Theoretical Investigations of Multivalent Reactions

Inaugural Dissertation
to obtain the academic degree
Doctor rerum naturalium (Dr. rer. nat.)

submitted to
the Department of Biology, Chemistry and Pharmacy
of Freie Universität Berlin

by
Andreas Johannes Achazi
from Berlin, Germany

2017

This thesis was developed under the supervison of Prof. Dr. Beate Paulus between March 2013 and October 2017 at the Department of Biology, Chemistry and Pharmacy of Freie Universität Berlin.

1st Referee: Prof. Dr. Beate Paulus
2st Referee: Prof. Dr. Christoph A. Schalley

Date of defence: 27 November 2017

"The underlying physical laws necessary for the mathematical theory of a large part of physics and the whole of chemistry are thus completely known, and the difficulty is only that the exact application of these laws leads to equations much too complicated to be soluble."

– P. A. M. Dirac, 1929

Acknowledgment

First and foremost I would like to thank Prof. Dr. Beate Paulus for the chance to work in her group on this exciting research topic. Furthermore I would like to express my deepest gratitude for her support and supervision throughout my thesis. Together with the freedom to try new approaches this provided an excellent working environment and allowed me to increase my scientific expertise.

I would also like to express my gratitude to Prof. Dr. Christoph Schalley for his supervision of the joint research project, his contributions to it and co-examining this thesis.

Without my collaboration partners this work would not have been possible. I would like to thank all of them, in particular Prof. Dr. Hans-Ulrich Reissig, Dr. Dirk Andrae, Dr. Doreen Mollenhauer, Marthe Solleder, Dr. Marcus Weber, Hendrik V. Schröder, Henrik Hupatz, Susanne Liese, Prof. Dr. Roland R. Netz, Prof. Dr. Jürgen P. Rabe, Mirko Lohse and Prof. Dr. Stefan Hecht. However, my thanks to Dr. Larissa K. S. von Krbek must be emphasized. Our discussions and her tremendous experimental contributions were crucial for this thesis.

I would also like to thank all the students I had the chance to supervise during their master thesis, bachelor thesis or research internship: Johannes Horst Budau, Stefan Mattsson, Felix Witte, Fabian Weber, Federica Maschietto and Anton Samuelis.

Moreover, I thank all current members and alumni of the AG Paulus , AG Tremblay and AG Keller for the productive working atmosphere and helpful discussions, especially Dr. Carsten Müller, Vincent Pohl, Matthias Berg, Christian Stemmle, Stefan Mattsson, Zeinab Kaawar and Yasmeen Qawasmeh.

For financial support I thank the Collaborative Research Center (SFB) 765 “Multivalency as chemical organization and action principle: new architectures, functions and applications” of the German Research Foundation (DFG) and the Freie Universität Berlin. I also thank all members and alumni of the SFB 765 and the Dahlem Research School for the great lectures and discussions, especially my fellow student representatives Susanne Liese and Valentin Reiter.

I am grateful for the high-performance computing resources and the support from the Zentraleinrichtung für Datenverarbeitung (ZEDAT).

I also thank the office assistants Julija Djordjevic, Lydia Alnajjar and Katharina Tebel, which were also crucial for an undisturbed work flow.

My fellow students Jan Driller, Moritz Maleschewski, Stefan Ellrodt and Sebastian

Gohr, thank you all for our time together during Bachelor, Master and Ph.D. degree courses.

Finally, I would like to express my gratitude to my family, especially my parents and siblings for their support.

Abstract

A strong but easily reversible connection is important on the molecular level in biological recognition, adhesion, and signaling processes but also for drug design or supramolecular complexes. This connection can be achieved by using multivalency: the connection is formed with a large number of weak interactions instead of only one strong interaction. Multivalency and the underlying cooperative effects are not yet fully understood.

Small multivalent crown ether/ammonium assemblies are used in this thesis as model system for multivalent interactions and as building blocks for functionalized supramolecular host-guest complexes. For analysis, a theoretical multilevel approach to determine the Gibbs energy of association in solution $\Delta G_{\mathrm{A}}^{\mathrm{sol}}$ of these assemblies *in silico* has been developed. Dispersion corrected density functional theory is used to obtain the association energies in the gas phase. The effects of translation, rotation and vibration are taken into account by a rigid-rotor-harmonic-oscillator approximation with a free-rotor approximation for low-lying vibrations. The solvation is included by the implicit solvent models COSMO-RS and D-COSMO-RS. Addionally counterions have to be included explicitly for a good agreement with the experiments. The average deviation from the experimental results was about 5 $\mathrm{kJ\,mol^{-1}}$, which is state-of-the-art.

In collaboration with specialists for supramolecular chemistry and for molecular dynamic simulations new insights into the bond formation of crown ether/ammonium assemblies and multivalency were gained. A delicate counterbalancing between the effects of the solvent methanol on the crown ether/ammonium assembly formation and the ammonium-guest/counterion complex dissociation was revealed. This renders a chloroform/methanol mixture with a medium methanol content best for achieving high association constants with crown ether/ammonium assemblies. The counterions cause negative allosteric cooperativity in the investigated guests, whereas the negative allosteric cooperativity in the investigated hosts is caused by the polarisable π-system of the spacer. Attractive spacer-spacer interactions yield a strong positive chelate cooperativity. The variation of the length and flexibility of the ammonium guest spacer gives interesting results. It was expected that guests with a rigid spacer, that approximates the spacing between the binding sites in the host, would have the strongest positive chelate cooperativities and highest association constants, because guests with flexible spacer would suffer from an entropic penalty. Surprisingly the ammonium guests with flexible spacer exhibited the strongest positive chelate cooperativities and one of these ammonium guests exhibited the highest association constant. The results were summarized to a general guideline to

create multivalent molecular assemblies with high binding affinity. Most important is a high complementarity of host and guest. A mix of preoganisation and adaptability is preferred over maximizing the preoganisation by lowering the adaptability.

Additionally this thesis includes the investigation of two functionlized supramolecular complexes: a photoswitch and a donor-acceptor complex.

The theoretical multilevel approach to determine $\Delta G_{\mathrm{A}}^{\mathrm{sol}}$ was then applied in a mechanistic investigation of the samarium diiodide mediated reductive coupling of the *N*-oxoalkyl-substituted methyl indole-3-carboxylates. In contrast to previous proposals, it was found that the high diastereoselectivity in the cyclization step is caused by the formation of an energetically highly favorable chelate complex.

Finally the stretching of polyethylene glycol in water was investigated in collaboration with experts for atomic force microscope experiments and molecular dynamic simulations. Surprisingly the stretching of polyethylene glycol in water is enthalpically unfavorable and not entropically unfavorable. This indicates that a multivalent host or guest in water with a flexible polyethylene glycol spacer suffers from an enthalpic penalty and not from an entropic penalty in complex formation. Together with the results of the crown ether/ammonium assemblies, this challenges the notion that flexible spacers cause in general an entropic penalty in complex formations. However, the situation is more complicated. Spacer length, intramolecular interactions, solvent and counterions can affect the conformational space of both the flexible spacer and the whole system or have strong enthalpic effects. This will be difficult to grasp experimentally. The *in silico* methods presented in this thesis are powerful tools to identify the important effects in chemical reactions and to comprehend the experimental findings.

Zusammenfassung

Feste, jedoch leicht reversible Verbindungen auf molekularer Ebene sind essentiell für biologische Erkennungs-, Adhäsions- und Signalprozesse und werden mittlerweile auch in der medizinischen Wirkstoffforschung und für supramolekulare Komplexe genutzt. Diese Art von Verbindung kann mittels Multivalenz erreicht werden, das heißt die Verbindung wird mit einer Vielzahl schwacher Wechselwirkungen geformt anstatt einer einzigen sehr starken. Multivalenz und die zugrundeliegenden kooperativen Effekte sind allerdings bis heute noch nicht vollständig verstanden. In der vorliegenden Dissertation werden kleine multivalente Kronenether/Ammonium Aggregate als Modellsystem für multivalente Wechselwirkungen und als Bausteine für funktionalisierte supramolekulare Wirt-Gast-Komplexe verwendet. Zur Analyse wurde eine theoretische Mehrstufen-Methode zur *in silico* Berechnung der Gibbs-Energie der Assoziation $\Delta G_{\mathrm{A}}^{\mathrm{sol}}$ solcher Aggregate im Lösungsmittel entwickelt. In dieser wird die Assoziationsenergie der Partner in der Gasphase mit Hilfe von dispersionskorrigierter Dichtefunktionaltheorie bestimmt. Die Effekte der Translation, Rotation und Vibration werden mittels einer Starrer-Rotator-Harmonischer-Oszillator-Näherung mit Freier-Rotator-Näherung für niedrig liegende Vibrationen berücksichtigt. Die Solvatisierung wird mittels der Lösungsmittelmodelle COSMO-RS und D-COSMO-RS einbezogen. Zusätzlich müssen Gegenionen explizit mit berücksichtigt werden, um eine gute Übereinstimmung mit dem Experiment zu erhalten. Die durchschnittliche Abweichung zu den experimentellen Ergebnissen lag bei circa 5 $\mathrm{kJ\,mol^{-1}}$. Dies entspricht dem derzeitigen Stand der Technik.

In Kooperation mit Spezialisten für supramolekuare Chemie und für Moleküldynamik-Simulationen konnten neue Erkenntnisse über die Bildung von Kronenether/Ammonium Aggregaten und über die Multivalenz erhalten werden. Zwischen den Effekten vom Lösungsmittel Methanol auf die Bildung des Kronenether/Ammonium Aggregates und auf die Dissoziation des Ammonium/Gegenion Komplexes zeigte sich ein fragiles Gleichgewicht. Daher ist ein Chloroform/Methanol Gemisch mit einer mittleren Methanolkonzentration am besten geeignet, um hohe Assoziationskonstanten bei Kronenether/Ammonium Aggregaten zu erzielen. In Bezug auf die negative allosterische Kooperativität zeigt sich, dass diese in den untersuchten Gästen von Gegenionen verursacht wird, in den untersuchten Wirten aber durch Polarisierung des π-Systems des Spacers (Verknüpfungsstruktur). Attraktive Spacer-Spacer-Wechselwirkungen erzeugen dagegen eine stark positive Chelatkooperativität. In den Untersuchungen wurden Länge und Flexibilität von den Spacern der Gastmoleküle variiert. Dabei wurde davon ausgegangen, dass Gäste

mit starren Spacern, die ungefähr den Abstand zwischen den Bindungsstellen des Wirtes überspannen, sowohl die stärkste positive Chelatkooperativität aufweisen, als auch die höchste Assoziationskonstante, weil die Gäste mit flexiblen Spacern entropisch ungünstig sind. Im Gegensatz zu dieser Annahme zeigten jedoch die Gäste mit flexiblen Spacer die stärksten positiven Chelatkooperativitäten, wobei einer sogar die höchste Assoziationskonstante aufwies. Die Ergebnisse lassen sich zu einer allgemeinen Leitlinie für die Herstellung multivalenter molekularer Aggregate mit hoher Assoziationskonstante zusammenfassen. Am wichtigsten ist ein hoher Grad an Komplementarität von Wirt und Gast. Zudem ist eine Mischung aus Vororganisation und Anpassungsfähigkeit besser als die Vororganisation durch eine Reduktion der Anpassungsfähigkeit zu maximieren.

In dieser Dissertation wurden darüber hinaus zwei funktionalisierte Wirt-Gast-Komplexe untersucht, ein Photoschalter und ein Elektronen-Donor-Akzeptor-Komplex.

Die theoretische Mehrstufen-Methode zur Berechnung $\Delta G_{\mathrm{A}}^{\mathrm{sol}}$, wurde ebenfalls in einer mechanistischen Untersuchung der Samariumdiiodid vermittelten reduktiven Kupplung von N-Oxoalkyl-substituiertem Methyl-indol-3-carboxylat angewendet. Im Gegensatz zu bisherigen Vorschlägen für den Reaktionsmechanismus, hat sich gezeigt, dass die hohe Diastereoselektivität im Zyklisierungsschritt durch die Bildung eines energetisch sehr vorteilhaften Chelatkomplexes hervorrufen wird.

Des weiteren wurde mit Experten für Rasterkraftmikroskopie und Moleküldynamik-Simulationen die Dehnung von Polyethylenglycol in Wasser untersucht. Anders als erwartet, war die Dehnung von Polyethylenglycol in Wasser enthalpisch und nicht entropisch ungünstig. Das bedeutet, dass ein multivalenter Wirt oder Gast mit einem flexiblen Polyethylenglycol-Spacer bei der Komplexbildung enthalpisch und nicht entropisch benachteiligt ist. Zusammen mit dem Ergebnissen für die Kronenether/Ammonium Aggregate widerspricht dies der Auffassung, dass flexible Spacer im allgemeinen entropisch ungünstig sind. In der Realität stellt sich die Situation deutlich komplizierter dar. So beeinflussen Spacerlänge, intramolekulare Wechselwirkungen, Lösungsmittel und Gegenionen sowohl den Konformationsraum des flexiblen Spacers als auch den des Gesamtsystems und können darüber hinaus auch starke enthalpische Effekte haben. In Experimenten ist es schwer diese Effekte auseinanderzuhalten. Die *in silico* Methoden, die in dieser Dissertation vorgestellt werden, stellen wirkungsvolle Werkzeuge dar, um die wichtigen Effekte in chemischen Reaktionen zu identifizieren und experimentelle Ergebnisse nachzuvollziehen.

List of Abbreviations

AFM	Atomic force microscope
BCH	Baker-Campbell-Hausdorff
BOA	Born–Oppenheimer approximation
BSSE	Basis set superposition error
CC	Coupled cluster
CCSD	Coupled Cluster Singles and Doubles
CCSDT	Coupled Cluster Singles, Doubles and Triples
CCSD(T)	Coupled Cluster Singles and Doubles with perturbative Triples
CGTO	Contracted Gaussian Type Orbital
CI	Configuration Interaction
CISD	Configuration Interaction with Singles and Doubles
COSMO	Conductor-like Screening Model
COSMO-RS	Conductor-like Screening Model for Real Solvents
CP	Counterpoise
DCM	Dielectric continuum solvation model
DCOSMO-RS	Direct Conductor-like Screening Model for Real Solvents
DF	Density fitting
DFT	Density Functional Theory
DMC	Double mutant cycle
ECP	Effective core potential
FCI	Full Configuration Interaction
GGA	Generalized Gradient Approximation
GTO	Gaussian Type Orbital
HF	Hartree-Fock
HO	Harmonic oscillator
HOMO	Highest occupied molecular orbital
ITC	Isothermal titration calorimetry
L	Local
LDA	Local Density Approximations
LMO	Localized molecular orbital
LUMO	Lowest unoccupied molecular orbital
MARI-J	Multipole accelerated resolution of identity for J
MD	Molecular dynamics
NAO	Natural atomic orbital
NMB	Natural minimal basis
NRB	Natural Rydberg basis
NPA	Natural Population Analysis
PAO	Projected Atomic Orbital
PEG	Polyethylene glycol
PES	Potential energy surface

PGTO	Primitive Gaussian Type Orbital
PNAO	Pre-natural atomic orbitals
RI	Resolution of the identity
RHF	Restricted Hartree-Fock
ROHF	Restricted open-shell Hartree–Fock
RR	Rigid-rotor
RRHO	Rigid-rotor-harmonic-oscillator
SASA	Solvent-accessible surface area
SCF	Self-consistent field
SI	International System of Units
THF	Tetrahydrofuran
TISE	Time-independent Schrödinger equation
UHF	Unrestricted Hartree-Fock

Contents

1 Introduction

A key principle in nature to create a connection between two components, which is strong but easily reversible at the same time, is to use a large number of weak interactions. A macroscopic example is the burr (*Arctium spec.*) and the hook-and-loop fastener derived from it (Figure 1.1). The hooks on the one surface catch on the loops on the opposite surface. This creats a large number of weak interactions which form a strong connection between the two surfaces. In contrast to two surfaces which are connected by only one strong interaction, these two surfaces can be easily separated from each other by sequentially detaching the individual hooks and loops. This principle is called "multivalency". One the molecular level, it is essential for biologic systems in processes such as recognition, adhesion, and signaling.[1,2] Multivalent interactions are interactions between systems consisting of multiple, interconnected binding sites (Figure 1.2). In the given macroscopic example (Figure 1.1), the binding sites are the hooks and the loops. In molecular systems these hooks and the loops are replaced with non-covalent bonds. Multivalency is already used in medicine (drug design),[1,3] biochemistry,[4] material science,[5] surface science[6] and supramolecular chemistry.[7,8] However, it is not yet fully understood.

To gain a better understanding, the effects of multivalency can be deconvoluted into contributions originating from allosteric cooperativity, chelate cooperativity, interannular

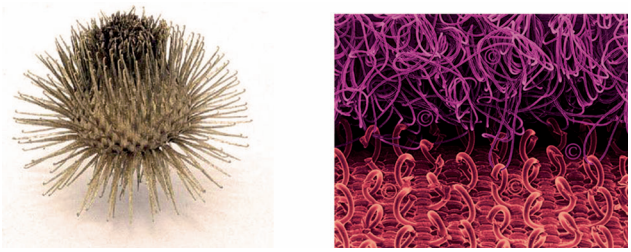

Figure 1.1: On the left side is a burr (*Arctium spec.*), one the right site is a hook-and-loop fastener inspired from it. Included with permission from Fasting *et al.*[1] (©2012 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim).

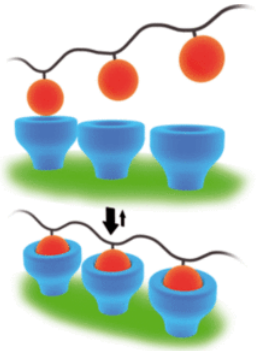

Figure 1.2: Mulivalent molecular binding units. The blue cylinders with cavity are the binding sites of the host. These encompass the binding sites of the guest, which are depicted as red balls. Included with permission from SFB765.[16]

cooperativity and symmetry effects.[9–14] Detailed overviews on cooperativity are written by Hunter and Anderson,[13] Ercolani and Schiaffino[14] and Schalley and co-workers.[15]

Allosteric cooperativity describes the effects that a bond formation between a monovalent guest and a multivalent host (and vice versa) has on the subsequent bond formation, that is on bond formation of the next monovalent guest with the multivalent host. Positive allosteric cooperativity is present if the first binding event amplifies the association strength of the subsequent binding event and negative allosteric cooperativity is present if the association strength of the subsequent binding event is reduced.

Chelate cooperativity is an effect that solely occurs in multivalent systems, for example in a complex of a divalent guest with a divalent host. For the rating of the strength of chelate cooperativity Hunter and Anderson[13] introduced one approach and Ercolani and Schiaffino[14] another. Negative chelate cooperativity, as defined by these two approaches, indicates that not all binding sites of the multivalent complex are closed and the complex is prone to oligo- or polymerization. Positive chelate cooperativity indicates that all binding sites of the multivalent complex are closed and oligo- or polymerization is unlikely to occur.

Interannular cooperativity does not occur in the systems investigated in this thesis. Therefore, it is not discussed here. A detailed explaination of interannular cooperativity is given by Ercolani and Schiaffino.[14]

A monovalent guest binds m-times more often to a m-valent host (with m binding sites) than to a monovalent host.[1] Such statistical effects are also called symmetry effects.[1,9–12] They are system dependent and have to be subtracted for the determination of the cooperativity.

The experimental determination of the cooperativity can be carried out using the double mutant cycle (DMC) analysis.[17–23] This consists of the measurement of the association constant not only for the multi- and monovalent complexes but also their intermediates, i.e. multivalent host with monovalent guest and vice versa. These association constants are used to define cooperativity factors,[13–15] which evaluate whether the allosteric or chelate cooperativity is positive or negative. The equations for the allosteric and chelate cooperativity factors will not be presented here, but can be found in the experimental part of the corresponding published papers.

Of special importance for the research on multivalency is host-guest chemistry, a branch of supramolecular chemistry defined in 1974 by Cram.[24] A host-guest complex consists of at least two molecules or ions. The binding sites of the host encompass usually the binding sites of the guest (Figure 1.2). The bonds between host and guest are only formed by non-covalent interactions, like ion-ion interactions, ion-dipole interactions, hydrogen bonds, $\pi - \pi$ stacking and van der Waals interactions (including London dispersion). Host-guest systems can be designed with a defined number of binding sites and can be used to investigate intermediate systems of monovalent and multivalent hosts and guests. Thus, host-guest systems are ideal for performing DMC analyses to quantify the multivalent binding amplification.[1] Additionally, the length, geometry and flexibility of the spacer which connects the binding sites can be systematically varied. In other words, the complementarity, preorganization and adaptability of the systems can be controlled and modified. These three are important characteristics of hosts and guests. A common approach, suggested by Cram,[25,26] is to use preorganized and complementary hosts and guests to obtain high binding affinities.[1,5,7,27,28] "Preorganization" implies that the conformation of the unbound host and the unbound guest should be as close as possible to their bound conformations in the host-guest complex.[25,26] "Complementarity" means that the binding sites of the host attract the binding sites of the guest (and vice versa). Cram[25,26] was only considering monovalent systems. In a multivalent system "complementarity" additionally means that the distance between the binding sites of the host corresponds to the distance between the binding sites of the guest (and vice versa). Whitesides and co-workers[2] basically extended Cram's approach to multivalent systems by recommending the use of guests with rigid (i.e. preorganized) spacers, that approximates the spacing between the binding sites in the host (i.e. complementarity spacers). The advantage of the rigid host-guest systems is, that they suffer from a lower entropic penalty in the formation of the host-guest complex compared to flexible systems, because

the number of available conformations reduces much less.[2,29,30] The disadvantage of rigid host-guest systems is, that even small spatial mismatches between host and guest can result in drastically reduced binding affinities.[2] Therefore, the complementarity is important. However, nature has created highly effective systems for molecular recognition based on flexible peptides.[1,30] An unbound flexible guest can adapt to a huge variety of different sized hosts (and vice versa). This characteristic of the guest or host can be called "adaptability" and could be an alternative concept to obtain high binding affinities. A guest whose binding sites are connected with a flexible spacer could compensate for small spatial mismatches with the host and vice versa, which might then negate the above mentioned possible entropic penalty for flexbile systems in the formation of the host-guest complex. Amazingly, except from foldamers,[31–33] adaptability has been barely used[34–37] in multivalent systems.[38] One part of this thesis will focus on determining to which extent complementarity, preorganization, and adaptability are needed to gain high binding affinities and to develop a general "guideline" to create multivalent molecular assemblies with high binding affinity.

Crown ether/cation complexes were discovered in 1967 by Pedersen.[39,40] Meanwhile, they are easily synthetically accessible and have been used extensively as host–guest complexes in supramolecular chemistry and to mimic biological systems.[20,41–52] Therefore, the 18-crown-6/primary ammonium ion and the 24-crown-8/secondary ammonium ion binding motif are used in this thesis as binding sites for host-guest model systems to investigate multivalency. The primary ammonium ion is too large for penetrating the center of the 18-crown-6 ether; therefore, it stays on top of the crown ether ring plane and both form a side-on complex. However, the secondary ammonium ion threads through the 24-crown-8 ether. They form a pseudo[2]rotaxane, which can be interlocked by adding large stopper units yielding a [2]rotaxane. Figure 1.3 shows schematic depictions of these three types of crown ether/ammonium ion assemblies.

The experimental supramolecular chemistry was mainly done by the AG Schalley, especially by Larissa K. S. von Krbek and the molecular dynamic simulations were performed by the AG Weber, especially by Marthe Solleder. In this thesis quantum chemical methods are applied to analyze multivalency. Quantum chemistry allows to make *in silico* reliable predictions of experimental outcomes, which offer new insights into the reaction processes and can be used to verify the assumptions drawn from experimental results. The AG Schalley uses isothermal titration calorimetry (ITC) to analyze the thermodynamic and measure the Gibbs energy of association of the multivalent crown ether/ammonium

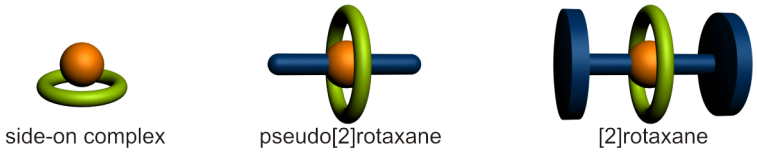

Figure 1.3: Schematic depictions from the left to the right of a side-on complex as formed by the 18-crown-6/primary ammonium ion binding motif, a pseudo[2]rotaxane as formed by the 24-crown-8/secondary ammonium ion binding motif and a [2]rotaxane. A [2]rotaxane can be gained from a pseudo[2]rotaxane by locking the the ring on the rod with stoppers. Included with permission from von Krbek.[38]

assemblies. At the same time, one of the main research objectives of this thesis is to develop a theoretical (multilevel) approach to determine the Gibbs energy of association in solution $\Delta G_{\mathrm{A}}^{\mathrm{sol}}$ of these assemblies *in silico*. The Gibbs energy of association in solution can be separated into three terms (Figure 1.4): the electronic energy in the gas phase, the contributions of translation, rotation and vibration (often called "finite temperature contribution") and the Gibbs free energy of solvation.[53] These terms can be calculated at different theoretical-chemical levels, which allow to find a balance between accuracy and computational cost. Coupled Cluster Singles and Doubles with perturbative Triples (CCSD(T)) is the "Gold Standard" in quantum chemistry to obtain accurate electronic energies of association in the gas phase.[54,55] However, this method is not feasible for the systems under investigation. Even with further approximations, like the local/domain[55] and the density fitting[55] approxiamtion, the Coupled Cluster method can only be used to get benchmarks for the electronic association energies of the smallest systems in this thesis. Therefore, the Kohn-Sham Density Functional Theory (DFT)[56,57] is applied. DFT includes electron correlation in a highly effectiv manner, which makes it a very fast and accurate method. It is less reliable than CCSD(T); however, with Coupled Cluster benchmarks for the association energies, an appropriate density functional can be chosen. Unfortunately, most density functionals fail in describing the long-range electron correlation, which is responsible for the London dispersion. Therefore, the disperion correction DFT-D3(BJ)[58–61] is employed. The effects of translation, rotation and vibration are taken into account by a rigid-rotor-harmonic-oscillator (RRHO) approximation with free-rotor approximation to account for low-lying vibrations. Even with DFT the solvent cannot be included explicitly. The Gibbs free energy of solvation is calculated as a seperate term with the COSMO-RS[62–65] model. COSMO-RS describes solvent effects with a model of interacting molecular surface segments based on quantum chemical

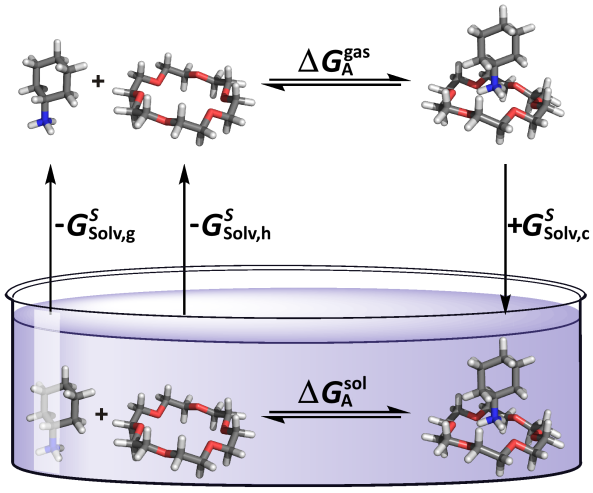

Figure 1.4: Schematic representation of the thermodynamic cycle for the calculation of the Gibbs energy of association $\Delta G_{\mathrm{A}}^{\mathrm{sol}}$ in solution. The Gibbs energy of association $\Delta G_{\mathrm{A}}^{\mathrm{gas}}$ in the gas phase contains the electronic energy of association in the gas phase and the contributions of translation, rotation and vibration. $G_{\mathrm{Solv,g}}^{S}$, $G_{\mathrm{Solv,h}}^{S}$ and $G_{\mathrm{Solv,c}}^{S}$ are the Gibbs energies of the solvation of a guest, host and complex molecule. Included with permission from Achazi *et al.*[72] (©2015 Wiley Periodicals, Inc.)

calculations with the dielectric continuum solvation model COSMO.[66,67] In contrast to pure dielectric continuum solvation models, COSMO-RS is able to take short-range interactions into account like hydrogen bonds and London dispersion interactions. In total this results in a DFT-based multilevel approach and the Gibbs energy of association in solution $\Delta G_{\mathrm{A}}^{\mathrm{sol}}$ is determined in a thermodynamical cycle as depicted in Figure 1.4. This kind of approach is also used by other groups.[68–71] Unlike the investigations of these other groups, in this thesis the systems of interest have charged binding sites and sometimes large flexible parts. Thus, it is may be necessary to consider several conformers for the same molecule or counterions which interfere in the binding process. An average deviation of 7.5 $\mathrm{kJ\,mol^{-1}}$ or less of the theoretical approach to the experiemental results can be seen as state-of-the-art according to a survey[70] of theoretical publications. As mentioned above the *in silico* determination of the Gibbs energy of association in solution $\Delta G_{\mathrm{A}}^{\mathrm{sol}}$ of the multivalent assemblies will offer new insights into the reaction processes and the underlying effects which render the cooperativity. The Gibbs energy of association in solution $\Delta G_{\mathrm{A}}^{\mathrm{sol}}$ also yields the association constant. Thus, the allosteric and chelate cooperativity factors can be determined quantum chemically in case the accuracy is sufficient. The quantum chemical data also give insight, for example, into the molecular

structure, which is important for evaluating the complementarity, or other characteristics like HOMO-LUMO gaps.

This thesis aims on achieving the research objectives:

- Development of a multilevel DFT-based approach to determine the Gibbs energy of association in solution of monovalent 18-crown-6/primary ammonium ion complexes. (**Paper A1**)

- Extension of the multilevel DFT-based approach to divalent 18-crown-6/primary ammonium ion complexes and 24-crown-8/secondary ammonium ion pseudo[2]rotaxanes to elucidate the association mechanisms as well as use of this approach and further quantum chemical approaches to analyze the impact of complementarity, preorganization and adaptability on the multivalent binding amplification. (**Paper A2-A5**)

- Development of a general "guideline" for the creation of multivalent molecular assemblies with high binding affinity based on the acquired knowledge.

- Investigation of functionalized supramolecular assemblies with quantum chemical methods. (**Paper A5-A6**)

As already mentioned, this thesis has been accomplished in strong collaboration with specialists for supramolecular chemistry (AG Schalley) and for molecular dynamic simulations (AG Weber). The results are the outcome of the joint efforts of the members of the just mentioned groups and me.

Another research objective is to use the acquired knowledge for the DFT-based determination of the Gibbs energy of association in solution for reaction processes in which the experimental determination the of the Gibbs energy is not possible. Thus, the reaction mechanisms for the samarium diiodide (SmI_2) mediated reductive coupling of the *N*-oxoalkyl-substituted methyl indole-3-carboxylates is investigated in this thesis, too. (**Paper A7**)

This thesis is part of the Collaborative Research Center (SFB) 765 "Multivalency as chemical organization and action principle: new architectures, functions and applications". As part of part of this Collaborative Research Center, this thesis includes also a collaboration investigating the stretching of polyethylene glycol (PEG) in water. PEG can be used as flexible spacer, which connects this investigation with the investigation of the crown ether/ammonium ion assemblies. (**Paper B1**)

Altogether, this thesis is a quantum chemical investigation of various systems to gain insights into multivalency and its underlying effects.

2 Theoretical Methods

2.1 Time-independent and dependent Schrödinger equation

In this thesis the thermodynamics of the formation of mono- and multivalent molecular assemblies in solution is investigated. The three main contributions are the electronic interactions within the molecular assemblies, effects of translation, rotation and vibration of the participating molecules, and the interaction with the solvent. They are to be investigated based on natural constants and laws of nature and not on experiments. The number of empirical models is kept as small as possible. However, simplifications can not be avoided.

The electronic interactions are described with quantum mechanics which describes a physical system by its quantum state. The evolution with time of the quantum state can be determined by the non-relativistic, time-dependent Schrödinger equation[73]

$$i\hbar\frac{\partial}{\partial t}\Psi = \hat{H}\Psi \tag{2.1}$$

where $\hat{H}$ is the Hamiltonian and the quantum state is expressed by a wave function Ψ. For the large systems investigated in this study the non-relativistic Schrödinger equation is only solvable in reasonable time by using various simplifications or approximations, respectively.

The first simplification is already done by using the non-relativistic Schrödinger equation omitting relativistic effects, the spin-orbit interactions, the spin of the atomic nuclei and further interactions which have negligible influence on the valence electrons (i.e. the chemical properties of a molecule).

The Hamiltonian $\hat{H}$ associated with the energies of interest for this thesis is given in

atomic units by

$$\hat{H} = \underbrace{-\sum_{i}^{n} \frac{\vec{\nabla}_i^2}{2}}_{\hat{T}_\mathrm{e}} \underbrace{-\sum_{A}^{C} \frac{\vec{\nabla}_A^2}{2M_A}}_{\hat{T}_\mathrm{C}} \underbrace{-\sum_{i}^{n}\sum_{A}^{C} \frac{Z_A}{\left|\vec{r}_i - \vec{R}_A\right|}}_{\hat{V}_\mathrm{eC}} + \underbrace{\sum_{i}^{n}\sum_{j>i}^{n} \frac{1}{\left|\vec{r}_i - \vec{r}_j\right|}}_{\hat{V}_\mathrm{ee}} + \underbrace{\sum_{A}^{C}\sum_{B>A}^{C} \frac{Z_A Z_B}{\left|\vec{R}_A - \vec{R}_B\right|}}_{\hat{V}_\mathrm{CC}} \,. \tag{2.2}$$

where the atomic nuclei of the number C have the indexes A and B, the spatial coordinates $\vec{R}$, the mass M, and the nuclear charge Z. The electrons of the number n have the indexes i and j, and the spatial coordinates $\vec{r}$. $\vec{\nabla} = \left(\frac{\delta}{\delta x}, \frac{\delta}{\delta y}, \frac{\delta}{\delta z}\right)$ is the nabla operator. The operators for the kinetic energy of the electrons and atomic nuclei are $\hat{T}_\mathrm{e}$ and $\hat{T}_\mathrm{C}$, respectively and the ones for the potential energy of the electron-nuclei-interaction, the electron-electron-interaction and the nuclei-nuclei-interaction are $\hat{V}_\mathrm{eC}$, $\hat{V}_\mathrm{ee}$ and $\hat{V}_\mathrm{CC}$, respectively. This Hamiltonian $\hat{H}$ is not explicitly time-dependent. This can be used for another simplification, because it allows to use a separation ansatz:

$$\Psi\left(\vec{R}, \vec{r}, \vec{s}, t\right) = \Phi\left(\vec{R}, \vec{r}, \vec{s}\right) \cdot P\left(t\right) \,, \tag{2.3}$$

with the time t and the spin coordinates $\vec{s}$ of the electrons. Putting the equation (2.3) in the time-dependent Schrödinger equation (2.1) leads to

$$\frac{i\hbar \frac{\partial}{\partial t} P\left(t\right)}{P\left(t\right)} = \frac{\hat{H}\Phi\left(\vec{R}, \vec{r}, \vec{s}\right)}{\Phi\left(\vec{R}, \vec{r}, \vec{s}\right)} = E \,, \tag{2.4}$$

which transforms to the non-relativistic, time-independent Schrödinger equation (TISE)

$$\hat{H}\Phi\left(\vec{R}, \vec{r}, \vec{s}\right) = E\Phi\left(\vec{R}, \vec{r}, \vec{s}\right) \,, \tag{2.5}$$

where E is the total energy of the system. A wave function $\Phi\left(\vec{R}, \vec{r}, \vec{s}\right)$ is a stationary state if $\Phi\left(\vec{R}, \vec{r}, \vec{s}\right)$ fulfills the TISE.

2.2 Born–Oppenheimer approximation

The next simplification is the Born–Oppenheimer approximation (BOA).[74] Its idea is to decouple the movement of the electrons and of the atomic nuclei. Atomic nuclei are much heavier and hence they move much slower than electrons. Therefore, the atomic nuclei in the BOA are considered to be fixed for the motion of the electrons. Thus, the electrons

are moving in a field created by fixed atomic nuclei. The wave function can than be expressed as

$$\Phi\left(\vec{R},\vec{r},\vec{s}\right)=\Phi_{\mathrm{e}}\left(\vec{r},\vec{s},\left\{\vec{R}\right\}\right)\cdot\Phi_{\mathrm{C}}\left(\vec{R}\right)\,, \tag{2.6}$$

with $\Phi_{\mathrm{e}}\left(\vec{r},\vec{s},\left\{\vec{R}\right\}\right)$ being the electronic wave function, which only parametrically depends on the coordinates of the atomic nuclei. $\Phi_C\left(\vec{R}\right)$ is the wave function of the atomic nuclei, but is not of interest in this thesis. As the atomic nuclei are assumed to be fixed, $\hat{T}_{\mathrm{C}}$ is getting zero and the potential energy for the nuclei-nuclei-interaction V_{CC} (calculated with $\hat{V}_{\mathrm{CC}}$) becomes an easy calculable constant. The BOA leads to the non-relativistic, time-independent electronic Schrödinger equation

$$\hat{H}_{\mathrm{e}}\Phi_{\mathrm{e}}\left(\vec{r},\vec{s},\left\{\vec{R}\right\}\right)=\left(\hat{T}_{\mathrm{e}}+\hat{V}_{\mathrm{eC}}+\hat{V}_{\mathrm{ee}}\right)\Phi_{\mathrm{e}}\left(\vec{r},\vec{s},\left\{\vec{R}\right\}\right)=E_{\mathrm{e}}\Phi_{\mathrm{e}}\left(\vec{r},\vec{s},\left\{\vec{R}\right\}\right)\,. \tag{2.7}$$

It should also be pointed out that for the positions of the atomic nuclei a classical description with point like atomic nuclei is used instead of a probability distribution which would arise from the actual wave function of the atomic nuclei $\Phi_{\mathrm{C}}\left(\vec{R}\right)$. The energy $E\left(\left\{\vec{R}\right\}\right)$ of a system is calculated as

$$E\left(\left\{\vec{R}\right\}\right)=E_{\mathrm{e}}\left(\left\{\vec{R}\right\}\right)+V_{\mathrm{CC}}\left(\left\{\vec{R}\right\}\right) \tag{2.8}$$

and depends parametrically on the coordinates $\vec{R}$ of the atomic nuclei. The solutions of $E\left(\left\{\vec{R}\right\}\right)$ in dependence of $\vec{R}$ create a multidimensional surface, the potential energy surface (PES), on which the atomic nuclei move in the BOA.

2.3 Hartree–Fock Method

To solve the non-relativistic, time-independent electronic Schrödinger equation (2.7) the wave function Φ_{e} is approximated with an antisymmetric product of one-electron wave functions $\chi_i\left(\vec{r}_i,\vec{s}_i\right)$ called Slater determinant[75] Φ_{SD},

$$\Phi_{\mathrm{SD}}=\frac{1}{\sqrt{n!}}\begin{vmatrix}\chi_1\left(\vec{r}_1,\vec{s}_1\right) & \chi_2\left(\vec{r}_1,\vec{s}_1\right) & \dots & \chi_n\left(\vec{r}_1,\vec{s}_1\right)\\ \chi_1\left(\vec{r}_2,\vec{s}_2\right) & \chi_2\left(\vec{r}_2,\vec{s}_2\right) & \dots & \chi_n\left(\vec{r}_2,\vec{s}_2\right)\\ \vdots & \vdots & \ddots & \vdots\\ \chi_1\left(\vec{r}_n,\vec{s}_n\right) & \chi_2\left(\vec{r}_n,\vec{s}_n\right) & \dots & \chi_n\left(\vec{r}_n,\vec{s}_n\right)\end{vmatrix}\,. \tag{2.9}$$

The one-electron wave functions $\chi_i(\vec{r}_i, \vec{s}_i)$ are called spin orbitals. They consist of a spatial function $\phi_i(\vec{r}_i)$ and a spin function $\sigma_{\text{sp}}(\vec{s}_i)$ which is either spin-up (α) or spin-down (β)

$$\chi_i(\vec{r}_i, \vec{s}_i) = \phi_i(\vec{r}_i)\,\sigma_{\text{sp}}(\vec{s}_i) \qquad \text{with} \qquad \sigma_{\text{sp}}(\vec{s}_i) = \alpha \text{ or } \beta\,. \tag{2.10}$$

The use of an antisymmetric product is necessary because electrons are fermions. A wave function of fermions changes its sign if two of the fermions are exchanged. This is known as Pauli's exclusion principle. A linear combination of all possible Slater determinants built from an infinite set of one-electron wave functions is needed to get the exact wave function of a many-electron system. However, this is impossible to handle in an actual calculation. Therefore, the wave function is approximated in a first step by using only one Slater determinant built from a finite set of one-electron wave functions.

The variation principle is applied to find the Slater determinant which is closest to the exact solution. It states that an approximation for the wave function will always have an energy higher or equal to the energy of the exact solution but never a lower energy. In case of normalized wave functions the variation principle can be written as:

$$\langle\Phi_{\text{SD}}|\,\hat{H}_{\text{e}}\,|\Phi_{\text{SD}}\rangle = E_{\text{HF}} \geq E_{\text{e}} = E_{\text{exact}} = \langle\Phi_{\text{exact}}|\,\hat{H}_{\text{e}}\,|\Phi_{\text{exact}}\rangle\ . \tag{2.11}$$

E_{HF} is the Hartree–Fock energy. Equality is only reached in case the approximation for the wave function is the exact solution ($\Phi_{\text{SD}} = \Phi_{\text{exact}}$). This lower limit can be used to find the best approximation for the exact wave function by varying the approximate wave function Φ_{SD} until the minimum energy is reached. In order to calculate and minimize the Hartree–Fock energy $\left(E_{\text{HF}} = \langle\Phi_{\text{SD}}|\,\hat{H}_{\text{e}}\,|\Phi_{\text{SD}}\rangle\right)$ the Hamiltonian $\hat{H}_{\text{e}}$ can be separated into the one-electron operator $\hat{h}_i$ which describes the kinetic energy of the electron i in the field created by all atomic nuclei and the two-electron operator $\hat{g}_{ij}$ which describes electron-electron interaction:

$$\hat{H} = \sum_{i}^{n} \hat{h}_i + \sum_{i}^{n}\sum_{j>i}^{n} \hat{g}_{ij}\,, \tag{2.12}$$

$$\hat{h}_i = -\frac{\vec{\nabla}_i^2}{2} - \sum_{A}^{C} \frac{Z_A}{\left|\vec{r}_i - \vec{R}_A\right|}\,, \tag{2.13}$$

$$\hat{g}_{ij} = \frac{1}{|\vec{r}_i - \vec{r}_j|}\,. \tag{2.14}$$

The matrix elements h_a of $\hat{h}$ and the matrix elements J_{ab} and K_{ab} of $\hat{g}$ can be expressed with the spin orbitals $\chi_i(\vec{r}_i, \vec{s}_i)$ by using the Slater-Condon rules. This can be applied to calculate the Hartree–Fock energy E_{HF}

$$E_{\mathrm{HF}} = \sum_{a=1}^{N} h_a + \frac{1}{2} \sum_{a=1}^{N} \sum_{b=1}^{N} (J_{ab} - K_{ab}) \,, \tag{2.15}$$

with

$$h_a = \int \chi_a^*(\vec{x}_1)\, \hat{h}_1 \chi_a(\vec{x}_1)\, \mathrm{d}\vec{x_1} \,, \tag{2.16}$$

$$J_{ab} = \int \chi_a^*(\vec{x}_1)\, \chi_b^*(\vec{x}_2)\, \hat{g}_{12} \chi_a(\vec{x}_1)\, \chi_b(\vec{x}_2)\, \mathrm{d}\vec{x_1} \mathrm{d}\vec{x_2} \,, \tag{2.17}$$

$$K_{ab} = \int \chi_a^*(\vec{x}_1)\, \chi_b^*(\vec{x}_2)\, \hat{g}_{12} \chi_a(\vec{x}_2)\, \chi_b(\vec{x}_1)\, \mathrm{d}\vec{x_1} \mathrm{d}\vec{x_2} \,. \tag{2.18}$$

N is the number of occupied spin orbitals, whereas a and b are the summation indices. For better clarity $\vec{r}_i$ and $\vec{s}_i$ are merged to $\vec{x}_i$. J_{ab} is called Coulomb integral and can be interpreted as classical repulsion between two charge distributions which are described by $|\chi_a(\vec{x}_1)|^2$ and $|\chi_b(\vec{x}_2)|^2$. K_{ab} is called an exchange integral, but in contrast to the Coulomb integral it has no analogue in classical physics.

The Hartree–Fock energy E_{HF} will then be minimized with respect to the orthonormal spin orbitals

$$\langle \chi_a | \chi_b \rangle = \delta_{ab} = \begin{cases} 1 & \text{for } a = b \\ 0 & \text{for } a \neq b \end{cases} \,. \tag{2.19}$$

The minimization can be done under this constraint with the method of Lagrange multipliers

$$\mathcal{L} = E_{\mathrm{HF}} - \sum_{a}^{N} \sum_{b}^{N} \lambda_{ab} \left(\langle \chi_a | \chi_b \rangle - \delta_{ab} \right) \,, \tag{2.20}$$

$$\delta\mathcal{L} = \delta E_{\mathrm{HF}} - \sum_{a}^{N} \sum_{b}^{N} \lambda_{ab} \delta \langle \chi_a | \chi_b \rangle = 0 \,. \tag{2.21}$$

$\mathcal{L}$ is the Lagrange functional, λ_{ab} are the Lagrange multipliers and δ_{ab} is the Kronecker delta. Solving equation (2.21) leads to the so called Hartree–Fock equations

$$\hat{F} |\chi_a\rangle = \sum_{b}^{N} \lambda_{ab} |\chi_b\rangle \tag{2.22}$$

with the Fock operator $\hat{F}$.

With the help of an unitary transformation the matrix of Lagrange multipliers can be diagonalized (this means $\lambda_{ab} = 0$ and $\lambda_{aa} = \varepsilon_a$)

$$\hat{F} \left| \tilde{\chi}_a \right\rangle = \varepsilon_a \left| \tilde{\chi}_a \right\rangle \, . \tag{2.23}$$

The spin oribtals χ_a transform to the so called canonical spin orbitals $\tilde{\chi}_a$ with the orbital energies ε_a. The Fock operator $\hat{F}$ is not changed due to the unitary transformation and can be described by

$$\hat{F}_i = \hat{h}_i + \sum_b^N \left(\hat{J}_b \left(\vec{x}_i \right) - \hat{K}_b \left(\vec{x}_i \right) \right) \, , \tag{2.24}$$

with the Coulomb operator $\hat{J}_b \left(\vec{x}_i \right)$

$$\hat{J}_b \left(\vec{x}_i \right) \tilde{\chi}_a \left(\vec{x}_i \right) = \left[\int \tilde{\chi}_b^* \left(\vec{x}_j \right) \hat{g}_{ij} \tilde{\chi}_b \left(\vec{x}_j \right) \mathrm{d}\vec{x}_j \right] \tilde{\chi}_a \left(\vec{x}_i \right) \tag{2.25}$$

and the exchange operator $\hat{K}_b \left(\vec{x}_i \right)$

$$\hat{K}_b \left(\vec{x}_i \right) \tilde{\chi}_a \left(\vec{x}_i \right) = \left[\int \tilde{\chi}_b^* \left(\vec{x}_j \right) \hat{g}_{ij} \tilde{\chi}_a \left(\vec{x}_j \right) \mathrm{d}\vec{x}_j \right] \tilde{\chi}_b \left(\vec{x}_i \right) \, . \tag{2.26}$$

The Coulomb and the exchange operator require that all other orbitals have to be known to solve the Hartree-Fock equation (2.23) for a specific orbital. Therefore, Hartree-Fock equations are pseudo-eigenfunction-equations and they can only be solved in an iterative procedure. A starting guess for the spin orbitals is used to construct the starting Fock operator and solve the Hartree-Fock equations. This gives a new set of spin orbitals which can be used for a new iteration. The procedure continues until the spin orbitals before and after an iteration step are identical within a given error margin. Therefore, the Hartree–Fock method is also called self-consistent field (SCF) method.

A common step is now to expand the spin orbitals $\tilde{\chi}_a$ linearly into a finite set of atom-centered basis functions ϕ_μ,

$$\tilde{\chi}_a = \sum_\mu^L C_{\mu a} \phi_\mu \, . \tag{2.27}$$

$C_{\mu a}$ are the expansion coefficients and will serve as optimization parameters. L is the number of basis functions and is finite. Therefore, this is an approximation as well. The equation (2.27) is now inserted into equation (2.23). Multiplying a specific basis function from the left and subsequent integrating yields for closed shell systems the so called

Roothaan–Hall[76,77] equations

$$\mathbf{FC} = \mathbf{SC}\boldsymbol{\varepsilon}\,, \tag{2.28}$$

while open shell systems will be discussed in the next section. The Fock matrix $\mathbf{F}$ contains the matrix elements $F_{\nu\mu} = \langle\phi_\nu|\,\hat{F}\,|\phi_\mu\rangle$. $\mathbf{S}$ is the overlap matrix with $S_{\nu\mu} = \langle\phi_\nu|\phi_\mu\rangle$. The overlap matrix shows up because the basis functions are not orthogonal. $\mathbf{C}$ represents the expansion coefficient matrix and $\boldsymbol{\varepsilon}$ is a diagonal matrix of the orbital energies ε_a. Details like the formation of the density matrix, which is the product of the expansion coefficients $C_{\mu a}$, and the diagonalization of $\mathbf{S}$, are not discussed here. A detailed overview of the Hartree-Fock method[78–82] can be found in modern textbooks on molecular quantum chemistry like "Modern Quantum Chemistry" by Szabo and Ostlund, 1982.[83] The most intuitive procedure for solving the non-relativistic, time-independent electronic Schrödinger equation (2.7) with the Hartree-Fock method is the so called "in-core" ("core" is referring to the computational procedure) SCF procedure:

a) Choose the atomic nuclei and their spatial coordinates $\vec{R}$, the number of electrons n and the basis set.

b) Calculate all required integrals.

c) Diagonalize $\mathbf{S}$.

d) Make a starting guess for the expansion coefficients matrix $\mathbf{C}$.

e) Calculate the density matrix from (the starting guess of) $\mathbf{C}$.

f) Calculate the Fock matrix.

g) Diagonalize the Fock matrix. The eigenvectors build the new expansion coefficient matrix $\mathbf{C}$ and the eigenvalues the new orbital energy matrix $\boldsymbol{\varepsilon}$.

h) The new expansion coefficient matrix $\mathbf{C}$ yields the new density matrix. If the new density matrix is identical (within the chosen accuracy) to the previous density matrix, the procedure has converged and finished. Otherwise continue with step f.

The number of two-electron integrals is proportional to the number of basis functions L to the power of four (i.e. L^4). For the systems in this thesis the amount of two-electron integrals is too large to be stored on the main memory of today's computer systems. Therefore, approximations have to be used which reduce the number of integrals (see

section 2.4.2). For the calculations done without these approximations, "direct SCF" was applied. This means the integrals are calculated in every iteration instead of only once for the whole SCF cycle. In the direct SCF procedure the calculation of the two-electron integrals is the most time-consuming part. Since their number is proportional to the number of basis functions to the power of four (i.e. L^4), the calculation time scales formally with L^4 as well.

As each electron experiences only an averaged repulsive potential from all other electrons, the Hartree-Fock method is a mean field approximation. For electrons with the same spin the Hartree-Fock method includes the so called "Fermi correlation" due to the exchange term which derives from the antisymmetry of the Slater determinant. All other correlation effects are not considered. The Hartree-Fock energy reaches the "Hartree-Fock limit" $E_{\mathrm{HF}}^{\mathrm{Limit}}$ as the basis set approaches completeness. The difference between the exact energy E^{Exact} and the Hartree-Fock limit $E_{\mathrm{HF}}^{\mathrm{Limit}}$ is called "correlation energy" E^{corr}

$$E^{\mathrm{corr}} = E^{\mathrm{Exact}} - E_{\mathrm{HF}}^{\mathrm{Limit}} \tag{2.29}$$

despite the Hartree-Fock method already covers the correlation of electrons with the same spin partly.

2.3.1 RHF and UHF

Closed shell systems (often) can be described reasonably well by using doubly occupied spatial orbitals. Thus, two spin orbitals $\chi_{i\alpha}$ and $\chi_{i\beta}$ are represented by the same spatial orbital ϕ_i multiplied with an α and a β spin function, respectively. The two spin orbitals will have the same orbital energy. This approach is known as restricted Hartree-Fock (RHF) and leads to the Roothaan–Hall equations (2.28). For open shell systems unrestricted Hartree-Fock (UHF) is applied, which uses different spatial orbitals for different spins. This yields the Pople-Nesbet-Berthier equations[84]

$$\begin{aligned} \mathbf{F}^{\alpha}\mathbf{C}^{\alpha} &= \mathbf{S}\mathbf{C}^{\alpha}\boldsymbol{\varepsilon}^{\alpha}\,, \\ \mathbf{F}^{\beta}\mathbf{C}^{\beta} &= \mathbf{S}\mathbf{C}^{\beta}\boldsymbol{\varepsilon}^{\beta}\,, \end{aligned} \tag{2.30}$$

which consist of a pair of coupled Roothaan equations (2.28). These equations are coupled since both Fock matrices $\mathbf{F}^{\alpha}$ and $\mathbf{F}^{\beta}$ depend on $\mathbf{C}^{\alpha}$ as well as on $\mathbf{C}^{\beta}$. The orbital energies for α and β electrons differ in UHF. Finally it must be added that the wave function in

the UHF approach is not an eigenfunction of the total spin operator $\hat{S}^2$. The spin state can get contaminated by other spin states. The expectation value of $\hat{S}^2$ can be used as a measure of the contamination which should be low.

Only restriced and unrestriced approaches were used in this thesis. An alternative to UHF for open shell systems is the restricted open-shell Hartree–Fock (ROHF) approach.[85] In the ROHF approach doubly occupied spatial orbitals are used as far as possible. Only for the rest different spatial orbitals are applied for different spins. ROHF is more difficult to implement and is not used in this thesis.

2.3.2 Fermi Smearing

"Fermi smearing"[86] can be used to find the lowest spin state in an unrestricted calculation. Instead of integer occupation numbers fractional occupation numbers are used for the orbitals. The fractional occupation number n_a of an orbital with energy ε_a is calculated at temperature T as

$$n_a = \frac{1}{2}\,\mathrm{erfc}\left(\frac{\varepsilon_a - \mu}{\frac{4k_\mathrm{B}T}{\sqrt{\pi}}}\right) . \tag{2.31}$$

where erfc is the complementary error function and k_B is the Boltzmann constant. The "Fermi level" μ gives the orbital energy which corrpondes to $n_a = 0.5$ with equation (2.31). However, μ is not the actual Fermi level because equation (2.31) is chosen to diminish the appearance of extreme small occupation numbers and not to represent exactly the Fermi–Dirac statistics.[87,88]

The occupation numbers n_a and the Fermi level μ are determined in every SCF step for the given set of orbitals with energies ε_a at temperature T provided that the sum over the occupation numbers n_a is equal to the total number of electrons. Due to this approach the occupation pattern changes during the SCF procedure and the minimization of the energy leads to the lowest spin state. The Fermi smearing calculations in this thesis were started with a temperature of 300 K. This temperature was reduced at every iteration step about 2% until 50 K was reached. Thus, integer occupation numbers are yielded at the end of the SCF procedure.

2.3.3 Correlation energy

The Hartree-Fock method covers in the Hartree-Fock limit $E_\mathrm{HF}^\mathrm{Limit}$ about 99% of the total energy. The remaining energy is the correlation energy E^corr (see equation (2.29)). Even

if the correlation energy amounts only to about 1% of the total energy, it is essential for binding energies. For a given basis set the Hartree-Fock method already gives the best result of the total energy achievable with a single determinant description of the wave function. Therefore, more determinants have to be included in the description of the wave function to cover the last 1% of the total energy or the correlation energy, respectively.

In the Hartree-Fock approach shown above, the spin orbitals are linearly expanded into a finite set of atom-centered basis functions and then the Roothaan–Hall equations are solved. This yields more molecular orbitals than needed to describe the electrons. The excess orbitals are unoccupied and are called virtual orbitals, but only the occupied orbitals are optimized and form the wave function (expressed with a Slater determinant) as well as determine the total energy. Starting from the Hartree-Fock approach new determinants can be built by replacing occupied orbitals with virtual orbitals in the forming an "excited" Slater determinant. Depending on the number of replaced orbitals (one, two, three, etc.) this yields singly, doubly, triply, etc. excited determinants relative to the Hartree-Fock determinant. The most intuitive approach would be a linear combination of these determinants. This is known as Configuration Interaction (CI) or Full Configuration Interaction (FCI) if all possible excitations are considered. With an infinitely large basis set FCI would exactly solve the non-relativistic, time-independent electronic Schrödinger equation (2.7). However, this is not achievable. FCI with a finite basis set yields the best solution possible of the non-relativistic, time-independent electronic Schrödinger equation (2.7) in the given basis set. Nevertheless, even this is only feasible for small molecules, but not for systems investigated in this thesis. Therefore, FCI must be truncated. For example only single and double excited determinants could be used. This is known as CISD (CI with Singles and Doubles). Unfortunately the CISD approach yields conceptually wrong binding energies. The binding energy of a dimer is calculated as difference between the total energy of the dimer and the total energy of the two monomers (see equation (2.129)). A truncated CI approach would include higher excited determinants for the combined system of two separated monomers than for the dimer. Hence, more of the correlation energy is covered for the monomers as for the dimer. Errors of this kind are summarized in the concept of size extensivity. A method is size extensive if two monomers calculated at a distance at which they do not interact (like 100 Å) have the same energy as the sum of the separately calculated monomers. The definition of size extensivity is not consistent in literature and it is sometimes interchanged with size

consistency. In this thesis the before mentioned definition of size extensivity is used.

Instead of a truncated CI the size extensive DF-LCCSD(T)[55] method was employed in this thesis. DF-LCCSD(T) is based on coupled cluster (CC)[54] which is size extensive at all levels of truncation. CC uses the excitation operator $\hat{T}$

$$\hat{T} = \sum_{i}^{n} \hat{T}_i \tag{2.32}$$

with the number of electrons n and the operator $\hat{T}_i$, acting on the Hartree Fock wave function/Slater determinant $\left|\Phi_{\mathrm{SD}}^{\mathrm{HF}}\right\rangle$ generating all i-order excited determinants. The operators for single and double excitation $\hat{T}_1$ and $\hat{T}_2$ yield

$$\hat{T}_1 \left|\Phi_{\mathrm{SD}}^{\mathrm{HF}}\right\rangle = \sum_{a}^{occ} \sum_{r}^{vir} t_a^r \left|\Phi_a^r\right\rangle \ , \tag{2.33}$$

$$\hat{T}_2 \left|\Phi_{\mathrm{SD}}^{\mathrm{HF}}\right\rangle = \sum_{a<b}^{occ} \sum_{r<s}^{vir} t_{ab}^{rs} \left|\Phi_{ab}^{sr}\right\rangle \ , \tag{2.34}$$

occ is the number of occupied oribtals with the summation indices a and b, *vir* is the number of virtual oribtals with the summation indices r and s. Φ_a^r and Φ_{ab}^{sr} are the single and double excited determinants, and t_a^r and t_{ab}^{rs} are the expansion coefficients called amplitudes.

The size extensivity at all levels of truncation is achieved by an exponential ansatz for the coupled cluster wave function $|\Psi_{\mathrm{CC}}\rangle$

$$|\Psi_{\mathrm{CC}}\rangle = \exp\left(\hat{T}\right) \left|\Phi_{\mathrm{SD}}^{\mathrm{HF}}\right\rangle \ . \tag{2.35}$$

The exponential operator $\exp\left(\hat{T}\right)$ can be written as

$$\exp\left(\hat{T}\right) = \underbrace{1}_{\mathrm{HF}} + \underbrace{\hat{T}_1}_{\text{single excitation}} + \underbrace{\left(\hat{T}_2 + \frac{1}{2}\hat{T}_1^2\right)}_{\text{double excitation}} + \underbrace{\left(\hat{T}_3 + \hat{T}_2\hat{T}_1 + \frac{1}{6}\hat{T}_1^3\right)}_{\text{triple excitation}} + \dots \ . \tag{2.36}$$

The terms are sorted in a way that the first term generates the Hartree Fock determinate, the second term generates all singly excited determinates etc. The product terms (like $\hat{T}_1^2$) are called disconnected and the pure terms (like $\hat{T}_2$) connected. For the truncated Coupled Cluster Singles and Doubles (CCSD) the exponential operator $\exp\left(\hat{T}_1 + \hat{T}_2\right)$ can

be written as

$$\exp\left(\hat{T}\right) = 1 + \hat{T}_1 + \left(\hat{T}_2 + \frac{1}{2}\hat{T}_1^2\right) + \left(\hat{T}_2\hat{T}_1 + \frac{1}{6}\hat{T}_1^3\right) + \dots \ . \tag{2.37}$$

Pluging equation (2.35) into the non-relativistic, time-independent electronic Schrödinger equation (2.7) yields

$$\hat{H}_\mathrm{e} \exp\left(\hat{T}\right) |\Phi_\mathrm{SD}\rangle = E_\mathrm{e} \exp\left(\hat{T}\right) |\Phi_\mathrm{SD}\rangle \ . \tag{2.38}$$

In contrast to CI the variational principle cannot be applied in the coupled cluster approach, because it results in an intractable set of nonlinear equations. Instead the coupled cluster Schrödinger equation (2.38) is projected onto the Hartree Fock and all excited determinants, to gain the energy and the amplitudes. It is beneficial to use the similarity transformed Hamiltonian $\exp\left(-\hat{T}\right) \hat{H}_\mathrm{e} \exp\left(\hat{T}\right)$. This is done by multiplying equation (2.38) first from the left with $\exp\left(-\hat{T}\right)$ and performing afterwards the projection. The $\exp\left(-\hat{T}\right)$ operator is the deexcitation operator and works on the function to the left. The resulting equations are

$$\left\langle \Phi_\mathrm{SD}^\mathrm{HF} \right| \exp\left(-\hat{T}\right) \hat{H}_\mathrm{e} \exp\left(\hat{T}\right) \left| \Phi_\mathrm{SD}^\mathrm{HF} \right\rangle = E_\mathrm{e} \tag{2.39}$$

$$\begin{aligned} \langle \Phi_a^r | \exp\left(-\hat{T}\right) \hat{H}_\mathrm{e} \exp\left(\hat{T}\right) \left| \Phi_\mathrm{SD}^\mathrm{HF} \right\rangle &= 0 \\ \langle \Phi_{ab}^{rs} | \exp\left(-\hat{T}\right) \hat{H}_\mathrm{e} \exp\left(\hat{T}\right) \left| \Phi_\mathrm{SD}^\mathrm{HF} \right\rangle &= 0 \\ &\vdots \end{aligned} \tag{2.40}$$

The similarity transformed Hamiltonian $\exp\left(-\hat{T}\right) \hat{H}_\mathrm{e} \exp\left(\hat{T}\right)$ can be simplified with the Baker-Campbell-Hausdorff (BCH) expansion. Due to the two-particle nature of the Hamiltonian the BCH expansion truncates conveniently after the quadruply nested commutator. Thus, the similarity transformed Hamiltonian $\exp\left(-\hat{T}\right) \hat{H}_\mathrm{e} \exp\left(\hat{T}\right)$ can be exactly expanded as

$$\begin{aligned} \exp\left(-\hat{T}\right) \hat{H}_\mathrm{e} \exp\left(\hat{T}\right) &= \hat{H}_\mathrm{e} + \left[\hat{H}_\mathrm{e}, \hat{T}\right] + \frac{1}{2!}\left[\left[\hat{H}_\mathrm{e}, \hat{T}\right], \hat{T}\right] + \frac{1}{3!}\left[\left[\left[\hat{H}_\mathrm{e}, \hat{T}\right], \hat{T}\right], \hat{T}\right] \\ &\quad + \frac{1}{4!}\left[\left[\left[\left[\hat{H}_\mathrm{e}, \hat{T}\right], \hat{T}\right], \hat{T}\right], \hat{T}\right] \ . \end{aligned} \tag{2.41}$$

With the BCH expansion the equations (2.40) can be solved in order to get the amplitudes which are then used to calculate the energy E_e of equation (2.39). For CCSD

only the two explicitly shown equations (2.40) have to be solved. In the calculation of the amplitudes for CCSD up to quadruple excitations are considered with $\hat{T}_2^2$, $\hat{T}_2\hat{T}_1^2$ and $\hat{T}_1^4$. The inclusion of these excitation is the reason for the size extensivity of the truncated coupled cluster.

Unfortunately the accuracy of CCSD is not high enough for chemical accuracy and the next truncation level (CCSDT) is not feasible for large molecules. Therefore, the perturbation theory[89] has to be applied to add a perturbative correction for the connected Triples $\hat{T}_3$ based on CCSD amplitudes. The forth and fifth order perturbation theory terms which contain the connected triples amplitudes are used, while for the fifth order only those which arise from the projection onto single and double exitated determinates are used. Coupled Cluster Singles and Doubles with perturbative Triples is abbreviated as CCSD(T). Its calculation time scales with the system size to the power of seven and has a large prefactor. Thus, CCSD(T) was not feasible for all systems in this thesis, and further approximations are utilized.

Electron correlation is a short range effect. A good approximation for the electron correlation can be gained by determining only the electron correlation of spatially close electrons with high accuracy, i.e. on coupled cluster level. This approach is called local (L) coupled cluster. The canonical orbitals from the Hartree-Fock approach are usually delocalized over the whole molecular system; therefore, localized molecular orbitals (LMOs) are used. LMOs are molecular orbitals which are concentrated at a small and defined spatial region like a specific bond or atom. They are gained in this thesis with the Pipek-Mezey[90,91] localization scheme by applying an unitary transformation onto the occupied canonical orbitals from the Hartree-Fock approach. In the applied LCCSD(T) method[55] the excitations from pairs of LMOs are restricted to a certain subspace called domain. The pairs of LMOs are divided into different classes according to their distance. Close pairs are treated with coupled cluster theory, more distant pairs with perturbation theory and pairs far away from each other are neglected. The domains are generated by projecting atomic orbitals onto the canonical virtual orbital space. The generated orbitals are called projected atomic orbitals[92] (PAOs). They are orthogonal to the LMOs but non-orthogonal to each other. PAOs spatially close to a LMO pair are used to build the corresponding domain.

Furthermore the density fitting (DF) approximation is used to reduce the computational costs. Density fitting is also called “resolution of the identity” and is explained in section 2.4.2 for the case of Density Functional Theory. In total it is possible to perform DF-

LCCSD(T) single point calculations with sufficiently large basis sets for the monovalent complexes of this thesis.

2.4 Density Functional Theory

Density Functional Theory (DFT) is an electronic structure method to describe the electronic ground state of a many-body system based on the spatially dependent electron density $\rho(\vec{r})$. The electron density of a system with n electrons is the absolute square of the wave function integrated over all spin coordinates $\vec{s}_i$ and all except one spatial coordinates

$$\rho(\vec{r}) = n \int \cdots \int \Phi^*(\vec{x}_1, \vec{x}_2 \ldots, \vec{x}_n)\, \Phi(\vec{x}_1, \vec{x}_2 \ldots, \vec{x}_n)\, \mathrm{d}\vec{s}_1 \mathrm{d}\vec{x}_2 \ldots \mathrm{d}\vec{x}_n \tag{2.42}$$

where possible $\vec{r}_i$ and $\vec{s}_i$ are merged to $\vec{x}_i$.

Early attempts to use the electron density were made by Thomas[93] and Fermi[94] in 1927. However, a firm physical foundation for the Density Functional Theory (DFT) was first established in 1964 with the two Hohenberg–Kohn theorems.[56] The first theorem proves that the ground state electronic energy is determined uniquely by the electron density $\rho(\vec{r})$. The second proves that the variation principle for the density holds true for DFT. Thus, the exact functional $E[\rho]$ which connects electron density and ground state electronic energy, gives the lowest energy only with the true ground state electron density. One year after the physical foundation Kohn and Sham published an orbital-based framework which is known as "Kohn-Sham DFT".[57] It has become the basis for modern DFT as other DFT methods failed in determining the kinetic energy. In Kohn-Sham DFT a system of non-interacting electrons which are moving in an effective potential $V_\mathrm{S}(\vec{r})$ is used to describe the real system of interacting electrons. The effective potential $V_S(\vec{r})$ is chosen in order that the electron density $\rho_S(\vec{r})$ of the non-interacting system is equal to the electron density $\rho_0(\vec{r})$ of the real system. The Hamiltonian $\hat{H}_\mathrm{KS}$ for the non-interacting system is

$$\hat{H}_\mathrm{KS} = -\frac{1}{2}\sum_i^n \vec{\nabla}_i^2 + \sum_i^n V_\mathrm{S}(\vec{r}_i) = \sum_i^n \hat{F}_i^\mathrm{KS}\,. \tag{2.43}$$

$\hat{H}_\mathrm{KS}$ can be written as sum of one-electron Hamiltonians called Kohn-Sham operators $\hat{F}_i^\mathrm{KS}$. The ground state of the non-interacting system is exactly represented by a single

Slater determinant Θ_{KS}, because such a system contains no electron-electron interaction (the electrons are just moving in an effective potential $V_{\mathrm{S}}(\vec{r})$)

$$\Theta_{\mathrm{KS}} = \frac{1}{\sqrt{n!}} \begin{vmatrix} \varphi_1(\vec{x}_1) & \varphi_2(\vec{x}_1) & \dots & \varphi_n(\vec{x}_1) \\ \varphi_1(\vec{x}_2) & \varphi_2(\vec{x}_2) & \dots & \varphi_n(\vec{x}_2) \\ \vdots & \vdots & \ddots & \vdots \\ \varphi_1(\vec{x}_n) & \varphi_2(\vec{x}_n) & \dots & \varphi_n(\vec{x}_n) \end{vmatrix} . \tag{2.44}$$

The spin orbitals φ are called Kohn-Sham orbitals to distinguish these from the orbitals in the Hartree-Fock methods. The orbitals are determined with the Kohn-Sham operator

$$\hat{F}^{\mathrm{KS}} |\varphi_a\rangle = \varepsilon_a |\varphi_a\rangle . \tag{2.45}$$

Inserting the Slater determinant Θ_{KS} (eqatuion (2.44)) into equation (2.42) yields

$$\rho_S(\vec{r}) = \sum_a^N |\varphi_a(\vec{r})|^2 , \tag{2.46}$$

where the spin is integrated out and $\varphi_a(\vec{r})$ is only a spatial orbital. The connection to the real system is established by choosing the effective potential $V_S(\vec{r})$ so that the sum of the squared moduli of the Kohn-Sham orbitals $\varphi_a(\vec{r})$ is equal the exact electron density $\rho_0(\vec{r})$ of the real system

$$\rho_S(\vec{r}) = \sum_a^N |\varphi_a(\vec{r})|^2 \stackrel{!}{=} \rho_0(\vec{r}) . \tag{2.47}$$

The effective potential $V_S(\vec{r})$ does not depend on the spin. However, the exact density functional is not known (see below) and the approximations often use the α- and β-spin density.[95]

In Kohn-Sham DFT the total energy as a functional of the electron density $E[\rho]$ is given by

$$E[\rho] = T_{\mathrm{S}}[\rho] + J[\rho] + E_{\mathrm{XC}}[\rho] + E_{\mathrm{eC}}[\rho] . \tag{2.48}$$

$T_{\mathrm{S}}[\rho]$ gives the kinetic energy of the non-interacting system which has the same electron density as the real system

$$T_{\mathrm{S}}[\rho] = -\frac{1}{2} \sum_a^N \langle\varphi_a| \vec{\nabla}^2 |\varphi_a\rangle . \tag{2.49}$$

$T_\mathrm{S}[\rho]$ is considered as functional of the electron density, because the Kohn-Sham orbitals are connected to the electron density by equation (2.47). $J[\rho]$ is the functional of the classical Coulomb interaction

$$\begin{aligned} J[\rho] &= \frac{1}{2}\iint \rho(\vec{r}_1)\frac{1}{|\vec{r}_1-\vec{r}_2|}\rho(\vec{r}_2)\,\mathrm{d}\vec{r}_1\mathrm{d}\vec{r}_2 \\ &= \frac{1}{2}\sum_a^N\sum_b^N \iint |\varphi_a(\vec{r}_1)|^2 \frac{1}{|\vec{r}_1-\vec{r}_2|}|\varphi_b(\vec{r}_2)|^2\,\mathrm{d}\vec{r}_1\mathrm{d}\vec{r}_2\,. \end{aligned} \tag{2.50}$$

$E_\mathrm{eC}[\rho]$ is the functional for the electron-nuclei-interaction

$$\begin{aligned} E_\mathrm{eC}[\rho] &= -\int\sum_A^C \frac{Z_A}{|\vec{r}_1-\vec{R}_A|}\rho(\vec{r}_1)\,\mathrm{d}\vec{r}_1 \\ &= -\sum_a^N\int\sum_A^C \frac{Z_A}{|\vec{r}_1-\vec{R}_A|}|\varphi_a(\vec{r}_1)|^2\,\mathrm{d}\vec{r}_1\,. \end{aligned} \tag{2.51}$$

$E_\mathrm{XC}[\rho]$ is the so called exchange-correlation functional. It contains the self-interaction correction, the exchange and correlation effects and the part of kinetic energy which is not covered by $T_\mathrm{S}[\rho]$.

In the next step the variational principle is applied like in the Hartee-Fock approximation. The energy is minimized with respect to the Kohn-Sham orbitals upon condition that the Kohn-Sham orbitals stay orthonormal $\langle\varphi_i|\varphi_j\rangle = \delta_{ij}$. This results in equation (2.45)[95,96] with the effective potential $V_S(\vec{r})$

$$V_\mathrm{S}(\vec{r}_1) = \int\frac{\rho(\vec{r}_2)}{|\vec{r}_1-\vec{r}_2|} + V_\mathrm{XC}(\vec{r}_1) - \sum_A^C\frac{Z_A}{|\vec{r}_1-\vec{R}_A|} \tag{2.52}$$

for the Kohn-Sham operator (equation (2.43)). V_S depends on the electron density or the orbitals; therefore, equation (2.45) has to be solved iteratively by a self-consistent field approach like Hartree-Fock. The exchange-correlation potential V_XC is the derivative of the exchange-correlation functional E_XC with respect to the electron density $\rho(\vec{r})$

$$V_\mathrm{XC}(\vec{r}) = \frac{\delta E_\mathrm{XC}[\rho]}{\delta\rho(\vec{r})}\,. \tag{2.53}$$

Up to this point no approximations are made in the Kohn-Sham DFT approach except the Born–Oppenheimer approximation, which is here applied as well. Thus, Kohn-Sham DFT

gives the exact energy or eigenvalue of the non-relativistic, time-independent electronic Schrödinger equation (2.7) in case the exchange-correlation functional E_{XC} is known, but unfortunately E_{XC} is unknown. As a result Kohn-Sham DFT contains approximations for the unknown exchange-correlation functional E_{XC}.

Like in Hartree-Fock the Kohn-Sham orbitals are now linearly expanded into a finite set of atom-centered basis functions. Hence the density $\rho(\vec{r})$ is calculated as

$$\rho(\vec{r}) = \sum_{a}^{N} |\varphi_a(\vec{r})|^2 = \sum_{a}^{N} \sum_{\mu}^{L} \sum_{\nu}^{L} C_{\mu a} C_{\nu a} \phi_\mu(\vec{r}) \phi_\nu(\vec{r}) = \sum_{\mu}^{L} \sum_{\nu}^{L} D_{\mu\nu} \phi_\mu(\vec{r}) \phi_\nu(\vec{r}) \tag{2.54}$$

with the elements $D_{\mu\nu}$ of the Density-matrix $\mathbf{D}$

$$D_{\mu\nu} = \sum_{a}^{N} C_{\mu a} C_{\nu a} \,. \tag{2.55}$$

Analogous to Hartree-Fock which yields in the Roothaan–Hall equations, a matrix equation is obtained for Kohn-Sham DFT which can be solved as described for Hartree-Fock. The largest part of the calculation are the four-center-two-electron integrals for the Coulomb term. As for Hartree-Fock a direct SCF procedure has to be applied and the calculation time scales formally with the number of basis functions to the power of four.

The approximations for the exchange-correlation potential V_{XC} are in general complicated mathematical constructs, which are not shown here. The corresponding integrals cannot be solved analytically and so they are calculated numerically on a grid. Nevertheless, these calculations are much faster than the calculation of the four-center-two-electron integrals for the Coulomb term.

2.4.1 DFT Functionals

The exchange-correlation functionals used in "Kohn-Sham DFT" can be characterized in accordance to the density related variables they take into acount. Perdew[97,98] suggested a hierarchic order by putting the different functional types on different rungs of a ladder, the so called "Jacob's ladder". The different rungs are:

1. Local Density Approximations (LDA),

2. Generalized Gradient Approximation (GGA),

3. meta-GGA,

4. Hybrid GGA and Hybrid meta-GGA,

5. Fully Non-Local.

With every rung more density related variables are added and the functionals should gain more accuracy. In practice the computational effort usually rises with every rung, however, not necessarily the accuracy, especially between functionals of the second, third and fourth rung.[99]

LDA functionals only use the electron density $\rho(\vec{r})$. For systems with an inhomogeneous electron density like molecules, they usually perform poorly. The second rung are the GGA functionals. They use additionally to the electron density $\rho(\vec{r})$ the gradient of the electron density $\nabla\rho(\vec{r})$; thus, enabling the calculation of molecules. The prefix "meta" means that in addition the second order gradient of the electron density $\nabla^2\rho(\vec{r})$ or the kinetic energy density $\tau(\vec{r}) = \frac{1}{2}\sum_a^N |\Delta\varphi_a(\vec{r})|^2$ will be taken into account. The "Hybrid" prefix points out that the Kohn-Sham orbitals are used to calculate the exact, non-local exchange as in the Hartree-Fock method. This non-local exchange is then usually mixed with the exchange calculated from a GGA or meta-GGA.

In this thesis mainly the non-empirical meta-GGA functional TPSS[100–103] was used. It is designed to satisfy exact constraints on the exchange-correlation functional E_{XC} and is not fitted to experimental data. Therefore, TPSS can be used for a broad set of systems. The constraints, used for the construction of non-empirical denstiy functionals can be found in common textbooks on computational chemistry like "Introduction to Computational Chemistry" by Jensen, 2007.[104] The GGA functional BP86[100,101,105–108] was used for the COSMO-RS approach, which is optimized for BP86. It consists of the exchange functional B88[106] and the correlation functional P86,[107,108] which both contain empirical parameters. For the samarium calculations the Hybrid GGA functional PBE0[100–102,109–111] was used. It replaces 25% of the exchange in the GGA functional PBE[100–102,109,110] with the Hartree-Fock exchange calculated from the Kohn-Sham orbitals. The TPSS functional is also based on PBE.

The Hybrid GGA fucntional B3LYP[100,101,105,112,113] was used to calculate excitation energies (see section 2.9.2). Finally the M06-2X[114] functional was applied for the investigation of polyethylene glycol (PEG). It is a Hybrid meta-GGA functional which is fitted for main-group thermochemistry and kinetics.

2.4.2 RI-DFT and MARI-*J*

The most time consuming step in the DFT approach is the calculation of the four-center-two-electron integrals. In the non-Hybrid functionals these integrals only appear for the Coulomb terms. The Coulomb $J[\rho]$ energy given by equation (2.50) can be written with equation (2.54) as

$$J[\rho] = \frac{1}{2}\sum_{\mu}^{L}\sum_{\nu}^{L}\sum_{\lambda}^{L}\sum_{\sigma}^{L} D_{\mu\nu}D_{\lambda\sigma}\iint \phi_\mu(\vec{r}_1)\,\phi_\nu(\vec{r}_1)\,\frac{1}{|\vec{r}_1-\vec{r}_2|}\,\phi_\lambda(\vec{r}_2)\,\phi_\sigma(\vec{r}_2)\,\mathrm{d}\vec{r}_1\mathrm{d}\vec{r}_2\,. \tag{2.56}$$

The computational demand can be drastically reduced by expanding the electron density $\rho(\vec{r})$ in terms of an atom-centered auxiliary basis set[115–117]

$$\rho(\vec{r}) \approx \tilde{\rho}(\vec{r}) = \sum_{\kappa}^{K} c_\kappa \eta_\kappa(\vec{r})\,. \tag{2.57}$$

The approximated auxiliary density is $\tilde{\rho}(\vec{r})$, c_κ are the extension coefficients, η_κ are the auxiliary basis functions, K is the finite number of auxiliary basis functions and κ is the summation index. The coefficients c_κ are determined by minimizing the Coulomb repulsion of the "density error" $\delta\rho(\vec{r}) = \rho(\vec{r}) - \tilde{\rho}(\vec{r})$

$$\iint \frac{\delta\rho(\vec{r}_1)\delta\rho(\vec{r}_2)}{|\vec{r}_1-\vec{r}_2|}\mathrm{d}\vec{r}_1\mathrm{d}\vec{r}_2 = \min. \tag{2.58}$$

This leads to a system of linear equations

$$\sum_{\upsilon}^{K} c_\upsilon \iint \eta_\kappa(\vec{r}_1)\,\frac{1}{|\vec{r}_1-\vec{r}_2|}\,\eta_\upsilon(\vec{r}_2)\,\mathrm{d}\vec{r}_1\mathrm{d}\vec{r}_2 = \gamma_\kappa\,. \tag{2.59}$$

with γ_κ being

$$\gamma_\kappa = \sum_{\lambda}^{L}\sum_{\sigma}^{L} D_{\lambda\sigma}\iint \eta_\kappa(\vec{r}_1)\,\frac{1}{|\vec{r}_1-\vec{r}_2|}\,\phi_\lambda(\vec{r}_2)\,\phi_\sigma(\vec{r}_2)\,\mathrm{d}\vec{r}_1\mathrm{d}\vec{r}_2\,. \tag{2.60}$$

In the first step equation (2.60) is calculated to yield the γ_κ elements. Then with γ_κ equation (2.59) can be solved to yield the coefficients c_κ. Now the Coulomb $J[\rho]$ energy (equation (2.56)) is approximated with $\tilde{J}[\rho]$

$$J[\rho] \approx \tilde{J}[\rho] = \frac{1}{2}\iint \tilde{\rho}(\vec{r}_1)\,\frac{1}{|\vec{r}_1-\vec{r}_2|}\,\rho(\vec{r}_2)\,\mathrm{d}\vec{r}_1\mathrm{d}\vec{r}_2 = \frac{1}{2}\sum_{\kappa}^{K} c_\kappa\gamma_\kappa\,. \tag{2.61}$$

Thus, four-center-two-electron integrals are approximated with three-center-two-electron integrals. The approach is abbreviated as "RI-J" or "RI-DFT" because equation (2.58) can be rewritten to formally resemble a "resolution of the identity" (RI).[115–117] The number of three-center-two-electron integrals is proportional to L^2K with L being the number of basis functions and K the number of auxiliary basis functions.[95] K is usually two to three times as large as L to have a reasonable balance between accuracy and feasibility. Compared to the number of four-center-two-electron integrals which is proportional to L^4, the number of integrals is drastically reduced. The reduction is large enough to switch from the direct SCF procedure to the "in-core" procedure. Therefore, RI-DFT is not only faster because the number of integrals is reduced, but also because the integrals only have to be solved once in the SCF procedure.[115,116] For the systems in this thesis the calculations were usually five to ten times faster with RI-DFT. The Coulomb energy in the RI-DFT approach typically deviates about $10^{-4}\,E_\mathrm{h}$ per atom from the corresponding energy from a DFT calculation without RI-J approximation.[115,116] For energy differences like binding energies these deviations are neglectable due to error cancellation.

Further reduction of the computational demand can be archived with the "multipole accelerated resolution of identity for J" (MARI-J) approach which is based on RI-DFT.[117] Coulomb interactions are divided into a "near-field" and a "far-field" part. In the near-field the Coulomb interactions are calculated with the RI-J approach. In the far-field a multipole expansion is used. The electron density $\rho(\vec{r})$ and the auxiliary density $\tilde{\rho}(\vec{r})$ are divided into atomic partitions and into atom-centered partitions, respectively. For these partitions spherical extents (= radii) are defined based on the exponents ζ of the Gaussian Type Orbitals[118] which are used as basis functions (see section 2.7). If the extent of a density partition and the one of an auxiliary density partition are overlapping, they are considered as near-field otherwise as far-field. A small exponent ζ yields a large extent and by this a larger near-field. The deviation of MARI-J to the pure RI-DFT calculation is typically below $10^{-7}\,E_\mathrm{h}$ per atom; hence, three orders of magnitude smaller than the deviation of RI-DFT to DFT.[117] For systems with about 1000 atoms MARI-J calculations are reported to be up to 6.5 times faster than the already fast pure RI-DFT calculations.[117] In this thesis with systems about 170 atoms, the MARI-J calculations were just slightly accelerated by a factor between 1.01 and 1.2 compared to the pure RI-DFT.

2.4.3 Dispersion correction DFT-D

Many density functionals are failing in describing the correct $-C_6R^{-6}$ dependence of the long-range electron correlation which is responsible for the London dispersion[119,120] interaction.[60] A non-local description of the correlation would be necessary to describe the London dispersion interactions correctly. LDA, GGA and meta-GGA functionals only use local informations of the density. Hybrid functions include the non-local Hartree-Fock exchange, however the description of the correlation is still local. To take the London dispersion into account and still use DFT (especially GGA and meta-GGA functionals) the disperion corrections DFT-D3[58–60] and DFT-D3(BJ)[61] were used which differ only in the damping function f_q^{damp} (see below). The non-additive three body term was not used, because it has only a minor effect on the energy of a molecular system.

In the DFT-D3 and the DFT-D3(BJ) approach a correction term E^{D} is added to the total energy E^{DFT} of a Kohn-Sham DFT calculation to get the corrected total energy $E^{\text{DFT-D}}$

$$E^{\text{DFT-D}} = E^{\text{DFT}} + E^{\text{D}} \,. \tag{2.62}$$

E^{D} is a semi-empirical atom-pairwise potential. For DFT-D3 it is given with

$$E^{\text{D}} = -\frac{1}{2} \sum_{A \neq B} \sum_{q=6,8} s_q \frac{C_q^{\text{AB}}}{R_{\text{AB}}^q} f_q^{\text{damp}} \tag{2.63}$$

with the so called zero-damping

$$f_q^{\text{damp}} = \frac{1}{1 + 6\left(R_{\text{AB}}(s_{\text{R},q} R_0^{\text{AB}})^{-1}\right)^{-(8+q)}} \,. \tag{2.64}$$

E^{D} for DFT-D3(BJ) is given with

$$E^{\text{D}} = -\frac{1}{2} \sum_{A \neq B} \sum_{q=6,8} s_q \frac{C_q^{\text{AB}}}{R_{\text{AB}}^n + f_q^{\text{damp}}} \tag{2.65}$$

with the Becke-Johnson-damping[61,121–123]

$$f_q^{\text{damp}} = \left(\beta_1 R_0^{\text{AB}} + \beta_2\right)^q \,. \tag{2.66}$$

R_{AB} is the distance between the atoms A and B. C_q^{AB} are the qth-order atom-pair AB depending isotropic dispersion coefficients and s_q are DFT functional depended scaling

factors. The s_6 scaling factor is usually set to unity for all functionals and only s_8 is used as empirical functional dependent parameter. C_6^{AB} is computed based on the Casimir–Polder formula[58,124]

$$C_6^{\mathrm{AB}} = \frac{3}{\pi}\int_0^\infty \alpha^{\mathrm{A}}(i\omega)\,\alpha^{\mathrm{B}}(i\omega)\,\mathrm{d}\omega \tag{2.67}$$

with $\alpha(i\omega)$ being the averaged dipole polarizability at imaginary frequency ω calculated with time-dependent DFT (TDDFT). The TDDFT calculations were done with the PBE functional mixed with 37.5% Hartree-Fock exchange for the evaluation of C_6^{AB}. Instead of free atoms different hydrides were used to account for different binding situations. The C_8^{AB} coefficients are calculated recursively from the C_6^{AB} coefficients with

$$C_8^{\mathrm{AB}} = 3C_6^{\mathrm{AB}}\sqrt{\sqrt{Z_A}Q_A\sqrt{Z_B}Q_B} \tag{2.68}$$

where Z_A and Z_B are the nuclear charges, and Q_A and Q_B are the quotients of averages of powers of the electron-nucleus distance calculated from the ground state electron density of atom A and B, respectively.[58,125]

R_0^{AB} is the cutoff radius. The cutoff radius is multiplied with the functional dependent scaling factor $s_{\mathrm{R},q}$ or β_1. The $s_{\mathrm{R},8}$ scaling factor is set to unity for all functionals. In total DFT-D3 with zero-damping has two functional dependent scaling factors s_8 and $s_{\mathrm{R},6}$. The successor, DFT-D3(BJ), has three functional dependent scaling factors s_8, β_1 and additionally β_2.

E^{D} in DFT-D3 with zero-damping approaches zero if the distance between the atoms A and B goes to zero. Therefore, atoms can experience a repulsive force for small and medium distances. With Becke-Johnson damping this problem is solved, because E^{D} approaches a constant finite value if the distance between the atoms A and B goes to zero. In this thesis mainly DFT-D3(BJ) was used. Calculating analytically vibrational frequencies with DFT-D3(BJ) yields spurious imaginary frequencies; therefore, vibrational frequencies are determined numerically with DFT-D3(BJ). DFT-D3 with zero-damping was used for analytic frequency calculations for systems which are not feasible with numerical calculations.

2.5 Thermal effects and thermodynamic properties

Until now only the electronic energy is taken into account, however, the change in the electronic energy makes only about 50% of the Gibbs energy of association for the com-

plexes in this thesis at room temperature. The other 50% are solvent effects and the "thermal effects" of translation, rotation and vibration. In this section the determination of the thermal effects in the gas phase is described. The experimental parts belonging to this study are performed on macroscopic samples considering an ensemble of molecules. To connect the results of the calculation of single molecules to an ensemble of molecules, statistical mechanics has to be used. Thermodynamic functions, macroscopic observables and their derivatives (enthalpy H, the entropy S and Gibbs energy G) for an ensemble are obtained from the partition function in statistical mechanics. The ensemble of molecules is regarded as ideal gas which allows to express the partition function in closed analytical form.[104]

The partition function q for a single molecule is the sum over the Boltzmann factors of all its quantum states n

$$q = \sum_{n}^{\infty} e^{-\epsilon_n (k_\mathrm{B} T)^{-1}}\,, \tag{2.69}$$

with the energy ϵ_n of quantum state n, the Boltzmann constant k_B and the temperature T (in this section the common symbols/characters from statistical mechanics for the variables and the International System of Units (SI) are used in contrast to the other sections). The energy ϵ_n is given with respect to the energy ϵ_0 of the lowest state, which will be chosen to be zero for the sake of convenience. Q is the corresponding partition function of an canonical ensemble of N indistinguishable non-interacting molecules at room temperature

$$Q = \frac{q^N}{N!}\,, \tag{2.70}$$

the number of molecules N, the volume V and the temperature T are constants. In the canonical ensemble the enthalpy H, the entropy S and the Gibbs energy G can be calculated from the partition function Q with

$$H = U + PV = k_\mathrm{B} T^2 \left(\frac{\partial \ln Q}{\partial T}\right)_{N,V} + k_\mathrm{B} T V \left(\frac{\partial \ln Q}{\partial V}\right)_{N,T}\,, \tag{2.71}$$

$$S = k_\mathrm{B} T \left(\frac{\partial \ln Q}{\partial T}\right)_{N,V} + k_\mathrm{B} \ln Q\,, \tag{2.72}$$

$$G = H - TS = k_\mathrm{B} T V \left(\frac{\partial \ln Q}{\partial V}\right)_{N,T} - k_\mathrm{B} T \ln Q\,. \tag{2.73}$$

U is the internal energy and P the pressure.

To calculate the partition function Q, and thus the enthalpy H, the entropy S and the Gibbs energy G, further approximations have to be made, assuming first the Hamiltonian $\hat{H}_{\text{tot}}$ characterizing the total energy ϵ_{tot} of the quantum state of a single molecule is the sum of the translational $\hat{H}_{\text{trans}}$, rotational $\hat{H}_{\text{rot}}$, vibrational $\hat{H}_{\text{vib}}$ and electronic $\hat{H}_{\text{e}}$ Hamiltonians[126]

$$\hat{H}_{\text{tot}} = \hat{H}_{\text{trans}} + \hat{H}_{\text{rot}} + \hat{H}_{\text{vib}} + \hat{H}_{\text{e}} \,. \tag{2.74}$$

The total wave function Φ_{tot} for a single molecule is then a product of the translational Φ_{trans}, rotational Φ_{rot}, vibrational Φ_{vib} and electronic Φ_{e} wave function

$$\Phi_{\text{tot}} = \Phi_{\text{trans}} \cdot \Phi_{\text{rot}} \cdot \Phi_{\text{vib}} \cdot \Phi_{\text{e}} \tag{2.75}$$

and the total energy ϵ_{tot} the sum over the corresponding individual energies

$$\epsilon_{\text{tot}} = \epsilon_{\text{trans}} + \epsilon_{\text{rot}} + \epsilon_{\text{vib}} + \epsilon_{\text{e}} \,. \tag{2.76}$$

It follows from equation (2.76) and (2.69) that the total partition function q_{tot} of a signal molecule can be calculated as a product of the individual translational q_{trans}, rotational q_{rot}, vibrational q_{vib} and electronic q_{elec} partition functions

$$q_{\text{tot}} = q_{\text{trans}} \cdot q_{\text{rot}} \cdot q_{\text{vib}} \cdot q_{\text{elec}} \,. \tag{2.77}$$

At atmospheric pressure and room temperature the translational wave function Φ_{trans} of an individual molecule is delocalized over the the whole volume V of the ensemble, whereas the rotational Φ_{rot}, vibrational Φ_{vib} and electronic Φ_{e} wave functions are localized at the corresponding molecule not overlapping with the corresponding wave functions of the other molecules. This absence of overlap makes the molecules distinguishable if only the rotational, vibrational and electronic states are considered.[127] Thus, indistinguishability of the molecules only effects the translational partition function Q_{trans} of the ensemble; therefore, the $\frac{1}{N!}$ term correcting the over-counting of states due to the indistinguishability is included only in Q_{trans}. It follows for the partition function Q of the ensemble from equation (2.70)

$$Q = \frac{q_{\text{tot}}^N}{N!} = \underbrace{\frac{q_{\text{trans}}^N}{N!}}_{Q_{\text{trans}}} \cdot \underbrace{q_{\text{rot}}^N}_{Q_{\text{rot}}} \cdot \underbrace{q_{\text{vib}}^N}_{Q_{\text{vib}}} \cdot \underbrace{q_{\text{elec}}^N}_{Q_{\text{elec}}} \cdot \tag{2.78}$$

Applying the natural logarithm leads to

$$\ln(Q) = N\ln(q_{\text{trans}}) - \ln(N!) + N\ln(q_{\text{rot}}) + N\ln(q_{\text{vib}}) + N\ln(q_{\text{elec}}) \,. \tag{2.79}$$

This can be simplified in case of a large N, with Stirling's formula truncated after the second term $\left(\ln(N!) \approx N\ln(N) - N = N\ln\left(N\frac{1}{\exp(1)}\right)\right)$, to

$$\ln(Q) \approx \underbrace{N\ln\left(\frac{q_{\text{trans}}}{N}\exp(1)\right)}_{\ln(Q_{\text{trans}})} + \underbrace{N\ln(q_{\text{rot}})}_{\ln(Q_{\text{rot}})} + \underbrace{N\ln(q_{\text{vib}})}_{\ln(Q_{\text{vib}})} + \underbrace{N\ln(q_{\text{elec}})}_{\ln(Q_{\text{elec}})} \,. \tag{2.80}$$

The total enthalpy H_{tot} (equation (2.71)) and entropy S_{tot} (equation (2.72)) are calculated from $\ln(Q)$ (equation (2.80)); therefore, they can be separated also into the individual contributions

$$\begin{aligned} H_{\text{tot}}(\ln(Q)) = {} & H_{\text{trans}}(\ln(Q_{\text{trans}})) + H_{\text{rot}}(\ln(Q_{\text{rot}})) + H_{\text{vib}}(\ln(Q_{\text{vib}})) \\ & + H_{\text{elec}}(\ln(Q_{\text{elec}})) \,, \end{aligned} \tag{2.81}$$

$$\begin{aligned} S_{\text{tot}}(\ln(Q)) = {} & S_{\text{trans}}(\ln(Q_{\text{trans}})) + S_{\text{rot}}(\ln(Q_{\text{rot}})) + S_{\text{vib}}(\ln(Q_{\text{vib}})) \\ & + S_{\text{elec}}(\ln(Q_{\text{elec}})) \,. \end{aligned} \tag{2.82}$$

2.5.1 Translation

The energy for the quantum states of translation can be determined from the "particle in a box" model with infinitely large potential walls. In the first step the translation in only one dimension is assumed, leading to the translational energy levels ϵ_{n_x},

$$\epsilon_{n_x} = \frac{n_x^2 h^2}{8mL_x^2} \tag{2.83}$$

with the mass m of the molecule under investigation, h the Planck constant and the length L_x in x-direction. At 298.15 K the difference between the Boltzmann factors is getting so small that the summation in equation (2.69) can be replaced by an integral leading to the partition function $q_{\text{trans},x}$ for the translation along x

$$q_{\text{trans},x} = \sum_{n_x=0}^{\infty} e^{-\epsilon_{n_x}(k_{\text{B}}T)^{-1}} = \int_0^{\infty} e^{-\epsilon_{n_x}(k_{\text{B}}T)^{-1}} \mathrm{d}n_x = \sqrt{\frac{2\pi m k_{\text{B}}T}{h^2}} L_x \,. \tag{2.84}$$

The lowest state in the "particle in a box" model is characterized with $n_x = 1$, nevertheless, $n_x = 0$ with $\epsilon_0 = 0$ is used in this thesis. It does not change the partition function because the energy difference is small but it allows to solve equation (2.84) in a closed analytic form. The total translational partition function q_{trans} for a single molecule is the product of the translational partition functions in the individual directions

$$q_{\text{trans}} = q_{\text{trans},x} q_{\text{trans},y} q_{\text{trans},z} = \left(\frac{2\pi m k_{\text{B}} T}{h^2}\right)^{\frac{3}{2}} V \tag{2.85}$$

with the volume $V = L_x L_y L_z$. Inserting q_{trans} in equations (2.80), (2.71) and (2.72) yields for $N = 1\,\text{mole} = N_{\text{A}}$ (N_{A}: Avogadro constant) and volume $V = V_{\text{m}}$ (V_{m}: molar volume) the translational H_{trans} and S_{trans}

$$H_{\text{trans}} = \frac{5}{2} RT\,, \tag{2.86}$$

$$S_{\text{trans}} = \frac{5}{2} R + R \ln\left(\frac{V_{\text{m}}}{N_{\text{A}}}\left(\frac{2\pi m k_{\text{B}} T}{h^2}\right)^{\frac{3}{2}}\right)\,, \tag{2.87}$$

where $R = k_{\text{B}} N_{\text{A}}$ is the ideal gas constant.

2.5.2 Rotation

The geometric structure of the molecules is assumed not to change or to stretch due to rotation. Hence the geometric structure stays fixed in the energetic minimum calculated within the Born–Oppenheimer approximation. This approximation is called "rigid-rotor" (RR) approximation. Even within the rigid-rotor approximation the solution for a non-linear polyatomic molecule is non-trival. However, every molecule (except hydrogen) which is stable at room temperature and atomospheric pressure has such larges moments of inertia (which leads to very small differences between the rotational energy levels) that at room temperature the rotations can be treated accurately with classical mechanics,[126] which yields the following partition function

$$q_{\text{rot}} = \frac{\sqrt{\pi}}{\sigma}\left(\frac{8\pi^2 k_{\text{B}} T}{h^2}\right)^{\frac{3}{2}} \sqrt{I_1 I_2 I_3}\,. \tag{2.88}$$

I_1, I_2 and I_3 are the moments of inertia and can be calculated from the optimized geometric structure. The number of rotations which transfer the molecule into itself yields the

symmetry index σ which corrects for the overcounting of rotational states. This is a consequence of Pauli's exclusion principle. The total wave function has to be symmetric with respect to the exchange of two identical nuclei with zero or integer spins (bosons) and antisymmetric if the nuclei have half integer spins (fermions). To rationalize this, the nuclear spin wave function Φ_{ns} has to be included in equation (2.75). It can be seen that only rotational states are accessible which have a symmetry fitting with the symmetry of the nuclear spin states. Detailed discussions on this are made by Mayer *et al.*[126] and Gilson *et al.*[128] Other than that, the nuclear spin wave function Φ_{ns} does not have to be taken into account, because the effects will be canceled out in chemical reactions.[126]

Equation (2.88) yields for $N = N_{\mathrm{A}}$ to

$$H_{\mathrm{rot}} = \frac{3}{2} RT\,, \tag{2.89}$$

$$S_{\mathrm{rot}} = \frac{3}{2} R + R \ln \left(\frac{\sqrt{\pi}}{\sigma} \left(\frac{8\pi^2 k_{\mathrm{B}} T}{h^2} \right)^{\frac{3}{2}} \sqrt{I_1 I_2 I_3} \right)\,. \tag{2.90}$$

2.5.3 Vibration

If the geometric structure of the molecule is at a stationary point, meaning the first derivatives of the electronic energy with respect to the nuclear coordinates are zero, then the vibrational modes of the molecule can be calculated with a harmonic oscillator (HO) approximation. In this approximation the nuclear Schrödinger equation is separable into $3N_{\mathrm{atom}}$ one-dimensional Schrödinger equations, with N_{atom} being the number of atoms in the molecule. The one-dimensional Schrödinger equations have the form of one-dimensional harmonic oscillators. The corresponding eigenvalues can be gained from the Hessian matrix, which contains all second derivatives of the energy $E\left(\left\{\vec{R}\right\}\right)$ (equation (2.8)) with respect to the coordinates.[104] The Hessian matrix has to be transformed to mass-dependent coordinates and then diagonalized. This yields $3N_{\mathrm{atom}}$ eigenvalues λ_i. Six of the eigenvalues are zero because they correspond to the three translations and the three rotations of the rigid molecule, respectively. The other eigenvalues λ_i yield the frequencies ν_i for the $3N_{\mathrm{atom}} - 6$ vibrational modes

$$\nu_i = \frac{1}{2\pi} \sqrt{\lambda_i}\,. \tag{2.91}$$

The corresponding energy levels $\epsilon_{n,i}$ are

$$\epsilon_{n,i} = \left(n + \frac{1}{2}\right) h\nu_i \quad \text{with} \quad n = 0, 1, 2, \ldots \infty \,. \tag{2.92}$$

The Planck constant h is used. If the temperature T is 0 K, only the ground-state $n = 0$ is occupied. In contrast to the translation and rotation, the vibrational ground-state energy E_{ZPVE} is different from zero

$$E_{\mathrm{ZPVE}} = \sum_{i=1}^{3N_{\mathrm{atom}}-6} \frac{1}{2} h\nu_i \,. \tag{2.93}$$

The vibrational ground-state energy E_{ZPVE} will be subtracted from the vibrational energy levels $\epsilon_{n,i}$ (equation (2.92)) and included as a separate term in the vibrational enthalpy H_{vib} (equation (2.97)). The new vibrational energy levels $\epsilon'_{n,i}$ are

$$\epsilon'_{n,i} = nh\nu_i \quad \text{with} \quad n = 0, 1, 2, \ldots \infty \,. \tag{2.94}$$

The spacing between the energy levels $\epsilon'_{n,i}$ is too large to replace the summation for $q_{\mathrm{vib},i}$ with an integral. Nevertheless, the summation yields a closed form, because the spacing is regular

$$q_{\mathrm{vib},i} = \sum_{n=0}^{\infty} e^{-\epsilon'_{n,i}(k_{\mathrm{B}}T)^{-1}} = 1 + e^{-h\nu_i(k_{\mathrm{B}}T)^{-1}} + e^{-2h\nu_i(k_{\mathrm{B}}T)^{-1}} + \cdots = \frac{1}{1 - e^{-h\nu_i(k_{\mathrm{B}}T)^{-1}}} \,. \tag{2.95}$$

The partition function q_{vib} of all vibrational modes derives from the partition function $q_{\mathrm{vib},i}$ of a single vibrational mode i (equation (2.95))

$$q_{\mathrm{vib}} = \prod_{i=1}^{3N_{\mathrm{atom}}-6} q_{\mathrm{vib},i} \,. \tag{2.96}$$

For $N = N_{\mathrm{A}}$, this yields to

$$H_{\mathrm{vib}} = R \sum_{i=1}^{3N_{\mathrm{atom}}-6} \left(\frac{h\nu_i}{2k_{\mathrm{B}}} + \frac{h\nu_i}{k_{\mathrm{B}}\left(e^{h\nu_i(k_{\mathrm{B}}T)^{-1}} - 1\right)} \right) , \tag{2.97}$$

$$S_{\mathrm{vib}} = R \sum_{i=1}^{3N_{\mathrm{atom}}-6} \left(\frac{h\nu_i}{k_{\mathrm{B}}T\left(e^{h\nu_i(k_{\mathrm{B}}T)^{-1}} - 1\right)} - \ln\left(1 - e^{-h\nu_i(k_{\mathrm{B}}T)^{-1}}\right) \right) . \tag{2.98}$$

The first addend in equation (2.97) is the zero-point vibrational energy E_{ZPVE}.

The approach described so far in this section is called “rigid-rotor-harmonic-oscillator”

(RRHO) approximation, although it contains some more approximations. The RRHO approximation performs reasonably well with one exception. Figure 2.1 shows the rotation of a methyl group in toluene. It is not a rotation of the whole molecule, it is an internal rotation. More precisely, it is a hindered internal rotation which has to overcome an energy barrier for full rotation. This kind of rotations are treated as vibrations with the harmonic oscillator (HO) approximation. As long as the barrier is high (which results in high frequencies) the harmonic oscillator approximation gives a reasonable description of the partition function of the hindered internal rotation. If the barrier becomes vanishingly small (which results in low frequencies) the partition function determined from the harmonic oscillator approximation goes to infinity, whereas the actual partition function of the hindered internal rotation converges to the partition function of a free rotator as shown in Figure 2.2.[129,130] Thus, for molecules with low-lying vibrations the RRHO approximation usually overestimates the entropy. This causes an overestimation of the binding energies of organic molecules, because a large product molecule usually has more hindered internal rotations than its smaller educts.

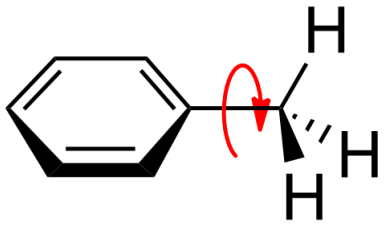

Figure 2.1: Internal rotation of a methyl group in toluene.

A cosine potential can be used to describe the hindered internal rotation in case it is a rotation along an isolated single bond like shown in Figure 2.1. Usually the hindered internal rotations are not isolated (for example in alkans). In this cases a simple description of the potential of a hindered internal rotation can not be given, because an accurate description of the potential needs to recognize that internal rotations are coupled.[130,131] The first attempts of theoretical investigations of hindered internal rotations were started in 1932 by Nielsen,[132] resulting in a series of publications from Pitzer[131,133,134] in the 1940s delivering accurate descriptions for the partition function of hindered internal rotations. Pitzer's work was adapted in 1997

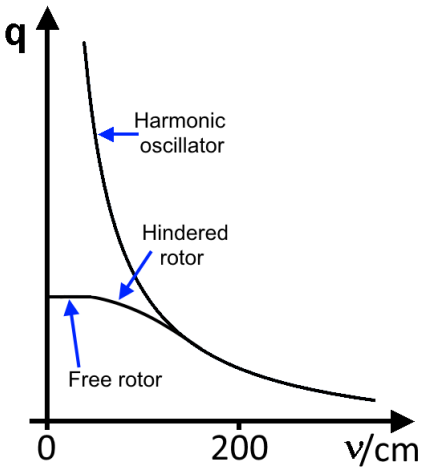

Figure 2.2: Partition function **q** for a harmonic oscillator, a hindered rotor, and a free rotor plotted against the vibration frequency ν.[129]

by East and Radom[130] to gain accurate entropies for molecules with hindered internal rotations with modern computational chemistry. East's and Radom's approach include an *ab inito* calculation of the multidimensional torsional potential (of the corresponding molecule) and the expression of this potential in terms of Fourier series. The partition function for the hindered internal rotations can then be calculated in high accuracy according to Pitzer's[131] proposal involving the numerical solution of multidimensional integrals built from the multidimensional torsional potential. Unfortunately this method is not feasible for large molecules. Therefore, more drastic approximations have to be applied. The easiest solution to account for the hindered internal rotations would be to treat all low-lying vibrational modes as independent free rotors with the rigid-rotor approximation and to empirically interpolate subsequently between this new approximation and the harmonic oscillator approximation. Truhlar[135] published such an approach in 1991 and Grimme[61] a similar one in 2012. The approach from Grimme is used in this thesis, but only for vibrational modes with a frequency below 500 cm^{-1} since this correction makes no sense for high-lying vibrational modes (Figure 2.2). First the harmonic oscillator approximation is used to calculate the frequencies ν_i of all "vibrational" modes. For the vibrational modes with frequencies below 500 cm^{-1} the moment of inertia μ_i is calculated, which corresponds to the moment of inertia for a rigid-rotor having the same frequency ν_i as the vibrational mode

$$h\nu_i = \frac{h^2}{8\pi^2\mu_i} \quad \Rightarrow \quad \mu_i = \frac{h}{8\pi^2\nu_i}\,. \tag{2.99}$$

To prevent too large μ_i values resulting from too small frequencies ν_i, the damped moment of inertia μ_i' is used

$$\mu_i' = \frac{\mu_i B}{\mu_i + B} \quad \text{with} \quad B = 10^{-44}\,\mathrm{kg\,m^2}\,. \tag{2.100}$$

The limiting value B in this thesis was always set to $10^{-44}\,\mathrm{kg\,m^2}$. The damping introduced due to the limiting value B has only noticeable effects if the frequency ν_i gets below 1 cm^{-1}, which did not occur in this thesis. The damped moment of inertia μ_i' is used to calculate the rotational entropy $S_{\text{int-rot},\nu_i}$

$$S_{\text{int-rot},\nu_i} = \frac{1}{2}R + R\ln\left(\left(\frac{8\pi^3\mu_i' k_\mathrm{B} T}{h^2}\right)^{\frac{1}{2}}\right)\,. \tag{2.101}$$

A damping function $f\,(\nu_i)$ from Head-Gordon[136] is used to interpolate between rotational

S_{rot,ν_i} and harmonic vibrational S_{vib,ν_i} (equation (2.98)) approximation

$$S_{\nu_i} = f(\nu_i)\, S_{\mathrm{vib},\nu_i} + (1 - f(\nu_i))\, S_{\text{int-rot},\nu_i} \quad \text{with} \quad f(\nu_i) = \frac{1}{1 + \left(\frac{\nu_0}{\nu_i}\right)^4}\,. \tag{2.102}$$

The parameter ν_0 is set to 100 cm^{-1}.

This correction for hindered rotations will be called “free-rotor approximation” in this thesis.

2.5.4 Electronic partition function

The energy difference between the electronic ground state and the excited states of the molecules in this thesis is large. Therefore, q_{elec} is set to one, S_{elec} to zero, and for H_{elec} the energy $E\left(\left\{\vec{R}\right\}\right)$ (equation (2.8)) is used to get the total values for equation (2.77), (2.81) and (2.82).

2.6 Solvent effects

Next to the electronic and thermal effects the solvent effects have to be considered. Therefore, the COSMO-RS[62–65] (Conductor-like Screening Model for Real Solvents) is employed, which is based on the dielectric continuum solvation model (DCM) COSMO[66,67] (Conductor-like Screening Model).

2.6.1 COSMO

In DCMs the investigated solute molecule forms a cavity in a dielectric continuum. This cavity yields the solvent-accessible surface area (SASA) and is usually constructed from the about 20% increased van der Waals radii of the atoms of the solute molecule. The SASA serves as an interface between solute molecule and dielectric continuum. The electric charge distribution in the solute molecule polarizes the dielectric continuum creating screening charges at the SASA. The screening charges from the dielectric continuum at the SASA affect subsequently the solute molecule. Solvents have a certain finite permittivity ε (or relative permittivity ε_r). In DCMs the solvents are simulated by modeling a dielectric continuum with the same permittivity ε. The basis for this approach is the Poisson equation, which usually has to be solved iteratively for a molecular shaped surface and a finite permittivity ε. COSMO is able to solve the problem non-iteratively by

describing the solvent as an ideal conductor ($\varepsilon_\mathrm{r} = \infty$). The total potential $\mathbf{\Phi}^\mathrm{tot}$ at the SASA becomes zero for $\varepsilon_\mathrm{r} = \infty$

$$\mathbf{\Phi}^\mathrm{tot} = \mathbf{\Phi}^\mathrm{X} + \mathbf{A}_X\mathbf{q} = 0\,. \tag{2.103}$$

The total potential $\mathbf{\Phi}^\mathrm{tot}$ is the sum of the electrostatic potential $\mathbf{\Phi}^\mathrm{X}$ of the molecule X in solvent S and the potential $\mathbf{A}_X\mathbf{q}$ arising from the screening charges. $\mathbf{\Phi}^\mathrm{X}$ derives from the atomic nuclei and electron density of the molecule X (determined in an electronic structure calculation). $\mathbf{A}_X\mathbf{q}$ consists of the matrix of the screening charge interactions $\mathbf{A}_X$ at solute molecule X and the vector of the screening charges $\mathbf{q}$ consisting of the scalar components q_i. The SASA is divided into segments. Each segment i has the area s_i and the screening charge density σ_i. Thus, the scalar components q_i can be expressed as

$$q_i = s_i\sigma_i\,. \tag{2.104}$$

Using the boundary equation (2.103), the screening charges $\mathbf{q}$ are derived with

$$\mathbf{q} = -\mathbf{A}_X^{-1}\mathbf{\Phi}^\mathrm{X}\,. \tag{2.105}$$

In an ideal conductor the surface charges $\mathbf{q}^\mathrm{solute}$ from the solute molecule are exactly nullified by the screening charges $\mathbf{q}$

$$\mathbf{q}^\mathrm{solute} = -\mathbf{q}\,. \tag{2.106}$$

The interaction energy (or "COMSO energy") E_S^X of a solute molecule X with the solvent S is the product of the potential $\mathbf{A}_X\mathbf{q}$ arising from the screening charges and the surface charges $\mathbf{q}^\mathrm{solute}$

$$E_S^X = \frac{1}{2}\mathbf{q}^\mathrm{solute}\mathbf{A}_X\mathbf{q}\,. \tag{2.107}$$

The factor "$\frac{1}{2}$" appears instead of "1" because the polarization of the dielectric continuum consumes half of the energy which is gained. The surface charges $\mathbf{q}^\mathrm{solute}$ can be replaced with the screening charges $\mathbf{q}$ (equation (2.106)) and the screening charges $\mathbf{q}$ are written as sum over the screening charges q_i on the surface segments i[137]

$$E_S^X = -\frac{1}{2}\sum_{i\in X} q_i \sum_{j\in X} A_{ij}^X q_j\,, \tag{2.108}$$

with equation (2.104) yielding

$$E_S^X = -\frac{1}{2}\sum_{i\in X}\sigma_i s_i \sum_{j\in X} A_{ij}^X \sigma_j s_j \,. \tag{2.109}$$

For actual solvents the correction factor $f(\varepsilon_\mathrm{r}) = (\varepsilon_\mathrm{r}-1)(\varepsilon_\mathrm{r}+0.5)^{-1}$ is used to correct for the finite permittivity of the solvents

$$\mathbf{q}^* = f(\varepsilon_\mathrm{r})\,\mathbf{q}\,, \tag{2.110}$$

$$E_S^X = -\frac{1}{2} f(\varepsilon_\mathrm{r}) \sum_{i\in X}\sigma_i s_i \sum_{j\in X} A_{ij}^X \sigma_j s_j = -\frac{1}{2}\frac{\varepsilon_\mathrm{r}-1}{\varepsilon_\mathrm{r}+\frac{1}{2}} \sum_{i\in X}\sigma_i s_i \sum_{j\in X} A_{ij}^X \sigma_j s_j \,. \tag{2.111}$$

COMSO is integrated into the SCF procedure of Hartree-Fock and DFT as an additive external potential V_S^X which is the derivative of E_S^X with respect to the electron density $\rho^X(\vec{r})$ of the solute molecule

$$V_S^X = \frac{\partial E_S^X}{\partial \rho^X(\vec{r})}\,. \tag{2.112}$$

2.6.2 COSMO-RS

DCMs do not take short-range interactions into account like hydrogen bonds and van der Waals interactions. COSMO-RS aims to overcome these just mentioned drawbacks as in this approach (Figure 2.3) the interaction between solute molecule and solvent is calculated based on pairwise interactions of surface segments. Each segment is characterized by its screening charge density σ (Figure 2.4 shows the charge distribution of typical solute and solvent molecules investigated in this thesis) and the atom type at this position.

From each species in solution, either solute or solvent, an electronic structure calculation (with COSMO) of one molecule in an ideal conductor ($\varepsilon_\mathrm{r} = \infty$) is performed. This yields the screening charge densities σ at the surface of the molecule. Subsequently, the surface is divided into segments. The gained information is collected for each molecule in a probability distribution $p_X(\sigma)$ of the appearance of specific screening charge densities σ at the surface of the molecule, called σ-profile. The σ-profile $p_S(\sigma)$ of the whole solution is the sum of individual σ-profiles of the involved molecules weighted by their corresponding mole fraction x_X

$$p_S(\sigma) = \sum_{X\in S} x_X p_X(\sigma)\,. \tag{2.113}$$

In this thesis the solute molecules are considered in infinite dilution; thus, only the mole

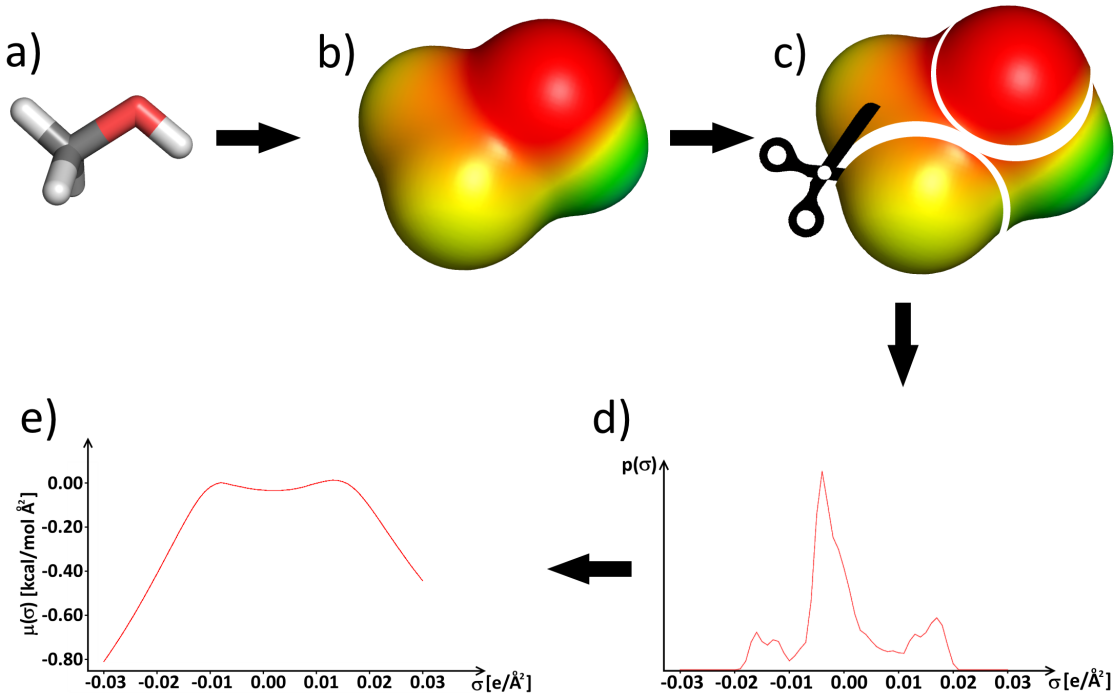

Figure 2.3: The COSMO-RS approach: a) A molecule is calculated in an ideal conductor; b) this yields the screening charges at the surface of the molecule; c) in a next step, the surface is divided into small segments; d) a probability distribution $p_X(\sigma)$ of the appearance of specific screening charge densities σ at the surface of the molecule is created, called σ-profile; e) the σ-potential is calculated from the σ-profile. The steps a) and b) are performed within the COSMO model[66,67] and c), d) and e) are performed within the COSMO-RS model.[62–65]

fractions of the solvent molecules are non-zero.

In the next step the chemical potential $\mu_S(\sigma)$ (called σ-potential) of a surface segment with the specific screening charge densities σ in the solvent S with the σ-profile $p_S(\sigma)$ is calculated. The σ-potential is a measure for the affinity of the solvent S to a surface segment with the screening charge density σ. It is assumed, that every surface segment is always in contact with another surface segment, but only pairwise interactions will be allowed. This means also the surface segments of the solvent are in contact with other surface segments of the solvent. Steric or geometrical aspects are neglected. The σ-potential $\mu_S(\sigma)$ can be derived from the partition function Q of this ensemble of surface

segments and the number of segments $N(\sigma)$

$$\mu_S(\sigma) = -k_B T \left(\frac{\delta \ln Q}{\delta N(\sigma)} \right)_{T,V} . \tag{2.114}$$

From equation (2.114) it follows for large number of segments $N(\sigma)$ that the sum of the chemical potentials $\mu_S(\sigma_k)$ and $\mu_S(\sigma_l)$ of the surface segments k and l can be expressed with the partition function Q and the partition function $Q_{-(\sigma_k,\sigma_l)}$ in which the two surface segments k and l are removed

$$\mu_S(\sigma_k) + \mu_S(\sigma_l) = -k_B T \left(\ln Q - \ln Q_{-(\sigma_k,\sigma_l)} \right) . \tag{2.115}$$

Thus, $Q_{-(\sigma_k,\sigma_l)}$ can be expressed as

$$Q_{-(\sigma_k,\sigma_l)} = \exp\left(\frac{\mu_S(\sigma_k) + \mu_S(\sigma_l)}{k_B T} \right) Q . \tag{2.116}$$

The partition function Q can also be written as

$$Q = N(\sigma) \sum_{\substack{l=1 \\ l \neq k}}^{N(\sigma)} \exp\left(\frac{-E_{\text{int}}(\sigma_k, \sigma_l)}{k_B T} \right) Q_{-(\sigma_k,\sigma_l)} , \tag{2.117}$$

where $N(\sigma)$ in the front accounts for the additional degeneracy in Q. $E_{\text{int}}(\sigma_k, \sigma_l)$ is the interaction energy of the the two surface segments k and l (see below). Putting the equation (2.116) in equation (2.117) yields

$$Q = N(\sigma) \sum_{\substack{l=1 \\ l \neq k}}^{N(\sigma)} \exp\left(\frac{-E_{\text{int}}(\sigma_k, \sigma_l) + \mu_S(\sigma_k) + \mu_S(\sigma_l)}{k_B T} \right) Q . \tag{2.118}$$

The next steps involve dividing by "$Q \exp\left(\frac{\mu_S(\sigma_k)}{k_B T}\right)$", switching to natural logarithm, replacing the sum with an integral, inculding the σ-profile $p_S(\sigma)$ and switching σ_k and σ_l to σ and σ'. This finally yields[138]

$$\mu_S(\sigma) = -\frac{RT}{a_{\text{eff}}} \ln \left[\int p_S(\sigma') \exp\left\{ -\frac{a_{\text{eff}}}{RT} \left[E_{\text{int}}(\sigma, \sigma') - \mu_S(\sigma') \right] \right\} \mathrm{d}\sigma' \right] \tag{2.119}$$

with the effective contact area a_{eff} of two surface segments. The σ-profile $p_S(\sigma)$ of the solvent S in equation (2.119) is normalized to sum up to 1 mol of surface segments. The

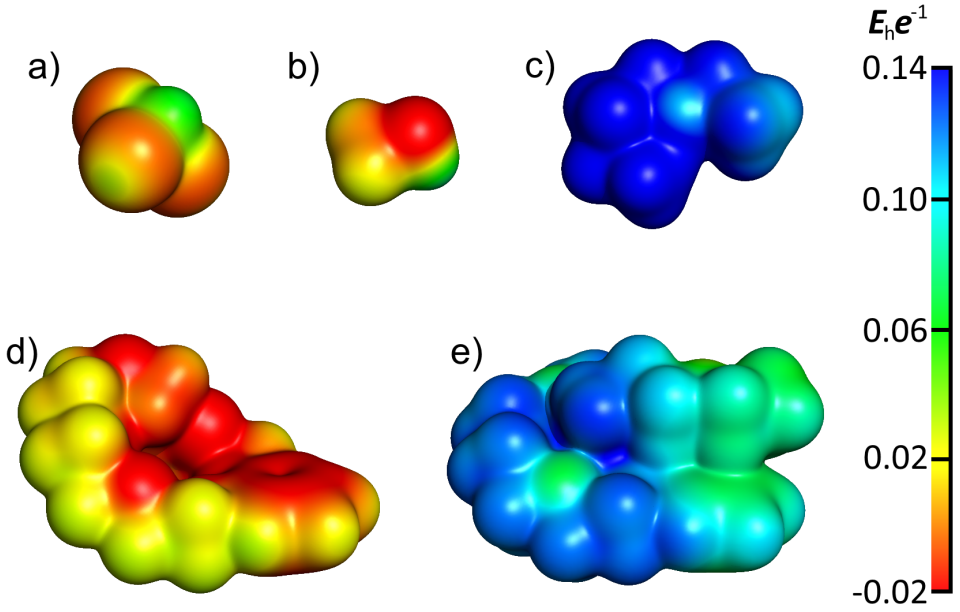

Figure 2.4: Electrostatic potential maps (which indicate the surface charges) for the solvent molecules chloroform (a) and methanol (b), the monovalent educts 3-ethoxypropylammonium (c) and benzo-18-crown-6 (d) and their product (e). The surface charges are indicated from negative to positive by red over yellow, green and cyan to blue. The screening charges in an ideal conductor are identical to the surface charges, but they have the opposite sign.

equation (2.119) has to be solved iteratively, because μ_S is on both sides of the equation.

The chemical potential μ_X^S of solute molecule X in solvent S is calculated by integrating the σ-potential $\mu_S(\sigma)$ of solvent S over the surface of solute molecule X. The σ-profile $p_X(\sigma)$ is normalized to sum up to exactly 1.

$$\mu_X^S = \int p_X(\sigma)\,\mu_S(\sigma)\,\mathrm{d}\sigma + \mu_X^{C,S}\,. \tag{2.120}$$

The empirical optimized, so called "combinatorial contribution" term $\mu_X^{C,S}$ accounts for effects based on differences between the volume and the size of the surface of the solute molecule compared to the solvent molecules.

The surface and the screening charges on the surface can significantly differ for various conformers of a molecule. This is especially the case for ion pairs and for molecules which conformers have varying numbers of intramolecular hydrogen bonds. It is necessary to include all the important conformers j of a molecule X to get an accurate description in

every solvent. Therefore, the σ-profile (equation (2.113)) of molecule X is formed by the σ-profile of the conformers j weighted by their corresponding mole fraction $x_{X,j}^{S}$ in solvent solvent S. The mole fraction $x_{X,j}^{S}$ of conformer j in the specific solvent S is calculated assuming a Boltzmann distribution

$$x_{X,j}^{S} = \frac{\exp\left[\left(-E_{\mathrm{COSMO}}^{X,j} + \mu_{X,j}^{S}\right)(k_{\mathrm{B}}T)^{-1}\right]}{\sum_j \exp\left[\left(-E_{\mathrm{COSMO}}^{X,j} + \mu_{X,j}^{S}\right)(k_{\mathrm{B}}T)^{-1}\right]} \,. \tag{2.121}$$

$E_{\mathrm{COSMO}}^{X,j}$ are the quantum chemically determined total energies of the conformer j in an ideal conductor calculated with COSMO, and $\mu_{X,j}^{S}$ is the σ-potential of the conformer j in solvent S.

The interaction energy $E_{\mathrm{int}}(\sigma,\sigma')$ in equation (2.119) consists of the "misfit" energy $E_{\mathrm{MF}}(\sigma,\sigma')$, the energy term for hydrogen bonds $E_{\mathrm{HB}}(\sigma,\sigma')$ and the energy term for van der Waals interactions $E_{\mathrm{vdW}}(\sigma,\sigma')$

$$E_{\mathrm{int}}(\sigma,\sigma') = \frac{E_{\mathrm{MF}}(\sigma,\sigma') + E_{\mathrm{HB}}(\sigma,\sigma') + E_{\mathrm{vdW}}(\sigma,\sigma')}{a_{\mathrm{eff}}} \,, \tag{2.122}$$

$$E_{\mathrm{MF}}(\sigma,\sigma') = a_{\mathrm{eff}}\frac{\alpha'}{2}(\sigma+\sigma')^2 \,, \tag{2.123}$$

$$E_{\mathrm{HB}}(\sigma,\sigma') = a_{\mathrm{eff}}c_{\mathrm{HB}}\min\left[0;\min\left(0;\sigma_{\mathrm{donor}}+\sigma_{\mathrm{HB}}\right)\max\left(0;\sigma_{\mathrm{acceptor}}-\sigma_{\mathrm{HB}}\right)\right] \,, \tag{2.124}$$

$$E_{\mathrm{vdW}}(\sigma,\sigma') = a_{\mathrm{eff}}\left(\tau_{\mathrm{vdW}}+\tau'_{\mathrm{vdW}}\right) \,. \tag{2.125}$$

The misfit energy $E_{\mathrm{MF}}(\sigma,\sigma')$ with the adjustable parameter α' arises from the electrostatic interaction of two surface segments. The COSMO-RS approach starts from molecules in an ideal conductor, in which the charges are perfectly screened by an equal charge with opposite sign. Thus, the misfit energy $E_{\mathrm{MF}}(\sigma,\sigma')$ describes the deviation from this ideal situation. For $\sigma = -\sigma'$ the misfit energy $E_{\mathrm{MF}}(\sigma,\sigma')$ vanishes, because this fulfills the situation in the ideal conductor. The energy term for hydrogen bonds $E_{\mathrm{HB}}(\sigma,\sigma')$ contains the adjustable parameters c_{HB} and σ_{HB}. $E_{\mathrm{HB}}(\sigma,\sigma')$ is atom type dependent, the donor atom for the hydrogen bond is always a hydrogen atom in COSMO-RS. Therefore, σ_{donor} and $\sigma_{\mathrm{acceptor}}$ are written instead of σ or σ', respectively. The energy for the hydrogen bond $E_{\mathrm{HB}}(\sigma,\sigma')$ will be only different from zero if the screening charge σ_{donor} at the hydrogen atom is more negative than σ_{HB} and the screening charge $\sigma_{\mathrm{acceptor}}$ at the acceptor atom is more positive than σ_{HB}. The energy term for van der Waals interactions $E_{\mathrm{vdW}}(\sigma,\sigma')$ does not use the screening charge at all and instead only uses the atom type specific adjustable

parameters τ_{vdW} and τ'_{vdW}, which are empirical optimized and tabulated.[64] Thus, the equations (2.113), (2.119) and (2.120) should also include the atom type dependency. This is not done in the presented notion for the sake of clarity.

The chemical potential μ_X^{Gas} of molecule X in the gas phase is derived from

$$\mu_X^{\mathrm{Gas}} = E_{\mathrm{Gas}}^X - E_{\mathrm{COSMO}}^X - \omega_{\mathrm{ring}} n_{\mathrm{ring}}^X + \eta_{\mathrm{Gas}} \,. \tag{2.126}$$

E_{Gas} and E_{COSMO}^X are the total energies of the molecule X in the gas phase and in an ideal conductor calculated with electronic structure methods. In general the correction term η_{Gas} is added which accounts for the errors in the electronic structure method. For ring shaped molecules the correction term $\omega_{\mathrm{ring}} n_{\mathrm{ring}}^X$ is used, too.

From μ_X^S and μ_X^{Gas} the Gibbs energy $G_{\mathrm{Solv},X}^S$ of the solvation of molecule X in solvent S can be calculated with

$$G_{\mathrm{Solv},X}^S = \left(\mu_X^S - \mu_X^{\mathrm{Gas}}\right) - RT \ln\left(\frac{\rho_{solv} V_{\mathrm{M}}}{MW_S}\right) , \tag{2.127}$$

where ρ_{solv} is the density of the solvent, MW_S is the molecular weight of the solvent and V_{M} is the molar volume of the ideal gas. Finally it should be added that the empirical fitted terms used in COSMO-RS imply the use of the BP86 functional for the quantum chemical part.

The Gibbs energy $G_{\mathrm{Solv},X}^S$ of the solvation can be used to calculate the Gibbs energies of association $\Delta G_{\mathrm{A}}^{\mathrm{sol}}$ in solution. For this purpose the Gibbs energy of association $\Delta G_{\mathrm{A}}^{\mathrm{gas}}$ in the gas phase is calculated

$$\Delta G_{\mathrm{A}}^{\mathrm{gas}} = \Delta E^{\mathrm{QM}} + \Delta G_{\mathrm{RRHO}} \,, \tag{2.128}$$

$$\Delta E^{\mathrm{QM}} = \sum_X^{\text{products}} E_X^{\mathrm{QM}} - \sum_X^{\text{edcuts}} E_X^{\mathrm{QM}} \,, \tag{2.129}$$

$$\Delta G_{\mathrm{RRHO}} = \sum_X^{\text{products}} G_X^{\mathrm{RRHO}} - \sum_X^{\text{edcuts}} G_X^{\mathrm{RRHO}} \,. \tag{2.130}$$

E_X^{QM} is the quantum chemically calculated electronic energy in the gas phase of the product or educt molecule. G_X^{RRHO} consists of the contributions of translation, rotation and vibration in the gas phase for the product or educt molecule. Afterwards the Gibbs energy $\Delta G_{\mathrm{Solv}}^S$ to transfer the educt in the gas phase and the product from the gas phase

in solution is added (Figure 1.4)

$$\Delta G_{\mathrm{A}}^{\mathrm{sol}} = \Delta G_{\mathrm{A}}^{\mathrm{gas}} + \Delta G_{\mathrm{Solv}}^{S} \,, \tag{2.131}$$

$$\Delta G_{\mathrm{Solv}}^{S} = \sum_{X}^{\mathrm{products}} G_{\mathrm{Solv},X}^{S} - \sum_{X}^{\mathrm{edcuts}} G_{\mathrm{Solv},X}^{S} \,. \tag{2.132}$$

The enthalpy and entropy can be calculated[68] similar to equation (2.128) and (2.131) with the use of the enthalpy $\Delta H_{\mathrm{Solv}}^{S}$ and entropy $\Delta S_{\mathrm{Solv}}^{S}$ to transfer the educt in the gas phase and the product from the gas phase in solution

$$\Delta H_{\mathrm{Solv}}^{S} = -T^2 \frac{\partial \left(\frac{\Delta G_{\mathrm{Solv}}^{S}}{T} \right)}{\partial T} \,, \tag{2.133}$$

$$\Delta S_{\mathrm{Solv}}^{S} = \frac{\Delta H_{\mathrm{Solv}}^{S} - \Delta G_{\mathrm{Solv}}^{S}}{T} \,. \tag{2.134}$$

2.6.3 D-COSMO-RS

COSMO-RS includes hydrogen bonds and van der Waals interactions in contrast to COSMO. However, COSMO-RS just yields an additive term to the total Gibbs energy. Direct COSMO-RS[139,140] (DCOSMO-RS) combines COSMO and COSMO-RS by including the chemical potential μ_X^S (equation (2.119)) and the "combinatorial contribution" $\mu_X^{C,S}$ (equation (2.120)) of solute molecule X in solvent S from the COSMO-RS approach in the SCF procedure of Hartree-Fock and DFT with the additive external potential $V_{S,\mathrm{DC\text{-}RS}}^{X}$

$$V_{S,\mathrm{DC\text{-}RS}}^{X} = \frac{\partial E_{S,\mathrm{DC\text{-}RS}}^{X}}{\partial \rho^{X}\left(\vec{r}\right)} \,, \tag{2.135}$$

where $\rho^{X}\left(\vec{r}\right)$ is the electron density and $E_{S,\mathrm{DC\text{-}RS}}^{X}$ the D-COMSO-RS energy

$$E_{S,\mathrm{DC\text{-}RS}}^{X} = -\frac{1}{2}\frac{\varepsilon_{\mathrm{r}}-1}{\varepsilon_{\mathrm{r}}+\frac{1}{2}} \sum_{i\in X} \sigma_i s_i \sum_{j\in X} A_{ij}^{X} \sigma_j s_j + f_{\mathrm{pol}} \sum_{i\in X} s_i \tilde{\mu}_S\left(\sigma_i\right) + \mu_X^{C,S} \,. \tag{2.136}$$

The first term is from the COMSO energy E_S^X (equation (2.111)). The second term uses the scaling factor f_{pol} which is set to unity and the reduced σ-potential $\tilde{\mu}_S\left(\sigma_i\right)$

$$\tilde{\mu}_S\left(\sigma_i\right) = \mu_S\left(\sigma_i\right) - \left(1 - \frac{\varepsilon_{\mathrm{r}}-1}{\varepsilon_{\mathrm{r}}+\frac{1}{2}}\right) c_0 \sigma_i^2 \,. \tag{2.137}$$

The c_0 parameter is calculated from the σ-potential $\mu_S(\sigma_i)$ of hexan. The term "$-\left(1-\frac{\varepsilon_r-1}{\varepsilon_r+\frac{1}{2}}\right)c_0\sigma_i^2$" is a correction for non-polar solvents (small relative permittivity ε_r), which subtracts the contributions from the misfit energy $E_{MF}(\sigma,\sigma')$ (2.123) because the electrostatic interactions are well enough modeled with COSMO for non-polar solvents.[140]

2.7 Basis Set

In both Hartree-Fock and DFT, orbitals are expanded in atomic basis functions ϕ (see equation (2.27)). Atom-centered Gaussian Type Orbitals[118] (GTOs) are used as basis functions in this thesis. In Cartesian coordinates they are given as

$$\phi_{\zeta,l_x,l_y,l_z}(x,y,z) = Nx^{l_x}y^{l_y}z^{l_z}\exp\left(-\zeta r^2\right). \tag{2.138}$$

N is the normalization constant, l_x, l_y and l_z are defining the type of orbital (s-,p-,d-,... orbital), $r = x^2+y^2+z^2$ is the radius and the exponent ζ defines the spherical extent. GTOs have a "$-r^2$" dependence in the exponential, although a "$-r$" dependence would be a better representation of the electronic wave function from molecules. Two problems arise from the "$-r^2$" dependency: GTOs have a zero slope at the position of the nucleolus instead of a cusp (discontinuous derivative) and far away from the nucleus the GTOs decay too rapidly. This disadvantage can be solved by using large numbers of GTOs. GTOs have the advantage that the product of any number of Gaussians can be rewritten as a single Gaussian, and their integrals can be solved in a closed analytical form. This speeds up the calculation drastically.

The generic GTO (equation (2.138)) is a primitive GTO (PGTO). A further speedup is achieved by using contracted GTOs (CGTOs) which are linear combinations of PGTOs with the fixed contraction coefficients a_i

$$\phi^{\mathrm{CGTO}} = \sum_i a_i\phi^{\mathrm{PGTO}}. \tag{2.139}$$

The number of integrals in the Hartree-Fock and DFT approach is not reduced by the use of CGTOs because the number of integrals depends on the number of PGTOs used to build the CGTO. Despite this, CGTOs reduce the calculation time because the variational problem (size of Fock matrix and number of iterations) gets smaller. The fixed linear combinations reduce the flexibility of the basis set. Therefore, CGTOs are used especially for the core orbitals because these need large numbers of basis functions due to the cusp,

but not necessarily large flexibility as they are less effected by the chemical bonding situation.

2.7.1 BSSE and Basis Set Extrapolation

The binding energy ΔE_{AB} of a dimer AB is calculated as difference between the electronic energy E_{A} of the dimer and the electronic energies E_{A} and E_{B} of the isolated monomers A and B (see equation (2.129)). If the calculation is not done in the basis set limit, an error can occur, which is called the basis set superposition error (BSSE). The dimer is calculated with the basis functions of both monomers, consequently more basis functions are used to describe the dimer. The electronic energy E_{AB} of the dimer is therefore closer to the basis set limit and the binding energy ΔE_{AB} gets too negative. The BSSE can be corrected with the counterpoise (CP) method[141]

$$\Delta E_{\mathrm{AB}} = E_{\mathrm{AB}} - E_{\mathrm{A}} - E_{\mathrm{B}} - E^{\mathrm{CP}}\,, \tag{2.140}$$

where E^{CP} is the counterpoise correction for the BSSE. It is gained from the not relaxed monomers

$$E^{\mathrm{CP}} = E_{\mathrm{A}}^{\mathrm{AB,AB}} - E_{\mathrm{A}}^{\mathrm{AB,A}} + E_{\mathrm{B}}^{\mathrm{AB,AB}} - E_{\mathrm{B}}^{\mathrm{AB,B}}\,. \tag{2.141}$$

$E_{\mathrm{A}}^{\mathrm{AB,AB}}$ is the energy of monomer A with the basis functions of dimer AB and the nuclear positions like in dimer AB, $E_{\mathrm{A}}^{\mathrm{AB,A}}$ is the energy of monomer A only with the basis functions of monomer A and the nuclear positions like in dimer AB, and $E_{\mathrm{B}}^{\mathrm{AB,AB}}$ and $E_{\mathrm{B}}^{\mathrm{AB,B}}$ are the corresponding energies for monomer B.

The counterpoise method can not be used to correct the BSSE in intramolecular reactions. Since the D-COSMO-RS approach was used for intramolecular reactions, a different approach was applied. In DFT calculations the basis set limit is often already achieved by using a quadruple-ζ basis set. The D-COSMO-RS approach is too time demanding to use a quadruple-ζ basis sets. Thus, the binding energy was calculated in gas phase with a quadruple-ζ ($\Delta E_{\mathrm{gas}}^{\text{q-}\zeta}$) and with a triple-$\zeta$ ($\Delta E_{\mathrm{gas}}^{\text{t-}\zeta}$). The difference between both was used as correction for the binding energy in solution $\Delta E_{\mathrm{sol}}^{\text{t-}\zeta}$ calculated with a triple-ζ basis set and the D-COSMO-RS approach, to gain the corrected binding energy in solution $\Delta E_{\mathrm{sol}}^{\mathrm{corr}}$

$$\Delta E_{\mathrm{sol}}^{\mathrm{corr}} = \Delta E_{\mathrm{sol}}^{\text{t-}\zeta} + \left(\Delta E_{\mathrm{gas}}^{\text{q-}\zeta} - \Delta E_{\mathrm{gas}}^{\text{t-}\zeta}\right)\,. \tag{2.142}$$

The third approach to correct the BSSE is to extrapolate the basis set limit, because in

the basis set limit the BSSE becomes zero. For correlation methods like coupled cluster the correlation energy in the basis set limit $E_{\text{Limit}}^{\text{correl}}$ can be extrapolated[142] with correlation consistent basis sets[143]

$$E_{\text{Limit}}^{\text{correl}} = \frac{Y^3 E_{\text{Y}}^{\text{correl}} - X^3 E_{\text{X}}^{\text{correl}}}{Y^3 - X^3} \,. \tag{2.143}$$

X and Y are the cardinal numbers of the basis sets (i.e. three for triple-ζ, etc.) with $Y = X+1$ and $E_{\text{Y}}^{\text{correl}}$ and $E_{\text{X}}^{\text{correl}}$ are correlation energies calculated with the corresponding correlation consistent basis sets.

At the Hartree-Fock and DFT level the energy converges exponential, and thus much faster than for coupled cluster.[104,142,144] Hence, a basis set extrapolation is not necessary at the Hartree-Fock and DFT level. Instead a calculation with quadruple-ζ basis sets is often good enough.

2.7.2 Effective core potential

In this thesis systems are calculated which include samarium and iodine. Samarium and iodine are heavy atoms (fifth and sixth row). These atoms contain a large amount of core electrons which are not effected by the chemical bonding situation. It is very time consuming to include these electrons in the calculation not only because of their large number, but also because relativistic effects cannot be ignored for atoms with heavy cores. In this cases the effective core potential (ECP) approximation[145] is used. In this approximation only the valence electrons are explicitly considered. The core electrons are not considered explicitly, instead their effects on the valence system are included in an effective core potential. The exact Hamiltonian is replaced with an pseudo-Hamiltonian which only acts on the valence electrons and includes the ECP. Further the scalar relativistic effects on the valence electrons can be included in the ECP. Thus, scalar relativistic effects can be included in formally non-relativistic calculations. The errors from the effective core potential approximation are usally much smaller than the errors from the regular post-Hartree-Fock or DFT methods, provided reasonable effective core potentials are selected, so that every electron that effects the chemical bonding is still treated explicitly.[145]

2.8 Population Analysis

2.8.1 Mulliken Population Analysis

The partial atomic charges Q_A can be calculated with a Mulliken population analysis.[146]

The total number of electrons n_{el} can be calculated as integral over the electron density $\rho(\vec{r})$ (equation (2.54))

$$n_{\mathrm{el}} = \int \rho(\vec{r}) = \sum_{a}^{N} \sum_{\mu}^{L} \sum_{\nu}^{L} C_{\mu a} C_{\nu a} \int \phi_{\mu}(\vec{r})\, \phi_{\nu}(\vec{r})\, \mathrm{d}\vec{r} = \sum_{\mu}^{L} \sum_{\nu}^{L} D_{\mu\nu} S_{\mu\nu}\,, \tag{2.144}$$

where $D_{\mu\nu}$ are the elements of the Density-matrix $\mathbf{D}$ and $S_{\mu\nu}$ the elements of the overlap matrix $\mathbf{S}$. The $\mathbf{DS}$ matrix is used in the Mulliken population analysis to determine the electron population ρ_A at atom A. This is done by summing up all contributions which contain basis functions centered at atom A

$$\rho_A = \sum_{\mu \in A}^{L} \sum_{\nu}^{L} D_{\mu\nu} S_{\mu\nu}\,. \tag{2.145}$$

The partial atomic charge Q_A or Mulliken charge at atom A is then calculated as difference between the nuclear charge Z_A and electron population ρ_A at atom A

$$Q_A = Z_A - \rho_A\,. \tag{2.146}$$

The Mulliken population analysis can also be used to determine the populations of specific functions, like all f functions at an atom A. In case of an "unrestricted" calculation, the difference between the electron population of electrons with α- and β-spin can be calculated.

2.8.2 Natural Population Analysis

Another method to determine the partial atomic charges Q_A is the Natural Population Analysis (NPA).[147] The gist of the NPA is to construct a set $\left\{\phi_i^{\mathrm{NAO}}\right\}$ of "natural atomic orbitals" (NAOs) for the molecule of interest. "Natural orbitals" (NOs)[148] are derived by diagonalizing the "first order reduced density matrix" $\Gamma(\vec{x}'_1, \vec{x}_1)$

$$\Gamma(\vec{x}'_1, \vec{x}_1) = n \int \cdots \int \Phi^*(\vec{x}'_1, \vec{x}_2 \ldots, \vec{x}_n)\, \Phi(\vec{x}_1, \vec{x}_2 \ldots, \vec{x}_n)\, \mathrm{d}\vec{x}_2 \ldots \mathrm{d}\vec{x}_n\,. \tag{2.147}$$

The corresponding eigenvectors are the NOs and the eigenvalues are the occupation numbers. However, NOs are completely delocalized. NAOs are similar to NOs, but they are localized on the individual atoms of the molecule. The first order reduced density matrix $\hat{\Gamma}(\vec{x}'_1, \vec{x}_1)$ can be partitioned in Alm subblocks associated with the (nonorthogonal)

atom-centered basis functions. For the Alm subblocks A denotes the atom, l the angular momentum (s, p, d, f, ...) and m the particular symmetry element of l (e.g. p_x etc.). The subblocks are diagonalized interdependently from each other, which leads to a set $\left\{\phi_i^{\mathrm{PNAO}}\right\}$ of "pre-natural atomic orbitals" (PNAOs). The PNAOs are only orthogonal towards PNAOs on the same atom A (i.e. $\left\langle \phi_i^{\mathrm{PNAO},A} | \phi_j^{\mathrm{PNAO},A} \right\rangle = \delta_{ij}$ but $\left\langle \phi_i^{\mathrm{PNAO},A} | \phi_j^{\mathrm{PNAO},B} \right\rangle \neq 0$). The complete set of PNAOs $\left\{\phi_i^{\mathrm{PNAO}}\right\}$ is orthogonalized while preserving their atomic nature, which yields the set $\left\{\phi_i^{\mathrm{NAO}}\right\}$ of orthogonal NAOs. This orthogonalization is done in four steps. First the PNAOs are divided in two classes. The highest occupied PNAOs form the "natural minimal basis" (NMB), compare equation (2.148) for occupation. The other PNAOs, with occupation numbers close to zero, form the "natural Rydberg basis" (NRB). In the second step the NMB set is orthogonalized with an occupancy-weighted procedure.[147] The next step is to orthogonalize the NRB set relative to the NMB set with the Schmidt orthogonalization. And in the fourth step the NRB set is orthogonalized with the occupancy-weighted procedure.

The electron population ρ_A at atom A is than calculated with

$$\rho_A = \sum_i w_i^A = \sum_i \int\int \phi_i^{*\mathrm{NAO},A}\left(\vec{x}'_1\right) \Gamma\left(\vec{x}'_1, \vec{x}_1\right) \phi_i^{\mathrm{NAO},A}\left(\vec{x}_1\right) \mathrm{d}\vec{x}'_1 \mathrm{d}\vec{x}_1 \,, \tag{2.148}$$

where w_i^A is the occupancy of orbital $\phi_i^{\mathrm{NAO},A}$. In the final step the partial atomic charge Q_A can be calculated with equation (2.146). The advantage of the NPA is that it depends less on the applied basis set than the Mulliken population analysis.[147]

2.9 Orbital energies

2.9.1 Hartree-Fock Method

The orbital energies ε_a for the canonical spin orbitals $\tilde{\chi}_a$ in the Hartree-Fock method can be connected approximatively with the experimentally accessible physical quantities ionization energy E_{IE} and electron affinity E_{EA}, receptively. The ionization energy E_{IE} is the energy needed to remove an electron from an n-electron system. E_{IE} is calculated exactly with

$$E_{\mathrm{IE}} = E_{n-1} - E_n \,. \tag{2.149}$$

E_n is the total energy of a n-electron system and E_{n-1} is the energy of the corresponding system with one electron less. Now two approximations are invoked. The first is to use

the Hartree-Fock method to calculate E_n. The second is to assume that the corresponding $(n-1)$-electron system has identical spin orbitals, except one electron is removed from the canonical spin orbital $\tilde{\chi}_c$. This is known as "frozen orbital" approximation. The energies for the n-electron and $(n-1)$-electron system are given with equation (2.15) as

$$E_n = \sum_a^N h_a + \frac{1}{2}\sum_a^N\sum_b^N (J_{ab} - K_{ab}) \;, \tag{2.150}$$

$$E_{n-1} = \sum_{a\neq c}^N h_a + \frac{1}{2}\sum_{a\neq c}^N\sum_{b\neq c}^N (J_{ab} - K_{ab}) \;. \tag{2.151}$$

N is the number of occupied spin orbitals. It follows

$$\begin{aligned} E_{\mathrm{IE}} &= E_{n-1} - E_n \\ &= -h_c - \frac{1}{2}\sum_a^N (J_{ac} - K_{ac}) - \frac{1}{2}\sum_b^N (J_{cb} - K_{cb}) \\ &= -h_c - \sum_b^N (J_{cb} - K_{cb}) \\ &= -\langle\tilde{\chi}_c|\,\hat{F}\,|\tilde{\chi}_c\rangle = -\varepsilon_c \end{aligned} \tag{2.152}$$

with the Fock operator $\hat{F}$ (see equation (2.24)). Thus, the ionization energy E_{IE} needed to remove a electron from the canonical spin orbital $\tilde{\chi}_c$ is equal to the negative orbital energy of $\tilde{\chi}_c$ within the Hartree-Fock method and the "frozen orbital" approximation. This is known as "Koopmans' theorem".[149] The electron affinity E_{EA} is the energy released if an electron is added. E_{EA} can be calculated like E_{IE}. The electron is added in the unoccupied canonical spin orbital $\tilde{\chi}_r$

$$\begin{aligned} E_{\mathrm{EA}} &= E_n - E_{n+1} \\ &= -h_r - \sum_b^N (J_{rb} - K_{rb}) \\ &= -\varepsilon_r \,. \end{aligned} \tag{2.153}$$

The unoccupied orbitals are not optimized and are just left-overs from the basis set expansion. For this reason, the use of unoccupied orbitals is disputed.[83,104]

The HOMO-LUMO gap, the difference in energy between the highest occupied molecular orbital (HOMO) and the lowest unoccupied molecular orbital (LUMO), corresponds

in the Hartree-Fock method to the difference between ionization energy E_{IE} and electron affinity E_{EA}.

2.9.2 DFT

The Kohn-Sham orbitals have a different connection to experimentally accessible physical quantities. For Kohn-Sham DFT with the exact exchange-correlation functional E_{XC}, the electron density $\rho^{\mathrm{KS}}(\vec{r})$ decays for large distances from all nuclei as

$$\rho^{\mathrm{KS}}(\vec{r}) \propto \exp\left[-2\sqrt{-2\varepsilon_{\mathrm{HOMO}}}\,|\vec{r}|\right], \tag{2.154}$$

where $\varepsilon_{\mathrm{HOMO}}$ is the orbital energy of the HOMO. The exact electron density $\rho(\vec{r})$ decays for large distances from all nuclei as

$$\rho(\vec{r}) \propto \exp\left[-2\sqrt{2E_{\mathrm{IE}}}\,|\vec{r}|\right], \tag{2.155}$$

and it follows

$$-\varepsilon_{\mathrm{HOMO}} = E_{\mathrm{IE}}\,. \tag{2.156}$$

This connection is exact. Kohn-Sham DFT with the exact exchange-correlation functional does not need a frozen orbital approximation and it includes the correlation energy. Unfortunately the exact exchange-correlation functional is not known; therefore, it is just an approximation when an actually Kohn-Sham DFT is applied.

The unoccupied orbitals in the Hartree-Fock approach and in the local Kohn-Sham DFT (like LDA, GGA and meta-GGA) are different. In the Hartree-Fock approach the unoccupied orbitals are determined in a field of n electrons because of the orbital-dependent exchange term.[150] Therefore, the orbital energy of an unocuppied canonical spin orbital in the Hartree-Fock method can be interpreted as the energy to add an electron (electron affinity E_{EA}). In local Kohn-Sham DFT the unoccupied orbitals are determined in a field of $n-1$ electrons because of the local exchange-correlation term. Thus, the unocuppied orbitals in local Kohn-Sham DFT represent excited electrons instead of additional electrons.[150] The HOMO-LUMO gap in local Kohn-Sham DFT is an approximation for the lowest excitation energy of a system.[150] This excitation energy can often be measured by ultraviolet–visible spectroscopy. Due to a self-interaction error the orbital energy of the orbital energy of the HOMO calculated with LDA and GGA functionals often becomes too positive.[104,151] The Coulomb term in Hartree-Fock and Kohn-Sham DFT includes a

Coulomb interaction of each electron with itself. This is known as self-interaction and is also an error. In the Hartree-Fock method the self-interaction is exactly canceled by the exchange term. Exchange-correlation functionals for Kohn-Sham DFT usually not fully cancel the self-interaction error.[104] Hybrid functionals can be used to improve the orbital energy of the HOMO,[150] but unfortunately they include the orbital-dependent exchange term from the Hartree-Fock method. The HOMO-LUMO gap of a hybrid functional which uses mainly the orbital-dependent exchange term can not be interpreted as an approximation for the lowest excitation energy of a system. In this thesis the functional B3LYP is used to calculate the lowest excitation energy by using the HOMO-LUMO gap. B3LYP is a hybrid functional, but it uses only 20% orbital-dependent exchange.

3 Summarized Results

This chapter summarizes the outcomes of this thesis which are published in eight scientific papers and presents additional results not yet published. The investigations were performed in close collaboration with specialists for supramolecular chemistry (AG Schalley) and for molecular dynamic simulations (AG Weber).

The investigation started with the development of a multilevel DFT-based approach to determine the Gibbs energy of association in solution of the monovalent binding motif, which consists of an 18-crown-6 ether and a primary ammonium ion (cation) with an organic residue/spacer. Figure 3.1 shows the investigated systems. The results are published in **Paper A1** (section 5). The first step of our approach is the determination of the electronic binding energy ΔE^{QM} in the gas phase. It was evaluated with LCCSD(T) and various DFT-D3(BJ) to be about $-260\ \mathrm{kJ\,mol^{-1}}$ mainly resulting from the formation of three ionic hydrogen bonds between the hydrogen atoms of the ammonium cation and the oxygen atoms of the crown ether. Dispersive interactions between the aromatic system of the crown ether and the organic spacer of the primary ammonium ion contribute only about $-20\ \mathrm{kJ\,mol^{-1}}$. Combining electronical, translational, rotational and vibrational contributions yield the Gibbs energy of association in the gas phase $\Delta G_{\mathrm{A}}^{\mathrm{gas}}$. The effects of translation, rotation and vibration were taken into account by a RRHO approximation with free-rotor approximation for low-lying vibrations.

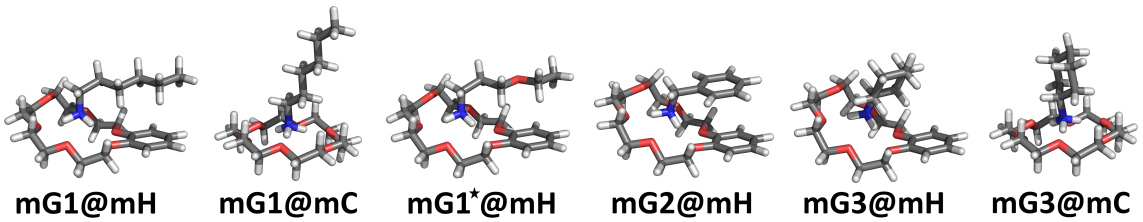

Figure 3.1: Crown ether/ammonium complexes investigated in **Paper A1**.[72] The tosylate counterions are not depicted.

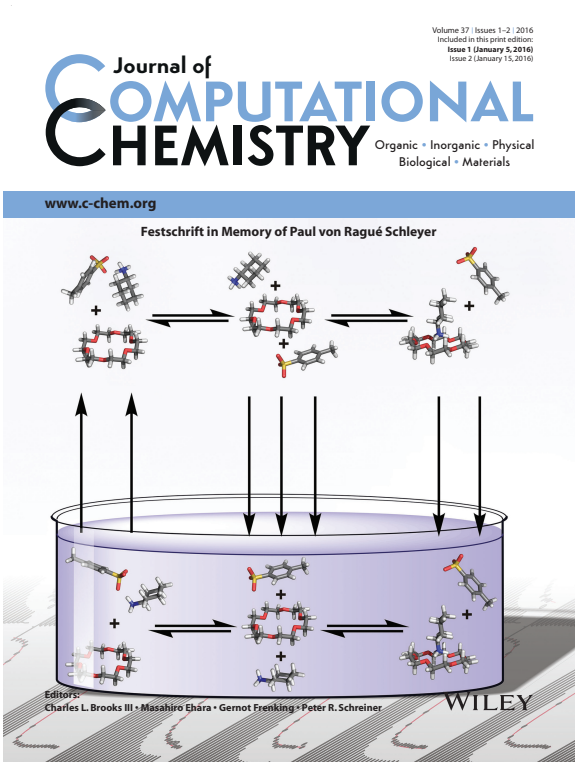

Figure 3.2: Inside cover for **Paper A1** adapted with permission from Achazi *et al.*[72] (©2015 Wiley Periodicals, Inc). It depicts the approach to calculate the Gibbs energy of association in solution $\Delta G_{\mathrm{A}}^{\mathrm{sol}}$. First the Gibbs energy of association in the gas phase $\Delta G_{\mathrm{A}}^{\mathrm{gas}}$ is calculated, then the Gibbs energy of solvation $\Delta G_{\mathrm{Solv}}^{\mathrm{S}}$ is included according to the arrows.

Their main impact on the complex formation is the loss in translational (about 50 kJ mol^{-1}) and rotational entropy (about 30 kJ mol^{-1}). The translational and the rotational entropy are proportional to the logarithm of the molecular mass and the logarithm of the square root of the moments of inertia, respectively. Due to the logarithmic dependencies, these entropies reduce when two small molecules merge to one heavier but similarly sized entity. It should also be pointed out that the vibrational entropy, including the entropy for internal rotations, has a binding effect (i.e. a gain in vibrational entropy occurs) of about -15 kJ mol^{-1}. The gain can most likely be assigned to changes in degrees of freedom. Three translational and three rotational degrees of freedom are transformed into six vibrational ones in the joint system. The gain in vibrational entropy (including internal rotations) can be almost exclusively determined from the vibrational modes below 600 cm^{-1}. Their number is growing by six. This would yield a gain of at least -11 kJ mol^{-1} if all new vibrational modes would be at 600 cm^{-1}. However, all vibrational modes are effected by the formation of the complex. Thus, the gain in vibrational entropy cannot be completely explained by the increased number of vibrational degrees of freedom. In comparison with an experimental measurement[152] of $\Delta G_{\mathrm{A}}^{\mathrm{gas}}$, the DFT functional TPSS with D3(BJ) dispersion correction was the most accurate giving a deviation of only 3 kJ mol^{-1} compared to the experiment.

The experiments to determine the Gibbs energy of association in solution $\Delta G_{\mathrm{A}}^{\mathrm{sol}}$ of the monovalent 18-crown-6/ammonium ion complex were performed in various chloroform/methanol solvent mixtures and tosylate was used as counterion for the ammonium

ion. In order to include the effects of the chloroform/methanol solvent mixtures in the calculation, the implicit solvent model COSMO-RS was applied. In this thesis several conformers of the same complex or molecule have been included in the COSMO-RS approach to predict the surface charges of the ensembles of solute molecules. However, it was only important to consider that the lowest energy structure in the gas phase and the one in solvent were usually not identical. The additional inclusion of complexes and molecules other than the lowest energy structures into the calculation changed the results by less than 1 kJ mol^{-1}. The agreement with the experimental Gibbs energy of association in solution $\Delta G_{\mathrm{A}}^{\mathrm{sol}}$ for the crown ether/ammonium complex could be increased by adding the Gibbs energy for the dissociation of the complex that is formed by the primary ammonium ion and the tosylate counterion. In conclusion, it has been assumed that the primary ammonium ion forms a complex with the tosylate counterion which has to be dissolved for the formation of the crown ether/ammonium complex as shown in Figure 3.2.

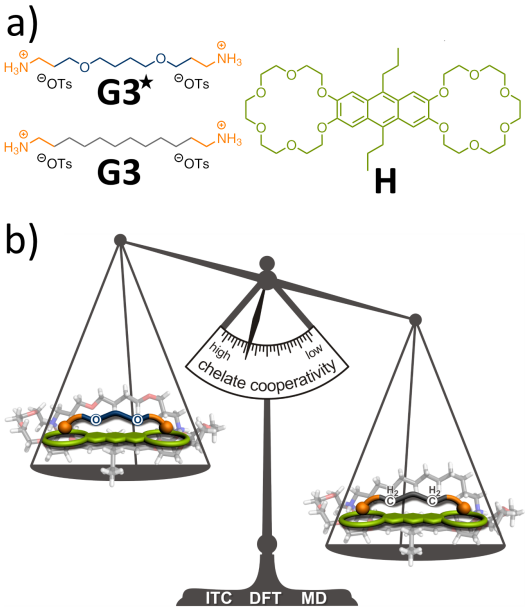

Figure 3.3: a) The investigated divalent crown ether/ammonium complexes. b) Graphical abstract included with permission from von Krbek *et al.*[153] (©2016 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim). It depicts the significantly higher chelate cooperativity of **G3@H** compared to **G3*****@H**.

The association thermodynamics (including the Gibbs energy of association in solution $\Delta G_{\mathrm{A}}^{\mathrm{sol}}$) for the complexes was measured with isothermal titration calorimetry (ITC) by the AG Schalley. The experiments showed that the unbound primary ammonium ion forms a complex with the tosylate counterion, too. Our investigation revealed that changing the methanol content of the solvent has two counterbalancing effects. Methanol aggravates the formation of the crown ether/ammonium complexes by forming hydrogen bonds to the free crown ether and the ammonium ion and it facilitates the formation of the crown ether/ammonium complexes by dissolving the ammonium-tosylate complexes. The extent of the cancellation between these two effects depends on the structural details of the actual crown ether and ammonium ion.

The described theoretical multilevel (DFT-based) approach yielded Gibbs energies of

association $\Delta G_{\mathrm{A}}^{\mathrm{sol}}$ with an average deviation to the experimental results of 4.3 kJ mol^{-1}, which is state-of-the-art.[70] In summary a suitable theoretical approach was developed for the determination of the thermochemistry of crown ether/ammonium complexes in organic solvents with interfering counterions.

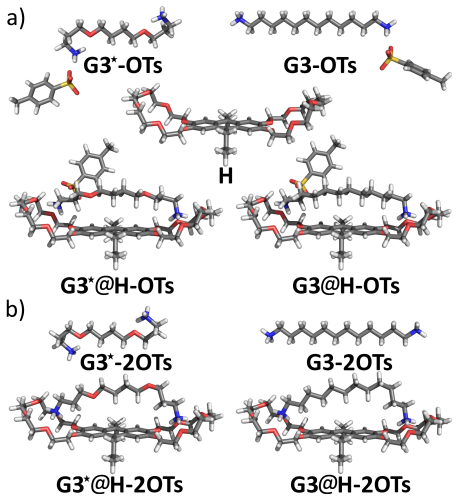

Figure 3.4: Lowest-energy structures with a) one counterion and b) none counterion. Calculated with BP86/def-TZVP and the implicit solvent model COSMO ($\varepsilon_{\mathrm{r}} = \infty$). Included with permission from von Krbek *et al.*[153] (©2016 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim).

Switching to the divalent systems, in **Paper A2** allosteric and chelate cooperativity effects were investigated in two virtually identical divalent crown ether/ammonium complexes **G3⋆@H** and **G3@H** (Figure 3.3). Their two binding sites are built by the monovalent crown ether/ammonium binding motif investigated in **Paper A1**. The two divalent primary ammonium ions **G3⋆** and **G3** have only slightly different spacers as two ether oxygen atoms of **G3⋆** are replaced by two isoelectronic methylene groups in **G3**. The association thermodynamics of the divalent complexes **G3⋆@H** and **G3@H** were measured in various chloroform/methanol mixtures with ITC. The allosteric and the chelate cooperativity were determined experimentally with a double mutant cycle (DMC) analyses.

Similar to the monovalent complexes, our investigation of divalent complexes showed a delicate counterbalancing between the effects of the solvent methanol on the host-guest complex formation and on the guest-counterion complex dissociation. Increasing the methanol content increases the association strength until a maximum is reached at the 1:1 (v/v) chloroform/methanol mixture. From that onward the association strength decreases with increasing methanol content, likely because the ions are then completely dissolved. **G3⋆@H** and **G3@H** both had a high association constant in the experiments with **G3@H** having the slightly higher one.

We successfully adjusted our multilevel DFT-based approach for the divalent systems. For medium methanol content the best agreement between theoretical and experimental Gibbs energy of association in solution $\Delta G_{\mathrm{A}}^{\mathrm{sol}}$ was achieved by including one tosylate counterion (Figure 3.4a) and for high methanol content by assuming the ions were completely

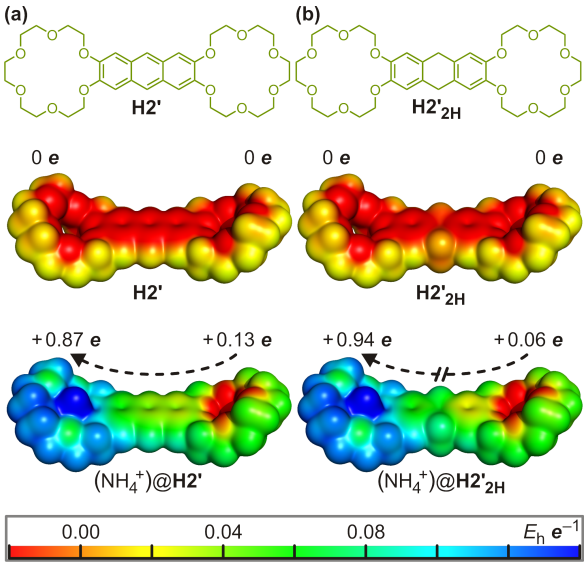

Figure 3.5: a) **H2'** is a simplified (side-chain-truncated) derivative of **H**. **H2'**$_{\mathbf{2H}}$ is a derivative of **H2'**, two additonal hydrogen atoms interrupt the conjugated π system. The flat molecular structure of the anthracene spacer is fixed with constraints in **H2'**$_{\mathbf{2H}}$. a) Structural formula of **H2'**, and electrostatic potential maps of **H2'** and (NH_4^+)@**H2'**. b) Structural formula of **H2'**$_{\mathbf{2H}}$, and electrostatic potential maps of **H2'**$_{\mathbf{2H}}$ and (NH_4^+)@**H2'**$_{\mathbf{2H}}$. The charges of each binding site are given in e. Included with permission from von Krbek *et al.*[153] (©2016 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim).

dissolved (Figure 3.4b). Especially for the 10:3, 2:1 and 1:2 (v/v) chloroform/methanol mixture the agreement was very good (< 5 kJ mol^{-1}). The experimental $\Delta G_{\mathrm{A}}^{\mathrm{sol}}$ value for the 1:1 (v/v) chloroform/methanol mixture was in between the two models (of one and none counterion). In case of **G3***, the intramolecular hydrogen bonds (Figure 3.4) have to be taken into account. They are most probably the reason for the lower association strength of **G3*** compared to **G3**, which was found in the DFT calculations as well as in the experiments. An additional proof for the effect of the intramolecular hydrogen bonds was obtained by MD simulations.

In the experiments negative allosteric cooperativities were discovered in the host **H** and the guests **G3** and **G3***. For the host, two effects rationalize the negative allosteric cooperativity. The first effect occurs because the second ammonium ion experiences charge-charge repulsion from the first, which makes the second binding event less favorable. Theoretically we determined an upper limit for this effect; however, the counterions lower the charge, and therefore the negative allosteric cooperativity is much smaller than our

suggestion. A higher methanol content dissolves the guest-counterion complex, and should consequently enhance the negative allosteric cooperativity. The experimental measured allosteric cooperativity factors of the host are confirming this trend. However, due to the uncertainty of the measurement the experimental measured allosteric cooperativity factors have to be seen as equal. Experiments with higher accuracy would be necessary to verify the theoretical considerations. The second effect contributing to the negative allosteric cooperativity is the high polarizability of the anthracene spacer. The binding of the first ammonium ion polarizes the anthracene spacer. Therefore, positive charge is transmitted to the unbound binding site of the host which repels the second ammonium ion. We elucidated the effect with electrostatic potential maps and partial charges evaluated with DFT (Figure 3.5). The negative allosteric cooperativity in the guest is an effect of the counterion. The first ammonium ion of the divalent guest binds to the crown ether. This liberates a counterion, which interferes with the binding of the second ammonium ion. Accordingly the negative allosteric cooperativity of the guest gets attenuated by increasing the methanol content. **G3*** experiences a stronger negative allosteric effect than **G3**. This can be explained by the lowest energy conformation of the free guest **G3*** found by DFT (Figure 3.4). Both ammonium groups form intramolecular hydrogen bonds with the oxygen atoms of the spacer. If the first ammonium ion binds to crown ether, the distance between the two ammonium groups increases and thereby their charge-charge repulsion decreases. This strengthens the intramolecular hydrogen bond of the second ammonium group (Figure 3.6). Hence, the second binding event experiences a stronger competition than the first one. It can be assumed that conformations of **G3*** with only one intramolecular hydrogen bond experience a similar but less pronounced effect.

As a result, allosteric cooperativity cannot explain the strong association constants of the two divalent complexes. The complexes must have a positive chelate cooperativity as confirmed actually by the experiments. Attractive spacer-spacer interactions are probably a reason for this, as shown for similar systems experimentally[20] and by us[154] in **Paper A4**. The MD simulations provide evidence for a very effective rebinding effect[15,155,156] in **G3*****@H** and **G3@H**, which adds to the positive chelate cooperativity. In addition to this, the MD simulations rationalize the lower chelate cooperativity of **G3*****@H** compared to **G3@H**: The free **G3** is mostly in a linear conformation close to its conformation in the complex, whereas the intramolecular hydrogen bonds in **G3*** cause a coiled conformation, which has to unfold prior to complex formation.

The high association constants and chelate cooperativities were surprising. Flexible

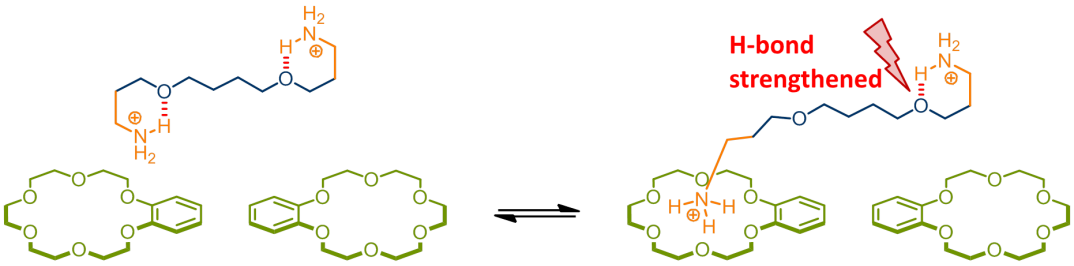

Figure 3.6: The opening of the intramolecular hydrogen bond in order to bind to the crown ether (first binding event) increases the distance between the charged ammonium groups. This strengthens the second intramolecular hydrogen bond, which reduces the association constant for the second binding event, and thus leads to negative allosteric cooperativity. Special thanks to Larissa K. S. von Krbek for creating the graphic.

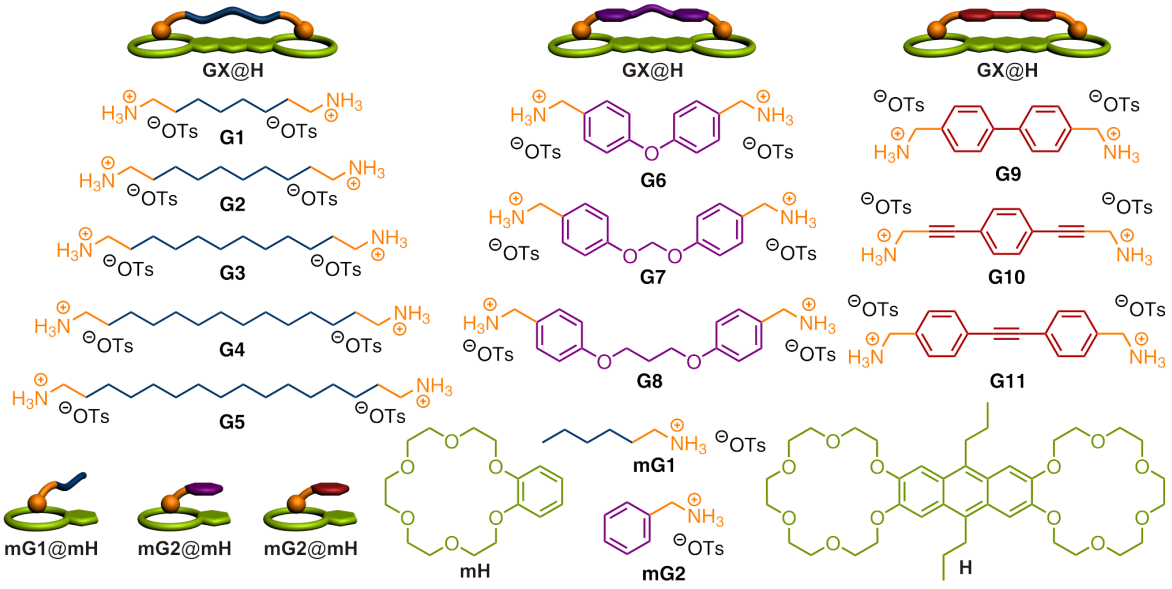

Figure 3.7: The mono- and divalent host and guest molecules investigated in **Paper A3**. Three kinds of guest were used: guests with flexible spacers (blue), semi-rigid spacers (violet) and rigid spacers (red). For the quantification of cooperative effects the monovalent complexes also have to be investigated. Included with permission from von Krbek *et al.*[157] (©2017 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim).

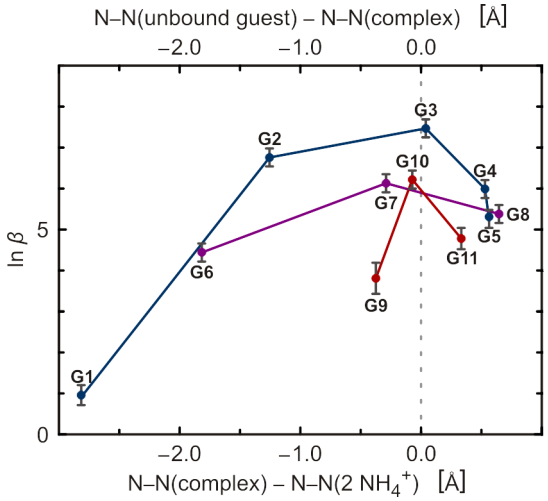

Figure 3.8: Depiction of the measured chelate cooperativity factor $\ln \beta$ against the calculated differences in N-N distance of the complexes **G***X*@**H** and the spacer free system $(NH_4^+)_2$@**H** for the guests with a flexible spacer (blue) and a semi-rigid spacer (violet) or the calculated differences in N-N distance of unbound guests **G***X* and the complexes **G***X*@**H** for guests with a rigid spacer (red). Included with permission from von Krbek *et al.*[157] (©2017 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim).

guests have the advantage to adapt to the host, in case they are not perfectly fitting. On the other hand it is expected that flexible guests experience a large entropic penalty because binding to a host will restrict their conformational space. Therefore, guests with a rigid spacer that approximates the spacing between the binding sites in the host are expected to have higher association constants,[2] which should result in a stronger positive chelate cooperativity. However, even small spatial mismatches between these rigid, highly preorganized guests and the host can heavily reduce chelate cooperativity. To get further insights into the balance of preorganization and adaptability, we investigated the association thermodynamics for a variety of additional divalent guests with spacers of different spacer lengths and flexibility (Figure 3.7) in **Paper A3**. Three kinds of guests were used: guests with flexible spacers (**G1**-**G5**), with semi-rigid spacers (**G6**-**G8**) and with rigid spacers (**G9**-**G11**). We introduced a DFT-based approach to estimate which of the guests fits the best and has the highest complementarity to the host, respectively. We distinguished between two cases. In case the guest (**G1**-**G8**) is more flexible than the host, we assumed that especially that guest will fit the best, which forms the complex with N-N distance closest to the N-N distance of the spacer free system ($(NH_4^+)_2$@**H**). Deviations from this optimal distance imply less favorable interactions between guest and host spacer

and an unfavorable deformation of the rigid host. In the other case the guest (**G9-G11**) is more rigid than the host. Actually, the spacer of the host is as rigid as the guests **G9-G11**, but the crown ethers are flexible compared to these guests. We assumed that the rigid guest which gets the least deformed due to binding to the host will fit best and has the highest complementarity to the host, respectively. This is estimated from the N-N distance in the unbound guests compared to the N-N distance in the complex. Figure 3.8 shows the guests which have according to our DFT-based approach the best complementarity to the host, have as well the most positive chelate cooperativity in the experiment. In addition the experiments show that guests with flexible spacers have by far the highest positive chelate cooperativities and the **G3@H** complex with the flexible guest **G3** has the highest association constant. From the experimental measured entropies for the ring closing step (called residual entropies), it can be seen that the guests with flexible spacers are not exhibiting in general higher entropic penalties than the guest with rigid spacer. Additionally Table 6.1 in the appendix shows the calculated change in vibrational entropy including inner rotations ($-T\Delta S_{\text{vib}}$) in the gas phase for the complex formation without counterions of the mono- and divalvent complexes of **Paper A3**. According to the calculations, the flexible guest **G3** is entropically more favorable than the semi-rigid guests **G7** and **G8** and entropically similar favorable to the rigid guest **G11**. These results are in contrast to the common notion,[2] that flexible spacers are unfavorable due to a high entropic penalty.

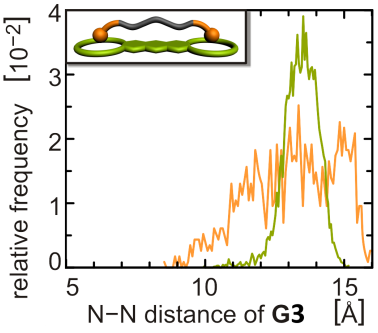

Figure 3.9: N-N distance of unbound (orange) and bound (green) **G3** during the MD simulation. Included with permission from von Krbek *et al.*[153] (©2016 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim).

Another reason for the strong positive chelate cooperativity and the high association constant achieved with **G3** is presumably its preorganization. **G3** is most probably not as preorganized as the guests with rigid spacers, but a significant fraction of the conformations of the unbound **G3** has the same N-N distance as **G3** in the **G3@H** complex as shown by the MD simulations in **Paper A2** (Figure 3.9). In this complex the guest **G3** is in a mainly linear molecular structure. The lowest energy structure for alkanes up to $C_{17}H_{36}$ in the gas phase is the linear molecular structure.[158] Beyond this, a coiled molecular structure is more favorable because of intramolecular dispersive interactions. Lipophobic solvents like water are stabilizing the coiled molecular structure because the

molecular surface is reduced. However, it is assumed that in chloroform the linear molecular structure is for alkanes up to $C_{142}H_{286}$ the lowest in energy.[159] Thus, the short size and the lipophilic solvent increase the preorganization of the guests with flexible alkane spacers investigated in **Paper A3**. Altogether, guests like **G3** which have both adaptability (compared to benzene and alkyne spacers) and preorganization can outperform the guests with semi-rigid and rigid spacers which achieve a maximum of preorganization at the price of low adaptability.

Finally it should be added that the experimental measured residual enthalpies indicate that attractive spacer-spacer interactions are important for strong positive chelate cooperativities and high association constants. The insights into multivalency gained from this research (**Paper A1-A3**) can be summarized in a "guideline" for designing highly stable supramolecular structures.

- Guest and host should have the best possible complementarity.
- Preorganization and adaptability should be both maximized, instead of maximizing the preorganization by lowering the adaptability.
- Favorable spacer–spacer interactions should be maximized.
- The preorganization of a flexible guest (or hosts) depends on its size, its ability to form intramolecular bonds/dispersive interactions and the solvent. Thus, the preorganization of a flexible guest (or hosts) can and should be optimized.
- The solvent and the counterions effect the association process and have to be optimized.

This guideline extends and refines a previously published guideline from Larissa K. S. von Krbek.[38]

In **Paper A4** we switched from the 18-crown-6 complexes to 24-crown-8 pseudorotaxanes in which the guest is threaded through the 24-crown-8 ether. There, the Gibbs energy of association $\Delta G_{\mathrm{A}}^{\mathrm{sol}}$ was calculated for mono and divalent crown ether/ammonium pseudorotaxane with different linkers in the guest molecule (Figure 3.10). Furthermore, the full double mutant cycle was investigated. The host molecules differ slightly from the ones of the experiment.[20] The 1,4-diazanaphthalene groups are replaced by phenyl groups, and the side chains of the anthracene spacer in the divalent host are replaced by hydrogen atoms. Our calculated results agree with the experimental findings. The

divalent guest with the shortest spacer has the highest association constant because of dispersion interaction between the spacers from the guest and the host. In case of the shortest spacer a nearly ideal $\pi - \pi$ stacking is achieved. With larger spacers this is not possible due to steric hindrance. This investigation supports the experimental finding that the spacer-spacer interactions can contribute significantly to the binding strength. The chelate cooperativity factors calculated with the DFT-based approach have the right trend and predict a positive chelate cooperativity as found in the experiment. However, the DFT-based approach yielded far too large chelate cooperativity factors due to the logarithmic connection of $\Delta G_{\mathrm{A}}^{\mathrm{sol}}$ to the assocaition constant. The average deviation of the calculataed $\Delta G_{\mathrm{A}}^{\mathrm{sol}}$ to the experiment is 7.4 kJ mol^{-1}. This is higher than in the other investigations but can still be seen as state-of-the-art.[70] For the accurate determination of cooperativity factors the deviation of the quantum chemical based approach has to achieve probably an average error of below 1 mol^{-1}.

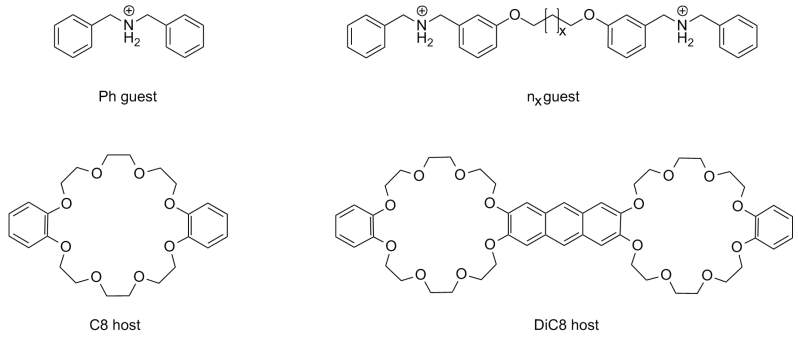

Figure 3.10: Structural formula of the mono- and divalent guest and host molecules investigated in **Paper A4** with $x = 0$, 1 or 2. Included with permission from Achazi *et al.*[154] (©2015 Achazi *et al.*; licensee Beilstein-Institut).

In **Paper A5** again a divalent 24-crown-8 host was considered. The host was used to control the photoisomerisation of a divalent guest. The divalent *E*-configured azobenzene guest **E-1** can undergo a light-induced *E*-*Z* isomerisation (Figure 3.11). Experiments by the AG Hecht and AG Schalley show that this photoisomerisation can be effectively inhibited by adding the complementary host **3**, which yields the complex **E-1@3**. Adding a base deprotonates **E-1**, which dissolves the **E-1@3** complex and reestablishes photochromism of the guest **E-1**. The process is reversible. Reprotonating the guest with acid will yield **E-1@3** again and switches "off" the *E*-*Z* photoisomerisation. To gain further insights the stilbene guests **E-2** and **Z-2** have been investigated, too.

The DFT-based calculated Gibbs energies of association $\Delta G_{\mathrm{A}}^{\mathrm{sol}}$ were in good agreement

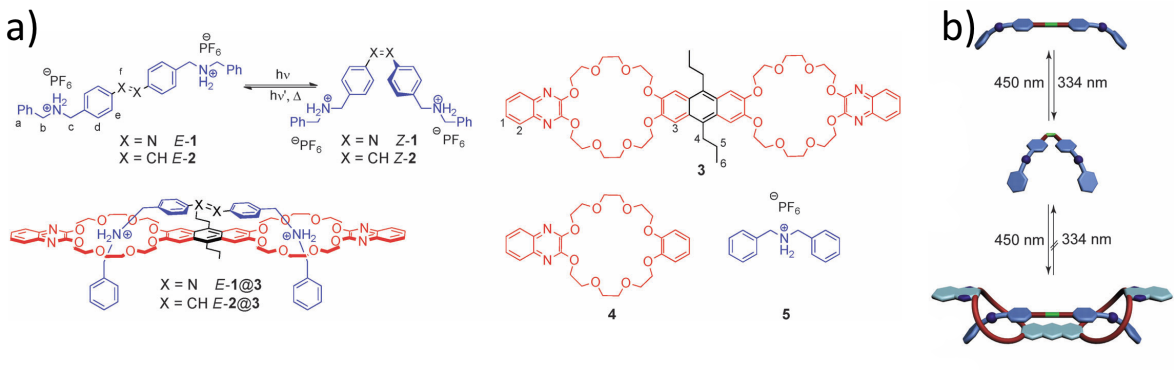

Figure 3.11: a) Structural formula of the mono- and divalent guest and host molecules investigated in **Paper A5** with $x = 0$, 1 or 2. Included with permission from Lohse *et al.*[160] (©2015 The Royal Society of Chemistry)

with the experiments (within 2 kJ mol^{-1} for **E-1@3**). **Z-1@3** and **Z-2@3** should not be formed according to the calculation and are not found in the experiment. The calculated $\Delta G_{\mathrm{A}}^{\mathrm{sol}}$ value of **E-2@3** seems reasonable but the solubility **E-2** was not sufficient for the ITC experiments. This study shows how we can construct a controllable photoswitch, in other words, a system in which the photochromism can be switched "on" (unbound guest) and "off" (guest bound to host). On the other side, the guest in the *Z*-configuration cannot bind to the host. Therefore, this system is also a model for photocontrolled association processes.

In **Paper A6** the binding motif of a 24-crown-8 ether with a secondary ammonium ion was used again to construct a divalent crown ether/ammonium pseudo[2]rotaxane but with a different application. The pseudo[2]rotaxane **P-1** (Figure 3.12) was built with spacers that form an electron-donor-acceptor complex. The host has a tetrathiafulvalene spacer, which serves as donor, and the guest has a naphthalene diamide spacer, which serves as acceptor. The investigation demonstrates that divalent crown ether/ammonium pseudo[2]rotaxane **P-1** is a straightforward tool to bring the donor and acceptor unit in close proximity to each other. These donor-acceptor complexes display partial π-electron sharing, which result in an intramolecular charge transfer measurable with UV/Vis spectroscopy. Assembly and disassembly of the pseudo[2]rotaxane **P-1** is fully reversible and controllable by adding chemical stimuli (like $NaBArF_{24}$). As a result, the intramolecular charge transfer properties can be switched on and off. Isothermal titration calorimetry (ITC) and a double mutant cycle (DMC) analysis showed a positive chelate cooperativity.

The pseudo[2]rotaxane **P-1** can be capped to the [2]rotaxane **R-1**. AG Schalley in-

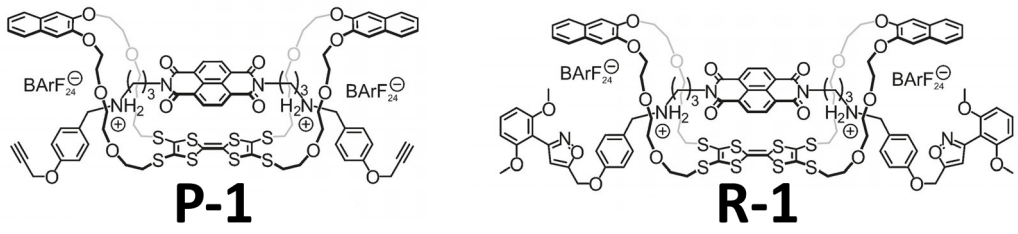

Figure 3.12: Pseudo[2]rotaxane **P-1** and [2]rotaxane **R-1**. Included with permission from Schröder *et al.*[161] (©2017 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim)

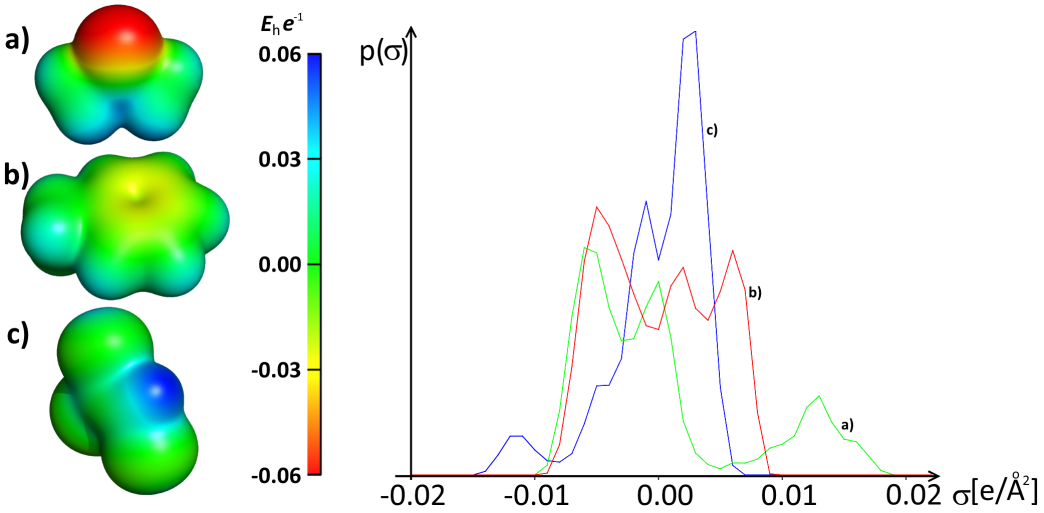

Figure 3.13: Left: electrostatic potential maps of a) acetone, b) toluene and c) trichloroethylene. Right: the σ-profiles of a) acetone (green), b) toluene (red) and c) trichloroethylene (blue).

vestigated the optoelectronic properties of **R-1** by UV/Vis and cyclic voltammetry experiments and could reveal **R-1** is redox-switchable and stable in five different redox states. The DFT calculations gave further insights into the molecular and electronic structure. The HOMO-LUMO-gap calculated with B3LYP/def2-TZVP and the implicit solvent model COSMO was in good agreement with the observed intramolecular charge transfer band. The ICT band was measured in several different solvents. It shows no connection to the relative permittivity ε_r of the solvents. However, the ICT band can be set in relation to the surface charge of the solvent. Trichloroethylene has an excess of positive surface charge (Figure 3.13) and no negative surface charge to counterbalance. Therefore, trichloroethylene can only stabilize molecules with negative surface charge,

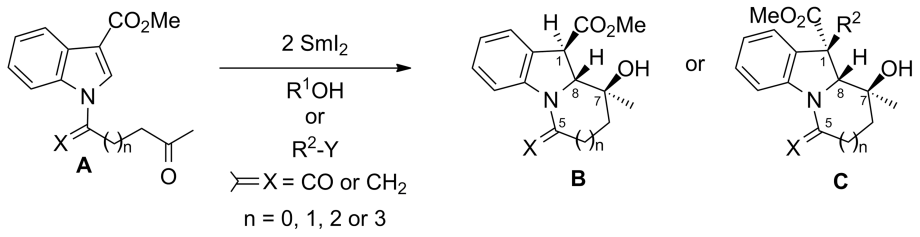

Figure 3.14: The SmI_2 mediated reductive coupling of *N*-oxoalkyl-substituted methyl indole-3-carboxylates investigated in **Paper A7**. Included with permission from Achazi *et al.*[162] (©2017 Wiley Periodicals, Inc).

but can not stabilize molecules with positive surface charge. Acetone is the exact opposite of trichloroethylene, what can be visualized with the σ-profiles of these solvents, as shown in Figure 3.13. It appears that solvent molecules (like trichloroethylene and dichloromethane) which can only stabilize negative surface charges lead to the highest wavelength for the ICT band of **R-1**, and solvent molecules (like dimethyl sulfoxide and acetone) which only stabilize positive surface charges lead to the lowest wavelength. Toluene displays no surface areas which have a highly positive or negative charge (compared to acetone and trichloroethylene). It has only a small stabilizing effect on highly negative or positive surface charges. The ICT band of **R-1** in toluene is situated between trichloroethylene and acetone. The solvents terahydrofuran, ethyl acetate and acetonitrile have σ-profiles similar to acetone, and these solvents shift the ICT band of **R-1** to a lower wavelength like acetone. Isopropanol is the only solvent in this study which has due to its hydroxyl group both, surface areas with highly positive and with highly negative charge. Thus, the wavelength of the ICT band of **R-1** should be situated between trichloroethylene and acetone as for toluene. Indeed the wavelength of the ICT band measured in isopropanol is similar to the wavelength measured in toluene. However, this is just a rough model to predict the solvent effects on the ICT band of **R-1**. To gain further insights, the electronic structures of the (photo-)excited states of **R-1** have to be calculated as well, which is beyond the scope of this thesis.

After investigating organic host-guest complexes with crown ether/ammonium binding motif in solution, we applied our DFT-based approach for a mechanistic investigation involving organic f-element complexes in solution. More precisely, we investigated the samarium diiodide (SmI_2) mediated reductive coupling of *N*-oxoalkyl-substituted methyl indole-3-carboxylates in tetrahydrofuran (THF) depicted in Figure 3.14. The reductive coupling yields benzannulated pyrrolizidines and indolizidines surprisingly with high di-

astereoselectivity at the adjacent positions C-1, C-7 and C-8. Based on our knowledge obtained in determining Gibbs energies in solution, in **Paper A7** we developed a detailed model for the reaction mechanism of the SmI_2 mediated reductive coupling. For this purpose, we determined the Gibbs energies for the intermediates of possible reaction pathways (Figure 3.15) at ambient temperature. The electronic energy was determined with a dispersion corrected density functional PBE0-D3(BJ) together with a small-core pseudopotential for samarium. Thus, the f shell was explicitly included. The solvent effects were incorporated by the implicit D-COSMO-RS method. It was crucial for the prediction of the molecular structure of the reaction intermediates to include solvent molecules explicitly, too. Translational, rotational and vibrational contributions were included similar to the previous studies (**Paper A1**-**A5**). In contrast to the two mechanisms which were proposed before,[163,164] our results predict the formation of an energetically highly favorable chelate complex (**4a** in Figure 3.15) in which the experimentally found product configuration at positions C-7 and C-8 is preformed. The corresponding chelate complex (**5a** in Figure 3.15) with the "wrong" configuration preformed is about 50 kJ mol^{-1} higher in energy. Thus, the reaction will follow the path indicated with green arrows in Figure 3.15. Based on experimental results, it was proposed[163] that the configuration at position C-1 for the protonation products is under thermodynamic control. Our calculations confirm this assumption.

The final but very interesting project (published in **Paper B1**) is a side project dedicated to the investigation of the stretching of of polyethylene glycol (PEG) in water with single-molecule atomic force microscope (AFM) experiments (AG Rabe), molecular dynamics (MD) simulations (AG Netz) and DFT. The PEG molecule was stretched in the sense that the ends were pulled apart. It was expected that the stretching would reduce the total entropy, because the number of available conformations decreases. Surprisingly, the investigations show that in case of high stretching the gain in entropy for the solvent molecules (water) equalizes almost exactly the decrease in conformational entropy of the PEG molecule. In total, the stretching mainly results in a decrease of enthalpy. Therefore, PEG can be considered an energetic spring and not an entropic spring at high forces. The finding can be explained with MD simulations. The oxygen atoms in the relaxed PEG are predominantly in *gauche* conformation (Figure 3.16) with one water molecule as bridge between each pair of adjacent oxygen atoms. Stretching the PEG turns the oxygen atoms in *anti* conformation. In this conformation each water molecule binds only with one hydrogen atom to the PEG. On the one hand this yields a gain in entropy

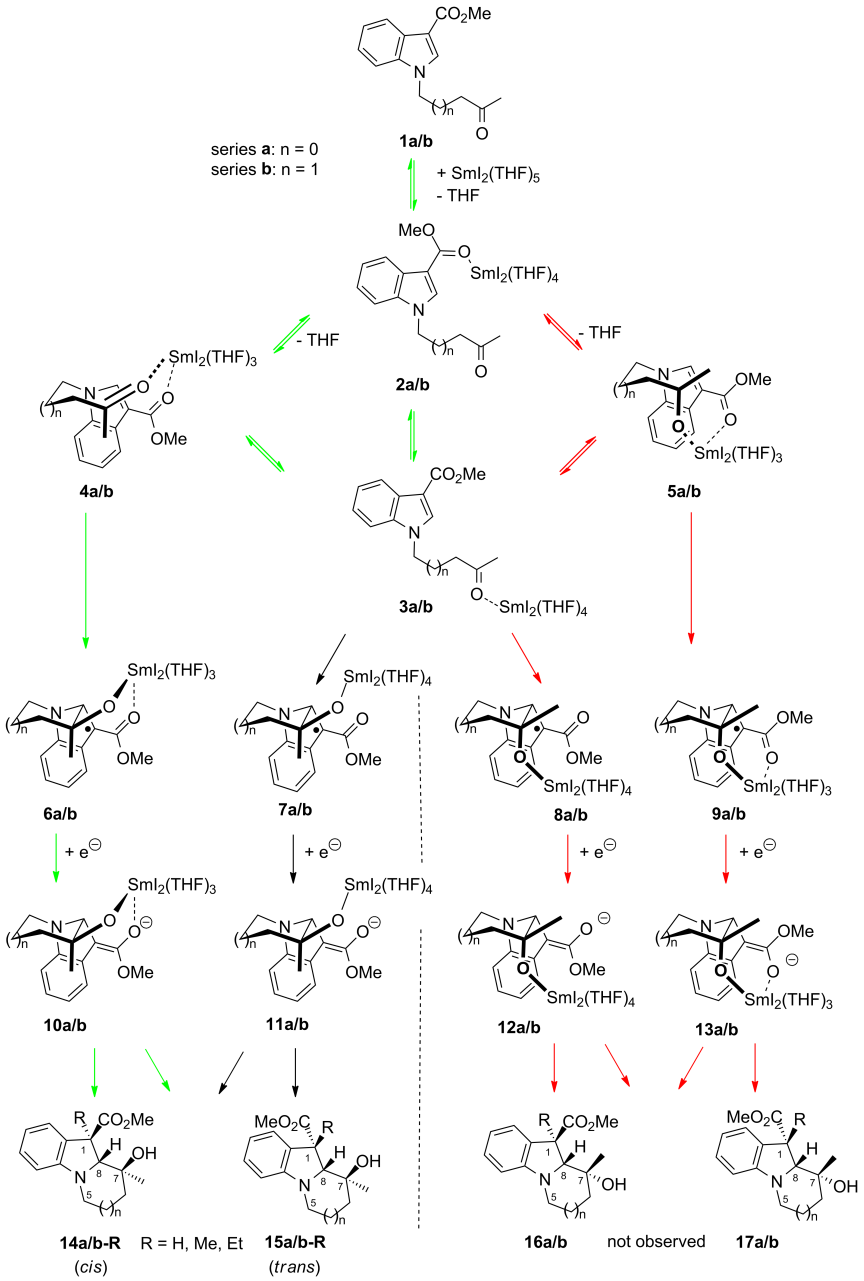

Figure 3.15: The investigated reaction pathways for the SmI_2 mediated reductive coupling of the *N*-oxoalkyl-substituted methyl indole-3-carboxylates **1a** and **1b**. Our results predict the reaction will follow the reaction pathway with green arrows. The reaction pathway with black arrows is possible but according to our determination unlikely to happen. The reaction pathway with red arrows yields the diastereomers which have not been found in experiments. Additionally these reaction pathways are also unlikely to happen based on our determination. Included with permission from Achazi *et al.*[162] (©2017 Wiley Periodicals, Inc).

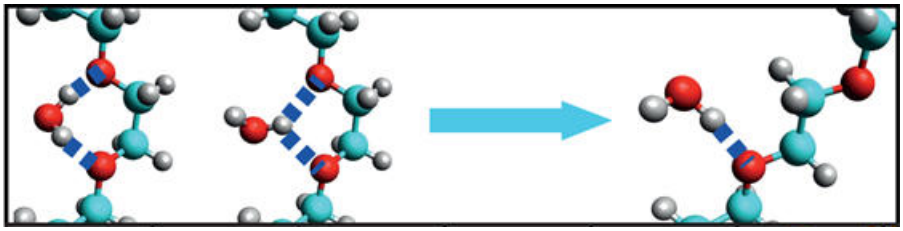

Figure 3.16: Depiction of a *gauche* conformation in PEG with a water bridge (left), a *gauche* conformation in PEG with only one hydrogen atom of the water molecule invloved in the hydrogen bond (center) and an *anti* conformation in PEG with one hydrogen bond to one water molecule. Included with permission from Liese *et al.*[165] (©2017 American Chemical Society).

for the water molecules, on the other hand the enthalpy decreases to a much larger extent. In the high-force limit (stretching force larger than 375 pN), the experimental data are in perfect agreement with the stretching force determined with DFT in the gas phase at zero temperature. This not only demonstrates that at the high-force limit conformational fluctuations and solvent effects are negligible, but it also allows to determine the vertical anchoring position in the AFM experiments. This yielded a considerable correction to the measured PEG length.

Summarized, this investigation is a case example for the counterbalancing interplay of conformational and solvent entropy. These insights are important for understanding the low solubility of PEG in water at high temperatures. The water bridges are enthalpically favorable but entropically unfavorable. Thus, they are only stable below a certain temperature. Future investigations should focus on this temperature dependence.

4 Conclusion and Outlook

This thesis presents a multilevel DFT-based approach for the accurate calculation of the Gibbs energy of association for supramolecular crown ether/ammonium host-guest systems. Dispersion corrected density functional theory was used to obtain the electronic energy. Solvent effects were included with the implicit solvent models COSMO-RS and D-COSMO-RS. The effects of translation, rotation and vibration were taken into account by a RRHO approximation with free-rotor approximation for low-lying vibrations. Crucial was the explicit consideration of the counterions. The average deviation from the experimental results of all presented investigations in this thesis combined was about 5 $\mathrm{kJ\,mol^{-1}}$, which is state-of-the-art.[70] In particular better models for the determination of the Gibbs energies of solvation and the vibrational contribution should be developed to achieve a higher accuracy and to replace experiments. This approach and further DFT-based data, for example molecular structures, were used to investigate 18-crown-6/ammonium complexes and 24-crown-8/ammonium (pseudo-)rotaxanes in a joint study with specialists for supramolecular chemistry (AG Schalley) and for molecular dynamic simulations (AG Weber). Our results can be combined to a "guideline" for highly stable supramolecular host-guest structures:

- Host and guest should have the best possible complementarity.
- Preorganization and adaptability should be both maximized, instead of maximizing the preorganization by lowering the adaptability.
- Favorable spacer–spacer interactions should be maximized.
- The preorganization of a flexible guest depends on its size, its ability to form intramolecular bonds and the solvent. Thus, the preorganization of a flexible guest can and should be optimized.
- The solvent and the counterions effect the association process and have to be optimized.

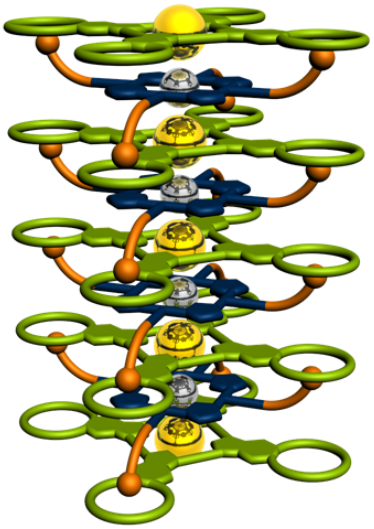

Figure 4.1: Schematic representation of a self-assembled[166] molecular fiber consisting of tetravalent host and guest molecules with the crown ether/ammonium binding motif. Special thanks to Larissa K. S. von Krbek for creating the graphic.

With this guideline in mind, more fuctionalized systems can be developed. Two examples are already presented in this thesis. A controllable photoswitch and an electron-donor-acceptor complex were constructed based on divalent 24-crown-8/ammonium pseudo[2]rotaxanes. The future goal will be the construction of self-assembling molecular fibers (Figure 4.1) consisting of host and guest molecules with crown ether/ammonium binding motifs and metallo-porphyrins as spacer in the center of these molecules. In case the fibers are able to conduct electricity, the fibers could be used as wires. An advantage of such a wire would be that its functionality could be switched "on" and "off" as in the switchable systems presented in this thesis or more sophisticated as described by Zhang *et al.*[43] AG Hecht and AG Schalley have already started with the preliminary investigations[167] like other groups,[47,48] too. Particularly interesting could be the use of ruthenium(II)-porphyrin as spacer. It is known from experimental research[168] that ruthenium(II)-porphyrins form dimers connected with a metal-metal bond, which are stable in solution. Additionally preliminary investigations performed by Stefan Mattsson from the AG Paulus showed that stacks of ruthenium(II)-porphyrin exhibit a particularly high binding energy, short ring distance and small band gap compared to other metallo-porphyrin stacks like a zinc(II)-porphyrin stack.

In the last step the multilevel DFT-based approach was modified for a mechanistic investigation of the SmI_2 mediated reductive coupling of *N*-oxoalkyl-substituted methyl indole-3-carboxylates. This enabled the development of a detailed model for the reaction mechanism, which explains the high diastereoselectivity of the reaction. It is now on experimental chemists to verify the proposed reaction mechanism. The smallest of the next steps on the theoretical side is applying this approach to the investigation of other SmI_2 mediated reductive coupling reactions like the formation of hexahydroisoquinoline from α-amino-ketones.

The side project with AG Rabe and AG Netz revealed that the flexible molecule PEG

acts as an energetic spring and not as an entropic spring at high forces. This knowledge can be used for future investigations of the solubility of PEG in water at high temperatures. However, these results have an additional consequence. If a multivalent host or guest in water with a flexible PEG spacer has to turn into the linear conformation (all oxygen atoms in *anti* conformation) in order to form a complex, the flexible PEG spacer would cause an enthalpic penalty and not an entropic penalty. This connects the investigation of PEG with the investigation of crown ether/ammonium assemblies. In case of the crown ether/ammonium assemblies, the flexibility of the spacer was not connected to the entropic penalty; instead, enthalpic spacer-spacer interactions played an important role. Both challenge the notion that flexible spacers in general cause an entropic penalty in complex formations. Spacer length, intramolecular interactions, solvent and counterions can affect the conformational space of both the flexible spacer and the whole system or have strong enthalpic effects. This thesis demonstrates how important joined experimental and computational investigations are to elucidate these effects and it presents powerful *in silico* methods to identify these effects.

5 Published Papers

5.1 Paper A1

"Theoretical and experimental investigation of crown/ammonium complexes in solution"
Andreas J. Achazi, Larissa K. S. von Krbek, Christoph A. Schalley and Beate Paulus
Journal of Computational Chemistry **2016**, **37**, 18-24.
DOI: 10.1002/jcc.23914
URL: http://dx.doi.org/10.1002/jcc.23914

Contributions

I did the development of the multilevel DFT-based approach to calculated the Gibbs energy for the formation of charged organic complexes in solvents and its application. The experimental investigations were done by Larissa K. S. von Krbek. The results were discussed by Larissa K. S. von Krbek and me. We wrote the manuscript together, the general concept as well as the main contributions were done by myself. All authors added contributions to the final version of the publication.

5.2 Paper A2

"Allosteric and chelate cooperativity in divalent crown ether/ammonium complexes with strong binding enhancement"
Larissa K. S. von Krbek, Andreas J. Achazi, Marthe Solleder, Marcus Weber, Beate Paulus and Christoph A. Schalley
Chemistry – A European Journal **2016**, **22**, 15475–15484.
DOI: 10.1002/chem.201603098
URL: http://dx.doi.org/10.1002/chem.201603098

Contributions

The experimental investigations were done by Larissa K. S. von Krbek and the classical molecular dynamics simulations were done by Marthe Solleder. I adjusted the multilevel DFT-based approach for the doubly charged species and then applied the new approach. The results were discussed by Marthe Solleder, Larissa K. S. von Krbek and me. Larissa K. S. von Krbek and I wrote the manuscript together, the main contributions were done by Larissa K. S. von Krbek. All authors added contributions to the final version of the publication.

5.3 Paper A3

"The delicate balance of preorganisation and adaptability in multiply bonded host–guest complexes"
Larissa K. S. von Krbek, Andreas J. Achazi, Stefan Schoder, Marius Gaedke, Tobias Biberger, Beate Paulus and Christoph A. Schalley
Chemistry – A European Journal **2017**, **23**, 2877–2883.
DOI: 10.1002/chem.201605092
URL: http://dx.doi.org/10.1002/chem.201605092

Contributions

The experimental investigations were done by Larissa K. S. von Krbek with assistance from Stefan Schoder, Marius Gaedke and Tobias Biberger. I did the DFT calculations. The results were discussed by Larissa K. S. von Krbek and me. Larissa K. S. von Krbek and I wrote the manuscript together, the main contributions were done by Larissa K. S. von Krbek. All authors added contributions to the final version of the publication.

5.4 Paper A4

"First principle investigation of the linker length effects on the thermodynamics of divalent pseudorotaxanes"
Andreas J. Achazi, Doreen Mollenhauer and Beate Paulus
Beilstein Journal of Organic Chemistry **2015**, **11**, 687–692.
DOI: 10.3762/bjoc.11.78
URL: http://dx.doi.org/10.3762/bjoc.11.78

Contributions

I did the DFT calculation and adjustment of the multilevel DFT-based approach. Doreen Mollenhauer did pre-optimizations with DFT for some of the molecules. The manuscript was written in cooperation with Beate Paulus. All authors added contributions to the final version of the publication.

BEILSTEIN JOURNAL OF ORGANIC CHEMISTRY

First principle investigation of the linker length effects on the thermodynamics of divalent pseudorotaxanes

Andreas J. Achazi[1], Doreen Mollenhauer[2] and Beate Paulus[*1]

Full Research Paper

Open Access

Address:
[1]Institut für Chemie und Biochemie, Freie Universität Berlin, Takustr. 3, 14195 Berlin, Germany and [2]Physikalisch-Chemisches Institut, Justus-Liebig-Universität Gießen, Heinrich-Buff-Ring 58, 35392 Gießen, Germany

Email:
Beate Paulus[*] - b.paulus@fu-berlin.de

* Corresponding author

Keywords:
density functional theory (DFT); dispersion correction; Gibbs energy; pseudorotaxanes; solvent effects; COSMO-RS

Beilstein J. Org. Chem. **2015,** *11,* 687–692.
doi:10.3762/bjoc.11.78

Received: 06 March 2015
Accepted: 29 April 2015
Published: 08 May 2015

This article is part of the Thematic Series "Multivalency as a chemical organization and action principle".

Guest Editor: R. Haag

Abstract

The Gibbs energies of association (Gibbs free (binding) energies) for divalent crown-8/ammonium pseudorotaxanes are determined by investigating the influence of different linkers onto the binding. Calculations are performed with density functional theory including dispersion corrections. The translational, rotational and vibrational contributions are taken into account and solvation effects including counter ions are investigated by applying the COSMO-RS method, which is based on a continuum solvation model. The calculated energies agree well with the experimentally determined ones. The shortest investigated linker shows an enhanced binding strength due to electronic effects, namely the dispersion interaction between the linkers from the guest and the host. For the longer linkers this ideal packing is not possible due to steric hindrance.

Introduction

If two or more binding sites of a molecular system are involved in the association process, the interaction energy can be significantly increased compared to the sum of the individual binding energies. This effect is called multivalency [1] and is mainly observed in biochemical systems [2-9]. But the concept of multivalency can be transferred to supramolecular assemblies with suitable building blocks [10-12] including (pseudo)rotaxanes [13-15] as well. One common building block for pseudorotaxanes is the crown/ammonium binding motif. In this motif ammonium can bind on top of small crown ethers, e.g., crown-6, or can pass through larger crown ethers, e.g., crown-8. Jiang et al. [16] have investigated the assembly thermodynamics and kinetics of divalent crown-8/ammonium pseudorotaxanes with different linkers. The shortest linker shows a much larger chelate cooperativity than the longer linkers due to non-innocent linkers that contribute to the binding. To analyze the individual contributions to the binding, we perform first principle calculations of the model system shown in Figure 1, which

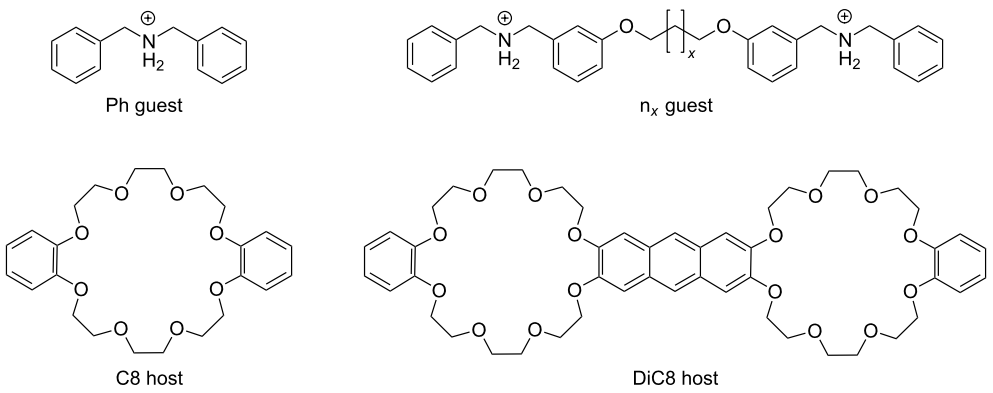

Figure 1: Structures of the mono- and divalent guest and host molecules. The linker in the divalent guest molecule is varied with x = 0, 1 or 2.

is strongly related to the experimentally investigated systems of Jiang et al. [16]. The only difference is that 1,4-diazanaphthalene groups of the host molecule are replaced by phenyl groups and the side chains of the anthracene bridge in the divalent host are neglected. In addition to the electronic contributions, enthalpic and entropic temperature effects as well as solvent effects are included in our simulations in order to compare to experimentally obtained Gibbs energy of association.

Results and Discussion

In order to investigate the cooperativity effects of the binding between divalent host molecules and divalent guest molecules it is important to firstly describe the monovalent binding motif computationally as accurately as possible and to understand the underlying effects that contribute to the binding. Three major terms have to be considered in the evaluation of the Gibbs energy of association ΔG to model the reaction in solution at finite temperature with reasonable accuracy. 1) The electronic association energy ΔE is calculated [17] with the DFT functional TPSS-D3(BJ) [18-20] and the basis set def2-TZVP [21,22]. A comparison with the electronic association energy determined with the DF-LCCSD(T) method [23,24] at the extrapolated basis set limit shows good agreement (see Table 1). Already the DF-LCCSD(T) with the cc-pVTZ basis set deviates only by 5% from the TPSS-D3(BJ) value, whereas the basis set extrapolated value is more or less equivalent to the TPSS-D3(BJ) value (deviation less than 0.3%). This very good agreement is somewhat fortunate, because a basis set extrapolation with DZ and TZ is only accurate to within a few percent. Additionally, the possible errors of the functional and the dispersion correction can also be in the range of 10% for the system under investigation. A more detailed analysis of the accuracy of the TPSS-D3(BJ) functional has been performed for the crown-6/ammonium complex in [25]. Another point to remark is that even for the monovalent system about 36% of the electronic interaction energy is due to the dispersion correction. 2) The finite temperature effects from translation, rotation and vibration are calculated with an approach from Grimme [26], which partially treats the low-lying vibrations as hindered rotations (TPSS-D3(BJ)/def2-SVP [22,27] for vibrations). 3) The influence of the solvent for the association process in solution is derived from the difference of the solvation effects of the product and the reagents, calculated with COSMO-RS [28,29]. For the COSMO-RS (BP_TZVP_C30_1301.ctd parameterization) calculation all structures have been optimized in an ideal conductor [30] and in vacuum with BP86/def-TZVP [31-34]. This procedure yields very good results for the Gibbs energy of association in the case of the crown-6/ammonium complex in comparison with experiment [25]. For the simulations of the crown-8/ammonium systems the same solvent as in the experiment [16] is used, namely a 2.2:1 mixture of chloroform/acetonitrile. The influence of the counter ion PF_6^- onto the Gibbs energy of association is taken into account explicitly.

Table 1: Electronic association energy ΔE for Ph@C8*.[a]

system	method	ΔE (kJ/mol)
Ph@C8*	TPSS/def2-TZVP	−134.9
Ph@C8*	TPSS-D3(BJ)/def2-TZVP	−210.5
Ph@C8*	DF-LCCSD(T)/cbs(DZ-TZ)	−210.0
Ph@C8*	DF-LCCSD(T)/cc-pVDZ	−174.7
Ph@C8*	DF-LCCSD(T)/cc-pVTZ	−199.9

[a]ΔE calculated at TPSS-D3(BJ)/def2-TZVP level of theory is not identical to the one in Table 2, because there another conformer (a slightly more stable one) is used. The Ph@C8* structure has been optimized with TPSS-D3(BJ)/def2-TZVP. For the other methods only single point calculations are done.

The divalent host molecules consist of two crown-8 ethers that are linked by an anthracene bridge. For the divalent guest molecule different flexible linkers, namely $–O(CH_2)_2O–$ (n_0), $–O(CH_2)_3O–$ (n_1) and $–O(CH_2)_4O–$ (n_2) have been investigated both experimentally in [16] and computationally. The results for the electronic association energy ΔE, the Gibbs energy of association ΔG in the gas phase and its enthalpic (ΔH) and entropic ($-T\Delta S$) contributions are given in Table 2. Comparing the electronic association energy for the n_0 guest in the divalent case with the doubled value of the monovalent (Ph@C8) system, an electronic cooperativity effect of 9.7 kJ/mol is discovered. When the linker length is increased, this electronic cooperativity effect is lost, and a lower electronic association energy (by 11.3 kJ/mol) is discovered for the divalent system with the n_1 linker compared to two monovalent systems. For the longer n_2 linker the electronic association energy is even lower by 24.2 kJ/mol for the divalent system compared to two monovalent systems. This is mainly due to the dispersive interaction of the linking unit (two phenyl rings and the linker), which in case of the n_0 guest fits perfectly on top of the anthracene linker of the DiC8 host. The distance between the linker of the host and the linker of the n_0 guest is around 3.7 Å, quite close to an ideal distance for the π–π stacking of two benzene rings. The n_1 and n_2 guest do not perfectly fit with the host (Figure 2). In the n_1-case the linker is folded away from the anthracene bridge, and for the n_2-case one phenyl ring is twisted away due to steric constraints.

Table 2: Electronic association energy ΔE and Gibbs energy of association ΔG in the gas phase at room temperature (T = 298.15 K).[a]

system	ΔE (kJ/mol)	ΔG (kJ/mol)	ΔH (kJ/mol)	$-T\Delta S$ (kJ/mol)
Ph@C8	−215.6	−130.2	−204.8 (+10.9)	+74.6
n_0@DiC8	−440.9	−339.3	−422.6 (+18.3)	+83.3
n_1@DiC8	−419.9	−317.5	−402.6 (+17.3)	+85.2
n_2@DiC8	−407.0	−299.8	−386.8 (+20.2)	+87.0

[a]The enthalpic (ΔH) and entropic ($-T\Delta S$) contribution to ΔG are given. The ΔH contribution resulting from finite temperatures is given in brackets.

The Gibbs energy of association ΔG in the gas phase of the divalent systems (Table 2) result in the same trend as observed for the electronic association energy ΔE, because the enthalpic (ΔH) and entropic ($-T\Delta S$) contributions are similar for n_0@DiC8, n_1@DiC8 and n_2@DiC8.

In Table 3, the Gibbs energies of association in solution with and without counter ion are compared to the calculated electronic association energies, Gibbs energies of association in the gas phase and to the measured experimental values. For the monovalently bound system Ph@C8 the computationally obtained value of ΔG (−12.6 kJ/mol) agrees well with the

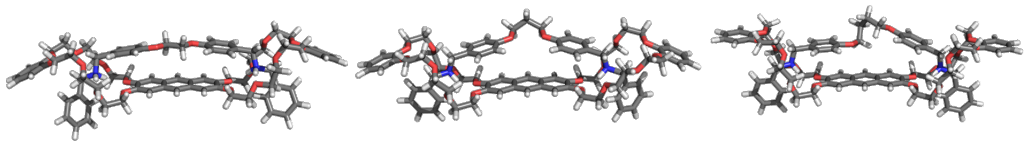

Figure 2: Optimized gas phase structures (TPSS-D3(BJ)/def2-TZVP) of the divalent complexes n_0@DiC8, n_1@DiC8 and n_2@DiC8.

Table 3: Gibbs energy of association ΔG in solution.[a]

system	ΔE (kJ/mol)	ΔG gas phase (kJ/mol)	ΔG solution (kJ/mol)	ΔG counter ion (kJ/mol)	ΔG experiment (kJ/mol)
Ph@C8	−215.6	−130.2	−1.1	−12.6	−15.0
n_0@DiC8	−440.9	−339.3	−42.5	−44.3	−25.1
n_1@DiC8	−419.9	−317.5	−24.2	−28.9	−17.4
n_2@DiC8	−407.0	−299.8	−11.5	−15.3	−16.2

[a]Electronic association energy ΔE, Gibbs energy of association ΔG in gas phase and in solution, in the latter case with and without inclusion of the counter ion PF_6^-, and experimentally determined ΔG for monovalent and divalent pseudorotaxanes in a 2.2:1 solvent mixture of chloroform/acetonitrile at room temperature (T = 298.15 K) are presented.

experimentally determined value (−15.0 kJ/mol). The Gibbs energies of association in gas phase and the Gibbs energies of association in solution show similar differences between n_0@DiC8, n_1@DiC8 and n_2@DiC8 as the electronic association energies. Hence, the dependence on the linker length is of electronic origin and not affected by temperature or solvent effects. Including the counter ion in the determination of ΔG has a much weaker effect in the divalent case compared to the monovalent one, because the guest molecule is larger and the positive charge of the amide group can be distributed better over the molecule. For the divalent pseudorotaxanes the absolute agreement between the calculated and the experimentally determined Gibbs energies is not as good as in the case of monovalent binding, but the same trends are observed in the simulations as in experiment. The divalent pseudorotaxane with the n_0 linker shows a significantly stronger binding than the longer molecules.

Additionally, the full double mutant cycle from [16] has been calculated (Figure 3 and Table 4). The Gibbs energy of association ΔG in case of Ph@DiC8 and n_0@2C8 is in good agreement with the experimental data. For 2Ph@DiC8 and n_0@C8 the deviation is larger just as for the divalent systems in Table 3. This deviation strongly affects the calculated equilibrium constants K, because ΔG is included exponentially in K. Therefore only a qualitatively discussion of the equilibrium constants is possible. With the determined equilibrium constants K, the effective molarity EM can be calculated [16]:

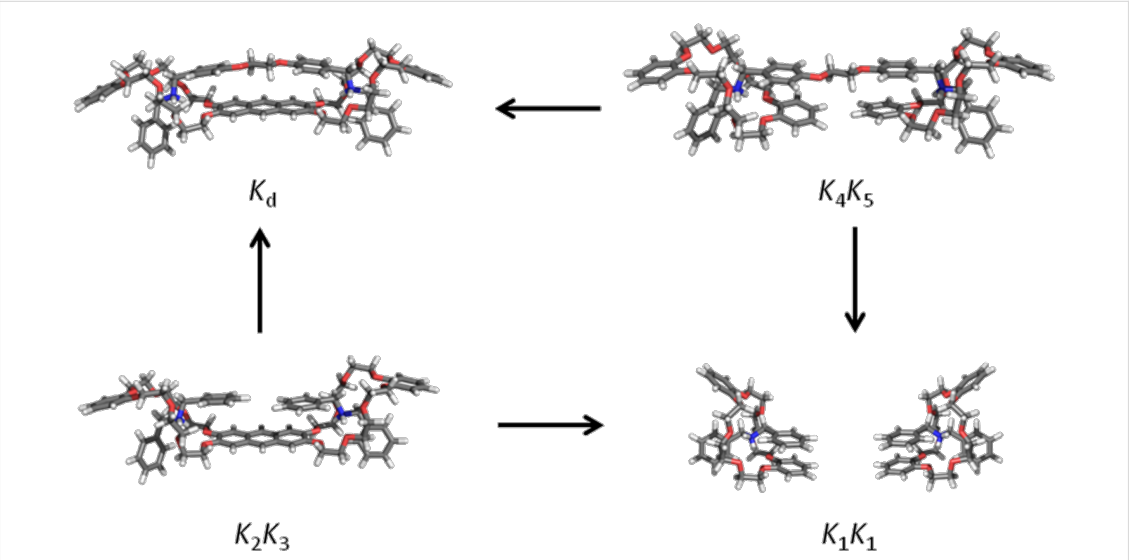

Figure 3: Double mutant cycle for n_0@DiC8. The K variables are declared in Table 4 and are used in Equation 1. Top left: n_0@DiC8, top right: n_0@2C8, bottom left: 2Ph@DiC8 and bottom right: two Ph@C8. The figures show the optimized gas phase structures.

Table 4: Gibbs energy of association ΔG in solution (2.2:1 chloroform/acetonitrile, 298.15 K) and equilibrium constant K for the systems from the double mutant cycle.[a]

system	ΔG counter ion (kJ/mol)	K ($mol^{-1}\cdot L^{-1}$)	#K	ΔG experimental (kJ/mol)	K experimental ($mol^{-1}\cdot L^{-1}$)
Ph@C8	−12.6	161.2	K_1	−15.0	420
Ph@DiC8	−16.2	677.8	K_2	−16.4	735
2Ph@DiC8	−5.11	7.9	K_3	−12.3	145
n_0@C8	+1.4	0.6	K_4	−16.3	714
n_0@2C8	−13.8	261.6	K_5	−13.3	220
n_0@DiC8	−44.3	57679927.3	K_d	−25.1	25000
n_1@DiC8	−28.9	115627.5	K_d	−17.4	1100
n_2@DiC8	−15.3	479.1	K_d	−16.2	700

[a]The effects of the counter ion PF_6^- are included in the calculation. #K declares the equilibrium constant K with regard to Equation 1 and Figure 3.

$$\mathrm{EM} = \frac{K_d K_1^2}{2K_2K_3K_4K_5}, \quad (1)$$

$$K = e^{\frac{-\Delta G}{RT}}. \quad (2)$$

According to Hunter and Anderson [35] EM·K_1 can be used to quantify cooperativity. If EM·$K_1 \approx 1$, the system shows no or small cooperativity, if EM·$K_1 >> 1$ the systems shows positive cooperativity and for EM·$K_1 << 1$ the opposite occurs. The data for the EM·K_1 values are all based on the double mutant cycle of n_0, because the experimental data are also using only the double mutant cycle of n_0 for n_1 and n_2. The experiment shows that n_0@DiC8 (EM·K_1(exp.) = 55.3) has a highly positive cooperativity while n_1@DiC8 (EM·K_1(exp.) = 2.4) and n_2@DiC8 (EM·K_1(exp.) =1.5) have no significant cooperativity. In contrast to the experiment, the calculations show that n_0@DiC8 (EM·K_1(cal.) = $1.6\cdot10^8$), n_1@DiC8 (EM·K_1(cal.) = $3.1\cdot10^5$) and n_2@DiC8 (EM·K_1(cal.) = $1.3\cdot10^3$) have highly positive cooperativity, but all calculated values are much too high compared to experiment due to the deviations of ΔG for 2Ph@DiC8 and n_0@C8. Despite these errors the calculation shows in agreement to experiment, that n_0@DiC8 has a much higher EM·K_1 value than n_1@DiC8 and n_2@DiC8. So the calculations confirm that the linkers contribute to the binding strength in the divalent pseudorotaxanes and can be called non-innocent as in [16].

Regarding the aforementioned deviations from experiment, the difference in the absolute Gibbs energies of association can be explained by the insufficient modeling of solvent effects. The solvent model assumes a uniform distribution of the two different solvents in the mixture. An explicit treatment of at least some solvent molecules would be desirable but is computationally not feasible at the required quantum mechanical level. A combined molecular mechanics/quantum mechanics treatment could be a solution to this problem in the future. Nevertheless, concerning the difference between ΔG in the gas phase and the experimental value, the solvent model that is used in this study yields a significant part of ΔG, but it cannot resolve details of the solvation effects.

At the end of this discussion it is worth mentioning that the most stable structure of the host molecule changes from gas phase to solution. Both the monovalent and the divalent host have a folded ground state structure the in gas phase (Figure 4). The electronic energy ΔE that is needed for unfolding the monovalent host is 29.7 kJ/mol. This value increases up to 72.3 kJ/mol for fully unfolding the divalent host (52.6 kJ/mol for the first step and 19.6 kJ/mol for the second step). In solution (2.2:1 chloroform/acetonitrile, 298.15 K) the monovalent host is more stable in the unfolded form with ΔG being 8.2 kJ/mol lower than that of the folded form. The divalent host stays in the folded structure, and ΔG is 6.5 kJ/mol lower than that of the unfolded form.

Conclusion

The Gibbs energies of association, including enthalpic and entropic temperature effects, solvent effects and the counter ions, have been determined for the divalent crown-8/ammonium pseudorotaxane with different linkers in the guest molecule. Additionally, a full double mutant cycle has been investigated in the same way. Our results agree with the experimental findings that the shortest investigated linker yields a strongly enhanced binding compared to the monovalent case due to the binding of the guest linker to the host linker. Our first principle calculations show clearly that this enhanced binding is due to electronic effects, namely the dispersion interaction of the two

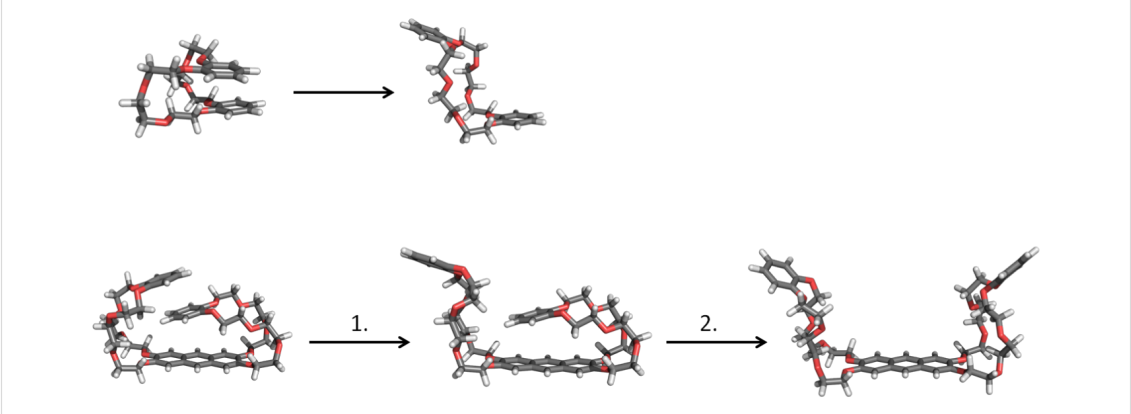

Figure 4: Optimized gas-phase structures for unfolding the monovalent (first row) and divalent (second row) host molecules. For the latter case a two-step process is found.

linkers. For the shortest linker this interaction results in a nearly ideal π–π stacking. For the two longer linkers ideal packing is not possible due to steric hindrance. These investigations proved that besides the primary binding sites in multivalent arrangements the interaction of the linkers can influence the binding process significantly. Therefore the term of non-innocent linkers introduced in [16] is well justified.

Acknowledgements

Support by the German Research Foundation (DFG) through the Collaborative Research Center (CRC) 765 'Multivalency as chemical organization and action principle: new architectures, functions and applications' and the High-Performance Computing facilities of the Freie Universität Berlin (ZEDAT) is acknowledged.

References

1. Fasting, C.; Schalley, C. A.; Weber, M.; Seitz, O.; Hecht, S.; Koksch, B.; Dernedde, J.; Graf, C.; Knapp, E.-W.; Haag, R. *Angew. Chem., Int. Ed.* **2012,** *51,* 10472–10498. doi:10.1002/anie.201201114
2. Mammen, M.; Choi, S. K.; Whitesides, G. M. *Angew. Chem., Int. Ed.* **1998,** *37,* 2754–2794. doi:10.1002/(SICI)1521-3773(19981102)37:20<2754::AID-ANIE2754>3.0.CO;2-3
3. Choi, S.-K. *Synthetic Multivalent Molecules;* Wiley: Hoboken, NJ, USA, 2004. doi:10.1002/0471578908
4. Kitov, P. I.; Sadowska, J. M.; Mulvey, G.; Armstrong, G. D.; Ling, H.; Pannu, N. S.; Read, R. J.; Bundle, D. R. *Nature* **2000,** *403,* 669–672. doi:10.1038/35001095
5. Kitov, P. I.; Bundle, D. R. *J. Am. Chem. Soc.* **2003,** *125,* 16271–16284. doi:10.1021/ja038223n
6. Hartmann, M.; Lindhorst, T. K. *Eur. J. Org. Chem.* **2011,** *2011,* 3583–3609. doi:10.1002/ejoc.201100407
7. Schwefel, D.; Maierhofer, C.; Beck, J. G.; Seeberger, S.; Diederichs, K.; Möller, H. M.; Welte, W.; Wittmann, V. *J. Am. Chem. Soc.* **2010,** *132,* 8704–8719. doi:10.1021/ja101646k
8. Cairo, C. W.; Gestwicki, J. E.; Kanai, M.; Kiessling, L. L. *J. Am. Chem. Soc.* **2002,** *124,* 1615–1619. doi:10.1021/ja016727k
9. Kramer, R. H.; Karpen, J. W. *Nature* **1998,** *395,* 710–713. doi:10.1038/27227
10. Badjić, J. D.; Nelson, A.; Cantrill, S. J.; Turnbull, W. B.; Stoddart, J. F. *Acc. Chem. Res.* **2005,** *38,* 723–732. doi:10.1021/ar040223k
11. Mulder, A.; Huskens, J.; Reinhoudt, D. N. *Org. Biomol. Chem.* **2004,** *2,* 3409–3424. doi:10.1039/b413971b
12. Röckendorf, N.; Lindhorst, T. K. *Top. Curr. Chem.* **2001,** *217,* 201–238. doi:10.1007/3-540-45003-3_6
13. Badjić, J. D.; Balzani, V.; Credi, A.; Lowe, J. N.; Silvi, S.; Stoddart, J. F. *Chem. – Eur. J.* **2004,** *10,* 1926–1935. doi:10.1002/chem.200305687
14. Badjić, J. D.; Balzani, V.; Credi, A.; Silvi, S.; Stoddart, J. F. *Science* **2004,** *303,* 1845–1849. doi:10.1126/science.1094791
15. Badjić, J. D.; Ronconi, C. M.; Stoddart, J. F.; Balzani, V.; Silvi, S.; Credi, A. *J. Am. Chem. Soc.* **2006,** *128,* 1489–1499. doi:10.1021/ja0543954
16. Jiang, W.; Nowosinski, K.; Löw, N. L.; Dzyuba, E. V.; Klautzsch, F.; Schäfer, A.; Huuskonen, J.; Rissanen, K.; Schalley, C. A. *J. Am. Chem. Soc.* **2012,** *134,* 1860–1868. doi:10.1021/ja2107096
17. *TURBOMOLE,* V6.5; TURBOMOLE GmbH: Karlsruhe, Germany, 2013, http://www.turbomole.com.
18. Tao, J.; Perdew, J. P.; Staroverov, V. N.; Scuseria, G. E. *Phys. Rev. Lett.* **2003,** *91,* 146401. doi:10.1103/PhysRevLett.91.146401
19. Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. *J. Chem. Phys.* **2010,** *132,* 154104. doi:10.1063/1.3382344
20. Grimme, S.; Ehrlich, S.; Goerigk, L. *J. Comput. Chem.* **2011,** *32,* 1456–1465. doi:10.1002/jcc.21759
21. Weigend, F.; Häser, M.; Patzelt, H.; Ahlrichs, R. *Chem. Phys. Lett.* **1998,** *294,* 143–152. doi:10.1016/S0009-2614(98)00862-8
22. Weigend, F. *Phys. Chem. Chem. Phys.* **2006,** *8,* 1057–1065. doi:10.1039/b515623h
23. Werner, H.-J.; Schütz, M. *J. Chem. Phys.* **2011,** *135,* 144116. doi:10.1063/1.3641642
24. *Molpro,* version 2010.1; 2010, http://www.molpro.net.
25. Achazi, A. J.; v. Krbek, L. K. S.; Schalley, C. A.; Paulus, B. *J. Comput. Chem.* **2015,** in press. doi:10.1002/jcc.23914
26. Grimme, S. *Chem. – Eur. J.* **2012,** *18,* 9955–9964. doi:10.1002/chem.201200497
27. Schäfer, A.; Horn, H.; Ahlrichs, R. *J. Chem. Phys.* **1992,** *97,* 2571–2577. doi:10.1063/1.463096
28. Eckert, F.; Klamt, A. *AIChE J.* **2002,** *48,* 369–385. doi:10.1002/aic.690480220
29. *COSMOtherm,* Version C3.0, Release 13.01; COSMOlogic GmbH & Co. KG: Leverkusen, Germany, 2013.
30. Klamt, A.; Schüürmann, G. D. *J. Chem. Soc., Perkin Trans. 2* **1993,** 799–805. doi:10.1039/P29930000799
31. Becke, A. D. *Phys. Rev. A* **1988,** *38,* 3098–3100. doi:10.1103/PhysRevA.38.3098
32. Perdew, J. P. *Phys. Rev. B* **1986,** *33,* 8822–8824. doi:10.1103/PhysRevB.33.8822
33. Schäfer, A.; Huber, C.; Ahlrichs, R. *J. Chem. Phys.* **1994,** *100,* 5829–5835. doi:10.1063/1.467146
34. Eichkorn, K.; Weigend, F.; Treutler, O.; Ahlrichs, R. *Theor. Chem. Acc.* **1997,** *97,* 119–124. doi:10.1007/s002140050244
35. Hunter, C. A.; Anderson, H. L. *Angew. Chem., Int. Ed.* **2009,** *48,* 7488–7499. doi:10.1002/anie.200902490

License and Terms

5.5 Paper A5

"Gating the photochromism of an azobenzene by strong host–guest interactions in a divalent pseudo[2]rotaxane"
Mirko Lohse, Karol Nowosinski, Nora L. Traulsen, Andreas J. Achazi, Larissa K. S. von Krbek, Beate Paulus, Christoph A. Schalley and Stefan Hecht
Chemical Communications **2015**, **51**, 9777-9780.
DOI: 10.1039/c5cc02811f
URL: http://dx.doi.org/10.1039/c5cc02811f

Contributions

The experimental investigations were done by Mirko Lohse, Karol Nowosinski, Nora L. Traulsen and Larissa K. S. von Krbek. I adjusted the multilevel DFT-based approach and applied it. Mirko Lohse and Karol Nowosinski wrote the manuscript together. All authors added contributions to the final version of the publication.

5.6 Paper A6

"A Divalent Pentastable Redox-Switchable Donor–Acceptor Rotaxane"
Hendrik V. Schröder, Henrik Hupatz, Andreas J. Achazi, Sebastian Sobottka, Biprajit Sarkar, Beate Paulus and Christoph A. Schalley
Chemistry – A European Journal **2017**, **23**, 2960-2967.
DOI: 10.1002/chem.201605710
URL: http://dx.doi.org/10.1002/chem.201605710

Contributions

The synthesis and experiments were done by Hendrik V. Schröder and Henrik Hupatz. Sebastian Sobottka assisted with the cyclic voltammetry experiments. Hendrik V. Schröder and Henrik Hupatz and I wrote the manuscript together, the main contributions and general concept were done by Hendrik V. Schröder. All authors added contributions to the final version of the publication.

5.7 Paper A7

"A Computational Study of Samarium Diiodide-Induced Cyclizations of *N*-Oxoalkyl-Substituted Methyl Indole-3-Carboxylates – A Rationale of the Diastereoselectivity"
Andreas J. Achazi, Dirk Andrae, Hans-Ulrich Reissig, Beate Paulus
Journal of Computational Chemistry **2017**.
DOI: 10.1002/jcc.25055
URL: http://dx.doi.org/10.1002/jcc.25055

Contributions

I did the DFT calculation and wrote the manuscript. Dirk Andrae helped with the proper description of the rare earth metal and Hans-Ulrich Reissig with the interpretation of the data from the synthetic organic view. All authors added contributions to the final version of the publication.

5.8 Paper B1

"Hydration Effects Turn a Highly Stretched Polymer from an Entropic into an Energetic Spring"
Susanne Liese, Manuel Gensler, Stefanie Krysiak, Richard Schwarzl, Andreas Achazi, Beate Paulus, Thorsten Hugel, Jürgen P. Rabe and Roland R. Netz
ACS Nano **2017**, **11**, 702-712.
DOI: 1010.1021/acsnano.6b07071
URL: http://dx.doi.org/10.1021/acsnano.6b07071

Contributions

The MD simulations were done by Susanne Liese and Richard Schwarzl. Manuel Gensler and Stefanie Krysiak did the Atomic-force microscopy (AFM). Susanne Liese and I did the DFT calculations. Susanne Liese wrote the manuscript. All authors added contributions to the final version of the publication.

6 Appendix

Molecular structures used to calculate the dispersion interaction between the aromatic system of the crown ether and the organic rest of the primary ammonium cation (section 3).

Energy = -1404.556861852
C -1.3581300 2.1533100 -1.5737700
C -0.0059200 2.5962800 -2.0790500
O -1.1836800 1.6059100 -0.2569600
C 0.8999800 0.3258600 3.7552600
C -0.5404800 0.0370900 3.4079500
O -0.7957400 0.5277400 2.0761900
C 4.7668400 -0.3991900 -0.1037500
C 4.4698600 0.4001600 -1.3513100
C 2.8869800 0.5458600 -3.1533000
C 2.0891900 1.8120500 -2.8990200
O 0.8434400 1.4526000 -2.2756200
O 3.2769600 -0.1213500 -1.9462100
C 3.1437800 -0.1592400 3.2038600
C 4.0177500 -0.7759900 2.1381700
O 3.7829700 -0.0848200 0.8978800
O 1.7692200 -0.4083600 2.8889200
C -4.5121500 0.5581100 2.0252300
C -4.7159900 1.1416500 0.7821300
C -3.6232500 1.5002200 -0.0139200
C -2.3255500 1.2884700 0.4423700
C -2.1202200 0.7006400 1.7081900
C -3.2111300 0.3299700 2.4875000
H -1.8037000 1.3969400 -2.2320500

H -2.0150500 3.0313900 -1.5317500
H -0.1417500 3.1153800 -3.0381100
H 0.4620100 3.2891800 -1.3655700
H 1.1009200 1.4030400 3.6591700
H 1.0698500 0.0313900 4.8024300
H -0.7574600 -1.0373500 3.4525700
H -1.1813500 0.5663900 4.1219800
H 5.7636400 -0.1315800 0.2735200
H 4.7495100 -1.4752300 -0.3267900
H 5.3113200 0.3010600 -2.0529000
H 4.3519300 1.4620400 -1.0923000
H 2.2707400 -0.1706400 -3.7038200
H 3.7710600 0.7872100 -3.7600200
H 2.6370300 2.5150500 -2.2543600
H 1.8858100 2.3118200 -3.8571900
H 3.3946200 -0.5989900 4.1809900
H 3.3253300 0.9245500 3.2501800
H 5.0712200 -0.6614200 2.4281000
H 3.8001800 -1.8468100 2.0180200
H -3.7951700 1.9366000 -0.9905600
H -3.0631900 -0.1434800 3.4510200
N 1.0839700 -0.3391300 -0.0190200
C 0.9610400 -1.7556700 -0.5111500
C -0.2664400 -1.9375900 -1.3986900
C -1.5727000 -2.0414000 -0.6184100
O -2.6364400 -1.8079300 -1.5255400
C -3.9212900 -2.0615900 -0.9300900
C -4.9924800 -1.6908000 -1.9360500
H 0.9984000 0.3224400 -0.8165000
H 0.3675900 -0.0886100 0.6839600
H 1.8764900 -1.9521100 -1.0693900
H 0.9376500 -2.4015400 0.3711600
H -0.1300800 -2.8393200 -2.0037300
H -0.3367200 -1.0961200 -2.0975300

H -1.6662000 -3.0399700 -0.1596300
H -1.6114200 -1.3022900 0.1990400
H -3.9844600 -3.1266200 -0.6571700
H -4.0193800 -1.4656300 -0.0112200
H -4.9316900 -0.6276800 -2.1878400
H -5.9837000 -1.8908000 -1.5168500
H -5.3549400 0.2679600 2.6435800
H -5.7215800 1.3149900 0.4135900
H -4.8798200 -2.2743800 -2.8543000

Energy = -328.5123785205
C -0.7149272 -0.4817255 -1.6246618
O -0.2631266 0.0537930 -0.3801864
C -0.2733548 1.4800075 -0.3709689
C 0.2181817 1.9640372 0.9854493
C 0.2394859 3.4972338 1.0911446
N 0.7042349 4.0412356 2.3770675
C -0.6718179 -1.9968431 -1.5365293
H 0.3482911 -2.3400138 -1.3391384
H -1.3192075 -2.3518265 -0.7289342
H -1.7399048 -0.1285174 -1.8289910
H -0.0689712 -0.1164105 -2.4409793
H -1.2961089 1.8467863 -0.5671519
H 0.3748451 1.8629449 -1.1785319
H 1.2268203 1.5677929 1.1598985
H -0.4312501 1.5490622 1.7669914
H -0.7673621 3.8944097 0.9112402
H 0.8837453 3.9133450 0.3065480
H 1.6469379 3.7022392 2.5712822
H 0.1175454 3.6839411 3.1316951
H -1.0131965 -2.4402616 -2.4780338

Energy = -1132.622899350
C -1.3581300 2.1533100 -1.5737700

C -0.0059200 2.5962800 -2.0790500
O -1.1836800 1.6059100 -0.2569600
C 0.8999800 0.3258600 3.7552600
C -0.5404800 0.0370900 3.4079500
O -0.7957400 0.5277400 2.0761900
C 4.7668400 -0.3991900 -0.1037500
C 4.4698600 0.4001600 -1.3513100
C 2.8869800 0.5458600 -3.1533000
C 2.0891900 1.8120500 -2.8990200
O 0.8434400 1.4526000 -2.2756200
O 3.2769600 -0.1213500 -1.9462100
C 3.1437800 -0.1592400 3.2038600
C 4.0177500 -0.7759900 2.1381700
O 3.7829700 -0.0848200 0.8978800
O 1.7692200 -0.4083600 2.8889200
C -4.5121500 0.5581100 2.0252300
C -4.7159900 1.1416500 0.7821300
C -3.6232500 1.5002200 -0.0139200
C -2.3255500 1.2884700 0.4423700
C -2.1202200 0.7006400 1.7081900
C -3.2111300 0.3299700 2.4875000
H -1.8037000 1.3969400 -2.2320500
H -2.0150500 3.0313900 -1.5317500
H -0.1417500 3.1153800 -3.0381100
H 0.4620100 3.2891800 -1.3655700
H 1.1009200 1.4030400 3.6591700
H 1.0698500 0.0313900 4.8024300
H -0.7574600 -1.0373500 3.4525700
H -1.1813500 0.5663900 4.1219800
H 5.7636400 -0.1315800 0.2735200
H 4.7495100 -1.4752300 -0.3267900
H 5.3113200 0.3010600 -2.0529000
H 4.3519300 1.4620400 -1.0923000
H 2.2707400 -0.1706400 -3.7038200

H 3.7710600 0.7872100 -3.7600200
H 2.6370300 2.5150500 -2.2543600
H 1.8858100 2.3118200 -3.8571900
H 3.3946200 -0.5989900 4.1809900
H 3.3253300 0.9245500 3.2501800
H 5.0712200 -0.6614200 2.4281000
H 3.8001800 -1.8468100 2.0180200
H -3.7951700 1.9366000 -0.9905600
H -3.0631900 -0.1434800 3.4510200
N 1.0839700 -0.3391300 -0.0190200
H 0.8426665 -1.2611610 -0.3833106
H 0.9984000 0.3224400 -0.8165000
H 0.3675900 -0.0886100 0.6839600
H -5.3549400 0.2679600 2.6435800
H -5.7215800 1.3149900 0.4135900

Energy = -56.58756117550
N 0.8030158 3.1167886 2.4684722
H 1.4113685 3.7113519 1.9058679
H 0.2049690 3.7439101 3.0062751
H 0.1925867 2.6281193 1.8136748

Table 6.1: Change in vibrational entropy S_{vib} (including the inner rotations) when the guest binds to the host (**mH** and **H**, respectively). The change is given in $\text{kJ}\,\text{mol}^{-1}$ at 298.15 K.

Guest	$-T\Delta S_{\text{vib}}$
mG1-OTs	-13.6
G1-2OTs	-1.4
G2-2OTs	0.1
G3-2OTs	-1.7
G4-2OTs	2.3
G5-2OTs	2.7
mG2-OTs	-12.3
G6-2OTs	-4.9
G7-2OTs	0.5
G8-2OTs	0.3
G9-2OTs	-5.1
G10-2OTs	-4.1
G11-2OTs	-2.0

Bibliography

[1] Fasting, C.; Schalley, C. A.; Weber, M.; Seitz, O.; Hecht, S.; Koksch, B.; Dernedde, J.; Graf, C.; Knapp, E.-W.; Haag, R. *Angewandte Chemie International Edition* **2012**, *51*, 10472–10498.

[2] Mammen, M.; Choi, S.-K.; Whitesides, G. M. *Angewandte Chemie International Edition* **1998**, *37*, 2754–2794.

[3] Joshi, A.; Vance, D.; Rai, P.; Thiyagarajan, A.; Kane, R. S. *Chemistry – A European Journal* **2008**, *14*, 7738–7747.

[4] Kiessling, L. L.; Gestwicki, J. E.; Strong, L. E. *Angewandte Chemie International Edition* **2006**, *45*, 2348–2368.

[5] Mulder, A.; Huskens, J.; Reinhoudt, D. N. *Organic and Biomolecular Chemistry* **2004**, *2*, 3409–3424.

[6] Huskens, J. *Current Opinion in Chemical Biology* **2006**, *10*, 537–543.

[7] Badjić, J. D.; Nelson, A.; Cantrill, S. J.; Turnbull, W. B.; Stoddart, J. F. *Accounts of Chemical Research* **2005**, *38*, 723–732.

[8] He, Z.; Jiang, W.; Schalley, C. A. *Chemical Society Reviews* **2015**, *44*, 779–789.

[9] Benson, S. W. *Journal of the American Chemical Society* **1958**, *80*, 5151–5154.

[10] Bishop, D. M.; Laidler, K. J. *The Journal of Chemical Physics* **1965**, *42*, 1688–1691.

[11] Ercolani, G. *Journal of the American Chemical Society* **2003**, *125*, 16097–16103.

[12] Ercolani, G.; Piguet, C.; Borkovec, M.; Hamacek, J. *Journal of Physical Chemistry B* **2007**, *111*, 12195–12203.

[13] Hunter, C. A.; Anderson, H. L. *An* **2009**, *48*, 7488–7499.

[14] Ercolani, G.; Schiaffino, L. *Angewandte Chemie International Edition* **2011**, *50*, 1762–1768.

[15] von Krbek, L. K. S.; Schalley, C. A.; Thordarson, P. *Chemical Society Reviews* **2017**, *46*, 2622–2637.

[16] Accessed on: 18th september 2017 from http://www.sfb765.de/.

[17] Carter, P. J.; Winter, G.; Wilkinso, A. J.; Fersht, A. R. *Cell* **1984**, *38*, 835–840.

[18] Cockroft, S. L.; Hunter, C. A. *Chemical Society Reviews* **2007**, *36*, 172–188.

[19] Fischer, F. R.; Wood, P. A.; Allen, F. H.; Diederich, F. *Proceedings of the National Academy of Sciences of the United States of America* **2008**, *105*, 17290–17294.

[20] Jiang, W.; Nowosinski, K.; Löw, N. L.; Dzyuba, E. V.; Klautzsch, F.; Schäfer, A.; Huuskonen, J.; Rissanen, K.; Schalley, C. A. *Journal of the American Chemical Society* **2012**, *134*, 1860–1868.

[21] Sun, H.; Hunter, C. A.; Navarro, C.; Turega, S. *Journal of the American Chemical Society* **2013**, *135*, 13129–13141.

[22] Kaufmann, L.; Traulsen, N. L.; Springer, A.; Schröder, H. V.; Mäkelä, T.; Rissanen, K.; Schalley, C. A. *Organic Chemistry Frontiers* **2014**, *1*, 521–531.

[23] Traulsen, N. L.; Traulsen, C. H.-H.; Deutinger, P. M.; Müller, S.; Schmidt, D.; Linder, I.; Schalley, C. A. *Organic and Biomolecular Chemistry* **2015**, *13*, 10881–10887.

[24] Cram, D. J.; Cram, J. M. *Science* **1974**, *183*, 803–809.

[25] Cram, D. J. *Angewandte Chemie International Edition* **1986**, *25*, 1039–1134.

[26] Cram, D. J. *Angewandte Chemie International Edition* **1988**, *27*, 1009–1112.

[27] Hogben, H. J.; Sprafke, J. K.; Hoffmann, M.; Pawlicki, M.; Anderson, H. L. *Journal of the American Chemical Society* **2011**, *133*, 20962–20969.

[28] Adams, H.; Chekmeneva, E.; Hunter, C. A.; Misuraca, M. C.; Navarro, C.; Turega, S. M. *Journal of the American Chemical Society* **2013**, *135*, 1853–1863.

[29] Mammen, M.; Shakhnovich, E. I.; Whitesides, G. M. *Journal of Organic Chemistry* **1998**, *63*, 3168–3175.

[30] Eblinger, F.; Schneider, H.-J. *Angewandte Chemie International Edition* **1998**, *37*, 826–829.

[31] Gellman, S. H. *Accounts of Chemical Research* **1998**, *31*, 173–180.

[32] Hill, D. J.; Mio, M. J.; Prince, R. B.; Hughes, T. S.; Moore, J. S. *Chemical Reviews* **2001**, *101*, 3893–4011.

[33] Zhong, Z.; Li, X.; Zhao, Y. *Journal of the American Chemical Society* **2011**, *133*, 8862–8865.

[34] Krishnamurthy, V. M.; Semetey, V.; Bracher, P. J.; Shen, N.; Whitesides, G. M. *Journal of the American Chemical Society* **2007**, *129*, 1312–1320.

[35] Sun, H.; Hunter, C. A.; Llamas, E. M. *Chemical Science* **2015**, *6*, 1444–1453.

[36] Stross, A. E.; Iadevaia, G.; Hunter, C. A. *Chemical Science* **2016**, *7*, 94–101.

[37] Stross, A. E.; Iadevaia, G.; Hunter, C. A. *Chemical Science* **2016**, *7*, 5686–5691.

[38] von Krbek, L. K. S. Cooperative effects in multivalent crown ether/ammonium assemblies. Ph.D. thesis, Freie Universität Berlin, 2016.

[39] Pedersen, C. J. *Journal of the American Chemical Society* **1967**, *89*, 2495–2496.

[40] Pedersen, C. J. *Journal of the American Chemical Society* **1967**, *89*, 7017–7036.

[41] Badjić, J. D.; Ronconi, C. M.; Stoddart, J. F.; Balzani, V.; Silvi, S.; Credi, A. *Journal of the American Chemical Society* **2006**, *128*, 1489–1499.

[42] Ashton, P. R.; Chrystal, E. J. T.; Glink, P. T.; Menzer, S.; Schiavo, C.; Spencer, N.; Stoddart, J. F.; Tasker, P. A.; White, A. J. P.; Williams, D. J. *Chemistry – A European Journal* **1996**, *2*, 709–728.

[43] Zhang, Z.-J.; Han, M.; Zhang, H.-Y.; Liu, Y. *Organic Letters* **2013**, *15*, 1698–1701.

[44] Jiang, W.; Sattler, D.; Rissanen, K.; Schalley, C. A. *Organic Letters* **2011**, *13*, 4502–4505.

[45] Jiang, W.; Winkler, H. D. F.; Schalley, C. A. *Journal of the American Chemical Society* **2008**, *130*, 13852–13853.

[46] Gokel, G. W.; Leevy, W. M.; Weber, M. E. *Chemical Reviews* **2004**, *104*, 2723–2750.

[47] Yamada, Y.; Okamoto, M.; Furukawa, K.; Kato, T.; Tanaka, K. *Angewandte Chemie International Edition* **2012**, *124*, 733–737.

[48] Yamada, Y.; aki Okada, M.; Tanaka, K. *Chemical Communications* **2013**, *49*, 11053–11055.

[49] Luo, Y.-H.; Song, W.-T.; Ge, S.-W.; Sun, B.-W. *Polyhedron* **2014**, *69*, 160–166.

[50] Jullien, L.; Lehn, J.-M. *Tetrahedron Letters* **1988**, *29*, 3803–3806.

[51] Fyles, T. M.; James, T. D.; Pryhitka, A.; Zojaji, M. *Journal of Organic Chemistry* **1993**, *58*, 7456–7468.

[52] Voyer, N.; Rohitaille, M. *Journal of the American Chemical Society* **1995**, *117*, 6599–6600.

[53] Hellweg, A.; Eckert, F. *American Institute of Chemical Engineers Journal* **2017**, *63*, 3944–3954.

[54] Bartlett, R. J. *The Journal of Physical Chemistry* **1989**, *93*, 1697–1708.

[55] Werner, H.-J.; Schütz, M. *The Journal of Chemical Physics* **2011**, *135*, 144116.

[56] Hohenberg, P.; Kohn, W. *Physical Review* **1964**, *136*, B864–B871.

[57] Kohn, W.; Sham, L. J. *Physical Review* **1965**, *140*, A1133–A1138.

[58] Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. *The Journal of Chemical Physics* **2010**, *132*, 154104.

[59] Grimme, S. *Journal of Computational Chemistry* **2004**, *25*, 1463–1473.

[60] Grimme, S. *Journal of Computational Chemistry* **2006**, *27*, 1787–1799.

[61] Grimme, S.; Ehrlich, S.; Goerigk, L. *Journal of Computational Chemistry* **2011**, *32*, 1456–1465.

[62] Klamt, A. *The Journal of Chemical Physics* **1995**, *99*, 2224–2235.

[63] Klamt, A.; Jonas, V.; Bürger, T.; Lohrenz, J. C. W. *Journal of Physical Chemistry A* **1998**, *102*, 5074–5085.

[64] Eckert, F.; Klamt, A. *American Institute of Chemical Engineers Journal* **2002**, *48*, 369–385.

[65] 2016; COSMO*therm*, Version C3.0, Release 16.02; COSMO*logic* GmbH & Co. http://www.cosmologic.de.

[66] Klamt, A.; Schüürmann, G. *Journal of the Chemical Society, Perkin Transactions 2* **1993**, *5*, 799–805.

[67] Schäfer, A.; Klamt, A.; Sattel, D.; Lohrenz, J. C. W.; Eckert, F. *Physical Chemistry Chemical Physics* **2000**, *2*, 2187–2193.

[68] Deglmann, P.; Schenk, S. *Journal of Computational Chemistry* **2012**, *33*, 1304–1320.

[69] Grimme, S. *Chemistry: A European Journal* **2012**, *18*, 9955–9964.

[70] Antony, J.; Sure, R.; Grimme, S. *Chemical Communications* **2015**, *51*, 1764–1774.

[71] Caldararu, O.; Olsson, M. A.; Riplinger, C.; Neese, F.; Ryde, U. *Journal of Computer-Aided Molecular Design* **2017**, *31*, 87–106.

[72] Achazi, A. J.; von Krbek, L. K. S.; Schalley, C. A.; Paulus, B. *Journal of Computational Chemistry* **2016**, *37*, 18–24.

[73] Schrödinger, E. *Physical Review* **1926**, *28*, 1049–1070.

[74] Born, M.; Oppenheimer, R. *Annalen der Physik* **1927**, *389*, 457–484.

[75] Slater, J. C. *Physical Review* **1929**, *34*, 1293–1322.

[76] Roothaan, C. C. *Reviews of Modern Physics* **1951**, *23*, 69–89.

[77] Hall, G. G. *Proceedings of the Royal Society A* **1951**, *205*, 541–552.

[78] Hartree, D. R. *Mathematical Proceedings of the Cambridge Philosophical Society* **1928**, *24*, 89–110.

[79] Hartree, D. R. *Mathematical Proceedings of the Cambridge Philosophical Society* **1928**, *24*, 111–132.

[80] Hartree, D. R. *Mathematical Proceedings of the Cambridge Philosophical Society* **1928**, *24*, 426–437.

[81] Fock, V. *Zeitschrift für Physik* **1930**, *61*, 126–148.

[82] Slater, J. C. *Physical Review* **1930**, *35*, 210–211.

[83] Szabo, A.; Ostlund, N. S. *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory*; Dover Publications, Inc., Minela, New York, 1996.

[84] Pople, J. A.; Nesbet, R. K. *The Journal of Chemical Physics* **1954**, *22*, 571–572.

[85] Roothaan, C. C. J. *Reviews of Modern Physics* **1960**, *32*, 179–185.

[86] TURBOMOLE V7.0.1 2015, a development of University of Karlsruhe and Forschungszentrum Karlsruhe GmbH, 1989-2007, TURBOMOLE GmbH, since 2007; available from http://www.turbomole.com.

[87] Fermi, E. *Rendiconti Lincei* **1926**, *3*, 145–149.

[88] Dirac, P. A. M. *Proceedings of the Royal Society, A* **1926**, *112*, 661–777.

[89] Møller, C.; Plesset, M. S. *Physical Review* **1934**, *46*, 618–622.

[90] Pipek, J.; Ladik, J. *Chemical Physics* **1986**, *102*, 445–458.

[91] Pipek, J.; Mezey, P. G. *The Journal of Chemical Physics* **1989**, *90*, 4916–4926.

[92] Pulay, P. *Chemical Physics Letters* **1983**, *100*, 151–154.

[93] Thomas, L. H. *Proceedings of the Cambridge Philosophical Society* **1927**, *23*, 542–548.

[94] Fermi, E. *Rend. Accad. Naz. Lincei* **1927**, *6*, 602–607.

[95] Koch, W.; Holthausen, M. C. *A Chemist's Guide to Density Functional Theory*, second edition ed.; Wiley-VCH Verlag GmbH, Weinheim, 2001.

[96] Parr, R. G.; Yang, W. *Density-Functional Theory of Atoms and Molecules*; Oxford University Press, New York, 1989.

[97] Perdew, J. P. In *Density Functional Theory and Its Application to Materials*; Doren, V. E. V., Alsenoy, C. V., Geerlings, P., Eds.; American Institute of Physics, 2001.

[98] Sousa, S. F.; Fernandes, P. A.; Ramos, M. J. *Journal of Physical Chemistry A* **2007**, *111*, 10439–10452.

[99] Medvedev, M. G.; Bushmarinov, I. S.; Sun, J.; Perdew, J. P.; Lyssenko, K. A. *Science* **2017**, *355*, 49–52.

[100] Dirac, P. A. M. *Proceedings of the Royal Society A* **1929**, *123*, 714–733.

[101] Slater, J. C. *Physical Review* **1951**, *81*, 385–390.

[102] Perdew, J. P.; Wang, Y. *Physical Review B* **1992**, *45*, 13244–13249.

[103] Tao, J.; Perdew, J. P.; Staroverov, V. N.; Scuseria, G. E. *Physical Review Letters* **2003**, *91*, 146401.

[104] Jensen, F. *Introduction to Computational Chemistry*, second edition ed.; John Wiley & Sons Ltd, 2007.

[105] Vosko, S. H.; Wilk, L.; Nusair, M. *Canadian Journal of Physics* **1980**, *58*, 1200–1211.

[106] Becke, A. D. *Physical Review A* **1988**, *38*, 3098–3100.

[107] Perdew, J. P. *Physical Review B* **1986**, *33*, 8822–8824.

[108] Perdew, J. P. *Physical Review B* **1986**, *34*, 7406.

[109] Perdew, J. P.; Burke, K.; Ernzerhof, M. *Physical Review Letters* **1996**, *77*, 3865–3868.

[110] Perdew, J. P.; Burke, K.; Ernzerhof, M. *Physical Review Letters* **1997**, *78*, 1396.

[111] Perdew, J. P.; Ernzerhof, M.; Burke, K. *The Journal of Chemical Physics* **1996**, *105*, 9982–9985.

[112] Lee, C.; Yang, W.; Parr, R. G. *Physical Review B* **1988**, *37*, 785–789.

[113] Becke, A. D. *The Journal of Chemical Physics* **1993**, *98*, 5648–5652.

[114] Zhao, Y.; Truhlar, D. G. *Theoretical Chemistry Accounts* **2008**, *120*, 215–241.

[115] Eichkorn, K.; Treutler, O.; Öhm, H.; Häser, M.; Ahlrichs, R. *Chemical Physics Letters* **1995**, *240*, 283–290.

[116] Eichkorn, K.; Treutler, O.; Öhm, H.; Häser, M.; Ahlrichs, R. *Chemical Physics Letters* **1995**, *242*, 652–660.

[117] Sierka, M.; Hogekamp, A.; Ahlrichs, R. *The Journal of Chemical Physics* **2003**, *118*, 9136–9148.

[118] Boys, S. F. *Proceedings of the Royal Society A* **1950**, *200*, 542.

[119] Eisenschitz, R.; London, F. W. *Zeitschrift für Physik* **1930**, *60*, 491–527.

[120] London, F. W. *Zeitschrift für Physik* **1930**, *63*, 245–279.

[121] Becke, A. D.; Johnson, E. R. *The Journal of Chemical Physics* **2005**, *123*, 154101.

[122] Johnson, E. R.; Becke, A. D. *The Journal of Chemical Physics* **2005**, *123*, 024101.

[123] Johnson, E. R.; Becke, A. D. *Chemical Physics Letters* **2006**, *432*, 600–603.

[124] Casimir, H. B. G.; Polder, D. *Physical Review* **1948**, *73*, 360–372.

[125] Starkschall, G.; Gordon, R. G. *The Journal of Chemical Physics* **1972**, *56*, 2801–2806.

[126] Mayer, J. E.; Brunauer, S.; Mayer, M. G. *Journal of the American Chemical Society* **1933**, *55*, 37–53.

[127] Lesk, A. M. *Journal of Physics A: Mathematical and General* **1980**, *13*, L111–L114.

[128] Gilson, M. K.; Irikura, K. K. *The Journal of Physical Chemistry B* **2010**, *114*, 16304–16317.

[129] Ayala, P. Y.; Schlegel, H. B. *The Journal of Chemical Physics* **1998**, *108*, 2314–2325.

[130] East, A. L. L.; Radom, L. *The Journal of Chemical Physics* **1997**, *106*, 6655–6674.

[131] Pitzer, K. S.; Gwinn, W. D. *The Journal of Chemical Physics* **1942**, *10*, 428–440.

[132] Nielsen, H. H. *Physical Review* **1932**, *40*, 445–456.

[133] Pitzer, K. S. *The Journal of Chemical Physics* **1946**, *14*, 239–243.

[134] Kilpatrick, J. E.; Pitzer, K. S. *The Journal of Chemical Physics* **1949**, *17*, 1064–1075.

[135] Truhlar, D. G. *Journal of Computational Chemistry* **1991**, *12*, 266–270.

[136] Chai, J.-D.; Head-Gordon, M. *Physical Chemistry Chemical Physics* **2008**, *10*, 6615–6620.

[137] Klamt, A.; Diedenhofen, M. *The Journal of Physical Chemistry A* **2015**, *119*, 5439–5445.

[138] Klamt, A. *Journal of Physical Chemistry* **1995**, *99*, 2224–2235.

[139] Sinnecker, S.; Rajendran, A.; Klamt, A.; Diedenhofen, M.; Neese, F. *Journal of Physical Chemistry A* **2006**, *110*, 2235–2245.

[140] Klamt, A.; Diedenhofen, M. *The Journal of Physical Chemistry A* **2015**, *119*, 5439–5445.

[141] van Duijneveldt, F. B.; van Duijneveldt-van de Rijdt, J. G. C. M.; van Lenthe, J. H. *Chemical Reviews* **1994**, *94*, 1873–1885.

[142] Helgaker, T.; Klopper, W.; Koch, H.; Noga, J. *The Journal of Chemical Physics* **1997**, *106*, 9639.

[143] Dunning Jr., T. H. *The Journal of Chemical Physics* **1989**, *90*, 1007–1023.

[144] Karton, A.; Martin, J. M. *Theoretical Chemistry Accounts* **2006**, *115*, 330–333.

[145] Schwerdtfeger, P. *ChemPhysChem* **2011**, *12*, 3143–3155.

[146] Mulliken, R. S. *The Journal of Chemical Physics* **1955**, *23*, 1833–1840.

[147] Reed, A. E.; Weinstock, R. B.; Weinhold, F. *The Journal of Chemical Physics* **1985**, *83*, 735–746.

[148] Löwdin, P.-O. *Physical Review* **1955**, *97*, 1474–1489.

[149] Koopmans, T. *Physica* **1934**, *1*, 104–113.

[150] Baerends, E. J.; Gritsenko, O. V.; van Meer, R. *Physical Chemistry Chemical Physics* **2013**, *15*, 16408–16425.

[151] van Meer, R.; Gritsenko, O. V.; Baerends, E. J. *Journal of Chemical Theory and Computation* **2014**, *10*, 4432–4441.

[152] Meot-Ner, M. *Journal of the American Chemical Society* **1983**, *105*, 4912–4915.

[153] von Krbek, L. K. S.; Achazi, A. J.; Solleder, M.; Weber, M.; Paulus, B.; Schalley, C. A. *Chemistry – A European Journal* **2016**, *22*, 15475–15484.

[154] Achazi, A. J.; Mollenhauer, D.; Paulus, B. *Beilstein Journal of Organic Chemistry* **2015**, *11*, 687–692.

[155] Mack, E. T.; Snyder, P. W.; Perez-Castillejos, R.; Whitesides, G. M. *Journal of the American Chemical Society* **2011**, *133*, 11701–11715.

[156] Weber, M.; Bujotzek, A.; Haag, R. *The Journal of Chemical Physics* **2012**, *137*, 054111.

[157] von Krbek, L. K. S.; Achazi, A. J.; Schoder, S.; Gaedke, M.; Biberger, T.; Paulus, B.; Schalley, C. A. *Chemistry – A European Journal* **2017**, *23*, 2877–2883.

[158] Lüttschwager, N. O. B.; Wassermann, T. N.; Mata, R. A.; Suhm, M. A. *Angewandte Chemie International Edition* **2013**, *52*, 463–466.

[159] Goodman, J. M. *Journal of Chemical Information and Computer Sciences* **1997**, *37*, 876–878.

[160] Lohse, M.; Nowosinski, K.; Traulsen, N. L.; Achazi, A. J.; von Krbek, L. K. S.; Paulus, B.; Schalley, C. A.; Hecht, S. *Chemical Communications* **2015**, *51*, 9777–9780.

[161] Schröder, H. V.; Hupatz, H.; Achazi, A. J.; Sobottka, S.; Sarkar, B.; Paulus, B.; Schalley, C. A. *Chemistry – A European Journal* **2017**, *23*, 2960–2967.

[162] Achazi, A. J.; Andrae, D.; Reissig, H.-U.; Paulus, B. *Journal of Computational Chemistry* **2017**, *38*, 2693–2700.

[163] Beemelmanns, C.; Blot, V.; Gross, S.; Lentz, D.; Reissig, H.-U. *Eur. J. Org. Chem.* **2010**, 2716–2732.

[164] Kise, N.; Mano, T.; Sakurai, T. *Org. Lett.* **2008**, *10*, 4617–4620.

[165] Liese, S.; Gensler, M.; Krysiak, S.; Schwarzl, R.; Achazi, A.; Paulus, B.; Hugel, T.; Rabe, J. P.; Netz, R. R. *ACS Nano* **2017**, *11*, 702–712.

[166] Lehn, J.-M. *Angewandte Chemie International Edition* **1990**, *29*, 1304–1319.

[167] Lohse, M.; von Krbek, L. K. S.; Radunz, S.; Moorthy, S.; Schalley, C. A.; Hecht, S. *Beilstein Journal of Organic Chemistry* **2015**, *11*, 748–762.

[168] Collman, J. P.; Barnes, C. E.; Swepston, P. N.; Ibers, J. A. *Journal of the American Chemical Society* **1984**, *106*, 3500–3510.

List of Figures

List of Tables