silva, N. A. Murugan, J. Kongsted, Z. Rinkevicius, S. Canuto, H. Ågren, *J. Phys. Chem. B* **2012**, *116*, 8169.
- [65] J. Olesiak-Banska, K. Matczyszyn, R. Zalesny, N. A. Murugan, J. Kongsted, H. Ågren, W. Bartkowiak, M. Samoc, *J. Phys. Chem. B* **2013**, *117*, 12013.
- [66] K. Matczyszyn, J. Olesiak-Banska, K. Nakatani, P. Yu, N. A. Murugan, R. Zalesny, A. Roztoczyńska, J. Bednarska, W. Bartkowiak, J. Kongsted, H. Ågren, M. Samoc, *J. Phys. Chem. B* **2015**, *119*, 1515.
- [67] T. N. Ramos, L. R. Franco, D. L. Silva, S. Canuto, *J. Chem. Phys.* **2023**, *159*, 024309.
- [68] M. Wielgus, R. Zalesny, N. A. Murugan, J. Kongsted, H. Ågren, M. Samoc, W. Bartkowiak, *ChemPhysChem* **2013**, *14*, 3731.
- [69] N. Arul Murugan, J. Kongsted, Z. Rinkevicius, K. Aidas, K. V. Mikkelsen, H. Ågren, *Phys. Chem. Chem. Phys.* **2011**, *13*, 12506.
- [70] A. H. Steindal, J. M. H. Olsen, K. Ruud, L. Frediani, J. Kongsted, *Phys. Chem. Chem. Phys.* **2012**, *14*, 5440.
- [71] N. A. Murugan, R. Zalesny, *J. Chem. Inf. Model.* **2020**, *60*, 3854.
- [72] M. Rossano-Tapia, J. M. H. Olsen, A. Brown, *Front. Mol. Biosci.* **2020**, *7*.
- [73] K. Sneskov, J. M. H. Olsen, T. Schwabe, C. Hättig, O. Christiansen, J. Kongsted, *Phys. Chem. Chem. Phys.* **2013**, *15*, 7567.
- [74] A. A. Attia, P. Reinholdt, J. Kongsted, *Adv. Theory Simul.* **2021**, *4*, 2000294.
- [75] J. Olsen, N. List, K. Kristensen, J. Kongsted, *J. Chem. Theory Comput.* **2015**, *11*, 1832.
- [76] K. Schwinn, N. Ferré, M. Huix-Rotllant, *Phys. Chem. Chem. Phys.* **2020**, *22*, 12447.
- [77] M. P. Kabir, Y. Orozco-Gonzalez, S. Gozem, *Phys. Chem. Chem. Phys.* **2019**, *21*, 16526.
- [78] R. K. Kar, A.-F. Miller, M.-A. Mrogiński, *WIREs Comput. Mol. Sci.* **2022**, *12*, e1541.
- [79] G. Palczewska, F. Vinberg, P. Stremplewski, M. P. Bircher, D. Salom, K. Komar, J. Zhang, M. Cascella, M. Wojtkowski, V. J. Kefalov, K. Palczewski, *Proc. Natl. Acad. Sci.* **2014**, *111*, E5445.
- [80] S. Grimme, *J. Chem. Phys.* **2013**, *138*, 244104.
- [81] C. Bannwarth, S. Grimme, *Comput. Theor. Chem.* **2014**, 1040–1041, 45.
- [82] S. Grimme, C. Bannwarth, *J. Chem. Phys.* **2016**, *145*, 054103.
- [83] M. de Wergifosse, S. Grimme, *J. Chem. Phys.* **2018**, *149*, 024108.
- [84] M. de Wergifosse, S. Grimme, *J. Chem. Phys.* **2019**, *150*, 094112.
- [85] M. de Wergifosse, C. Bannwarth, S. Grimme, *J. Phys. Chem. A* **2019**, *123*, 5815.
- [86] M. de Wergifosse, J. Seibert, S. Grimme, *J. Chem. Phys.* **2020**, *153*, 084116.
- [87] J. Seibert, B. Champagne, S. Grimme, M. de Wergifosse, *J. Chem. Phys. B* **2020**, *124*, 2568.
- [88] P. Beaujean, B. Champagne, S. Grimme, M. de Wergifosse, *J. Chem. Phys. Lett.* **2021**, *12*, 9684.
- [89] M. de Wergifosse, S. Grimme, *J. Phys. Chem. A* **2021**, *125*, 3841.
- [90] M. de Wergifosse, P. Beaujean, S. Grimme, *J. Phys. Chem. A* **2022**, *126*, 7534.
- [91] S. Löffelsender, P. Beaujean, M. de Wergifosse, *WIREs Comput. Mol. Sci.* **2024**, *14*, e1695.
- [92] M. de Wergifosse, S. Grimme, *J. Chem. Phys.* **2024**, *160*, 204110.
- [93] M. de Wergifosse, *J. Phys. Chem. Lett.* **2024**, *15*, 12628.
- [94] D. Oesterhelt, W. Stoekenius, *Nat. New Biol.* **1971**, *233*, 149.
- [95] U. Haupts, J. Tittor, D. Oesterhelt, *Annu. Rev. Biophys.* **1999**, *28*, 367.
- [96] M. P. Krebs, T. A. Isenbarger, *Biochim. et Biophys. Acta Bioenergetics* **2000**, *1460*, 15.
- [97] J. K. Lanyi, *Annu. Rev. Physiol.* **2004**, *66*, 665.
- [98] K. Hayashi, M. Tanigawara, J. Kishikawa, *Biophysics* **2012**, *8*, 67.
- [99] K. Yanai, K. Ishimura, A. Nakayama, J. Hasegawa, *J. Chem. Theory Comput.* **2018**, *14*, 3643.
- [100] G. N. Kovács, J.-P. Colletier, M. L. Grünbein, Y. Yang, T. Stensitzki, A. Batyuk, S. Carbajo, R. B. Doak, D. Ehrenberg, L. Foucar, R. Gasper, A. Gorel, M. Hilpert, M. Kloos, J. E. Koglin, J. Reinsteint, C. M. Roome, R. Schlesinger, M. H. Seaberg, R. L. Shoeman, M. Stricker, S. Boutet, S. Haacke, J. Heberle, K. Heyne, T. Domratheva, T. R. M. Barends, I. Schlichting, *Nat. Commun.* **2019**, *10*, 3177.
- [101] S. Hayashi, E. Tajkhorshid, E. Pebay-Peyroula, A. Royant, E. M. Landau, J. Navarro, K. Schulten, *J. Phys. Chem. B* **2001**, *105*, 10124.
- [102] R. R. Birge, C.-F. Zhang, *J. Chem. Phys.* **1990**, *92*, 7178.
- [103] M. B. Masthay, D. M. Sammeth, M. C. Helvenston, C. B. Buckman, W. Li, M. J. Cde-Baca, J. T. Kofron, *J. Am. Chem. Soc.* **2002**, *124*, 3418.
- [104] D. Rhinow, M. Imhof, I. Chizhik, R.-P. Baumann, N. Hampf, *J. Phys. Chem. B* **2012**, *116*, 7455.
- [105] S. Chapman, C. Faulkner, E. Kaiserli, C. Garcia-Mata, E. I. Savenkov, A. G. Roberts, K. J. Oparka, J. M. Christie, *Proc. Natl. Acad. Sci.* **2008**, *105*, 20038.
- [106] L. M. Antill, S.-Y. Takizawa, S. Murata, J. R. Woodward, *Mol. Phys.* **2019**, *117*, 2594.
- [107] S. S. Daz, H. Al-Zubaidi, A. C. Ross-Obare, S. O. Obare, *Phys. Sci. Rev.* **2020**, *5*, 2020007.
- [108] X. Ran, Q. Zhang, Y. Zhang, J. Chen, Z. Wei, Y. He, L. Guo, *Molecules* **2020**, *25*, 2544.
- [109] J. Torra, C. Lafaye, L. Signor, S. Aumonier, C. Flors, X. Shu, S. Nonell, G. Gotthard, A. Royant, *Sci. Rep.* **2019**, *9*, 2428.
- [110] S. Salzmann, M. R. Silva-Junior, W. Thiel, C. M. Marian, *J. Phys. Chem. B* **2009**, *113*, 15610.
- [111] C. Neiss, P. Saalfrank, M. Parac, S. Grimme, *J. Phys. Chem. A* **2003**, *107*, 140.
- [112] R. J. Stanley, H. Jang, *J. Phys. Chem. A* **1999**, *103*, 8976.
- [113] M. P. Kabir, P. Ghosh, S. Gozem, *J. Phys. Chem. B* **2024**, *128*, 7545.
- [114] A. Onufriev, D. Bashford, D. A. Case, *Proteins: Struct. Funct. Bioinf.* **2004**, *55*, 383.
- [115] G. Sigalov, A. Fenley, A. Onufriev, *J. Chem. Phys.* **2006**, *124*, 124902.
- [116] A. W. Lange, J. M. Herbert, *J. Chem. Theory Comput.* **2012**, *8*, 1999.
- [117] Y. De Coene, S. Van Cleuvenbergen, N. Van Steerteghem, V. Baekelandt, T. Verbiest, C. Bartic, K. Clays, *J. Phys. Chem. C* **2017**, *121*, 6909.
- [118] E. Winder, M. H. Griep, D. Lueking, C. Friedrich, in *Proc. of the 26th Army Science Conf.*, Orlando, Florida, Vol. 14, **2008**.
- [119] R. J. Homans, R. U. Khan, M. B. Andrews, A. E. Kjeldsen, L. S. Natrajan, S. Marsden, E. A. McKenzie, J. M. Christie, A. R. Jones, *Phys. Chem. Chem. Phys.* **2018**, *20*, 16949.
- [120] P. Pracht, F. Bohle, S. Grimme, *Phys. Chem. Chem. Phys.* **2020**, *22*, 7169.
- [121] J. W. Furness, A. D. Kaplan, J. Ning, J. P. Perdew, J. Sun, *J. Phys. Chem. Lett.* **2020**, *11*, 8208.
- [122] S. Grimme, A. Hansen, S. Ehlert, J.-M. Mewes, *J. Chem. Phys.* **2021**, *154*, 064103.
- [123] E. Caldeweyher, C. Bannwarth, S. Grimme, *J. Chem. Phys.* **2017**, *147*, 034112.
- [124] S. Grimme, F. Bohle, A. Hansen, P. Pracht, S. Spicher, M. Stahn, *J. Phys. Chem. A* **2021**, *125*, 4039.
- [125] Universität Bonn, M. C. T. C., Bonn, Germany, F. Bohle, CENSO - Commandline Energetic Sorting of Conformer-Rotamer Ensembles, Version 1.2.0, <https://github.com/grimme-lab/censo/releases/>, (accessed: March 2022).
- [126] V. Åsgerisson, C. A. Bauer, S. Grimme, *Chem. Sci.* **2017**, *8*, 4879.
- [127] T. J. Zuehlsdorff, A. Montoya-Castillo, J. A. Napoli, T. E. Markland, C. M. Isborn, *J. Chem. Phys.* **2019**, *151*, 074111.
- [128] J. Cerezo, D. Aranda, F. J. Avila Ferrer, G. Prampolini, F. Santoro, *J. Chem. Theory Comput.* **2020**, *16*, 1215.
- [129] R. Crespo-Otero, M. Barbatti, *Theor. Chem. Acc.* **2012**, *131*, 1237.
- [130] V. Lukeš, R. Šolc, M. Barbatti, H. Lischka, H.-F. Kauffmann, *J. Theor. Comput. Chem.* **2010**, *09*, 249.
- [131] P. Löwdin, *J. Chem. Phys.* **1950**, *18*, 365.
- [132] K. Nishimoto, N. Mataga, *Z. Phys. Chem.* **1957**, *12*, 335.
- [133] K. Ohno, *Theor. Chim. Acta* **1964**, *2*, 219.
- [134] G. Klopman, *J. Am. Chem. Soc.* **1964**, *86*, 4550.
- [135] R. J. Buenker, S. D. Peyerimhoff, *Theor. Chim. Acta* **1974**, *35*, 33.
- [136] R. J. Buenker, S. D. Peyerimhoff, *Theor. Chim. Acta* **1975**, *39*, 217.
- [137] S. Born, M. Waletzke, *J. Chem. Phys.* **1999**, *111*, 5645.
- [138] F. Neese, *J. Chem. Phys.* **2003**, *119*, 9428.
- [139] P. Monson, W. McClain, *J. Chem. Phys.* **1970**, *53*, 29.
- [140] M. J. Wirth, A. Koskelo, M. J. Sanders, *Appl. Spectrosc.* **1981**, *35*, 14.
- [141] S. Chakrabarti, K. Ruud, *J. Phys. Chem. A* **2009**, *113*, 5485.
- [142] R. Nifos, Y. Luo, *J. Phys. Chem. B* **2007**, *111*, 14043.
- [143] D. H. Fries, M. T. Beerepoot, M. Ringholm, K. Ruud, *J. Chem. Theory Comput.* **2015**, *11*, 1129.
- [144] S. Gholami, L. Pedraza-González, X. Yang, A. A. Granovsky, I. N. Ioffe, M. Olivucci, *The J. Phys. Chem. Lett.* **2019**, *10*, 6293.

- [145] A. Rizzo, S. Coriani, K. Ruud, in *Response Function Theory Computational Approaches To Linear And Nonlinear Optical Spectroscopy*, John Wiley and Sons, Hoboken, NJ **2012**, pp. 77–135.
- [146] M. de Wergifosse, A. L. Houk, A. I. Krylov, C. G. Elles, *J. Chem. Phys.* **2017**, *146*, 144305.
- [147] J.-D. Chai, M. Head-Gordon, *J. Chem. Phys.* **2008**, *128*, 084106.
- [148] J.-D. Chai, M. Head-Gordon, *Phys. Chem. Chem. Phys.* **2008**, *10*, 6615.
- [149] R. Ditchfield, W. J. Hehre, J. A. Pople, *J. Chem. Phys.* **1971**, *54*, 724.
- [150] M. M. Francl, W. J. Pietro, W. J. Hehre, J. S. Binkley, M. S. Gordon, D. J. DeFrees, J. A. Pople, *J. Chem. Phys.* **1982**, *77*, 3654.
- [151] M. S. Gordon, J. S. Binkley, J. A. Pople, W. J. Pietro, W. J. Hehre, *J. Am. Chem. Soc.* **1982**, *104*, 2797.
- [152] P. C. Hariharan, J. A. Pople, *Theor. Chim. Acta* **1973**, *28*, 213.
- [153] W. J. Hehre, R. Ditchfield, J. A. Pople, *J. Chem. Phys.* **1972**, *56*, 2257.
- [154] B. P. Pritchard, D. Altarawy, B. Didier, T. D. Gibbs, T. L. Windus, *J. Chem. Inf. Model.* **2019**, *59*, 4814.
- [155] C. Bannwarth, S. Ehlert, S. Grimme, *J. Chem. Theory Comput.* **2019**, *15*, 1652.
- [156] S. Grimme, *J. Comput. Chem.* **2006**, *27*, 1787.
- [157] S. Grimme, S. Ehrlich, L. Goerigk, *J. Comput. Chem.* **2011**, *32*, 1456.
- [158] F. Weigend, R. Ahlrichs, *Phys. Chem. Chem. Phys.* **2005**, *7*, 3297.
- [159] S. Sinnecker, A. Rajendran, A. Klamt, M. Diedenhofen, F. Neese, *J. Phys. Chem. A* **2006**, *110*, 2235.
- [160] M. Renz, M. Kess, M. Diedenhofen, A. Klamt, M. Kaupp, *J. Chem. Theory Comput.* **2012**, *8*, 4189.
- [161] S. Grimme, *Chemistry* **2012**, *18*, 9955.
- [162] S. Spicher, S. Grimme, *J. Phys. Chem. Lett.* **2020**, *11*, 6606.
- [163] S. Spicher, S. Grimme, *J. Chem. Theory Comput.* **2021**, *17*, 1701.
- [164] A. Klamt, *J. Phys. Chem.* **1995**, *99*, 2224.
- [165] A. Klamt, V. Jonas, T. Bürger, J. C. W. Lohrenz, *J. Phys. Chem. A* **1998**, *102*, 5074.
- [166] S. G. Balasubramani, G. P. Chen, S. Coriani, M. Diedenhofen, M. S. Frank, Y. J. Franzke, F. Furche, R. Grotjahn, M. E. Harding, C. Hättig, A. Hellweg, B. Helmich-Paris, C. Holzer, U. Huniar, M. Kaupp, A. Marefat Khah, S. Karbalaeei Khani, T. Müller, F. Mack, B. D. Nguyen, S. M. Parker, E. Perl, D. Rappoport, K. Reiter, S. Roy, M. Rückert, G. Schmitz, M. Sierka, E. Tapavicza, D. P. Tew, C. van Wüllen, V. K. Voora, F. Weigend, A. Wodyński, J. M. Yu, *J. Chem. Phys.* **2020**, *152*, 184107.
- [167] F. Furche, R. Ahlrichs, C. Hättig, W. Klopper, M. Sierka, F. Weigend, *WIREs Computat. Mol. Sci.* **2014**, *4*, 91.
- [168] Universität Karlsruhe & Forschungszentrum Karlsruhe GmbH, Karlsruhe, Germany, TURBOMOLE 7.6, <https://www.turbomole.org/>, (accessed: March 2022).
- [169] F. Eckert, A. Klamt, *AIChE J.* **2002**, *48*, 369.
- [170] E. F. Pettersen, T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E. C. Meng, T. E. Ferrin, *J. Comput. Chem.* **2004**, *25*, 1605.
- [171] M. Brehm, B. Kirchner, *J. Chem. Inf. Model.* **2011**, *51*, 2007.
- [172] M. Brehm, M. Thomas, S. Gehrke, B. Kirchner, *J. Chem. Phys.* **2020**, *152*, 164105.
- [173] C. Hättig, F. Weigend, *J. Chem. Phys.* **2000**, *113*, 5154.
- [174] J. Dunning, H. Thom, *J. Chem. Phys.* **1989**, *90*, 1007.
- [175] Universität Karlsruhe & Forschungszentrum Karlsruhe GmbH, Karlsruhe, Germany, TURBOMOLE 7.7.1, <https://www.turbomole.org/>, (accessed: October 2023).
- [176] C. Bannwarth, S. Grimme, sTDA-xTB for Ground State Calculations, <https://github.com/grimme-lab/xtb4stda>, **2016**.
- [177] std2, A Program for Computing Excited States and Response Functions Via Simplified TD-DFT methods (sTDA, sTD-DFT, SF-sTD-DFT, XsTDA, XsTD-DFT, and SF-Xs-TD-DFT), <https://github.com/grimme-lab/std2>, (accessed: January 2025).
- [178] N. Nunthaboot, A. Nueangaudom, K. Lugsanangarm, S. Pianwanit, S. Kokpol, F. Tanaka, *Comput. Biol. Chem.* **2018**, *72*, 96.

Manuscript received: December 16, 2024
Revised manuscript received: April 7, 2025
Version of record online:

Molecular Folding Governs Switchable Singlet Oxygen Photoproduction in Porphyrin-Decorated Bistable Rotaxanes

Jan Riebe¹, Benedikt Bädorf², Sarah Löffelsender², Matias Ezequiel Gutierrez Suburu³, María Belén Rivas Aiello³, Cristian Alejandro Strassert³, Stefan Grimme², Jochen Niemeyer¹

Received: March 15, 2024

First published: August 07, 2024

Reprinted with permission from:

J. Riebe, B. Bädorf, S. Löffelsender, M. E. Gutierrez Suburu, M. B. Rivas Aiello, C. A. Strassert, S. Grimme, and J. Niemeyer, *Communications Chemistry* **7** (2024), DOI: 10.1038/s42004-024-01247-7
Copyright © 2024 The Authors. Published under the license CC BY 4.0

Own manuscript contributions:

- Geometry preparations
- Preparation and performing parts of calculations
- Supervision for all calculations
- Interpretation of results
- Writing and revising parts of the manuscript

¹ Faculty of Chemistry (Organic Chemistry) and Center for Nanointegration Duisburg-Essen (CENIDE), University of Duisburg-Essen, Universitätsstrasse 7, 45141 Essen, Germany

² Mulliken Center for Theoretical Chemistry, Clausius Institute for Physical and Theoretical Chemistry, Rheinische Friedrich-Wilhelms-Universität Bonn, Beringstr. 4, 53115 Bonn, Germany

³ Institut für Anorganische und Analytische Chemie, CeNTech, CiMIC, SoN, Universität Münster, Heisenbergstr. 11, 48149 Münster, Germany

<https://doi.org/10.1038/s42004-024-01247-7>

Molecular folding governs switchable singlet oxygen photoproduction in porphyrin-decorated bistable rotaxanes



Jan Riebe¹, Benedikt Bädorf², Sarah Löffelsender², Matias E. Gutierrez Suburu³,
María Belén Rivas Aiello³, Cristian A. Strassert³, Stefan Grimme²✉ & Jochen Niemeyer¹✉

Rotaxanes are mechanically interlocked molecules where a ring (macrocycle) is threaded onto a linear molecule (thread). The position of the macrocycle on different stations on the thread can be controlled in response to external stimuli, making rotaxanes applicable as molecular switches. Here we show that bistable rotaxanes based on the combination of a Zn(II) tetraphenylporphyrin photosensitizer, attached to the macrocycle, and a black-hole-quencher, attached to the thread, are capable of singlet oxygen production which can be switched on/off by the addition of base/acid. However, we found that only a sufficiently long linker between both stations on the thread enabled switchability, and that the direction of switching was inversed with regard to the original design. This unexpected behavior was attributed to intramolecular folding of the rotaxanes, as indicated by extensive theoretical calculations. This evidences the importance to take into account the conformational flexibility of large molecular structures when designing functional switchable systems.

Molecular switches are molecules that exist in different states which can be reversibly addressed. Most switches feature two stable switching states that can be selectively addressed by an external stimulus^{1–4}. The change in molecular properties which results from the switching process can be used to deliver a function, which has allowed the construction of a plethora of functional materials based on molecular switches, such as molecular muscles and elevators^{5,6}, nanovalves and other systems for controlled release^{7,8}, materials with switchable surface properties⁹, switchable catalysts^{10–14} and materials with switchable photophysical properties^{15–17}.

Mechanically interlocked molecules, such as rotaxanes (and, to a lesser extent, catenanes) provide an excellent platform for the construction of functional molecular switches^{18–22}. In a bistable rotaxane, featuring a macrocycle that encircles a thread with two binding stations, the preferred position of the macrocycle is determined by its relative affinity to the binding stations. However, when an external stimulus can be employed to modify one station, this can invert the relative affinities of the two binding stations. This infers a large amplitude motion of the macrocycle along the thread, which moves away from the (modified) first station and binds to the second station. Reversing this process leads to re-binding of the macrocycle at the first station. Like molecular switches in general, switchable bistable rotaxanes have been based on a number of switching mechanisms, most

commonly (but not limited to) changes in temperature, irradiation with light, application of an electric potential, addition/removal of other chemicals or a change in pH. One of the most widely used types of bistable rotaxanes features a crown-ether-based macrocycle and a thread with an amine/ammonium station plus a triazolium station (see Fig. 1)^{23,24}. Deprotonation of the ammonium station, which is the preferred binding station when protonated, leads to relocation of the macrocycle towards the triazolium station (and vice versa, with the respective co-conformer being the predominant one by >99%, making the switching effectively binary²⁵). This motion has been used to shield/reveal the amine/ammonium station by the macrocycle (e.g., for the design of switchable catalysts^{26–29}, see Fig. 1a), to change the co-conformation between stretched/contracted conformers (e.g., for the design of molecular muscles^{30,31}, see Fig. 1b) or to change the distance and/or environment of chromophores attached to the sub-components (e.g., for the design of molecules with tunable photophysical properties, such as fluorescence, see Fig. 1c)^{32–38}. However, it was also realized that the prediction of the molecular conformations in each switching state (and with this, the function in each state) is not trivial. For example, Leigh and coworkers found that in the protonated state, their rotaxane (see Fig. 1a) can adopt a folded conformation that brings both triazolium units in close proximity. This allows the use of the protonated rotaxane as a catalyst

¹Faculty of Chemistry (Organic Chemistry) and Center for Nanointegration Duisburg-Essen (CENIDE), University of Duisburg-Essen, Universitätsstrasse 7, 45141 Essen, Germany. ²Mulliken Center for Theoretical Chemistry, Rheinische Friedrich-Wilhelms-Universität Bonn, Beringstrasse 4, 53115 Bonn, Germany. ³Institut für Anorganische und Analytische Chemie, CeNTech, CiMIC, SoN, Universität Münster, Heisenbergstr. 11, 48149 Münster, Germany.

✉ e-mail: grimme@thch.uni-bonn.de; jochen.niemeyer@uni-due.de

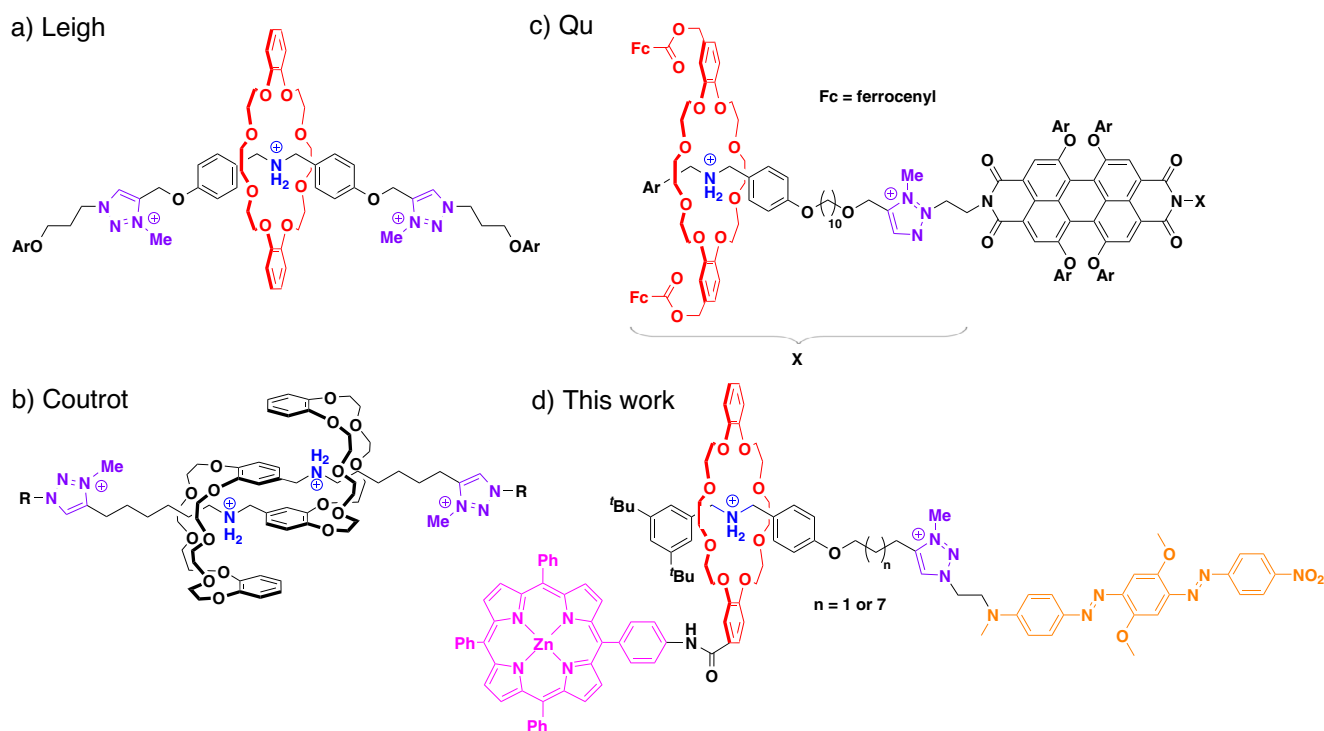


Fig. 1 | Selected molecular switches based on bistable rotaxanes with ammonium/amine plus triazolium stations. **a** Switchable catalysts²⁹, **b** molecular muscles³¹, **c** switchable chromophors³⁴, **d** switchable systems for ¹O₂ photoproduction (this work).

via halide-abstraction, enabled by chelate-type binding of halides by the triazolium units²⁶.

The development of switchable photosensitizers for singlet dioxygen (¹O₂) photoproduction (for simplicity, referred to hereafter as: singlet oxygen production) is an important field of application for functional materials. Singlet oxygen can be used for medical purposes, such as photodynamic therapy^{39,40} or disinfection^{41,42}, but also as a reagent for organic synthesis^{43–46}. However, due to its high reactivity and cell toxicity, there is an ongoing interest in smart materials that produce singlet oxygen only after switching to an on-state, while the singlet oxygen production is suppressed in an off state^{47–52}. Various approaches have been used to design such systems, especially in the context of photodynamic therapy. Most systems rely on the quenching of photosensitizers in the off-state (suppressing singlet oxygen production), which can be realized via different mechanisms. Importantly, this includes systems that can be activated in a non-reversible or in a reversible form. For example, non-reversible activation has been realized by integration of photosensitizers into hybrid organic/inorganic materials based on layered double hydroxides or calcium phosphate (leading to quenching), which release the photosensitizer in its non-quenched on-state at low pH^{53,54}. In addition, photosensitizer-quencher conjugates with cleavable linkers have shown to undergo fragmentation in a redox-dependent or enzyme-mediated fashion, leading to a non-reversible activation of singlet oxygen production^{55–60}. Systems that, in principle, allow for reversible activation include polymers that feature photosensitizer-substituents, allowing for an on/off-switchable singlet oxygen production based on pH-dependent conformational changes of the polymer-backbone^{61,62}. Photosensitizers or photosensitizer dyads with Lewis-basic substituents have been used to enable the on-switching of singlet oxygen production at low pH, based on the suppression of photoinduced electron transfer and/or modulation of acceptor/donor energy levels^{63–65}. In a related approach, photosensitizers conjugated with diarylethene-photoswitches make use of the UV/Vis-absorption of the closed photoswitch, which leads to a quenching of the photosensitizers (off-state) and vice versa^{66,67}. Photosensitizers and quenchers conjugated with complementary DNA strands have been used to achieve on/off-switching of singlet oxygen production by

reversible DNA-hybridization⁶⁸. Finally, rotaxanes have also successfully been used to construct systems for biomedical applications in general, and for the design of switchable photosensitizers specifically^{69–71}. Huang and coworkers developed a pH-switchable pseudorotaxane based on an AIEgen-based thread (AIE: aggregation-induced emission) and a pillar[5]arene macrocycle that undergoes a pH-dependent co-conformational change, leading to a switchable singlet oxygen production⁷². Lin and coworkers reported a pH-switchable [1]rotaxane-based on a crown-ether macrocycle and a multifunctional thread containing amine/ammonium and triazolium stations, plus a diarylethene-photoswitch and an AIEgen photosensitizer. This system undergoes dual switching of photochromic behavior and singlet oxygen production controlled by pH and light⁷³. Shinmori and coworkers synthesized a [2]rotaxane featuring a gold-nanoparticle-stopper (which acts as a quencher), a crown-ether macrocycle substituted with a porphyrin photosensitizer and a single amine/ammonium station on the thread⁷⁴. This system shows pH-dependent switching, although reversibility is low, and the singlet oxygen quantum yields of the switching states are quite similar.

Based on this precedence, we envisaged that bistable [2]rotaxanes containing a macrocycle substituted with a suitable photosensitizer, and a thread functionalized with a suitable quencher, should allow for the construction of a reversibly switchable system for singlet oxygen production. Ideally, singlet oxygen production in the off-mode should be fully suppressed, to avoid undesired side-effects in a possible application. To this end, we designed rotaxanes **1a** and **1b** (see Fig. 1d), which commonly contain Zn(II) tetraphenylporphyrin⁷⁵ as a photosensitizer (attached to the macrocycle) and a black-hole-quencher (attached to one end of the tread), but differ in the length of the linker between both stations on the thread. We found that rotaxane **1a** indeed allows for an on/off-switching (¹O₂ quantum yield, Φ_{Δ} = 15%/3%) by pH-induced translocation of the macrocycle on the thread, whereas rotaxane **1b** remains in an off-state independent of macrocycle position. Surprisingly, we discovered that the protonated state of **1a** represents the off-state of this molecular switch (and vice versa). Based on a series of control-rotaxanes and extensive theoretical analysis of the rotaxane conformations, we found that the switchable singlet oxygen production is

governed by molecular folding, which in turn is influenced by structural variations in the rotaxane-structures.

Results and discussion

Synthesis of the rotaxanes

The design of the pH-switchable rotaxanes for controllable singlet oxygen production was based on the following considerations: We employed Zn(TPP) tetraphenylporphyrin (ZnTPP) as a photosensitizer, due to its strong absorption at ca. 420 nm ($\epsilon = 560,000 \text{ cm}^{-1} \text{ M}^{-1}$)⁷⁶ and its high quantum yield of singlet oxygen production ($\Phi_{\Delta} \approx 0.7$)⁷⁷. As a quencher, we employed the black-hole-quencher **6** (BHQ-2)⁷⁸, which features a dialkylaniline, a central dimethoxybenzene and a terminal nitrobenzene which are linked by two diazo groups. The BHQ unit shows a broad absorption band in the range of 350–650 nm is thus suitable as a FRET-acceptor (FRET: Förster resonance energy transfer) for the ZnTPP unit, which shows fluorescence emission in the range of 550–720 nm. This should suppress singlet oxygen production when both units are in close proximity.

For the combination of ZnTPP and BHQ-2, the calculated Förster-radius amounts to 2.8 nm (see SI section 1.4.2 for details). We envisaged that a sufficiently large conformational switching, which allows for moving ZnTPP and BHQ within/out of their Förster radius, should be possible by employing a bistable rotaxane with a sufficiently long linker between both stations. Here, we opted for a pH-switchable rotaxane consisting of a thread with an ammonium/amine- and a triazolium station and a dibenzo[24]crown[8] macrocycle. Attaching the ZnTPP to the macrocycle and using the BHQ as one of the stopper groups on the thread (in close proximity to the triazolium station), the rotaxane was designed to allow switching as follows: In the protonated state, the macrocycle is located around the ammonium station, sufficiently increasing the distance between ZnTPP and BHQ, resulting in $^1\text{O}_2$ production of the ZnTPP (on-state). In turn, deprotonation should locate the macrocycle at the triazolium station in closer proximity to the BHQ unit, leading to efficient FRET-based quenching and suppression of the $^1\text{O}_2$ production (off-state).

The synthesis of these functionalized rotaxanes **1a/b** (differing in the length of the alkylene spacer between both stations) was realized as follows

(see Fig. 2): First, an amino-substituted ZnTPP⁷⁹ was coupled to a carboxylic-acid functionalized dibenzo[24]crown[8]⁸⁰ via amide coupling to yield the ZnTPP-appended macrocycle **5**. Then, the dibenzylammonium-based half-threads **H-4a⁺/H-4b⁺** featuring a di-*tert*-butylphenyl-stopper on one side and an alkyne-terminated chain on the other side (undecynyl for **a**, pentynyl for **b**) were constructed by reductive amination, followed by protonation with HPF₆. Upon mixing of **5** with **H-4a⁺/H-4b⁺**, formation of the pseudorotaxanes was observed in a slow-exchange regime, with association constants of 61/142 L mol⁻¹ (see SI Figs. S19 and S20). This allowed for synthesis of the rotaxanes by stoppering of the pseudorotaxanes, which was achieved by Cu-catalyzed alkyne-azide click reaction using the BHQ-azide **6** as the coupling partner. Finally, methylation of the triazole with methyl iodide and ion-exchange with ammonium-hexafluorophosphate yielded rotaxanes **H-1a²⁺/H-1b²⁺** in their protonated ammonium form as the bis-hexafluorophosphate salts (obtained in 10/9% yield over two steps from the half-threads **H-4a⁺/H-4b⁺**).

Acid/base inducing conformational switching

The protonated rotaxanes **H-1a²⁺/H-1b²⁺** undergo conformational switching by reversible deprotonation/reprotonation with base/acid. Change of the protonation state leads to distinct chemical shift changes in the ¹H-NMR spectrum, in line with other amine-/triazolium-rotaxanes described earlier²⁴ (see Fig. 3 for **1a**, see SI Fig. S14 for **1b**): Upon deprotonation with NaOH, the signals of the methylene protons H-3 and H-4 (see Fig. 2 for numbering) next to the ammonium/amine group move upfield ($\Delta\delta = -1.2 \text{ ppm}$) not only due to loss of the positive charge on the nitrogen atom after deprotonation, but also because of the loss of hydrogen-bonding interactions with the crown-ether's oxygen atoms. Similarly, the signal for proton H-6 on the electron-rich phenylene near the amine station experiences the expected small upfield shift of -0.1 ppm , due to loss of the positive charge on the ammonium. Successful relocation of the macrocycle onto the triazolium station is indicated by a drastic downfield shift of the signal of the aromatic triazolium proton H-8 by $+1.5 \text{ ppm}$ as well as the adjacent signal of the methylene protons H-7 by $+0.7 \text{ ppm}$, caused by hydrogen-bonding to the crown-ether's oxygen atoms.

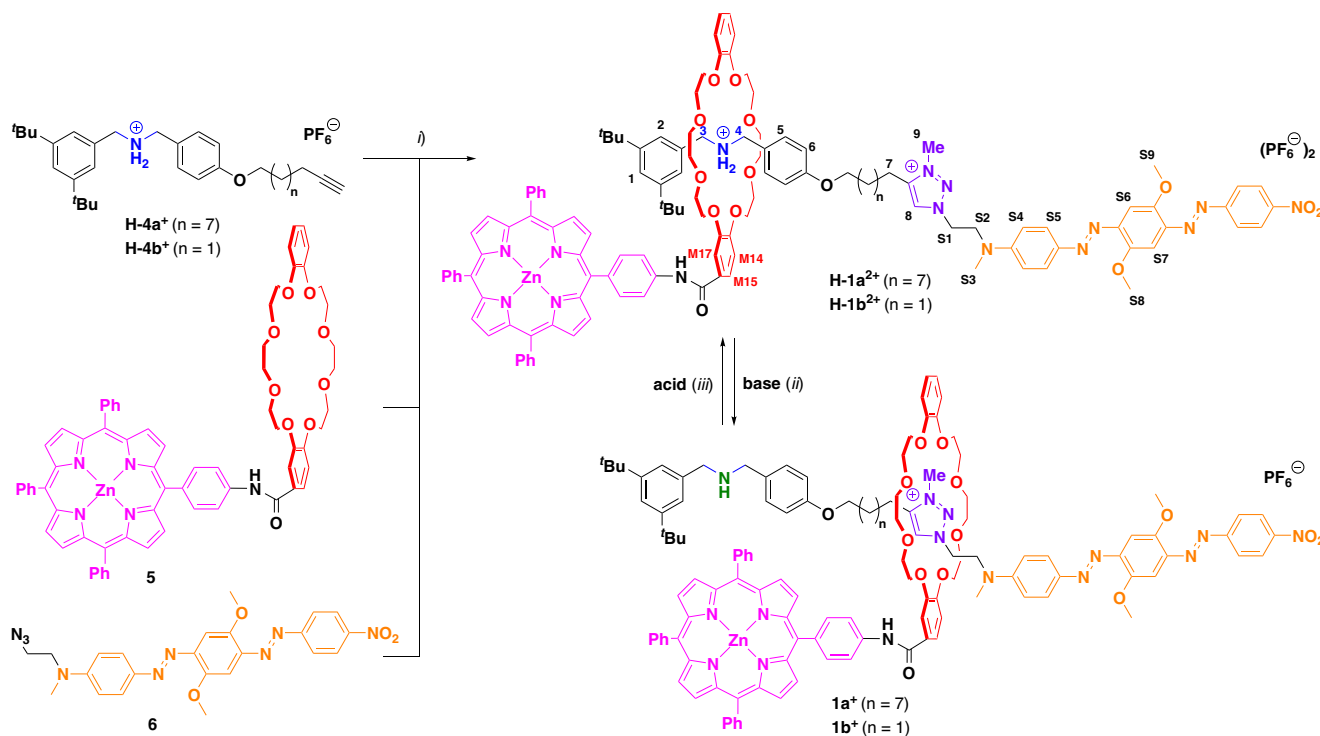


Fig. 2 | Synthesis of the rotaxanes **1a/b.** i) Cu(MeCN)₄PF₆, CH₂Cl₂, r. t., 16 h, then methyl iodide, r. t., 24–48 hours, 10%/9% yield over two steps; ii) aq. NaOH (2 eq.), acetone, followed by removal of excess NaOH; iii) CF₃COOH (1 eq.). In situ acid/base switching leads to the introduction of additional cations/anions, which are not shown.

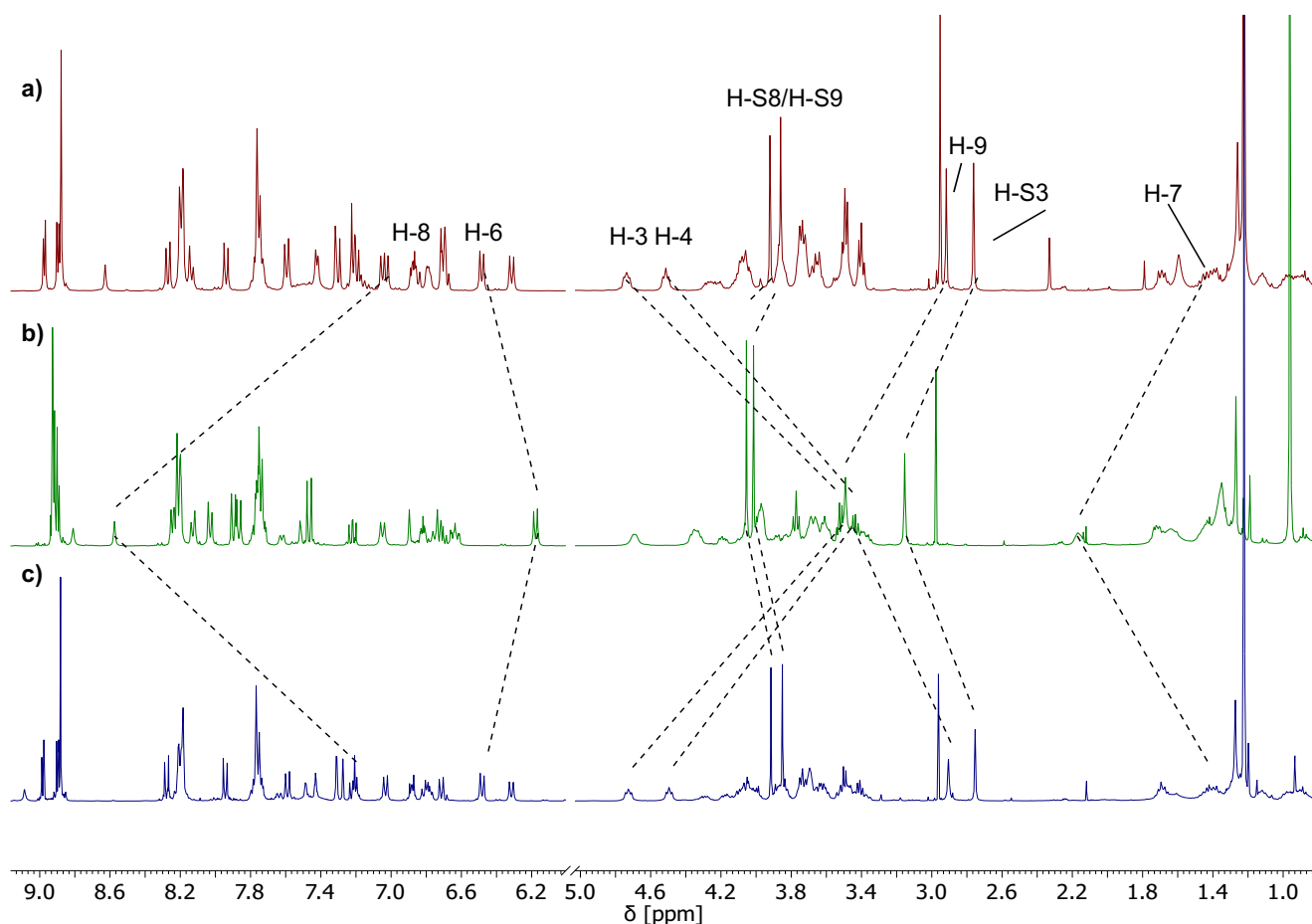


Fig. 3 | Switching between rotaxanes H-1a²⁺ and 1a⁺ by deprotonation/reprotonation, followed by ¹H-NMR. a Rotaxane H-1a²⁺ (as synthesized), **b** rotaxane 1a⁺ after deprotonation. The signals at around 4.6 ppm in this spectrum could not be

unambiguously assigned, but do not correspond to H-3/H-4, which can clearly be assigned. **c** Solution from **b** after addition of CF₃COOH (1 eq.) in situ. All spectra: CD₂Cl₂, 400 MHz, 298 K. For numbering of the positions, see Fig. 2.

However, some shifts were unexpected: The signals of the N-methyl group H-S3 and the methoxy groups H-S8 and H-S9 on the BHQ moiety are also shifted ($\Delta\delta = +0.4$ ppm and $+0.1$ ppm, respectively), despite their large distance from the triazolium station. Furthermore, the downfield shift of the triazolium methyl group H-9 ($\Delta\delta = +0.6$ ppm) upon deprotonation is counterintuitive, because an upfield shift would be expected²⁴. Such a downfield shift for H-9 is observed in all rotaxanes (1a/b⁺, 2⁺, and 3⁺) in this study, but not with an analog of **3** possessing no ZnTPP substituent on the macrocycle (S20-H(PF₆)₂, see SI Fig. S17). Additionally, the switching of this rotaxane and H-1a²⁺ was also investigated in THF-*d*₈, showing almost identical shifts of characteristic signals as in DCM-*d*₂ (for example, H-3 and H-4 move upfield by -1.3 ppm, see SI Figs. S13 and S18). Unfortunately, an attempt to prepare the non-interlocked thread for comparison was unsuccessful. In summary, these unexpected changes in chemical shift could indicate secondary interactions resulting from molecular folding of the rotaxane-structures (see theoretical investigation of the conformation space, *vide infra*).

The reverse switching can be achieved by the addition of CF₃COOH, which restores the original NMR spectrum with only minor perturbations (see Fig. 1c), probably caused by the newly introduced trifluoroacetate counter-anion.

Investigation of ¹O₂ photoproduction

With both pH-switchable rotaxanes in hand, we investigated whether the acid/base-induced conformational change influences the ¹O₂ production upon irradiation. To follow the ¹O₂ production over time, solutions of the

rotaxanes in THF were irradiated at 420 nm in the presence of excess diphenylisobenzofuran (DPBF), which acts as a ¹O₂ scavenger⁸¹. The decrease in DPBF-fluorescence intensity was monitored over time (see SI Figs. S21–S23) and used to calculate the singlet oxygen quantum yields (from triplicate measurements, after subtraction of the background reaction, see SI section 1.4.1 for details). For validation of this method, the Φ_{Δ} of unsubstituted ZnTPP was also measured by quantification of the singlet oxygen phosphorescence, giving a near-identical result ($\Phi_{\Delta} = 0.66 \pm 0.03$ via ¹O₂-phosphorescence, $\Phi_{\Delta} = 0.72 \pm 0.12$ via DPBF method). Comparison of the singlet oxygen production of ZnTPP with that of rotaxanes in both protonation states (i.e. H-1a²⁺ and 1a⁺; H-2b²⁺ and 1b⁺) delivered the following surprising results (see Fig. 4a): In all four cases, the rotaxanes show significantly reduced ¹O₂ production in comparison to free ZnTPP, however with marked differences between the systems. For **1a**, which features the longer alkylene linker, the protonated H-1a²⁺-state shows only minor ¹O₂ production ($\Phi_{\Delta} < 0.03$), while the deprotonated 1a⁺-state gives a significantly higher ¹O₂ quantum yield ($\Phi_{\Delta} = 0.15 \pm 0.03$). Thus, the deprotonated state, which is characterized by the macrocycle being located around the triazolium station, is ca. 500% more active in ¹O₂ production, which is counterintuitive based on the original design of the molecular switch (i.e., the distance between ZnTPP and BHQ should be smaller for 1a⁺, leading to an *off*-switching). As a comparison, rotaxane **1b** with the shorter alkylene spacer shows virtually no ¹O₂ production independent of protonation state ($\Phi_{\Delta} < 0.01$ for H-1b²⁺, $\Phi_{\Delta} < 0.01$ for 1b⁺), although the pH-responsive translocation of the macrocycle was clearly demonstrated by ¹H-NMR spectroscopy.

Fig. 4 | $^1\text{O}_2$ production of rotaxanes in comparison to free ZnTPP. **a** $^1\text{O}_2$ production of rotaxanes **H-1a $^{2+}$** (red circles) and **1a $^+$** (green triangles); **H-1b $^{2+}$** (blue triangles) and **1b $^+$** (turquoise diamonds) in comparison to free ZnTPP (black squares) (uncertainty bars for **H-1a $^{2+}$** , **H-1b $^{2+}$** and **1b $^+$** too small to be visible); **b** $^1\text{O}_2$ production of control-rotaxanes **H-2 $^{2+}$** (purple hexagons) and **2 $^+$** (ochre hexagons); **H-3 $^{2+}$** (blue pentagons) and **3 $^+$** (orange triangle) in comparison to free ZnTPP (black squares). All values determined by decrease in DPBF fluorescence (all: THF, concentration of photosensitizers adjusted to absorbance of 0.8 at 424 nm, excitation at 420 nm; values given as average of triplicates, error bars are standard deviations).

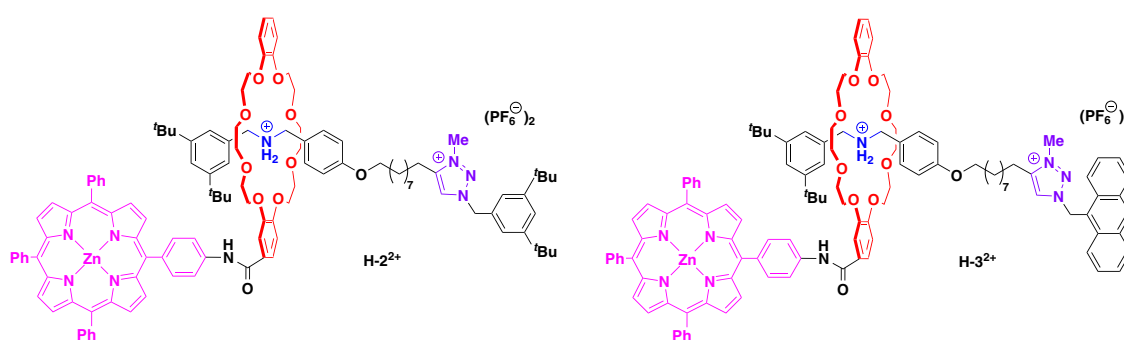
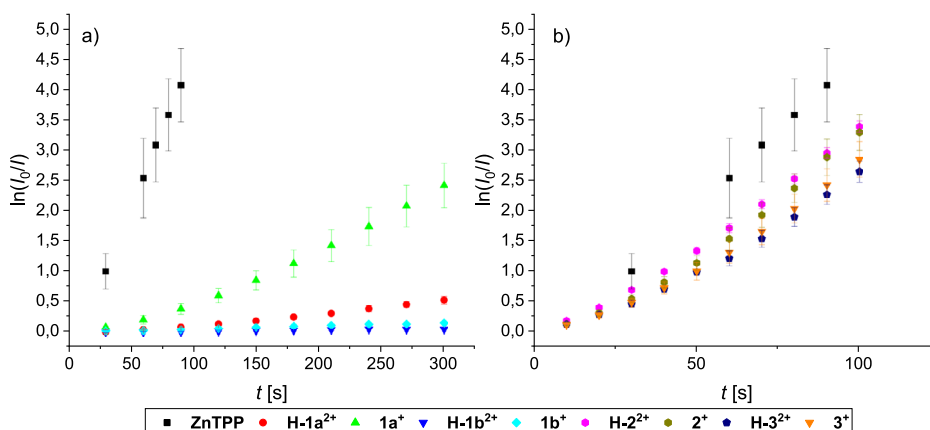


Fig. 5 | Structural formulas of control rotaxanes. The control rotaxanes **H-2 $^{2+}$** and **H-3 $^{2+}$** (only protonated states shown) were synthesized analogously to rotaxanes **H-1a $^{2+}$** and **H-1b $^{2+}$** (see SI section 1.2.2).

While these results show that bistable rotaxanes can be used to generate an *on/off*-switchable material for $^1\text{O}_2$ photosensitization, the following questions needed to be answered for an in-depth understanding:

Why is the deprotonated state **1a $^+$** the *on*-state of the molecular switch (while **H-1a $^{2+}$** is *off*)?

Why is the $^1\text{O}_2$ quantum yield low in comparison to the free ZnTPP, even in the *on*-state?

Why does the shorter spacer in **H-1b $^{2+}$** /**1b $^+$** lead to suppression of $^1\text{O}_2$ production, independent of protonation state?

To answer these questions, we concluded that a better picture of the influence of the rotaxane architecture on the $^1\text{O}_2$ production is needed. Thus, we synthesized two control rotaxanes **2/3**, which lack the BHQ unit but instead feature a di-*tert*-butylphenyl-stopper or an anthracenyl-stopper in proximity to the triazolium station (see Fig. 5). These were synthesized in analogous fashion to **1a/b** by employing the corresponding stopper azides and were obtained in yields of 60/15% after methylation and anion-exchange. Furthermore, we attempted to synthesize analogs of **1a/b** with even longer or shorter alkylene spacers and also a variant of **1a** with an inverted thread that features the quencher in proximity to the ammonium/amine station, but these syntheses could not be realized.

With the control rotaxanes in hand, we investigated their ability to produce $^1\text{O}_2$ in both protonation states (i.e. **H-2 $^{2+}$** and **2 $^+$** ; **H-3 $^{2+}$** and **3 $^+$**). We found that all four systems showed nearly identical results, with $^1\text{O}_2$ quantum yields in a range of 0.48–0.60 (see Fig. 4b). Firstly, this demonstrates that the quenching in **1a/b** is indeed affected by the BHQ unit and not by other parts of the rotaxane-architecture. Secondly, the results show that the protonation state of the ammonium/amine station alone does not influence $^1\text{O}_2$ production, but that indeed, the conformational change that places the ZnTPP and BHQ in different spatial arrangements must play a crucial role in **1a/b**. Thirdly, we found that the anthracenyl-stopper, which is

known to act as an acceptor for triplet energy transfer from Zn(II) porphyrins⁸², is not suitable for modulation of $^1\text{O}_2$ photosensitization by conformational switching, showing that the nature of the BHQ quencher is crucial for excitation energy transfer.

Fluorescence lifetimes

To obtain information about the influence of structures of rotaxanes **1a/b**, **2**, and **3** on their excited state behavior, we investigated their photophysical properties. The absorption and emission spectra of all rotaxanes are governed by the ZnTPP chromophore and show only minor differences (see SI Figs. S25–S29). However, the differences observed in the $^1\text{O}_2$ production are reflected by differences in the corresponding fluorescence lifetimes (see Table 1): Here, both the ZnTPP reference and the control rotaxanes in both protonation states (i.e. **H-2 $^{2+}$** and **2 $^+$** ; **H-3 $^{2+}$** and **3 $^+$**) show almost identical lifetimes of ca. 1.7–1.8 ns with monoexponential fluorescence decays. This substantiates the finding that in the absence of the BHQ unit, the rotaxane structure does not significantly influence the photophysical behavior of the ZnTPP unit, independent of conformational switching. For the BHQ-substituted rotaxanes, we observed multiexponential decay profiles, which might be indicative of different conformations that influence fluorescence lifetimes. For the rotaxane with the shorter alkylene spacer, both protonation states **H-1b $^{2+}$** and **1b $^+$** show a biexponential decay featuring a smaller component with a lifetime similar to free ZnTPP (1.76/1.57 ns, 19–29% contribution) and a dominating component with a significantly shortened lifetime (0.50/0.39 ns, 81–71% contribution), resulting in almost identical average values (0.73/0.74 ns for **H-1b $^{2+}$** /**1b $^+$**). For the rotaxane with the longer alkylene spacer (**H-1a $^{2+}$** /**1a $^+$**), we find triexponential decay kinetics, which only shows a minor component similar to free ZnTPP (1.69/1.40 ns, 5–7% contribution). The major contributions are given by two shorter lifetimes, which differ depending on the rotaxane protonation state: The protonated rotaxane **H-1a $^{2+}$** shows components at 0.87 ns (61%) and

Table 1 | Singlet oxygen quantum yield (Φ_{Δ}) and amplitude-weighted average fluorescence lifetimes (τ_{av}) of rotaxanes **1a/b, **2**, and **3** in both protonation states in comparison to free ZnTPP^a**

Compound	Φ_{Δ} ^b	lifetime components τ [ns] (relative amplitudes in %) ^c	Amplitude-weighted average lifetimes τ_{av} [ns]
ZnTPP	0.72 ± 0.12	– ^d	1.84 ± 0.004
H-1a ²⁺	<0.03	1.69 (5), 0.87 (61), 0.30 (33)	0.727 ± 0.016
1a ⁺	0.15 ± 0.03	1.40 (7), 0.57 (15), 0.11 (78)	0.263 ± 0.015
H-1b ²⁺	<0.01	1.76 (19), 0.50 (81)	0.73 ± 0.03
1b ⁺	<0.01	1.57 (29), 0.39 (71)	0.74 ± 0.05
H-2 ²⁺	0.60 ± 0.02	– ^d	1.705 ± 0.003
2 ⁺	0.57 ± 0.06	– ^d	1.712 ± 0.003
H-3 ²⁺	0.48 ± 0.03	– ^d	1.712 ± 0.002
3 ⁺	0.53 ± 0.05	– ^d	1.712 ± 0.002

^aAll measured in THF solution. ^bConcentration based on absorbance of 0.8 at 424 nm, Φ_{Δ} determined from fluorescence decrease of DPBF scavenger in a linearized $\ln(I_0/I)$ vs. time plot from triplicate measurements, after subtraction of the background reaction (see section 1.4.1 of the SI for details). ^c10 μ M solutions, λ_{exc} = 407.2 nm, λ_{em} = 605 nm, uncertainty: ±0.01 ns. ^dOnly one component (given as average lifetime).

0.30 ns (33%), resulting in an average lifetime of 0.73 ns (similar to that of H-1b²⁺/1b⁺). However, for the deprotonated case 1a⁺, a smaller component at 0.57 ns (15%) and a dominating component at 0.11 ns (78%) lead to a decreased average lifetime of 0.26 ns. These findings show that the photophysical properties of 1a/b are strongly influenced by the presence of the BHQ unit and that the properties of 1a⁺ are unique within this series. This qualitatively reflects the behavior of the systems in ¹O₂ production (1a⁺ behaves differently than H-1a²⁺, H-1b²⁺, and 1b⁺); however, there is no direct connection between the fluorescence lifetimes (decay from the singlet state) and the ¹O₂ production (which occurs from the triplet state). Unfortunately, attempts to further characterize rotaxanes 1a/b, 2, and 3 by transient absorption spectroscopy were hampered by significant photobleaching of the ZnTPP upon irradiation. This issue resulted in a notable lack of reproducibility, so that no data regarding the triplet lifetimes could be obtained (see SI Table S1).

Theoretical investigation of the conformational space of the rotaxanes

Although a pH-induced switching of singlet oxygen production for the H-1a²⁺/1a⁺ pair can be observed, it is reversed compared to what was initially envisaged in the design. Additionally, no switching is observed for the H-1b²⁺/1b⁺ pair even though NMR-spectroscopic analysis clearly shows that both systems do undergo a pH-induced relocation of the macrocycle. Thus, a better understanding of the molecular conformations of the rotaxanes 1a/b in both switching states was needed in order to explain their photophysical behavior. With the aid of computational chemistry, the conformational spaces of all four systems H-1a²⁺/1a⁺ and H-1b²⁺/1b⁺ were studied, focusing on possible intramolecular interactions involving the ZnTPP unit. In the first step, the potential energy surface of all variants was scanned using the fast semiempirical methods GFN n -xTB ($n = 1, 2$)^{83–85} and the force field GFN-FF⁸⁶ in combination with the program CREST⁸⁷. Subsequently, relevant (low-lying) and structurally representative conformations were selected and their Gibbs free energies (consisting of electronic gas-phase energy, thermal correction, and solvation contribution) were computed at a higher DFT level of theory (DFT: density functional theory)^{88–90} (<https://github.com/grimme-lab/xtb>). For easier comparison, these conformers were divided into three categories: *open*, *half-open*, and *closed*. The division into those categories is based on binding to the Zn(II) center in its axial (octahedral) position and the folding of the whole system (see the methods sections or the SI for more details). Note that we could not consider explicit THF molecules in the calculations as this

would be computationally too demanding. The coordination of explicit THF molecules to the Zn(II) center is expected to be in competition with the O₂ molecules in the solution. However, the influence on the singlet oxygen production is most probably similar for both protonation states, so this effect cancels out for the presented comparison.

Conformational search and Gibbs free energies for rotaxanes H-1a²⁺/1a⁺

In the investigation of the conformational spaces, almost all performed conformer searches for H-1a²⁺ identified the same low-lying *closed* conformational motif indicating limited conformational flexibility. This is attributed to the high charge (+2) of the system, which leads to strong (intramolecular) electrostatic interactions that stabilize *closed* conformations. Finding the most relevant low-lying conformers of 1a⁺ proved to be more challenging. A variety of distinct structural motifs were observed, with many of them featuring a different segment of the rotaxane coordinating with the Zn-atom. Moreover, several less folded (half-open) conformations comparable in energy to those of *closed* conformers were found, a phenomenon not observed for H-1a²⁺. These results suggest a greater degree of flexibility for 1a⁺.

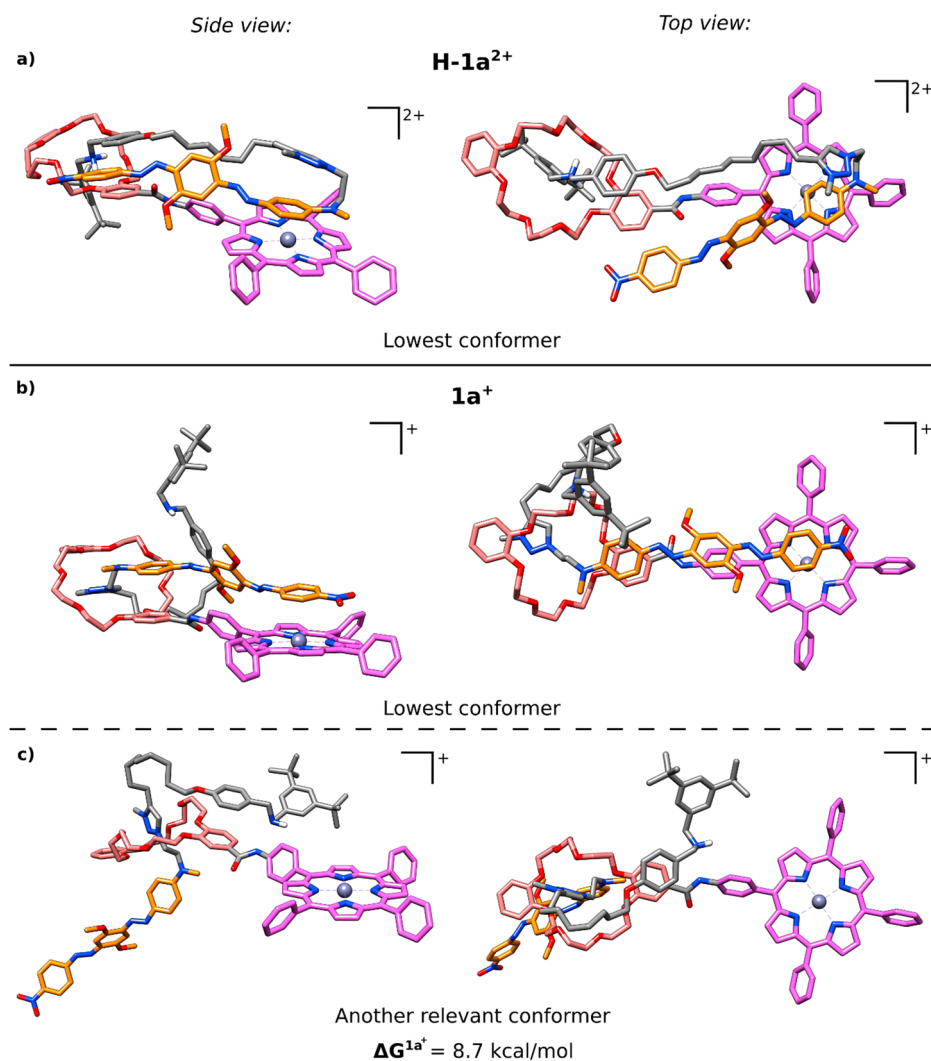
Based on the results of the conformational searches, we selected relevant (low-lying) and structurally representative conformers and calculated the Gibbs free energies of a total of 31 conformers of H-1a²⁺ and 26 conformers of 1a⁺ (see SI Tables S4 and S5).

The computed electronic gas-phase, thermal, and solvation contributions varied strongly between closed, half-open, and open conformations, resulting in a considerable reranking of the investigated 1a⁺ conformations (based on their Gibbs free energies), while the ranking stayed almost the same for H-1a²⁺.

For H-1a²⁺, primarily one *closed* structural motif was identified. Despite unfavorable solvation and thermal contributions, this conformation exhibited the lowest free energy (see Table S4 in the SI) and is depicted in Fig. 6a. As expected for this protonation state, the crown-ether macrocycle encircles the dibenzylammonium station, mediated by two NH⁺⋯O hydrogen bonds ($d(\text{NH}^+\cdots\text{O}) = 1.84/1.95 \text{ \AA}$) and three additional CH₂⋯O contacts ($d(\text{CH}\cdots\text{O}) = 2.45\text{--}2.86 \text{ \AA}$). Tight molecular folding is enabled by a turn at the C₂H₄NMe linker between the triazolium group and the BHQ unit. Interestingly, this leads to a triple π -stack, involving the ZnTPP unit, the aniline group of the BHQ unit, and the triazolium group, which probably stabilizes this conformation. This unusual stacking, which places the triazolium group and the BHQ unit in close proximity to the ZnTPP, might also be responsible for the unusual NMR chemical shifts observed for the NMe group of the linker, the methoxy group of the BHQ moiety and the methyl group of the triazolium unit (vide supra). In general, most of the investigated *closed* conformations of this protonation state were thermodynamically preferred over (half-)open conformations (see SI Table S4). This result substantiates the previous finding that H-1a²⁺ tends to adopt a strongly folded conformation in which the ZnTPP unit is coordinated intramolecularly.

However, for the deprotonated 1a⁺ state, half-open conformations tend to be more stable than *closed* conformations (see Table S5 in the SI). Moreover, several distinct conformers were thermodynamically significant. This reaffirms the higher flexibility observed in the conformer searches. Two relevant conformers are depicted in Fig. 6b, c. These structures are energetically only separated by 8.7 kcal/mol, although their conformations are quite different: In both conformers, the crown-ether encircles the methyltriazolium station, as expected for the deprotonated rotaxane state. Here, the methyltriazolium group interacts with the macrocycle by π -stacking (for the lower-energy conformer only) and/or by CH⁺⋯O interactions of the aromatic CH proton (both conformers, $d(\text{CH}^+\cdots\text{O}) = 2.30\text{--}2.82 \text{ \AA}$). In the lower-energy conformer, all three aromatic units of the BHQ moiety showed π -contacts with other aromatic units. The terminal nitrophenyl group is located above the ZnTPP unit (as opposed to H-1a²⁺, where the aniline group of BHQ is located above ZnTPP), whereas the dimethoxybenzene and the aniline moieties of the BHQ unit stack on top of the benzene rings of

Fig. 6 | Side and top views of relevant conformers. **a** Conformers with the lowest Gibbs free energies of **H-1a²⁺**. **b** Conformers with the lowest Gibbs free energies of **1a⁺**. **c** Another relevant **1a⁺** conformer with its free energy relative to the lowest conformation. The coloring of the carbon atoms is equivalent to Fig. 2. For better clarity, the triazolium station is not colored, and all H-atoms (except for NH₂) are excluded.



ZnTPP and the dibenzo[24]crown[8] macrocycle, respectively. In contrast, the higher-energy conformer shows no intermolecular interactions involving the ZnTPP group, and the entire BHQ unit is pointed away from the ZnTPP (see Fig. 6c).

In conclusion, **H-1a²⁺** is mostly found in one dominant, strongly folded (*closed*) conformation, with the ZnTPP unit being coordinated by the BHQ unit. In contrast, **1a⁺** also adopts conformers that are in a *half-open* conformation and proved to be more flexible. As a consequence, the ZnTPP unit in **1a⁺** is less frequently in close spatial proximity to the BHQ moiety. This may influence the ¹O₂ production of the rotaxane in two ways: First, a close proximity of ZnTPP and the quencher will facilitate energy transfer from the ZnTPP triplet excited state, thus inhibiting ¹O₂ production. Second, a direct contact of the BHQ with the ZnTPP might also inhibit diffusion of ³O₂ towards ZnTPP, thus preventing energy transfer to ³O₂. The computational data is therefore consistent with experimental findings: The strongly folded conformation of **H-1a²⁺** leads to an almost complete shut-down of the ¹O₂ production ($\Phi_{\Delta} = 0.03$, c.f. $\Phi_{\Delta} = 0.72$ for free ZnTPP, see Table 1). In **1a⁺** there are multiple accessible conformations, only some of which feature a close contact between ZnTPP and the BHQ unit. This leads to a considerable, but not complete quenching of the ¹O₂ production ($\Phi_{\Delta} = 0.15$) for this protonation state. Together this results in the reversed direction of switching compared to the original design concept.

To further substantiate these qualitative correlations, we also briefly investigated whether the computed structures of **H-1a²⁺** and **1a⁺** lead to a Förster radius comparable to what is experimentally expected. Due to the size of the BHQ unit, which features three aromatic rings connected by two

diazo groups, it was unclear which position of the BHQ unit should be used to measure intramolecular distances. Thus, this attempt was discarded (for more details, see SI section 2.3).

Conformational search and Gibbs free energies for rotaxanes **H-1b²⁺**/**1b⁺**

We also examined the short variants **H-1b²⁺** and **1b⁺** using the same approach described above. The conducted simulations showed a reduced flexibility for both variants, which likely resulted from the shorter alkylene spacer. Mostly *closed* conformations were found, with many binding motifs being similar for both protonation states. Gibbs free energies of 16 conformers of **H-1b²⁺** and 17 conformers of **1b⁺** were calculated (see SI Table S6 and Table S7). We found that *closed* conformations are thermodynamically favorable for both **H-1b²⁺** and **1b⁺**, and their Gibbs free energy distributions are considerably more similar than those of **H-1a²⁺** and **1a⁺**. Thus, for both **1b** protonation states, the ZnTPP is in close contact with the BHQ unit, suggesting a low singlet oxygen quantum yield. This result is in line with experimental findings, where **H-1b²⁺** and **1b⁺** showed the same vanishing ¹O₂ production ($\Phi_{\Delta} < 0.01/0.01$), similar to that of **H-1a²⁺**.

Conclusions

Based on our original concept for a pH-switchable rotaxane for singlet oxygen production, we have conducted an in-depth study that highlights the importance of molecular folding in the design of rotaxane-based switches.

We synthesized rotaxanes **1a/b**, which consist of a ZnTPP-appended macrocycle and a thread featuring a BHQ unit as one of the stoppers.

Both systems underwent pH-induced switching leading to reversible relocation of the macrocycle between an amine/ammonium and a triazolium station, as evidenced by NMR. For rotaxane **1b**, with a shorter C₃ linker, no ¹O₂ production was observed independent of the switching state. However, for rotaxane **1a**, which features a longer C₉ linker between both stations, the switching has a strong effect on ¹O₂ production, with the deprotonated rotaxane **1a**⁺ representing the *on*-state ($\Phi_{\Delta} = 0.15$), while the protonated rotaxane **H-1a**²⁺ represents the *off*-state ($\Phi_{\Delta} = 0.03$). Thus, we achieved the important goal of generating a system that can be switched *off* almost completely, minimizing unwanted side-effects of ¹O₂, while the *on*-state is significantly more active (500% increase ¹O₂ production). However, the direction of switching is inverted with regard to the original design, where we placed the BHQ unit close to the triazolium station, assuming that in the deprotonated rotaxane **1a**⁺ the positioning of the macrocycle around the triazolium would lead to suppressed ¹O₂ production (and vice versa).

This unexpected behavior (no ¹O₂ production for either **1b**⁺/**H-1b**²⁺, ¹O₂ production for **1a**⁺ but not for **H-1a**²⁺) was explained by in-depth theoretical calculations of the conformational space of these systems: In the protonated state **H-1a**²⁺, the ZnTPP unit does not have a larger distance from the BHQ-stopper (as intended), but intramolecular folding actually leads to one energetically preferred closed conformation with a strong interaction between the ZnTPP unit and the BHQ moiety. In contrast, **1a**⁺ adopts several energetically similar conformations, including half-open conformations without intramolecular contacts with ZnTPP. This explains why **1a**⁺ represents the *on*-state of the system, based on the ability of the ZnTPP to mediate ¹O₂ production without interference from the quencher. As for **1b**⁺/**H-1b**²⁺, the conformational analysis also showed mostly closed conformations with intramolecular interactions between ZnTPP and BHQ, thus explaining suppressed ¹O₂ photosensitization for these rotaxanes.

These studies highlight the importance of taking into account molecular flexibility when designing functional switchable systems, such as switchable photosensitizers (as in this study), as well as other applications. Especially for highly charged species (such as **H-1a**²⁺/**H-1b**²⁺) that can additionally interact via favorable intramolecular interactions such as π -stacking, strongly folded conformations can be preferred. In the present case, even seemingly small structural units, such as the C₂H₄NMe linker between the triazole and the BHQ unit, can have a significant impact if they allow the formation of energetically relevant folded structures.

Thus, we have not only presented the synthesis of pH-switchable rotaxanes for singlet oxygen production, but also developed a deeper understanding of the design of bistable rotaxanes for the development of functional switchable systems in general.

Methods

Materials

For a list of the commercially available chemicals used, see SI section 1.1.1. Known compounds were prepared according to literature procedures (see SI section 1.2.1, Figs. S1–S3).

Standard analytical methods

NMR spectra were recorded with a Bruker Avance NEO 400 or a Bruker DRX 600 spectrometer at 298 K using CDCl₃, CD₂Cl₂, THF-*d*₈, or DMSO-*d*₆ as the solvent. The chemical shifts are referenced relative to the residual proton signals of the solvent in ¹H-NMR or the signal of the solvent in ¹³C-NMR. Further instrumental details are given in section 1.1.2 of the SI. Full characterization details of all new compounds are given in section 1.2.2 of the SI. For NMR spectra of all new compounds, see section 3 of the SI (Figs. S49–S68). UV/Vis-absorption spectra were recorded on a Varian Cary 300 Bio UV-Vis spectrophotometer in spectrophotometric grade tetrahydrofuran. Fluorescence spectra were recorded on a Varian Eclipse fluorescence spectrophotometer. Time-resolved measurements were carried out on a FluoTime 300 spectrometer from PicoQuant.

Synthesis of rotaxanes

All rotaxanes (**1a/b**, **2** and **3**) were prepared by the following general procedure: Dibenzylammonium half-thread **4a/b-HPF₆** (1 eq.), crown-ether macrocycle **5** (1.2 eq.) and the corresponding azide (1.3 eq.) were dissolved in DCM (4 mL per mmol of **4a/b-HPF₆**), degassed by purging with argon for five minutes and stirred with tetrakis(acetonitrile)copper(I) hexafluorophosphate (1.5 eq.) overnight. The rotaxanes were purified by flash column chromatography (SiO₂, DCM:MeOH). For methylation of the triazole, the rotaxanes were stirred in methyl iodide (0.1 mL/mg) until full conversion of the starting material was observed by thin layer chromatography (2–10 days). Excess methyl iodide was removed in vacuo and the methylated compound was purified by flash column chromatography (SiO₂, DCM:MeOH) if necessary. The methylated rotaxane was then dissolved in DCM (2 mL) and stirred over solid ammonium-hexafluorophosphate (20 eq.) for 20 hours, filtered over a polyamide syringe filter, and dried under reduced pressure. For details, see section 1.2.2 of the SI.

Switching of rotaxanes

The protonated rotaxane-bis(hexafluorophosphate) was dissolved in acetone (2 mL/10 mg), and 50 mM aqueous sodium hydroxide (2 eq.) was added. Volatiles were removed *in vacuo* and the residue was taken up in dichloromethane and filtered over a polyamide syringe filter. The filtrate was evaporated to yield the deprotonated rotaxane (see SI Figs. S12–S18).

Investigation of singlet dioxygen photoproduction

Approximately 10 μ M solutions of free porphyrin or the rotaxanes were prepared in THF under ambient conditions and the absorbance of the most intense band (*Soret* band around 424 nm) was adjusted to 0.8. Then, 70 μ L DPBF solution (860 μ M) was added in the dark. Then the cuvette was irradiated with 420 nm light in intervals of 10 or 30 seconds and fluorescence emission spectra ($\lambda_{\text{Exc}} = 420$ nm) were recorded to monitor the decrease in fluorescence intensity (*I*) of DPBF at 457 nm over time (see SI Fig. S19).

Photophysical characterization

Steady-state excitation and emission spectra were recorded on a FluoTime 300 spectrometer from PicoQuant (a full description of the equipment can be found in section 1.4.3 of the SI). Steady-state spectra and photoluminescence lifetimes were recorded in TCSPC mode by a PicoHarp 300 (minimum base resolution 4 ps). Emission and excitation spectra were corrected for source intensity (lamp and grating) by standard correction curves. An instrument response function calibration was performed using a diluted Ludox[®] dispersion. Lifetime analysis was performed using the commercial EasyTau 2 software (PicoQuant). The quality of the fit was assessed by minimizing the reduced chi-squared function (χ^2) and visual inspection of the weighted residuals and their autocorrelation (for further details, see SI section 1.4.3, Figs. S25–S42).

Computational details

We employed the Conform-Rotamer Ensemble Sampling Tool CREST⁸⁷ (v. 2.12) in combination with the semi-empirical quantum mechanical (SQM) methods GFN n -xTB ($n = 1, 2$)^{83–85} and the force field GFN-FF⁸⁶, applying the implicit solvation model ALPB (THF)⁹¹. Chosen structures were optimized at the GFN2-xTB [ALPB:THF] level of theory. Thermal corrections (G_{thermo}) were computed at the same level by employing the modified rigid-rotor-harmonic-oscillator approximation (mRRHO)⁸⁸. This was done using the *xtb* (v. 6.5.1) program (<https://github.com/grimme-lab/xtb>). Utilizing the TURBOMOLE (v. 7.5.1) (<https://www.turbomole.org>)⁹² program package, electronic gas-phase energies (E_{gas}), and solvation contributions (∂G_{solv}) were computed using the PBEh-3c composite DFT method⁸⁹. The solvation contributions were computed

with the COSMO-RS⁹⁰ implicit solvation model. Finally, Gibbs free energies (G_{Gibbs}) were calculated using:

$$G_{\text{Gibbs}} = E_{\text{gas}} + G_{\text{thermo}} + \partial G_{\text{solv}}$$

To facilitate comparison, only relative values ($\Delta X = X - X_{\text{ref}}$) are discussed, with the smallest value being defined as the respective reference value (X_{ref}).

Division of the found conformations into three categories *open*, *half-open*, and *closed* is based on the degree of intramolecular folding exhibited by the structures. As a quantification the computed Solvent Accessible Surface Area (SASA) of each conformer was used (<https://github.com/grimme-lab/numsa>). In the *open* conformers no or minimal folding is observed, and the Zn atom in the ZnTPP unit is not coordinated in its axial (octahedral) plane. The *closed* conformations exhibit a high degree of intramolecular folding, and (mostly) a coordination of the Zn atom in the aforementioned plane. *Half-open* conformers show a moderate degree of folding, but in some cases still exhibit a coordinating group on the Zn atom. For more details, see section 2.1 of the SI.

Data availability

The Supplementary information contains all details of the synthesis and characterization of novel compounds, photophysical measurements, determination of singlet oxygen production, and details regarding the theoretical calculations and their evaluations. Supplementary data 1 include the xyz-files of the calculated conformers of the rotaxanes. Supplementary data 2 include the numerical source data for the determination of the singlet oxygen quantum yields.

Received: 15 March 2024; Accepted: 18 July 2024;

Published online: 07 August 2024

References

- Feringa, B. L. & Browne, W. R. (Eds.) Molecular switches. Wiley-VCH, Weinheim (2011).
- Erbas-Cakmak, S., Leigh, D. A., McTernan, C. T. & Nussbaumer, A. L. Artificial molecular machines. *Chem. Rev.* **115**, 10081–10206 (2015).
- Zhang, L., Marcos, V. & Leigh, D. A. Molecular machines with bio-inspired mechanisms. *Proc. Natl Acad. Sci. USA* **115**, 9397–9404 (2018).
- Wang, H., Bisoyi, H. K., Zhang, X., Hassan, F. & Li, Q. Visible light-driven molecular switches and motors: recent developments and applications. *Chem. Eur. J.* **28**, e202103906 (2022).
- Bruns, C. J. & Stoddart, J. F. Rotaxane-based molecular muscles. *Acc. Chem. Res.* **47**, 2186–2199 (2014).
- Antoine, A., Moulin, E., Fuks, G. & Giuseppone, N. [c2]Daisy chain rotaxanes as molecular muscles. *CCS Chem.* **1**, 83–96 (2019).
- Karimi, M., Mirshekari, H., Aliakbari, M., Sahandi-Zangabad, P. & Hamblin, M. R. Smart mesoporous silica nanoparticles for controlled-release drug delivery. *Nanotechnol. Rev.* **5**, 195–207 (2016).
- Chen, W., Glackin, C. A., Horwitz, M. A. & Zink, J. I. Nanomachines and other caps on mesoporous silica nanoparticles for drug delivery. *Acc. Chem. Res.* **52**, 1531–1542 (2019).
- Dattler, D. et al. Design of collective motions from synthetic molecular switches, rotors, and motors. *Chem. Rev.* **120**, 310–433 (2019).
- Leigh, D. A., Marcos, V. & Wilson, M. R. Rotaxane catalysts. *ACS Catal.* **4**, 4490–4497 (2014).
- Blanco, V., Leigh, D. A. & Marcos, V. Artificial switchable catalysts. *Chem. Soc. Rev.* **44**, 5341–5370 (2015).
- Martinez-Cuezva, A., Saura-Sanmartin, A., Alajarin, M. & Berna, J. Mechanically interlocked catalysts for asymmetric synthesis. *ACS Catal.* **10**, 7719–7733 (2020).
- Kwamen, C. & Niemeyer, J. Functional rotaxanes in catalysis. *Chem. Eur. J.* **27**, 175–186 (2021).
- Heard, A. W., Suárez, J. M. & Goldup, S. M. Controlling catalyst activity, chemoselectivity and stereoselectivity with the mechanical bond. *Nat. Rev. Chem.* **6**, 182–196 (2022).
- Kwan, C.-S., Chan, A. S. C. & Leung, K. C.-F. A fluorescent and switchable rotaxane dual organocatalyst. *Org. Lett.* **18**, 976–979 (2016).
- Ma, X. et al. A room temperature phosphorescence encoding [2] rotaxane molecular shuttle. *Chem. Sci.* **7**, 4582–4588 (2016).
- Ghosh, A., Paul, I., Adlung, M., Wickleder, C. & Schmittl, M. Oscillating emission of [2]rotaxane driven by chemical fuel. *Org. Lett.* **20**, 1046–1049 (2018).
- Neal, E. A. & Goldup, S. M. Chemical consequences of mechanical bonding in catenanes and rotaxanes: isomerism, modification, catalysis and molecular machines for synthesis. *Chem. Commun.* **50**, 5128–5142 (2014).
- van Dongen, S. F. M., Cantekin, S., Elemans, J. A. A. W., Rowan, A. E. & Nolte, R. J. M. Functional interlocked systems. *Chem. Soc. Rev.* **43**, 99–122 (2014).
- Lewis, J. E., Galli, M. & Goldup, S. M. Properties and emerging applications of mechanically interlocked ligands. *Chem. Commun.* **53**, 298–312 (2016).
- Sluysmans, D. & Stoddart, J. F. The burgeoning of mechanically interlocked molecules in chemistry. *Trends Chem.* **1**, 185–197 (2019).
- Wu, P., Dharmadhikari, B., Patra, P. & Xiong, X. Rotaxane nanomachines in future molecular electronics. *Nanoscale Adv.* **4**, 3418–3461 (2022).
- Bukhtiarova, N., Credi, A. & Corra, S. Controlling molecular shuttling in a rotaxane with weak ring recognition sites. *Chem. Commun.* **59**, 13159–13162 (2023).
- Coutrot, F. A focus on triazolium as a multipurpose molecular station for pH-sensitive interlocked crown-ether-based molecular machines. *ChemistryOpen* **4**, 556–576 (2015).
- Ragazzon, G., Credi, A. & Colasson, B. Thermodynamic insights on a bistable acid–base switchable molecular shuttle with strongly shifted co-conformational equilibria. *Chem. Eur. J.* **23**, 2149–2156 (2017).
- Eichstaedt, K. et al. Switching between anion-binding catalysis and aminocatalysis with a rotaxane dual-function catalyst. *J. Am. Chem. Soc.* **139**, 9376–9381 (2017).
- Blanco, V., Leigh, D. A., Marcos, V., Morales-Serna, J. A. & Nussbaumer, A. L. A switchable [2]rotaxane asymmetric organocatalyst that utilizes an acyclic chiral. *Secondary Amine J. Am. Chem. Soc.* **136**, 4905–4908 (2014).
- Blanco, V., Leigh, D. A., Lewandowska, U., Lewandowski, B. & Marcos, V. Exploring the activation modes of a rotaxane-based switchable organocatalyst. *J. Am. Chem. Soc.* **136**, 15775–15780 (2014).
- Blanco, V., Carlone, A., Hänni, K. D., Leigh, D. A. & Lewandowski, B. A rotaxane-based switchable organocatalyst. *Angew. Chem. Int. Ed.* **51**, 5166–5169 (2012).
- Du, G., Moulin, E., Jouault, N., Buhler, E. & Giuseppone, N. Muscle-like supramolecular polymers: integrated motion from thousands of molecular machines. *Angew. Chem. Int. Ed.* **51**, 12504–12508 (2012).
- Coutrot, F., Romuald, C. & Busseron, E. A new pH-switchable dimannosyl[c2]daisy chain molecular machine. *Org. Lett.* **10**, 3741–3744 (2008).
- Li, H. et al. A switchable bis-branched [1]rotaxane featuring dual-mode molecular motions and tunable molecular aggregation. *ACS Appl. Mater. Interfaces* **6**, 18921–18929 (2014).
- Li, H., Li, X., Ågren, H. & Qu, D.-H. Two switchable star-shaped [1] rotaxanes with different multibranched cores. *Org. Lett.* **16**, 4940–4943 (2014).
- Cao, Z. Q. et al. A perylene-bridged switchable [3]rotaxane molecular shuttle with a fluorescence output. *Asian J. Org. Chem.* **4**, 212–216 (2014).

35. Zhou, W. et al. A bis-spiropyran-containing multi-state [2]rotaxane with fluorescence output. *Tetrahedron* **69**, 5319–5325 (2013).
36. Li, H. et al. Dual-mode operation of a bistable [1]rotaxane with a fluorescence signal. *Org. Lett.* **15**, 3070–3073 (2013).
37. Yang, W. et al. Synthesis of a [2]rotaxane operated in basic environment. *Org. Biomol. Chem.* **9**, 6022 (2011).
38. Jiang, Q., Zhang, H.-Y., Han, M., Ding, Z.-J. & Liu, Y. pH-controlled intramolecular charge-transfer behavior in bistable [3]rotaxane. *Org. Lett.* **12**, 1728–1731 (2010).
39. Maharjan, P. S. & Bhattarai, H. K. Singlet oxygen, photodynamic therapy, and mechanisms of cancer cell death. *J. Oncol.* **2022**, 1–20 (2022).
40. Correia, J. H., Rodrigues, J. A., Pimenta, S., Dong, T. & Yang, Z. Photodynamic therapy review: principles, photosensitizers, applications, and future directions. *Pharmaceutics* **13**, 1332 (2021).
41. Alessandri, I. & Vassalini, I. Oxygen-mediated surface photoreactions: exploring new pathways for sustainable chemistry. *ChemPhotoChem* **7**, e202300069 (2023).
42. García-Fresnadillo, D. Singlet oxygen photosensitizing materials for point-of-use water disinfection with solar reactors. *ChemPhotoChem* **2**, 512–534 (2018).
43. Kalaitzakis, D., Sofiadis, M., Tsopanakis, V., Montagnon, T. & Vassilikogiannakis, G. Merging singlet-oxygen induced furan oxidations with organocatalysis: synthesis of enantiopure cyclopentanones and hydrindanes. *Org. Biomol. Chem.* **18**, 2817–2822 (2020).
44. Bayer, P., Pérez-Ruiz, R. & Jacobi von Wangelin, A. Stereoselective photooxidations by the schenck ene reaction. *ChemPhotoChem* **2**, 559–570 (2018).
45. Ghogare, A. A. & Greer, A. Using singlet oxygen to synthesize natural products and drugs. *Chem. Rev.* **116**, 9994–10034 (2016).
46. Montagnon, T., Kalaitzakis, D., Triantafyllakis, M., Stratakis, M. & Vassilikogiannakis, G. Furans and singlet oxygen – why there is more to come from this powerful partnership. *Chem. Commun.* **50**, 15480–15498 (2014).
47. Li, Z. et al. Photoswitchable diarylethenes: from molecular structures to biological applications. *Coord. Chem. Rev.* **497**, 215451 (2023).
48. Pham, T. C., Nguyen, V.-N., Choi, Y., Lee, S. & Yoon, J. Recent strategies to develop innovative photosensitizers for enhanced photodynamic therapy. *Chem. Rev.* **121**, 13454–13619 (2021).
49. Liu, J. et al. Recent progress in the development of multifunctional nanoplatfrom for precise tumor phototherapy. *Adv. Healthc. Mater.* **10**, 2001207 (2021).
50. Li, X., Lee, S. & Yoon, J. Supramolecular photosensitizers rejuvenate photodynamic therapy. *Chem. Soc. Rev.* **47**, 1174–1188 (2018).
51. Bonnet, S. Why develop photoactivated chemotherapy? *Dalton Trans.* **47**, 10330–10343 (2018).
52. Wu, W., Shao, X., Zhao, J. & Wu, M. Controllable photodynamic therapy implemented by regulating singlet oxygen efficiency. *Adv. Sci. Sci.* **4**, 1700113 (2017).
53. Li, X. et al. A tumor-pH-responsive supramolecular photosensitizer for activatable photodynamic therapy with minimal in vivo skin phototoxicity. *Theranostics* **7**, 2746–2756 (2017).
54. Nomoto, T. et al. Calcium phosphate-based organic–inorganic hybrid nanocarriers with pH-responsive on/off switch for photodynamic therapy. *Biomater. Sci.* **4**, 826–838 (2016).
55. Chen, W.-H. et al. Mesoporous silica-based versatile theranostic nanoplatfrom constructed by layer-by-layer assembly for excellent photodynamic/chemo therapy. *Biomater* **117**, 54–65 (2017).
56. Tong, H. et al. Glutathione activatable photosensitizer-conjugated pseudopolyrotaxane nanocarriers for photodynamic theranostics. *Small* **12**, 6223–6232 (2016).
57. Fan, H. et al. A smart photosensitizer–manganese dioxide nanosystem for enhanced photodynamic therapy by reducing glutathione levels in cancer cells. *Angew. Chem. Int. Ed.* **55**, 5477–5482 (2016).
58. Liu, F., Ma, Y., Xu, L., Liu, L. & Zhang, W. Redox-responsive supramolecular amphiphiles constructed via host–guest interactions for photodynamic therapy. *Biomater. Sci.* **3**, 1218–1227 (2015).
59. Kim, W. L., Cho, H., Li, L., Kang, H. C. & Huh, K. M. Biarmed poly(ethylene glycol)-(pheophorbidea)2 conjugate as a bioactivatable delivery carrier for photodynamic therapy. *Biomacromolecules* **15**, 2224–2234 (2014).
60. Cho, Y., Kim, H. & Choi, Y. A graphene oxide–photosensitizer complex as an enzyme-activatable theranostic agent. *Chem. Commun.* **49**, 1202 (2013).
61. Yuan, Y., Kwok, R. T. K., Tang, B. Z. & Liu, B. Cancer therapy: smart probe for tracing cancer therapy: selective cancer cell detection, image-guided ablation, and prediction of therapeutic response in situ. *Small* **11**, 4606–4606 (2015).
62. Park, S. Y. et al. A smart polysaccharide/drug conjugate for photodynamic therapy. *Angew. Chem. Int. Ed.* **50**, 1644–1647 (2011).
63. Erbas-Cakmak, S., Bozdemir, O. A., Cakmak, Y. & Akkaya, E. U. Proof of principle for a molecular 1:2 demultiplexer to function as an autonomously switching theranostic device. *Chem. Sci.* **4**, 858–862 (2013).
64. Jiang, X.-J., Lo, P.-C., Yeung, S.-L., Fong, W.-P. & Ng, D. K. P. A pH-responsive fluorescence probe and photosensitizer based on a tetraamino silicon(iv) phthalocyanine. *Chem. Commun.* **46**, 3188 (2010).
65. Ozlem, S. & Akkaya, E. U. Thinking outside the silicon box: molecular AND logic as an additional layer of selectivity in singlet oxygen generation for photodynamic therapy. *J. Am. Chem. Soc.* **131**, 48–49 (2008).
66. Liu, G. et al. A highly efficient supramolecular photoswitch for singlet oxygen generation in water. *Chem. Commun.* **52**, 7966–7969 (2016).
67. Hou, L., Zhang, X., Pijper, T. C., Browne, W. R. & Feringa, B. L. Reversible photochemical control of singlet oxygen generation using diarylethene photochromic switches. *J. Am. Chem. Soc.* **136**, 910–913 (2014).
68. Cló, E., Snyder, J. W., Voigt, N. V., Ogilby, P. R. & Gothelf, K. V. DNA-programmed control of photosensitized singlet oxygen production. *J. Am. Chem. Soc.* **128**, 4200–4201 (2006).
69. Pairault, N. et al. Rotaxane-based architectures for biological applications. *C. R. Chim.* **19**, 103–112 (2016).
70. Riebe, J. & Niemeyer, J. Mechanically interlocked molecules for biomedical applications. *Eur. J. Org. Chem.* **2021**, 5106–5116 (2021).
71. Yu, G., Yung, B. C., Zhou, Z., Mao, Z. & Chen, X. Artificial molecular machines in nanotheranostics. *ACS Nano* **12**, 7–12 (2017).
72. Shao, L. et al. Constructing adaptive photosensitizers via supramolecular modification based on pillararene host–guest interactions. *Angew. Chem. Int. Ed.* **59**, 11779–11783 (2020).
73. Khang, T. M. et al. Dual and sequential locked/unlocked photochromic effects on FRET controlled singlet oxygen processes by contracted/extended forms of diarylethene-based [1]rotaxane nanoparticles. *Small* **19**, 2205597 (2022).
74. Shinohara, A., Pan, C., Wang, L. & Shinmori, H. Acid–base controllable singlet oxygen generation in supramolecular porphyrin–gold nanoparticle composites tethered by rotaxane linkers. *J. Porph. Phthalocyan.* **24**, 171–180 (2020).
75. Hewson, S. W. & Mullen, K. M. Porphyrin-containing rotaxane assemblies. *Eur. J. Org. Chem.* **2019**, 3358–3370 (2019).
76. Taniguchi, M., Lindsey, J. S., Bocian, D. F. & Holten, D. Comprehensive review of photophysical parameters (ϵ , Φ_f , τ_S) of tetraphenylporphyrin (H2TPP) and zinc tetraphenylporphyrin (ZnTPP) – critical benchmark molecules in photochemistry and photosynthesis. *J. Photochem. Photobiol. C: Photochem. Rev.* **46**, 100401 (2021).
77. Wilkinson, F., Helman, W. P. & Ross, A. B. Quantum yields for the photosensitized formation of the lowest electronically excited singlet

- state of molecular oxygen in solution. *J. Phys. Chem. Ref. Data* **22**, 113–262 (1993).
78. Chevalier, A., Renard, P.-Y. & Romieu, A. Straightforward synthesis of bioconjugatable azo dyes. Part 2: black hole quencher-2 (BHQ-2) and blackberry quencher 650 (BBQ-650) scaffolds. *Tetrahedron Lett.* **55**, 6764–6768 (2014).
 79. Konev, A. S., Khlebnikov, A. F., Levin, O. V., Lukyanov, D. A. & Zorin, I. M. Photocurrent in multilayered assemblies of porphyrin–fullerene covalent dyads: evidence for channels for charge transport. *ChemSusChem* **9**, 676–686 (2016).
 80. Feng, D.-J., Li, X.-Q., Wang, X.-Z., Jiang, X.-K. & Li, Z.-T. Highly stable pseudo[2]rotaxanes co-driven by crown ether–ammonium and donor–acceptor interactions. *Tetrahedron* **60**, 6137–6144 (2004).
 81. Entradas, T., Waldron, S. & Volk, M. The detection sensitivity of commonly used singlet oxygen probes in aqueous environments. *J. Photochem. Photobiol. B: Biol.* **204**, 111787 (2020).
 82. Gray, V. et al. Porphyrin–anthracene complexes: potential in triplet–triplet annihilation upconversion. *J. Phys. Chem. C* **120**, 19018–19026 (2016).
 83. Bannwarth, C. et al. Extended tight-binding quantum chemistry methods. *WIREs Comput. Mol. Sci.* **11**, e1493 (2020).
 84. Bannwarth, C., Ehlert, S. & Grimme, S. GFN2-xTB—an accurate and broadly parametrized self-consistent tight-binding quantum chemical method with multipole electrostatics and density-dependent dispersion contributions. *J. Chem. Theory Comput.* **15**, 1652–1671 (2019).
 85. Grimme, S., Bannwarth, C. & Shushkov, P. A robust and accurate tight-binding quantum chemical method for structures, vibrational frequencies, and noncovalent interactions of large molecular systems parametrized for all spd-block elements (Z = 1–86). *J. Chem. Theory Comput.* **13**, 1989–2009 (2017).
 86. Spicher, S. & Grimme, S. Robust atomistic modeling of materials, organometallic, and biochemical systems. *Angew. Chem. Int. Ed.* **59**, 15665–15673 (2020).
 87. Pracht, P., Bohle, F. & Grimme, S. Automated exploration of the low-energy chemical space with fast quantum chemical methods. *Phys. Chem. Chem. Phys.* **22**, 7169–7192 (2020).
 88. Spicher, S. & Grimme, S. Single-point Hessian calculations for improved vibrational frequencies and rigid-rotor-harmonic-oscillator thermodynamics. *J. Chem. Theory Comput.* **17**, 1701–1714 (2021).
 89. Grimme, S., Brandenburg, J. G., Bannwarth, C. & Hansen, A. Consistent structures and interactions by density functional theory with small atomic orbital basis sets. *J. Chem. Phys.* **143**, 054107 (2015).
 90. Klamt, A. The COSMO and COSMO-RS solvation models. *WIREs Comput. Mol. Sci.* **1**, 699–709 (2011).
 91. Ehlert, S., Stahn, M., Spicher, S. & Grimme, S. Robust and efficient implicit solvation model for fast semiempirical methods. *J. Chem. Theory Comput.* **17**, 4250–4261 (2021).
 92. Balasubramani, S. G. et al. TURBOMOLE: modular program suite for ab initio quantum-chemical and condensed-matter simulations. *J. Chem. Phys.* **152**, 184107 (2020).

Acknowledgements

J.N. would like to thank the Deutsche Forschungsgemeinschaft DFG (project NI1273/2-2 and Heisenberg-Professorship NI1273/4-1) for funding. C.A.S. gratefully acknowledges funding from the Deutsche

Forschungsgemeinschaft (DFG, German Research Foundation)—project-ID 433682494 - SFB 1459 “Intelligent Matter”. C.A.S. gratefully acknowledges the generous financial support for the acquisition of an “Integrated Confocal Luminescence Spectrometer with Spatiotemporal Resolution and Multiphoton Excitation” (DFG/Land NRW: INST 211/915-1 FUGG; DFG EXC1003: Berufungsmittel) and of a “Laser-Induced Transient Absorption and Photoacoustic Spectrometer” (DFG/Land NRW: INST 211/1033-1 FUGG; DFG EXC1003: Berufungsmittel).

Author contributions

J.R. performed the synthetic chemistry. J.R., M.E.G.S., and M.B.R.A. performed the photophysical characterization and singlet dioxygen measurements. B.B. and S.L. performed the theoretical calculations. J.R. and J.N. devised the project. C.A.S., S.G., and J.N. directed the research. All authors contributed to the writing of the manuscript.

Funding

Open Access funding enabled and organized by Projekt DEAL.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s42004-024-01247-7>.

Correspondence and requests for materials should be addressed to Stefan Grimme or Jochen Niemeyer.

Peer review information *Communications Chemistry* thanks Stefano Crespi and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. A peer review file is available.

Reprints and permissions information is available at <http://www.nature.com/reprints>

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2024

It's Complicated: On Relativistic Effects and Periodic Trends in the Melting and Boiling Points of the Group 11 Coinage Metals

Sarah Löffelsender¹, Peter Schwerdtfeger², Stefan Grimme¹, and Jan-Michael Mewes¹

Received: October 14, 2021

First published: December 29, 2021

Reprinted with permission from[‡]:

S. Löffelsender, P. Schwerdtfeger, S. Grimme, J.-M. Mewes, *Journal of the American Chemical Society*, **144**(1) (2022), 485-494,

DOI: 10.1021/jacs.1c10881

Copyright © 2021 American Chemical Society

Own manuscript contributions:

- Preparation and performance of all calculations excluding non-relativistic potential generation
- Evaluation and interpretation of the results
- Visualization of all results
- Writing the initial draft of the manuscript
- Revising of the manuscript

¹ Mulliken Center for Theoretical Chemistry, Rheinische Friedrich-Wilhelms-Universität Bonn, Beringstr. 4, 53115 Bonn, Germany

² Centre for Theoretical Chemistry and Physics, The New Zealand Institute for Advanced Study, Massey University, Auckland Campus, 0632 Auckland, New Zealand

[‡] Permission requests to reuse material from this chapter should be directed to American Chemical Society

It's Complicated: On Relativistic Effects and Periodic Trends in the Melting and Boiling Points of the Group 11 Coinage Metals

Sarah Löffelsender, Peter Schwerdtfeger, Stefan Grimme, and Jan-Michael Mewes*

Cite This: *J. Am. Chem. Soc.* 2022, 144, 485–494

Read Online

ACCESS |



Metrics & More



Article Recommendations



Supporting Information

ABSTRACT: While the color of metallic gold is a prominent and well-investigated example for the impact of relativistic effects, much less is known regarding the influence on its melting and boiling point (MP/BP). To remedy this situation, this work takes on the challenging task of exploring the phase transitions of the Group 11 coinage metals Cu, Ag, and Au through nonrelativistic (NR) and scalar/spin–orbit relativistic (SR/SOR) Gibbs energy calculations with λ -scaled density-functional theory (λ DFT). At the SOR level, the calculations provide BPs in excellent agreement with experimental values (1%), while MPs exhibit more significant deviations (2–10%). Comparing SOR calculations to those conducted in the NR limit reveals some remarkably large and, at the same time, some surprisingly small relativistic shifts. Most notably, the BP of Au increases by about 800 K due to relativity, which is in line with the strong relativistic increase of the cohesive energy, whereas the MP of Au is very similar at the SOR and NR levels, defying the typically robust correlation between MP and cohesive energy. Eventually, an inspection of thermodynamic quantities traces the trend-breaking behavior of Au back to phase-specific effects in liquid Au, which render NR Au more similar to SOR Ag, in line with a half-a-century-old hypothesis of Pyykkö.



1. INTRODUCTION

The quantum-chemical simulation of phase transitions in extended or infinite systems is a notoriously difficult yet tempting problem whose solution promises detailed insights into the underlying mechanisms. Paradoxically, although melting and boiling are very common physical processes for which we assume an intuitive understanding, there exists no generally accepted microscopic picture. The main reason (besides the complex nature of phase transitions per se) lies in the difficulty of correctly describing the interaction between the atoms and molecules involved,¹ and in the underlying, resource-intensive algorithms for simulating such phase transitions,^{2–6} taking care of different solid state structures, defects, and multiple reaction paths, superheating or supercooling, extrapolations from cluster melting to the infinite system, and large volume changes encountered in liquid-to-gas phase transitions.⁵ As the computational problem becomes very soon intractable, it comes as no surprise that phase transition simulations using molecular dynamics or Monte Carlo algorithms^{1,7} are mostly based on empirical interaction potentials (force field methods) such as Lennard–Jones or more sophisticated model potentials containing effective many-body terms.^{2,8–11} Only for van der Waals type of interactions, where the expansion of the total interaction energy in terms of many-body potentials converges relatively fast,¹² cohesive energies and melting temperatures can be obtained to high accuracy as shown recently for the rare gas solids in Monte

Carlo simulations using many-body potentials derived from relativistic coupled-cluster theory.^{13–16}

For covalently bonded or metallic solids, the correct description of bond weakening and breaking in phase transitions requires an accurate treatment of electron correlation in ab initio methods, which still remains a major challenge for bulk systems.^{17–19} Here, the main quantum chemical working horse for the accurate simulation of phase transitions is density functional theory (DFT).^{20–22} However, the plethora of density functionals currently available has its own consistency problems.^{23–25} To overcome the well-known dilemma of the rather different and element-specific performance of density functionals for bulk properties, Mewes et al. have recently composed a scaled density-functional approach (λ DFT). This approach includes both scalar and spin–orbit relativistic contributions to the Gibbs (free) energy by combining molecular dynamics simulations (MD), thermodynamic integration (TI), and perturbation theory (TPT) as well as an empirical scaling of the Hamiltonian to mitigate

Received: October 14, 2021

Published: December 29, 2021



Table 1. Experimental Dissociation Energies D_e (in eV) and Equilibrium Distances r_e (in Å) for the Group 11 Dimers,^{60–62} Experimental Cohesive Energies E_{coh} (in eV) and Nearest Neighbor Distances r_{NN} (in Å) for the Solid State,⁶³ and Melting T_m and Boiling Points T_b (in K)⁶⁴ for all Group 11 Elements^a

M	D_e	$\Delta_R D_e^{\text{DFT}}$	r_e	$\Delta_R r_e^{\text{DFT}}$	E_{coh}	r_{NN}	T_m	T_b
Cu	2.017	0.155	2.219	−0.032	3.49	2.56	1358	2835
Ag	1.66	0.246	2.48	−0.082	2.95	2.89	1235	2435
Au	2.278	0.938	2.472	−0.245	3.81	2.95	1337	3129

^aRelativistic effects, e.g., $\Delta_R D_e = D_e^{\text{SOR}} - D_e^{\text{NR}}$, from nonrelativistic (NR) and spin-orbit relativistic (SOR) DFT/PBESol calculations of this work.

systematic errors. This TI-MD- λ DFT approach can pin down melting and boiling points (MPs and BPs) by locating the intersection between the Gibbs energies of the respective phases.^{26–28} Previous applications have demonstrated that the approach produces phase transition temperatures, and in particular BPs, with only a few percent deviation from the experimental values,^{27–29} clearing out previous uncertainties concerning the relativistic shift of the MP of Hg.^{22,27}

This work employs the TI-MD- λ DFT method²⁸ to investigate the melting and boiling behavior of the Group 11 metals Cu, Ag, and Au. These elements occur naturally in their elemental form and have been known since prehistoric times,^{30,31} with gold being the most corrosion-resistant and noble metal.^{32,33} Gold is therefore used in many technologies,^{34–37} with most recent applications in nanoscience³⁸ such as in colloidal gold-based immuno-chromatographic devices for COVID-19 rapid tests.³⁹ The Group 11 elements, and in particular Au, show some very interesting and well-known anomalies in the periodic trends of chemical and physical properties.^{36,40} It is now well established that this is due to a particularly strong relativistic 6s contraction/stabilization in Au, the so-called “inert-pair effect”. This originates from the direct action of the relativistic perturbation operator on the 6s density near the nucleus.⁴¹ In addition, the resulting Z^2 scaling of the relativistic ns stabilization for the Group 11 elements is significantly enhanced by the filling of the energetically lower lying soft $(n-1)d$ shell,^{42–44} resulting in a well-known peak of the impact of relativistic effects in Group 11.⁴⁵

As a result, gold has the highest electron affinity (~ 2.4) of any metal,⁴⁶ and is therefore regarded as a pseudohalide, adopting a negative oxidation state of -1 ,⁴⁷ which renders Au an amphoteric oxide.⁴⁰ Furthermore, the proximity of the 5d and 6s shells due to relativistic effects leads to a high stability of the oxidation state $+3$,^{48,49} to the occurrence of the oxidation state $+5$,⁵⁰ to aurophilic interactions,⁵¹ and many more interesting chemical features.^{52–54} Concerning the solid state of Au, we mention the high specific resistivity compared to the lighter congeners,⁴⁴ the ionic nature of AuCs (two metals become a nonmetal by mixing!),^{55,56} the yellow color,⁵⁷ the unusual chain-like structures of Au halides,^{58,59} and the rather large cohesive energy, which is reflected in the bond energy and distance of the Au-dimer⁴⁴ (cf. Table 1).

Taking a step step back from Group 11 and regarding the transition metals as a whole, we notice the typical increase of the cohesive energy for most transition metals when moving down a group in the Periodic Table.

There are three notable exceptions, and all of them are at the end of the transition metal series (see Figure 1). The cohesive energy of the Group 12 metals decreases with increasing nuclear charge, akin to main group elements. For Group 10 and Group 11, the cohesive energy decreases from period 4 to 5 and increases from period 5 to 6. For Group 11, the behavior of the cohesive energy is thus in line with the melting (and boiling)

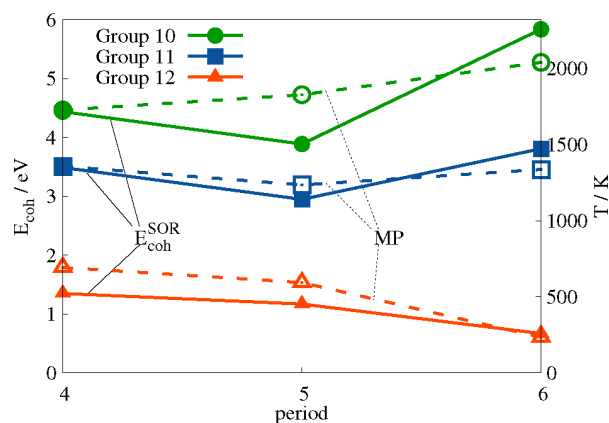


Figure 1. Experimental cohesive energies (solid lines, in eV,⁶³) and melting points (dashed lines, in K) for Group 10 (Ni, Pd, Pt), 11 (Cu, Ag, Au), and 12 (Zn, Cd, Hg).

points, as is evident from Figure 1. It has been argued that the trends observed in the melting and boiling points (see Table 1: $\text{Ag} < \text{Au} \approx \text{Cu}$ for T_m and $\text{Ag} < \text{Cu} < \text{Au}$ for T_b) are due to relativistic effects.⁶⁵ To put this hypothesis to the test, and to explore in detail periodic trends in the phase transitions of the coinage metals, we report relativistic and nonrelativistic Gibbs energy calculations using thermodynamic integration within a DFT framework.

2. APPROACH AND METHODOLOGY

Phase transition occurs when the Gibbs energy of another phase becomes lower than the Gibbs energy of the prevailing phase. Hence, the MP and BP correspond to the intersections of the respective Gibbs energies of the solid, liquid, and gas phase, which are a nonlinear function of the temperature T :

$$G(T) = U(T) - TS(T) + pV \approx U - TS \quad (1)$$

However, over small temperature ranges (tens of K), it is typically a reasonable approximation to neglect the temperature dependence of U and S , rendering $G(T)$ a linear function with the slope S .

To obtain the Gibbs energies and entropies of the solid and liquid phases at a given temperature T_{sim} , we employ the recent TI-MD- λ DFT approach of Mewes et al.^{26,28} This method augments and combines the upsampled thermodynamic integration using Langevin dynamics (UP-TILD) approach of Neugebauer et al.^{66–68} for the solid and an approach of Kresse et al.²⁹ for the liquid by a treatment of relativistic effects, and moreover introduces the so-called λ -scaling. This scaling is based on the ratio of the experimental cohesive energy E_{coh} and the respective calculated DFT value.²⁶ Only through this scaling, a meaningful discussion of periodic trends and comparison between elements becomes possible, since the trends are otherwise hidden behind systematic and element-

specific errors of the density-functional approximation. For a detailed discussion and formal proof of this scaling, the reader is referred to refs 27 and 28.

All Gibbs energy calculations for the solid and liquid are conducted at the predetermined equilibrium volumes (see the Supporting Information (SI) for details), such that the pV term (eq 1) becomes very small (<1 meV), and the volume work (upon melting) can be neglected. For the gas phase, or in other words, the volume work upon evaporation, the pV term is substantial (tens of meV/atom) and explicitly taken into account. The calculation of the Gibbs energy of the solid G_{sol} starts at the ideal crystal at the respective equilibrium volume $V_{\text{eq,sol}}$. Vibrational (phonon) contributions are calculated in the harmonic approximation and anharmonic effects are accounted for via thermodynamic integration (TDI).^{67,68} The influence of spin–orbit coupling (SO) and converged numerical accuracy is obtained using thermodynamic perturbation theory (TPT). The liquid Gibbs energy G_{liq} calculation begins with a noninteracting reference at the liquid equilibrium volume $V_{\text{eq,liq}}$. Integration over the interaction strength λ from the reference to the (fully interacting) DFT liquid²⁹ gives a first scalar relativistic (SR) Gibbs energy. Including SO coupling and refining numerical settings via TPT²⁶ leads to the final G_{liq} . The effects of scaling the forces, which couples the noninteracting reference to the fully interacting DFT liquid, is illustrated using the pair correlation function (PCF) of liquid Ag for several λ values in the SI.

The Gibbs energy of the gas phase G_{gas} is calculated analytically from the ideal gas law at normal pressure (101.3 kPa). The isolated atoms of the Group 11 elements have one unpaired electron and thus an electronic degeneracy of 2. Cu, Ag, and Au moreover have low-lying electronically excited states, which are significantly populated near their BPs. To account for this, we use fractional degeneracies Θ of 2.17 and 2.49 for Cu and Au, respectively, which reproduce the literature-known Gibbs energy of the gas phase at the BP.^{28,69–71} It is known that the vapor of the Group 11 elements at their BPs contain dimers (about 3% for Cu⁶⁹ and Ag⁷⁰ and 4% for Au⁷¹). However, Mewes and Smits²⁸ have shown that this does not alter the BP noticeably, as the gain in internal energy U is approximately canceled by the loss in entropy S . To check if the strong interaction between single atoms has a notable impact, we calculated the virial two-body correction based on a Lennard–Jones potential as described in ref 28. At the experimental BPs, the contributions amount to 2, 3, and 6 meV/atom for Cu, Ag, and Au, respectively. Although this is substantial compared to previously investigated elements, the effects on the BP remain negligible (<10 K).

Having calculated the solid and liquid Gibbs energies, the entropy is obtained as the difference between G and U , the latter of which is obtained from a normal NVT simulation at the equilibrium volume. Subsequently, linear extrapolation of solid and liquid Gibbs energies to their intersection provides the MP, while extrapolation of the liquid Gibbs energy to the intersection with the analytically calculated Gibbs energy of the gas phase provides the BP.

Since the MPs of the Group 11 elements are rather similar, we decided, for the sake of comparability, to use the same simulation temperatures (and settings, see the SI) for all initial Gibbs energy calculations. Simulations for relativistic solids are conducted at 1200 K and relativistic liquids at 1450 K (for the MP) and 2600 K (for the BP). NR solids were simulated at 900 K, and NR liquids at 1150 K (MP) and 2400 K (BP).

Since these first simulations afforded MPs that deviate significantly from the employed simulation temperatures, certain Gibbs energy calculations are repeated at the predicted MP to confirm the results. Those additional calculations moreover enable a nonlinear (quadratic) fit of the Gibbs energy, which can further validate the calculated MPs (see the SI for further information). The provided theoretical best estimates given in the results section take into account all of these data points, as well as their variation. Apart from the error due to the (linear) extrapolation, there exists a statistical error in the calculated Gibbs energies to due finite simulations lengths. This error amounts to 2–2.5 meV/atom in the most relevant solid and liquid free energies, which conservatively translated into an error of ± 30 K in the MP and ± 5 K in the BP.

3. RESULTS AND DISCUSSION

Setting the Stage. A fundamental quantity concerning phase transitions that is strongly correlated with the MPs and BPs is the cohesive energy (E_{coh}).²⁷ This is illustrated in Figure 3, which shows a correlation plot of the MPs and BPs of all elements against their respective cohesive energies. An overview over experimental and calculated values for E_{coh} and equilibrium volumes at 0 K (V_{eq}^0) for all Group 11 elements obtained with different density-functionals (PBE^{73,74} with and without the D3 atom-pairwise dispersion correction with the default Becke–Johnson damping,^{75,76} PBEsol,^{77,78} and SCAN⁷⁹) is provided in Figure 2. Note that we consistently give cohesive energies as positive values. Inspection of Figure 2 reveals that the cohesive energy and the volume are most consistently reproduced (for all Group 11 elements) with the PBEsol and the SCAN functionals compared to the experimental reference.^{63,72} Due to the lower computational effort and for the sake of consistency with a previous investigation into Group 12 phase transitions,²⁷ we chose PBEsol for all further simulations. Nevertheless, we also compute MPs and BPs with SCAN via TPT from the PBEsol simulations, which can be found in the SI.

For the NR case, the most stable solid structure has to be determined. To this end, we consider the face-centered cubic (*fcc*), body-centered cubic (*bcc*), and hexagonal close-packed (*hcp*) lattices at 0 K with PBEsol; see Table 2. For both cases, SOR and NR, the *fcc* lattice is the most stable, which is in agreement with Takeuchi et al. who investigated Ag and Au using the local density approximation (LDA).⁸⁰ We therefore used the *fcc* lattice for all subsequent NR and SOR calculations for the solid. For Au, we note that the difference between the energetically lowest (*fcc*) and the next higher phase (*hcp*) is smaller at the NR level compared to SOR.

To further study the impact of relativistic effects on the cohesive energy, the experimental and calculated values for the energetically lowest structures at the spin–orbit relativistic (SOR), scalar relativistic (SR), and nonrelativistic (NR) levels of Groups 11 and 12 are collected in Table 3. Inspection shows that relativistic effects are strongest for Au and decrease to Ag and Cu as one would naively expect from the relativistic effects in the dissociation energies of the corresponding dimer (cf. Table 1). In fact, the relativistic effect in the bond energy of Au_2 per atom is $\Delta_R D_e^{\text{PBEsol}} = 0.93/2 = 0.46$ eV. This $\Delta_R D_e^{\text{PBEsol}}$ value of 0.46 eV is in reasonable agreement with the LDA value of Takeuchi et al. of 0.59 eV,⁸⁰ and agrees even better with 0.43 eV (0.861/2) obtained at the coupled-cluster level⁸¹ with a spin–orbit contribution of 0.056 eV added to the

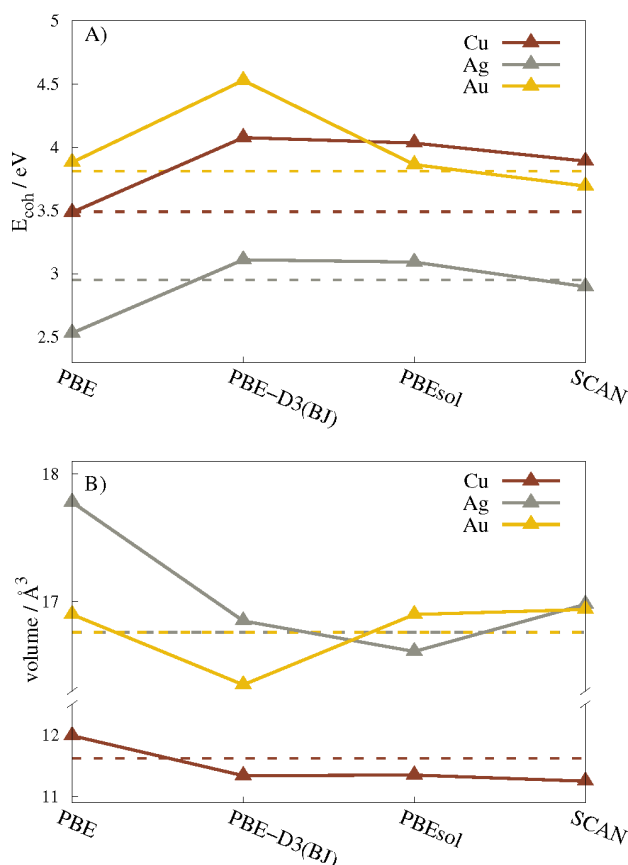


Figure 2. Calculated (A) cohesive energies E_{coh} (in eV) and (B) volumes V (in $\text{\AA}^3/\text{atom}$) at 0 K for all Group 11 elements using different density functionals in the spin-orbit relativistic (SOR) framework. Experimental E_{coh} from ref 63, experimental volumes are corrected for zero-point vibrations (shown as dashed lines) and are from ref 72.

Table 2. Cohesive Energies (in eV) for the *fcc*, *bcc*, and *hcp* Lattices of the Group 11 Elements at the SOR and NR Level of Theory Conducted with the PBEsol Functional

	SOR			NR		
	Cu	Ag	Au	Cu	Ag	Au
<i>fcc</i>	4.033	3.092	3.863	3.833	2.765	2.787
<i>bcc</i>	3.998	3.055	3.843	3.797	2.735	2.749
<i>hcp</i>	4.026	3.083	3.849	3.823	2.760	2.780

Table 3. Cohesive Energies (in eV) for all Group 11 and Group 12^a Elements (SOR = Spin-Orbit Relativistic, SR = Scalar-Relativistic, NR = Nonrelativistic, $\Delta_{\text{NR}}^{\text{SOR}}(\%) = E_{\text{calc}}^{\text{SOR}}/E_{\text{calc}}^{\text{NR}}$, Calculated with the PBEsol Functional^b)

element	$E_{\text{coh}}^{\text{Exp}}$	$E_{\text{coh}}^{\text{SOR}}$	λ^{SOR}	$E_{\text{coh}}^{\text{SR}} (\Delta_{\text{SR}}^{\text{SOR}})$	$E_{\text{coh}}^{\text{NR}} (\Delta_{\text{NR}}^{\text{SOR}})$	$\Delta_{\text{NR}}^{\text{SOR}}(\%)$	$\lambda E_{\text{coh}}^{\text{NR}} (\lambda \Delta_{\text{NR}}^{\text{SOR}})$
Group 11							
Cu	3.490	4.033	0.865	4.031 (0.002)	3.833 (0.200)	−5.2	3.314 (0.176)
Ag	2.950	3.092	0.954	3.077 (0.015)	2.765 (0.327)	−11.8	2.638 (0.312)
Au	3.810	3.863	0.986	3.719 (0.144)	2.787 (1.076)	−38.6	2.748 (1.062)
Group 12							
Zn	1.350	1.572	0.859	1.570 (0.002)	1.661 (−0.089)	5.4	1.426 (−0.076)
Cd	1.169	1.178	0.985	1.169 (0.009)	1.445 (−0.267)	18.5	1.423 (−0.263)
Hg	0.670	0.618	1.084	0.546 (0.072)	1.336 (−0.718)	53.7	1.448 (−0.778)

^aFrom ref 27. ^bAll values refer to the *fcc* lattice and are calculated with an energy cutoff of 600 eV with at least 27^3 k -points. $\Delta_{\text{X}}^{\text{SOR}}$ provides the difference between SOR and level X of theory.

dissociation energy.⁸² Moving from the dimer to the solid state, relativistic effects increase by a factor of $1.076/0.46 = 2.34$ (on a per-atom basis).

Further, compared to Group 12, the absolute relativistic shift of the cohesive energy (SOR vs NR level, $\Delta_{\text{NR}}^{\text{SOR}} E_{\text{coh}}$) is more than twice as large for Cu as for its direct neighbor Zn (0.20 eV vs 0.09 eV), while it is somewhat larger for Ag and Au compared to Cd and Hg (Ag vs Cd: 0.33 eV vs 0.27 eV and Au vs Hg: 1.08 eV vs 0.72 eV, see Table 3 for details). To rationalize these differences and connect them to phase transition temperatures, we put them in the context of the aforementioned linear correlation between E_{coh} and the MP/BP, i.e., $\text{MP/BP} = \gamma \cdot E_{\text{coh}}$ ²⁷ illustrated in Figure 3. The characteristic slope γ can be determined by fitting the respective data either (1) for each element ($\gamma[\text{e}]$), (2) for each Group ($\gamma[\text{g}]$), or, most generally, (3) for all elements ($\gamma[\text{all}]$), excluding those for which this relation does not apply, i.e., molecular gases. Using the resulting γ 's, we obtain a first estimate for the relativistic shift of the phase transition temperatures of the Group 11 elements. Note that, similar to ref 27, we use λ -scaled energy differences for these estimates, i.e., $\lambda \Delta_{\text{NR}}^{\text{SOR}}$ to be consistent with the λ -scaled Gibbs-energy-based calculations discussed thereafter.

The relativistic shifts resulting from this simple linear ansatz are summarized in Table 4. Inspection shows that the impact of relativity increases with increasing nuclear charge, regardless which γ is used. Additionally, it becomes evident that relativistic effects play a larger role for the BP than for the MP due to the generally larger slopes (γ)²⁷ (see Figure 3). Regarding Group 11, it bears pointing out that the element-specific γ 's for the MP of Cu, Ag, and Au differ significantly, whereas in Group 12 the slopes for Zn and Cd are virtually identical (they lie on a line in Figure 3), and only Hg steps out of the line.²⁷ This is a first hint toward a nonlinear relation between the MP and E_{coh} in Group 11, and it therefore becomes questionable if linear predictions of NR shifts of the MP (and BP) based on this relation are reliable. To obtain a more reliable picture, we will in the following discuss MPs and BPs obtained with the Gibbs energy approach as detailed above.^{26–28}

MP and BP via Gibbs Energy. Calculated phase transition temperatures are plotted in Figure 4 and summarized in Table 4 (see the SI for a breakdown of all contributions to the Gibbs energies and further details). Let us begin with the BPs calculated at the SOR level, which agree very well with the experimental references (absolute deviations <1%). This is consistent with previous studies by some of us for Group 12

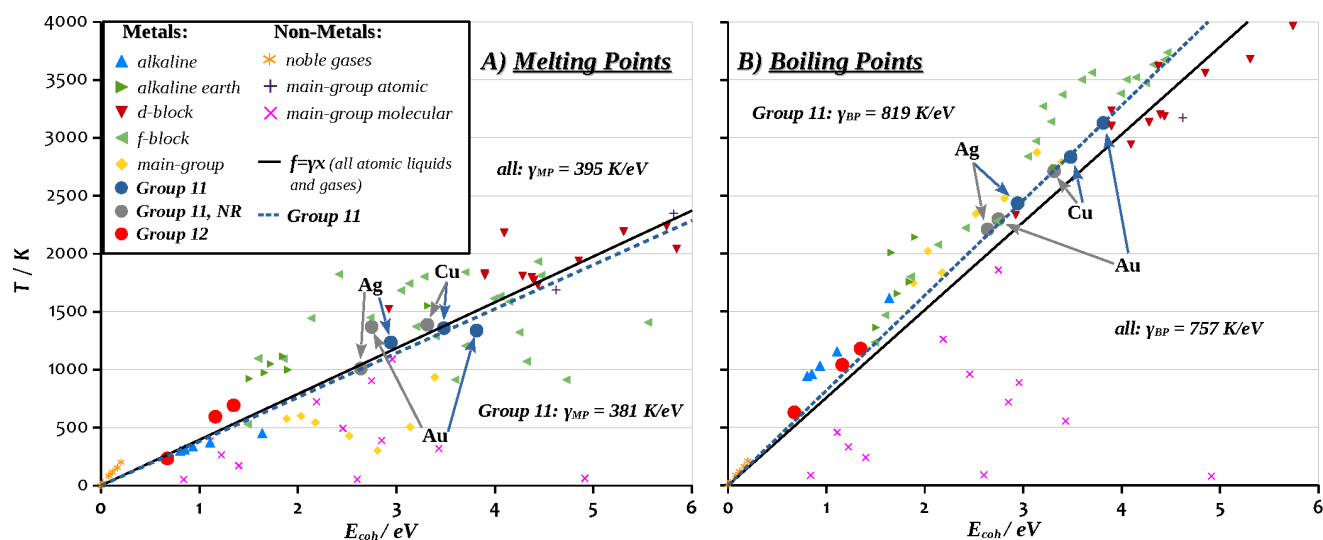


Figure 3. Melting and boiling points (in K) for the elements of the periodic table versus E_{coh} (in eV) up to 4000 K and 6 eV (data taken from refs 63 and 83) as well as a linear regression with forced ordinate intersection. Although we did not include all elements in the Figure, all are taken for the linear regression, except for molecular gases, where the lattice energy is the relevant quantity. For the NR Group 11 elements, the calculated shifts for E_{coh} and MPs/BPs (difference between NR and SOR level) are combined with the experimental data.

Table 4. Experimental (Exp) MPs and BPs, Linear Estimates for Relativistic Shifts Based on E_{coh} ($\lambda\Delta_{\text{NR}}^{\text{SOR}}\gamma$) Using (a) Element-Specific Slopes $\gamma[e]$,^a (b) Group-Specific Slopes $\gamma[g]$,^b and (c) Global Slopes $\gamma[\text{all}]$,^c as Well as the Respective Values Calculated with the Gibbs-Energy-Based Approach Using the PBEsol Functional^d

element	Exp	linear estimates via E_{coh}			Gibbs-energy-based approach					Exp ^{NR}
		$\Delta_{\text{NR}}^{\text{SOR}}(\gamma[\text{e}])$	$\Delta_{\text{NR}}^{\text{SOR}}(\gamma[\text{g}])$	$\Delta_{\text{NR}}^{\text{SOR}}(\gamma[\text{all}])$	SOR	SR	NR	$\Delta_{\text{SR}}^{\text{SOR}}$	$\Delta_{\text{NR}}^{\text{SOR}}$	
<i>melting</i>										
Cu	1358	−68	−67	−70	1220	1220	1260	0	40	1388
Ag	1235	−131	−119	−123	1240	1230	1030	−10	−210	1025
Au	1337	−373	−404	−419	1210	1190	1240	−20	30	1367
<i>boiling</i>										
Cu	2835	−143	−144	−133	2859	2857	2736	−2	−123	2721
Ag	2435	−257	−256	−236	2447	2438	2222	−9	−225	2210
Au	3129	−872	−870	−804	3110	2996	2279	−114	−831	2298

^a $\gamma[e]/\text{MP}$: Cu: 389 K/eV, Ag: 419 K/eV, Au: 351 K/eV; $\gamma[e]/\text{BP}$: Cu: 812 K/eV, Ag: 825 K/eV, Au: 821 K/eV.²⁷ ^b $\gamma[g]/\text{MP}$: 381 K/eV; $\gamma[g]/\text{BP}$: 819 K/eV.²⁷ ^c $\gamma[\text{all}]/\text{MP}$: 395 K/eV; $\gamma[\text{all}]/\text{BP}$: 757 K/eV.²⁷ ^dMPs calculated with the Gibbs energy approach are rounded to 10 K due to the error bars of ± 30 K (see the SI for details). In the last column, we combined experimental MPs and BPs with the calculated total relativistic shift $\Delta_{\text{NR}}^{\text{SOR}}$ to provide a best estimate for the NR case. All values are given in K.

and a representative set of elements (including Cu studied with DFT/PBE).^{27,28} It moreover gives us confidence that the calculated liquid Gibbs energies are reasonably accurate and shows that λ -scaling improves the agreement, most notably in case of Cu where λ significantly differs from unity.

In contrast, deviations for the MPs are somewhat larger: Only the SOR MP of Ag fits very well to the experimental reference (<2%), whereas the MPs of Cu and Au deviate by around 10%. In a recent study of Group 12, the largest deviation was 8.4% for Zn, while the MPs of Cd and Hg were recovered to within 3% and 1%, respectively.²⁷ One possible reason for these larger deviations could be the approximate linear extrapolation to the intersection of the Gibbs energies of the solid and liquid phases. In fact, some of the initial simulations gave (unscaled) MPs that deviate significantly from the simulation temperatures (K). To rule this out as a source of error, we verified the calculated phase transition temperatures by repeating simulations near the predicted MPs (note that λ -unscaled MPs have to be used for this purpose). These additional calculations provided another independent

and, in part, significantly different predicted MP. Moreover, the new data points enable a quadratic instead of a linear fit for the $G(T)$ curves and, in turn, another predicted MP. The final MPs provided in Table 4 and Figure 4 are theoretical best estimates (TBE) based either on the simulations closest to the intersection point, and/or the quadratic fits. Let us illustrate this procedure for the example of NR Cu (similar considerations for the other elements are provided in the SI), which was initially simulated at 900 K (solid) and 1150 K (liquid), providing an (unscaled) MP of 1278 K. Since this is too far above the solid simulation temperature for an accurate linear extrapolation, we conducted new Gibbs energy calculations at 1200 K (solid) and 1450 K (liquid). This resulted in a new linearly extrapolated (unscaled) MP of 1385 K, and, via a quadratic fit using all points, an (unscaled) MP of 1441 K. Since this is again more than 200 K above the highest calculation of the solid, we confirmed the new value with a final solid simulation at 1450 K, which provided an unscaled MP of 1476 K. λ -scaling these last two values provides 1246 and 1276 K, which we combine to a theoretical best estimate for the MP

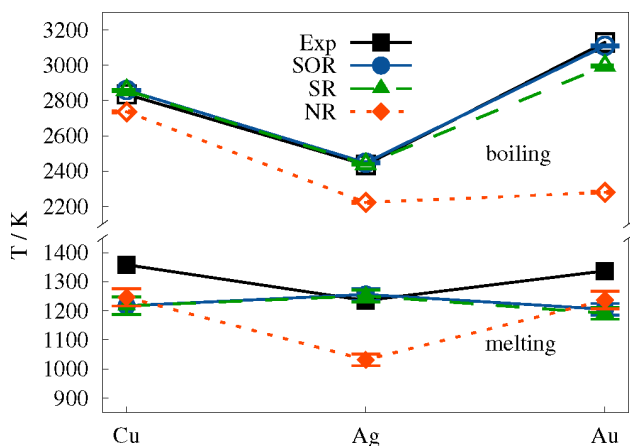


Figure 4. Experimental and calculated phase transition temperatures using the TI Gibbs energy PBEsol approach. Experimental reference in black,⁶⁴ and calculated SOR, SR, and NR results (and their error bars) as solid, dashed, and dotted lines, respectively. Note that the errors in the BP are so small that they are barely visible.

of NR Cu of 1260 ± 30 K, where the error-bars include the statistical uncertainty in the calculated Gs as well as the variation between simulations.

After all, the deviations between calculated and experimental MPs of Cu and Au remain significant despite all our efforts. We therefore discuss several further sources of error apart from the linear extrapolation: First, we note that density functional approximations have difficulties in correctly describing the polarization of the lower-lying *d*-shell for the late transition metals (i.e., of Group 11).^{84,85} For the Au₁₀ cluster, for example, the 6s orbital lies deep in the 5*d* space, and the different performance of the various density functionals compared to coupled cluster has been demonstrated.⁸⁶ Such errors are presumably much less relevant for the closed-shell elements of Group 12, which could explain why the agreement for Group 11 is significantly worse. Moreover, the agreement between DFT/PBEsol and the experimental cohesive energy of Cu is not very good. This is similar for Zn, which showed a deviation and a λ similar to Cu. For both, Cu and Zn, λ -scaling slightly worsens the agreement (unscaled MP of Cu is 1431 K corresponding to 5.4% deviation), whereas it strongly improves the BPs (unscaled BP Cu: 3224 K/13.7% vs scaled BP Cu: 2859 K/0.85%) and in case of Group 12 also the MPs of Cd and Hg.^{27,28} Second, and perhaps most importantly, the MP is generally more susceptible to errors in the Gibbs energies since the solid and liquid curves run almost parallel near their intersection, whereas the liquid–gas intersection is steeper. In consequence, an error of 1 meV/atom in the Gibbs energy translates to an error of about 10 K in the MP, but only to an error of about 1 K in the BP. With statistical errors of up to 3 meV/atom in our free energies, this translates into conservative error-bars of ± 30 K in the MP (and ± 5 K in the BP). In other words, the errors in the MP of Cu and Au correspond to moderate deviations of about 12–13 meV/atom in the Gibbs energies. Finally, we explore if the error originates from the description of the solid, the liquid, or both phases. Comparing the calculated entropies to experimental references^{69–71} (a plot is available in the SI, Figure S1) shows that calculated liquid entropies fit rather well to the literature values (consistent with the accurately predicted BPs), whereas solid entropies, especially for Cu, deviate stronger. This indicates that the

error in the MP originates in the description of the solid rather than the liquid.

Despite these apparent deviations between experiment and theory for the MP, our calculated MP for Cu is in good agreement with a previously reported theoretical investigation by Zhu et al., who employed a similar Gibbs-energy-based approach.⁶⁸ With the PBE functional, they obtained an MP of 1251 ± 15 K, in excellent agreement with our value of 1260 ± 30 K, whereas the LDA provided a distinctly larger MP of 1494 ± 5 K, comparable to our unscaled PBEsol result of 1456 K. Although Zhu et al. did not apply λ -scaling, at least their PBE results are directly comparable since λ for Cu with PBE is very close to unity (see Figure 2A). Discussing their results, Zhu et al. argue that LDA and PBE give an upper and a lower boundary for the MP, respectively. However, having in mind the linear γE_{coh} relationship and considering the strong overbinding of LDA for *fcc* Cu ($E_{\text{coh}}^{\text{LDA}}/E_{\text{coh}}^{\text{exp}} = \lambda = 0.77$, which is comparable to PBEsol with $\lambda = 0.87$), we think the higher MP of LDA is merely a result of the strong overbinding. Note that a very similar argument has been made by Zhu et al. in a later work on Al and Ni.⁸⁷ Correcting for the systematic deviation through λ -scaling, the MP of Cu with LDA becomes 1150 K, which is even further away from the experimental value. Also, the more recent SCAN functional does not improve the agreement, neither does increasing the number of atoms in the simulations, nor using a harder PAW potential, which we explore in the SI. Hence, we conclude that DFT-based Gibbs energy calculations are missing some important contributions to the Gibbs energy and, in turn, the MP of Cu (and presumably also Au). This appears to be rather independent of the employed density functional, and probably most severe for the solid phase. We speculate that a higher-level treatment for the electronic structure is required to restore the accuracy, e.g., via the random-phase approximation as in case of Si,²⁹ which would be a formidable task for the Group 11 metals due to the larger size of the valence space. However, even if the absolute values of the predicted MPs are not as accurate as, e.g., for Group 12,²⁷ the changes due to relativistic effects will profit from some form of error-compensation and are thus presumably more accurate.

Impact of Relativistic Effects. Having established the agreement between experiment and relativistic calculations, let us now explore the impact of relativity by considering calculations conducted in the SR and NR limits. To this end, we first need to establish how relativistic shifts and absolute relativistic MPs and BPs provided in Table 4 are derived. This is particularly relevant for the MP, since here the deviation between the SOR calculations and the experiment is substantial, such that the definition of relativistic shifts is ambiguous. To eliminate this ambiguity and make the best use of any error-compensation in the theoretical values, we evaluate relativistic shifts between the calculated SOR, SR, and NR values. Furthermore, to provide consistent absolute NR MPs, we combine the calculated relativistic shifts with the experimental MPs (last column of Table 4). For the BPs, the deviation between experimental and calculated values is much smaller, such that differences between the possible schemes are insignificant ($<1\% \approx 20$ K). However, for the sake of consistency, we use the same approach as for the MPs.

In general, we note that the lion's share of relativistic changes is due to SR effects, whereas SOR effects only have a rather small influence on both the MPs and BPs. This is consistent with findings reported for Group 12.²⁷ Comparing

Table 5. Comparison of Calculated MPs and BPs for Group 11 and Group 12 (in K) with the Gibbs-Energy-Based Approach at the SOR and NR Levels^a

	melting				boiling			
	SOR	NR	$\Delta_{\text{NR}}^{\text{SOR}}$	$\Delta_{\text{NR}}^{\text{SOR}\%}$	SOR	NR	$\Delta_{\text{NR}}^{\text{SOR}}$	$\Delta_{\text{NR}}^{\text{SOR}\%}$
<i>Group 11</i>								
Cu	1220	1260	40	3.2	2859	2736	−123	−4.3
Ag	1240	1030	−210	−17	2447	2222	−225	−9.2
Au	1210	1240	30	2.5	3110	2279	−831	−27
<i>Group 12</i>								
Zn	635	658	23	3.6	1197	1259	62	5.2
Cd	616	678	62	10	1060	1268	208	20
Hg	231	648	417	181	630	1236	606	96

^a $\Delta_{\text{NR}}^{\text{SOR}}$ is the difference between NR and SOR phase transition temperatures, given in absolute and in relative (%) form (with respect to the SOR result).

the total relativistic shift (SOR to NR) for each of the elements, we note that the BP of Au exhibits the largest impact, which decreases for Ag and even further for the BP of Cu. This behavior is expected, as it is known that relativistic effects scale with Z^2 , and moreover obeys the linear relation between the BP and the cohesive energy (cf. Table 4 and Figure 3).

For the MP, however, a strikingly different picture emerges. First of all, the MP of Cu and of Au in the NR limit is higher than in the SOR limit, which is in contrast to Ag and the BPs. However, we note that these shifts are just within the combined uncertainties of our calculations and are thus not strongly significant. Second, Ag has the highest total relativistic shift $\Delta_{\text{NR}}^{\text{DEF}}$ of −210 K, whereas the shift for Au is essentially the same as for Cu. The absence of a large shift for Au is even more surprising since the relativistic shift of the cohesive energy is the largest for Au with more than 1.0 eV, compared to 0.3 eV and <0.2 eV for Ag and Cu, respectively (cf. Table 3 and Figure 3). To elucidate the origin of this unexpected behavior of Au, we compared several thermodynamic quantities obtained from our simulations. Most notably, we found a surprising behavior of the internal energy U_{pot} (kinetic part cancels out) of the solid and liquid phases, which we confirmed with further simulations in the SOR (SR) and NR limit at the same simulation temperature of 1200 K. From basic considerations detailed below, which are confirmed by our experience with Group 12 and further elements, it can be expected that any (relativistic) change of E_{coh} should exert a smaller influence on U_{pot} of disordered phases (the liquid) than on ordered phases (the solid) and, accordingly, the impact should decrease with increasing temperature. The underlying reason is that in ordered phases at low temperatures, more atoms are located at their most favorable position near the minimum, such that an increase of the depth of the potential (reflected in E_{coh}) will have a larger overall effect on U_{pot} . This is confirmed by our simulations for Cu and Ag: Here, the stabilization of the solid over the liquid phase when turning on relativistic effects is visible in the ratio $U_{\text{SOR}}/U_{\text{NR}}$ at 1200 K, which is smaller for liquid Ag and Cu (1.12 and 1.05) than for solid Ag and Cu (1.14 and 1.06). However, in the case of Au, the internal energy of the liquid benefits more from relativistic effects as evident from a ratio $U_{\text{SOR}}/U_{\text{NR}}$ of 1.43 compared to 1.41 for the solid. We speculate that this can be related to the energy differences between the various solid structures of the respective relativistic and nonrelativistic elements (see Table 2). Inspection shows that in particular the *fcc* and *hcp* phases are distinctly closer in NR Au compared to SOR Au, whereas the differences are about the same in NR and

SOR Ag. When picturing the liquid as a mixture of several different solid phases, changes in the relative energies of the solid phases of Au could explain the lower stability of the liquid phase of NR Au and, in turn, the surprisingly high NR MP. Further analysis of differences in the average coordination number of Au atoms in the SOR and NR solid and liquid states and/or the presence of dimers in the liquid did not produce any significant results. This is consistent with the observation that the relativistic shift of the cohesive energy of Au (Table 3) is, in absolute terms, similar to that of the dissociation energy of the Au dimer (Table 1).

In any case, our simulations show that the liquid phase of SOR Au is particularly stable compared to the NR situation, explaining why the large change in the cohesive energy does not lead to a similarly large change in the MP, as is suggested by the linear γ -relation to the cohesive energy (cf. Table 4). This is also evident from the significantly different element-specific γ of $1260/2.748 = 459 \text{ K/eV}$ of NR Au compared to the SOR case with 351 K/eV . Interestingly, this renders NR Au much more similar to SOR Ag (419 K/eV), which is also visible when regarding the E_{coh} of NR Au, as well as the absolute values of its MP and BP as shown in Figure 3). As such, these results corroborate an old hypothesis put forward by Pyykkö almost 50 years ago, who speculated that the difference between Ag and Au is mainly of relativistic nature.⁸⁸

In general, comparing the linearly predicted relativistic shifts obtained with $\gamma[\text{e}]$, $\gamma[\text{g}]$, and $\gamma[\text{all}]$ to those obtained with the Gibbs energy approach shows that neither of them agrees well for the MPs. These differences show that the linear relation is much less applicable to the MPs of Group 11 elements than to those of Group 12, where it provided much more accurate estimates that were consistent with the Gibbs-energy-based results.²⁷ We speculate that this can be traced back to the more complicated electronic structure of the Group 11 elements, leading to phase-specific effects and, in turn, nonlinear changes in the phase transition temperatures. For the BP, however, the linear estimates with the global γ agree reasonably well with the Gibbs-energy-derived BPs.

Finally, let us compare trends and relativistic effects observed for Group 11 to those reported for Group 12. For the latter, the periodic decrease in the MPs and BPs from Zn to Cd and Hg, and eventually to Cn, is exclusively the result of relativistic effects.²⁷ In turn, in the NR limit, all Group 12 elements become very similar to Zn. This is fundamentally different in Group 11—here relativity increases the MP of Ag and all BPs compared to the NR limit, whereas it decreases the MP of Cu and Au. In addition and contrary to Group 12, the

NR Group 11 elements behave differently from their lightest pendant Cu. In general, despite larger absolute changes in the cohesive energies (compare Table 3) and in the BPs, the relative impact of relativistic effects is distinctly larger for most Group 12 elements as evident from Table 5. It can be seen that only for the MPs of the fifth Period (Ag and Cd) relativistic effects are stronger in Group 11 (−17% vs 10%), whereas in all other cases Group 12 elements exhibit larger relativistic shifts (compare Table 5). Hence, at least concerning the impact on phase transition temperatures, and presumably also other bulk properties, we find a peak of relativistic effects in Group 12 and not in Group 11.⁴⁵

4. SUMMARY AND CONCLUSIONS

We reported an extensive investigation of the melting and boiling points (MP and BP) of the Group 11 elements with spin–orbit relativistic (SOR), scalar-relativistic (SR), and nonrelativistic (NR) density-functional theory, using a recently composed Gibbs-energy-based TI-MD- λ DFT approach.^{26–28} While the SOR calculations afford BPs that agree very well with experimental references (absolute deviations %), the calculated MPs show larger deviations with respect to experimental data (2–10%). This generally larger deviation for the MP compared to the BP is consistent with previous applications of the method,²⁷ and it was traced back to a generally more challenging prediction of MPs, and to the difficult electronic structure of the Group 11 elements, specifically the solid. Despite these deviations, results for the MP of Cu (and its deviation from the experiment) were found to be in very good agreement with previous studies.⁶⁸

Exploring the impact of relativistic effects on the phase transitions of Group 11, it was established that the relativistic shift of the BPs generally increases with increasing nuclear charge, leading to the largest SOR–NR shifts for Au (−831 K), followed by Ag (−225 K) and eventually Cu (−123 K). Accordingly, these shifts are in close agreement with linear estimates based on the change in the cohesive energies. However, this is remarkably different for the MP, where the relativistic shifts are surprisingly large for Ag and much smaller than predicted for Au: Here, the MP remains essentially the same in the SOR and NR limits, defying the typically strong linear relation with the cohesive energy, which is obeyed by most elements, including the neighboring Group 12. This surprisingly small relativistic shift of the MP was traced back to phase-specific effects, leading to a particularly strong stabilization of the liquid phase of SOR Au, and, in turn, a much smaller relativistic shift than expected from the linear relation between MP and cohesive energy. Altogether, the higher slope γ in the MP vs E_{coh} plot (Figure 3) as well as its MP and BP render NR Au very similar to SOR Ag, which nicely falls into place with an old hypothesis of Pyykkö, who stated that the difference between Ag and Au is essentially due to relativistic effects.⁸⁸

Comparing the impact of relativity for Group 11 and 12, it was shown that while absolute relativistic shifts in the cohesive energy and BPs are larger for Group 11, the relative changes are more substantial for the Group 12 elements. In general, the proposed linear relation between phase transition temperature and the cohesive energy, which was quite accurate for the Group 12 elements,²⁷ was found to be much less useful for any predictions concerning Group 11. As such, the results hold general implications about the origin and progression of periodic trends in the phase transition temperatures, revealing

dramatic differences to the neighboring Group 12 and illustrating that the results of the simple linear extrapolation schemes for complex properties should always be taken with several grains of salt.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/jacs.1c10881>.

Details for DFT calculations and MD simulations, protocol for Gibbs energy calculations and resulting numerical values for all relevant phases (G, U, S, and a breakdown of the contributions), how they are extrapolated to their intersections, approach to determine equilibrium volumes for the solid and liquid phases, and calculations of errors (PDF)

■ AUTHOR INFORMATION

Corresponding Author

Jan-Michael Mewes – Mulliken Center for Theoretical Chemistry, University of Bonn, 53115 Bonn, Germany; orcid.org/0000-0002-4669-8091; Email: janmewes@janmewes.de

Authors

Sarah Löffelsender – Mulliken Center for Theoretical Chemistry, University of Bonn, 53115 Bonn, Germany; orcid.org/0000-0003-3146-7810

Peter Schwerdtfeger – Centre for Theoretical Chemistry and Physics, The New Zealand Institute for Advanced Study, Massey University, 0632 Auckland, New Zealand; orcid.org/0000-0003-4845-686X

Stefan Grimme – Mulliken Center for Theoretical Chemistry, University of Bonn, 53115 Bonn, Germany; orcid.org/0000-0002-5844-4371

Complete contact information is available at: <https://pubs.acs.org/doi/10.1021/jacs.1c10881>

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We acknowledge financial support from the Marsden Fund (17-MAU-021) of the Royal Society of New Zealand (Wellington). We are grateful to Dr. M. Bursch (MPI Kohlenforschung) for help with designing the Abstrac graphic, and to Dr. P. Jerabek (Helmholtz-Zentrum Hereon) for interesting discussions.

■ REFERENCES

- (1) Haile, J. M.; Johnston, I.; Mallinckrodt, A. J.; McKay, S. Molecular dynamics simulation: elementary methods. *Comput. Phys.* **1993**, *7*, 625–625.
- (2) Belonoshko, A. B. Molecular dynamics of silica at high pressures: Equation of state, structure, and phase transitions. *Geochim. Cosmochim. Acta* **1994**, *58*, 1557–1566.
- (3) Pedersen, U. R. Direct calculation of the solid-liquid Gibbs free energy difference in a single equilibrium simulation. *J. Chem. Phys.* **2013**, *139*, 104102.
- (4) Pedersen, U. R.; Hummel, F.; Kresse, G.; Kahl, G.; Dellago, C. Computing Gibbs free energy differences by interface pinning. *Phys. Rev. B: Condens. Matter Mater. Phys.* **2013**, *88*, 094101.

- (5) Allen, M. P.; Tildesley, D. J. *Computer simulation of liquids*; Oxford University Press: Oxford, 2017; 640 pp.
- (6) Weck, G.; Recoules, V.; Queyroux, J.-A.; Datchi, F.; Bouchet, J.; Ninet, S.; Garbarino, G.; Mezouar, M.; Loubeyre, P. Determination of the melting curve of gold up to 110 GPa. *Phys. Rev. B: Condens. Matter Mater. Phys.* **2020**, *101*, 014106.
- (7) Binder, K. The Monte Carlo method for the study of phase transitions: A review of some recent progress. *J. Comput. Phys.* **1985**, *59*, 1–55.
- (8) Ercolessi, F.; Parrinello, M.; Tosatti, E. Simulation of gold in the glue model. *Philos. Mag. A* **1988**, *58*, 213–226.
- (9) Kaplan, I. G. *Intermolecular interactions: physical picture, computational methods and model potentials*; John Wiley & Sons: New York, 2006; 384 pp.
- (10) Watanabe, H.; Ito, N.; Hu, C.-K. Phase diagram and universality of the Lennard-Jones gas-liquid system. *J. Chem. Phys.* **2012**, *136*, 204102.
- (11) Liu, Z.-L.; Sun, J.-S.; Li, R.; Zhang, X.-L.; Cai, L.-C. Comparative Study on Two Melting Simulation Methods: Melting Curve of Gold. *Commun. Theor. Phys.* **2016**, *65*, 613–616.
- (12) Hermann, A.; Krawczyk, R. P.; Lein, M.; Schwerdtfeger, P.; Hamilton, I. P.; Stewart, J. J. P. Convergence of the many-body expansion of interaction potentials: From van der Waals to covalent and metallic systems. *Phys. Rev. A: At., Mol., Opt. Phys.* **2007**, *76*, 013202.
- (13) Pahl, E.; Calvo, F.; Koči, L.; Schwerdtfeger, P. Accurate melting temperatures for neon and argon from ab initio Monte Carlo simulations. *Angew. Chem., Int. Ed.* **2008**, *47*, 8207–8210.
- (14) Schwerdtfeger, P.; Tonner, R.; Moyano, G. E.; Pahl, E. Towards J/mol Accuracy for the Cohesive Energy of Solid Argon. *Angew. Chem., Int. Ed.* **2016**, *55*, 12200–12205.
- (15) Smits, O. R.; Jerabek, P.; Pahl, E.; Schwerdtfeger, P. A Hundred-Year-Old Experiment Re-evaluated: Accurate Ab Initio Monte Carlo Simulations of the Melting of Radon. *Angew. Chem., Int. Ed.* **2018**, *57*, 9961–9964.
- (16) Smits, O. R.; Jerabek, P.; Pahl, E.; Schwerdtfeger, P. First-principles melting of krypton and xenon based on many-body relativistic coupled-cluster interaction potentials. *Phys. Rev. B: Condens. Matter Mater. Phys.* **2020**, *101*, 104103.
- (17) Fulde, P. *Electron correlations in molecules and solids*; Springer Science & Business Media: Berlin, 1995; Vol. 100, 480 pp.
- (18) Shepherd, J. J.; Booth, G.; Grüneis, A.; Alavi, A. Full configuration interaction perspective on the homogeneous electron gas. *Phys. Rev. B: Condens. Matter Mater. Phys.* **2012**, *85*, 081103.
- (19) Booth, G. H.; Grüneis, A.; Kresse, G.; Alavi, A. Towards an exact description of electronic wavefunctions in real solids. *Nature* **2013**, *493*, 365–370.
- (20) Chen, M.; Hung, L.; Huang, C.; Xia, J.; Carter, E. A. The melting point of lithium: an orbital-free first-principles molecular dynamics study. *Mol. Phys.* **2013**, *111*, 3448–3456.
- (21) Seko, A.; Maekawa, T.; Tsuda, K.; Tanaka, I. Machine learning with systematic density-functional theory calculations: Application to melting temperatures of single- and binary-component solids. *Phys. Rev. B: Condens. Matter Mater. Phys.* **2014**, *89*, 054303.
- (22) Steenbergen, K. G.; Pahl, E.; Schwerdtfeger, P. Accurate, Large-Scale Density Functional Melting of Hg: Relativistic Effects Decrease Melting Temperature by 160 K. *J. Phys. Chem. Lett.* **2017**, *8*, 1407–1412.
- (23) Cohen, A. J.; Mori-Sánchez, P.; Yang, W. Insights into Current Limitations of Density Functional Theory. *Science* **2008**, *321*, 792.
- (24) Jones, R. O. Density functional theory: Its origins, rise to prominence, and future. *Rev. Mod. Phys.* **2015**, *87*, 897–923.
- (25) Medvedev, M. G.; Bushmarinov, I. S.; Sun, J.; Perdew, J. P.; Lyssenko, K. A. Density functional theory is straying from the path toward the exact functional. *Science* **2017**, *355*, 49–52.
- (26) Mewes, J.-M.; Smits, O. R.; Kresse, G.; Schwerdtfeger, P. Copernicium: A Relativistic Noble Liquid. *Angew. Chem., Int. Ed.* **2019**, *58*, 17964–17968.
- (27) Mewes, J.-M.; Schwerdtfeger, P. Exclusively Relativistic: Periodic Trends in the Melting and Boiling Points of Group 12. *Angew. Chem., Int. Ed.* **2021**, *60*, 7703–7709.
- (28) Mewes, J.-M.; Smits, O. R. Accurate elemental boiling points from first principles. *Phys. Chem. Chem. Phys.* **2020**, *22*, 24041–24050.
- (29) Dorner, F.; Sukurma, Z.; Dellago, C.; Kresse, G. Melting Si: Beyond Density Functional Theory. *Phys. Rev. Lett.* **2018**, *121*, 195701.
- (30) Norman Greenwood, E. G. *Alan Earnshaw Chemistry of the Elements*, 2nd ed.; Butterworth-Heinemann, 1997; 1173 pp.
- (31) Aldenderfer, M.; Craig, N. M.; Speakman, R. J.; Popelka-Filcoff, R. Four-thousand-year-old gold artifacts from the Lake Titicaca basin, southern Peru. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105*, 5002–5005.
- (32) Song, J.; Wang, L.; Zibart, A.; Koch, C. Corrosion protection of electrically conductive surfaces. *Metals* **2012**, *2*, 450–477.
- (33) Samsonov, G. *Handbook of the Physicochemical Properties of the Elements*; Springer, 1968; pp 387–446.
- (34) Corti, C. W.; Holliday, R. J.; Thompson, D. T. Developing new industrial applications for gold: Gold nanotechnology. *Gold Bulletin* **2002**, *35*, 111–117.
- (35) Laguna, A. *Modern supramolecular gold chemistry: gold-metal interactions and applications*; Wiley-VCH: Weinheim, 2008; 525 pp.
- (36) Mohr, F. *Gold Chemistry - Applications and Future Directions in the Life Sciences*; Wiley-VCH: Weinheim, 2009; 424 pp.
- (37) Moore, J. A.; Chow, J. C. L. Recent progress and applications of gold nanotechnology in medical biophysics using artificial intelligence and mathematical modeling. *Nano Express* **2021**, *2*, 022001.
- (38) Bond, G. C.; Louis, C.; Thompson, D. *Catalysis by gold*; World Scientific: Singapore, 2006; Vol. 6, 384 pp.
- (39) Li, G.; Wang, A.; Chen, Y.; Sun, Y.; Du, Y.; Wang, X.; Ding, P.; Jia, R.; Wang, Y.; Zhang, G. Development of a Colloidal Gold-Based Immunochromatographic Strip for Rapid Detection of Severe Acute Respiratory Syndrome Coronavirus 2 Spike Protein. *Frontiers in Immunology* **2021**, *12*, 560.
- (40) Puddephatt, R. J. *The Chemistry of Gold*; Elsevier: Amsterdam, 1978; 274 pp.
- (41) Baerends, E. J.; Schwarz, W. H. E.; Schwerdtfeger, P.; Snijders, J. G. Relativistic atomic orbital contractions and expansions: magnitudes and explanations. *J. Phys. B: At., Mol. Opt. Phys.* **1990**, *23*, 3225–3240.
- (42) Schwarz, W. H. E.; van Wezenbeek, E. M.; Baerends, E. J.; Snijders, J. G. The origin of relativistic effects of atomic orbitals. *Journal of Physics B: Atomic, Molecular and Optical Physics* **1989**, *22*, 1515–1529.
- (43) Autschbach, J.; Siekierski, S.; Seth, M.; Schwerdtfeger, P.; Schwarz, W. H. E. Dependence of relativistic effects on electronic configuration in the neutral atoms of d- and f-block elements. *J. Comput. Chem.* **2002**, *23*, 804–813.
- (44) Schwerdtfeger, P.; Lein, M. *Gold Chemistry*; Wiley-VCH: Weinheim, 2009; pp 183–247.
- (45) Pyykkö, P.; Desclaux, J. P. Relativity and the periodic system of elements. *Acc. Chem. Res.* **1979**, *12*, 276–281.
- (46) Schwerdtfeger, P. Relativistic and electron-correlation contributions in atomic and molecular properties: benchmark calculations on Au and Au₂. *Chem. Phys. Lett.* **1991**, *183*, 457–463.
- (47) Jansen, M. The chemistry of gold as an anion. *Chem. Soc. Rev.* **2008**, *37*, 1826–1835.
- (48) Schwerdtfeger, P. Relativistic effects in gold chemistry. 2. The stability of complex halides of gold (III). *J. Am. Chem. Soc.* **1989**, *111*, 7261–7262.
- (49) Schwerdtfeger, P.; Boyd, P. D. W.; Burrell, A. K.; Robinson, W. T.; Taylor, M. J. Relativistic effects in gold chemistry. 3. Gold(I) complexes. *Inorg. Chem.* **1990**, *29*, 3593–3607.
- (50) Schwerdtfeger, P.; Boyd, P. D. W.; Brienne, S.; Burrell, A. K. Relativistic effects in gold chemistry. 4. Gold(III) and gold(V) compounds. *Inorg. Chem.* **1992**, *31*, 3411–3422.
- (51) Pyykkö, P. Strong Closed-Shell Interactions in Inorganic Chemistry. *Chem. Rev.* **1997**, *97*, 597–636.

- (52) Pyykkö, P. Relativistic effects in structural chemistry. *Chem. Rev.* **1988**, *88*, 563–594.
- (53) Schwerdtfeger, P. Relativistic effects in properties of gold. *Heteroat. Chem.* **2002**, *13*, 578–584.
- (54) Pyykkö, P. Theoretical Chemistry of Gold. *Angew. Chem., Int. Ed.* **2004**, *43*, 4412–4456.
- (55) Christensen, N.; Kollar, J. Electronic structure of CsAu. *Solid State Commun.* **1983**, *46*, 727–730.
- (56) Fossgaard, O.; Gropen, O.; Eliav, E.; Saue, T. Bonding in the homologous series CsAu, CsAg, and CsCu studied at the 4-component density functional theory and coupled cluster levels. *J. Chem. Phys.* **2003**, *119*, 9355–9363.
- (57) Glantschnig, K.; Ambrosch-Draxl, C. Relativistic effects on the linear optical properties of Au, Pt, Pb and W. *New J. Phys.* **2010**, *12*, 103048.
- (58) Söhnle, T.; Hermann, H.; Schwerdtfeger, P. Towards the Understanding of Solid-State Structures: From Cubic to Chainlike Arrangements in Group 11 Halides. *Angew. Chem., Int. Ed.* **2001**, *40*, 4381–4385.
- (59) Söhnle, T.; Hermann, H.; Schwerdtfeger, P. Solid State Density Functional Calculations for the Group 11 Monohalides. *J. Phys. Chem. B* **2005**, *109*, 526–531.
- (60) Huber, K. P.; Herzberg, G. *Molecular Spectra and Molecular Structure IV. Constants of Diatomic Molecules*; Van Nostrand Reinhold Company: New York, 1979; 716 pp.
- (61) Bishea, G. A.; Morse, M. D. Spectroscopic studies of jet-cooled AgAu and Au₂. *J. Chem. Phys.* **1991**, *95*, 5646–5659.
- (62) Bornhauser, P.; Beck, M.; Zhang, Q.; Knopp, G.; Marquardt, R.; Gourlaouen, C.; Radi, P. P. Accurate ground state potential of Cu₂ up to the dissociation limit by perturbation assisted double-resonant four-wave mixing. *J. Chem. Phys.* **2020**, *153*, 244305.
- (63) Kittel, C.; McEuen, P.; McEuen, P. *Introduction to solid state physics*; Wiley: New York, 1996; Vol. 8, 673 pp.
- (64) CRC Handbook, *CRC Handbook of Chemistry and Physics*, 88th ed.; CRC Press: Boca Raton, FL, 2007; 2640 pp.
- (65) Schwerdtfeger, P.; Dolg, M.; Schwarz, W. H. E.; Bowmaker, G. A.; Boyd, P. D. W. Relativistic effects in gold chemistry. I. Diatomic gold compounds. *J. Chem. Phys.* **1989**, *91*, 1762–1774.
- (66) Grabowski, B.; Ismer, L.; Hickel, T.; Neugebauer, J. Ab initio up to the melting point: Anharmonicity and vacancies in aluminum. *Phys. Rev. B: Condens. Matter Mater. Phys.* **2009**, *79*, 134106.
- (67) Duff, A. I.; Davey, T.; Korbmayer, D.; Glensk, A.; Grabowski, B.; Neugebauer, J.; Finnis, M. W. Improved method of calculating ab initio high-temperature thermodynamic properties with application to ZrC. *Phys. Rev. B: Condens. Matter Mater. Phys.* **2015**, *91*, 214311.
- (68) Zhu, L.-F.; Grabowski, B.; Neugebauer, J. Efficient approach to compute melting properties fully from ab initio with application to Cu. *Phys. Rev. B: Condens. Matter Mater. Phys.* **2017**, *96*, 224202.
- (69) Arblaster, J. Thermodynamic properties of copper. *J. Phase Equilib. Diffus.* **2015**, *36*, 422–444.
- (70) Arblaster, J. Thermodynamic properties of silver. *J. Phase Equilib. Diffus.* **2015**, *36*, 573–591.
- (71) Arblaster, J. Thermodynamic properties of gold. *J. Phase Equilib. Diffus.* **2016**, *37*, 229–245.
- (72) Schimka, L.; Gaudoin, R.; Klimeš, J. c. v.; Marsman, M.; Kresse, G. Lattice constants and cohesive energies of alkali, alkaline-earth, and transition metals: Random phase approximation and density functional theory results. *Phys. Rev. B: Condens. Matter Mater. Phys.* **2013**, *87*, 214102.
- (73) Perdew, J. P.; Burke, K.; Ernzerhof, M. Generalized Gradient Approximation Made Simple. *Phys. Rev. Lett.* **1996**, *77*, 3865–3868.
- (74) Perdew, J. P.; Burke, K.; Ernzerhof, M. Generalized Gradient Approximation Made Simple [Phys. Rev. Lett. 77, 3865 (1996)]. *Phys. Rev. Lett.* **1997**, *78*, 1396–1396.
- (75) 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.
- (76) Grimme, S.; Ehrlich, S.; Goerigk, L. Effect of the damping function in dispersion corrected density functional theory. *J. Comput. Chem.* **2011**, *32*, 1456–1465.
- (77) Perdew, J. P.; Ruzsinszky, A.; Csonka, G. I.; Vydrov, O. A.; Scuseria, G. E.; Constantin, L. A.; Zhou, X.; Burke, K. Restoring the Density-Gradient Expansion for Exchange in Solids and Surfaces. *Phys. Rev. Lett.* **2008**, *100*, 136406.
- (78) Perdew, J. P.; Ruzsinszky, A.; Csonka, G. I.; Vydrov, O. A.; Scuseria, G. E.; Constantin, L. A.; Zhou, X.; Burke, K. Erratum: Restoring the Density-Gradient Expansion for Exchange in Solids and Surfaces [Phys. Rev. Lett. 100, 136406 (2008)]. *Phys. Rev. Lett.* **2009**, *102*, 039902.
- (79) Sun, J.; Ruzsinszky, A.; Perdew, J. P. Strongly Constrained and Appropriately Normed Semilocal Density Functional. *Phys. Rev. Lett.* **2015**, *115*, 036402.
- (80) Takeuchi, N.; Chan, C. T.; Ho, K. First-principles calculations of equilibrium ground-state properties of Au and Ag. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1989**, *40*, 1565.
- (81) Wesendrup, R.; Laerdahl, J. K.; Schwerdtfeger, P. Relativistic effects in gold chemistry. VI. Coupled cluster calculations for the isoelectronic series AuPt-, Au₂, and AuH+. *J. Chem. Phys.* **1999**, *110*, 9457–9462.
- (82) Tu, Z.; Chen, A.; Xia, C.; Li, Z.; Yang, M.; Wang, C.; Wang, W. Theoretical investigation of the spectroscopic constants for the ground-state diatomic species Cu₂, Ag₂, and Au₂. *Comput. Theor. Chem.* **2017**, *1112*, 88–93.
- (83) Lide, D. R. *CRC handbook of chemistry and physics*; CRC Press, 2004; Vol. 85, pp 4–98.
- (84) Thierfelder, C.; Schwerdtfeger, P.; Saue, T. ⁶³Cu and ¹⁹⁷Au nuclear quadrupole moments from four-component relativistic density-functional calculations using correct long-range exchange. *Phys. Rev. A* **2007**, *76*, 034502.
- (85) Cramer, C. J.; Truhlar, D. G. Density functional theory for transition metals and transition metal chemistry. *Phys. Chem. Chem. Phys.* **2009**, *11*, 10757–10816.
- (86) Götz, D. A.; Schäfer, R.; Schwerdtfeger, P. The performance of density functional and wavefunction-based methods for 2D and 3D structures of Au₁₀. *J. Comput. Chem.* **2013**, *34*, 1975–1981.
- (87) Zhu, L.-F.; Körmann, F.; Ruban, A. V.; Neugebauer, J.; Grabowski, B. Performance of the standard exchange-correlation functionals in predicting melting properties fully from first principles: Application to Al and magnetic Ni. *Phys. Rev. B* **2020**, *101*, 144108.
- (88) Desclaux, J.; Pyykkö, P. Dirac-Fock one-centre calculations. The molecules CuH, AgH and AuH including p-type symmetry functions. *Chem. Phys. Lett.* **1976**, *39*, 300–303.

Acknowledgements

This thesis would not have been possible without the help of many people who supported and inspired me throughout my Ph.D. studies. First, I would like to thank my supervisor Prof. Dr. Stefan Grimme for the opportunity to work in his group. I have learned a lot for my life and I am grateful for the financial support for my interesting scientific work in Bonn and at conferences.

Additional, I want to thank Prof. Dr. Marc de Wergifosse, my second examiner, for his constant and professional support over long parts of my Ph.D. studies. Thank you for inspiring and encouraging me to follow my path and continue exploring the exciting, yet challenging world of proteins. It has not been an easy path for me, but your knowledge and guidance certainly made a difference.

I am also grateful to Prof. Dr. Nikolay Kornienko and Prof. Dr. Gregor Hagelücken for their time to serve as chairperson and fourth member of my doctoral committee. Many thanks to Benedikt Bädorf, and Julius Stückrath, for valuable remarks on parts of this thesis, as well as for fruitful scientific and non-scientific discussions.

Furthermore, I want to thank my former and current colleagues for their scientific and non-scientific support, explicitly Benedikt Bädorf, Fabian Bohle, Markus Bursch, Eike Caldeweyher, Sebastian Dohm, Marvin Friede, Thomas Froitzheim, Thomas Gasevic, Johannes Gorges, Andreas Hansen, Christian Hölzer, Abylay aka. Albert Katbashev, Julia Kohn, Lukas Kunze, Jan Mewes, Marcel Müller, Katja Ostermeir, Christoph Plett, Dr. Philipp Pracht, Thomas Rose, Karola Schmitz, Jakob Seibert, Sebastian Spicher, Marcel Stahn, and Lukas Wittmann. For organizational and IT-related questions, I want to thank Claudia Kronz and Jens Mekelburger for their straightforward help.

I would also like to thank my former office buddies Karola Schmitz, and Julius Stückrath for many enjoyable and interesting conversations. A thank you goes to my collaborators Pierre Beaujean, Benoît Champagne, Gregor Hagelücken, Jochen Niemeyer, Marilù Maraldi, Martin Peter, and Jan Riebe for interesting scientific projects, even if not all of them could be brought to a successful end. Many thanks to all members of the 2nd St. Céré Quantum Chemistry Seminar in France. It was a pleasure to discuss with you, and I have learned so much personally and scientifically during my stay. My gratitude goes out to everyone I have referred to or might have unintentionally left out. You all helped me at some point to create this thesis.

Additionally, many people outside the academic world constantly supported me, and without them, this thesis would never have been possible. Therefore, special thanks go to my family, in particular my mother, my foster mother, and my parents-in-law for their loving and encouraging help, for their time and their support.

Finally, I would like to thank my husband Sven and my daughter Nia from the bottom of my heart. Your love, your patience, your positive attitude kept me going, your joy and your energy gave me the strength I needed during all this time. You are my greatest motivation. Thank you.