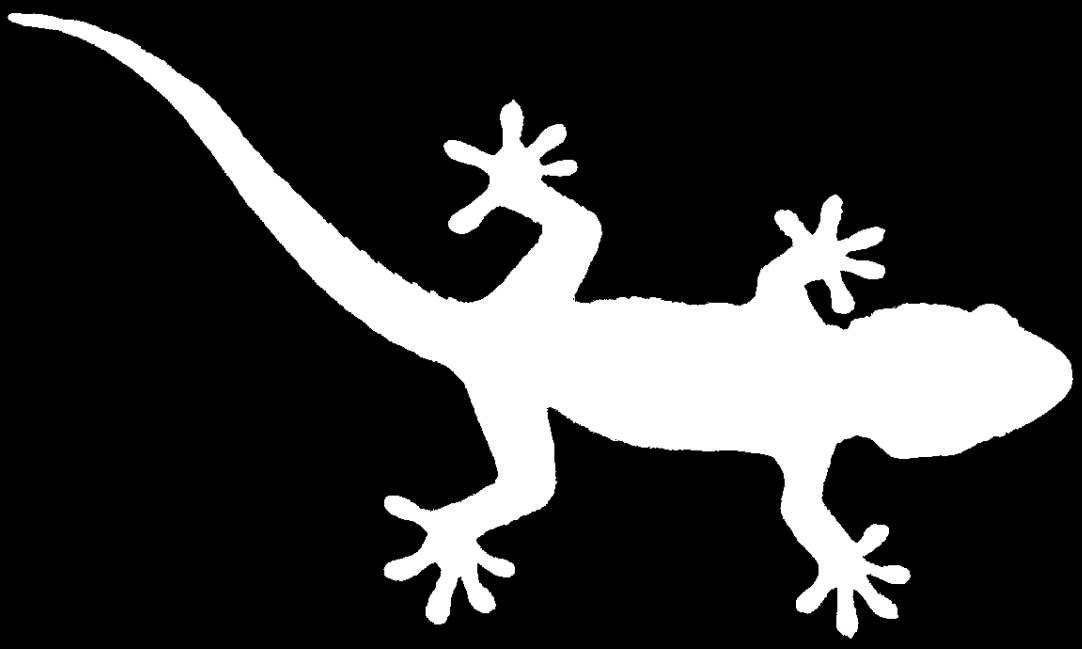
ROBERT POLLICE

Diss. ETH NO. 26208

LONDON DISPERSION IN MOLECULAR SYSTEMS



2019

Diss. ETH NO. 26208

LONDON DISPERSION IN MOLECULAR SYSTEMS

A thesis submitted to attain the degree of

DOCTOR OF SCIENCES

(Dr. sc. ETH Zurich)

presented by

ROBERT POLLICE

MSc, Vienna University of Technology, Austria

born on 28.03.1992
citizen of Austria

accepted on the recommendation of

Prof. Dr. Peter Chen, examiner
Prof. Dr. Sereina Riniker, co-examiner

2019

”Istud quod tu summum putas gradus est.”
— Seneca, Epistulae morales

Acknowledgements

Before going into scientific matters, I would like to thank all the people who shared parts of this long journey and helped me to reach my destination.

First of all, I want to express my deepest gratitudes to **Prof. Dr. Peter Chen**. I am very grateful that he allowed me to join his research group and work very independently on all my research projects and pursue my own ideas. His seemingly endless knowledge about chemistry is very inspiring and his calm and entertaining character are exemplary.

In addition, I would also like to thank **Prof. Dr. Sereina Riniker** for accepting to be my co-examiner without hesitation.

Big thanks goes to **Dr. Ilya Shenderovich** from the University of Regensburg who provided me with numerous suggestions for my research and specialized NMR equipment for crucial measurements.

I appreciate the support of **Dr. Olivier Byrde**, **Dr. Samuel Fux** and **Dr. Urban Borstnik** with the Euler cluster. They dealt with all my many special requests very fast and helped me to set up many different kinds of programs and computations. Additionally, **Dr. Andreas Bach** is acknowledged for IT support and for taking care of the group cluster.

Thanks goes to **Fritz Kastner**, **Florian Hastreiter** (both University of Regensburg), **Dr. Marc-Olivier Ebert**, **René Arnold**, **Rainer Frankenstein** and **Stephan Burkhardt** (all ETH Zürich) for help with the variable-temperature and DOSY NMR measurements. In addition, I am grateful to **Dr. Nils Trapp** and **Michael Solar** for performing the XRD structure determinations. Moreover, I thank **Peter Kälin** (ETH Zürich) for help with the ITC measurements and **Dr.**

Marco Zahner (ETH Zürich) for help with the static dielectric permittivity measurements.

Furthermore, I am grateful to **Anke Witten**, **Vlorina Ademi**, **Gabriela Schenk** and **Barbara Loepfe** for all their help they provided me with respect to organizational matters. Their continuous support always allowed me to focus on my research.

Armin Limacher is acknowledged for always lending me a helping hand in matters of IT, laboratory equipment and Swiss bureaucracy. In addition, I am very thankful to him for supporting me as a driver when I had to move within Zurich.

I also need to mention all the former group members who joined me on my journey. I appreciated all the late discussions over drinks and board games with **Dr. Ilya Kobylianskii**, **Dr. Krista Kobylianskii** and **Dr. Mihai Raducan**. I really enjoyed all the football talk with **Dr. Augustin Tchawou**, who was a regular lab companion on weekends, and who introduced me to the local football scene of Hardhof. I thank **Inesa Hadrovic** for leaving me her Milka blanket and **Ilchul Yoon** for many late night discussions. **Dr. Raphael Oeschger** deserves tribute for likely being the G.O.A.T. football player of the Chen group and I appreciate that he joined to take on the Zürcher Hochschulmeisterschaft in 2015, which we miserably failed at despite his participation. **Dr. Lukas Fritsche** deserves thanks for organizing the SOLA runs three times during my stay, and I really enjoyed all our parties together.

I would also like to mention all the current group members of the Chen group. A thanks goes to **Dr. Larisa Miloglyadova** for some insight into Russian politics and to **Yanan Miao** for some insight into Chinese lifestyle. I also thank **Adélaïde Savoy** and **Maurice Andrey** for revitalizing the Chen group and for a great group outing in Fribourg. I appreciate all the discussions about Swiss politics with **Joël Gubler** and **Stefan Jungen**. It was always fun to rant about Bavaria and to win against some of the Swiss group members in Jass with **Kevin Breitwieser**. I enjoyed several late night discussions with **Raphael Bissig**, who is definitely among the G.O.A.T. skiers of the Chen group. I would like to thank **Yannick**

Stöferle for always being a reliable companion in the lab. **Vladimir Gorbachev** is acknowledged for fun times in the lab and during group outings, especially late in the night. Thanks also goes to **Stefan Künzi** for being a regular lab companion on weekends and for great times out over beers.

Big thanks belong to **Felix Fleckenstein**, **Lotte Stindt** and **Paul Türtscher**, the students I supervised, for making my time in the lab very entertaining. In particular, I would like to point out that **Lotte Stindt** enriched me with science. Additionally, I had a great time with **Felix Fleckenstein** in Regensburg measuring NMR and testing the local restaurants and I would like to wish him all the best as PhD student in our group. Furthermore, I really enjoyed pondering for hours about kinetic data and their mechanistic implications with **Paul Türtscher**.

Moreover, I would like to thank **Armin Limacher**, **Marek Bot** and **Dr. Alexandra Tsybizova** for the great atmosphere in the Witcher lab. I really enjoyed our (fun) discussions about European history and politics, especially about Brexit, and about both old and new computer games.

I am very grateful to **Dr. Renana Poranne**, who became a very close friend of mine during my stay in the Chen group. Thank you for all the (late night) discussions about OC4 and for all the fun times over drinks. It was really great to have a regular companion for the GRC Physical Organic Chemistry and I appreciate your constant input to help me to improve my English and myself.

Additionally, I am very thankful to **Eno Paenurk** for becoming a very close friend of mine. Thank you for all our discussions about (bad) grammar and (bad) physics, whether it was over beers, or not. I am sure that, while reading this, you already found numerous mistakes in this very thesis and I am looking forward to have them all corrected by you.

Special thanks goes to my brothers **Marcus** and **Reinhard** for supporting me in IT matters and for providing me with a continuous supply of new computer games.

Finally, I owe my parents **Elisabeth** and **Giancarlo** everything in my life. They always supported me in every matter I had and, hence, they made it possible for me to be a scientist.

Published Parts

Peer-Reviewed Articles

R. Pollice, F. Fleckenstein, I. G. Shenderovich, P. Chen:
Compensation of London Dispersion in the Gas Phase and in Aprotic Solvents
Angew. Chem. Int. Ed. **2019**, *58*, 14281 – 14288.

R. Pollice, P. Chen:
A Universal Quantitative Descriptor of the Dispersion Interaction Potential
Angew. Chem. Int. Ed. **2019**, *58*, 9758 – 9769.

R. Pollice, P. Chen:
Origin of the Immiscibility of Alkanes and Perfluoroalkanes
J. Am. Chem. Soc. **2019**, *141*, 3489 – 3506.

R. Pollice, M. Bot, I. J. Kobylianskii, I. G. Shenderovich, P. Chen:
Attenuation of London Dispersion in Dichloromethane Solutions
J. Am. Chem. Soc. **2017**, *139*, 13126 – 13140.

Poster Presentations

R. Pollice, P. Chen:
A Universal Quantitative Descriptor of Dispersive Interaction Potential
GRC Physical Organic Chemistry 2019, Holderness, NH, USA

R. Pollice, P. Chen:
A Universal Quantitative Descriptor of Dispersive Interaction Potential
GRS Physical Organic Chemistry 2019, Holderness, NH, USA

R. Pollice, P. Chen:
Trends in Intermolecular Dispersion Explain Differences in Properties of Alkanes and Perfluoroalkanes
International Symposium on Fluorine Chemistry 2018, Oxford, UK

R. Pollice, P. Chen:
Competing Bond Dissociation Between Covalent Bonds and A Dispersion-Supported Hydrogen Bond
International Symposium on Reactive Intermediates and Unusual Molecules 2018, Ascona, CH

R. Pollice, I. G. Shenderovich, P. Chen:
Attenuation of London Dispersion in Dichloromethane Solutions
GRC Physical Organic Chemistry 2017, Holderness, NH, USA

R. Pollice, I. G. Shenderovich, P. Chen:
Attenuation of London Dispersion in Dichloromethane Solutions
GRS Physical Organic Chemistry 2017, Holderness, NH, USA

Oral Presentations

R. Pollice, P. Chen:
A Universal Quantitative Descriptor of Dispersive Interaction Potential
GRC Physical Organic Chemistry 2019, Holderness, NH, USA

R. Pollice, P. Chen:
A Universal Quantitative Descriptor of Dispersive Interaction Potential
GRS Physical Organic Chemistry 2019, Holderness, NH, USA

Abstract

London dispersion is one of the basic interactions present between all atoms. Efficient theoretical models have been devised to estimate the dispersive interaction energies of atoms and molecules in the gas phase and it has been demonstrated that it is crucial even for the conformational stability of molecules possessing less than 200 atoms both in the gas phase and in solution. However, the knowledge about quantitative contributions of dispersion in small molecular systems largely stems from computational rather than experimental data due to the paucity of the latter. In addition, it is not clear how important dispersion is for interactions in solution because of a lack of suitable experimental studies.

We begin this thesis with an in-depth study of the importance of London dispersion for the BDE of proton-bound dimers in the gas phase and in DCM. First, we establish that proton-bound dimers with alkyl and aryl substituents are good test systems for the strength of dispersion both in the gas phase and in solution. Using this class of compounds, our main finding is that compensation of dispersive interactions is about 70 % with respect to the gas phase, and that widely employed continuum solvation models underestimate this compensation in the investigated proton-bound dimers significantly.

Following up on the previous chapter, we extend the investigation of compensation of London dispersion both in the gas phase and in solution by accounting for conformational flexibility of the molecules in the gas phase and by extending the experimental data set to a range of both polar and apolar aprotic organic solvents. Based on the extended data gathered, we demonstrate that the total compensation of dispersion is essentially independent of the solvent bulk polarizability and we are able to decompose the total compensation into gas-phase compensation, which amounts to 0 %–40 %, and solvent compensation, which reaches about 40 %–80 % in a temperature range of −100 °C–80 °C. Further-

more, we confirm the hypothesis that we developed in the previous chapter that implicit solvent models underestimate solute-solvent dispersion for a wide range of solvents systematically.

Our results of the compensation of dispersion in 1H-perfluorohexane were surprising and made us think about fluorous solvents in more detail. Hence, we decided to perform an in-depth computational investigation of the interactions of alkanes and perfluoroalkanes. Fluorocarbons are widely assumed to have low intermolecular attraction compared to alkanes in order to rationalize their low ability to dissolve organic molecules of similar size, their low surface tensions and their low boiling points. While the current description of perfluoroalkanes can explain many of their properties in a qualitative manner, many trends in their properties as chain length is increased cannot be understood. We carry out an extensive computational investigation of intermolecular complexes of alkanes and perfluoroalkanes and find that perfluoroalkanes inherently have a higher propensity for intermolecular dispersion per interaction compared to analogous alkanes. In addition, we also show that the largest part of the dispersion originates from close fluorine–fluorine contacts. Furthermore, we demonstrate that the preferred structures of perfluoroalkanes are not well-suited for intermolecular complexes with either perfluoroalkanes or alkanes. The combination of these properties and characteristics results in a dependence of the physical properties of fluorocarbons on their chain length as geometries for intermolecular interactions get increasingly poor and the relative content of fluorine in the corresponding fluorocarbons decreases.

We finish this thesis with introducing a generally applicable descriptor of dispersive interaction potential and its application to various chemical problems. The recognition of London dispersion among chemists is still behind electrostatic interactions. This can be rationalized by the lack of simple tools to understand London dispersion, whilst there are numerous simple ways to understand electrostatics like partial charges and electrostatic potential (ESP) maps. To overcome this problem, we introduce a quantitative descriptor of dispersion interaction potential, which can be used to construct London dispersion potential (LDP) maps. We show that this measure can be readily computed and applied to numerous problems related to NCIs in organic and organometallic systems.

Zusammenfassung

London Dispersion ist eine der fundamentalen Wechselwirkungen zwischen allen Atomen. Effiziente theoretische Modelle wurden entwickelt um dispersive Wechselwirkungsenergien zwischen Atomen und Molekülen in der Gasphase abzuschätzen und es wurde gezeigt, dass Dispersion auch für die Stabilität von Molekülen mit weniger als 200 Atomen wichtig ist. Der Wissensstand über den quantitativen Anteil der Dispersion stammt jedoch hauptsächlich von Berechnungen und nicht von Experimenten, da letztere diesbezüglich nur spärlich verfügbar sind. Zusätzlich dazu ist es nicht klar, wie wichtig Dispersion in Lösung ist, da entsprechende experimentelle Studien ebenfalls noch fehlen.

Wir beginnen diese Arbeit mit einer Detailstudie zur Wichtigkeit von London Dispersion für die Bindungsdissoziationsenergien von Proton-gebundenen Dimeren in der Gasphase und in Dichlormethan. Zuallererst zeigen wir, dass Proton-gebundene Dimere mit Alkyl- und Arylsubstituenten gute Testsysteme für Dispersion sowohl in der Gasphase als auch in Lösung sind. Unsere wichtigste Beobachtung ist, dass die Kompensation der dispersiven Wechselwirkungen in Dichlormethan etwa 70 % im Vergleich zur Gasphase beträgt. Weiters beobachten wir, dass vielfach verwendete Kontinuumssolvatisierungsmodelle die Kompensation von Dispersion in den untersuchten Molekülen signifikant unterschätzen.

Im nächsten Kapitel vertiefen wir die Untersuchungen bezüglich Kompensation von London Dispersion in derselben Klasse von Molekülen, indem wir die Flexibilität der Moleküle in der Gasphase explizit berücksichtigen, und indem wir die experimentellen Daten auf eine Reihe von polaren und apolaren Lösungsmitteln ausweiten. Auf Basis dieser Daten zeigen wir, dass die Gesamtkompensation der Dispersion quasi unabhängig von der Polarisierbarkeit der kondensierten Phase ist. Weiters können wir die Gesamtkompensation in einen Gasphasenanteil, der etwa 0 %–40 % beträgt, und einen Solvatisierungsanteil, der etwa 40 %–80 % be-

trägt, zerlegen.

Wir setzen unsere Reise mit einer detaillierten rechnerunterstützten Untersuchung der Wechselwirkungen von Alkanen und Perfluoralkanen fort. Im Allgemeinen wird angenommen, dass Fluorkohlenstoffe schwache intermolekulare Anziehungskräfte im Vergleich zu Alkanen besitzen, was wiederum ihre geringe Fähigkeit Stoffe ähnlicher Grösse im Allgemeinen aufzulösen wie auch ihre geringen Oberflächenspannungen und Siedepunkte erklärt. Obwohl die gegenwärtige Beschreibung von Perfluoralkanen viele ihrer Eigenschaften qualitativ erklären kann, so können viele Trends in ihren Eigenschaften, zum Beispiel wenn die Kettenlänge steigt, nicht zufriedenstellend verstanden werden. Wir führen umfangreiche Berechnungen von intermolekularen Komplexen von Alkanen und Perfluoralkanen durch und beobachten, dass letztere eine inhärent höhere Neigung für intermolekulare dispersive Wechselwirkungen besitzen. Weiters zeigen wir, dass der Grossteil der Dispersion von kurzen Fluor-Fluor-Kontakten stammt. Zusätzlich dazu legen wir dar, dass die bevorzugte Struktur von Perfluoralkanen nicht gut für intermolekulare Komplexe sowohl mit Perfluoralkanen als auch mit Alkanen geeignet ist. Die Kombination dieser Eigenschaften führt dazu, dass die physikalischen Eigenschaften von Fluorkohlenstoffen von ihrer Kettenlänge abhängen weil ihre Strukturen immer ungeeigneter für gute Wechselwirkungen werden und ihr relativer Fluorgehalt abnimmt.

Wir schliessen diese Arbeit mit der Einführung eines allgemein anwendbaren Deskriptors von dispersivem Wechselwirkungspotential und dessen Anwendung. Die allgemeine Anerkennung der London Dispersion unter Chemikern is weit hinter der klassischen Elektrostatik. Das kann durch einen Mangel an einfachen Werkzeugen für ihr Verständnis erklärt werden, wohingegen einfache Konzepte zum Verständnis von Elektrostatik wie zum Beispiel Partialladungen und das molekulare elektrostatische Potential weit verbreitet sind. Um diese Probleme zu überwinden führen wir einen quantitativen Deskriptor für dispersives Interaktionspotential ein, mit dessen Hilfe auch das molekulare dispersive Potential definiert werden kann. Wir zeigen, dass dieser Deskriptor mit geringem Aufwand ermittelt werden kann und, dass er dazu verwendet werden kann unzählige Probleme im Bereich der nichtkovalenten Wechselwirkungen zu studieren.

Contents

1. Introduction **1**

1.1. London Dispersion 1

1.1.1. Classical Models of London Dispersion 2

1.1.2. Modern Treatment of London Dispersion 7

1.1.3. Manifestations of London Dispersion 14

1.1.4. Compensation of London Dispersion 17

1.2. Energy Decomposition Analysis 18

1.2.1. Deformation Energy 18

1.2.2. Symmetry Adapted Perturbation Theory 19

1.2.3. Local Energy Decomposition 20

1.3. Visualization of Interactions 22

1.3.1. Electrostatic Potential 22

1.3.2. Noncovalent Interactions 23

1.3.3. Dispersion Interaction Density 23

1.3.4. Hirshfeld and Becke Surfaces 24

1.4. Solvent Models 25

1.4.1. Implicit Solvent Models 26

1.4.2. Explicit Solvent Models 28

2. Compensation of Dispersion in Dichloromethane **31**

2.1. Introduction 31

2.2. Results 32

2.2.1. Model System 32

2.2.2. Theoretical Investigation of Gas-Phase Dissociation 35

2.2.3. Experimental Investigation of Dissociation in Solution . . . 45

2.2.4. Comparison of Computed and Experimental Energies . . . 49

2.3. Discussion 51
2.3.1. Theoretical Investigation of Gas-Phase Dissociation 52
2.3.2. Experimental Investigation of Dissociation in Solution . . . 56
2.3.3. Comparison of Computed and Experimental Energies . . . 61
2.4. Conclusions 65

3. Compensation of Dispersion in Aprotic Solvents 67
3.1. Introduction 67
3.2. Results 69
3.2.1. Model Compounds and Solvents 69
3.2.2. Gas-phase Compensation of Dispersion 71
3.2.3. Compensation of Dispersion in Solution 73
3.2.4. Separation of Compensation of Dispersion 81
3.2.5. Estimation of Gibbs Free Energies in Solution 83
3.3. Discussion 86
3.3.1. Gas-phase Compensation of Dispersion 86
3.3.2. Compensation of Dispersion in Solution 89
3.4. Conclusions 99

4. Immiscibility of Alkanes and Perfluoroalkanes 101
4.1. Introduction 101
4.2. Results 102
4.2.1. Test Systems 102
4.2.2. Benchmarking 103
4.2.3. Energy Decomposition Analysis 107
4.2.4. Simple Interaction Models 112
4.2.5. Interaction Descriptors 114
4.2.6. Interaction Geometries 117
4.2.7. Computed Mixing Behavior 120
4.2.8. Trends in Experimental Properties 123
4.3. Discussion 126
4.3.1. Benchmarking 126
4.3.2. Energy Decomposition Analysis 130
4.3.3. Simple Interaction Models 132

4.3.4. Interaction Descriptors 133
4.3.5. Interaction Geometries 134
4.3.6. Computed Mixing Behavior 137
4.3.7. Trends in Experimental Properties 139
4.4. Conclusions 141

5. Universal Descriptor of Dispersive Potential 143
5.1. Introduction 143
5.2. Results 144
5.2.1. Methodology 144
5.2.2. Applications 145
5.3. Discussion 165
5.3.1. Methodology 166
5.3.2. Applications 169
5.4. Conclusions 172

6. Summary and Outlook 175

A. Experimental Details 179
A.1. Compensation of Dispersion in Dichloromethane 179
A.1.1. General Experimental Methods 179
A.1.2. Qualitative ESI-MS Measurements 180
A.1.3. NMR Measurements 180
A.1.4. Synthesis of Test Substances 183
A.1.5. X-ray Crystal Structure of the Proton-bound Dimer of 2,6-Dibenzylpyridine 219
A.2. Compensation of Dispersion in Aprotic Solvents 219
A.2.1. General Experimental Methods 219
A.2.2. Isothermal Titration Calorimetry Evaluation 221
A.2.3. NMR Measurements 223
A.2.4. Variable-Temperature NMR Evaluation 226
A.2.5. Synthesis of Test Substances 227

B. Computational Details 235
B.1. Compensation of Dispersion in Dichloromethane 235

B.2. Compensation of Dispersion in Aprotic Solvents 240
B.3. Immiscibility of Alkanes and Perfluoroalkanes 243
B.3.1. Geometry Optimizations for Benchmarking 244
B.3.2. Single Points 244
B.3.3. Energy Decomposition Schemes 248
B.3.4. Visual Analyses 249
B.3.5. Additional Analyses 249
B.4. Universal Descriptor of Dispersive Potential 250

C. Mathematical Derivations **255**
C.1. Compensation of Dispersion in Dichloromethane 255
C.1.1. Master Equation of Proton-bound Dimer Equilibria 255
C.2. Immiscibility of Alkanes and Perfluoroalkanes 259
C.2.1. Homomolecular Cuboid Model for Interaction Surfaces . . . 259
C.2.2. Heteromolecular Cuboid Model for Interaction Surfaces . . 260

D. Curriculum Vitae **309**

1. Introduction

We begin this thesis with a short introduction about London dispersion and a few theoretical models and techniques we will be using throughout this work.

1.1. London Dispersion

London dispersion, the attraction between instantaneous dipoles and induced dipoles, is omnipresent in all molecules and is one of numerous so-called non-covalent interactions.[1] The interplay of all these non-covalent interactions determines molecular structure and the associated energies. Based on the London dispersion formula,[2] it can be derived that it depends on the distance between atoms, on atomic polarizabilities and on atomic ionization potentials. In molecules, it can be approximated by the sum of atom-pairwise interaction terms,[3,4] which directly implies that the corresponding interaction count scales with the square of the number of atoms. Alternatively, dispersion has also been approximated as linear function of interaction surface as the steep R^{-6}-distance dependence renders all but direct dispersive contacts essentially unimportant.[5,6] Therefore, it is to be expected that London dispersion reaches substantial energies of several kilocalories per mole in molecules of considerable sizes with several tens or hundreds of atoms as both atom count and interaction surfaces are typically higher in larger molecules. Dispersion was recently invoked to explain the surprising thermal stability of bis(diamondoid) hydrocarbons with carbon-carbon bonds of up to 1.70 Å.[7] This test system showcases that organic molecules of relatively moderate size can benefit from significant dispersive stabilization.

The fact that the scaling of dispersive interactions is somewhat higher than linear and their ubiquitousness make them attractive as general design elements

for both molecular structure and property. This has been first proposed by Grimme invoking the concept of dispersion energy donors (DEDs).[8] Dispersion is crucial for the accurate description of matter and its phase changing behavior in general,[9,10] which is evidenced by the recognition of the Nobel Prize of physics in 1910 for the van der Waals equation of state.[11] Despite that, for a long time the attractive dispersive forces between molecules in solution were largely neglected. The general consensus was that large groups at short distances in solution lead solely to steric repulsion. Therefore, it has been pointed out in the literature that considerable dispersive contributions to interaction energies are only expected in the gas phase.[7,12–15] The main reason for that is that attractive intramolecular dispersion in solute molecules competes with solute-solvent dispersion as dispersive interaction surface freed from intramolecular interactions is immediately covered by intermolecular interactions with the solvent. Effectively, this makes dispersion less important in solution compared to the gas phase.[14] The main question in that regard is whether the compensation is complete or not, as a full compensation would render DEDs for the design of intermolecular interactions in solution null. Furthermore, if compensation is not complete, the question is to what extent dispersion is compensated in solution.

In this section, we will investigate classical models of London dispersion and then see how they developed into the currently used models. We will also have a short section about manifestations of London dispersion, in which we will study experimental observations closely related to dispersion. We will close this section about London dispersion with the investigation of effective compensation of dispersion, where we will study effects that can result in dispersion being weaker or less important in a real system than one would estimate from simplified theoretical models.

1.1.1. Classical Models of London Dispersion

A simple model of London dispersion was provided by London.[1,2,16] It is based on the interaction between two three-dimensional isotropic spherically symmetric harmonic oscillators with no permanent dipole moment in the rest position, each having a polarizability α. Let us assume that the charges e of the two oscillators are displaced from the rest position by the following displacement vectors:

$$\vec{r_I} = (x_I, y_1, z_I) \text{ and } \vec{r_{II}} = (x_{II}, y_{I1}, z_{II}) \tag{1.1}$$

The potential energy of the system after displacement can be expressed in terms of the elastic energy and the dipole interaction energy as follows:

$$V = \frac{e^2 r_I^2}{2\alpha} + \frac{e^2 r_{II}^2}{2\alpha} + \frac{e^2}{R^3}(x_I x_{II} + y_I y_{II} - 2 z_I z_{II}) \tag{1.2}$$

In a classical description, the two oscillators at finite distance in their respective equilibrium positions would not interact with each other. In a quantum mechanical description, both oscillators have a zero-point energy corresponding to a zero-point vibration. These zero-point vibrations of the two harmonic oscillators can interact with each other even as the oscillators remain on average in their equilibrium positions. The zero-point energy can be expressed based on the frequency of the harmonic oscillator:

$$E_0 = \frac{1}{2} h\nu \tag{1.3}$$

We change the coordinate system to reduced coordinates:

$$\vec{r_+} = \begin{cases} x_+ = \frac{1}{\sqrt{2}}(x_I + x_{II}) \\ y_+ = \frac{1}{\sqrt{2}}(y_I + y_{II}) \\ z_+ = \frac{1}{\sqrt{2}}(z_I + z_{II}) \end{cases} \text{ and } \vec{r_-} = \begin{cases} x_- = \frac{1}{\sqrt{2}}(x_I - x_{II}) \\ y_- = \frac{1}{\sqrt{2}}(y_I - y_{II}) \\ z_- = \frac{1}{\sqrt{2}}(z_I - z_{II}) \end{cases} \tag{1.4}$$

The potential energy expression in the reduced coordinate system can be derived by substituting the expression from Equation 1.4 into Equation 1.2:

$$V = \frac{e^2}{2\alpha}(r_+^2 + r_-^2) + \frac{e^2}{2R^3}(x_+^2 + y_+^2 - 2z_+^2 - x_-^2 - y_-^2 + 2z_-^2) \tag{1.5}$$

Rearrangement of the terms gives the following expression for the potential energy:

$$V = \frac{e^2}{2\alpha}\Bigg[\left(1+\frac{\alpha}{R^3}\right)(x_+^2+y_+^2)+\left(1-\frac{\alpha}{R^3}\right)(x_-^2+y_-^2) + \left(1-2\frac{\alpha}{R^3}\right)z_+^2+\left(1+2\frac{\alpha}{R^3}\right)z_-^2\Bigg] \tag{1.6}$$

Effectively, this is the sum of the potential energies of six independent harmonic oscillators. The corresponding frequencies can be expressed in terms of the frequency $\nu_0 = \frac{e}{\sqrt{m\alpha}}$ of the two systems, in case they are at infinite distance and m is the corresponding reduced mass:

$$\begin{aligned} \nu_{x_\pm} = \nu_{y_\pm} &= \nu_0\sqrt{1\pm\frac{\alpha}{R^3}} \\ \nu_{z_\pm} &= \nu_0\sqrt{1\mp 2\frac{\alpha}{R^3}} \end{aligned} \tag{1.7}$$

For $\alpha \ll R^3$, these expressions can be developed into a power series:

$$\begin{aligned} \nu_{x_\pm} = \nu_{y_\pm} &\approx \nu_0\left(1\pm\frac{\alpha}{2R^3}-\frac{\alpha^2}{8R^6}\pm\ldots\right) \\ \nu_{z_\pm} &\approx \nu_0\left(1\mp\frac{\alpha}{R^3}-\frac{\alpha^2}{2R^6}\mp\ldots\right) \end{aligned} \tag{1.8}$$

The ground state of the whole system can be derived as follows:

$$E_0 = \frac{1}{2}h\left(\nu_{x_+}+\nu_{y_+}+\nu_{z_+}+\nu_{x_-}+\nu_{y_-}+\nu_{z_-}\right) \tag{1.9}$$

By using the expression of Equation 1.8 the following expression for the total energy of the system can be derived:

$$E_0 = 3h\nu_0 - \frac{3}{4}\frac{h\nu_0\alpha^2}{R^6}+\ldots \tag{1.10}$$

The first term on the right hand side is simply the zero-point energy of the total system, which is the zero-point energy for six independent harmonic oscillators.

The second term is an interaction energy and can be identified as the London dispersion energy E_{disp} between the two three-dimensional harmonic oscillators at large distances R:

$$E_{disp} = -\frac{3}{4} h\nu_0 \frac{\alpha^2}{R^6} \tag{1.11}$$

In case the two interacting systems do not have the same polarizabilities, the corresponding formula can be written in terms of the respective polarizabilities α_1 and α_2. In addition, as a crude approximation, the zero-point energy $\frac{h\nu_0}{2}$ can be approximated using the ionization potentials of the two systems I_1 and I_2. This approximation can be shown to be an upper boundary for the intermolecular dispersion between two atoms, and was empirically shown to perform better than an analogous lower boundary for the intermolecular dispersion based on electronic excitation energies.[1] As a result, the well-known London dispersion formula is obtained:

$$E_{disp} = -\frac{3}{2} \frac{I_1 I_2}{I_1 + I_2} \frac{\alpha_1 \alpha_2}{R^6} \tag{1.12}$$

Equation 1.12 describes the interaction in the asymptotic limit of large distances. To derive a formula for the short-range limit, a different approach needs to be taken. Koide derived the London dispersion formula using a different model which provides a convergent series in physical properties of the separated atoms.[17] The main difference to the London derivation is that the perturbation operator for the electric interaction between the two atoms is expanded into interactions between spherical waves. Accordingly, polarizabilities can also be attributed to these spherical waves. The mathematical derivation is outside the scope of this introduction but we will summarize the corresponding results. Using that approach, it can be shown that the dispersion interaction energy converges to the well known London dispersion formula (Equation 1.12) in the limit of infinite distance. When the distance approaches zero, the dispersion energy converges to a parabola. The mathematical limits are shown in the following equations:

$$E_{disp}(R) = \begin{cases} R \to 0 : -A + BR^2 \\ R \to \infty : -\frac{C}{R^6} \end{cases} \tag{1.13}$$

Hence, the interatomic dispersion energy converges to a finite value at a distance of zero. The dispersion energy as a function of interatomic separation is illustrated schematically in Figure 1.1.

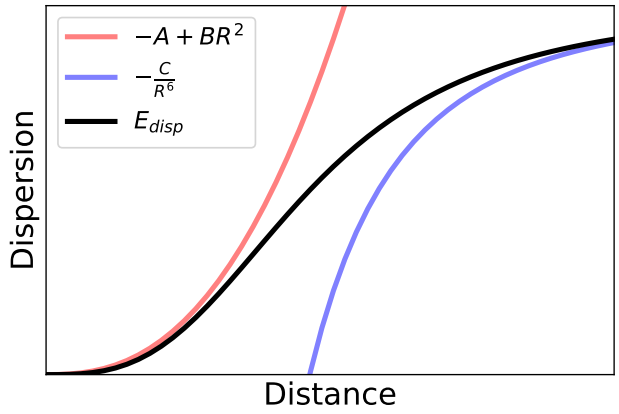


Figure 1.1.: Asymptotic behavior of distance dependence of interatomic dispersion energy.

Finally, we will briefly investigate an alternative model of London dispersion, the Feynman model. It can be considered as an alternative to the fluctuating dipole picture of the London model. In the Feynman model, as two atoms approach each other, at an internuclear separation that is large compared to the radii of the atoms, the electron density will redistribute so that it accumulates between the nuclei. The charge distribution of each atom changes from spherical symmetry to σ-symmetry. This induces a dipole moment on the order of $\frac{1}{R^7}$ in each of the two atoms. However, in the Feynman model, it is not the interaction between the induced dipoles that is responsible for the interatomic interaction. The attractive force arises from the attraction of the nucleus for its own electrons that results in the attractive $\frac{1}{R^7}$ force, which in turn is responsible for the attractive $\frac{1}{R^6}$ energy term. Both models provide the same distance dependence of dispersion in the long-range regime. However, the physical interpretation of the two models is different. Both models have their merits and can be used to interpret dispersion forces in particular systems.

1.1.2. Modern Treatment of London Dispersion

Having investigated classical models of London dispersion, we move on to a brief investigation of modern methods to account for London dispersion in computations. London dispersion is in its essence an electron correlation effect as it arises from the correlated fluctuations in electron density. Since HF neglects electron correlation completely, it does not account for London dispersion at all. In addition, standard functionals employed in DFT do not account for a large portion of dispersion because the energy functionals are usually described as function of local density properties like the value of the local density in LDA and both the value and the gradient of the local density in GGA functionals. These approximations result in systematic underestimation of the electron density at larger distances to the nuclei. As hybrid functionals make use of both HF and standard exchange-correlation functionals, they cannot account for London dispersion either. Therefore, both standard HF and standard DFT need additional ways to account for London dispersion.[18]

The default way to account for London dispersion naturally is by using so-called post-HF methods in the framework of WFT, which approximate the neglected electron correlation as accurately as possible. One of the computationally most affordable methods, which is also the reason it is one of the most popular WFT methods, is MP2. It starts with the HF wavefunction as zero-order approximation and treats the changes upon inclusion of electron correlation perturbatively. The first-order perturbation is zero and, hence, the second-order perturbation is the first non-trivial term in the perturbation expansion making MP2 the cheapest possible MP method. MP is known to give good results for the dispersion between saturated hydrocarbons but largely overestimates the dispersion between unsaturated systems.[19,20] To alleviate the problems of MP2, SCS-MP2 was developed as empirical correction for systematic improvement over MP2.[21] The MP2 correlation energy is separated into same-spin and opposite-spin contributions and these two energy components are scaled with empirical scaling factors derived from benchmark calculations of higher level of theory.[21] It was found that SCS-MP2 also improves noncovalent interaction energies.[22] It leads to significantly better results for interactions between unsaturated systems but gives worse results for saturated systems which leads to an underestimation of dispersion by

about 10 %–20 % compared to benchmark calculations.[18,21,22] Higher order MP methods can also be used to describe dispersion, but their typically only minor improvement of accuracy is not justified by the associated significant increase in computational cost.[23] One of the drawbacks of MP2 is that it typically requires rather large basis sets to be sufficiently accurate, which makes it quite expensive compared to HF and standard DFT calculations.

The WFT methods of choice for benchmarking of noncovalent interactions, and hence also to compute London dispersion, are CC methods with the current gold standard being CCSD(T).[24,25] For benchmark results, proper account of the BSSE and the BSIE are absolutely vital to obtain good results and, hence, either large basis sets combined with focal point analysis are required or explicitly correlated methods for treating basis set errors need to be carried out.[24–28] The main problem is that even the cheapest approaches giving accurate results require CCSD(T) single points at the augmented double zeta or at the triple zeta level, which becomes prohibitively expensive as system size exceeds about 20 non-hydrogen atoms. Therefore, accurate benchmark data is only available for small molecules, which is still a major challenge when scaling of dispersion with size wants to be investigated accurately. Due to their large cost, the conventional approach for the development of accurate methods to describe dispersion is to obtain accurate benchmark data for small molecules and intermolecular complexes with a wide variety of dispersive binding motifs, and use this benchmark data for the calibration of cheaper computational approaches which are then used to study large systems. Especially when new systems are investigated, this is the preferred approach. However, despite careful benchmarking, small deviations cannot be avoided when cheaper methods are used. In addition, when parameters are used, there are always uncertainties in the calibrated parameters. This means that seemingly small errors can accumulate to chemically significant deviations when the same binding motifs are combined multiple times to form large intermolecular complexes.

While WFT methods like MP2 and CCSD(T) offer a systematic approach to treat London dispersion in molecules, their computational cost is too prohibitive even for state-of-the-art computers to use them for large systems with more than 100 non-hydrogen atoms. This was realized already some time ago and lead

to the development of local correlation methods,[29] as it has been realized that the electron correlation covered by MP and CC methods is rather short-range.[30] They are based on localized molecular orbitals, which are as much localized in a minimum volume as possible. The drawback is that the resulting orbitals do not have well-defined energies.[29] The theoretical formulation of the WFT methods is retained; the only difference is the different orbital space used. The main difference of local correlation methods is that the localized orbital representation allows reduction in computational cost by neglecting correlation of pairs between distant orbitals and by reducing the dimension of the virtual space, only allowing excitations of localized orbitals to spatially close virtual levels.[29] The advantage of being able to treat larger systems due to the significant computational savings comes at the cost of systematic accuracy as arbitrary cut-off parameters need to be introduced. The field of local correlation methods developed quickly and many methods have been devised. An exhaustive summary is beyond the scope of this thesis. However, one particular method has emerged from all these developments, namely DLPNO-CCSD(T), which is a local variant of CCSD(T) promising similar results to the canonical method at a fraction of the computational cost with significantly lower computational size scaling.[31–36] In a series of two papers first DLPNO-CCSD[30] and then DLPNO-CCSD(T)[37] were described in 2013. Recently, an improved triples correction was developed to alleviate some of the problems observed in DLPNO-CCSD(T).[38] DLPNO-CCSD(T) relies on a small set of empirically derived thresholds, which determine the accuracy of the method. Currently, it is considered the method of choice for benchmarking noncovalent interactions of rather large intermolecular complexes. However, its agreement with canonical CCSD(T) can only be tested based on small systems, which begs again the question how the approximations inherent in the DLPNO method scale with system size.

Alternatively, a large amount of research was dedicated in recent years to DFT methods with proper account of London dispersion.[39,40] Recently, the developed approaches have been grouped according to level of sophistication in the ”stairway to heaven”, analogously to the Jacob’s ladder to classify functionals, and we will follow this grouping in the following account of DFT methods with treatment of London dispersion.[41]

The first level is occupied by methods accounting for dispersion locally with wrong asymptotic behavior.[41] They typically provide a reasonable description of dispersion in short-range binding situations close to the respective equilibrium geometries but the distance dependence as distances are increased is not properly accounted for. Popular examples of methods taking this approach are Minnesota functionals like M06-L and the complete suite of M06 functionals,[42,43] which perform surprisingly well for many systems.[44] These functionals account for London dispersion by making use of an appropriate semilocal functional form. They use the local electron density and its gradients to describe the correlation energy. As a consequence, the correlation energy shows an exponential decay as distance is increased and, hence, the long-range distance dependence of this energetic contribution does not correspond to R^{-6}.[3] For M06-L, the corresponding exchange-correlation functional is expressed in terms of the local spin density, its derivative and the local kinetic energy. The main idea is to have a very flexible functional form so that the functional parameters, there are 32 in the case of M06-L, can be fitted to a wide range of benchmark energies including noncovalent interactions. Overall, M06-L reproduces the interaction energies of noncovalent complexes at equilibrium energies reasonably well.[3]

The second level corresponds to DFT methods including an additional simple energy term for dispersion, which is determined from atom-pairwise dispersive interactions over all atom pairs in a molecule or molecular complex.[41] The corresponding approach can be described using the following two formulas:[41]

$$E_{tot} = E_{DFT} + E_{disp} \tag{1.14}$$

$$E_{disp}^{A,B} = -\sum_{A,B} \frac{C_6^{A,B}}{R_{A,B}^6} \tag{1.15}$$

Dispersion methods at the second level are limited to both isotropic and environment-independent dispersion coefficients and the most popular example of such a method is the DFT-D approach.[45] These methods improve relative to the first level by providing an appropriate asymptotic long-range description of dispersion with the correct distance dependence. This method gained large popularity due to its extremely low cost and reasonable results. Also its im-

mediate successor, the DFT-D2 approach is a method of the second level on the ”stairway to heaven” of DFT methods accounting for London dispersion. An important characteristic of dispersion corrections at the second level is the divergence of dispersion energies when interatomic distances approach zero.[41] This singularity needs to be removed and for that purpose damping schemes were developed. When damping functions (typically denoted as f) are included, the expression for the dispersion energy changes formally from Equation 1.15 to the following:

$$E_{disp}^{A,B} = -\sum_{A,B} \frac{C_6^{A,B}}{R_{A,B}^6} f(A, B, R_{A,B}) \tag{1.16}$$

The damping function needs to be parametrized against the specific functional used in conjunction with the dispersion correction. The standard approach in that regard is that the damping function parameters are optimized to reproduce benchmark results for dispersion-dominated complexes obtained from CCSD(T) calculations. This is important as the dispersion correction results are very sensitive to the particular damping function used because the net interaction energies in the relevant distance regimes are the result of small differences between large numbers.[41]

The third level is a direct improvement of the second level by allowing the dispersion coefficients to be environment-dependent and currently represents the most widely used approach. The most popular methods in that regard are the Tkatchenko-Scheffler[46] dispersion model, the exchange-dipole moment[47–51] (XDM) dispersion model and the DFT-D3[52] approach. A somewhat less popular example is the density-based dDsC dispersion model.[53,54] All of these methods account for the local changes in dispersion coefficients in a different way. However, the qualitative effect is that dispersion coefficients get smaller as the effective volume of the atoms decrease. In addition to the C_6 dispersion coefficients employed in the second level, higher order dispersion coefficients like the C_8 dispersion coefficient (DFT-D3, XDM) or the C_{10} dispersion coefficient (XDM) are also included in the third level which leads to the following general formula for the dispersion energy where n accounts for the dispersion coefficients included in the respective expansion:

$$E_{disp}^{A,B} = -\sum_{A,B}\sum_{n}\frac{C_n^{A,B}}{R_{A,B}^n} f(A,B,R_{A,B}) \tag{1.17}$$

The advantage of the DFT-D3 approach over all the other ones is that it is only dependent on atomic positions and hence, does not require information about the electron density in any way making it the computationally cheapest approach. It determines the number of direct neighbors of each atom and estimates the resulting dispersion coefficient as a function of direct neighbors from a set of reference values for each element.[41,52] The TS and XDM methods account for local changes in dispersion coefficients by computing the volume of atoms in molecules from electron densities to scale the reference free atomic dispersion coefficients. Hence, both these methods require also the electron density as input. Furthermore, recently, DFT-D4 was developed as successor of DFT-D3.[55,56] DFT-D3 does not account properly for charged systems. This is illustrated by the fact that the exact same geometry of a molecule with different total charge gives exactly the same dispersion correction. This is non-physical as it would also give the same dispersion energy if all the electrons were removed from the corresponding system. While this extreme scenario is usually not a system of interest, it does illustrate the inherent problems of DFT-D3. This was alleviated in DFT-D4 which, based on molecular geometries, determines classical atomic charges from an electronegativity equalization procedure and scales the atomic dispersion coefficients based on these atomic charges.[55,56] It improves on the accuracy of DFT-D3 without requiring electron density as input making it only slightly more expensive than DFT-D3 but still cheaper than TS, XDM or dDsC. The biggest improvements of DFT-D4 over DFT-D3 are found for metal complexes, especially for transition metal complexes.[57]

Beyond the third level of the ”stairway to heaven” for density functionals, there are two different approaches improving on the accuracy of the previous levels.[41] The first one is going beyond pairwise additivity. The second and third level only include atom-pairwise dispersion energy terms. When going to ever larger systems approaching condensed matter, many-body dispersion terms become significant. For three-body dispersion interactions, the Axilrod-Teller-Muto[58,59] formula is well-established and the corresponding energy contribution can be

calculated using the following formula:

$$E_{disp}^{A,B,C} = C_9^{A,B,C} \frac{3cos(\alpha)cos(\beta)cos(\gamma)+1}{R_{A,B}^3 R_{B,C}^3 R_{A,C}^3} \tag{1.18}$$

The angles α, β and γ are defined as follows:

$$\begin{aligned} \alpha &= \measuredangle CAB \\ \beta &= \measuredangle ABC \\ \gamma &= \measuredangle BCA \end{aligned} \tag{1.19}$$

The three-body dispersion formula can be readily used with the dispersion correction methods of the third level to improve upon the results and it exists as an optional energy contribution in DFT-D3.[3] For small molecules, the three-body contribution, as well as additional many-body contributions are negligible. However, for large intermolecular complexes and molecular crystals, the contribution becomes significant and including three-body dispersion improves the corresponding results.[3] This inspired the development of the many-body dispersion (MBD) correction, which is based on the model of coupled dipoles.[60] It describes the atoms as collection of three-dimensional, isotropic quantum harmonic oscillators and was found to yield accurate results for intermolecular dispersion interactions for both small and large systems.[3] Due to the excellent performance of the MBD correction for large systems[3,4,56,60] the DFT-D4 approach offers full account of many-body dispersion effects arising from dipole-dipole interactions up to infinite order, which should be used as additional correction on top of the atom-pairwise contribution and the three-body contribution and is recommended for large intermolecular complexes.[56] A more expensive method going beyond pairwise additivity promising even higher accuracy is RPA, which is based on the adiabatic-connection fluctuation-dissipation theorem and makes use of orbitals.[41] It can be considered a post-HF method and its scaling with the fourth power of the system size makes it a very expensive approach to account for dispersion.[41] Notably, DLPNO-CCSD(T) with NormalPNO parameters was shown to asymptotically scale like HF, which is close to the second power of the system size.[34] However, it was shown to provide excellent accuracy and is considered

one of the most accurate approaches to describe dispersion.[4] Notably, MBD can be considered as a semiclassical approximation of direct RPA.[3]

The second approach going beyond the third level of the "stairway to heaven" is using non-local density functionals. These approaches are direct DFT approaches as they add a non-local correlation functional to the description of the system. The two most notable approaches in that regard are the so-called vdW-DFs[61] and the VV10[62] methods. Due to their nonlocal nature in the DFT framework, they are quite expensive compared to most of the previously discussed dispersion correction methods, vdW-DFs are at least as expensive as GGA functionals for their evaluation, which is quite cheap, VV10 is somewhat more expensive than GGA functionals.[41] The vdW-DFs have been used quite frequently in the physics community for solid-state problems.[3] VV10 has experienced more widespread use in the chemistry community as it was shown that it can be readily combined with a wide variety of standard functionals in very much the same way as the D3 dispersion correction.[63] Hence, this approach was also termed DFT-NL.[3,63] Several functionals in conjunction with the VV10 dispersion correction provide very accurate results for the interaction energies of dispersion-bound systems.[63] Two notable developments in that regard are B97M-V and ω-B97M-V developed by Head-Gordon by systematically finding the best compromise between number of functional parameters and performance in a wide variety of standard benchmark data sets. They were shown to outperform many commonly used functionals, providing some of the most accurate results for DFT functionals with comparable computational cost.[64,65]

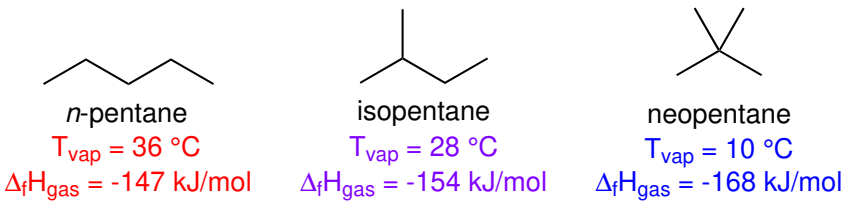
It should be noted that in this short section about modern ways to treat London dispersion in computations, we only mentioned the most popular approaches. For a more detailed overview explaining the underlying models in more detail, recently published comprehensive reviews on dispersion corrections should be consulted.[3,4]

1.1.3. Manifestations of London Dispersion

One of the direct manifestations of London dispersion that was recognized early is the description of real gases by means of the van der Waals equation of state, for which Johannes Diderik van der Waals was awarded the Nobel Prize

in Physics in 1910.[11] Notably, London dispersion is only one part of the van der Waals forces but it is the main contribution to the interactions present in gases without a permanent dipole moment like the noble gases or homoatomic diatomic gases. In that respect, London dispersion is responsible that noble gases have a boiling point and a melting point.

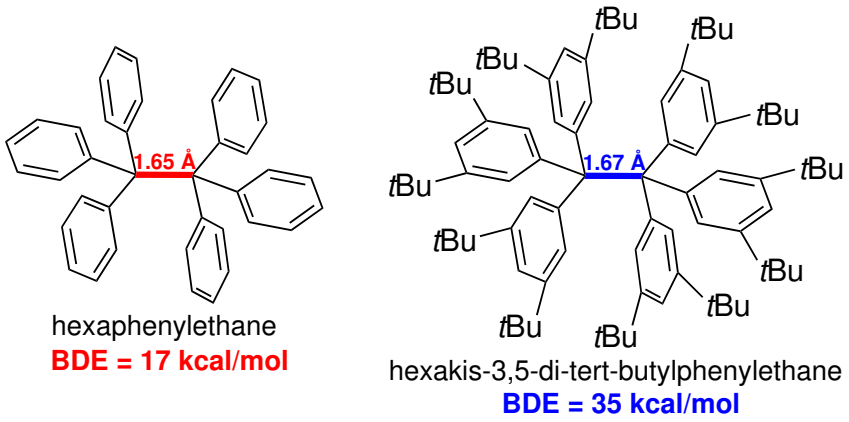
Another manifestation that is commonly taught in basic organic chemistry courses is the difference in boiling points of linear and branched alkanes.[18] This is illustrated based on the boiling points and the thermodynamic stabilities of the three pentane isomers in Scheme 1.1. The boiling point decreases as branching is increased. This can be explained because *n*-pentane has the largest surface area, which gradually decreases as branching increases and the molecule approaches a spherical shape. At the same time, it is usually emphasized accordingly that branched alkanes have higher thermodynamic stability.[18] It is important to point out that the trend in boiling points and thermodynamic stabilities of linear and branched alkanes are directly related. The decreasing surface area of the molecules is caused by an increase in intramolecular dispersive contacts in the branched alkanes. These dispersive contacts make branched alkanes thermodynamically more stable than the unbranched analogues and manifests itself in lower enthalpies of formation of the respective alkanes, which is also illustrated in Scheme 1.1.



Scheme 1.1: Boiling points and enthalpies of formation of pentanes as a function of branching. Data was taken from the NIST webbook.[66,67]

London dispersion can be used to stabilize otherwise labile molecules. In particular, *tert*-butyl groups can be very effective in doing so. This is reflected in the so-called ”hexaphenylethane riddle” concerning the molecules hexaphenylethane and hexakis-3,5-di-tert-butylphenylethane.[68] While the former has not

been isolated to date and claimed isolations have been shown to actually correspond to an unsymmetrical quinoid structure,[68] the latter has been successfully crystallized.[69] The molecules are illustrated in Scheme 1.2.



Scheme 1.2: Hexaphenylethane and hexakis-3,5-di-tert-butylphenylethane.

Based on calculations, hexaphenylethane is a stable molecule in the sense that it corresponds to a local minimum on the potential energy surface with a BDE of 17 kcal/mol and an estimated C-C bond length of 1.65 Å.[7,18] However, if ever present, at ambient temperatures it isomerizes rapidly to the quinoid structure.[18] The reason for its instability is the large steric repulsion around the central C-C bond caused by six phenyl groups in close proximity.[18] However, introducing 12 *tert*-butyl groups, which are sterically demanding, results in a much more persistent molecule, and the substituted analogue could be crystallized. The C-C bond length in hexakis-3,5-di-tert-butylphenylethane is 1.67 Å, which is significantly longer than the C-C bond in ethane of 1.54 Å.[69] This shows that, depending on the respective interatomic distance regime, steric interactions can be repulsive or attractive. The large *tert*-butyl groups on the outside of hexaphenylethane are ideally placed to hold the molecule together making use of what has been termed the ”corset effect” in literature.[70] Computational estimates based on DFT calculations for the corresponding BDE are approximately 35 kcal/mol, which is significantly stronger than in the parent hexaphenylethane demonstrating the strong attractive dispersion contribution of the 12 *tert*-butyl groups.[7]

As final example of manifestations of London dispersion, which is recently a very popular one, we look into the toes of geckos. Geckos can walk up walls of both hydrophilic and hydrophobic surfaces using what is called dry adhesion.[71] It has been demonstrated only relatively recently, that the dominating mechanism of dry adhesion is by attractive vdW forces between the geckoe toes and the surface rather than capillary forces or surface polarity.[71] It is maximized by the toes having millions of setae of high surface area, which are in contact with the surface and, hence, maximize the overall contact area.[71] Based on this finding, dry adhesive microstructures were successfully developed after the toes of geckos and they found immediate application in the so-called gecko tape.[18,71] This demonstrates that even animals evolved to make direct use of London dispersion.

1.1.4. Compensation of London Dispersion

There are (almost) omnipresent mechanisms that effectively counteract London dispersion. While it appears that they attenuate dispersive effects, they should be described as compensating rather than attenuating effects because dispersion still has the same strength, there is just an additional counteracting energy component. The one truly universal effect compensating dispersion is entropy. The second one is universal in solution, solvation. In general, the relative importance of these major compensating effects is only poorly studied as experimental data are sparse and computational methods for properly accounting for both of them are expensive, especially for large systems.

Both entropy of dissociation of a dispersive contact and the dissociation energy of the dispersive contact itself generally increase with molecular size. When the number of atoms increases, so does the number of dispersive contacts and the intermolecular dispersive interaction energy between the newly formed fragments. In addition, as the number of atoms increases, the available degrees of freedom in the dissociated fragments increases generally more strongly than in the intermolecular complex. This makes the intermolecular complex entropically increasingly unfavorable as the molecular size gets bigger. Hence, entropy of dissociation and dispersive interaction energy show parallel trends and attractive dispersion is partially compensated by unfavorable entropic contributions. In solution, interaction surface freed upon dissociation of an intermolecular complex

is used for new solute-solvent interactions.[12–15,72,73] The larger the complex, the more surface is freed upon dissociation which in turn gives rise to more solute-solvent interactions. Therefore, both intermolecular dispersion and solute-solvent interactions tend to increase with increasing molecular size. As a result, the solvation energy compensates the dispersive interactions. Less important compensating effects are zero-point energy, which also tends to increase with interaction strength, and the thermal enthalpy correction, which tends to increase with molecular size.

1.2. Energy Decomposition Analysis

In this section, we will explain the commonly employed EDA schemes to investigate the physical origin of particular interactions.

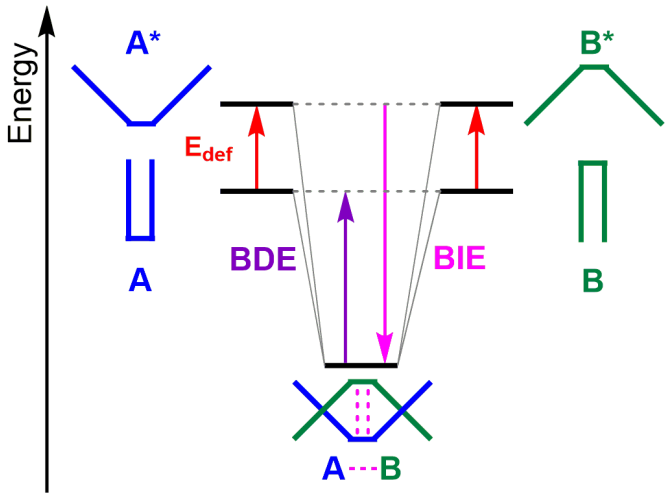
1.2.1. Deformation Energy

Let us consider the interaction between two fragments A and B. The BDE is the energy difference between the ground state of the complex between A and B (denoted AB) and the respective ground states of the isolated fragments. The deformation energy is the energy required to deform the isolated fragments from their respective ground state energies to their ground state geometries in the complex (denoted A* and B*, respectively). The BIE is the energy difference between the ground state of the complex and the isolated fragments at the interaction geometry. The BAE is the negative of the BDE. These definitions are illustrated in Scheme 1.3. The corresponding formulae are shown below:

$$BDE = E(AB) - E(A) - E(B) = -BAE \tag{1.20}$$

$$BIE = E(AB) - E(A^*) - E(B^*) \tag{1.21}$$

$$E_{def} = E(A^*) - E(A) + E(B^*) - E(B) \tag{1.22}$$



Scheme 1.3: Definitions of BDE, BIE and the deformation energy (E_{def}) for a complex between molecules A and B.

The deformation energy is useful to deconvolute the part of the energy invested to have ideal arrangement for interaction from the actual interaction itself. Minimizing deformation energies can be a useful design principle to maximize intermolecular interactions. In the context of EDA schemes it is also carried out because the energy decomposition typically is defined for the interaction of the isolated fragments at the interaction geometry with each other. If the deformation energy is to be decomposed further into separate energy components, molecular fragmentation approaches are typically necessary.

1.2.2. Symmetry Adapted Perturbation Theory

In supermolecular approaches, the BIE of intermolecular complexes is computed using Equation 1.21. SAPT is different from supermolecular approaches as it computes the BIE based on a perturbation of the Hamiltonian of the interacting molecules.[74] However, in addition to computing the BIE directly, SAPT also provides energy components of the total BIE, which are physically well-defined terms. The corresponding components are electrostatics (E_{elst}), exchange (E_{ex}), induction (E_{in}) and dispersion (E_{disp}, Equation 1.23).[74]

$$BIE = E_{elst} + E_{ex} + E_{in} + E_{disp} \tag{1.23}$$

Electrostatics consists of multipole-multipole interactions and interpenetration of charge densities. Exchange is repulsive and accounts for the antisymmtry requirement of the dimer wavefunction and the overlap of the monomer wavefunctions. Induction accounts for charge transfer and for polarization of the monomer wavefunctions due to the field of the respective other. Dispersion is the dynamic correlation between the electrons of the two interaction molecules.[74] The main drawback of the perturbative description of the BIE is that, inherently, only weak intermolecular interactions can be treated properly.[18] Strong interactions like chemical bonds are outside the scope of SAPT. SAPT0 is the simplest method of SAPT. It treats the monomers at the HF level of theory and carries out a second-order perturbation of the HF wavefunction of the dimer.[74] Hence, SAPT0 is comparable in computational cost to MP2.[74] Higher order SAPT methods treat the respective monomers also at higher level of theory and include higher than second-order perturbations.[74]

1.2.3. Local Energy Decomposition

LED is in its essence a supermolecular approach for EDA and is based on DLPNO-CCSD(T).[75] The main advantage of LED over SAPT is that the former, in principle, can treat both strong interactions like chemical bonds and weak interactions like noncovalent complexes at the same level of theory.[75] In a first step, the BDE of an intermolecular complex is decomposed, as outlined before, into BIE and E_{def}. The next step is decomposition of the BIE into an HF contribution (BIE^{HF}) and a correlation contribution (BIE^{C}):

$$BIE = BIE^{HF} + BIE^{C} \tag{1.24}$$

The HF term includes electrostatic interactions, polarization contributions and donor-acceptor interactions. The correlation term includes a dynamic correction to all these terms also accounted for in the HF contribution, and also accounts for dispersion. Due to the use of localized orbitals, both the HF term and the correlation term can be decomposed into intrafragment, which are called electronic preparation energies (E_{prep}), and interfragment contributions, which are called genuine interaction energies (E_{int}) if the localization was appropriate:

$$BIE = E_{int}^{HF} + E_{prep}^{HF} + E_{int}^{C} + E_{prep}^{C} \tag{1.25}$$

The preparation energy terms are obtained from LED calculations and DLPNO-CCSD(T) single point calculations of the isolated fragments in the interaction geometries using the basis set of the full molecule. The genuine HF interaction is decomposed further into electrostatics (E_{elst}) and exchange (E_{ex}). Electrostatics corresponds to the interactions of the distorted electron density of the fragments. The exchange in this decomposition scheme is always attractive and lowers the repulsion between same-spin electrons. The corresponding decomposition is shown in the following formula:

$$E_{int}^{HF} = E_{elst} + E_{ex} \tag{1.26}$$

The genuine correlation interaction is composed of the strong pairs (E_{SP}^{C}), the weak pairs (E_{WP}^{C}) and the perturbative triples contributions ($E_{(T)}^{C}$):

$$E_{int}^{C} = E_{SP}^{C} + E_{WP}^{C} + E_{(T)}^{C} \tag{1.27}$$

The strong pairs contribution is decomposed further based on the respective electronic excitations into dispersion (E_{disp}) and charge-transfer (E_{ct}). The weak pairs contribution is not decomposed further but typically makes a significant total energy contribution. Since these contributions arise from electron pairs far from each other, the corresponding charge-transfer contribution can be assumed to be negligible and it is reasonable to assume that it corresponds solely to dispersion. Hence, it is important that at least TightPNO parameters are chosen so that the weak pairs do not contain significant charge-transfer contributions. The perturbative triples contributions are not decomposed further as they are typically small. It should be noted that a significant portion of the triples correction is already part of the electronic preparation energy from correlation (E_{prep}^{C}). However, it has been argued that $E_{(T)}^{C}$ can be split into dispersion and charge-transfer contributions in the same ratio the strong pairs do.[75] Summarizing this decomposition of the genuine correlation interaction, the following full decomposition of the BIE is obtained from LED:

$$BIE = E^{HF}_{prep} + E^{C}_{prep} + E_{elst} + E_{ex} + E_{disp} + E_{ct} \tag{1.28}$$

Due to the inherently different theoretical framework of LED compared to SAPT, not all the energy components can be directly compared. The component that is most comparable between the two EDA schemes is dispersion. While exchange is always repulsive in SAPT, it is always attractive in LED due to the differing definition. In addition, electrostatics are also in principle comparable between SAPT and LED.

1.3. Visualization of Interactions

The visualization of interactions is a common tool to connect the corresponding physical models with spatial understanding. The human brain typically processes visual stimuli quickly and is very efficient in recognizing patterns. The main reason to visualize interactions is to gain further understanding of their origins. In this section we will investigate a few commonly employed methods to visualize particular energy components and discuss the corresponding information that can be obtained from them.

1.3.1. Electrostatic Potential

The ESP of a molecule allows to compute the attraction or repulsion of a test charge at an arbitrary position r in space by the corresponding molecule. It does not account for polarization of the molecule by the test charge and can be computed from the molecular wavefunction Ψ as follows:[76]

$$V_{ESP}(r) = \sum_{i}^{nuclei} \frac{Z_i}{|r - r_i|} - \int \Psi(r') \frac{1}{|r - r'|} \Psi(r') dr' \tag{1.29}$$

In Equation 1.29, the index i runs over all the nuclei in the molecule and Z_i and r_i are the charge and position of the nuclei. Notably, V_{ESP} can be, in principle, experimentally observed. The ESP map of molecules is typically visualized by projection onto the vdW surface of the molecule. This provides information

about the local polarity within a molecule and allows to predict where electrostatic interactions would be most attractive or most repulsive and can even guide predictions for chemical reactivity.[76]

1.3.2. Noncovalent Interactions

NCI analysis is based on the electron density and its derivatives. The reduced density gradient is defined as follows:[77,78]

$$s = \frac{1}{2(3\pi^2)^{1/3}} \frac{|\nabla\rho|}{\rho^{4/3}} \tag{1.30}$$

NCIs are characterized by regions in space having both a small reduced density gradient and low electron densities. These regions are plotted using isosurfaces of s with low values of s.[77,78] The values of the electron density ρ of the visualized surfaces are indicative of the strength of the interactions, high densities correspond to strong interactions, low densities correspond to weak interactions. Furthermore, the sign of the second eigenvalue of the Hessian of the electron density (λ_2) distinguishes attractive (negative sign) from repulsive (positive sign) interactions.[77,78] Therefore, the product of the sign of λ_2 and ρ is projected onto the reduced density gradient isosurfaces. Negative regions indicate attractive interactions, positive regions indicate repulsive regions.[77,78] The strength allows to distinguish between weak noncovalent interactions like dispersion, dipole interactions or weak exchange, and strong noncovalent interactions like hydrogen bonds or strong exchange.

1.3.3. Dispersion Interaction Density

DID is based on localized orbital spaces and provides the spatial density of dispersion interactions in a specified interfragment interaction energy.[79] The localization scheme merely needs to localize orbitals (both occupied and virtual) within the fragments so that no orbital spans across both fragments. After localization, we have occupied localized molecular orbitals, which are referred to with indices i and j and virtual localized orbitals, which are referred to with indices r and s. To obtain the interfragment dispersion, we only look at excitations of orbitals i on

fragment A to virtual orbitals r on fragment A simultaneous with excitations of orbitals j on fragment B to virtual orbitals s on fragment B. Hence, we only consider simultaneous excitations of two electrons in the separate interacting fragments.[79] Notably, this disregards exchange dispersion, which would correspond to excitations of orbitals i on fragment A to virtual orbitals r on fragment B simultaneous with excitations of orbitals j on fragment B to virtual orbitals s on fragment A.[75]

Using this partitioning, only the corresponding dispersion pair energies are used. Using these pair energies, the intermolecular dispersive contribution of each orbital can be obtained by multiplying with the respective orbital densities.[79] The dispersion interaction density as a function of the position in space can then be obtained by multiplication with the respective orbital densities as function of space.[79] Typically, the DID is visualized using a voxel plot and each position in space is colored according to its local intermolecular interaction density. Since the DID is defined as interaction density of fragment A for interaction with fragment B based on localized orbital excitations, it directly visualizes the origin of the dispersive interaction.

1.3.4. Hirshfeld and Becke Surfaces

The Hirshfeld surface is based on the question how much space a molecule occupies in a crystal. The idea was to partition the electron density into separate molecules. In addition, an alternative to the definition based on QTAIM was desired. The Hirshfeld surface is based on the stockholder partitioning scheme which assigns every point in space a weight w_a describing how much it belongs to a certain atom a in a molecule (mol):[80]

$$w_a(r) = \frac{\rho_a^{at}(r)}{\sum_{i \in \text{mol}} \rho_i^{at}(r)} \tag{1.31}$$

In Equation 1.31, $\rho_a^{at}(r)$ are spherically-averaged densities of atom a and are typically reference densities for each atom type. Based on this definition, the electron density of each atom can be determined from the molecular electron density ($\rho^{\text{mol}}(r)$) as follows:[80]

$$\rho_a(r) = w_a(r)\, \rho^{\text{mol}}(r) \tag{1.32}$$

The main difference between stockholder partitioning and QTAIM partitioning is that in the latter every point in space only belongs to atom or the boundary between atoms, while in the former every point in space belongs to several atoms at the same time as described by the respective atomic weights. In an analogous way, the weight function for a molecule A in a crystal can be defined as follows:[80]

$$w_A(r) = \frac{\sum_{i \in A} \rho_i^{at}(r)}{\sum_{i \in \text{crystal}} \rho_i^{at}(r)} \tag{1.33}$$

The Hirshfeld surface of a molecule A is now defined as the amount of points r where $w_A(r) = 0.5$. Initially, this was defined for molecules in a crystal.[80] However, in the same way this definition can be extended to intermolecular complexes where the weight is defined with respect to a specific fragment of the total complex. Since the Hirshfeld weights require atomic electron densities and the molecular electron density, they are tedious to compute. Hence, the Becke surface was proposed as alternative.[81] The Becke weight is determined based on the atomic coordinates and a set of atomic reference radii alone, making it operationally and computationally much simpler.[81] Nevertheless, it was shown to provide similar results. Both Hirshfeld and Becke surfaces can be used for projection of various properties providing information about intermolecular interactions and other molecular properties of interest.[80]

1.4. Solvent Models

The proper account of the solvent in computer simulations of physical and chemical processes is one of the main challenges in computational chemistry. Due to the large number of solvent molecules, the inherently complicated dynamic interplay between solute and solvent, and the highly disordered arrangement of solute and solvent molecules, it is intrinsically very complicated to tackle. While it is clear that the best model would be a full quantum mechanical treatment of a large number of solvent molecules, which surround the solute with appropriate treatment of the dynamics, it is also clear that the large number of solvent molecules required for such an approach is not computationally feasible with the current computers and the most accurate quantum mechanical methods. Hence,

in order to still be able to compute the corresponding systems in solution, either the physical model used to account for all the explicit solvent molecules must be cheap, or the number of solvent molecules explicitly included is reduced to a number that is still tractable. In this section we will briefly investigate both these approaches and discuss some of the more widely employed techniques to treat solvation in computations.

The quantity of interest, either implicitly or explicitly, is the Gibbs free energy of solvation ΔG°_{sol}. It describes the change in Gibbs free energy of a molecule in the gas-phase when going into a specific solution.[76] Care must be taken regarding the change in standard state definitions of a pure ideal gas and a solute in solution. From experimental data it is known that the range of Gibbs free energies of solvation spans from -10^2 kcal/mol to 10^1 kcal/mol in terms of orders of magnitude.[76]

1.4.1. Implicit Solvent Models

Implicit solvent models are continuum models. Hence, the charge distribution of the solvent molecules is replaced by a continuous electric field surrounding the solute, that represents the statistical average of the field effected by the solvent molecules. The simplest model of solvation is the Born model, which describes the solvation of monoatomic ions in the continuous field of the solvent. In a more general model, the work G necessary to create the solvent charge distribution at an arbitrary point in space is derived from the interaction of the charge density ρ of the solute with the electrostatic potential. The electric field corresponds to the gradient of the electrostatic potential ϕ. G can be computed according to the following equation:[76]

$$G = -\frac{1}{2}\int \rho(r)\ \phi(r) dr \tag{1.34}$$

The polarization free energy (G_P) is then the difference between the work to create the solvent charge in the gas phase (i.e. dielectric constant of 1) and in solution:

$$G_P = G_{gas} - G_{sol} \tag{1.35}$$

In implicit solvent models, the charge distribution of the solute is considered to be inside a cavity, which displaces the homogeneous dielectric medium of the solvent. Hence, there are regions inside the cavity and regions outside the cavity. If the medium has zero ionic strength, the Poisson equation is used to express the electrostatic potential in terms of the charge density and the dielectric constant. For the case of continuum solvent models, it can be written as follows:[76]

$$\nabla\epsilon(r)\ \nabla\phi(r) = -4\pi\rho(r) \tag{1.36}$$

In order to compute the electrostatic potential in solvents, Equation 1.35 needs to be solved based on the charge density of the solute, which is known, and the cavity determining the dielectric constant around the solute, which needs to be defined.[76] The definition of the cavity is vital in obtaining an appropriate model of the solute in a solvent. This was one of the main historic problems with implicit solvent models as one could get any numerical answer by the choice of the cavity size. After the electrostatic potential is obtained, the so-called reaction field V in the space occupied by solute is determined using Equation 1.34 as difference between gas phase and solution. Based on the reaction field, the non-linear Schrödinger equation that needs to be solved for the solute can be formulated as:[76]

$$\left(H - \frac{1}{2}V\right)\ \Psi = E\,\Psi \tag{1.37}$$

This non-linear Schrödinger equation can be solved in the same way as ordinary DFT and HF calculations in a self-consistent way. After the Schrödinger equation is solved, the electrostatic component of the solvation free energy (G_{ENP}), which accounts for electronic, nuclear and polarization components of the free energy, is determined as difference between the ground-state energy in the gas phase and in the continuum solvent:[76]

$$\Delta G_{ENP} = \langle\Psi_{sol}|H + G_P|\Psi_{sol}\rangle - \langle\Psi_{gas}|H|\Psi_{gas}\rangle \tag{1.38}$$

This is solely an enthalpic model of solvation, entropy is not included. Hence, it should be noted that Gibbs free energies and electronic energies are mixed in this definition. Typically, this is ignored in continuum solvent models.[76] Contin-

uum solvent models making use of this self-consistent procedure including the reaction field in the Hamiltonian are generally referred to as PCM models.[76] So far, we have described the solvent continuum as dielectric medium with a finite dielectric constant. An alternative approach is to describe the solvent continuum as conductor with infinite dielectric constant. This simplifies the calculation of the polarization free energy significantly. This approximation is made use of in the COSMO solvent model.[76,82]

Non-electrostatic contributions to the Gibbs free energy of solvation are typically included with exposed surface area (A) proportional contributions together with atom-type specific parameters σ describing the strength of the interaction between solute and solvent. The energy components accounted for with this approach include cavitation, dispersion and solvent structure. This energy component is denoted as G_{CDS} and can be computed as follows:[76]

$$\Delta G_{CDS} = \sum_{k \in solute} A_k \, \sigma_k \tag{1.39}$$

This approach is completely empirical and is based on observations for simple systems like alkanes, for which it works reasonably well.[76] However, it is important to note that the exposed surface area again depends strongly on the cavity definition. The standard approach in the development of continuum solvent models is to define the cavity generation procedure in a specific way and parametrize the atom-type specific parameters in order to minimize the differences between computed Gibbs free energies of solvation and experimental ones. It was found that continuum solvent models defined that way can reproduce experimental Gibbs free energies of solvation for neutral solutes to within errors of about 0.5 kcal/mol and for charged solutes to within errors of about 5 kcal/mol, which is not far from the respective experimental uncertainties.[76] Notably, these errors are valid for simple solutes with a relatively low number of atoms per molecule as experimental numbers for more complicated molecules are very sparse.

1.4.2. Explicit Solvent Models

Explicit solvent models rely on an atomistic description of the solvent molecules surrounding a solute. As the number of solvent molecules is typically large re-

lative to the number of solute molecules, the associated number of degrees of freedom of the complete system explodes compared to an isolated molecule in the gas phase.[76] In particular, for any single conformation of a solute there is an enormous number of possible solvent configurations. Hence, proper description of explicit solvent models relies on statistical thermodynamics, which relies on the partition function.

To compute explicit solvent models, a large number of configurations of solute surrounded by solvent together with their respective energies are required to estimate partition functions which then give access to all thermodynamic properties of interest for a given system. Typically, rather than generating a large number of configurations for a given system and compute their energies, molecular dynamics approaches are used in which the molecules are propagated according to the equations of motion. In any case, not only the number of atoms that need to be included is large, the number of configurations that need to be explicitly considered is even larger. That is the reason explicit solvent models are orders and orders of magnitude more expensive than implicit solvent models if treated at the same level of theory. Only the cheapest QM approaches are currently to be considered for explicit solvent models and the main method of choice for explicit solvent models are MM approaches. Explicit solvent models offer the possibility of studying the dynamics of physical and chemical processes in solution, something that is inherently impossible in implicit solvent models because the solvent dynamics cannot be accounted for. A detailed description of explicit solvent models is outside the scope of this thesis.

2. Compensation of Dispersion in Dichloromethane

Marek Bot and Ewa Pietrasiak are acknowledged for the synthesis of 1-aza[5]helicene and 1-aza[6]helicene and for T-CID measurements. Dr. Ilya Kobylianskii is acknowledged for the synthesis of azatrypticene and for T-CID measurements. This chapter is adapted with permission from "Attenuation of London Dispersion in Dichloromethane Solutions," Robert Pollice, Marek Bot, Ilia J. Kobylianskii, Ilya Shenderovich, and Peter Chen *J. Am. Chem. Soc.* **2017**, *139*, 13126 – 13140. Copyright © 2017 American Chemical Society.

2.1. Introduction

The main purpose of this chapter is to study the compensation of London dispersion in proton-bound dimers in DCM solutions. We do that by studying the bond dissociation of a systematic set of proton-bound dimers of nitrogen bases. These model compounds are both well-behaved and can be well-characterized. The first objective is to estimate the impact of London dispersion in these compounds in gas-phase BDEs based on experimental and in-depth computational studies. Then, we determine the Gibbs free energy of dissociation in DCM solution for the same compounds. Consequently, we compare the results from computation in the gas phase and experiment in solution to obtain estimates of the compensation of dispersion as a function of temperature. Finally, we also apply implicit solvent models to estimate the Gibbs free energies of proton-bound dimer dissociation in DCM and compare the computational numbers with experiment. While our model compounds are only a small subset of test systems for dispersion, our approach allows us to extract only the contributions of alkyl-alkyl, alkyl-aryl and aryl-aryl groups, which are ubiquitous in organic chemistry, making

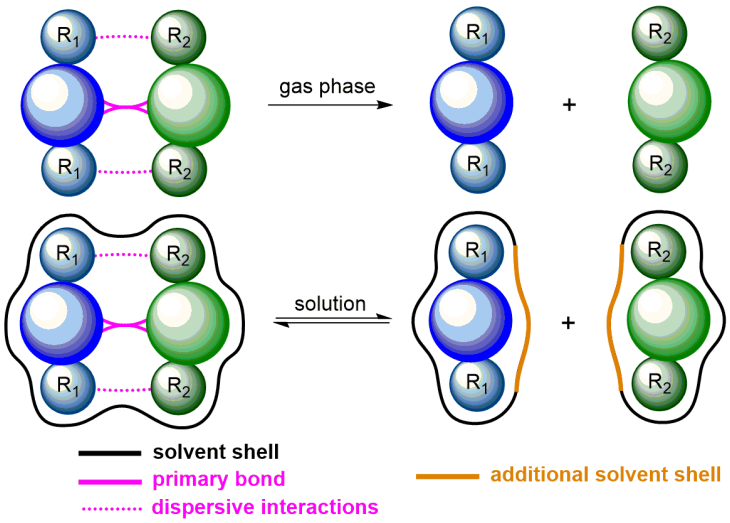
the derived conclusions valid for a wide range of organic molecules. Additionally, our consistent set of experimental Gibbs free energies of reaction constitutes a valuable benchmark set for London dispersion in solution that can be employed to test new solvation methods accounting for dispersion.

2.2. Results

2.2.1. Model System

First, it is important to design a suitable set of model compounds for our study. The set should contain structures with a common interaction motif. The idea was to have bimolecular complexes made up from the respective monomers and held together by a relatively strong, but not too strong, central bonding interaction. The strength of this interaction should be high enough so that the 3-dimensional arrangement of substituents around that central interaction is fixed, but weak enough so that this interaction can be readily broken. The respective interacting fragments are decorated with additional substituents that are sufficiently close to the central bonding interaction so that the substituents of the interacting fragments also interact with each other. These substituents should interact mainly by London dispersion. In addition, the substituents should only allow a limited number of low-energy conformers to minimize the entropic penalty of association. The corresponding design is illustrated in Scheme 2.1.

This design is based on the idea that cleavage of the central primary bond results in sacrifice of all the secondary dispersive interactions between the substituents on the fragments. When the primary bond is kept constant within a set of model compounds but the substituents are systematically varied, the total BDE is a function of the substituent. Since the substituents are designed to interact mainly by dispersion, the total BDE is a function of dispersive interaction strength. In solution, the interaction surface used for interfragment interactions in the complex will be freed upon dissociation and covered by additional solvent molecules. Hence, the loss in total interaction energy is reduced in solution compared to the gas phase. When the total interaction energy is determined as a function of substituents in the fragments, both in the gas phase and in solution,



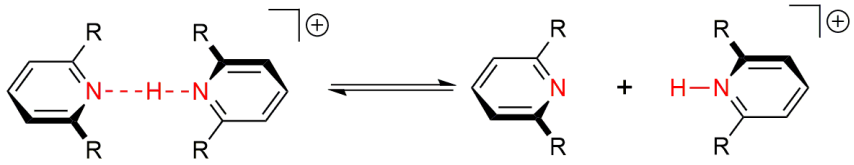
Scheme 2.1: General outline of the design philosophy for a good test system to study the compensation of dispersion in solution.

the degree of compensation of interfragment dispersion can be estimated.

Additional criteria for suitable model systems are well-behaved electronic structures so that they can be modelled readily by standard DFT and WFT methods and ready synthetic access of a range of differently substituted compounds. Furthermore, the bond dissociation process should be amenable to experimental studies both in the gas phase and in solution. In the gas phase, the primary bond should be breakable by CID utilizing ESI-MS/MS. In solution, the dissociated fragments should be in equilibrium with the complex so that the equilibrium constant of the corresponding reaction can be determined using spectroscopic or calorimetric techniques.

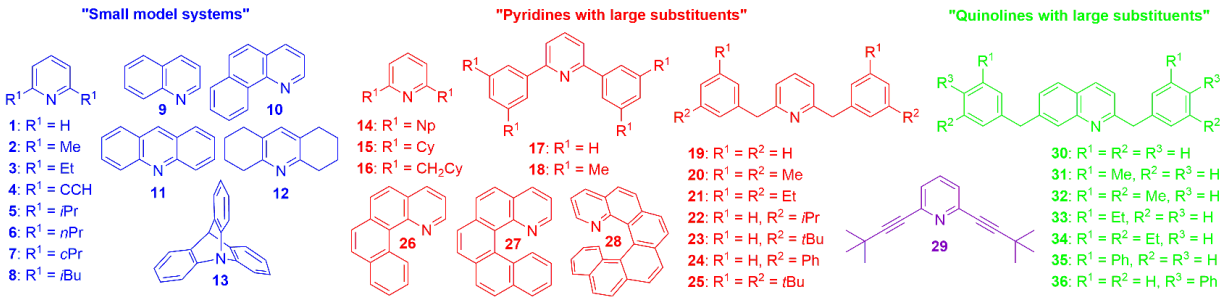
To construct a systematic set of model compounds we selected symmetric hydrogen-bonded dimers of substituted pyridines, quinolines, acridines and tertiary amines. The process under study is the dissociation of the central ionic hydrogen bond (Scheme 2.2).

Several reasons made us select this specific type of model compounds. First, hydrogen bonds are well-known in literature both from a computational and an experimental point of view. They can be described sufficiently well using standard electronic structure methods, as demonstrated by numerous computational benchmark studies,[74,83,84] and multireference character is typically negligible. We tested for multireference character using the T1-test[85] from our DLPNO-



Scheme 2.2: Proton-bound dimer dissociation of 2,6-disubstituted pyridines.

CCSD(T)[30,37,86] results. The highest value observed was 0.014 and the majority of structures had a T1 measure close to 0.010, which is far from a value of 0.020 which is considered to indicate incipient multireference character. Additionally, these test systems can be studied using both ^{1}H-NMR and calorimetry in solution. Notably, the structure of some of the hydrogen-bonded dimers investigated has already been studied in the literature using NMR spectroscopy at low temperatures.[87,88] Furthermore, since we study the interaction between a nitrogen base and its conjugated acid, which is protonated and charged, the proton-bound dimer is also charged, which makes it amenable to MS analysis, allowing us to perform gas-phase T-CID measurements. Moreover, standard cross-coupling chemistry provides ready access to most of the model compounds from commercially available building blocks. The compounds studied in this chapter are depicted in Scheme 2.3.



Scheme 2.3: Structures and numbering of nitrogen bases studied in this chapter.

2.2.2. Theoretical Investigation of Gas-Phase Dissociation

The general computational strategy used is based on the decomposition of the interaction energy as outlined in chapter 1. We started by carrying out geometry optimizations of all the model compounds under study using several functionals with differing treatment of dispersion. The main idea was to use affordable levels of theory for both geometry optimizations and frequency calculations to be able to treat all the model compounds at the same theoretical footing. Our methods of choice were BP86,[89,90] M06-L[42,43] and B97-D3.[52,91] We selected pc-2-sp(d)[92–94] as a truncated triple-zeta basis set, which was shown to be sufficiently large for accurate computations of non-covalent interactions.[95] To test the methods employed, the computed geometries of the proton-bound dimer of **19** were compared to the corresponding X-ray structure (Figure 2.1 and Table 2.1).

Property	Experiment	BP86/pc-2-sp(d)	M06-L/pc-2-sp(d)	B97-D3/pc-2-sp(d)
N–N distance [Å]	2.80	2.83	2.79	2.78
RMSD [Å]	0.00	0.38	0.17	0.14

Table 2.1.: Experimental X-ray structural parameters compared to computed structures of the proton-bound dimer of **19**. The RMSD is computed from atomic positions excluding H atoms.

Figure 2.1 illustrates the C_2-symmetry of the dimer, which is close to the S_4 point group. The four benzyl groups of the dimer surround the central N-H–N motif and are ideally placed for interfragment interactions. Notably, there are relatively close contacts (smaller than the sum of the respective vdW radii[96,97]) between each of the methylene groups and one phenyl ring. Table 2.1 illustrates that inclusion of dispersion is vital to obtain good estimates of the structures in the proton-bound dimers of the model compounds, as both M06-L and B97-D3 result in geometries very close to the X-ray structure while BP86 results in the largest discrepancies. Similar observations regarding non-covalent complexes are well documented in the literature.[98] The structural differences between the three DFT approaches are demonstrated by Figure 2.2. It reveals that the pyridine core

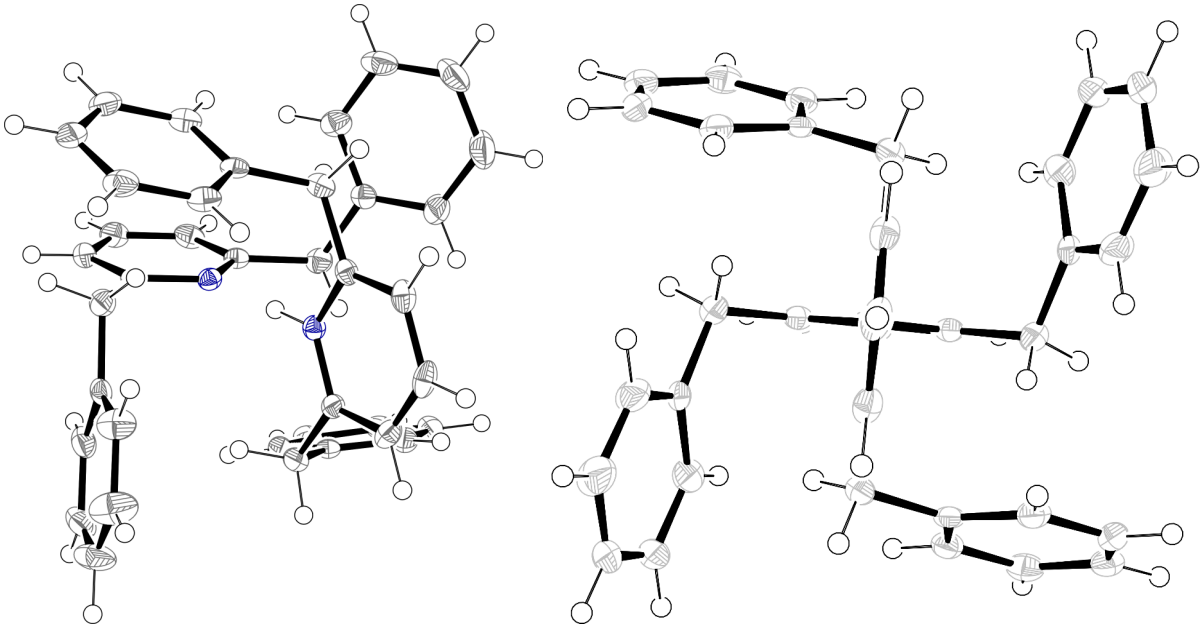


Figure 2.1.: ORTEP wtih thermal ellipsoids of 50 % of hydrogen-bonded dimer of **19** and the $BArF_{24}^-$ salt of the conjugate acid of **19**. The counterion is hidden to have a clear view of the dimer.

is reasonably close throughout all three methods but the spatial arrangement of the benzyl substituents is similar with M06-L and B97-D3 but differs considerably with BP86, which results in more extended benzyl groups.

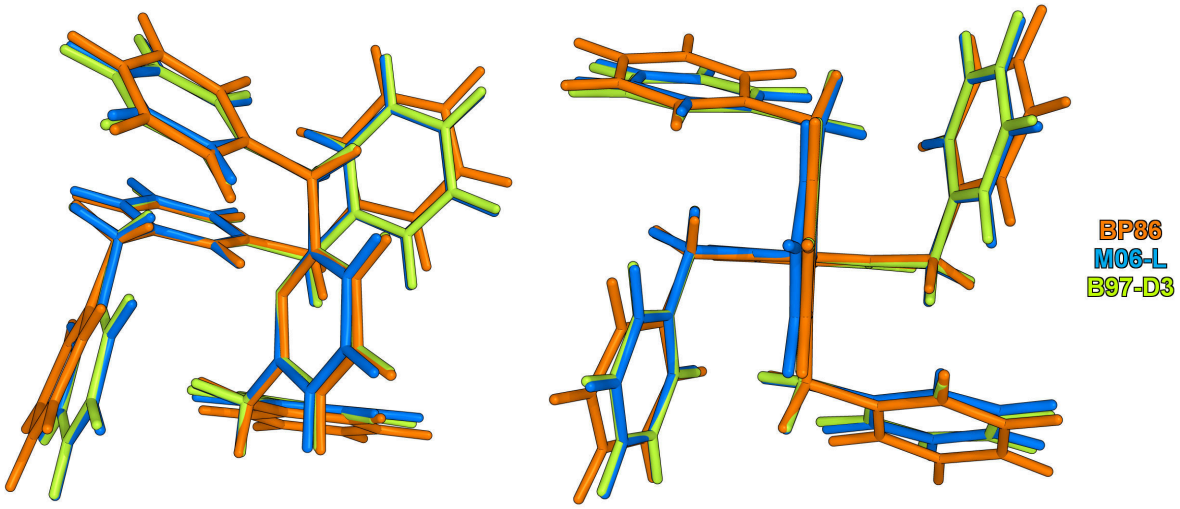


Figure 2.2.: Comparison of geometries from BP86, M06-L and B97-D3.

Based on the respective DFT geometries, further refinement based on single point calculations with WFT approaches was also carried out. We were interested in comparing the BDEs from the three computational approaches to single points at the DLPNO-CCSD(T)/CBS level of theory using the respective geometries. The corresponding comparison is illustrated in the left scatter plot in Figure 2.3. All the BDEs compared do not include ZPE corrections because they are unavailable for DLPNO-CCSD(T). Notably, BP86 and M06-L computations were not carried out for the entire set of molecules but rather for a subset spanning the whole BDE energy range. The results reveal that BP86 energies at BP86 geometries deviate considerably from DLPNO-CCSD(T) energies at the same geometries. For both M06-L and B97-D3 the trends are qualitatively reproduced, but only B97-D3 shows close agreement with the corresponding DLPNO-CCSD(T) energies. Additionally, DLPNO-CCSD(T) results from either B97-D3 or M06-L geometries are for most of our model compounds within 1.0 kcal/mol indicating again that both methods result in comparable geometries. Based on these results, B97-D3/pc-2-sp(d) was chosen as the method of choice for geometry optimizations as its computational cost is considerably lower than for M06-L. Next, we compared the energies from M06-L/CBS, B97-D3/CBS and DLPNO-CCSD(T)/CBS against experimental results from either GEQ[99] or T-CID measurements. The corresponding comparison is depicted in the right scatter plot of Figure 2.3. Overall, we had a set of nine experimental gas-phase data points for seven of our model compounds. All the experimental data points agree well with M06-L/CBS//B97-D3/pc-2-sp(d) BDEs. The six data points for proton-bound dimers with small substituents agree also well with B97-D3/CBS//B97-D3/pc-2-sp(d) and DLPNO-CCSD(T)/CBS//B97-D3/pc-2-sp(d).

We wanted to test the disagreement between experiment and DLPNO-CCSD(T) energies by an independent control experiment because, *a priori*, DLPNO-CCSD(T) would be expected to yield the best results of all the computational approaches employed.[31,33,34] Hence, we attached a group with an alternative dissociation pathway of well-known BDE to one of the proton-bound dimers that did not agree with DLPNO-CCSD(T). The control experiment would be a competition between two CID channels. By introducing a bond with a BDE exactly in between the predicted numbers from M06-L and DLPNO-CCSD(T),

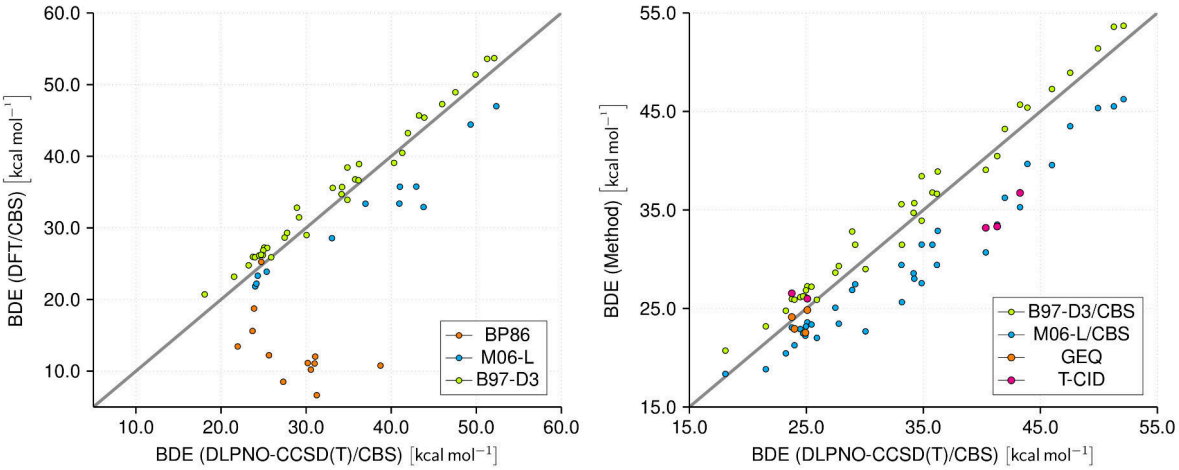
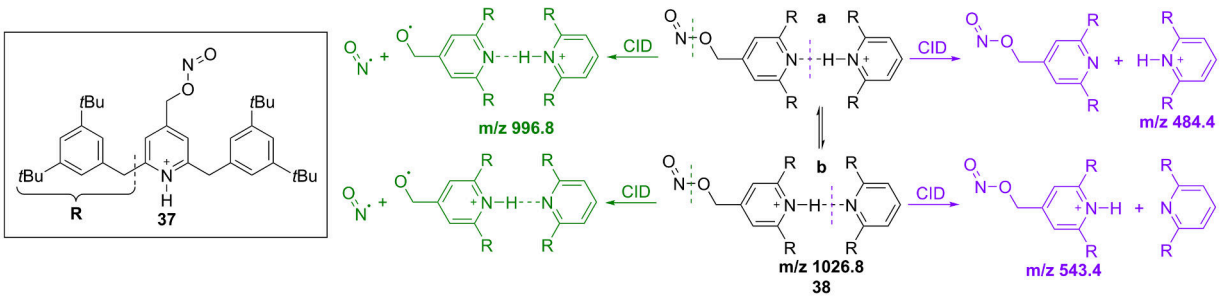


Figure 2.3.: Left: Comparison of non ZPE-corrected DFT/CBS//DFT/pc-2-sp(d) BDEs against the respective DLPNO-CCSD(T)/CBS//DFT/pc-2-sp(d) results. Right: Comparison of B97-D3/CBS, M06-L/CBS and DLPNO-CCSD(T)/CBS at B97-D3/pc-2-sp(d) geometries against experimental results from T-CID and GEQ measurements.

performing the CID experiment provides information about the relative bond strengths of the two competing dissociations. We decided to introduce a remote benzylnitrite group at the para-position of **25** and synthesized **37**, which is depicted in Scheme 2.4.



Scheme 2.4: CID competition in **38** between H-bond cleavage and O–NO bond cleavage.

The reasons for introducing a nitrite are that their experimental BDEs are known from independent methods (including gas-phase dissociation under thermal, but unimolecular conditions), that the NO fragment produced in the

dissociation is very small and, hence, not affected by London dispersion with the other fragment and that they are experimentally readily accessible from the corresponding alcohols. Experimental values for BDEs of various alkyl nitrites are available from gas-phase pyrolysis studies and are 34 kcal/mol–40 kcal/mol.[100–103] This puts the corresponding BDE exactly in between the computed estimates for the hydrogen bond in the proton-bound dimer of **25** which range from 34 kcal/mol using M06-L to 42 kcal/mol using DLPNO-CCSD(T). In addition, the fact that the O-NO BDEs from all the computational methods are within 2 kcal/mol while the NH-N BDEs are much farther apart suggests that dispersion is treated differently by the methods used. Hence, we synthesized **37** and prepared the proton-bound heterodimer between protonated **37** and **25** *in situ* and observed it by ESI-MS. CID of the heterodimer resulted in cleavage of the hydrogen bond but not the O-NO bond, while CID of **37** resulted in selective dissociation of the O-NO bond. All the results of this control experiment are summarized in Figure 2.4 and Table 2.2 and they show that the M06-L/CBS results agree best with all the experimental observations. This is independent support for the previous observation that M06-L/CBS energies best reproduce experimental BDEs from GEQ and T-CID measurements (cf. Figure 2.3).

Reaction	B97-D3	M06-L	DLPNO-CCSD(T)
		[kcal/mol]	
H-bond isomerization (**38b** to **38a**)	0.3	0.4	0.6
H-bond cleavage **38a**	44.3	34.2	44.2
H-bond cleavage **38b**	45.5	35.1	45.3
O–NO bond cleavage **38a**	40.7	40.1	38.5
O–NO bond cleavage **38b**	41.0	40.5	39.1

Table 2.2.: Computational estimates of BDEs for the competition between H-bond cleavage and O–NO bond cleavage.

Having established the computational methodology, we wanted to analyze the computed geometries of the proton-bound dimers of all the model compounds in more detail, focussing on geometry of the N-H–N motif. We looked into the N–N

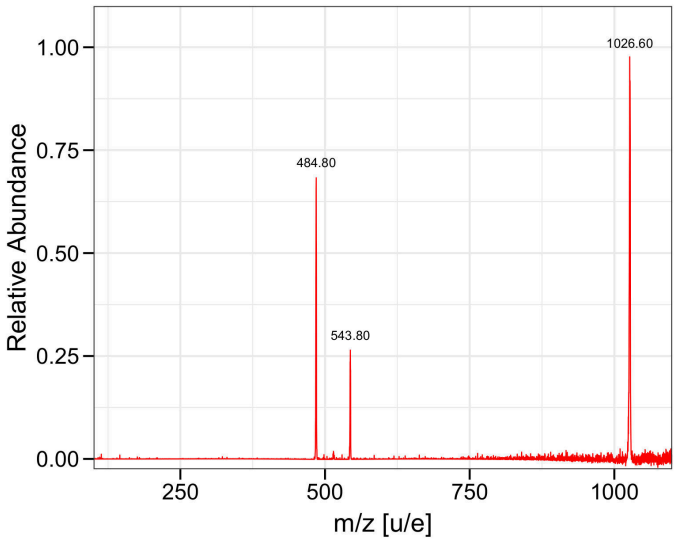


Figure 2.4.: CID of **38** shows only H-bond cleavage.

distances and the N-H–N angles and the corresponding results are illustrated in Figure 2.5. The proton-bound dimer of **1**, which is unsubstituted pyridine, has an N-H–N angle of 180° and the shortest of all N–N distances. Already introducing methyl groups at the 2 and 6 positions increases the N–N distance significantly and leads to an N-H–N angle slightly lower than 180° as can be seen in the data of **2**. While there are several substituted pyridines with an N-H–N angle of 180°, there is no substituted pyridine with an equally short N–N distance as in **1**.

As an important part of the computational investigation of the proton-bound dimers, we performed SAPT calculations using B97-D3/pc-2-sp(d) geometries to decompose the overall BIEs of the studied complexes into electrostatics, exchange, induction and dispersion. The corresponding results are depicted in Figure 2.6. It illustrates that the most sensitive energy components with respect to change in substitution are London dispersion and Pauli repulsion. Additionally, dispersion and exchange have opposite slopes indicating that they are counteracting. The sum of dispersion and exchange is negative for all the studied model compounds which shows that at least for this system repulsive exchange is outweighed by attractive dispersion. Hence, the model compounds chosen are suitable test systems to probe the sensitivity of the bond dissociations with respect to London dispersion between the substituents.

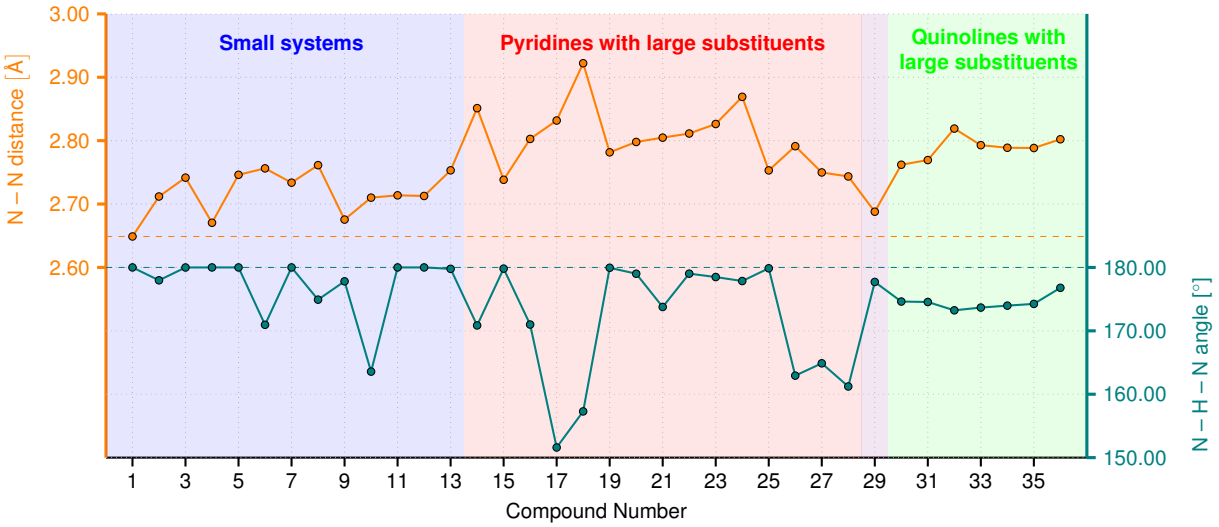


Figure 2.5.: N-H–N angles and N–N distances of all nitrogen bases using B97-D3/pc-2-sp(d) geometries. The dashed lines depict the values of unsubstituted pyridine as reference.

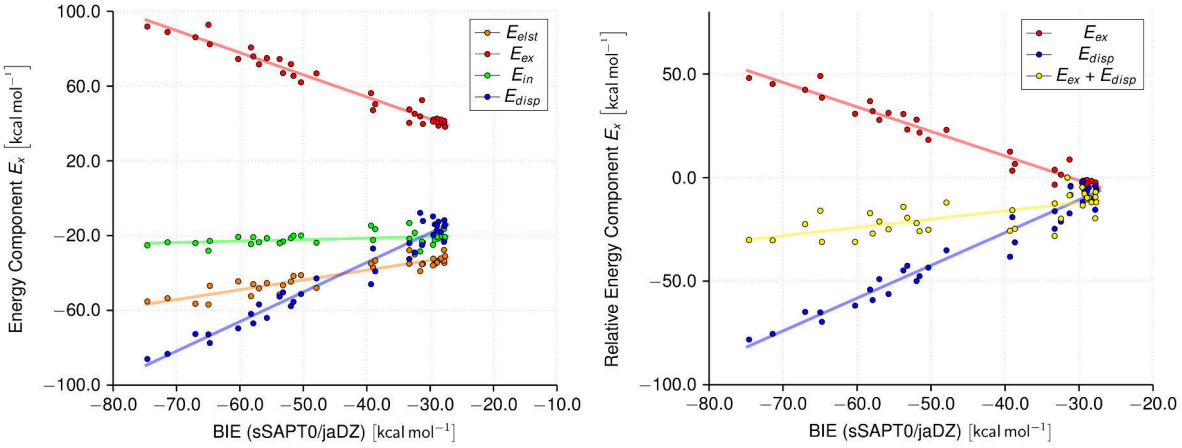


Figure 2.6.: Left: Decomposition of the BIEs of proton-bound dimers using sSAPT0/jaDZ. Right: Comparison of exchange and dispersion components with their sum.

We were also interested to test whether our model system design with respect to the positioning of the substituents on the pyridines and quinolines does indeed lead to close dispersive contacts. To do that, we computed NCI plots[77,78,104] of the proton-bound dimers of **19**, **25** and **30** at the B97-D3/pc-2-sp(d) geometries to visualize the inter-fragment interactions. The corresponding results are depicted in Figure 2.7. They show that inter-fragment dispersion in these structures is significant and that substituents at the 3,4 and 5 positions of the benzyl rings are ideally placed for attractive interactions. The benzyl groups surround the central N-H–N bond in order to maximize attractive contacts. In addition, it can be seen that the substituted quinolines are not packed as densely as the corresponding pyridines, which suggests that the corresponding dimers are not as rigid and that there may be internal voids or cavities.

Finally, we were interested to estimate the Gibbs free energies of bond dissociation of our model compounds in DCM solution. To do that, thermodynamic corrections assuming an ideal gas and the Gibbs free energies of solvation need to be computed. Thermodynamic corrections were performed using B97-D3/pc-2-sp(d), Gibbs free energies of solvation were estimated using COSMO-RS[82,105] and SMD,[6] which are both dispersion-corrected implicit solvent models and still considered state of the art and, hence, widely used.[106] The complete computational approach to arrive at estimates of the Gibbs free energies of bond dissociation for all our model compounds is delineated in Scheme 2.5.

The approach outlined in Scheme 2.5 was used to compute $\Delta_R G^\circ_{diss}$ at 298 K in DCM for all test systems. Notably, the conversion of reference states from of an ideal gas at a pressure of 1 atm to an ideal gas at a concentration of 1 M accounts for a difference of −1.5 kcal/mol at 298 K. Additionally, the Gibbs free energy was decomposed into gas-phase $\Delta_R H^\circ_{diss}$, gas-phase $-T\Delta_R S^\circ_{diss}$ and $\Delta_R G^\circ_{sol}$ in DCM and the corresponding Gibbs free energy components as a function of the total Gibbs free energy is shown in Figure 2.8 for both SMD and COSMO-RS. It demonstrates that the estimated overall trends in Gibbs free energy reflect the corresponding trends in gas-phase enthalpy of dissociation. The Gibbs free energy of solvation and the entropic component show opposing trends but with significantly smaller slopes so that gas-phase enthalpy dominates.

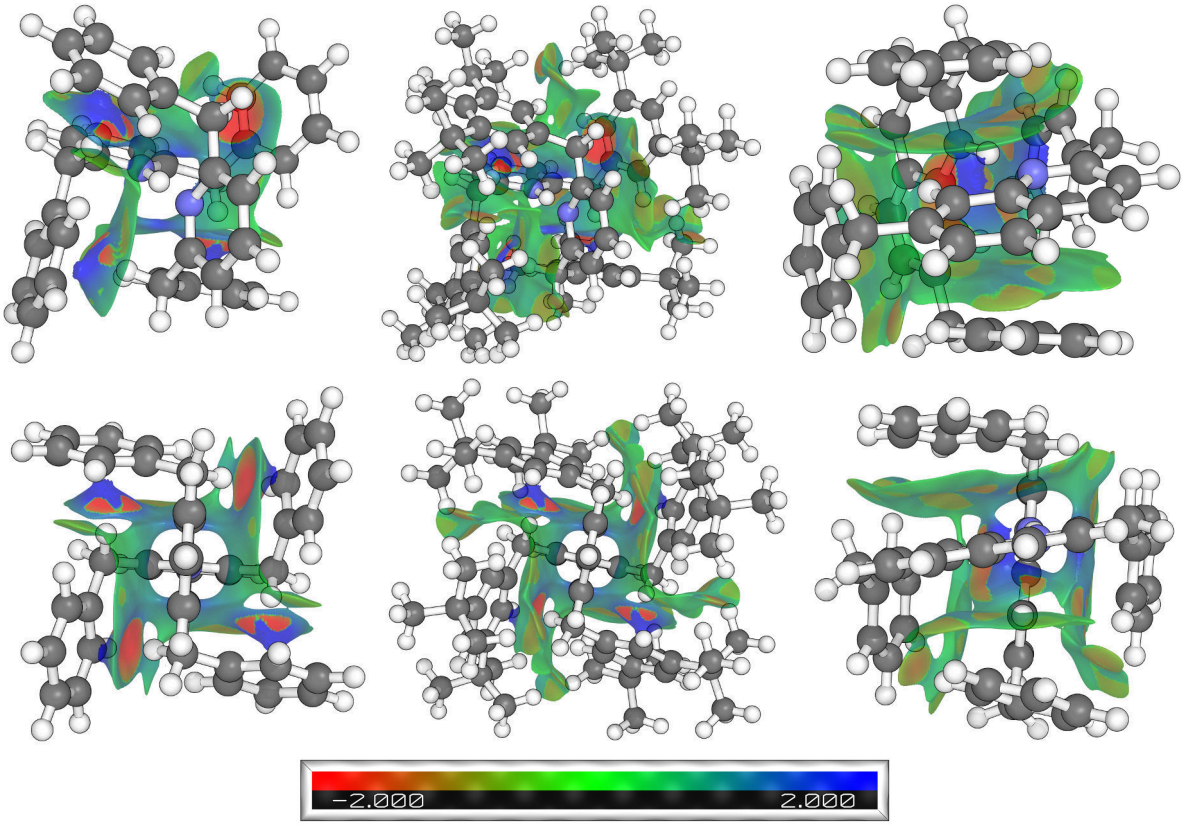


Figure 2.7.: NCI plots of the B97-D3/pc-2-sp(d) structures of the hydrogen-bonded dimers of **19** (left), **25** (middle) and **28** (right). Attractive interactions are shown in red, repulsive interactions are shown in blue.

$$\Delta_R G^\circ_{diss}(CH_2Cl_2) = \Delta_R E + \Delta_R ZPE + \Delta_R H^\circ_{corr} - 298.15\ \Delta_R S^\circ + \Delta_R G^\circ_{sol}(CH_2Cl_2)$$

$\Delta_R E$: Difference in electronic energies from single point calculations at M06-L/CBS level of theory using B97-D3/pc-2-sp(d) geometries

$\Delta_R ZPE$: Difference in zero-point energies from B97-D3/pc-2-sp(d) level of theory assuming strictly harmonic oscillations

$\Delta_R H^\circ_{corr}$: Difference in enthalpy corrections from B97-D3/pc-2-sp(d) level of theory assuming ideal gas behavior on the basis of a single minimum conformer

$\Delta_R S^\circ$: Difference in entropies from B97-D3/pc-2-sp(d) level of theory assuming ideal gas behavior and using the Q-RRHO approach for low-frequency modes based on a single conformer

$\Delta_R G^\circ_{sol}(CH_2Cl_2)$: Difference in Gibbs free energies of solvation from either SMD (M05-2X/6-31g* level of theory) or COSMO-RS (BP86/TZP level of theory)

Scheme 2.5: Summary of the computational approach to obtain Gibbs free energies of dissociation of hydrogen-bonded dimers in DCM at 298 K.

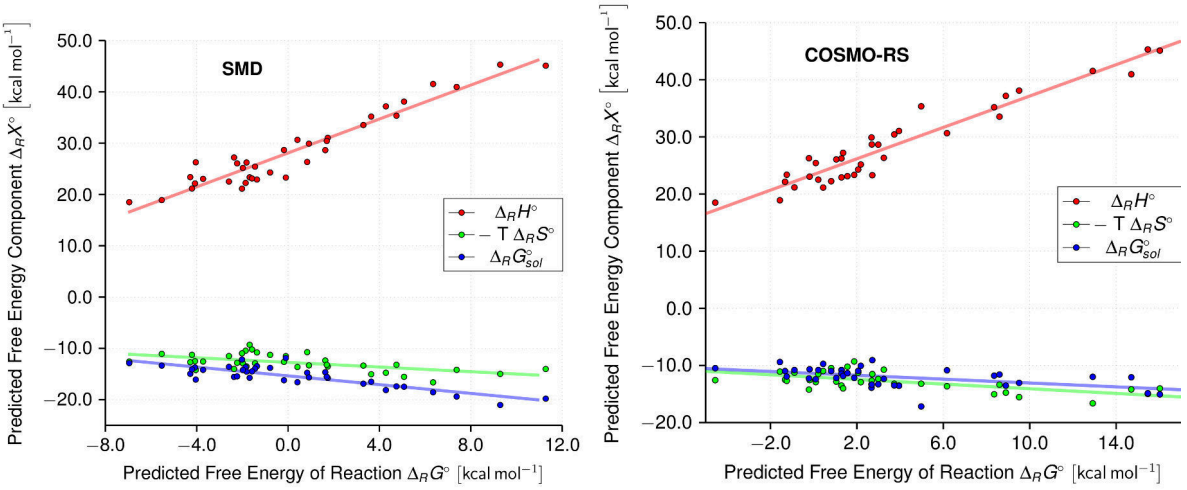


Figure 2.8.: Gibbs free energy of dissociation in DCM decomposed into gas-phase enthalpy, gas-phase entropy and Gibbs free energy of solvation for all the model compounds using either SMD (left) or COSMO-RS (right) as solvent model.

2.2.3. Experimental Investigation of Dissociation in Solution

To complement the computational investigation presented in the previous section, the Gibbs free energies of bond dissociation of the proton-bound dimers of all our model compounds was determined experimentally at −100 °C–30 °C with DCM as solvent. To carry out these experiments successfully, the solvent must be very dry. Using molecular sieves activated at 700 °C and 10^{-2} mbar we dried deuterated DCM with two successive stages of static drying to achieve a residual water content of smaller than or equal to 0.1 ppm as tested by means of ^{1}H-NMR analysis. We could not observe any residual peak for water anymore. Another important aspect is that the relative amounts of nitrogen base and the corresponding protonated salt needs to be known precisely. To minimize errors originating from weighing or measuring volumes, the ratios were determined based on ^{1}H-NMR analysis of the NMR samples with long relaxation delays utilizing the integrals of the signals of $BArF_{24}^{-}$ to quantify the amount of salt by comparing it to the integrals of the signals of the nitrogen base moiety.

In addition, we also looked into the effect of potential ion pairing in solution because both the protonated nitrogen base salt and the proton-bound dimer present in solution are monocations and, hence, potentially ion paired in a specific way with the counterion. Typically, ion pairing is also temperature-dependent and, hence, the temperature-dependence of the equilibrium constants of proton-bound dimer dissociation could be affected by it. This is one of the main differences of the solution experiments from the gas-phase measurements mentioned earlier which utilize the charged species directly. To minimize the corresponding effects, a non-coordinating counterion was used. We chose $BArF_{24}^{-}$ for that purpose as it was demonstrated not to affect protonated pyridines and proton-bound dimers of pyridines in NMR experiments under similar conditions.[87] Another recent study showed based on DOSY NMR studies that $BArF_{24}^{-}$ is still ion pairing in rather non-polar solvents but the difference to so-called coordinating anions is that $BArF_{24}^{-}$ coordinates in an unspecific way to the cation as opposed to coordinating anions.[107] Hence, we also used DOSY NMR to test for the presence of ion pairing between the protonated nitrogen base cations and $BArF_{24}^{-}$ using a small subset of all the model compounds covering both the small and large extremes with respect to cation size, as size usually also influences ion pairing.[108]

We found no indication for ion pairing for the large cation of **25** at room temperature and we found residual ion pairing to a very small extent for the small cation of **1**. Hence, ion pairing is widely absent in our test system and differential effects with respect to temperature are therefore negligible.

Having accounted for all these issues, the VT-NMR experiments could be performed. A typical data set is shown in Figure 2.9. The data fit allows estimation of both enthalpy and entropy from the temperature-dependent equilibrium constant. Using this method, we determined the thermodynamic parameters for the model compounds and visualized the data in a so-called compensation plot[109] in Figure 2.10.

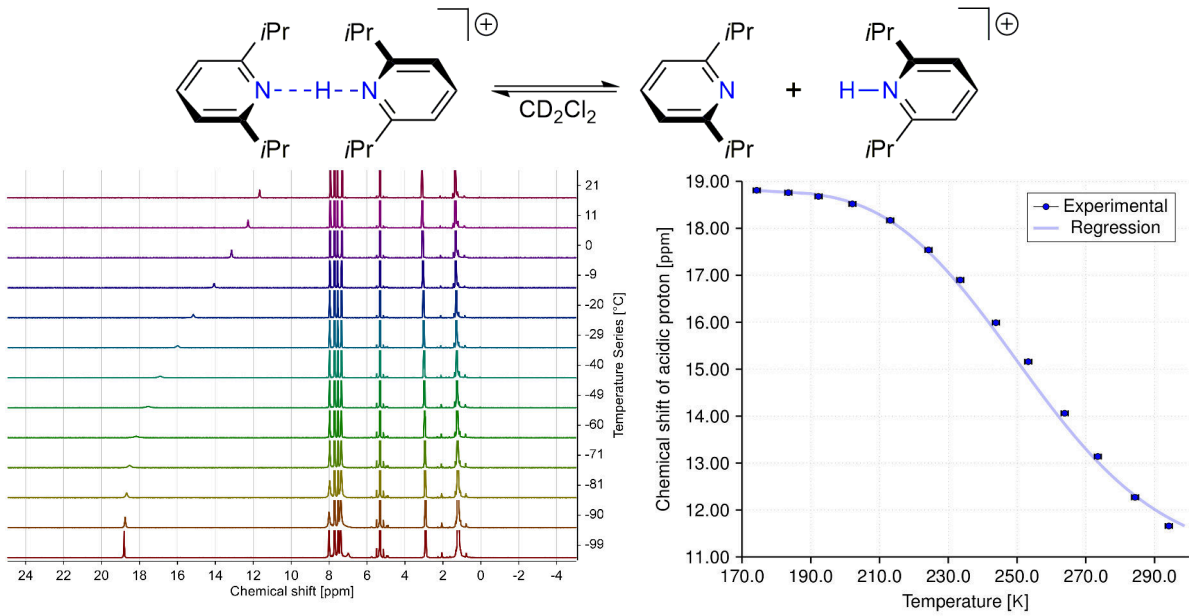


Figure 2.9.: VT-NMR data (bottom left and right) of the proton-bound dimer equilibrium of **5** (top). Bottom left: Overlay of the temperature series of ^{1}H-NMR spectra. Bottom right: ^{1}H-NMR shift of the N-H proton as a function of temperature with corresponding data fit.

Figure 2.10 indicates the presence of an enthalpy-entropy compensation (EEC). Treating all data points equally, the corresponding slope, which is called the compensation temperature,[109] was (234 ± 25) K. This is close to the experimental mean temperature. After separating the data points with respect to

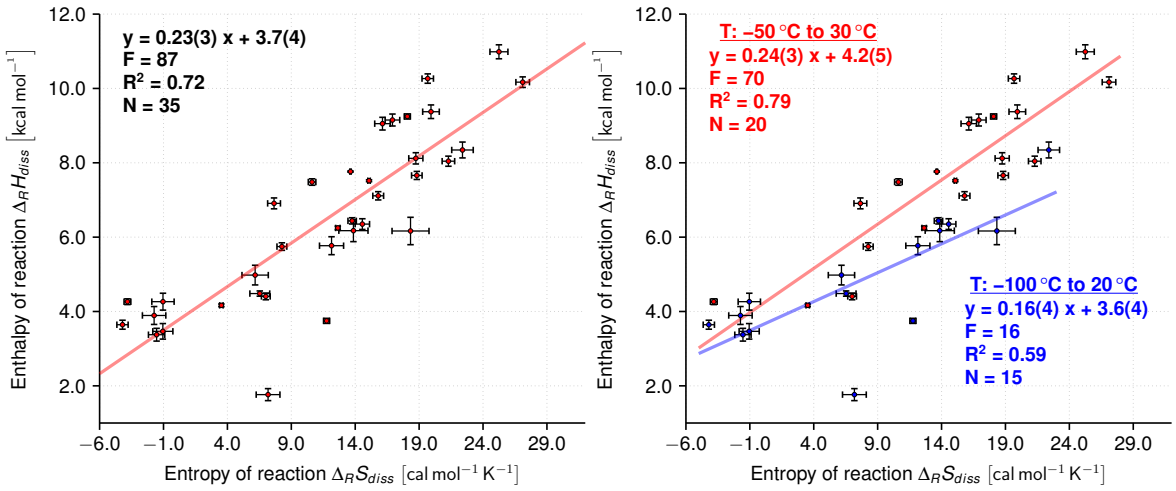


Figure 2.10.: Compensation plots of proton-bound dimer dissociation equilibria in DCM. Left: All data points are analyzed with one linear regression. Right: The data points are grouped with respect to their respective experimental temperature ranges.

their specific temperature ranges, the higher temperature data points resulted in a compensation temperature of (238 ± 28) K and the lower temperature data points resulted in a compensation temperature of (155 ± 36) K. The former compensation temperature remains unchanged but the latter changes significantly and is actually outside the experimental temperature range. We used several statistical tests developed in the literature[109–113] to determine the statistical significance of the EEC observed. Perhaps the simplest test is plotting enthalpies and entropies of reaction with their respective error bars as in Figure 2.10, which suggested that the correlation is statistically valid. An alternative test developed by Krug[111–113] determines the correlation of the enthalpy of reaction against the Gibbs free energy of reaction at the harmonic mean of the experimental temperatures used. The rationale for that test is that it can be shown that these quantities are not correlated at the harmonic mean of the temperature.[111–113] Notably, presence of a significant correlation in that test suggests the EEC not to be caused by correlated experimental errors. However, no significant correlation in that test does not show that the EEC is caused by correlated experimental errors. Hence, we determined these correlations for both the low temperature and the high temperature data sets. The results are depicted in Figure 2.11. Both subsets of the data do not have any considerable correlation meaning that

this plot does not provide evidence for an EEC which is not caused by correlated statistical errors.

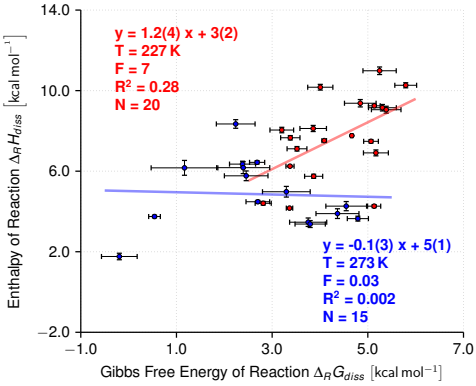


Figure 2.11.: Correlation of the enthalpy of reaction against the Gibbs free energy of reaction at the harmonic mean of the experimental temperatures for the proton-bound dimer dissociation equilibria.

Recently, a more rigorous statistical test has been developed, which is based on an error simulation approach. It provides a measure of the probability that an EEC observed in experiment can be explained by random and correlated experimental errors. The results of this p-test for both the lower and higher temperature subsets are illustrated in Figure 2.12. We find probabilities of 0.01 and 0.03 for the low and high temperature subsets, respectively, that the correlation is caused by random and correlated errors. Hence, based on this test it is unlikely that the EEC observed is the consequence of correlated errors.

Summarizing all the results of these statistical tests, it is unclear whether the observed EEC is physically meaningful or not. In addition to these statistical tests, we also benchmarked the VT-NMR results with ITC measurements of a small subset of the model compounds. While Gibbs free energies from ITC and VT-NMR agreed within less than 1 kcal/mol, both enthalpies and entropies of reaction showed a significant deviation. Hence, we decided to only use Gibbs free energies of dissociation to compare to values obtained from computations. In addition, these findings suggest that the EEC we found is likely not statistically significant.

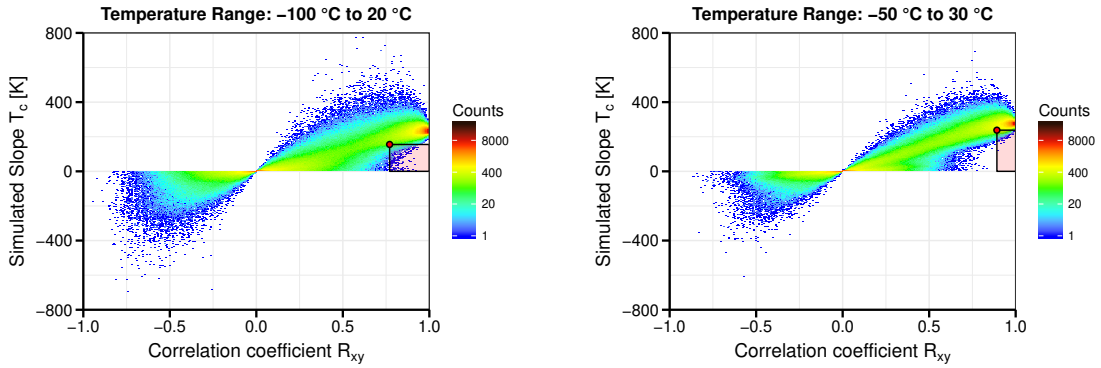


Figure 2.12.: Statistical p-tests of random experimental points to reproduce the experimentally observed compensation temperatures. The color of the pixels indicates the abundance of the respective combination of correlation coefficient and slope in the simulated data points. The pixels within the red rectangle correspond to simulation data points that are consistent with the experimentally observed correlation, which is indicated by the red circle.

2.2.4. Comparison of Computed and Experimental Energies

By comparing the results from the computations and the experiments, we wanted to gain understanding of the proton-bound dimer equilibria. Hence, we constructed plots of the Gibbs free energies of reaction of all the model compounds against the corresponding computed BDEs from M06-L/CBS//B97-D3/pc-2-sp(d) at various temperatures. The corresponding results are depicted in Figures 2.13-2.15.

At 180 K, the data is grouped into three subsets which is equivalent to the grouping into small systems, large pyridines and large quinolines depicted in Scheme 2.5. All the small systems are within a small region. All the Gibbs free energies of the large pyridines correlate significantly with the computed BDEs. This suggests that the corresponding trend in Gibbs free energies in DCM at low temperatures is governed by the respective trend of the BDEs in the gas phase. This suggests that the solution trend is affected significantly by London dispersion because we demonstrated before that the gas-phase BDEs are mainly influenced by inter-fragment dispersion (cf. Figure 2.6). At higher temperatures, the correlation becomes less significant and 300 K it is hardly statistically significant.

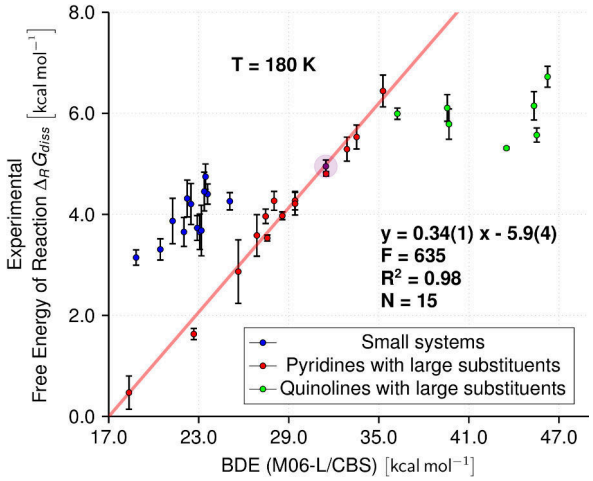


Figure 2.13.: Plot of the Gibbs free energies of dissociation of the model compounds at 180 K against their computed gas-phase BDEs.

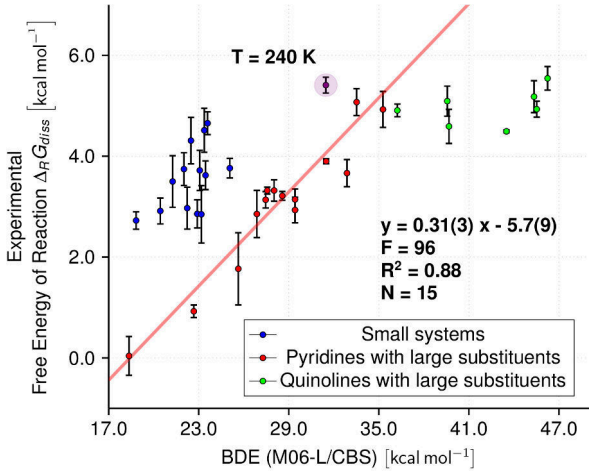


Figure 2.14.: Plot of the Gibbs free energies of dissociation of the model compounds at 240 K against their computed gas-phase BDEs.

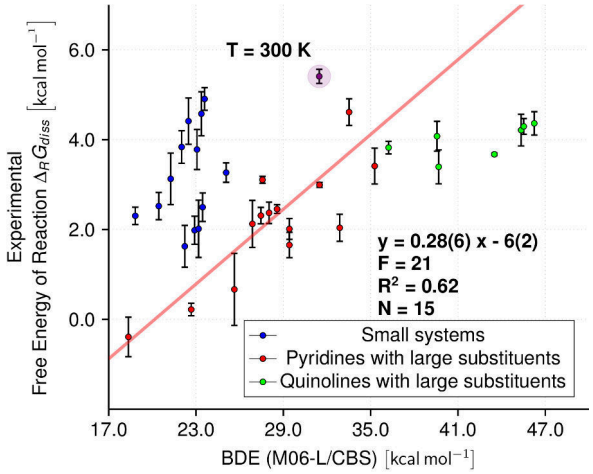


Figure 2.15.: Plot of the Gibbs free energies of dissociation of the model compounds at 300 K against their computed gas-phase BDEs.

Notably, all the data points of the large quinolines show some overlap with the large pyridines. However, at 180 K the corresponding data points are essentially invariant with respect to the gas-phase BDEs. At higher temperatures, the Gibbs free energies of the large quinolines correlates somewhat with the corresponding gas-phase BDEs but with a smaller slope than the data points of the large pyridines.

As final test, we wanted to compare the computed Gibbs free energies of proton-bound dimer dissociation at 298 K with all the experimental results. The corresponding comparison utilizing both SMD and COSMO-RS is depicted in Figure 2.16. For both COSMO-RS and SMD only a limited number of data points is within 1 kcal/mol, which is usually denoted as within chemical accuracy. Several data points deviate by several kilocalories per mole showing that both solvent models fail to reproduce the experimental numbers accurately.

2.3. Discussion

Summarizing our findings results in the three main conclusions of this chapter:

(i) London dispersion is important for the BDEs of proton-bound dimers with increasingly large substituents in the gas phase and can contribute as much as 10 kcal/mol for molecules with 100 – 200 atoms.

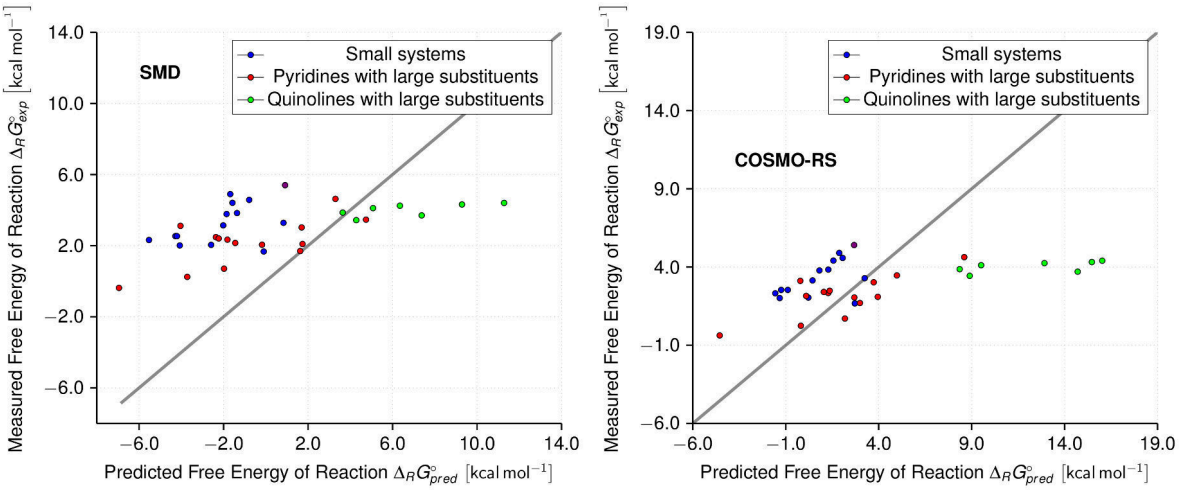


Figure 2.16.: Computed Gibbs free energies of dissociation of proton-bound dimers at 298 K in DCM compared to experimental results using SMD (left) and COSMO-RS(right) for the Gibbs free energies of solvation.

(ii) London dispersion is compensated by roughly 70 % in DCM solution compared to the gas phase.

(iii) Dispersion-corrected continuum solvation models are not suitable to reproduce the experimental Gibbs free energies of bond dissociation in DCM accurately.

In this section these three claims will be investigated separately in detail.

2.3.1. Theoretical Investigation of Gas-Phase Dissociation

EDA methods provide a means to decompose interaction energies in molecular systems into physically well-defined terms to gain insight into underlying trends.[114] SAPT[115,116] is one of the more widely used approaches in that respect. It starts from the wavefunctions of the isolated fragments and treats the interaction between these two fragments as perturbation to the corresponding isolated fragment Hamiltonians. The overall interaction is partitioned into electrostatics, exchange, induction and dispersion. While the definition of the respective energy components is theoretically clean, the main drawback of SAPT, due to its perturbative nature, is that it can only treat weak intermolecular complexes. Therefore, it is well-suited for the proton-bound dimers studied in this

chapter. SAPT confirms our model system design, which was aimed at changing the overall BIEs by tuning inter-fragment dispersion. Notably, non-negligible contributions also stem from electrostatics, which is likely due to quadrupole-quadrupole interactions between the substituents, especially in the case of (partially) aromatic substituents. However, these interactions will increase in parallel with dispersion and provide additional stabilization.

In the field of DFT, proper treatment of NCIs, especially in intermediate distance regimes, is one of the main research areas of interest.[3,4] Particularly relevant for our model compounds, various DFT approaches accounting for NCIs in molecules with hydrogen bonds and dispersion were tested in detail by Sherrill and coworkers.[83] Recently, an upsurge in local correlation methods resulted in the development of new affordable WFT approaches to treat molecules with in the range of hundreds of atoms. WFT approaches are generally considered less empirical compared to DFT methods and there are usually well-defined ways to arrive at better accuracy. One particularly notable method in that respect is DLPNO-CCSD(T),[30,37,86] which is quite affordable for molecules of up to 200 atoms using state-of-the-art computer hardware. In addition, several benchmarks demonstrated that DLPNO-CCSD(T) provides very good accuracy, which is claimed to be comparable to canonical CCSD(T).[3,31,33,34] Notably, CCSD(T), at the basis set limit, is considered the gold standard for single reference computations. DLPNO-CCSD(T) promises similar accuracy at a signifcantly lower computational expense and especially at a significantly reduced computational size scaling.

Our computational methods were selected as a trade-off between computational cost and accuracy. Hence, B97-D3, in conjunction with a sufficiently large basis set, allowed us to obtain good geometries and good thermodynamical corrections of all our model compounds within a reasonable time frame. B97-D3 is a GGA functional and has been shown to provide both good energies and geometries for organic systems.[83] For our model compounds, it provided very comparable geometries as the significantly more expensive M06-L functional, which is a meta-GGA functional. However, single points with M06-L were still computationally cheap and, hence, could be readily performed for all our model compounds. Notably, M06-L was optimized for main-group and organometallic

thermochemistry, as well as noncovalent interactions. To complement these DFT methods with a WFT approach, we also performed DLPNO-CCSD(T) single points using the B97-D3 geometries and compared all the results.

We were surprised to find that the DLPNO-CCSD(T) energies did not perform best compared to the experimental results in the gas phase. DLPNO-CCSD(T) resulted in significant overestimation of the BDEs of the large pyridines. While B97-D3 energies closely agreed with DLPNO-CCSD(T) results and, hence, did not agree well with experiment either, M06-L showed best agreement with all our experimental findings. That not only includes comparison of the computed BDEs of seven proton-bound dimers against nine experimental numbers from two distinct methods, but also the control experiment with two competitive CID channels. The control experiment was designed to minimize systematic errors from the modeling of the T-CID experiments. Both bond dissociations, *i.e.* both homolysis of O–NO and heterolysis of N-H–N dissociation, have a so-called loose transition state. Hence, this experiment does not require modeling of the kinetic shift and all the associated complications are avoided.[117–119] Based on statistical rate theory, one would expect the weaker of two bonds to break selectively if the corresponding difference in BDE is at least 2 kcal/mol.[120,121] Hence, the only assumption inherent to that competition experiment is statistical behavior of the two bond dissociations. The assumption of statistical behavior is inherent to quantitative measurements of BDEs in the gas-phase and was studied in great detail. The outcome of a vast amount of research is summarized by Armentrout and Baer,[120,121] with the main conclusion that molecular ions having more than a few heavy atoms behave statistically with only very rare exceptions. Furthermore, it is important to note that tables of BDEs for both charged and neutral small molecules like the NIST-JANAF thermochemical tables,[122,123] are largely based on ion dissociations and ion-molecule reactions. For all these measurements, statistical behavior needs to be assumed. Moreover, a recent database developed by Ruscic called ATcT[124,125] interconnects thermochemical data through thermochemical cycles to demonstrate consistency of the whole database. Gas-phase experiments requiring the assumption of statistical behavior are one of the major data sources of this database.

Compared to relatively small molecules studied in the literature, all our mo-

del compounds, and especially the corresponding proton-bound dimers, can be considered large which makes the assumption of statistical behavior even more robust. However, given the notion that CCSD(T) is considered the gold standard computational method for predicting thermochemistry of single-reference systems, the disagreement of DLPNO-CCSD(T) estimates for the BDEs of the proton-bound dimers of **25**, **27** and **28** is disturbing. There are several potential sources of this error. We used both NormalPNO and TightPNO to compute the DLPNO-CCSD(T) BDEs for our model compounds. These threshold could simply be not accurate enough for our model compounds.[35,126] Additionally, we perform CBS extrapolations based on double zeta and triple zeta basis sets, which could be insufficient for the systems under study. Efforts to perform quadruple zeta single points were unsuccessful due to the too high computational demands. In addition, the triples in DLPNO-CCSD(T) we used are not full iterative triples as in canonical CCSD(T) but a more crude approximation. Hence, the method we employed should rather be termed DLPNO-CCSD(T_0). Only very recently, DLPNO-CCSD(T_1) with full iterative triples has been developed and it could be tested whether this new method will yield better results.[38] However, DLPNO-CCSD(T_0) has only been observed to be problematic in rare cases, especially in systems with a low band gap.[38] Therefore, it is possible that the DLPNO approximation to CCSD(T) is not the main source of error but rather CCSD(T) itself.

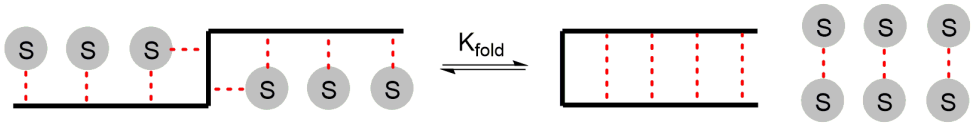
All the computational methods employed reproduce the expected BDE of the O–NO bond reasonably well. Computational estimates from B97-D3, M06-L and DLPNO-CCSD(T) for the O–NO dissociation in **38** are 38 kcal/mol–41 kcal/mol. Experimental results for comparable alkylnitrites are around 34 kcal/mol–42 kcal/mol and agree very well with the computed estimates for all of the methods used. For that bond dissociation, the contribution of dispersion is negligible because the dissociated NO is a very small and not very polarizable fragment. For the BDE of the N-H–N bond, the deviation between computational estimates and experimental results appear to scale with size and hence also with number of dispersive contacts. Since the computational DLPNO-CCSD(T) results agree very well with experiment for all the small pyridines but not with the big ones, it seems plausible that overestimation of dispersion by DLPNO-CCSD(T), and possibly even CCSD(T), is responsible for the observed

discrepancy. While such a systematic discrepancy has not yet been observed, the simple explanation for that is that molecules as large our test compounds have not been studied experimentally before. Notably, the excellent mutual agreement between B97-D3, sSAPT0 and DLPNO-CCSD(T) might seem to mutually support the validity of all the corresponding results given that these are three inherently very different theoretical approaches to determine interaction energies. However, this argument is misleading because all these methods were constructed and benchmarked extensively against the same reference methods, which is in most cases canonical CCSD(T).[3,26,33,74,83,127] Hence, rather than supporting the validity of all these methods, it casts some doubts on the common reference methods employed.

Based on all these results, we decided to use M06-L as method of choice to estimate the BDEs of our model compounds in the gas phase. However, it needs to be mentioned that all conclusions concerning the extent of compensation of dispersion in DCM and the poor performance of implicit solvent models to reproduce Gibbs free energies of dissociation of our model compounds in DCM remain valid when DLPNO-CCSD(T) or B97-D3 energies are used instead. The only difference is that implicit solvent models would perform even worse and compensation of dispersion in DCM would be higher by a few percent.

2.3.2. Experimental Investigation of Dissociation in Solution

Dispersion is at least partially compensated in solution because surrounding solvent molecules provide interactions to open interaction surfaces of solute molecules as illustrated in Scheme 2.6.[12,14] While in the gas phase the competition is between intramolecular interactions in the closed conformer and no intramolecular interactions in the open conformer, in solution the competition is between solute-solvent interactions in the open conformer and the combination of solute-solute and solvent-solvent interactions in the closed conformer. As solvent molecules interact with solute molecules, especially with nonpolar molecules, to a considerable extent by London dispersion, the relative strength of these interactions is similar in magnitude. Effectively, this leads to a compensation of dispersion in solution compared to the gas phase making it a less important energy component.



Scheme 2.6: Schematic depiction of folding and unfolding of a solute.

Based on these arguments and experimental observations,[14] it is well established that compensation of dispersion is very significant. However, the extent of compensation is yet unknown. Already in 1999, Houk studied the compensation of intramolecular dispersion in solution based on molecular balances using force fields and a dispersion-corrected continuum solvation representation of the solvent.[12] The main result was that solute-solvent dispersion have a considerable impact on folding equilibria of several organic molecules. Additional contributions come from Cockroft and coworkers based on both molecular torsion balances and association equilibria of urea hosts and naphthyridine guests in organic solvents.[14,72,73,128,129] Summarizing the corresponding results, it was found that both solvophobic contributions and dispersion are still important in solution and can be of similar magnitude in organic solvents. Overall, these results suggest that London dispersion is not fully compensated, but no estimate as to the extent of compensation in any solvent has been suggested.

First, the EEC we found needs to be addressed. EEC is defined as linear correlation between enthalpy and entropy of a reaction and has been found in numerous studies investigating chemical reactions or physical processes.[109,130–135] Accordingly, it has also enjoyed widespread discussion in various scientific communities. Various explanations have been provided as origin of EECs and the most thoroughly discussed is the presence of highly correlated errors in the experimental estimates of both entropy and enthalpy. Alternatively, based on thermodynamic models, environmental processes like solvent reorganization can be demonstrated to give rise to EECs and, hence, provide a universal mechanism for the occurrence of EECs, with the premise that the corresponding effect is large enough to be observable.[109,134] The bottom-line of either of these explanations for an EEC in a specific data set is that direct interpretation of enthalpies and entropies of reaction is extremely non-straightforward and difficult at best.[134] However, regardless of whether EECs originate from correlated errors

or from environmental processes, the Gibbs free energies of reaction are unaffected.[109,134] Therefore, while the presence of EECs in given data is curious by itself, when quantitative insight into intermolecular interactions is sought it is best to rely on Gibbs free energies of reaction and their corresponding trends.

As a final remark regarding the enthalpies and entropies of reaction, the corresponding results from VT-NMR and ITC measurements, which were carried out for a subset of all the model compounds, did not agree well. Notably, it has been found before in a number of cases that enthalpies and entropies of reaction from calorimetric methods and from the temperature-dependence of equilibrium constants did not agree with each other.[136–138] The origin lies in the inherently different methodology to derive enthalpies and entropies. Calorimetry measures enthalpy directly as released heat of the reaction, both Gibbs free energies of reaction and entropy of reaction are obtained indirectly through data fitting. Hence, enthalpies from calorimetric methods like ITC are the most reliable. When the temperature dependence of the equilibrium constant is used to determine enthalpy and entropy of reaction, both parameters are determined by data fitting, only the equilibrium constant is determined directly. In addition, to perform the data fitting, both enthalpy and entropy are assumed to be independent of temperature. In our case, the main problem is that across a temperature range of –100 °C–30 °C, which was investigated in the VT-NMR experiments, this assumption is likely not good leading to the observed discrepancies. However, it needs to be mentioned that also in ITC data EECs have been observed. One of the problems is that entropy of reaction cannot be measured directly and always has considerable uncertainty and errors in entropy from ITC are still correlated with errors in enthalpy.

To avoid difficulties in interpreting enthalpies and entropies of reaction, we focus on the Gibbs free energies of dissociation for our model compounds and its variation with temperature. We compared our experimental results to computed Gibbs free energy components and found the most reliable trends only when comparing experimental Gibbs free energies of dissociation with computed gas-phase BDEs. Similar trends are observed when the experimental values are compared to gas-phase enthalpies of reaction but with lower statistical confidence and more data spread. Notably, we obtain the same type of data classifi-

cation as observed in Figures 2.13–2.15 when B97-D3 or DLPNO-CCSD(T) are used as computational methods for the electronic energies.

Three groups of data points are revealed by Figures 2.13–2.15. The blue data corresponds to the small test compounds and pyridines with small substituents. They cluster together in a very narrow area because the gas-phase BDEs are only within a narrow range as are the Gibbs free energies of dissociation in DCM. In these compounds, the gas-phase BDEs have reduced electrostatics compared to **1** because of non-ideal N-H–N geometries caused by avoiding steric repulsion but increased dispersion due to the introduction of substituents which are spatially close. Overall, these effects roughly cancel each other and, hence, these compounds are not useful for estimating the extent of compensation of London dispersion in DCM.

Pyridines with large substituents are depicted in red in Figures 2.13–2.15. In all these compounds, the N-H–N geometry is distorted to a similar extent, rendering the H-bond strength roughly constant across the whole series of compounds. As increasingly large substituents are introduced, London dispersion leads to a significant increase in gas-phase BDE and also in an increase in Gibbs free energy of dissociation in DCM, especially at lower temperatures. The good linear correlation observed between experimental Gibbs free energy in solution and computed BDE in the gas phase alludes to the corresponding trend in solution being caused by the trend in gas-phase BDEs. The slope of this correlation is a direct measure of the extent of compensation of London dispersion in solution. We find compensation of dispersion to be 65 %–78 % at 180 K–300 K in DCM. Hence, it can be directly deduced that compensation is large, as has been suggested in the literature,[12,14] but far from complete. The slope is technically an estimate of the compensation of all intermolecular interactions in DCM in our model compounds. However, since the trend in gas-phase BDE of the pyridines with large substituents is governed by dispersion, the estimate of compensation also corresponds to an estimate of compensation of dispersion.

Notably, our estimate of compensation of dispersion includes both enthalpic and entropic energy components, and the gas-phase entropic contribution is unlikely to be insignificant and should not be disregarded. Correlating our computed gas-phase Gibbs free energies of dissociation against the corresponding

gas-phase BDEs results in an estimate of compensation of about 20 %. Hence, the main portion of the overall compensation of about 70 % is due to solvation effects, but a significant part is due to gas-phase entropic contributions. This implies that, in order to maximize dispersion in solution, especially when going to higher temperatures than studied in this data, entropic contributions need to be minimized. This is demonstrated by the data point of **29**, which is depicted in purple in Figures 2.13–2.15. At 180 K, the corresponding data point fits well to the pyridines with large substituents. However, at higher temperatures, its Gibbs free energy of dissociation is considerably larger than expected from the other data points. **29** has rigid alkynyl substituents in the 2 and 6 positions and, hence, the substituents have only minimal conformational freedom. Therefore, the corresponding data point corresponds also to the strongest proton-bound dimer in solution at room temperature.

Quinolines with large substituents are depicted in green in Figures 2.13–2.15. While these compounds show the largest gas-phase BDEs across all the model compounds, the corresponding Gibbs free energies of dissociation in DCM are not accordingly large. At 180 K, their Gibbs free energies of dissociation are essentially independent of gas-phase BDEs suggesting a very large extent of compensation, larger than in the pyridines with large substituents. One explanation for that is offered by the difference in molecular structure of the corresponding proton-bound dimers, which is demonstrated in Figure 2.7. As mentioned previously, the large quinolines are not as compact as the large pyridines and are therefore also more flexible. The additional free space within the proton-bound dimer could be occupied by solvent molecules. This would lead to significantly increased compensation of dispersion and justifies excluding the corresponding data points from the correlation to estimate compensation of dispersion. Notably, we found the salts and dimers of the large quinolines to possess significantly higher solubility compared to the corresponding large pyridines which could also be a consequence of solvent molecules inside the dimers. To confirm this hypothesis, additional studies would be necessary.

The compensation of dispersion can also be roughly estimated without grouping the data points. The total range of gas-phase BDEs in our model compounds is 26 kcal/mol. The corresponding Gibbs free energies in DCM span

5.5 kcal/mol–6.3 kcal/mol at 180 K–300 K. The ratio of these ranges provides an estimate of 76 %–79 % of compensation of dispersion at 180 K–300 K, which is reasonably close to the more sophisticated estimate based on grouping of the data points and linear correlation.

2.3.3. Comparison of Computed and Experimental Energies

One of the holy grails in computational chemistry is accurately predicting rate constants and equilibrium constants for physical and chemical processes in solution.[139–141] There are two main methodological philosophies to achieve this goal. First, the system under study can be directly computed in solution, both with and without accounting for molecular dynamics. Secondly, the system under study can be computed indirectly by the use of thermodynamic cycles adding Gibbs free energies of solvation on top of the corresponding parameters in the gas phase. These philosophies are independent with respect to the choice of solvent model and, accordingly, implicit,[106,142,143] explicit[144,145] and hybrid models[146] of solvation[141] can be applied for that purpose. Of all the types of solvent models listed here, explicit solvation in conjunction with molecular dynamics simulations of the full system under study would be the most rigorous method from a conceptual point of view. However, it is also computationally the most expensive approach and, additionally, requires expert knowledge of the methodology. Therefore, it is not widely used for predicting rate constants and equilibrium constants in organic solvents.[144,145] Notably, it is more popular in biomolecular simulations, but, for organic solvents, only a limited number of reliable models exists. The most popular approach in organic and organometallic chemistry is to carry out QM computations of the systems and processes under study in the gas phase and accounting for the solvent with implicit solvent models, because the corresponding methods usually strike a reasonable balance between computational cost and accuracy. In addition, the corresponding methods are implemented in a wide range of electronic structure calculation programs and are much simpler from an operational point of view compared to the molecular dynamics approaches because the corresponding computations can be performed in the same way as standard DFT and WFT calculations. Furthermore, several benchmarks suggest that QM methods in conjunction with continuum solvation provide

reasonable accuracy.[6,147–150] Hybrid approaches are somewhat in between fully explicit and implicit solvent models in terms of cost and technical difficulty. The main disadvantage is that no black-box program exists and the corresponding computations need to be done by hand. This explains the limited number of reports adopting a hybrid approach.[151,152] Additionally, it has been shown recently that the number of solvent molecules treated explicitly is absolutely crucial to obtain accurate results.[153] However, the corresponding number is very system specific and, hence, considerable effort to find the necessary number of explicit solvent molecules is required.

For modeling our test systems, we decided to use QM computations in conjunction with continuum solvation models. We used SMD and COSMO-RS, which both explicitly account for London dispersion between solute and solvent. We strictly followed the thermodynamic cycle approach, which was also employed for parametrization of both solvent models. The advantage of that approach is that it directly provides a decomposition of the Gibbs free energy of dissociation in solution in gas-phase enthalpy, entropy and Gibbs free energy of solvation. This decomposition could be used to understand the origin of the trends in the data. In addition, comparing the experimental Gibbs free energies of dissociation in DCM against the corresponding computational estimates is a direct test of the accuracy of the implicit solvent models used with respect to molecules with significant contributions of London dispersion. Notably, since we utilized deuterated DCM for our measurements, a solvent equilibrium isotope effect could contribute to our experimental results but for NCIs it is expected to be very close to 1 and, hence, insignificant.[154]

Figure 2.16 clearly shows that the overall computational approach adopted cannot account properly for the observed experimental trends. The most significant deviations between experiment and computation are 12 kcal/mol, which is equivalent to a deviation of eight orders of magnitude in the equilibrium constant at 298 K. It is also important to point out that the deviations show systematic behavior for both solvent models. The slopes of the correlations of the experimental with the computational Gibbs free energies of dissociation differ significantly, indicating that the corresponding trends with respect to London dispersion and compensation thereof are not properly reproduced. Considering all the possible

sources of error, the main problem could lie in three different areas. First, the error could be caused by errors in the electronic structure computations. However, since these values are individually benchmarked against experimental results, the corresponding errors are expected to be insignificant. Secondly, the gas-phase contribution to the entropy of reaction could lead to systematic errors. In principle, this is a plausible source of error since we did not properly account for full conformational ensembles of all the model compounds and their proton-bound dimers. Notably, this error is expected to increase with the flexibility of the model compounds. Since the computed Gibbs free energies of dissociation in DCM are governed by the gas-phase enthalpy, the most important source of error should also correlate with the enthalpy to counteract it. As the number of accessible conformers generally increases with molecular size in terms of number of atoms and as the dispersive attraction also tends to increase with molecular size of our model compounds, it is conceivable that this error would also correlate with the gas-phase enthalpic component.

Investigating the gas-phase entropic contribution as source of the deviations between computed and experimental Gibbs free energies of dissociation in DCM in more detail, the data point of **14** is very insightful. Steric repulsion makes the corresponding dimer adopt a very unfavorable N-H–N geometry and the overall gas-phase BDE is small because dispersion is still not big enough to make up for the lost stabilization of the H-bond. In solution, the corresponding Gibbs free energy of dissociation is significantly underestimated by our computational approach. The flexibility of the substituents of **14**, however, suggests that only accounting for a single conformer is not adequate for that compound. Because the conformational space accounted for in our computations is smaller than it is expected to be, our computational methodology should underestimate the corresponding gas-phase entropy of dissociation and, hence, overestimate the Gibbs free energy of dissociation rather than underestimate it. While this argument does not exclude a significant contribution of the gas-phase entropic component to the overall observed discrepancy between computation and experiment, it suggests that there is a different energy component with an error that is more significant.

Having closely examined all other likely sources of the observed discrepancy

between computation and experiment, the only remaining possible source of error is Gibbs free energy of solvation. The deviations observed can be rationalized by solute-solvent dispersion being underestimated significantly, which in turn leads to significantly underestimated compensation of dispersion. It can even explain the discrepancy observed for **14**. Solute-solvent dispersive interactions are expected to alter the preferred geometry of the proton-bound dimer of **14** in solution significantly compared to the gas phase. In solution, a more ideal N-H–N geometry would be preferred at the expense of intramolecular dispersive interactions, which are compensated by dispersive solute-solvent interactions of similar magnitude. This would lead to a Gibbs free energy of dissociation in DCM which is larger than predicted by our computational approach. We investigated the influence of the implicit solvent models on the equilibrium geometries of all our model compounds. The Gibbs free energies of solvation did not differ significantly between geometries optimized in the gas phase and geometries optimized in solution. Hence, the geometries in solution determined in conjunction with implicit solvent models likely do not properly account for the effect of solvation, which could also be a consequence of the underestimation of solute-solvent dispersion.

From a conceptual perspective, in implicit solvent models like COSMO-RS and SMD, London dispersion is accounted for with a term proportional to the solute-solvent interaction surface area.[6,105] Atom-pairwise dispersion corrections are essentially based on interaction volume. Merely considering the difference in scaling between these two approaches, it seems reasonable that systematic differences between the two approaches cannot be avoided when going to increasingly large systems. The expected discrepancies were demonstrated based on MD simulations of ubiquitin to estimate Gibbs free energies of solvation. An explicit account of solvent yielded results which differed by as much as 45 kcal/mol from an implicit solvent model approach.[155] In addition, current continuum solvent models are unable to treat entropy of solvation properly as configurational entropy is not treated at all, which results in significant systematic deviations in computed solvation energies. However, very recently, a statistical thermodynamic model has been proposed to alleviate this problem and it remains to be seen whether this new model is able to reduce the methodological problems of implicit

solvent models, at least with respect to the entropic free energy component.[156] In addition, parametrization of solvent models needs to be done at a range of temperatures to properly account for the temperature-dependence of the Gibbs free energy of solvation.

Altogether, the approach we adopted requires both accurate Gibbs free energies of dissociation in the gas phase and accurate Gibbs free energies of solvation of all the involved species in order to deliver good results for the Gibbs free energies of dissociation in solution. The gas-phase component can be computed quite accurately when state-of-the-art computational methods are employed, which are benchmarked against proper experimental data. For best results, multiple conformers should be considered using Boltzmann averaging or MD should be performed for direct account of canonical ensembles to treat both enthalpy and entropy of reaction in a proper way. Currently, implicit solvent models to account for solvation are able to produce good results for many systems but are merely a compromise between computational cost and numerical accuracy. Hence, they are likely the weakest link for the computation of accurate Gibbs free energies of chemical reactions in solution.

2.4. Conclusions

In this chapter we discussed a comprehensive study of bond dissociations of a systematic set of proton-bound dimers both in the gas phase and in solution by a combination of experimental and theoretical methods. From our results it can be deduced that London dispersion becomes important in the gas phase already for molecules of relatively moderate size of less than 100 atoms. In DCM, dispersion is compensated by about 70 % due to competitive solute-solvent interactions which are intrinsically absent in the gas phase. While the extent of compensation is very significant, it is still far from complete which allows dispersion to be used to tune interaction strengths in solution as well as reactivity and selectivity of chemical reactions. Furthermore, state-of-the-art continuum solvent models are unable to reproduce the compensation of dispersion properly which leads to a significant overestimation of dispersion in solution.

All our findings are in principle only valid for the model compounds studied.

However, they can be extrapolated to a wide range of systems. By careful computational and experimental data analysis, we were able to minimize any contributions from the hydrogen bond and only look at the trends which are a direct result of the change in substitution of the proton-bound dimers. The substituents were limited to alkyl and aryl groups without any heteroatoms and, hence, our results only probe dispersion in these structural motifs. However, given the ubiquity of alkyl and aryl groups in organic chemistry and the interactions between them, our results can be transferred to all systems with similar dispersive interactions. Furthermore, the observation of the systematic problems of implicit solvent models with respect to properly accounting for London dispersion are likely due to methodological problems rather than parametrization problems and, hence, they are expected to fail similarly for other organic and even organometallic systems. Given the widespread use of DFT+SMD or DFT+COSMO-RS computational approaches in the chemical community, we believe that our results have wide implications for homogeneous catalysis and supramolecular chemistry even in polar organic solvents.

3. Compensation of Dispersion in Aprotic Solvents

Felix Fleckenstein is acknowledged for the synthesis of several pyridinium salts and for a few VT-NMR measurements. This chapter is adapted with permission from "Compensation of London Dispersion in the Gas Phase and in Aprotic Solvents," Robert Pollice, Felix Fleckenstein, Ilya Shenderovich, and Peter Chen *Angew. Chem. Int. Ed.* **2019**, *58*, 14281 – 14288. Copyright © 2019 John Wiley and Sons.

3.1. Introduction

It has been established in the previous chapter that compensation of London dispersion in DCM solution is very significant, amounting to about 70 % but far from complete. While other studies appeared in the literature showing that compensation of dispersion is incomplete, the extent of compensation was not estimated.[14,72,73,129,157–159] Hence, it is clear that London dispersion can be used as a design element to affect stability and reactivity of molecules in solution and that the introduction of DEDs[8,18,160] can be a valuable tool in that regard. In principle, an independent way to tune the strength of dispersive interactions is to change the environment of a system. Currently, steric effects are one of the main tools used to design stereoselective organic reactions. Notably, steric effects were classically treated as intrinsic to the molecules of interest. However, now it has been shown that they are rather extrinsic and can be tuned by the environment.[18]

In solution, based on the current understanding of London dispersion, it seems plausible that the extent of compensation can be altered by changing the solvent. This would offer attractive means for rational design of intermolecular in-

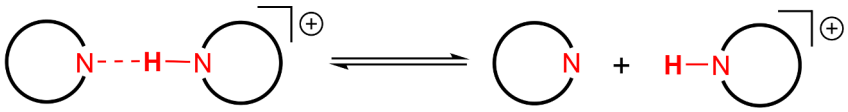
teractions, chemical reactivity and selectivity, which are orthogonal to the introduction of DEDs. In a recent study, Cockroft tackled the solvent influence on dispersion in molecular torsion balances having alkyl-alkyl interactions and having perfluoroalkyl-perfluoroalkyl interactions.[72] In that study, a statistically significant correlation of bulk solvent polarizability and Gibbs free energy of folding was observed. However, the Gibbs free energy could at most be changed by 0.37 kcal/mol, which is not a very big extent. In addition, the main contribution to the change in Gibbs free energy of folding as a function of solvent was caused by solvophobic interactions rather than bulk solvent polarizability. Therefore, further studies in that regard are required to find out, how much dispersion forces can really be tuned by rationally changing the solvent. Based on the previous literature findings, we expected that significant tuning of dispersive interactions by solvent is indeed possible and would lead to very significant changes in the extent of overall compensation of dispersion leading to differences in Gibbs free energies of dispersion-dominated proton-bound dimers of several kilocalories per mole.

The main purpose of this chapter is to extend on the findings of the previous one by performing the same type of measurements in several additional solvents. Furthermore, ITC control experiments were performed for a larger subset of proton-bound pyridines to provide a independent support for the VT-NMR data. The main goal is to establish the total compensation of dispersion as a function of bulk solvent polarizability. In addition, we improve on the computational approach developed in the first chapter by accounting for the presence of multiple low energy conformers in all the model compounds and their proton-bound dimers. This leads us to a solid estimate of the compensation of dispersion in the gas phase by entropy as a function of temperature, which in turn allows us to separate the total compensation of dispersion in gas-phase compensation and compensation by solvent, both as a function of temperature. Furthermore, the improved computational model provides us with estimates of the compensation of dispersion estimated from computations alone which can be directly compared to our experimental findings.

3.2. Results

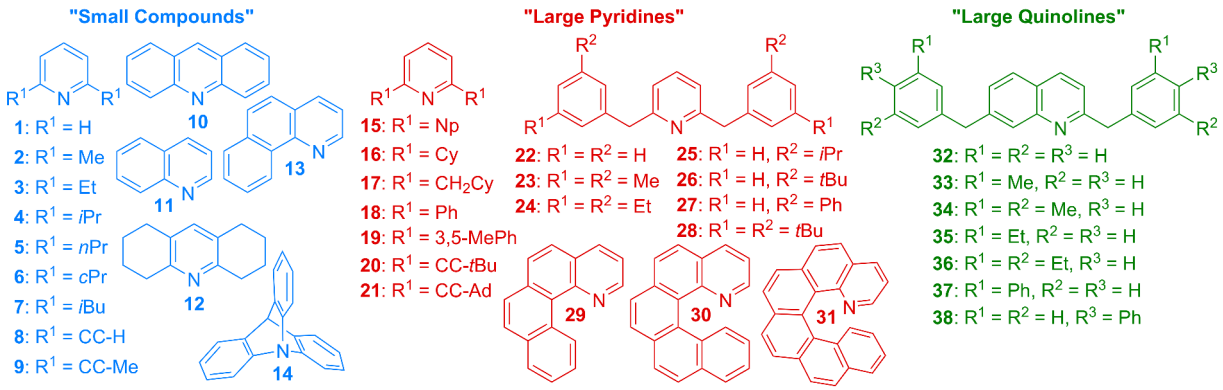
3.2.1. Model Compounds and Solvents

We started this study with the proton-bound dimers established in the previous chapter as test system to probe the influence of London dispersion on intermolecular interactions in solution. The bond dissociation equilibrium under study is depicted in Scheme 3.1.



Scheme 3.1: Bond dissociation equilibrium of proton-bound dimers.

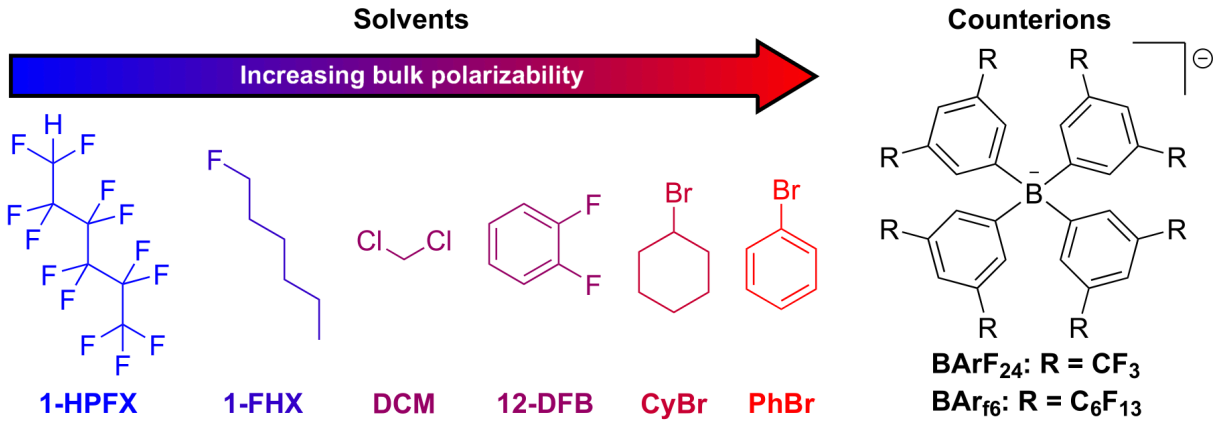
We extended the set of model compounds used in the previous chapter by two additional rigid alkynyl-substituted pyridines. The complete set of compounds studied in this chapter is depicted in Scheme 3.2. As established in the previous chapter, the main energy component of the total interaction energy that is changing with the substituents in the model compounds is London dispersion.



Scheme 3.2: Model compounds studied in this chapter and their classification.

In the previous chapter, we established the compensation of London dispersion

in DCM. In this chapter, we extend the solvents employed to a range of polar and nonpolar solvents with a considerable range of bulk polarizabilities (0.17–0.32) and refractive indices (1.27–1.56).[161] The rationale behind this choice of solvent is that bulk polarizability and refractive index are measures of the total dispersive ability of the solvent surrounding a solute.[72,161] All the solvents employed in this chapter are illustrated in Scheme 3.3 together with their corresponding abbreviations that we will use to refer to them. In order to achieve reasonable solubilities of the salts of our model compounds in 1-HPFX as fluorous solvent, we also synthesized the $BArF_{f6}^-$ salts[162–164] of all our protonated nitrogen bases as addition to the library of $BArF_{24}^-$ salts established in the previous chapter. We determined the Gibbs free energies of bond dissociation of the new model compounds in DCM and of all our model compounds in PhBr and 1-HPFX by VT-NMR experiments. In addition, we also determined the Gibbs free energies of bond dissociation for part of the model compounds in 12-DFB, 1-FHX and CyBr by VT-NMR and we extended the subset of compounds investigated by ITC in DCM. Notably, when we started this project, we thought that the solvents with low bulk polarizability have weak dispersive ability and the solvents with high bulk polarizability have strong dispersive ability.



Scheme 3.3: Counterions and solvents used in this chapter.

3.2.2. Gas-phase Compensation of Dispersion

To understand the total compensation of London dispersion in solution it is necessary to understand how much is already compensated in the gas phase. In the previous chapter, we attempted to estimate the compensation of dispersion already. However, we did not account properly for the presence of conformational ensembles and, hence, could not provide reliable estimates of the compensation of dispersion as a function of temperature. We preferred to carry out computational studies to tackle the compensation in the gas phase as experimentally it would require measuring the temperature dependence of gas-phase equilibrium constants for our model compounds, which would be very tedious to perform and would likely require too high temperatures for most of our compounds to still be stable.[99,165] As in the previous chapter, we calibrated the calculations against experiment. To carry out the estimations of compensation in the gas phase, we employed the recently developed GFN-xTB methods, which are semi-empirical calculations in the theoretical framework of DFT with approximations analogous to Hückel theory.[166,167] The advantage of the GFN-xTB methods over the DFT approach chosen in the previous chapter is that extensive conformer explorations can be carried out in order to generate a representative set of all the low energy conformations of each compound. In addition, BDEs from both GFN1-xTB[166] and GFN2-xTB[167] agreed very well both with M06-L/CBS BDEs and with the experimental BDEs. The comparison of ZPE-corrected BDEs from the GFN-xTB methods with M06-L/CBS and the comparison of entropies of reaction from the GFN-xTB methods with B97-D3/pc-2-sp(d) are illustrated in Figure 3.1. These results show that computations with the GFN-xTB methods agree closely with the computational approach established in the previous chapter. Based on these findings, we used GFN-xTB conformational ensembles to correct gas-phase entropies from B97-D3/pc-2-sp(d) directly by applying a correction for the presence of multiple conformers without reoptimizing the corresponding conformers at the B97-D3/pc-2-sp(d) level of theory. Notably, this is consistent with what we said about gas-phase entropy before.

After generating extensive conformational ensembles of all our model compounds, we performed Boltzmann averaging to obtain enthalpies, entropies and Gibbs free energies of dissociation as a function of temperature for the complete

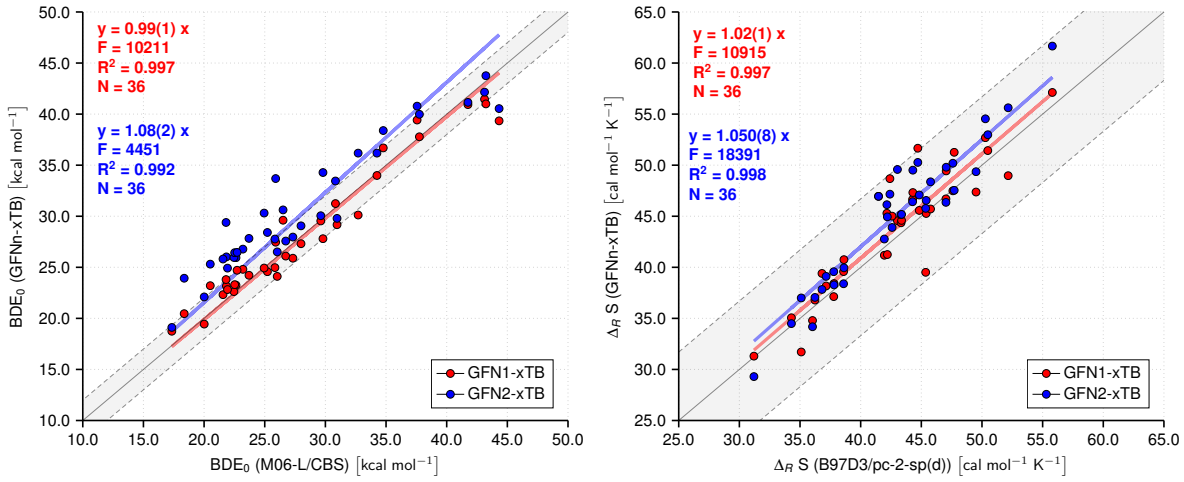


Figure 3.1.: Left: Benchmark of ZPE-corrected BDEs from GFN-xTB methods with M06-L/CBS. Right: Benchmark of entropies of reaction from GFN-xTB methods with B97-D3/pc-2-sp(d).

set of test systems. With this data in hand, we could derive also the temperature-dependent compensation of dispersion in the gas phase by linear correlation between the computed Gibbs free energies against the corresponding BDEs. The respective results for both GFN1-xTB and GFN2-xTb at 298 K are depicted in Figure 3.2.

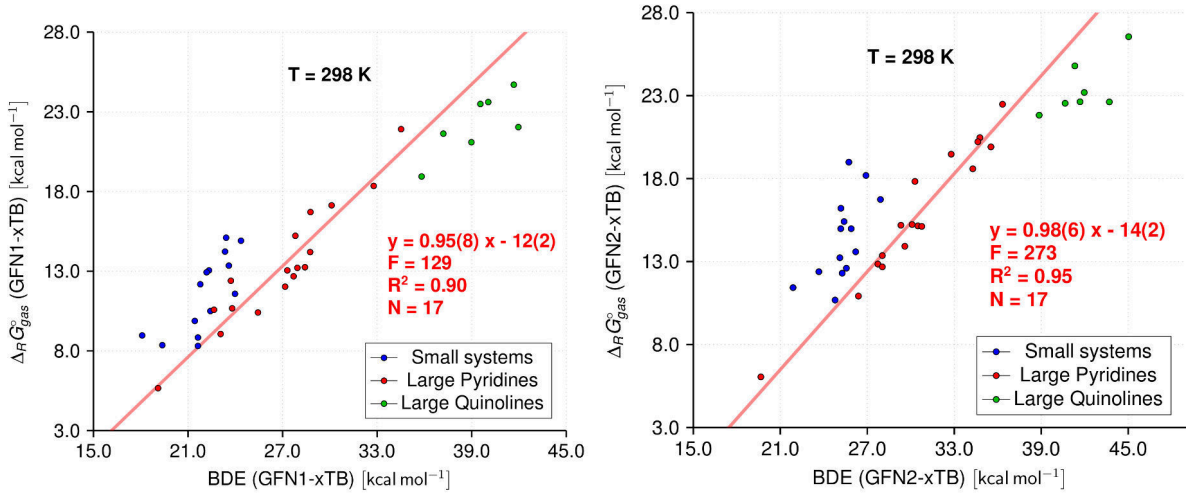


Figure 3.2.: Compensation of dispersion in the gas phase by correlation of gas-phase Gibbs free energies of dissociation against the respective BDEs using GFN1-xTB (left) and GFN2-xTB (right), respectively.

First, it is important to note that the classification of the data points of our model compounds, that we observed in our experimental numbers in DCM, can be reproduced also in the gas phase using the GFN-xTB methods. Secondly, the data of the Gibbs free energies of the large pyridines, unlike the data points of either the large quinolines or the small systems, correlates significantly with the corresponding gas-phase BDEs. Based on these correlation, the compensation of dispersion can be gauged from the respective slopes. Hence, both GFN-xTB methods estimate compensation in the gas-phase to be only a few percent of the total BDE with the caveat that the estimates are statistically not significantly different from a compensation of zero percent. Relative to the position of the large pyridines, the small systems are more endergonic and the large quinolines are more exergonic. Hence, the large quinolines show a significantly higher compensation of dispersion compared to the large pyridines, and this compensation can be estimated to be approximately 40 %. The reason for the significantly higher compensation in the large quinolines lies in the more positive entropies of reaction of proton-bound dimer dissociation. Overall, these results indicate that compensation of London dispersion is small for rigid compounds but can be very significant for more flexible compounds. Hence, compensation of London dispersion in the gas phase is highly dependent on molecular structure and in general smaller for more rigid systems.

3.2.3. Compensation of Dispersion in Solution

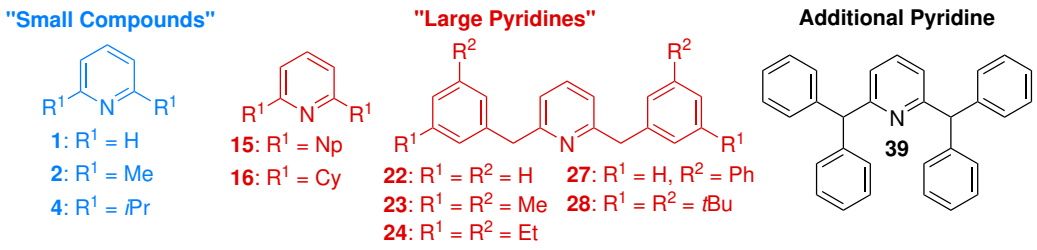
First, we studied the importance of ion pairing in DCM, PhBr and 1-HPFX using DOSY NMR measurements and found the propensity for ion pairing to increase with a decreasing solvent dielectric constant. The corresponding results are shown in Table 3.1. In addition to the dependence of solvent dielectric constant, the data also reveals a dependence on cation size, namely that smaller cations tend to favor ion pairing.

To study the dependence on cation size more closely, we needed a solvent that showed full ion pairing only for the smallest cation, and an intermediate extent of ion pairing for larger cations. We found $CDCl_3$ to be a good solvent for that purpose as it caused the pyridinium salt of unsubstituted pyridine to be completely ion paired and showed significant, but not full, ion pairing for pyridinium salts of

	Extent of Ion Pairing as a Function of Solvent			
Salt	DCM (ε = 8.9[168])	PhBr (ε = 5.4[169])	$CDCl_3$ (ε = 4.8[168])	1-HPFX (ε = 3.5)
1H^+ $BArF_{24}^-$	low	low	full	insoluble
28H^+ $BArF_{24}^-$	none	none	high	insoluble
1H^+ $BArF_{f6}^-$	n.d.	n.d.	n.d.	full
28H^+ $BArF_{f6}^-$	none	n.d.	n.d.	full

Table 3.1.: Extent of ion pairing of small (**1**H^+) and large (**28**H^+) pyridinium salts as a function of solvent; n.d. means not determined.

larger pyridines. Consequently, we selected a subset of our model compounds to investigate the dependence of the Gibbs free energy of ion pairing in $CDCl_3$ as a function of cation size. The model compounds used for that purpose are depicted in Scheme 3.4. The corresponding results are shown in the left diagram of Figure 3.3. In addition, we were also interested in the general trend of the temperature dependence of ion pairing. Hence, we determined the Gibbs free energy of ion pairing as a function of temperature for the pyridinium salt of **39**. From that data, we wanted to find out whether the Gibbs free energy of ion pairing increases or decreases with temperature and hence, get a rough estimate of the corresponding enthalpies and entropies of ion pairing. The respective results are depicted in the right diagram of Figure 3.3.



Scheme 3.4: Model compounds to study size dependence of ion pairing.

Figure 3.3 demonstrates that ion pairing correlates linearly with the hydrodynamic radius of the cation and confirms that the extent of ion pairing tends to decrease with increasing size. In addition, it also shows that the extent of ion pairing tends to decrease with increasing temperature for our model compounds in

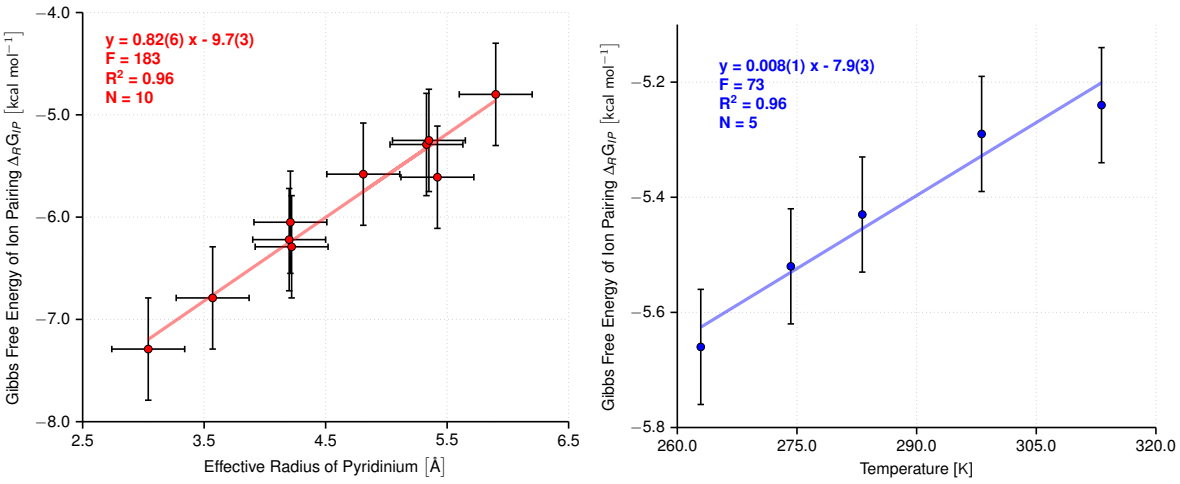


Figure 3.3.: Left: Gibbs free energies of ion pairing of $BArF_{24}^-$ salts in $CDCl_3$ plotted against the effective radius of the pyridinium ion in solution. Right: Gibbs free energies of ion pairing of **39**H^+ $BArF_{24}^-$ in $CDCl_3$ as a function of temperature.

$CDCl_3$ and that the corresponding Gibbs free energy of ion pairing is enthalpically driven. From these findings, we concluded that ion pairing in our test systems is controlled by electrostatics, which is an enthalpic component.[108,170,171] Since we observed full ion pairing in 1-HPFX for both small and big cations, we needed to correct our experimental results in 1-HPFX to remove a potential differential ion pairing effect. The advantage is that knowing that ion pairing in our systems is governed by electrostatics allows us to estimate this differential effect using the Eigen-Fuoss equation.[172,173] To do that, we estimated the hydrodynamic radii of all the protonated model compounds and their proton-bound dimers from COSMO computations because the corresponding estimates showed excellent agreement with the hydrodynamic radii obtained experimentally for the subset of model compounds studied by DOSY. The Gibbs free energy corrections for all the model compounds merely spanned 0.26 kcal/mol indicating that differential effects are very small and close to the experimental uncertainties. Hence, any differential effect of ion pairing is minimal and will not affect our results to a significant extent, which is due to the large anion we chose.

It is well established that enthalpies and entropies of reaction obtained from the temperature dependence of equilibrium constants are typically subject to large

correlated experimental errors resulting in poor estimates for enthalpies and entropies of reaction.[136–138] To obtain reliable estimates of enthalpy and entropy of proton-bound dimer dissociation in solution, we carried out ITC measurements in DCM. These measurements were performed on a subset containing 16 of the model compounds and served as a benchmark for the enthalpies and entropies of dissociation obtained from VT-NMR studies. The corresponding subset was selected to have all the subgroups of model compounds reasonably covered. In the previous chapter, only four control measurements were carried out but already there we observed that while Gibbs free energies agreed well between VT-NMR and ITC, both entropy and enthalpy of reaction did not. With a more comprehensive data set of ITC values in hand, we could confirm the initial findings of the previous chapter. These results confirm that the Gibbs free energies as a function of temperature from our VT-NMR results can be safely interpreted but the enthalpies and entropies of reaction cannot. Furthermore, from the ITC results we obtained interpretable enthalpies and entropies of proton-bound dimer dissociation in DCM. The corresponding results are depicted in Figure 3.4. In general, the dissociations are dominated by enthalpy and unfavorable for all but one compound by entropy. Notably, at least between 5 °C–25 °C, the enthalpies and entropies of reaction were found to be independent of temperature. However, comparing the temperature-dependence of the Gibbs free energy from VT-NMR with the corresponding results from ITC suggests that both enthalpy and entropy of reaction are not independent of temperature over the full temperature range studied by VT-NMR. Additionally, this is also one of the reasons the enthalpies and entropies from VT-NMR and ITC do not agree with each other.

We also wanted to investigate the presence of an EEC[109] in our VT-NMR data more closely. Hence, we combined all the experimental VT-NMR data in a single compensation plot, which is depicted in Figure 3.5. In addition, we also depicted the experimentally accessible data range as grey area, which shows that a wide area of the measurable range of data points is well covered. This suggests that the weak EEC observed is actually a spurious correlation rather than a genuine physical effect. Independent results for that observation comes from the poor agreement between ITC and VT-NMR results for enthalpies and entropies of reaction (*vide supra*).

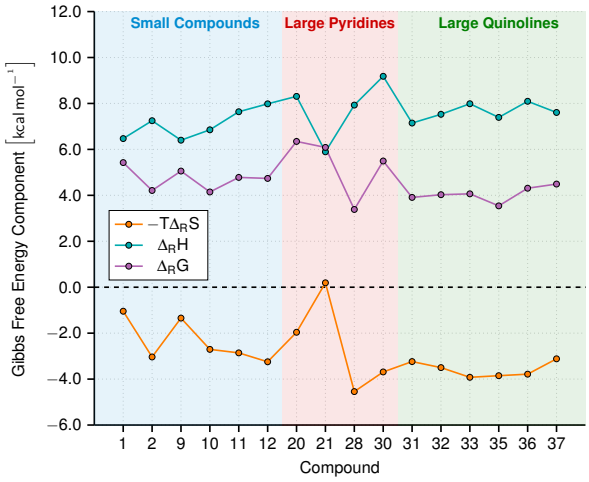


Figure 3.4.: Enthalpy, entropy and Gibbs free energy of dissociation of proton-bound dimers in DCM at 5 °C–25 °C from ITC measurements.

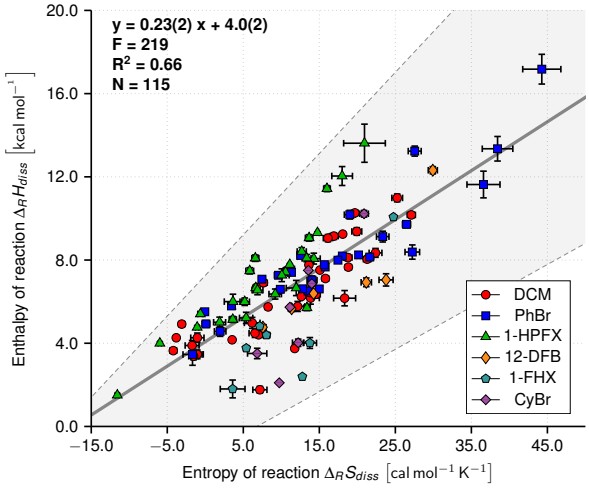


Figure 3.5.: Compensation plots of proton-bound dimer dissociations in various solvents. The experimentally accessible value range is shown as grey area.

Next, we moved on to determine the total compensation of dispersion in all the employed solvents. First, we compared the correlation of Gibbs free energies of dissociation against corresponding gas-phase BDEs for DCM, PhBr and 1-HPFX in Figures 3.6-3.8.

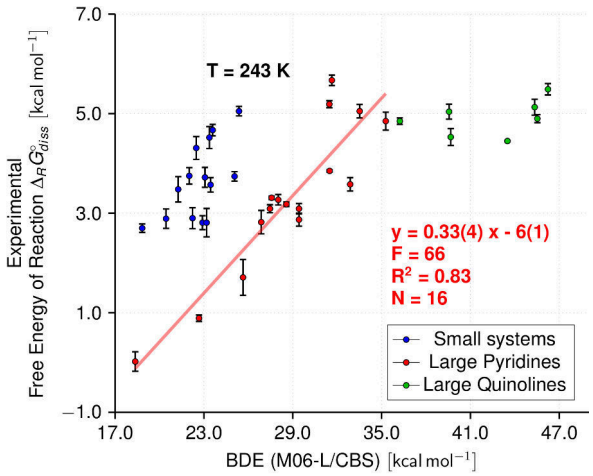


Figure 3.6.: Gibbs free energy of dissociation of proton-bound dimers in DCM at 243 K correlated against gas-phase BDE.

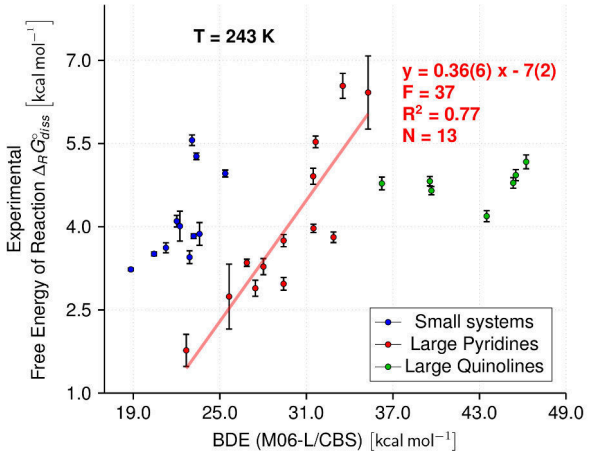


Figure 3.7.: Gibbs free energy of dissociation of proton-bound dimers in DCM at 243 K correlated against gas-phase BDE.

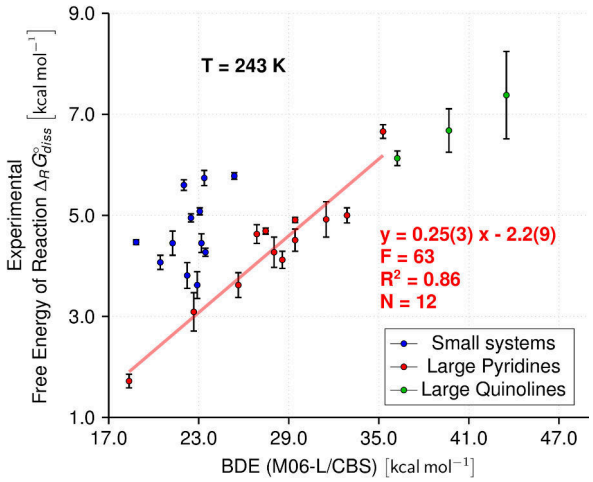


Figure 3.8.: Gibbs free energy of dissociation of proton-bound dimers in DCM at 243 K correlated against gas-phase BDE.

It should be emphasized that we observed the same type of grouping of data points in PhBr and 1-HPFX that we already observed in DCM in the previous chapter and in the gas-phase using GFN-xTB methods before (cf. Figure 3.2). Looking into more detail into the differences between the solvents, the large quinolines are not very significant outliers with respect to the large pyridines in 1-HPFX. On the contrary, they are more severe outliers in PhBr. In addition, in all three solvents we observe significant correlations between Gibbs free energies of bond dissociation and the corresponding gas-phase BDEs. Hence, we can estimate the total compensation of dispersion in all these solvents using the same methodology established in the previous chapter.

To cover a wider range of solvents, we decided to use a selected subset of seven of the large pyridines and determined the corresponding Gibbs free energies of dissociation in 12-DFB, 1-FHX and CyBr. The subset was chosen to cover the entire Gibbs free energy range with roughly evenly spaced data points so that the total compensation of dispersion can be determined as reliably as possible from linear correlations. To have consistent data of compensation of dispersion in all the investigated solvents, we also used the same data points to repeat estimation of compensation in DCM, PhBr and 1-HPFX. The corresponding correlations for all the solvents studied are depicted in Figure 3.9. For all solvents investigated, statistically significant correlations were obtained.

From all these correlations at various temperatures we were able to determine

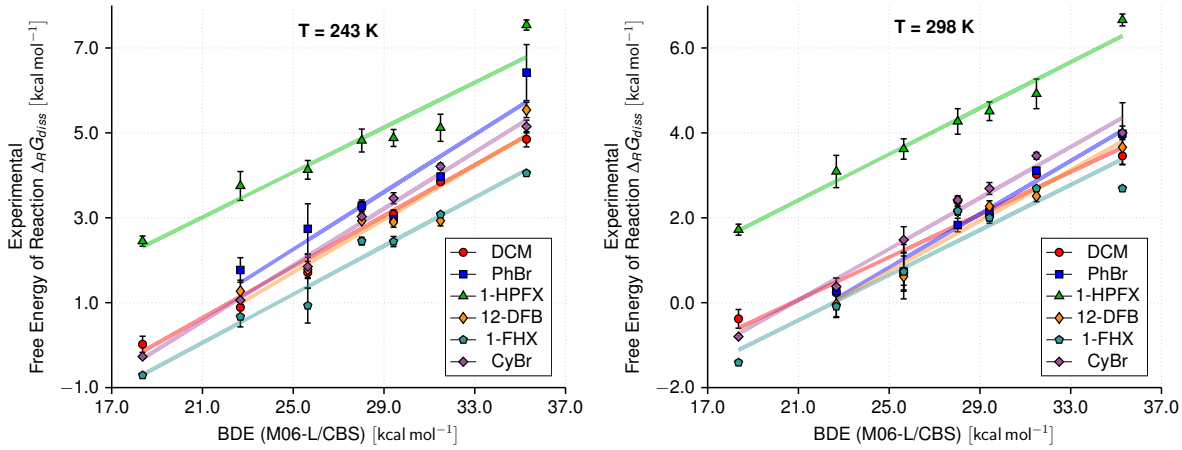


Figure 3.9.: Gibbs free energy of dissociation of a selected subset of the large pyridines in all studied solvents at 243 K (left) and at 298 K (right) correlated against gas-phase BDE.

the compensation of dispersion in all six solvents as a function of temperature. To gain insight into solvent-mediated differences, we correlated the corresponding results against dielectric permittivity, cohesive energy density,[72,129,174] refractive index, bulk polarizability[161] and Hansen dispersion solubility parameter[175] of the solvent. It should be noted that in the range of refractive indices investigated, the bulk polarizability is essentially a linear function of the refractive index and hence, the corresponding correlations are basically identical. The best correlations of compensation of dispersion against solvent parameter were obtained using refractive indices, bulk polarizabilities and Hansen dispersion solubility parameters. The corresponding correlation with respect to solvent bulk polarizability at 243 K and 298 K is shown in Figure 3.10. The quality of the correlation with respect to the Hansen dispersion solubility parameter is only marginally worse.

From Figure 3.10 it can be deduced that, at ambient temperature, compensation of dispersion is effectively invariant with respect to the solvent. All the data points obtained are within each others experimental errors, irrespective of solvent bulk polarizability. The situation changes slightly when going to lower temperatures, and, at –30 °C, it is observed that a higher solvent bulk polarizability tends to lead to lower total compensation of dispersion. Overall, the range of Gibbs free energies of dissociation can be changed by 1.4 kcal/mol at –30 °C and by 1.0 kcal/mol at room temperature. Hence, rationally tuning London dispersion by

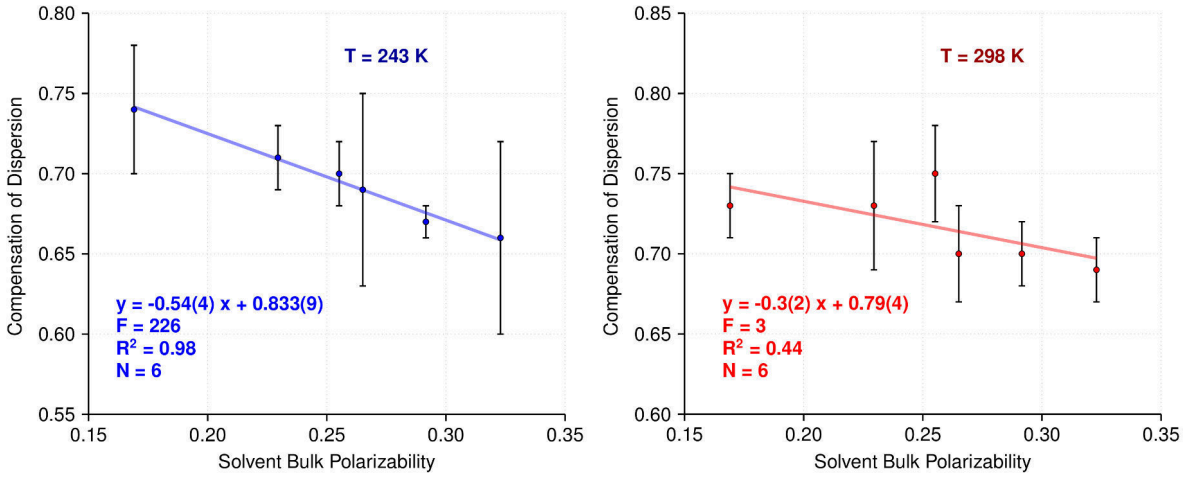


Figure 3.10.: Correlation of solvent bulk polarizability against total compensation of dispersion at 243 K (left) and at 298 K (right).

the choice of solvent is only possible to a very modest extent.

After looking into the solvent dependence of compensation of dispersion, we also looked more closely into the temperature dependences in all the solvents. Plotting the temperature-dependent total compensation of dispersion of all the solvents in one diagram revealed them to overlap significantly. This was to be expected since we just demonstrated that solvent effects are rather small. Therefore, we decided to combine all the temperature dependencies to a rough global estimation of the total compensation of London dispersion in our model compounds in solution as a function of temperature. The corresponding results are depicted in Figure 3.11. Notably, all solvents but 1-HPFX show total compensation to increase with temperature and, hence, this is also the global trend derived from all our data. The yellow area in Figure 3.11 can be used as a general approximate scheme to estimate the compensation of dispersion in aprotic solvents. We find that at 170 K–340 K compensation is estimated to be about 60 %–80 %.

3.2.4. Separation of Compensation of Dispersion

In the previous section, we established the total compensation of dispersion as a function of solvent and as a function of temperature. In this section, we want to go one step further by separating the total compensation of dispersion into a

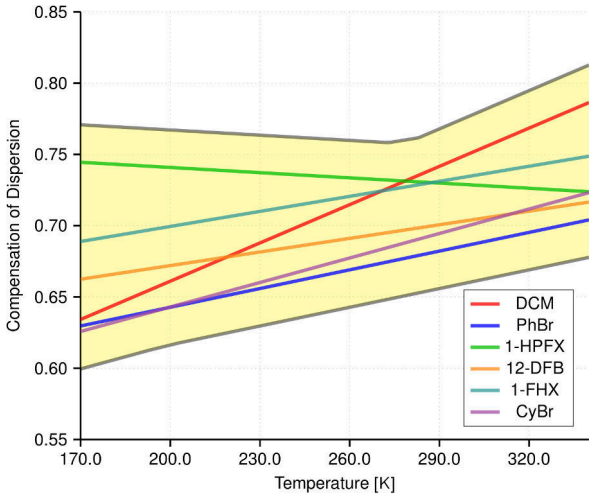


Figure 3.11.: Temperature-dependence of total compensation of dispersion for all solvents investigated condensed into a rough general estimate in aprotic solvents.

gas-phase compensation and a solvent compensation. To do that, we used the same subset of the large pyridines that we selected for the experimental determination of compensation and correlated the corresponding computed Gibbs free energies of dissociation from the GFN-xTB methods against the corresponding BDEs at the same level of theory. Doing that, we obtained the compensation of dispersion in the gas phase as a function of temperature. It should be noted, that the gas-phase compensation from the focussed subset of the large pyridines was somewhat larger than the corresponding estimate from the complete set of the large pyridines. This is a consequence of the structure-dependence of the compensation of dispersion in the gas phase. Having the compensation of dispersion in the gas phase from both GFN1-xTB and GFN2-xTB, we used the corresponding average of the results from the two methods. To extract the solvent compensation as a function of temperature from the total compensation we simply substracted the computed gas-phase compensation at every temperature. The temperature dependences of gas-phase compensation and solvent compensation are shown in Figure 3.12.

Figure 3.12 demonstrates that the total compensation of dispersion is largely governed by the solvent compensation over the full range of temperatures investigated. Only at temperatures higher than the ones experimentally covered in this study is the gas-phase compensation expected to rise to similar levels

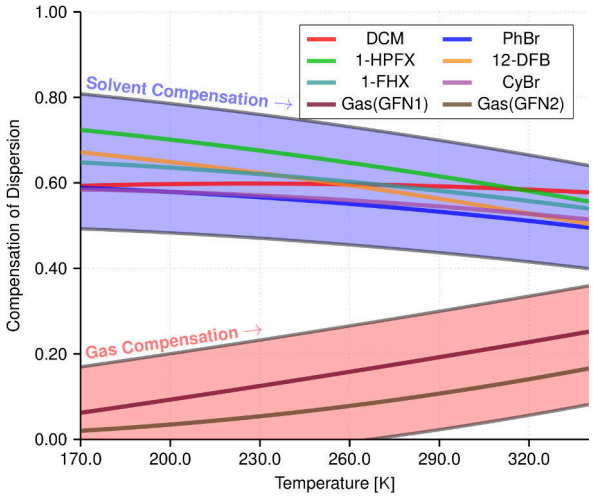


Figure 3.12.: General estimates of the temperature-dependence of gas-phase and solvent compensations of dispersion.

as the solvent compensation. It is important to note that the solvent compensation of dispersion shows a decreasing trend with temperature, while both the gas-phase compensation and the total compensation show an increasing trend with temperature. This suggests that the observed temperature trend in the total compensation originates from the corresponding trend in the gas phase. It should be noted that the gas-phase temperature dependence of dispersion is not linear but slightly curved. This is a consequence of Boltzmann averaging in the conformational ensembles over a large range of temperatures.

3.2.5. Estimation of Gibbs Free Energies in Solution

The final task in this chapter was to test the ability of computational methods to reproduce our Gibbs free energies of dissociation in various solvents. The important addition compared to the previous chapter was that having data for a range of solvents allows to assess the problems of the implicit solvent models more systematically. The computational approach established in the previous chapter was refined by including a correction for the presence of multiple low energy conformers for all the species involved in the bond dissociation equilibria. This correction was determined using GFN-xTB methods. The corresponding refined approach to compute Gibbs free energies of dissociation in solution is delineated in Table 3.2. It improves upon the approach chosen in the previous

chapter by improving the estimates of gas-phase entropies of dissociation.

$$\Delta_R G^\circ_{diss}(sol) = \Delta_R E + \Delta_R ZPE + \Delta_R H^\circ_{corr} - 298.15\ \Delta_R S^\circ + \Delta_R G^\circ_{conf} + \Delta_R G_{sol}$$

$\Delta_R E$	Electronic energy difference at M06-L/CBS level of theory using B97-D3/pc-2-sp(d) geometries.
$\Delta_R ZPE$	Difference in ZPE-corrections at the B97-D3/pc-2-sp(d) level of theory assuming harmonic oscillations.
$\Delta_R H^\circ_{corr}$	Difference in enthalpy corrections at the B97-D3/pc-2-sp(d) level of theory assuming harmonic oscillations.
$\Delta_R S^\circ$	Difference in entropy corrections at the B97-D3/pc-2-sp(d) level of theory assuming ideal gas heavier using the Q-RRHO approach.
$\Delta_R G^\circ_{conf}$	Gibbs free energy correction (in both enthalpy and entropy) at the GFNn-xTB (n = 1, 2) level of theory for the presence of multiple low-energy conformers.
$\Delta_R G_{sol}$	Gibbs free energy of solvation at the SMD (M05-2X/6-31g*) or COSMO-RS (CRS) (BP86/TZP) level of theories.

Table 3.2.: Computational approach to estimate $\Delta_R G^\circ_{diss}$ in solution at 298 K.

Using this refined approach, we estimated the Gibbs free energies of dissociation for all our model compounds with SMD and COSMO-RS as implicit solvent models. It should be noted that SMD could only be used for DCM and PhBr because parameters for the other solvents were not available. COSMO-RS could be used for all the solvents studied. We compared the computed estimates against all our experimental results and the corresponding comparison is depicted in Figure 3.13. It should be noted that in this comparison we used GFN1-xTB for the multiple conformer correction but the results using the GFN2-xTB corrections instead do not differ considerably. As can be seen in Figure 3.13, the large deviations between computed and experimental values are not limited to one particular solvent but are general for all the solvents investigated.

We also wanted to gain insight into the origin of the large deviations between computations and experiment observed. In the previous chapter we established that the weakest link in the computational protocol is the implicit solvent model, which underestimates the compensation of dispersion. However, it was not clear how much compensation of dispersion is underestimated. Hence, in this chapter, we wanted to quantify how much the total compensation of dispersion is underes-

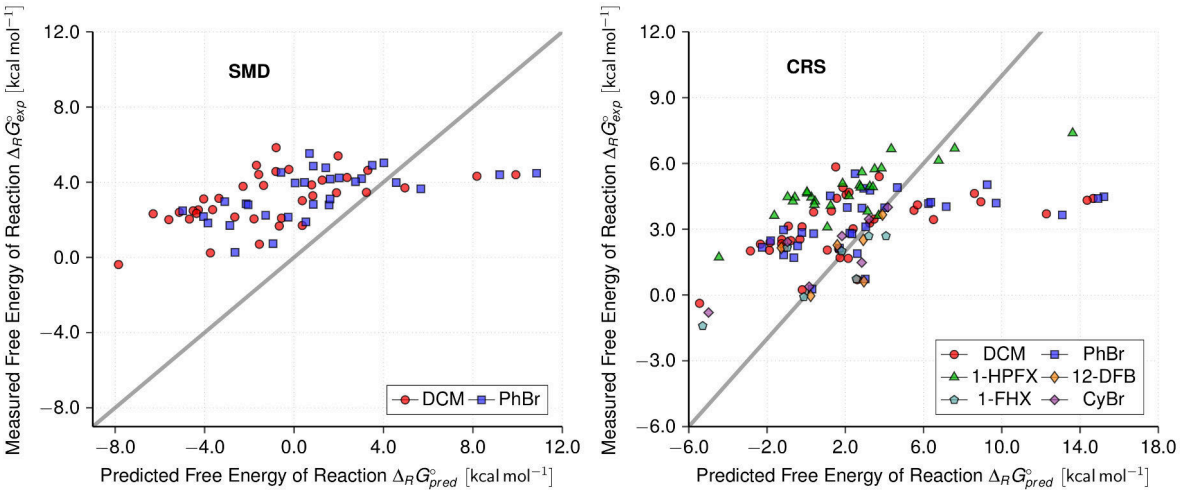


Figure 3.13.: Experimental Gibbs free energies of dissociation of proton-bound dimers in various solvents at 298 K plotted against the corresponding computational estimates using SMD (left) and COSMO-RS (right) as implicit solvent model. In both plots the multiple conformer corrections from GFN1-xTB are employed.

timated with our computational approach. Hence, we used the selected subset of the large pyridines to estimate the corresponding compensation of dispersion from the computed Gibbs free energies of dissociation. To obtain estimates of total compensation, we evaluated the corresponding computational results in the same way as the experimental numbers. For comparison to experiment, we determined the average of the total compensation of dispersion over all solvents both from the experimental and the computational numbers. In addition, we tested both SMD and COSMO-RS with and without the correction for the presence of multiple conformers to show the corresponding effects. The results are illustrated in Figure 3.14.

Figure 3.14 confirms that the computational approach underestimates the experimentally found total compensation of dispersion significantly. We also find that COSMO-RS provides somewhat better estimates of compensation compared to SMD. Furthermore, we find that the correction for the presence of multiple conformers improves the computational estimates of total compensation of dispersion.

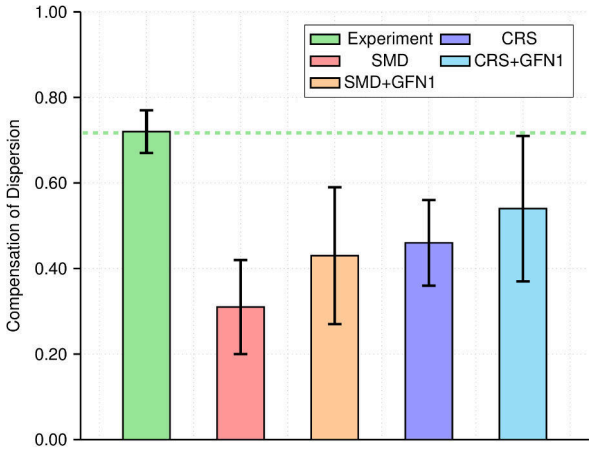


Figure 3.14.: Total compensation of dispersion from experiment and computation compared.

3.3. Discussion

Summarizing our findings results in the three main conclusions of this chapter:

(i) Gas-phase compensation of dispersion depends significantly on molecular structure and can be 0 % in rigid compounds and can reach up to 40 % in flexible molecules at ambient temperatures.

(ii) Total compensation of London dispersion is largely invariant in both polar and nonpolar aprotic solvents with a wide range of bulk polarizabilities. At 170 K–340 K it is approximately 60 %–80 %.

(iii) Continuum solvation approaches result in significant underestimation of the total compensation of dispersion, which in turn leads to large systematic discrepancies between computed Gibbs free energies of dissociation in solution and experimental values.

In this section these three claims will be investigated separately in detail.

3.3.1. Gas-phase Compensation of Dispersion

London dispersion is a universally attractive interatomic force and is present in all molecules.[3,16,18,176] At intermediate interatomic distances of 3 Å–5 Å it dominates over Pauli repulsion and considerably contributes to the overall interaction

energy. However, in the gas phase, London dispersion is partially compensated by a similarly omnipresent energy contribution: entropy.[3,177] They both generally show an increasing trend with number of atoms. Considering the dissociation of an intermolecular complex of increasing size into two molecules, which also increase in size accordingly, the entropy of dissociation tends to increase and becomes more favorable. In addition, increasing the size of the two fragments also increases the inter-fragment dispersion, which makes bond dissociation less favorable. Overall, the gain in intermolecular interaction energy caused by dispersion is partially compensated by the increased entropic cost. Additional compensatory contributions originate in ZPEs of intermolecular interactions, which tend to increase with interaction energy, and enthalpy corrections. All these energy contributions to the total Gibbs free energy are well known and well defined. However, compensation of dispersion in the gas phase has not been given much attention.[177] Hence, we were interested to carry out a detailed quantitative study of the compensation of dispersion in the gas phase considering all these effects by correlating the estimated Gibbs free energies of dissociation in the gas phase against the corresponding non-ZPE corrected BDEs. Notably, the negative contribution to the total Gibbs free energy, which comes from the entropic gain of dissociation of one molecule into two, is not included in this compensation. This free energy component is about 9 kcal/mol for unsubstituted pyridine and is roughly constant for all dissociations studied. Hence, we are only estimating the compensation of dispersion that comes about with the substituents in the proton-bound dimers.

Our methods of choice for estimation of total compensation of dispersion in the gas phase were the GFN-xTB methods.[166,167,178] The acronym GFN is an abbreviation for ”Geometries, Vibrational Frequencies and Noncovalent Interactions,”[166] and this name specifies the properties these methods were developed for. Their main application is to describe molecular structures, noncovalent interactions and thermochemical properties at a good cost-accuracy ratio, which can be readily applied to molecular systems with even more than 1000 atoms.[166,167] For our test systems, the important difference between GFN1-xTB and GFN2-xTB is the account of multipole electrostatics in GFN2-xTB. GFN1-xTB only uses partial charges and pairwise interactions between them. The consequence of

only including monopole electrostatics is that the description of the H-bond geometry is rather poor in GFN1-xTB because it leads to largely symmetric proton-bound dimers with somewhat shorter N-H bond distances. This problem is alleviated in GFN2-xTB, which prefers the N-H–N geometry to be asymmetric, which is in good agreement to more sophisticated computations[179] and to experimental observations.[87] Despite of this intrinsic problem of GFN1-xTB, when applied to our model compounds, both GFN1-xTB and GFN2-xTB reproduce the results of more sophisticated dispersion-corrected DFT methods very well, at least the corresponding BDEs. The corresponding geometries, interaction energies and thermodynamic corrections are all very similar to higher level of theory. Not only do the GFN-xTB results for our model compounds agree well with other computational methods, but they also show close agreement with our experimental benchmark results, which we discussed in great detail in the previous chapter. This excellent agreement gave us sufficient confidence in the method to use it directly for the corrections for the presence of multiple conformers, even on top of single point energies and thermodynamic corrections from different computational methods. Notably, there is a systematic offset between the compensation of dispersion from GFN-xTB methods and from M06-L/CBS. However, this offset is not statistically very significant as they are typically within the statistical errors of the estimates. In addition, also the results from both GFN1-xTB and GFN2-xTB are not statistically very different. Therefore, we decided to use the average of the compensation estimates of both values for the final results.

One of the recurring observations for the compensation of dispersion in the gas phase was that it is quite significantly structure-dependent. At room temperature, it can be as high as 40 % and as low as essentially 0 %. The reason for that is that rigid substituents in noncovalent complexes do not experience a significant restriction in their conformational space when bound to another molecule and, hence, the corresponding entropies of dissociation will be less favorable compared to noncovalent complexes with more flexible substituents. The same intrinsically lower number of low energy conformers manifests itself also in the isolated fragments which also contributes to a less favorable entropy of dissociation. Hence, based on this considerations, we expect compensation of London dispersion to be highly structure-dependent in general.

Furthermore, it is also notable that we observe the same grouping of our model compounds in our GFN-xTB calculations, that we also observe in solutions in various solvents. On the one hand, this gives some credence to the computational approach. On the other hand, this provides insight into the reason why the persistent grouping of the data points occurs. It suggests that the structural differences between the groups of molecules is manifested in the entropy of dissociation, which in turn influences the Gibbs free energy of dissociation.

3.3.2. Compensation of Dispersion in Solution

Before going into more detail into the compensation of dispersion itself, we need to discuss the results of our control experiments. We will start with the investigations regarding ion pairing of our pyridinium salts. A variety of experimental methods have been used to study ion pairing in solution, especially the structure and the extent of ion pairs.[108,170] Currently, DOSY-NMR is one of the most attractive methods for the study of ion pairs both qualitatively and quantitatively in a wide variety of structures.[180,181] Compared to most other spectroscopic techniques, the advantage of DOSY-NMR is that it directly probes the dynamic behavior molecules in solution rather than relying on the occurrence of new peaks or the shift of peaks in a spectrum. A recent DOSY-NMR study of ion pairing in cationic chromium(VI) complexes varying both polar aprotic solvents and counterions found that so-called non-coordinating counterions do not have a particularly low tendency to form ion pairs at all but they do not have large preference for a specific conformer of the ion pair.[107,182] Hence, they do not complex to a specific site of the cation but complex with a large number of possible coordination sites to a similar extent.[107] This property of non-coordinating counterions makes the corresponding ion pairing harder to study because, typically, no significant peak shifts are observed and no new peaks occur in the corresponding spectra. In addition, this property of non-coordinating ions is typically not considered. Therefore, non-coordinating ions like $BArF_{24}^-$ still form ion pairs in non-polar solvents but they do not coordinate to specific site of the cation.

We could confirm that finding also for our model compounds and found ion pairing to be very significant for $BArF_{24}^-$ and $BArF_{f6}^-$ salts both in 1-HPFX and chloroform. Based on numerous experiments varying the size of the cation, the

solvent and the temperature, we found that in our systems ion pairing is governed by electrostatic attraction between cation and anion. In addition, these results demonstrate that ion pairing can hardly be avoided in very nonpolar solvents like 1-HPFX. The only viable option to decrease the propensity for ion pairing even further is to make the non-coordinating anions even larger, because the concomitant increase in interaction distance between cation and anion decreases the electrostatic attraction between them. However, since not even $BArF_{f6}^-$ is sufficiently large to disfavor ion pairing enough to avoid it altogether, it seems that this handle is not very effective. Therefore, we looked closely into the differential effects of ion pairing to the Gibbs free energy of dissociation of proton-bound dimers and found that they are minimized for very large counterions. In addition, because ion pairing in our systems is dominated by electrostatics, the upper boundary of the differential effect could be estimated using the Eigen-Fuoss equation.[172,173] The Eigen-Fuoss equation provides upper boundaries of the differential effect because it is derived based on the assumption that the ion pairs are so-called contact ion pairs, i.e. ion pairs at the closest possible distance.[108] The corresponding estimates for the differential effect were negligible for all our model compounds. Hence, we decided not to add them on top of our experimental numbers. The main reason for not adding these corrections was to avoid overestimating the effects of ion pairing. The bottom line of our DOSY-NMR control experiments is that any significant differential effects of ion pairing to the Gibbs free energies of dissociation of our proton-bound dimers in solution can be excluded with confidence.

In addition, we want to address again the presence of a weak EEC, which still remains even when data from all the solvents investigated are combined. As discussed also in the previous chapter, EEC denotes a statistical significant correlation between entropies and enthalpies of reaction in a set of thermodynamic data of a specific process with small variations.[109] EECs were observed in many different fields[9,183–191] and various attempts to explain them have been made with popular explanations being ubiquitous redistribution of subspecies and solvent reorganization, and correlated experimental errors in both enthalpies and entropies causing a spurious correlation.[131,133–135,192–194] A number of statistical procedures to test for the presence of spurious correlations in thermodynamic

data have been published.[109–113,195] Alternatives to test for the presence of spurious correlations is to check whether the enthalpies and entropies, that show a correlation with each other, can be independently measured by different experimental methods to confirm their values. Hence, enthalpies and entropies of reaction originating from the temperature-dependence of equilibrium constants can be compared to enthalpies and entropies of reaction determined by calorimetric methods like ITC experiments.[109,185,188,190] In case the thermodynamic parameters cannot be confirmed by independent experimental measurements, they are likely not good estimates of the actual values. It is important to note in that regard, that enthalpies of reaction obtained from ITC experiments are expected to be more robust because the heat of reaction is directly measured rather than inferred like in temperature-dependent equilibrium studies.[109,134] However, the errors of enthalpies and entropies of reaction are also significantly correlated from ITC experiments and, accordingly, the entropies of reaction are similarly unreliable from either ITC or temperature-dependent equilibrium studies and spurious EECs can also be observed in ITC data sets.[134] We observed in our proton-bound dimer dissociations in solution that enthalpies and entropies of reaction show significant discrepancies between our VT-NMR measurements and ITC measurements and, hence, we concluded that the EEC we observe in the VT-NMR data is spurious. In addition, we found that the data points determined coincide with the experimentally accessible data range. Hence, we propose that the EEC in our data set is merely a demonstration of the experimentally measurable subset of thermodynamic parameters.[109] It is very important to point out again that all the models trying to explain EECs are consistent in that the correlations do not affect the Gibbs free energies in any way and this is sometimes overlooked in literature studies of EECs. The take-home message of these studies is that the Gibbs free energies of dissociation can be safely interpreted.

In this chapter, we expanded the work of the previous one from measurements in DCM to various aprotic solvents, which cover a significant range of bulk polarizabilities. First, we will look into the different Gibbs free energies of dissociation that come about by different solvents. One of the most notable differences is that all the Gibbs free energies of dissociation in 1-HPFX are consistently higher. Two characteristics of 1-HPFX can explain these qualitatively observed differences.

Of all the solvents employed, 1-HPFX is the weakest dielectric and hence, inter-fragment electrostatics are screened the least, leading to stronger H-bonds in 1-HPFX compared to all the other solvents. Based on that explanation, one would expect an offset between the Gibbs free energies in 1-HPFX and in all the other solvents. However, the offset observed seems rather large given the relatively moderate difference in dielectric permittivity between 1-HPFX ($\leq$ 3.5 at 298 K) and the other solvents (4–13 at 298 K). Since the other solvents all have Gibbs free energies within a relatively narrow margin, dielectric permittivity cannot explain the complete difference observed. Additionally, the entropies of reaction obtained from VT-NMR are consistently closer to zero in 1-HPFX compared to all the other solvents. While the absolute values are likely not very accurate, as discussed in great detail in the previous paragraph, the global trends in all the values of one solvent likely are a qualitative manifestation of true physical differences. This suggests that the entropies of dissociation in 1-HPFX are systematically lower compared to the other solvents. This is also reflected in our observation that the offset in Gibbs free energies of dissociation between 1-HPFX and all the other solvents is smaller at low temperatures and more pronounced at higher temperatures. Perfluoroalkanes are very rigid and, hence, 1-HPFX is also expected to be quite rigid.[196,197] This rigidity in the solvent molecules can affect also the flexibility of the solutes so that the solutes have a reduced conformational space compared to other solvents. This reduced conformational space would directly result in a reduced entropy of the dissociated molecules and would explain the experimentally observed reduced entropies of dissociation in 1-HPFX.

Another aspect that has to be discussed is the possible impact of temperature-dependent dielectric permittivities of the solvents on the temperature dependence of the Gibbs free energies of bond dissociation. The temperature-dependence of dielectric permittivities of solvents is well-known. Typically, dielectric permittivities increase with decreasing temperature and, typically, rational functions are used to model the corresponding dependence.[198–201] Given that we carried out our VT-NMR experiments over a very wide range of temperatures for all investigated solvents, the dielectric permittivity of the solvents varied accordingly. The temperature dependence of the dielectric permittivity will induce a contribution to the temperature-dependence of the Gibbs free energy of

dissociation as it will affect the importance of electrostatic interactions. Hence, this additional contribution to the temperature-dependence of the Gibbs free energy of dissociation could contribute to the observed discrepancies between enthalpies and entropies of dissociation from VT-NMR and ITC because the latter measurements were only carried out in a very narrow temperature range. However, this contribution does not affect the estimated compensation of dispersion as a function of temperature, because the compensation is derived from correlations of the Gibbs free energies of dissociation against BDEs at every temperature separately. At every temperature, the dielectric permittivity of the solvent is constant and equal for all the model compounds studied and hence, this contribution will cancel in the estimation of the compensation.

Looking into more detail into the solvent effects to the three subsets of the model compounds, especially the large quinolines behaved quite differently in DCM, PhBr and 1-HPFX. Compared to DCM and relative to the small systems and the large pyridines, the Gibbs free energies of dissociation were consistently lower in PhBr. Yet, in both DCM and PhBr the Gibbs free energies of dissociation of the large quinolines were essentially independent of the gas-phase BDE. In contrast, in 1-HPFX the large quinolines showed very comparable trends as the large pyridines. Hence, the data of the large quinolines showed a significant solvent dependence. It can be explained by a suggestion that we already mentioned in the previous chapter. The large quinolines have a considerable amount of additional space within the proton-bound dimers. Hence, this space could at least be partially occupied by additional solvent molecules. This would directly result in an increased extent of compensation. Accordingly, one would expect it to depend on the solvent used whether solvent molecules can enter the inside of a proton-bound dimer or not. From the solvents used, 1-HPFX occupies the largest volume and hence, it is likely too large to enter the inside of the proton-bound dimers of large quinolines which would explain why the large quinolines are not significantly different in 1-HPFX from the large pyridines.

Going back to the discussion about dispersion compensation, solvation is one of the most important effects accounting for a large portion of the total compensation of dispersion.[3] As discussed in great detail in the previous chapter, the intramolecular interaction surface freed upon dissociation of an intermolecular com-

plex is occupied by surrounding solvent molecules which provides new interactions that reduce the total amount of interaction energy lost.[12–15,72,73] Increasing the size of the interacting fragments not only increases the total inter-fragment interaction energy lost upon dissociation but also the solute-solvent dispersion gained in the process. Hence, both energy components increase with the number of atoms. Hence, in absolute terms, differential solvation and intramolecular dispersion both grow with molecular size. Effectively, solvation compensates part of the interaction energy introduced by dispersion into the molecule. From a qualitative point of view, this is well understood but there are not any quantitative data available in that respect.

In the previous chapter we presented a first consistent data set providing numerical estimates of the total compensation of dispersion in DCM solution. In this chapter, we extended the corresponding data to series of aprotic solvents. In general, we observe that solvent-mediated differences in the compensation of dispersion are minimal. We find that at 243 K the compensation of dispersion correlates with the solvent bulk polarizability and the refractive index. However, the corresponding differences only amount to about 1.4 kcal/mol showing that the differences brought about by the solvent are small. Notably, bulk polarizability has been shown to describe the polarizability experienced by a solute in the solvation shell.[161,202] The polarizability experienced by a solute is also indicative of solute-solvent dispersive interactions, which can be deduced from the London dispersion formula.[2] In addition, we also correlated the total compensation of London dispersion against the solvent dielectric permittivity, the solvent cohesive energy density[72,129,174] and the Hansen dispersion solubility parameter.[175] While, statistically, the most significant correlation was obtained with bulk polarizability or with refractive index, a similarly significant correlation was obtained with the Hansen dispersion solubility parameter. Given that we also observed that the Hansen dispersion solubility parameter correlated significantly with the refractive index of the solvent, this was not overly surprising but it provides independent support that we are really probing the dispersive ability of the solvent with these parameters. Notably, while solvent bulk polarizability was directly derived from solvent refractive index, the Hansen dispersion solubility parameters are derived using a completely independent procedure.

Notably, a correlation of Gibbs free energies of binding with solvent bulk polarizabilities has been observed before. Cockroft and coworkers used a Wilcox-type[203] torsion balance to study the effect of solvent on intramolecular interactions between alkyl groups and between perfluoroalkyl groups.[72] They established that the interaction between alkyl groups and between perfluoroalkyl groups can be correlated acceptably to both solvent cohesive energy density and solvent bulk polarizability using multiple linear regression. However, cohesive energy density was by far the major descriptor in that correlation. Notably, they unveiled that the contribution of the perfluoroalkyl groups becomes more attractive with increasing bulk polarizability, while the opposite is observed for alkyl groups.

In our model compounds, the total compensation of dispersion showed a decreasing trend with increasing bulk polarizability of the solvent. That means in turn that inter-fragment dispersion gets more important as bulk polarizability gets larger. Since we are probing inter-fragment alkyl-alkyl, alkyl-aryl and aryl-aryl interactions, we observe the opposite trend observed by Cockroft for the interactions between alkyl groups.[72] It is important to note that the solvent trends of the molecular balances are largely governed by solvophobic effects.[72] In our compounds, we did not observe a correlation with cohesive energy density, suggesting that our system is not dominated by solvophobic effects.[174] We could also exclude electrostatics to be the major energy component because our data do not correlate with the dielectric permittivity of the solvent. The best descriptors for the trends of total compensation of dispersion as a function of solvent were solvent bulk polarizability or the Hansen dispersion solubility parameter. Therefore, dispersive ability of the solvent is the major contributor to the observed solvent mediated effect at lower temperatures. The observed trend of compensation as a function of bulk polarizability can be explained by the fact that solvents with large bulk polarizabilities prefer to maximize interactions with themselves. Hence, solute-solvent interactions with solutes of lower dispersive ability are disfavored and therefore avoided. This can be achieved in proton-bound dimers, because solute-solute interactions are maximized therein. Overall, proton-bound dimers with large dispersive substituents are therefore favored in solvents with high bulk polarizability or with a high Hansen dispersion solubility parameter.

In the beginning of this chapter, we stated that we expected London dispersion in solution to be tunable by changing the solvent so that the trends in dispersion can be altered by several kilocalories per mole. However, to our surprise, we found that the solvent-induced effects are rather minimal. The largest effect we could find was a change of 1.4 kcal/mol at 243 K, which corresponds to about a factor of ten in the corresponding equilibrium constant. If one was to use these solvent-induced differences to tune enantioselectivity in asymmetric catalysis, a completely unselective reaction could be transformed into a reaction with 88 % ee. This illustrates that practically useful degrees of enantioselectivity are not within reach solely on the basis of this concept. Therefore, we believe that changing the solvent to tune rationally the extent of intramolecular dispersion is unlikely to become a general new paradigm to tune rationally chemical reactivity and selectivity. Furthermore, we also observed that the absence of a significant solvent-mediated effect on the overall equilibrium constant of a molecular torsion balance was invoked as argument to disregard dispersion as an important energy component in a series of molecules with significant chalcogen bonding interactions.[204] Considering our findings, even in dispersion dominated systems solvent tuning can be ineffective showing that the conclusions derived in that respect need to be reconsidered.[204]

The major aspect of interest in this chapter was to decompose the total compensation of dispersion into contributions from the gas phase and from solution and derive them as a function of temperature. While total compensation of dispersion and the corresponding gas-phase component tend to increase with increasing temperature, the solution component shows a decreasing trend with increasing temperature. Therefore, in our model compounds, the trend in total compensation of dispersion is dominated by the trend in gas-phase compensation at the temperatures studied experimentally. The solution trend is counteracting but the gas-phase trend dominates. Therefore, it can be inferred, that in solution a non-negligible contribution of the entropy of solvation counteracts the corresponding gas-phase entropic component. The gas-phase entropy is a manifestation of the preference of the proton-bound dimers to obtain more flexibility by dissociating. The contribution of the entropy of solvation counteracts this trend and reduces the impact of conformational flexibility for Gibbs free energies in so-

lution. To rationalize this observation a general and well-established model of rotational and translation motion in solution can be applied. In general, molecules dissolved in a solvent have reduced mobility with respect to their gas-phase state. This causes an unfavorable entropy contribution, which is least unfavorable in case of minimal contact between solvent molecules and solute. This causes proton-bound dimers in solution to be entropically favored over their solvated dissociated fragments relative to the gas phase, i.e. the equilibrium is shifted towards the proton-bound dimers in solution compared to the gas phase by that entropic component. The corresponding shift, however, does not remove the entire gas-phase entropic contribution and, hence, the total compensation of dispersion is still increasing with increasing temperature. The restriction of rotational and translational degrees of freedom in solution compared to the gas-phase state has been studied over several decades.[205–208] Compared to the gas phase, the restriction of motion in solution leads to a decreased translational and rotational entropy. Based on a very approximate model, it has been estimated that the total entropy of molecules dissolved in water is diminished by about 50 %.[208] In addition, a closely related phenomenon has also been studied in some detail in the field of protein-ligand binding and coined cratic free energy, which is the loss of rotational and translational flexibility upon formation of an intermolecular complex from two molecules.[207,209–212] Using this concept to explain our observations, it can be argued that the difference in cratic free energies in the gas phase and in aprotic solvents is the origin of the decreasing trend of the compensation of dispersion with increasing temperature. This difference in cratic free energy corresponds to an entropic free energy component and its main contribution comes from entropy of solvation. Hence, we effectively invoke a model similar to the currently accepted explanation for the hydrophobic effect.[213,214] The difference between them is that in our model the main contribution comes from the dynamics of the solute and not of the solvent, as in the hydrophobic effect, and in our case the energetic contribution is significantly smaller.

We close this section by looking again at the comparison of Gibbs free energies of dissociation in solution from computations and from experiment. In the previous chapter, we discussed the problems associated with implict solvent models in great detail. However, they are still among the most computationally afford-

able theoretical models to treat solvation in computations.[215,106,139,140,141,142] The main idea behind continuum solvation models is to treat the electrostatic component of solvation as a mean solvent field. We made use of COSMO-RS[82,105] and SMD,[6] which are two of the more popular implicit solvent models. Our main motivation to select them was because they also include a dispersive contribution to the solute-solvent interaction, which is described as a term proportional to molecular surface area. In the previous chapter, we already discussed the systematic deviations between the computational approach adopted and the experimental results. We identified an improper account of gas-phase entropy as one of the possible main sources of error. That was the main reason that we improved on the composite computational approach established in the previous chapter. We used GFN-xTB methods to correct the gas-phase entropy of dissociation for the presence of multiple conformers. Including these multiple conformer corrections only slightly improved the computational estimates of the Gibbs free energies in solution, confirming our conclusion in the previous chapter that the gas-phase entropic component was not the main source of the observed discrepancies. In addition, we established that the systematic deviations are present for all the investigated solvents. This confirms that the problem is not limited to one specific type of solvent, but is more general. Furthermore, we also confirmed the conclusion of the previous chapter regarding compensation of dispersion being significantly underestimated by the computational approach adopted. In this chapter, we estimated the compensation of dispersion both from experiment and from computations in all the solvents investigated. The corresponding results showed directly that compensation is underestimated significantly. Additionally, they also demonstrated that including the corrections for the presence of multiple conformers does improve the computational estimates significantly. One of the main problems of implicit solvent models is that they cannot properly treat entropy of solvation. However, in the previous paragraph we established that we do observe a significant contribution of the entropy of solvation which could explain part of the systematic differences between computation and experiment observed. Based on these findings, it seems likely that the problem is the treatment of the solvent as continuum, which cannot treat configurational entropy in the solvent properly and which treats dispersion as a surface-based rather than

a volume based energy contribution. It also suggests that hybrid[146,205] or explicit[216] solvent models might be more suitable for proper theoretical modeling of Gibbs free energies of dissociation of our test compounds in aprotic solvents. Furthermore, it also shows the importance of our set of experimental data as systematic benchmark for computational solvent models.

3.4. Conclusions

In this chapter, we expanded the findings of the previous chapter regarding compensation of London dispersion in dissociation reactions of proton-bound dimers in solution. We performed the same type of measurements in a range of aprotic solvents and found that total compensation of dispersion approximately equals 60 %–80 % at temperatures typically employed in organic chemistry. We established that total compensation of dispersion is largely invariant with respect to the solvent. This finding can likely be generally transferred to solvents that do not show specific binding interactions with themselves like H-bonds. Therefore, we deem it unlikely that rationally tuning the importance of inter-fragment London dispersion by the choice of solvent will become a general new paradigm to control chemical reactivity, selectivity and intermolecular interactions. Furthermore, based on the findings of the previous chapter, we gathered a substantial amount of evidence that implicit solvent models do not properly account for solute-solvent dispersive interactions, which results in a significant underestimation of the compensation of dispersion in solution. Therefore, our findings raise doubt as to the general applicability of COSMO-RS and SMD to compute Gibbs free energies of reaction in solution, especially when going to larger systems in which dispersion becomes relevant. Alternative solvent models are required to solve that issue and to obtain better computational estimates.

4. Immiscibility of Alkanes and Perfluoroalkanes

This chapter is adapted with permission from "Origin of the Immiscibility of Alkanes and Perfluoroalkanes," Robert Pollice, and Peter Chen *J. Am. Chem. Soc.* **2019**, *141*, 3489 — 3506. Copyright © 2019 American Chemical Society.

4.1. Introduction

The differences between alkanes and perfluoroalkanes are manifold.[217–221] From a fundamental perspective, perfluoroalkanes occupy a larger volume than alkanes with the same number of atoms.[217,222] For instance, several steric descriptors would describe a trifluoromethyl substituent as large or even larger as an isopropyl substituent.[220] Another major difference is the high electronegativity of fluorine, which makes the atoms on the outside of fluorocarbons negatively polarized as opposed to alkanes, which have slightly positively polarized hydrogen atoms on the outside. Additionally, alkanes are more flexible than perfluoroalkanes.[196] Most importantly, despite the fact that both alkanes and perfluoroalkanes are apolar molecules, their mutual miscibilities are generally low, which is widely known as the fluorophobic effect,[223–226] which is sometimes abbreviated simply as fluorous effect. The fluorophobic effect has been explained with fluorocarbons having low intermolecular interaction energies relative to alkanes and these low interaction energies would explain the fluorous effect, as well as the low cohesive energy density and the comparably low boiling points of perfluoroalkanes.[218–220,223,227,228] It is widely assumed that the reason for the low tendency of perfluoroalkanes for intermolecular interactions is the small atomic polarizability of fluorine. The low atomic polarizability leads

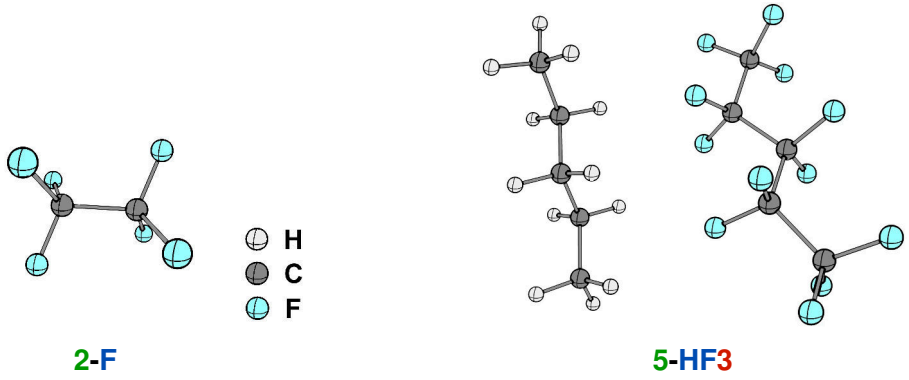
to relatively weak intermolecular London dispersion[16] so that perfluoroalkanes prefer to minimize contact with non-fluorous molecules in order to maximize contact of non-fluorous molecules with each other.[219,223] While that interaction model explains many of the properties of fluorocarbons, it fails to properly describe the trends of their bulk properties with increasing chain length. One particularly important observation in that regard is that the boiling points of linear alkanes and perfluoroalkanes invert their relative magnitudes when the chain is elongated.[220,227] Furthermore, a few studies with rigorous ab initio computations of the intermolecular complexes of methane and perfluoromethane estimate that the dimer of perfluoromethane is bound considerably stronger than the dimer of methane. In addition, our results in the previous chapter with 1-HPFX as solvent were surprising. These findings prompted us to critically reassess the general notion that perfluoroalkanes have a weak propensity for intermolecular interactions.[229–231]

4.2. Results

4.2.1. Test Systems

In this chapter, our test systems are homomolecular and heteromolecular intermolecular complexes of linear alkanes (**H**) and perfluoroalkanes (**F**) of equal chain length. Alkane-alkane complexes are generally abbreviated as **HH**, alkane-perfluoroalkane complexes are generally abbreviated as **HF** and perfluoroalkane-perfluoroalkane complexes are generally abbreviated as **FF**. The chain length is varied from one to six. For better reading, we defined a specific set of abbreviations to refer to specific molecules and intermolecular complexes and the corresponding abbreviations are explained in Table 4.1 based on two particular examples. Furthermore, Figures and Tables use a consistent color code when a particular subset of all the intermolecular complexes are studied. **HH** is shown in magenta, **HF** is shown in cyan and **FF** is shown in yellow.

In addition, for benchmarking, we used several common statistical parameters which assess the deviations of a given set of data with values x_i from a set of reference data with values x_i^{ref}. These parameters are defined in Equations



Number of carbon atoms	One or two letters for the molecules (H = alkane, F = fluorocarbon)	Optional number to distinguish conformers

Table 4.1.: Abbreviations used throughout this chapter.

4.1-4.4. Notably, the relative mean absolute deviation (MAD_{rel}) is alternatively sometimes also termed mean absolute relative deviation ($MARE$).

$$\text{Mean Absolute Deviation: } MAD = \sum_i \left| x_i - {x_i}^{ref} \right| \tag{4.1}$$

$$\text{Relative Mean Absolute Deviation: } MAD_{rel} = \sum_i \left| \frac{x_i - {x_i}^{ref}}{{x_i}^{ref}} \right| \tag{4.2}$$

$$\text{Mean Signed Deviation: } MSD = \sum_i \left(x_i - {x_i}^{ref} \right) \tag{4.3}$$

$$\text{Relative Mean Signed Deviation: } MSD_{rel}(X) = \sum_i \frac{x_i - {x_i}^{ref}}{{x_i}^{ref}} \tag{4.4}$$

Another measure that will recur in this chapter is the relative molar content of a certain type of atom X in a molecule, which is defined in Equation 4.5.

$$\text{Relative Molar Content of X: } n_{rel}(X) = \frac{\text{Number of X Atoms}}{\text{Total Number of Atoms}} \tag{4.5}$$

4.2.2. Benchmarking

To get started, we carried out extensive literature research to find computational benchmark studies of complexes of **H** and **F**, which should serve as a reference

for choosing good methods for geometries and energies. However, we realized quickly that only a very limited amount of benchmark data for **F** has been published.[229–232] Hence, we started our investigations with benchmarking computational methods against BIEs from computational methods, which are considered standard benchmark quality. We used DSD-PBEP86[233]/def2-TZVP(spd)[177] for geometry optimizations as the corresponding functional performed extremely well on a wide variety of systems, showing that it is a very robust method[233] and because the basis set provides a good cost-accuracy ratio. We also tested the geometries obtained from DSD-PBEP86/def2-TZVP(spd) against WFT methods by performing rigid scans of the intermolecular distance between interacting molecules, both with MP2/CBS and with DLPNO-CCSD(T)/QZ.[30,34,37] We found the minima of the DFT geometries to correspond very well to the minima in the scans and, hence, we concluded that DSD-PBEP86 geometries treat all three classes of intermolecular complexes consistently well.

For benchmarking of BIEs, we selected CCSD(T)/CBS and CCSD(T)-F12/VTZ as reference methods as they are considered the gold standard for noncovalent interactions and especially for London dispersion.[24,25] However, since these methods are extremely costly, we wanted to find a computationally more affordable method that reproduces the corresponding results of CCSD(T)/CBS and CCSD(T)-F12/VTZ and which would be best able to cover a wider range of our test compounds. Therefore, we computed the BIEs of all the methane and ethane complexes with high level WFT computations, most of which are considered benchmark quality. We selected CCSD(T)-F12/VTZ as reference method due to its good performance in literature benchmarks.[24,25] All the results are shown in Table 4.2.

We found that CCSD(T)-F12/VDZ gave very comparable results to CCSD(T)-F12/VTZ and CCSD(T)/CBS at a significantly lower computational cost. Our approach was to benchmark the smallest intermolecular complexes with CCSD(T)/CBS and CCSD(T)-F12/VTZ, and use CCSD(T)-F12/VDZ for a larger subset of all the model compounds. Hence, our main body of benchmark data consisted of intermolecular complexes of **H** and **F**, which were optimized at the DSD-PBEP86/def2-TZVP(spd) level of theory with single points at the CCSD(T)-F12/VDZ. It should be noted that we selected only a few of the lowest

Method	MAD [kcal mol^{-1}]	MAD$_{rel}$ [%]	MSD [kcal mol^{-1}]	MSD$_{rel}$ [%]
CCSD(T)-F12/VDZ	0.02(1)	2(1)	0.02(1)	2(1)
CCSD(T)/CBS(34)	0.03(3)	3(3)	0.02(3)	2(4)
SAPT2(CCD)δMP2/aTZ	0.03(3)	3(3)	0.02(4)	2(4)
SAPT2+(3)δMP2/aTZ	0.06(3)	6(3)	0.06(3)	6(3)
CCSD(T)/CBS(LPNO-CEPA/1)	0.07(4)	8(4)	-0.05(6)	-4(8)
CCSD(T)/CBS(MP2)	0.13(9)	14(9)	-0.13(9)	-14(9)
MP2/CBS(34)	0.13(8)	14(7)	0.12(10)	13(9)
MP3/CBS(34)	0.14(10)	20(10)	0.14(10)	20(10)

Table 4.2.: Benchmark of WFT methods for the BIEs of methane and ethane complexes of our test compounds against CCSD(T)-F12/VTZ as reference.

energy conformers for consecutive single points at higher level of theory. The lowest energy conformers were selected using DSD-PBEP86/def2-QZVP. Consequently, we tested a large number of DFT, WFT and semiempirical methods for their performance to reproduce the benchmark BIEs of our model complexes. The corresponding results are summarized in Table 4.3.

Table 4.3 reveals that B97M-V[64] provides the most robust results. Using ExtremePNO thresholds, DLPNO-CCSD(T)/CBS performs similarly well. While DLPNO-CCSD(T)/CBS with TightPNO or tighter thresholds give good results, MP2[234]/CBS is not considerably worse than DLPNO-CCSD(T)/CBS with TightPNO but much more computationally affordable, especially using RI-MP2.[235,236] However, the most efficient methods to obtain accurate results compared to the benchmark methods are B97M-V and ωB97M-V[65] as they are much cheaper than any of the DLPNO-CCSD(T) methods. Furthermore, all the DFT methods using the XDM[47,49–51] dispersion correction perform quite well. B3LYP-XDM is the best method after DLPNO-CCSD(T)/CBS with ExtremePNO parameters and the DFT methods B97M-V and ωB97M-V. To cover the whole range of systems, B97M-V is the best and most efficient method, B3LYP-XDM is somewhat better for **HF** and B3LYP-D3 is slightly better for **HH**.

Method	HH		HF		FF		Overall	
	MAD_{rel} [%]	MSD_{rel} [%]	MAD_{rel} [%]	MSD_{rel} [%]	MAD_{rel} [%]	MSD_{rel} [%]	MAD_{rel} [%]	MSD_{rel} [%]
B97M-V	7(5)	5(7)	3(2)	3(3)	3(1)	0(3)	4(3)	3(5)
ExtremePNO	2(1)	-1(2)	3(4)	3(4)	9(4)	9(4)	4(4)	3(4)
ωB97M-V	5(3)	5(3)	4(2)	-4(2)	7(4)	-7(4)	5(3)	-2(6)
B3LYP-XDM	3(1)	1(4)	2(2)	2(2)	14(3)	14(3)	5(5)	4(6)
VeryTightPNO	2(1)	0(2)	5(2)	5(2)	13(4)	13(4)	6(5)	5(5)
TightPNO	1(1)	1(1)	6(2)	6(2)	17(5)	17(5)	7(6)	7(6)
CCSD(T)/CBS(LPNO-CEPA/1)	3(3)	1(5)	11(2)	-11(2)	7(3)	-7(3)	8(4)	-7(6)
RI-MP2	4(3)	-3(4)	5(5)	5(5)	17(7)	17(7)	8(7)	5(9)
PBE-XDM	10(13)	-8(15)	7(8)	-6(9)	16(8)	12(14)	10(10)	-3(13)
SAPT2+	4(2)	4(3)	9(3)	9(3)	20(5)	20(5)	10(7)	10(7)
DSD-PBEP86	5(5)	-5(5)	12(2)	12(2)	28(1)	28(1)	13(9)	10(12)
BLYP-XDM	9(7)	5(10)	10(4)	10(4)	26(3)	26(3)	13(8)	12(10)
M06L	9(12)	8(13)	13(13)	13(13)	19(15)	19(15)	13(13)	13(13)
B3LYP-NL	22(10)	22(10)	7(4)	-5(6)	20(4)	-20(4)	14(9)	-1(17)
CCSD(T)/CBS(MP2)	3(2)	-2(3)	19(5)	-19(5)	19(3)	-19(3)	14(9)	-14(9)
PBE-D3	22(16)	-22(16)	8(8)	1(11)	23(14)	22(15)	15(14)	-1(21)
GFN2-xTB	24(13)	23(14)	9(8)	-6(10)	19(12)	2(24)	15(12)	4(19)
M06-2X-D3	6(4)	-4(6)	15(5)	15(6)	27(5)	27(5)	15(9)	12(13)
BLYP-NL	28(18)	28(18)	9(6)	-5(10)	20(6)	-20(6)	17(13)	1(22)
B3LYP-dDsC	22(15)	-20(18)	8(7)	7(8)	29(8)	29(8)	17(13)	4(21)
NormalPNO	4(4)	4(4)	18(4)	18(4)	33(5)	33(5)	17(11)	17(11)
M06L-D3	18(9)	-15(14)	19(6)	-17(11)	14(4)	-3(15)	18(7)	-14(14)
B2PLYP-D3	11(4)	11(4)	18(3)	18(3)	30(2)	30(2)	18(7)	18(7)
TPSS-D3	10(10)	-3(14)	20(8)	19(9)	39(18)	39(18)	21(15)	17(19)
B3LYP-D3	3(1)	1(3)	23(3)	23(3)	40(3)	40(3)	21(14)	20(15)
GFN1-xTB	27(18)	27(18)	14(5)	14(5)	37(5)	37(5)	23(14)	23(14)
ωB97X-D3	13(7)	-13(7)	20(5)	20(5)	65(13)	65(13)	28(21)	20(29)
vdW-DF10	14(9)	-14(9)	36(7)	-36(7)	32(12)	-32(12)	29(13)	-29(13)
PBE-dDsC	37(6)	-37(6)	34(6)	-34(6)	22(6)	-22(6)	32(8)	-32(8)
BLYP-D3	12(9)	12(9)	36(6)	36(6)	58(5)	58(5)	34(18)	34(18)
M06-2X	16(6)	16(6)	42(10)	42(10)	46(9)	46(9)	35(15)	35(15)
B97-D3	18(9)	-18(9)	33(8)	33(8)	67(22)	67(22)	36(21)	26(33)
SCS-MP2	38(3)	38(3)	41(6)	41(6)	50(6)	50(6)	42(7)	42(7)
sSAPT0	53(13)	53(13)	64(11)	64(11)	79(14)	79(14)	64(15)	64(15)

Table 4.3.: Benchmarking of DFT and WFT methods against CCSD(T)-F12/VDZ BIEs using DSD-PBEP86/def2-TZVP(spd) geometries. The respective standard deviations are given in parenthesis.

4.2.3. Energy Decomposition Analysis

Having completed the extensive benchmarking studies, we moved on to understanding the intermolecular interactions of **H** and **F**. We started our investigations with EDA using SAPT,[115,116,237,238] and our findings are reported in Figure 4.1. It shows that the general trends are roughly consistent across all the various complexes under study. This shows that the relative importance of all the energy components is approximately equal in **HH**, **HF** and **FF**. Notably, while induction is basically zero, electrostatics do show a significant contribution. The most important components are exchange and dispersion, and their sum is overall still attractive, showing that dispersion is the most important interaction energy component.

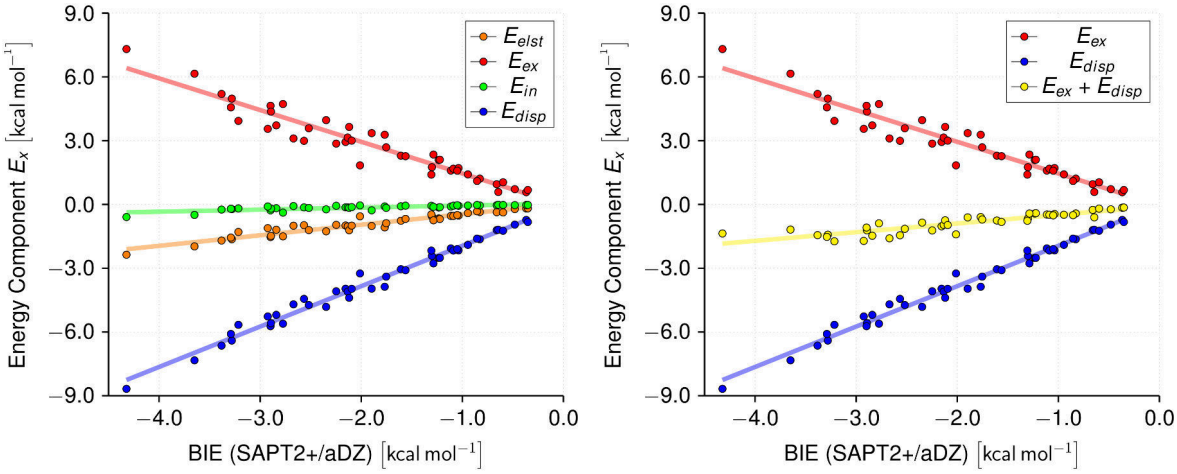


Figure 4.1.: Decomposition of BIEs of complexes using SAPT2+. Left: All energy components plotted against BIE. Right: Exchange, dispersion and their sum plotted against BIE.

By looking more closely into the sum of exchange and dispersion for the three subsystems **HH**, **HF** and **FF**, we found that in **FF** the sum of exchange and dispersion is slightly more attractive than in **HF** or **HH**. The corresponding results are depicted in the top diagrams of Figure 4.2. In contrast, investigating the electrostatic component for the three subsystems reveals that it is slightly more important in **HH** than in **HF** or **FF**. These results are illustrated in the bottom

diagrams of Figure 4.2. Hence, based on the SAPT results, this suggests that dispersion is more important in **FF** than in **HH**.

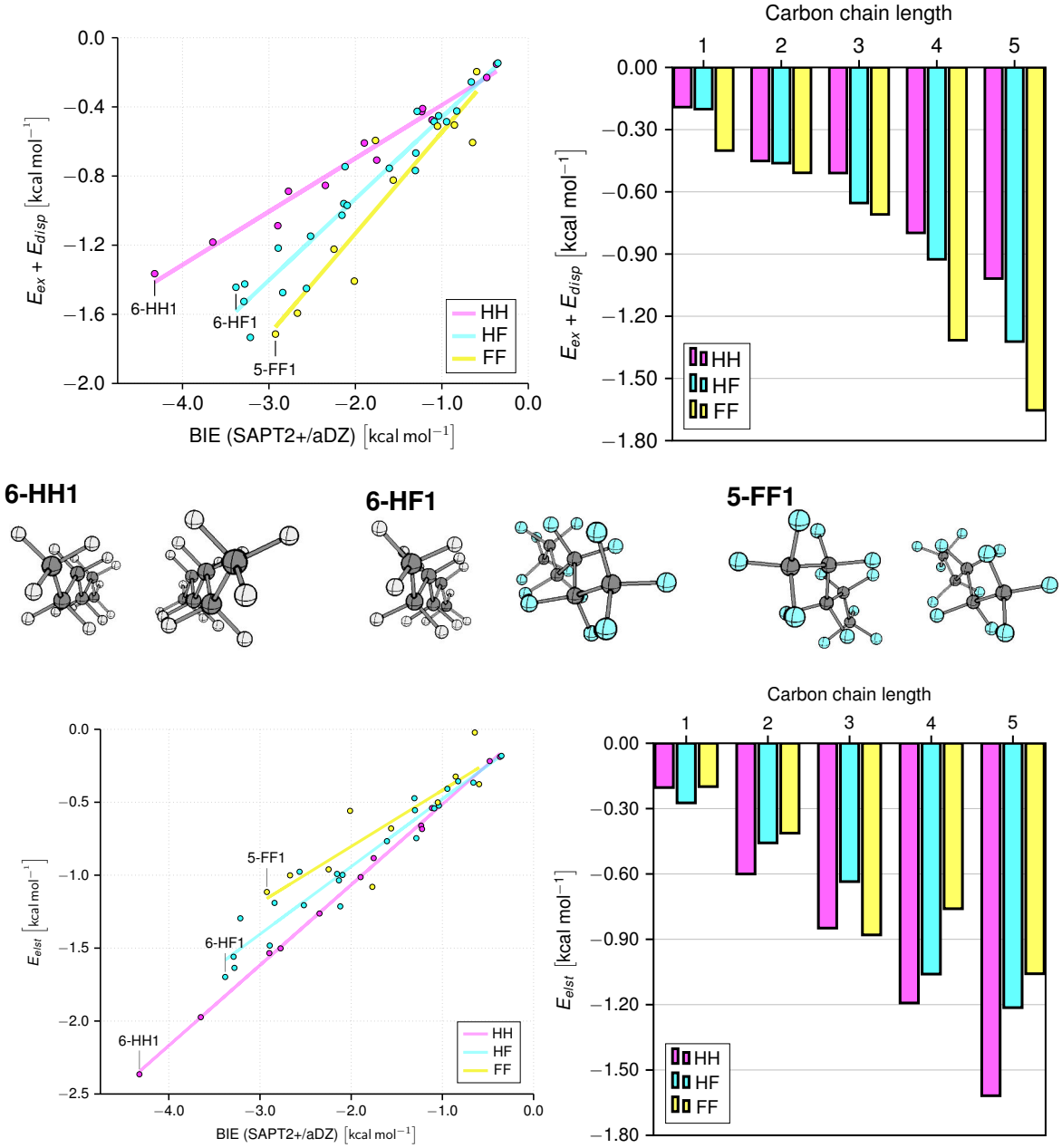
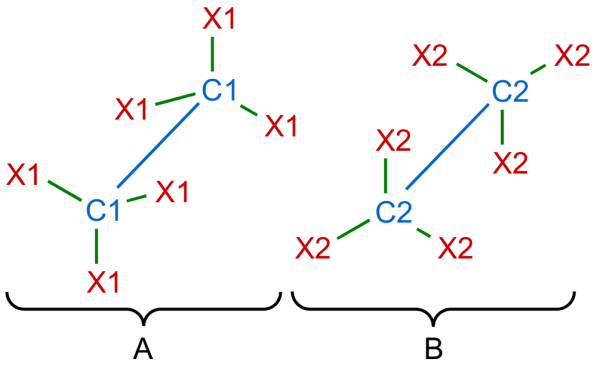


Figure 4.2.: Top: Sum of exchange and dispersion from SAPT in **HH**, **HF** and **FF**. Bottom: Electrostatics from SAPT in **HH**, **HF** and **FF**.

We wanted to gain more insight into the trends in BIEs by performing an alter-

native EDA method, namely LED.[75,239] First, we compared the dispersion component from both SAPT2+ and LED-DLPNO-CCSD(T)/QZ and found a very significant linear correlation suggesting that trends across all the molecular complexes studied are treated equally in both methods. We moved on to decompose total intermolecular dispersion further into contributions of bonds and atom types. This decomposition is illustrated in Scheme 4.1.



$$E_{disp}(A-B) = E(C1-C2) + E(C1-X2) + E(X1-C2) + E(X1-X2) + E(C1-CX2) + E(CX1-C2) + E(X1-CX2) + E(CX1-X2) + E(CX1-CX2)$$

Scheme 4.1: Dispersion energy decomposition of a complex A-B into bond and atom type components.

Hence, the total intermolecular dispersion is decomposed into interactions of the X atoms, where X can be either F or H, the C-X bonds and the carbon chains. Using that decomposition, we investigated all molecular complexes under study. The corresponding findings for the complexes with one carbon atom are depicted in Figure 4.3 as a representative result.

Figure 4.3 illustrates that the dispersion in **1-H** originates roughly equally from the carbon atom and the C-H bonds. In **1-F**, the dispersion originates almost entirely from the fluorine atoms with a very minor contribution from the C-F bond. Therefore, in **1-FF**, almost all the intermolecular dispersion is caused by F–F contacts. In **1-HF**, most dispersion comes from C–F contacts, from interactions

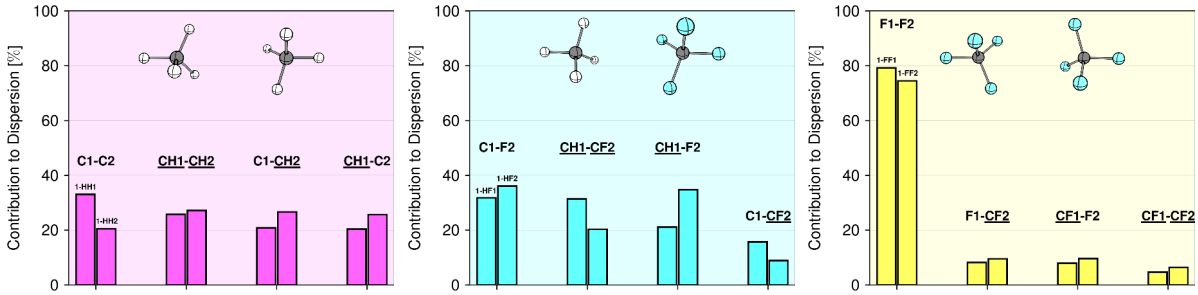


Figure 4.3.: Partitioning of total intermolecular dispersion into bond and atom components using LED for methane complexes.

between C-H bonds and C-F bonds and from interactions of fluorine atoms with C–H bonds. In **1-HH**, dispersion comes from C–C contacts, from interactions between carbon atoms with C-H bonds and from interactions between C-H bonds. For **HH** and **HF**, the results of the complexes with longer carbon chains are not significantly different from the methane complexes. However, for **FF** they reveal that the relative contribution of the F–F contacts towards the total dispersion decreases with increasing chain length, it even shows a statistically significant linear correlation with the relative molar fluorine content.

Next, we wanted to find support for the findings of LED using an alternative computational method. Therefore, we used the recently developed DID methodology.[79] DID is based on localized orbitals from WFT and determines the orbital excitations that correspond to inter-fragment dispersion interactions. The corresponding orbitals, which are excited to provide dispersion, are highlighted to visualize the origin of the biggest contribution to dispersion.[79] We used voxel plots to visualize the DID results of all our intermolecular complexes and the corresponding results for complexes with one carbon atom, which are representative for all the longer complexes as well, are depicted in Figure 4.4. These DID plots provide independent support for the findings of LED that dispersion in **1-F** originates mainly from the fluorine atoms and that it originates mainly from the C-H bonds and the carbon atom in **1-H**.

An alternative way to investigate the origin of dispersion in molecules is to study their distribution of molecular polarizabilities. This can be done using a QTAIM method, which partitions the molecular polarizabilities into atomic polarizabilities based on molecular basins and bond paths.[240] This method can be car-

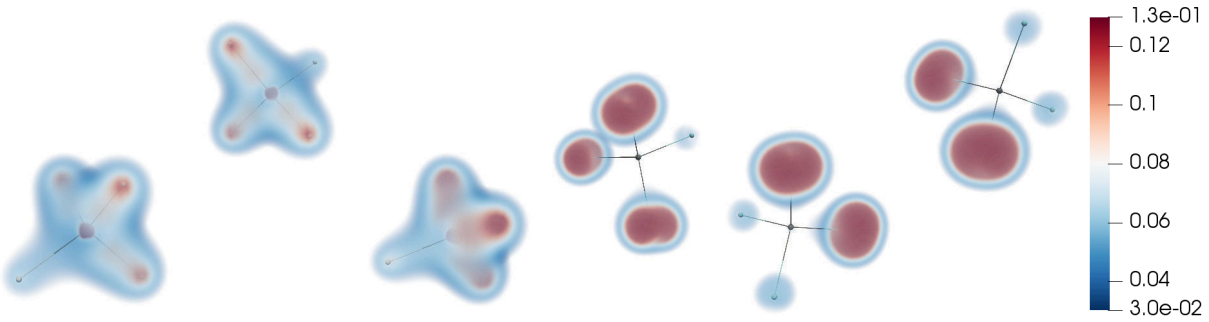


Figure 4.4.: DIDs of methane complexes: Left: **1-HH1**. Middle: **1-HF1**. Right: **1-FF2**. Strong dispersive contributions are shown in red, weak ones in blue.

ried out using AIMAll[241] and PolaBer.[242] Employing that approach, we computed relaxed MP2/def2-TZVPPD densities for all **H** and **F** molecules and determined the atomic polarizabilities for all of the single molecules of our test systems. PolaBer visualizes the resulting anisotropic atomic polarizabilities using ellipsoids to represent the respective anisotropy with their volumes representing the respective total isotropic polarizabilities. The corresponding results for **1-H** and **1-F** are illustrated in Figure 4.5. The corresponding numerical results compared to the atomic polarizabilities of isolated hydrogen, carbon and fluorine atoms are provided in Table 4.4.

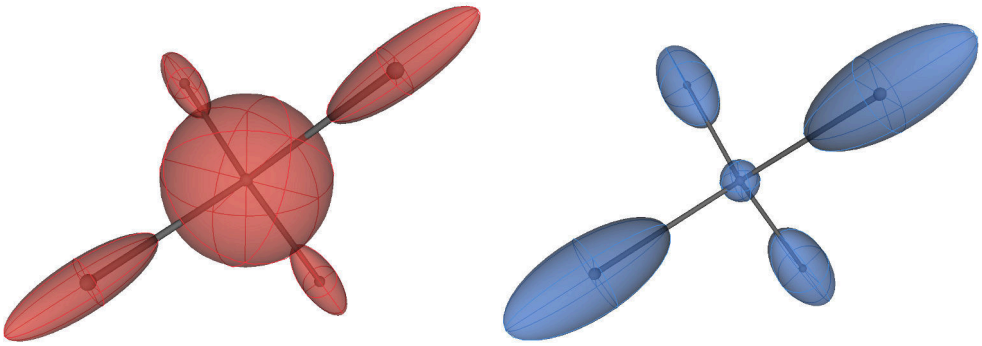


Figure 4.5.: Distributed atomic polarizabilities of **1-H** (left) and **1-F** (right) obtained using PolaBer.

The findings of Figure 4.5 and Table 4.4 demonstrate that in **1-F** fluorine has a larger polarizability compared to carbon, which is the reversed order compared

Atom Type	Polarizabilities [Å^3]		
	Free Atom	CH_4	CF_4
C	1.76[243]	0.87	0.28
H	0.67[244]	0.38	–
F	0.56[243]	–	0.62

Table 4.4.: Atomic polarizabilities of free hydrogen, carbon and fluorine atoms compared to their polarizabilities in CH_4 and CF_4, respectively.

to the polarizabilities in the free atoms.[245] In **1-H**, carbon has the largest atomic polarizability. These findings are consistent with London dispersion in **FF** stemming mainly from contacts between fluorine atoms because they are both the most polarizable atoms and they are closest to each other. Hence, in this section, we found independent support from three fundamentally very different methods that dispersion in fluorocarbons originates mainly from the fluorine atoms.

4.2.4. Simple Interaction Models

Having established from which atoms dispersion originates in the molecular complexes under study, we wanted to find simplified models to describe the intermolecular interactions and the corresponding trends. Hence, we used the London dispersion formula,[2] which is reproduced in Equation 4.6, to describe the BIEs of all our model complexes by treating the two molecules as interacting spheres, which are separated by the interaction distance R, with the ionization potentials (I_i) and molecular polarizabilities (α_i) of the actual interacting molecules.

$$E_{disp} = -\frac{3}{2}\frac{I_1 I_2}{I_1 + I_2}\frac{\alpha_1 \alpha_2}{R^6} \tag{4.6}$$

To determine the molecular ionization potentials, we used MP2/CBS computations and molecular polarizabilities at the MP2/def2-TZVPD level of theory. The only remaining question was which distance to use for the intermolecular separation. When we used the average distance between the carbon chains in the interacting molecules, the resulting estimated energies did not show good correlation to the BIEs obtained from benchmark calculations. When we only used the types of atoms, which are mainly responsible for London dispersion, namely carbon for **H** and fluorine for **F**, we observed a linear correlation between the

benchmark BIEs and the estimated intermolecular London dispersion energies. It should be noted, that we used the R^{-6}-weighted distance between all the atoms of the corresponding atom type in the intermolecular complex. The corresponding comparison is depicted in the left diagram of Figure 4.6.

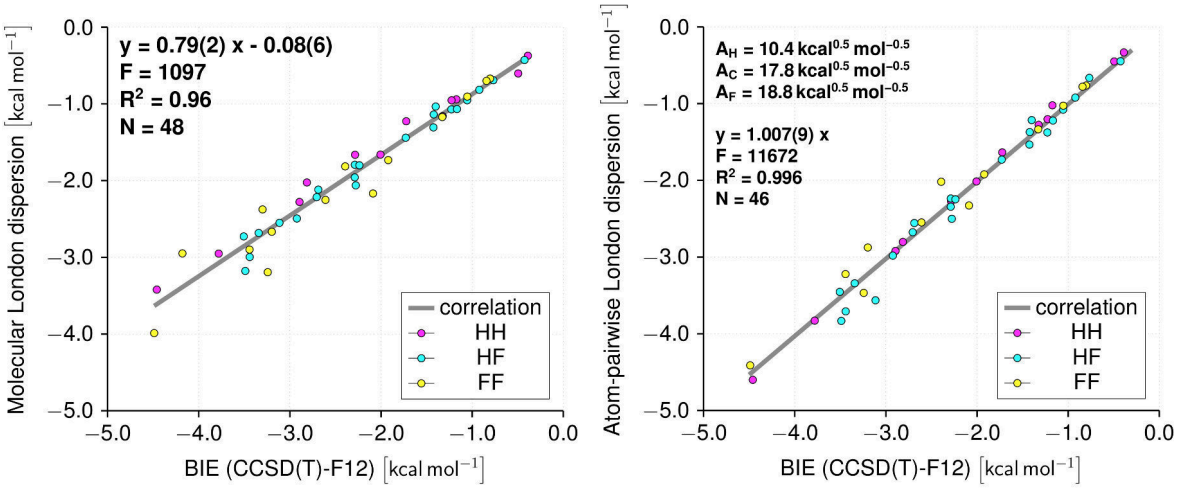


Figure 4.6.: Left: Comparison of CCSD(T)-F12 BIEs and intermolecular London dispersion between interacting molecules using molecular polarizabilities and ionization potentials. Right: Comparison of CCSD(T)-F12 BIEs and atom-pairwise London dispersion by fitting atomic pre-factors A to benchmark BIEs.

To refine this crude interaction model, which nevertheless provided surprisingly good results, we used an atom-pairwise London dispersion model, which is very similar to atom-pairwise dispersion corrections.[3] To do that, we used the optimized geometries from DSD-PBEP86 and the atomic polarizabilities determined on the basis of MP2/def2-TZVPD with PolaBer. In addition, we approximated the intermolecular London dispersion using an approach, which is explained and justified in the following section. It approximates the pre-factor containing the ionization potentials of the interacting atoms (termed A_{12} by us) as product of pre-factors A_1 and A_2, which are only atom-type dependent. Based on this approximation, we optimized these atomic pre-factors to minimize the deviations with respect to benchmark BIEs. The respective outcome is illustrated in the right-hand diagram of Figure 4.6. It demonstrates that the optimum set of atom-type

dependent pre-factors leads to excellent agreement between benchmark BIEs and intermolecular dispersion estimated based on this model. Furthermore, this model also allows decomposition of the total intermolecular London dispersion into components of certain pairs of atom types, which is very much akin to the decomposition performed using LED (*vide supra*). We find that the corresponding results using this simple atom-pairwise model agree even quantitatively well with the findings obtained using LED.

4.2.5. Interaction Descriptors

In the previous section, we could demonstrate that simple interaction models can reproduce both relative strength and trends in the BIEs of our model compounds. The next step was the development of simple descriptors of interaction, which would properly reflect the dispersive interaction potential of **H** and **F**. In 2004, Dunitz introduced the ratio Q of molecular polarizability (α) and molecular volume (V) for that purpose, which is dimensionless and could quantify differences in abilities of molecules to undergo dispersion:[219]

$$Q = \frac{\alpha}{V} \tag{4.7}$$

The main problem, which was already pointed out by Dunitz himself,[219] is that this Q measure disregards the spatial dimensions of the molecules. In addition, his definition neglects the spatial distribution of polarizability within a molecule. We compared the definition of the Q measure to the London dispersion formula and noticed, after some minor rearrangements, that two quotients, which are very similar to the Q measure, appear in the original formula:

$$E_{disp} = -\frac{3}{2}\frac{I_1 I_2}{I_1 + I_2}\frac{\alpha_1\alpha_2}{R^6} = -\frac{3}{2}\frac{I_1 I_2}{I_1 + I_2}\frac{\alpha_1\alpha_2}{R^3 R^3} = A_{12}Q_1Q_2 \approx A_1A_2Q_1Q_2 \tag{4.8}$$

Comparing the definition of the Q measure provided by Dunitz[219] to the rearranged London dispersion formula, there is a fundamental difference between the two measures. While in the former the molecular volume is used, the latter uses the interaction distance cubed, which can be regarded as an interaction

volume. This shows that a Q measure properly accounting for the dispersive ability of a molecule needs to use the interaction volume rather than the molecular volume. The advantage of the interaction volume is that it does account for the dimensions of the molecules and also for the distribution of polarizability in the molecules. Based on this derivation, we introduced a new Q metric, which we termed Q_{atom}, as the quotient between the polarizability of an atom in a molecule and the third power of the distance of that atom with respect to a variable point in space. Hence, depending on the position of that reference point, Q_{atom} will give different results. Q_{atom} is designed to quantify the dispersive ability of a specific atom in a molecule with respect to a reference point. We applied this definition to determine the maximum Q_{atom} of F, C and H in **H** and **F** in our model compounds by only considering points on the respective vdW surface of the molecules. The plot of maximum Q_{atom} as a function of carbon chain length is depicted in the left diagram of Figure 4.7. It demonstrates that fluorine in **F** has a higher dispersive ability than either H or C in **H**. Notably, in **F** the Q_{atom} value of C is much smaller than in **H**.

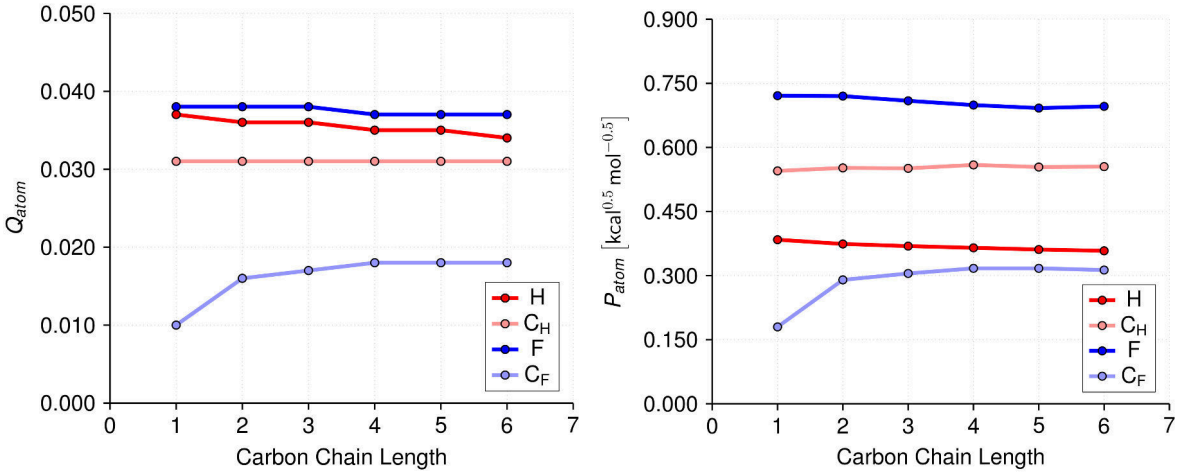


Figure 4.7.: Left: Maximum Q_{atom} of C, H and F atoms in **H** and **F**. Right: Maximum P_{atom} of C, H and F atoms in **H** and **F**. C_X refers to C atoms in **X** (X = H,F).

Q_{atom} only describes the dispersive ability of one atom. To get the total dispersive ability of a molecule, we simply sum all Q_{atom} values of a molecule and

obtain Q_{tot}. Again, it provides the dispersive ability with respect to a specific point in space. Equivalent to Q_{atom}, Q_{tot} accounts for the dimensions of molecules and the polarizability distribution of a molecule. We used the newly introduced measure to determine the average Q_{tot} for all our model compounds using points on their vdW surface as the reference points. We compared the corresponding Q_{tot} values of **H** and **F** of equal carbon chain length and found that, on average, the Q_{tot} of **F** is about 83 % of the corresponding value of **H** when carbon chain lengths are equal. The reason why Q_{tot} estimates the dispersive ability of **F** weaker than **H** is the low value of Q_{atom} of C in **F**. If we used Dunitz's Q measure, we would find that Q of **F** is about 64 % of **H**. Therefore, proper account of the dimensions of molecules and the spatial distribution of polarizability therein are important properties to take into account when dispersive interaction ability is to be quantified.

However, based on Q_{tot} one would expect the dispersive ability of **F** to be weaker than in **H**. This is counter to our findings in the previous sections. Therefore, we revisited our derivation leading to the new Q measures (cf. Equation 4.8) and found that Q alone is not sufficient to obtain dispersion interaction energies because the pre-factor A_{12} is neglected. This term is determined based on the atomic ionization potentials. In case the atomic ionization potentials in different molecules are equal, the Q_{tot} should be a reliable measure of dispersive interaction ability. However, to compare the dispersive ability of molecules with largely differing atomic ionization potentials, this term needs to be taken into account. The problem of A_{12} is that it depends on the ionization potential of both interacting atoms in an inseparable way. However, our goal is to have a measure that quantifies the dispersive abilities of the interacting molecules in a way that is only dependent on that one single molecule. Therefore, we made an approximation and assumed that A_{12} can be approximated as product of two factors A_1 and A_2, which are only dependent on the atom type (cf. Equation 4.8). These factors were estimated from the atom-pairwise London dispersion model discussed in the previous section by finding a set of parameters that best reproduces benchmark BIEs. The atom-pairwise London dispersion model reproduces benchmark BIEs with an MAD_{rel} of 5 %. In addition, all three subsystems are described well. With these pre-factors, we defined a new descriptor of dispersive ability, which

we termed P. It is the product of A and Q. It has the dimension of the square root of energy and hence, the dispersion energy between two interactors can be estimated as product of two P values:

$$P = AQ = A\frac{\alpha}{R^3} \tag{4.9}$$

Next, we introduced P_{atom} and P_{tot} analogous to Q_{atom} and Q_{tot} and determined the corresponding P_{atom} values of F, C and H in **F** and **H** with respect to the vdW surface. The corresponding outcome is illustrated in the right diagram of Figure 4.7. Comparing the corresponding results of Q_{atom} and P_{atom}, we found that the relative values of F and C are increased in P_{atom} because the corresponding prefactors A(C) and A(F) are significantly higher than A(H). In the same way, the estimated P_{tot} of all the model compounds were compared and we found that **F** has 116 % of the dispersive ability of **H** with the same number of carbon atoms. Furthermore, we found a linear correlation between the ratio of P_{tot} of **F** and **H** with the relative molar carbon content in the molecules. The corresponding ratio decreases as the carbon chain gets longer, which means that the dispersive ability of **F** compared to **H** decreases with increasing chain length. However, based on the correlation, the dispersive ability of **F** should never get below 100 % because the relative molar carbon content of saturated fluorocarbons and hydrocarbons only lies in a very narrow value range (between $\frac{1}{5}$ and $\frac{1}{3}$). The lowest extrapolated ratio of P_{tot} values is 108 %. Based on these findings, in the respective extended conformations, every saturated linear **F** should have a higher intrinsic dispersive ability than the corresponding **H**, which corresponds to what our EDA analysis had found.

4.2.6. Interaction Geometries

Having established an interaction descriptor that reproduces our findings in the previous sections, we compared them also to the BIEs of all the studied molecular complexes. It is important to point out that the intrinsic dispersive ability of the fluorocarbons compared to their hydrocarbon analogues is only manifested in the BIEs and BAEs of **1-FF** and **1-HF** compared to **1-HH**. As soon as the carbon chain gets longer, this is not true anymore. Therefore, based on these findings,

there should be an alternative reason for why **FF** and **HF** have relatively weak intermolecular interactions compared to their **HH** counterparts. Our hypothesis was that the reason lies in the geometries of the complexes and hence, we moved on to investigate the interaction geometries in more detail.

To do that, we first looked at NCI[78] plots of our model complexes. NCI plots are based on the mathematical analysis of electron densities. They visualize portions of the electron density, which are weakly attractive or weakly repulsive.[78] It is important to note that the study of NCIs using electron densitites is inherently not able to distinguish all the subtypes of vdW interactions.[78] Additionally, due to their inherently very different approach compared to the previously presented methods to visualize weak interactions, they provide additional information as they are also able to directly investigate the geometrical arrangement of the intermolecular complex. The corresponding 3D NCI plots of **5-HH1**, **5-HF4** and **5-FF1** are depicted in Figure 4.8. It can be deduced that the portion of the electron density accounting for the NCI shows a very uniform shape, which is equally distributed over the entire molecular complex, for **5-HH1**. In contrast, the corresponding portions of the electron density for **5-HF4** and **5-FF1** are very non-uniform and not equally distributed over the entire interaction surface. This illustrates that in **HF** and **FF**, only part of the total interaction surface is utilized for intermolecular interactions, while in **HH** the entire interaction surface is made use of. Based on these observations, it seems that the ground-state 3-dimensional structures of **F** are not optimal to form strong noncovalent complexes with either **F** or **H**. The consequence is that BIEs in the corresponding complexes are smaller than they could be if intermolecular arrangement were to be ideal.

Based on these findings, we wanted to find independent support for the poor interaction geometries in complexes of **F**. To do that, we investigated the increase of both total surfaces of **H** and **F** and interaction surfaces of the three subclasses of the studied complexes as the carbon chain is elongated. The molecular surfaces were derived based on COSMO[82] calculations and the interaction surfaces were determined as difference of the sum of the molecular surfaces of the interacting molecules and the total surface of the intermolecular complex. The surface area of **F** and **H** show a linear increase with carbon chain length. The interaction surfaces of the three subclasses of complexes also show a linear in-

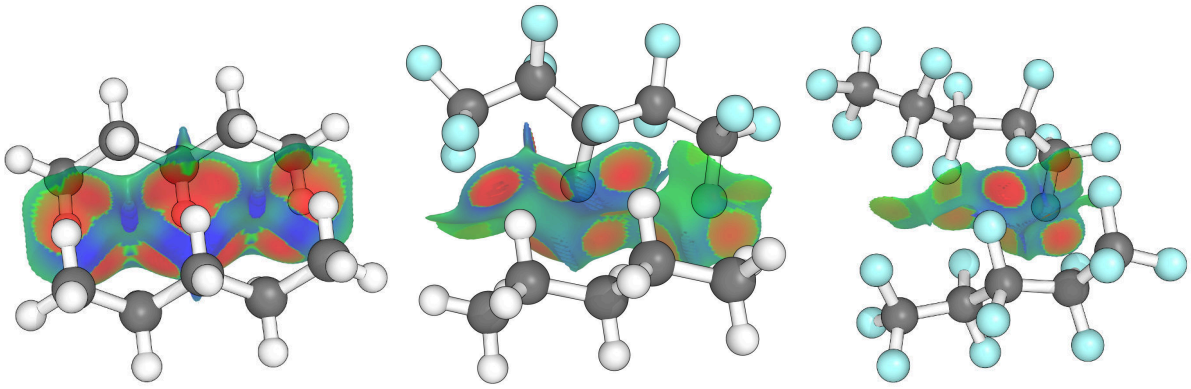


Figure 4.8.: 3D NCI plots of intermolecular complexes of **5-H** and **5-F**. Left: **5-HH1**. Middle: **5-HF4**. Right: **5-FF1**. Red shows attractive interactions, green depicts null interactions, blue shows repulsive interactions.

crease with their carbon chain length in the interacting molecules. Using a simple geometric model viewing the interacting molecules as cuboids it can be derived from the corresponding slopes of the linear correlations, that in **HF** merely about 85 % and in **FF** merely about 81 % of the added interaction surface area is used for noncovalent interactions when the carbon chain length is increased. Hence, not all the interaction surface, that would be available, is used for intermolecular interactions. These findings provide independent support for the observations in the NCI plots.

We sought to expand the evidence for poor interaction geometries even further. We wanted to characterize the quality of the interaction geometries of the three subclasses of intermolecular complexes directly. Hence, we defined a measure quantifying the alignment between interacting molecules or molecule parts. We termed the new measure the number of vdW contacts ($n_{contacts}$). One contact was defined between two atoms, if the separation between the two atoms is equal to the sum of the vdW radii (R_{vdW}) of the atoms. The vdW radii were taken from literature.[97] We wanted to have a continuous scale and hence, we introduced the following equation to compute the number of contacts, which explicitly accounts for the long-range distance dependence of London dispersion:

$$n_{contacts} = \frac{R_{vdW}^6}{R^6} \tag{4.10}$$

To determine the total number of contacts in an intermolecular complex, the sum of all atom-pairwise number of contacts needs to be determined. We used this new quantity to characterize the interaction geometries of all our molecular complexes and investigated how the number of contacts increases as the carbon chain is elongated. The corresponding findings are depicted in the left diagram of Figure 4.9. There are several observations to be pointed out. The number of contacts in the complexes of the methanes are very similar, which, as expected, shows that the interaction geometries do not differ significantly for the corresponding structures. In addition, for all three subclasses of complexes, the number of contacts correlate linearly with the carbon chain length. Looking more closely at the corresponding linear correlations, it can be seen that for **HH**, a significantly steeper increase in the number of contacts is observed as compared to the corresponding slopes for **HF** and **FF**, with **FF** showing a slightly smaller slope than **HF**. These findings provide additional support for the poor interaction geometries of **F** with both **H** and **F**. We were also interested to test whether the combination of P_{tot} and $n_{contacts}$ could be used to reproduce trends in the BIEs of all our model complexes and, hence, we correlated the benchmark BIEs of all our model complexes against the product of the number of contacts and the P values of the two interacting molecules. We observed a decent linear correlation, which is a first indication that our newly introduced parameters are indeed applicable for the approximate study of trends in intermolecular interaction energies.

4.2.7. Computed Mixing Behavior

In the beginning of this study, we benchmarked computational methods to reproduce benchmark BIEs. Having reliable methods to compute BAEs and BIEs in our model complexes, we looked into the relative BAEs of the three classes of model complexes, and how they change with the number of carbon atoms. The corresponding results are depicted in the left diagram of Figure 4.10. As the molecules get longer, the relative BAEs of **HH** become stronger compared to both **HF**, which shows weakest BAEs for molecules with long carbon chains, and **FF**. In addition, we were also interested into the corresponding results for bulk BAEs. We estimated the bulk BAEs based on the assumption that all the subclasses of intermolecular complexes studied have similar interaction geometries in bimole-

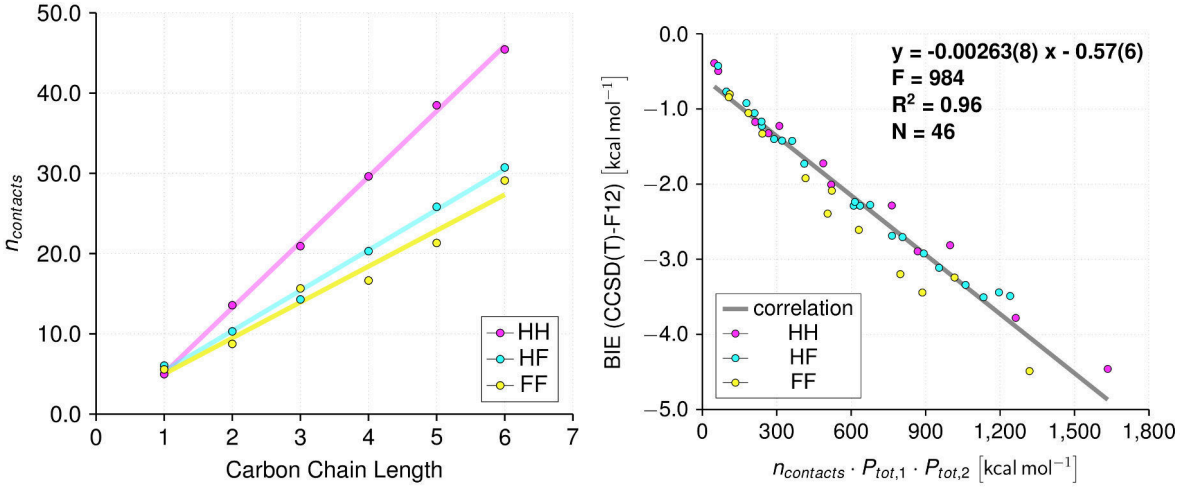


Figure 4.9.: Left: Plot of number of contacts against number of carbon atoms in **HH**, **HF** and **FF**. Right: Correlation of the BIE of all studied complexes against the product of P_{tot} and $n_{contacts}$ of the interacting molecules.

cular complexes and in the bulk phase, which also includes packing. Therefore, we computed the bulk BAEs in pure bulk phases of **F** and **H**, and also in mixed bulk phases of **H** and **F**. This was done by multiplying the bimolecular BAEs with the average coordination numbers, which we derived from the total molecular and interaction surface areas. This is a treatment of the bulk phase using Flory-Huggins solution theory.[246,247] The respective findings are depicted in the right diagram of Figure 4.10. The trends of the bulk BAEs differ from the trends of the bimolecular BAEs. While for **HF** the trend is similar from both types of BAEs, i.e. the relative BAEs become weaker as the carbon chain is elongated, for **FF** the trend differs between the bulk BAEs and microscopic BAEs. In the bulk BAEs, the relative strength of **HH** and **FF** seems to be more or less invariant with respect to the carbon chain length.

Furthermore, we were also interested in the mixing energies of **HH** and **FF** to form two complexes of **HF**. We estimated the corresponding energies for both macroscopic and microscopic mixing using various levels of theory. The corresponding results are summarized in Table 4.5. In the benchmark BAEs at the CCSD(T)-F12 level of theory, mixing is only favorable in complexes with one carbon atom but becomes unfavorable as chain length is increased, and increasingly so, the longer the chain. When instead of the benchmark BAEs the interaction energies from the London dispersion model are used, the mixing properties are

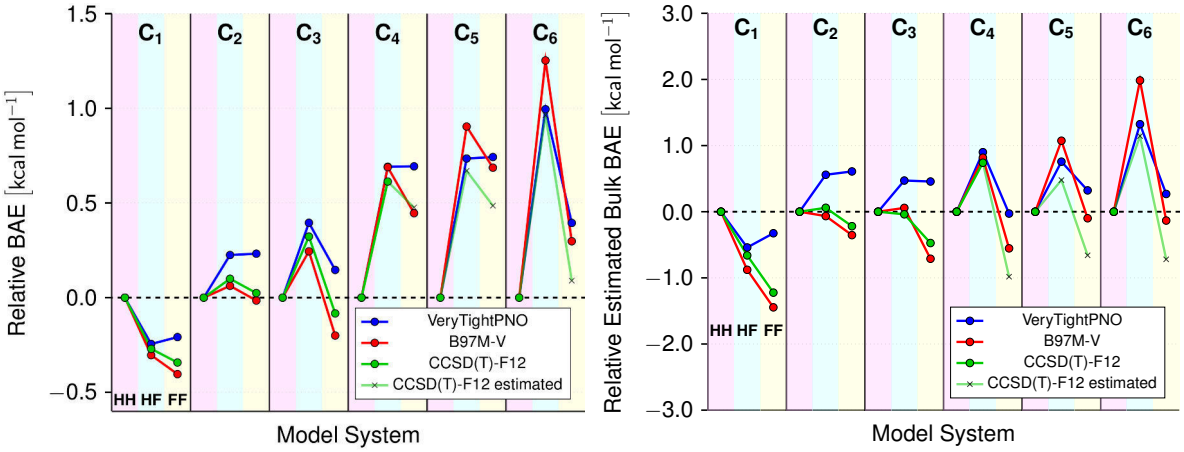


Figure 4.10.: Trend of the BAEs of the strongest binding conformers of **FF** and **HF** relative to **HH** as the carbon chain is elongated. Left: Microscopic BAEs. Right: Macroscopic BAEs.

estimated to be essentially equal. The trends in mixing energies from the London dispersion model have their origin in the respective interaction distances. When using the interaction distances only, the estimated mixing is the same as in the full London dispersion model.

Chain Length	Accounted Isomers	Relative Energies of Mixing [kcal mol^{-1}] CCSD(T) Microscopic	CCSD(T) Macroscopic	London Microscopic	London Macroscopic	Distance Microscopic	Distance Macroscopic
1	Strongest	-0.20	-0.19	-0.06	1.29	$-4.0 \cdot 10^{-5}$	$33 \cdot 10^{-5}$
	Average	0.07	0.32	0.04	0.14	$1.9 \cdot 10^{-5}$	$11 \cdot 10^{-5}$
2	Strongest	0.18	0.67	0.17	0.24	$4.9 \cdot 10^{-5}$	$9.5 \cdot 10^{-5}$
	Average	0.24	0.93	0.12	0.15	$3.7 \cdot 10^{-5}$	$3.8 \cdot 10^{-5}$
3	Strongest	0.73	0.79	0.81	1.85	$10 \cdot 10^{-5}$	$16 \cdot 10^{-5}$
	Average	0.62	0.95	0.68	1.53	$8.2 \cdot 10^{-5}$	$13 \cdot 10^{-5}$
4	Strongest	0.75	6.31	0.37	0.21	$3.2 \cdot 10^{-5}$	$0.02 \cdot 10^{-5}$
	Average	0.25	2.91	-0.02	0.25	$-0.2 \cdot 10^{-5}$	$1.2 \cdot 10^{-5}$
5	Strongest	0.85	3.23	0.67	1.98	$3.7 \cdot 10^{-5}$	$6.8 \cdot 10^{-5}$
	Average	0.72	3.01	0.36	0.75	$1.8 \cdot 10^{-5}$	$2.0 \cdot 10^{-5}$
6	Strongest	1.84	6.01	0.93	0.93	$3.5 \cdot 10^{-5}$	$3.4 \cdot 10^{-5}$
	Average	0.99	6.08	0.46	1.66	$1.5 \cdot 10^{-5}$	$4.4 \cdot 10^{-5}$

”Strongest” means that only the strongest conformer was used. ”Average” means that all conformers were used. ”CCSD(T)” are CCSD(T)-F12/VDZ values. ”London” values are based on the molecular London dispersion model. ”Distance” values are using the R^{-6}-weighted average interaction distance with all other factors in the dispersion formula to be 1 with the appropriate unit.

Table 4.5.: Energies of mixing: HH + FF $\rightleftharpoons$ 2 HF.

4.2.8. Trends in Experimental Properties

The final piece of this study is to investigate experimental trends of **H** and **F** to put the computational findings into perspective. All the experimental properties that we investigated and their information content regarding intermolecular interactions are presented in Table 4.6.

Symbol	Bulk Property	Unit	Information
$\Delta_{vap}H^{\circ}$	Enthalpy of Vaporization	kcal mol^{-1}	Strength of intermolecular interactions and packing in the liquid.
T_{boil}	Boiling Point	K	Strength of intermolecular interactions, packing in the liquid and conformational flexibility.
a/b	Critical Parameters Ratio	kcal mol^{-1}	Strength of intermolecular interactions and dependence on intermolecular orientation in the fluid.
γ	Surface Tension	mN m^{-1}	Strength of intermolecular interactions and packing in the liquid.
π_T	Internal Pressure	MPa	Dependence of strength of intermolecular interactions on incremental volume changes and packing in the liquid.
K	Bulk Modulus	MPa	Dependence of strength of intermolecular interactions on incremental volume changes and packing in the liquid.
β_T	Isothermal Compressibility	GPa^{-1}	Dependence of strength of intermolecular interactions on incremental volume changes and packing in the liquid.
η	Dynamic Viscosity	mPa s	Strength of intermolecular interactions and shape of molecules.
c_p	Molar Heat Capacity	cal mol^{-1} K^{-1}	Number of internal degrees of freedom and barrier height between them.
B_{ij}	Second Virial Coefficients	dm^3 mol^{-1}	Strength of homomolecular and heteromolecular interactions and orientational dependence of energies in the gas.

Table 4.6.: Experimental properties of **H** and **F** with explanations about their contained information.

Table 4.6 shows that the majority of the experimental properties studied contain information about the intermolecular interaction strength or about the packing in the bulk phase. While they provide similar information, the properties are based on inherently very different experimental measurements, which gives credence as to the validity of the observed trends in case they agree between different experimental measures. We looked into trends of all the experimental properties provided in Table 4.6 when the carbon chain is elongated, and most of the corresponding results are visualized in Table 4.7. The only properties not showing an inversion in their relative magnitude when **H** and **F** are compared are dynamic viscosity and molar heat capacity. For the boiling points, an inversion is observed between three and four C atoms, for the enthalpies of vaporization it is found between seven and eight C atoms. For all other experimental properties

with inversion it is found between one and two C atoms. It should be pointed out that the differences in the critical parameter ratio show a linear correlation with the relative molar carbon content, as do the differences in the boiling points. Additionally, the ratio of surface tensions shows a weak correlation with the relative molar carbon content and the differences between enthalpies of vaporization show a correlation with the number of C atoms in the chain.

Legend: F > H F = H F < H

Property	Comparison	Carbon Chain Length								
		1	2	3	4	5	6	7	8	9
$\Delta_{vap}H°$	Difference	0.88	0.35	0.50	0.46	0.31	0.29	0.17	-0.17	-0.33
T_{boil}	Difference	34.1	10.4	2.9	-2.0	-6.6	-11.6	-16.5	-19.7	n.a.[a]
a/b	Difference	0.25	-0.08	-0.17	-0.26	-0.32	-0.39	-0.44	-0.45	n.a.[a]
γ	Ratio – 1	0.28	-0.22	-0.23	-0.30	-0.37	-0.38	-0.34	-0.34	n.a.[a]
π_T	Difference	21	-26	-41	-67	-67	-59	-46	-26	n.a.[a]
K	Ratio – 1	0.48	-0.33	-0.54	-0.57	-0.52	-0.42	-0.40	-0.38	n.a.[a]
β_T	Ratio – 1	-0.32	0.21	0.94	0.76	1.35	0.92	0.58	0.39	n.a.[a]
η	Difference	0.05	0.11	0.19	0.38	0.62	0.97	1.31	2.02	n.a.[a]
c_p	Difference	5.8	13.6	19.9	30.0	35.4	37.7	45.1	n.a.[a]	n.a.[a]

[a]n.a. means data not available.

Table 4.7.: Trends in experimental properties of **H** and **F**. It is specified in column two from the left how the properties of **H** and **F** were compared.

Furthermore, we also looked into second interaction virial coefficients from experiment[248–250] to derive bulk energies of mixing of **H** and **F** with equal number of carbon atoms. The corresponding trend comparison is depicted in the left-hand diagram of Figure 4.11. It is important to notice that all the experimental virial coefficients are negative. Hence, they were determined below the corresponding Boyle temperatures for all the systems, which is the regime of predominantly attractive interactions in the gas. In addition, the sum of the homomolecular virial coefficients is for the entire range of carbon chain length considered more negative compared to the heteromolecular values. This shows that attractions between **F** and **H** are unfavorable compared to interactions of **F** with themselves

and **H** with themselves. In addition, Knobler and coworkers[248–250] also pointed out that the intermolecular interactions in mixtures of fluorocarbons and hydrocarbons are lower than the corresponding estimate from the geometric mean of the homomolecular interactions by about 10 %. We tested, whether this is also reflected in the computed BIEs and found that, on average, results from CCSD(T)-F12 energies, from the London dispersion interaction model and from the R^{-6}-weighted distances show also about the same deviation from the geometric mean.

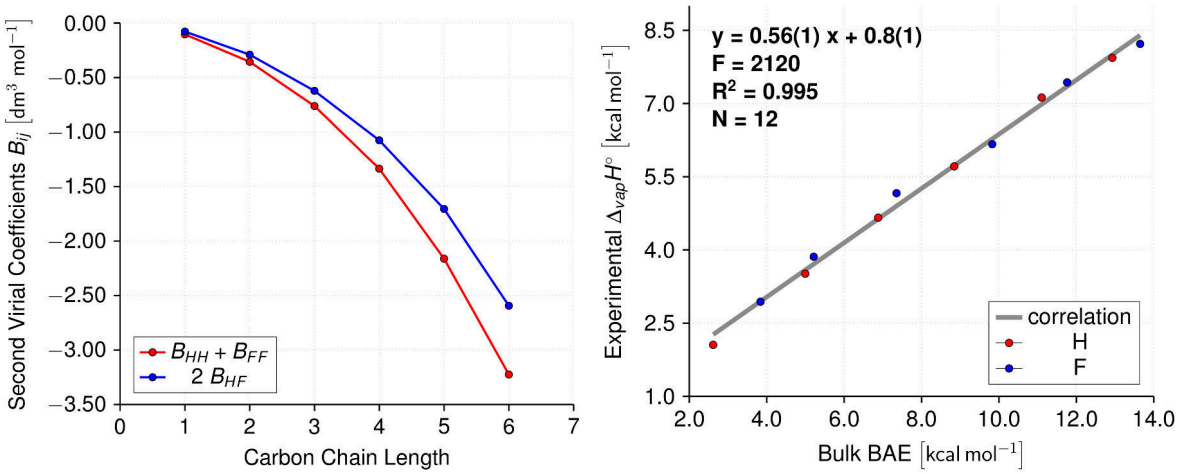


Figure 4.11.: Left: Trend in experimental second virial coefficients of heteromolecular and homomolecular interactions plotted against number of carbon atoms. Right: Bulk BAEs of **H** and **F** compared to experimental enthalpies of vaporization.

As a final comparison of our computed results and experimental bulk properties we compared the experimental enthalpies of evaporation to the estimated bulk BAEs. The corresponding results are shown in the right-hand diagram of Figure 4.11. It should be noted that it was ensured that the experimental enthalpies of evaporation were taken at as comparable temperatures as possible and lower rather than higher temperatures. We find an excellent linear correlation between the computed bulk BAEs and the experimental enthalpies of evaporation, which suggests that the bulk BAEs determined are a good measure of the corresponding strength of intermolecular interactions. We would like to point out, in

particular, that it was not necessary to use separate linear regressions for both **H** and **F**. Notably, we completely neglect entropy in this comparison and only look at enthalpies.

4.3. Discussion

Summarizing our findings results in the four main conclusions of this chapter:

(i) Several computational methods, which are expected to be good for dispersion-bound complexes, show significant deviations from benchmark energies for hydrocarbon-fluorocarbon complexes and for fluorocarbon-fluorocarbon complexes.

(ii) London dispersion in fluorocarbons originates largely from fluorine atoms.

(iii) Intrinsically, fluorocarbons have a higher dispersive interaction ability than analogous hydrocarbons. However, fluorocarbons have ground-state geometries, which are non-optimal for intermolecular interaction energies with either hydrocarbons or fluorocarbons.

(iv) The trends of experimental bulk properties of fluorocarbons and hydrocarbons as the carbon chain is elongated can be rationalized by the interaction geometries becoming more increasingly unsuitable and by the corresponding increase in relative molar carbon content.

In this section these four claims will be investigated separately in detail.

4.3.1. Benchmarking

In the past two decades, the accurate modeling of NCIs including London dispersion has been among the main areas of interest in computational chemistry.[3,4,26,251–253] For DFT calculations, several methods to correct for London dispersion have been devised[3,4] so that interaction energies of dispersion-dominated intermolecular complexes can be computed with good accuracy. Many of these methods have been benchmarked extensively and

demonstrated to reproduce the dispersive interactions of ”normal” organic molecules quite well.[44,84,127,177,252,254–261] In the context of this chapter, we used the D3 correction,[45,52] the XDM correction,[47–51] the dDsC correction,[53,54,262] the VV10 correction,[62] Minnesota functionals[42,43] and the vdW-DF10[61] functional for computing intermolecular interactions of **F** and **H** and compared the corresponding results to benchmark results at the CCSD(T)-F12/VDZ level of theory. In methods based on WFT, CCSD(T) extrapolated to the CBS limit is regarded as gold standard to treat systems with NCIs, especially systems with London dispersion.[26,230,260,263–267] In recent years, the introduction of local WFT methods has led to several low-scaling computational approaches, which are suitable to treat larger systems. DLPNO-CCSD(T) is one particular example in that regard and it was developed to provide accuracy close to canonical CCSD(T) with significantly reduced system size scaling, which makes it attractive for molecules with about 20–200 atoms or more.[30–34,36,37] In the course of this study, we were surprised that BIEs of intermolecular complexes of fluorocarbons have not been studied computationally in great detail.[229–232,268] We could not find any comprehensive systematic benchmark data for linear fluorocarbons with various chain lengths. Therefore, we provided this data in this chapter because we needed it to gain insight into the trends of dispersion and fluorocarbons compared to hydrocarbons. The benchmark set that is closest to the interactions that we are interrogating includes a series of small halogenated molecules with the most similar intermolecular complex being between methane and trifluoromethane.[265] We could not find any CCSD(T) benchmark energies at the CBS limit of dimers of fluorocarbons with more than one carbon atom, presumably due to the large computational cost.

To carry out our benchmark calculations, we chose CCSD(T)-F12, which is an explicitly correlated method providing energies very close to the CBS limit even with relatively small basis sets. In our case, the cc-pVDZ-F12 (VDZ-F12) basis set gave good results, essentially identical to CCSD(T)/CBS, and we chose CCSD(T)-F12/VDZ-F12 as most efficient benchmark method, which also allowed us to cover most of the model complexes under study. Explicitly correlated coupled cluster methods[24–28] currently provide the computationally most efficient methods to obtain benchmark quality results. Notably, using commonly employed

approximate approaches to obtain CBS extrapolations of CCSD(T) single points with cheaper methods like MP2 or LPNO-CEPA/1 did not yield good results for **HF** and **FF**. This is particularly important as these are methods that are commonly used for benchmark calculations of somewhat larger molecules.[265,266] Extrapolation with MP2 is much more common than the recently proposed improved extrapolation scheme with LPNO-CEPA/1.[269]

We were surprised, that numerous computational approaches do not provide accurate results for the BIEs of **FF** and **HF**. We would like to point out that even DLPNO-CCSD(T), which is regarded as excellent trade-off between cost and accuracy, is not able to yield very accurate results using the standard set of thresholds. For **FF**, TightPNO parameters did not provide satisfactory results and using tighter parameters only resulted in very slow convergence to our benchmark results. The deviations in the DLPNO-CCSD(T) calculations likely come from the very significant contribution of the so-called weak pairs, which are not considered further in the CCSD step and only calculated at the Local-MP2 level of theory.[33] It has already been pointed out before that for certain systems tighter parameters than the usually recommended TightPNO settings are required.[35,126,270] The downside of using tighter parameters is that the computational cost increases significantly making DLPNO-CCSD(T) a much less attractive method to obtain accurate results. We required ExtremePNO parameters, which make the method very expensive, and we still found a systematic deviation of 9 % for **FF**. This demonstrates that there is still room for improvement in efficient and accurate computational methods to describe dispersion-bound molecular systems.

We also inspected all the DFT results in more detail. It is interesting to point out that M06-L shows acceptable performance over all the intermolecular complexes studied, but larger deviations for **HF** and **FF** were still found. All the DFT methods in conjunction with the dDsC correction show very functional dependent deviations. The direction of the deviations is also not uniform, making it hard to assess the potential problems. The D3 corrections provide good results for the **HH** subset but it underestimates the interactions in both **HF** and **FF** very significantly, and the biggest deviations are observed for **FF**. It has been pointed out before that the standard parameters of D3 have troubles to reproduce the interaction energies

for **1-FF**.[229] In order to assess the impact of potential systematic uncertainties in the D3 corrections, we employed the recently developed BootD3[271] program. Uncertainties in the BIEs were generally below 10 % indicating that the deviations from the benchmark values cannot be solely explained by accumulation of uncertainties. The most likely reason for the failure of D3 is that changes in dispersion coefficients due to redistribution of electron density in the molecule are not considered. The polarized C-F bond results in significant transfer of polarizability from carbon to fluorine. The better performance of XDM compared to D3 can be explained because it does account for these changes in dispersion coefficients due to electron density deformation. The neglect of this effect in D3 is the reason D4 was developed recently as its successor.[55,272] However, when this study was carried out, D4 computations could only be carried out in conjunction with the semiempirical GFN2-xTB method.[167] The corresponding results for all our model compounds did improve compared to GFN-xTB but systematic deviations were still observed. Hence, the performance of D4 in conjunction with commonly employed functionals remains to be tested. VV10 provides the best results of all dispersion corrections in the B97M-V and ωB97M-V methods. However, with BLYP and B3LYP it significantly overestimates dispersion in **HF** and **FF**. Hence, the VV10 correction needs to be included in the DFT parametrization procedure for best results.

In addition, it is also important to discuss the geometries used for our benchmark calculations. We selected DSD-PBEP86, which is a double-hybrid functional, based on its good performance for common organic molecules.[233] While it is not among the best methods for the BIEs in our benchmarks, our rigid energy scans using DLPNO-CCSD(T)/cc-pVQZ suggests that the geometries of all complexes are very close to the corresponding minima. After having the benchmark results, B97M-V is probably the better functional for our benchmark calculations as it gives significantly better BIEs and is also computationally cheaper. However, we needed to carry out all these benchmark calculations to find that out.

Summarizing the benchmark results, the best results, both in terms of computational efficiency and in terms of accuracy, were obtained using B97M-V.[64] It yields the best results for **FF** and its overall performance is only matched by DLPNO-CCSD(T) with extremely tight thresholds. A similar performance is pro-

vided by ωB97M-V, but since it is a hybrid functional, it is also significantly more expensive. Next in line come DFT methods with the XDM dispersion correction, the best performer being B3LYP-XDM, which are excellent for **HF** and still acceptable for **FF** despite significant deviations. The problem of XDM is that it still understimates dispersion in **FF** compared to **HH**. Summarizing all our benchmark results, many computational methods, both DFT and WFT approaches, have problems to reproduce the benchmark BIEs for **HH**, **HF** and **FF** simultaneously and consistently. They all show distinct problems, which makes it hard to provide one universal rationalization for all the deviations observed. Most of the methods tested provide good results for **HH**. However, this is not surprising as some of the corresponding interaction energies have already been tested before and included in benchmark sets which are used to parametrize the methods or the dispersion corrections employed. In contrast, both **HF** and **FF** are much more problematic. One of the reasons is likely the absence of good benchmark data for these types of systems. Hence, we believe that the benchmark data set developed in the context of this chapter will be valuable for benchmarking computational methods to treat NCIs in the future. They can contribute to the development of cheap and efficient approaches to treat dispersion-bound systems. For our purpose, the benchmarking was necessary to establish reliable trends and have a solid foundation for consecutive EDA analyses.

4.3.2. Energy Decomposition Analysis

Decomposing interaction energies is one of the most popular approaches providing insight into molecular structure and chemical reactivity.[114] SAPT is among the more readily applied approaches because it provides decomposition of weak intermolecular interaction energies into physically well-defined and ready to interpret energy components.[115,116,237,238] It partitions BIEs into induction, exchange, electrostatics and dispersion. LED is an alternative approach, which has been introduced recently, and is based on DLPNO approaches.[75,239] Its main advantage is that it can offer BIEs and its components at the local CCSD(T) level of theory, at the expense of somewhat more difficult to interpret energy components. Due to differing theoretical foundations, only the dispersion energy components obtained from SAPT and LED are directly comparable. In recent years, to com-

plement EDA, numerous approaches for the visualization of interaction energies, their components and their origins have been published. In that regard, PolaBer,[242] NCI,[77,78] DID[79] and DORI[273] are only a few examples. Their advantage is that molecular structure and the corresponding physical consequences can be more readily connected by the human brain based on visual images.

We used SAPT to decompose the BIEs in all our model complexes and found dispersive attraction to be dominant. We also found Pauli repulsion to be very significant, but consistently lower in magnitude than dispersion. Notably, differences in the trends of the interaction energy components between the three subsystems are only observable when they are inspect in close detail. Dispersion is found to be most important in **FF** while electrostatics is found to be most important in **HH**. Our findings differ from previous EDA studies of complexes of alkanes and perfluoroalkanes, which found dispersion to be most important in **HH** and least important in **HF**.[219] It is interesting to note that electrostatics are largest in **HH** and even weakest in **HF**. A similar observation has been pointed out previously.[219] One might expect that the different polarities of the atoms on the outside of hydrocarbons and fluorocarbons lead to increased electrostatic attraction compared to homomolecular complexes. However, every attractive contribution between hydrogen and fluorine is accompanied by repulsive contributions between hydrogen and carbon as well as carbon and fluorine and, overall, no significant electrostatic stabilization is gained.[219]

The most important observation that we made, which is based on LED results, DID plots and PolaBer plots, is that most dispersion in **F** originates from the fluorine atoms rather than the carbon atoms, which are usually invoked as the main source of dispersion in fluorocarbons.[219] Fluorine and carbon have a strong difference in electronegativity which leads to redistribution of electron density causing *Umpolarisierung* relative to the free atomic polarizabilities. This effect is well-established in computational chemistry but the corresponding consequences are typically not stated in a concise and clear way.[3,4,46,47,51,52,55] *Umpolarisierung* can be rationalized based on Sanderson's electronegativity equalization principle.[274–278] The electronegativity difference between carbon and fluorine leads to electronegativity equalization by redistribution of electron density from carbon to fluorine until the local electronegativities are identical. Hence, fluorine bonded to

carbon will have a decreased electronegativity while carbon bonded to fluorine will have an increased electronegativity.[279] It has been demonstrated that electronegativity shows a good correlation with polarizability which leads to the conclusion that the corresponding fluorine atom has a higher polarizability compared to its free atomic state and the carbon atom has a lower polarizability compared to its free atomic state. However, in addition to that effect, the presence of the bond between carbon and fluorine leads to overall decreased polarizabilities. An alternative rationalization is based on the hardness and softness from conceptual DFT.[277,278,280,281] Hardness and softness were introduced as they provide a direct relation to the concept of hard and soft acids and bases (HSAB) and because they provide insight into basic properties of the electron density and the corresponding connection to chemical reactivity.[282,283] The equalization of electronegativity of both carbon and fluorine, because of the transfer of electron density from carbon to fluorine, results locally in a negative partial charge on fluorine and a positive partial charge on carbon. These partial charges lead to an increase in the local hardness of carbon and to a decrease in the local hardness of fluorine. The opposite is true for the corresponding localized softness. Softness shows a significant correlation with polarizability,[277,278] which leads to the estimation that the polarizability of fluorine increases and the polarizability of carbon decreases. Overall, *Umpolarisierung*, which describes the transfer of polarizability from one bonding partner to another, is expected to occur generally between bonding partners of unequal electronegativity and, hence, in all polar covalent bonds. In addition, it can be readily rationalized using well-established chemical principles.

4.3.3. Simple Interaction Models

The results of the simple interaction models provide independent support for the findings of EDA regarding the origin of dispersion. To reproduce the benchmark BIEs reasonably well, it was necessary to assume that dispersion in **H** originates mainly from the carbon atoms and in **F** mainly from the fluorine atoms. In addition, the extremely crude molecular London dispersion interaction model showcases that general trends in dispersive interactions can be estimated and understood using very approximate methods. It is particularly notable that this

approach predicts the relative interaction energies in **HH**, **HF** and **FF** very consistently, as opposed to many more sophisticated computational approaches, that we tested in the benchmark studies. With the atom-pairwise London dispersion model, we are able to reproduce the benchmark BIEs with an MAD_{rel} of only 5 % using only three adjustable parameters. This is not much worse than the best computational approaches used. Since also the partitioning into contributions of atoms gives very comparable results compared to LED, we believe that the atom-pairwise London dispersion model can provide direct insight into the trends of London dispersion in the studied model compounds. Its main drawback it that instead of determining total dispersion, it determines the effective intermolecular dispersion, which is the sum of attractive dispersion and all repulsive components. This makes the atomic pre-factors somewhat more complicated to interpret directly as both the attractive and the repulsive component will directly influence their magnitudes.

4.3.4. Interaction Descriptors

Recently, the recognition of the importance of London dispersion for the stability of molecules and molecular complexes[8,18,160] led to the advent of the concept of DEDs.[8] This also triggered a search for reliable parameters to quantify the ability of substituents and molecules to undergo dispersive interactions. Already in 2004, Dunitz suggested Q, the ratio between atomic or molecular polarizability and atomic or molecular volume, to estimate the propensity of atoms and groups of atoms to interact by London dispersion.[219] Based on his measure, he proposed that fluorine has a low propensity to undergo dispersion because the Q measure of atomic fluorine is significantly lower than the Q measure of either carbon or hydrogen. The same was true when he estimated the Q measure of **F** and **H**, **F** had a significantly lower Q measure than **H**. Dunitz proposed the low ratio of polarizability and volume of fluorocarbons to be the origin of their odd behavior.[219]

In this chapter, we reinvestigated the Q measure proposed by Dunitz and proposed an alternative definition based on the interaction volume, which is the interaction distance cubed, rather then the molecular volume. By using the interaction distance, the distribution of polarizability within molecules is treated adequately.

In addition, it accounts for the fact that dispersion is larger, the smaller the interatomic distance between the interaction partners. This is particularly important for comparison of dispersive abilities of **H** and **F**, as the largest portion of the total polarizability of **H** is on the carbon atoms, which are located on the inside of the molecule, while the largest portion of the total polarizability of **F** is on the fluorine atoms, which are located on the outside of the molecule. Therefore, the original Q measure underestimates the dispersive ability of **F** relative to **H**. In addition, the Q measure does not account for the effect of atomic ionization potential on the ability to form attractive dispersive interactions. To alleviate this shortcoming of the Q measure, we proposed the P measure, which accounts for the spatial distribution of polarizability in the molecule and for the effect of ionization potentials. Based on the London dispersion formula, larger ionization potentials and higher polarizabilities lead to stronger dispersive interactions. This is particularly interesting as the trends of polarizabilities and ionization potentials of free atoms in the periodic table are opposing and, hence, they will partially counteract. It also illustrates that only considering polarizability, which was largely done in the literature when the dispersive potential of **F** was assessed, is insufficient.

We introduced P_{tot} with respect to the vdW surface as a quantitative parameter to describe the dispersive interaction potential of molecules. Applied to **F**, it estimates linear fluorocarbons to have a higher intrinsic dispersive ability than analogous hydrocarbons. This is the opposite answer regarding the dispersive ability of **F** compared to Q measure proposed by Dunitz[219] or the modified Q measure proposed in the context of this work. Furthermore, the results of the P measure are conflicting with the general belief that **F** have a low tendency for interaction with anything.[219,223] This general belief is generally accepted in the field of fluorocarbons, even though both computational[229–231] and experimental[220,227] evidence suggests that at least for **1-HH** the BIEs and BAEs are significantly weaker than for **1-FF**.

4.3.5. Interaction Geometries

It is important to point out that P_{tot} quantifies how strong a molecule could interact by dispersion in case of ideal arrangement. However, the intermolecular interaction geometries are not ideal in **FF** and **HF** with two or more carbon atoms.

The reason for that is that the preferred geometries of **F** are not well-suited to interact with either **H** or **F**. In addition, the rigidity of **F**[196] makes it energetically unfeasible for them to deform significantly from the preferred geometries. In contrast, the preferred geometries of **H**, at least as long as the extended conformer is favored,[284] are well-suited for intermolecular interactions and no deformation is required.

In addition to geometrical differences between **H** and **F**, there is a difference with respect to what kind of interaction geometry they prefer. When two molecules of **H** interact with each other, the best arrangement would maximize the intermolecular dispersive interactions between the carbon atoms. In case of two molecules of **F**, the best arrangement would maximize the interactions between the fluorine atoms. Since the carbon atoms are in the center of the molecule and the fluorine atoms are on the outside, **HH** tend to prefer different types of intermolecular arrangements compared to analogous **FF**. This difference in preferred geometry can be illustrated based on the interaction geometries of **2-FF** and **3-FF** in comparison to **2-HH** and **3-HH**, respectively. The corresponding most favorable dispersion-bound complexes of **3-HH** and **3-FF** are depicted in Figure 4.12. It illustrates that for **3-HH** the preferred conformer has the carbon atoms of the two molecules facing each other while for **3-FF** the preferred conformer has as many fluorine atoms as possible facing each other. This difference in preferred interaction geometries also influences the ability of **F** to have ideal intermolecular arrangement as it is harder for **F** to get as many fluorine atoms as possible close than it is for **H** to get as many carbon atoms as close as possible.

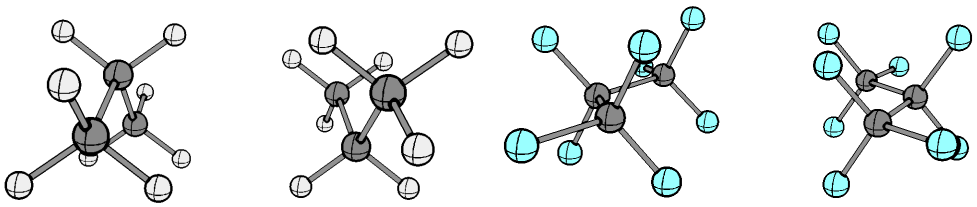


Figure 4.12.: Most stable conformers of **3-HH** (left) and **3-FF** (right).

In the context of this chapter, we introduced the number of contacts as direct

measure to assess and compare interaction geometries of noncovalent complexes. The main idea is to define that two atoms, which are separated by the sum of their respective vdW radii, have a single vdW contact. This idea is similar to the normalized contact distance, which is widely employed to analyze crystal structures with respect to intermolecular interactions.[285] One of the important differences between the normalized contact distance and the number of contacts is that we use the sixth power to account explicitly for the distance-dependence of vdW interactions, which makes it particularly suitable for the study of NCIs. First of all, it is important to notice that the number of contacts gives very close results for the complexes with one carbon atom in the interacting molecules. These intermolecular complexes indeed have very similar interaction geometries. This suggests that the newly introduced measure is suitable for comparison of interaction geometries across different classes of compounds, and does not show systematic deviations between different types of atoms. The linear scaling of the number of contacts with the carbon chain length suggests that it also scales accordingly with interaction surface area, which shows linear scaling with carbon chain length. This makes the number of contacts an attractive measure, as it is purely based on interaction geometries.

The combination of P_{tot} and $n_{contacts}$ is particularly attractive as it could be demonstrated that they can be used to estimate trends in the intermolecular BIEs in our model complexes. P_{tot} is used to assess the intrinsic strength of a molecule for dispersion and $n_{contacts}$ is used to assess a particular intermolecular arrangement. This demonstrates that these parameters have potential as quantitative parameters to assess dispersive interaction ability in molecules in general and are a first step towards a quantitative dispersion energy donor scale.

On the basis of these findings, we think that the reason for the weak intermolecular interactions in **F** compared to **H** is that they are unable to arrange in an ideal way. There is independent support for the poor arrangement in **F** from experiment. It is well-established that molecular size is crucial to determine whether a molecule is soluble in **F**. In literature, it is usually explained that the weak dispersive interactions cause bulk **F** to have cavities. This model of cavities in **F** is then used to explain the poor solvent properties of fluorocarbons for similarly large molecules but the excellent solvent properties of fluorocarbons for small,

non-polar gas molecules like O_2. Based on our findings, we propose that the cavities are formed because the ground-state geometries of **F** are unsuitable for interaction and not all the molecular surface can be arranged in an ideal way for interaction with surrounding molecules in the bulk. Therefore, we would invert the typical logical chain invoked in the literature. Rather than using the weak dispersive forces to explain the presence of cavities we propose to use the presence of cavities to explain the weak dispersive forces in **F**. Cavities can refer to surface parts that are not in close contact with another molecule in intermolecular complexes and they can refer to volume not occupied by molecules in the bulk because the 3-dimensional structure and arrangement of **F** are incompatible with higher overall volume occupation.

4.3.6. Computed Mixing Behavior

The fluorophobic effect, in short fluorous effect, is defined as the low miscibility of **F** with **H** below a critical mixing temperature.[223–226] It is well-known that the longer the carbon chain in **F**, the lower the mutual solubility of **F** and **H** becomes. It is important to note, in that regard, that typical fluorous properties are only apparent at a certain minimum carbon chain length. Attaching a perfluorohexyl group to a molecule usually is already sufficient to get at least some solubility of the molecule in fluorous solvents.[219] Another important property in that regard is the critical solution temperature of binary mixtures. Above this temperature, binary mixtures appear as homogeneous phase over the entire range of binary compositions. For **5-F** and **5-H**, the critical solution temperature is –8 °C.[286] For **6-F** and **6-H**, it is 22.7 °C.[287] Based on these experimental observations, it is fair to classify linear fluorocarbons with five to six carbon atoms as fluorous. This is important because **5-F** and **6-F** are the two longest fluorocarbons studied in this chapter and, hence, trends in the interaction properties of these compounds are relevant for the fluorophobic effect. Based on the experimental observations regarding the fluorophobic effect, it seems reasonable to suggest that the important property of a molecule is not the number of fluorine atoms but rather the number of carbon atoms bound to fluorine making a molecule fluorous.

The common explanation for the fluorophobic effect in the literature is that the low polarizability of atomic fluorine renders fluorocarbons insusceptible to

London dispersion with respect to any other molecule, and phase separation is driven by the minimization of intermolecular contact between fluorous matter and non-fluorous matter.[223] Our computational results for the relative BAEs of **HH**, **HF** and **FF** and the energies of mixing do not support this explanation. Except for **1-HF**, all **HF** complexes have lower BAEs than the analogous **HH** and **FF**. These results are in favor of preferential homo-aggregation being the reason for phase separation rather than **F** being indifferent about interaction with any other molecule.

Based on our results, we can look into even more detail into the mixing properties of **H** and **F**. The mixing behavior is the sum of two distinct effects, the first being the effect of the interaction geometries and the second being the effect of the dispersive ability of the molecules. With respect to the interaction geometries, **H** has a good interaction geometry only with itself, while **F** has poor interaction geometries with both **H** and with itself. For mixing properties, it is important to compare the difference between the mixed state and the unmixed state. In the mixed state, there are two combinations of **H** and **F**, which both have poor interaction geometries. In the unmixed state, there is **HH**, which has a good interaction geometry, and **FF**, which has a poor interaction geometry. Overall, one good and one poor interaction geometry is preferred over two poor interaction geometries. The result is that **F** has indifference for interaction with anything based on the interaction geometries and **H** prefers to interact with itself rather than with **F**. With respect to the dispersive ability of the molecules, **F** has a higher dispersive ability than **H**. Therefore, **F** prefers to interact with itself rather than with **H**. **H** would have a stronger interaction with **F** than with itself. However, since the interaction energy is the product of two P measures, the combination of **FF** and **HH** is preferred over two interactions between **H** and **F**. The result is that **F** prefers unmixing because of preferential interactions with itself. Both contributions to the total mixing behavior result in the preference for phase separation, even though the reasons are very different. It should be noted that these arguments are solely based on enthalpy. They compete with the entropy of mixing, which makes mixing more favorable. It is the cause of the existence of a critical solution temperature. Below the critical solution temperature, **H** and **F** phase-separate, above that temperature, they mix and form one homogeneous phase.

At the critical solution temperature, the entropic gain of mixing is equal in terms of Gibbs free energy to the enthalpic gain of phase-separation.

The origin of the mixing behavior of fluorocarbons has wide implications as it contributes towards finding solutions for the problems associated with perfluorooctanoic acid (POFA). PFOA is used as fluorous surfactant in aqueous solution to produce fluoropolymers like polytetrafluoroethylene (PTFE)[288] and fluoroelastomers. It is also applied in paper and textile production as a dirt and oil repellent coating.[289] All its applications are based on its fluorous properties, either due to miscibility with fluorous phases or due to immiscibility with non-fluorous phases. However, the problem is that PFOA is persistent in nature because no efficient biological or chemical decomposition pathways exist under environmental conditions.[290] In addition, due to its fluorophilicity, PFOA is bioaccumulative in lipophilic environments.[291] As a consequence, PFOA has been detected in humans worldwide because it accumulates in serum.[290] This is a problem because in recent years it has been established that PFOA is carcinogenic and, among other toxicities, toxic to reproduction, to the immune system and to the liver.[288] As a consequence, the European Union decided to add PFOA to the restricted substances of REACH, which will stop the production or distribution of PFOA in the European Union by July 2020.[289] As a consequence, industry is looking for efficient substitutes for PFOA, that retain the miscibility properties as much as possible but avoid the toxicity properties. The new understanding of fluorous phases established in this chapter can contribute to finding new compounds to replace PFOA. We are currently exploring promising candidates based on these considerations.

4.3.7. Trends in Experimental Properties

In addition to studying the mixing properties, we looked into the trends of various experimental bulk properties of **H** and **F** as the carbon chain is elongated. We observe that many of the properties studied change from being smaller in **H** to being larger in **H** as size increases, or *vice versa*. Two of the properties studied do not show an inversion. One of these is the molar heat capacity. The main reason it was included in this study was that a previous study proposed the presence of an inversion.[292] However, with new experimental data we could not confirm the

existence of an inversion, which was proposed previously. The second property lacking an inversion is viscosity, which can be explained as coming from differences in molecular shape of **H** and **F** rather than differences in intermolecular interactions.[293] All properties for which an inversion was observed quantify the strength of intermolecular interactions (surface tension, critical parameter ratio, boiling point and enthalpy of evaporation) or the change of interaction energy with volume (compressibility, bulk modules and internal pressure). Both of these classes of properties are related to the intrinsic strength of dispersive interactions in the respective molecules and the corresponding interaction geometries.

The only inversion of experimental properties of linear **H** and **F** pointed out before is the boiling point inversion, which was published already in 1947.[227] This paper reports the physical properties of fluorocarbons studied in great detail during World War II in the course of the Manhattan Project. Until now, the observed boiling point inversion could not be satisfactorily explained. The inversions of experimental properties of **H** and **F** that we observed can be rationalized by two synergistic trends. First of all, as the carbon chain is elongated, the relative molar carbon content in both **H** and **F** increases. In **H**, most dispersion comes from the carbons and the C-H bonds. In **F**, most dispersion comes from the fluorines. Therefore, the dispersive interaction potential of **F** relative to **H** decreases with the carbon chain length and we could directly show that based on P_{tot}. On the other hand, the interaction geometries of **F** become increasingly unsuitable as the chain length increases. The interaction geometries of linear **H** remain essentially unchanged. Therefore, the interaction geometries of **F** relative to **H** become poorer as the carbon chain length is increased, which we could demonstrate based on $n_{contacts}$. This explains also why all of the properties with inversion indicate stronger interactions in **1-F** compared to **1-H**, as **1-F** can still arrange in an ideal way, and it intrinsically has a higher ability to form attractive dispersive interactions. As the chain is elongated, the advantage in dispersive interaction potential of **F** over **H** becomes smaller and the interaction geometries in **F** get increasingly poor.

4.4. Conclusions

In this chapter, we outlined an extensive study of the intermolecular interactions of alkanes and perfluoroalkanes. Surprisingly, we find that many well-established computational approaches tailored to describe dispersion accurately have problems to reproduce the benchmark energies of perfluoroalkane-containing complexes well. In addition, we find that the most efficient method to obtain accurate BIEs and BAEs of fluorocarbons, with sufficient accuracy and uniform treatment of **HH**, **HF** and **FF**, is B97M-V. We also provide a first systematic benchmark dataset for computational methods based on perfluoroalkanes, which can be used to test new approaches for the description of dispersion in the future and can also contribute to force field parametrization of perfluoroalkanes.

We also established that in fluorocarbons, the biggest part of the intermolecular dispersion originates from the fluorine atom. The reason for that is threefold. First, the fluorine atoms are on the outside of the molecule and, hence, they are closest to any interaction partner. Second, due to the electronegativity differences of carbon and fluorine, redistribution of electron density towards fluorine results in *Umpolarisierung*, which makes fluorine in fluorocarbons more polarizable. Third, fluorine has a higher atomic ionization potential which results in stronger dispersion. In addition, we introduced a measure termed P_{tot} to quantify the dispersive interaction potential of any molecule and applied it to alkanes and perfluoroalkanes. We find that linear perfluoroalkanes intrinsically have a higher potential to undergo dispersive interactions than analogous alkanes. However, as the carbon chain is elongated, perfluoroalkanes have poor interaction geometries which diminishes the intermolecular interactions. The consequence is that, except for perfluoromethane, the interaction energies in the longer analogues up to six carbon atoms are comparable in alkanes and perfluoroalkanes. Another consequence is that several physical properties of perfluoroalkanes depend on the carbon chain length and, in particular, that fluorous properties only emerge at longer chain lengths. We propose that the poor intermolecular arrangement is the origin of the general notion that perfluoroalkanes interact weak with anything. Furthermore, our findings contribute to the understanding of London dispersion in molecules in general. The introduction of P_{tot} is the first step towards a quan-

titative dispersion energy donor scale, something that is highly desired in the field.[8,18]

Furthermore, our results provide new insight into fluorous phases as they suggest a mechanism for the transition of a non-fluorous molecule to a fluorous molecule as fluorine atoms are introduced. The fluorine atoms introduced will continuously change the preferred ground state geometry of the molecule. After a certain threshold is reached, the molecular shape is changed in a way which does not allow efficient packing anymore and, hence, promotes phase-separation from non-fluorous molecules due to indifference for interaction. In addition, the fluorine atoms introduced prefer to aggregate with other fluorine atoms rather than with non-fluorous matter due to its high intrinsic dispersive ability. Again, after a certain threshold of fluorine content, phase separation is favored. The synergism between these two properties causes the fluorophobic effect.

5. Universal Descriptor of Dispersive Potential

This chapter is adapted with permission from "A Universal Quantitative Descriptor of the Dispersion Interaction Potential," Robert Pollice, and Peter Chen *Angew. Chem. Int. Ed.* **2019**, *58*, 9758 – 9769. Copyright © 2019 John Wiley and Sons.

5.1. Introduction

The concept of DEDs is getting more attention in the literature. Recently, it was used to stabilize hexaphenylethane derivatives with unusually long C–C bonds,[7,158–160] to isolate Breslow intermediates[294,295] and to improve ligands for CuH-catalyzed hydroamination reactions.[296,297] However, there is still no quantitative scale to assess the ability of a group or a molecule to participate in dispersive interactions. If such a scale could be introduced, the rational design of dispersion would greatly benefit from it and could lead to numerous applications in supramolecular chemistry and catalysis. In addition, a robust descriptor for dispersion would also be of interest for applications making extensive use of multiple linear free energy relationships (MLFER) and statistical models to streamline catalyst screening and mechanistic studies.[298–306]

Investigating parameters for dispersion that have been used so far, molecular polarizability was used on a few occasions but did not enjoy widespread use.[158,301,307,308] In addition, Schneider derived dispersion increments for a number of substituents from experimental data in water, which showed a good correlation with the respective polarizabilities.[309] However, a comparable correlation is also found with respect to the surface area of the groups. This suggests that the increments could merely be a measure of substituent size. In

addition, Sigman and coworkers developed interaction parameters of aromatic substituents, which are based on computed interaction energies.[310] Consequently, the developed parameters were used in a few case studies to explain and predict stereoselectivity of catalytic reactions.[311,312] However, while their utility was demonstrated, these parameters are costly to compute, limited to substituents with aromatic rings and quantify not only dispersion. As discussed in the previous chapter, in 2004, Dunitz introduced the ratio of polarizability and volume, termed Q, to gauge the dispersion potential of molecules.[219] We have shown, that Q is not appropriate to quantify dispersion as it does not account for the distribution of polarizability within molecules, and it ignores the ionization potential contribution. We improved on his proposal, by specifically accounting for these effects. However, while our approach in the last chapter had potential as we successfully quantified the dispersive ability of alkanes and perfluoroalkanes, the procedure to obtain the P parameters was computationally too costly.

In this chapter, we redefine the P parameter developed in the previous chapter to be able to compute it readily for any substituent without resorting to benchmarking against interaction energies. By doing that, we obtain a universally applicable measure of dispersive interaction potential and use it for making a DED scale of the elements and of common substituents, for constructing London dispersion potential (LDP) maps and apply the newly introduced tools to numerous problems in chemistry.

5.2. Results

5.2.1. Methodology

In the previous chapter, we defined the P measure based on the London dispersion formula.[2] This made it necessary to compute both the atomic polarizabilities and the ionization potentials to be able to compute it, both of which were involved. When thinking about how to circumvent the computation of these parameters, we realized that, instead of the London dispersion formula, the formula used in modern atom-pairwise dispersion corrections could be employed to define P.

Instead of atomic polarizabilities and ionization potentials, it uses heteroatomic dispersion coefficients. The corresponding dispersion formula is shown in equation 5.1 and the alternative definition of P is provided in equation 5.2.

$$E_{disp}^{1,2} = \sum_n \frac{C_n^{1,2}}{R_{1,2}^n} \approx P_1 P_2 \tag{5.1}$$

$$P_i = \sum_n \sqrt{\frac{C_n^{i,i}}{R_{1,2}^n}} \tag{5.2}$$

In these equations, $E_{disp}^{1,2}$ is the intermolecular dispersion between atoms 1 and 2, $C_n^{1,2}$ is the heteroatomic dispersion coefficient of order n between atoms 1 and 2, $R_{1,2}$ is the distance between atoms 1 and 2 and $C_n^{i,i}$ is the homoatomic dispersion coefficient of order n of atom i. By using this definition, we avoid the pre-factors A by using the well-defined C_n coefficients, which can be readily computed for most elements based on dispersion corrections like the TS correction,[46] the XDM correction[47–49,51] or the D3/D4 correction.[52,55,56] Our method of choice to compute P was the recently published D4 model, which, based on the literature, should provide the most accurate C_n.[55,56] In addition, we introduce P_{int} as the average P of an atom or molecule with respect to the corresponding vdW surface, i.e. with respect to the isosurface of the electron density at 0.001 e Bohr^{-3},[313] as a descriptor of dispersive interaction ability of the corresponding atom or molecule. By overcoming the main problem of the initial definition of P, we could now test whether it is applicable to a broader range of systems.

5.2.2. Applications

This section details numerous applications of the P parameter for understanding the role of London dispersion.

Dispersion Scale

Our first target to apply the newly introduced parameter was to establish the trends of P_{int} in the periodic table. As the P_{int} of atoms is defined based on the

atomic dispersion coefficients and the atomic radii, the corresponding periodic trends are, accordingly, a combination of the periodic trends of the dispersion coefficients[52] and the atomic radii.[314] These periodic trends are well-known, but the net effect on P_{int} still has to be established. We determined P_{int} for all elements up to Z = 86. The corresponding results are illustrated in a periodic table in Figure 5.1.

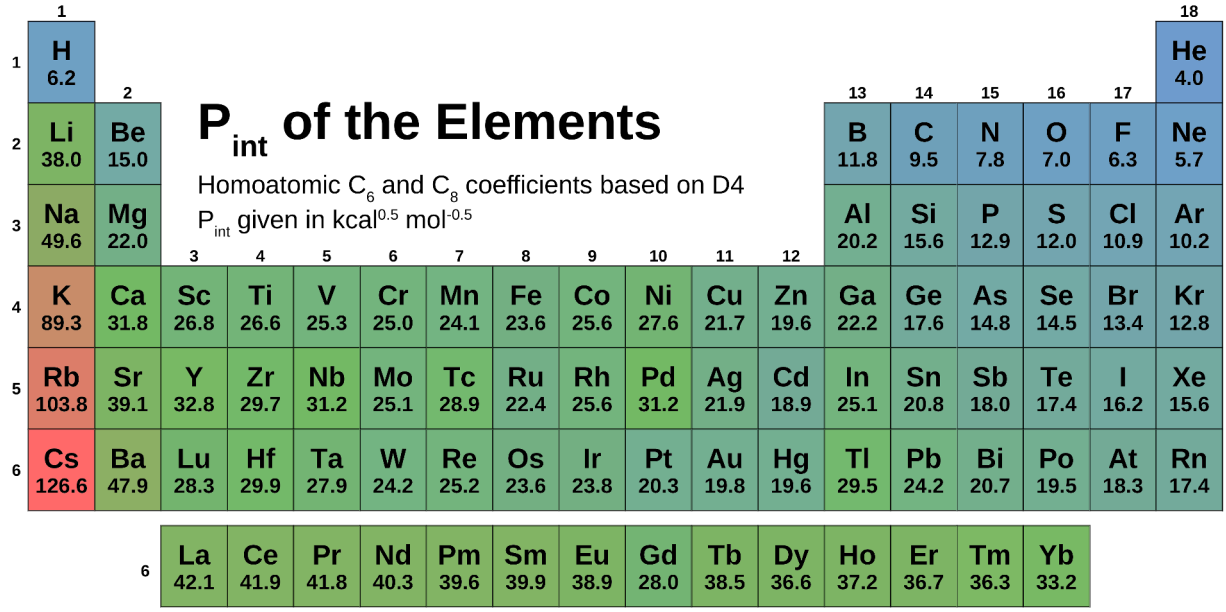


Figure 5.1.: P_{int} of elements 1–86 based on homoatomic dispersion coefficients from D4 and atomic radii taken from the literature.[314]

There are several things to be noted from Figure 5.1. First, in the main group elements P_{int} tends to decrease within a period from left to right and tends to increase within a group from top to bottom. Second, in the transition metals P_{int} also tends to decrease from left to right within a period, but within a group P_{int} does not change very much. Third, in the lanthanides the same general trend with decreasing P_{int} from left to right is observed. Overall, the alkaline metals show significantly higher P_{int} values than all the remaining elements within a given period, and except for the alkaline metals, P_{int} essentially only ranges within one order of magnitude. It is important to note that these P_{int} values are based on the electronic ground states of the free atoms and, hence, do not ne-

cessarily reflect the P_{int} values in molecules. We tried to assess the difference in the main group elements, except for the noble gases, by investigating the hydrides and fluorides of all the elements having a coordination number such that closed-shell compounds were obtained in all cases. The corresponding results are provided in Table 5.1.

We do observe significant changes, but they are less than the total P_{int} range observed in all the elements. Let us investigate CH_4 and CF_4 as exemplary cases. The P_{int} of carbon decreases from 9.5 to 7.6 in CH_4, in CF_4 it decreases to 6.4. The P_{int} of hydrogen decreases from 6.2 to 4.9 in CH_4, the P_{int} of fluorine increases from 6.3 to 7.3 in CF_4. Hence, these results illustrate that changes typically are significant but typically not dramatic. The largest change is seen for lithium. The corresponding P_{int} in LiH and LiF is much larger than in atomic Li. The corresponding change is governed by the significant reduction in size from atomic Li to Li bound to H or F. For Na the observed changes are similar but somewhat less. However, for K, Rb and Cs, P_{int} in the hydrides and fluorides is lower than in the element. The effect tends to be bigger for less electronegative elements, as illustrated by Al and Po. Most elements show an increase in P_{int} when going from the free atom to hydride or fluoride, the only exceptions here being C, K, Rb and Cs, which show a decrease for both hydrides and fluorides. This is suggestive of the volume decrease generally dominating over the decrease in dispersion coefficient when bonds are formed.

The next step was to generate a DED scale of common organic substituents. To do that, we selected typical substituents and computed average P_{int}, maximum P_{int} and volume (V_{rel}) of all the substituents. The volume was added to the table both to assess the room needed by the substituent and to test whether P_{int} correlates with volume. We also added the product of P_{int} and surface (A_{rel}) to the table. P_{int} is a measure of the dispersive strength of a single contact, the product of P_{int} with A_{rel} estimates the dispersive strength of the substituent, when all its surface is covered with close contacts to other molecules. All the results are provided in Table 5.2 and the substituents are ordered with increasing P_{int}. Notably, the volume is somewhat correlated to P_{int} but they are not interchangeable. Particular outliers are the halogens and substituents with atoms from higher periods like sulfur or selenium. We also compared our parameters to the incre-

Number	Element	Valence	P_{int}(Element)	P_{int}(Hydride) [$kcal^{0.5}\,mol^{-0.5}$]	P_{int}(Fluoride)
1	H	1	6.2	6.5	6.4
2	He	0	4.0	4.0	4.0
3	Li	1	38.0	139.8	168.2
4	Be	2	15.0	12.0	17.2
5	B	3	11.8	18.5	15.1
6	C	4	9.5	7.6	6.4
7	N	3	7.8	9.3	7.6
8	O	2	7.0	8.6	8.1
9	F	1	6.3	7.5	8.2
10	Ne	0	5.7	5.7	5.7
11	Na	1	49.6	63.1	106.0
12	Mg	2	22.0	23.5	59.0
13	Al	3	20.2	32.5	44.6
14	Si	4	15.6	17.2	21.9
15	P	3	12.9	16.0	16.9
16	S	2	12.0	16.2	16.6
17	Cl	1	10.9	15.1	17.5
18	Ar	0	10.2	10.2	10.2
19	K	1	89.3	44.0	72.6
20	Ca	2	31.8	56.5	102.2
31	Ga	3	22.2	39.9	39.7
32	Ge	4	17.6	19.4	25.1
33	As	3	14.8	19.5	21.7
34	Se	2	14.5	20.3	23.2
35	Br	1	13.4	19.2	23.3
36	Kr	0	12.8	12.8	12.8
37	Rb	1	103.8	46.4	68.7
38	Sr	2	39.1	53.4	88.5
49	In	3	25.1	35.3	41.6
50	Sn	4	20.8	23.7	32.1
51	Sb	3	18.0	23.3	25.4
52	Te	2	17.4	24.3	27.5
53	I	1	16.2	23.5	27.9
54	Xe	0	15.6	15.6	15.6
55	Cs	1	126.6	43.4	64.2
56	Ba	2	47.9	55.1	85.5
81	Tl	3	29.5	37.8	42.1
82	Pb	4	24.2	27.1	38.8
83	Bi	3	20.7	27.5	32.8
84	Po	2	19.5	29.9	38.2
85	At	1	18.3	29.0	36.2
86	Rn	0	17.4	17.4	17.4

Table 5.1.: Comparison of P_{int} of the elements with the atomic P_{max} of the corresponding hydrides and fluorides with the respective valence.

ments estimated by Schneider and found only a weak correlation between his increments and the product of P_{int} and A_{rel}.[309]

Looking in the values of Table 5.2, the most dispersive group from the selected ones is the triphenylsilyl (TPS) substituent. It has the highest P_{int} and also the highest product of P_{int} and A_{rel}. For practical use, TPS is interesting because it is relatively easy to introduce in molecules and, hence, could provide an interesting option to introduce a large amount of additional dispersion at a specific site of a molecule.

Dispersion Maps

The next idea was to use the P parameter to construct LDP maps of molecules. LDP maps are designed to be the equivalent of ESP maps for dispersive interactions and show at which position on the surface of the molecule a new interacting group would receive most dispersion. LDP maps are constructed by computing P on a 3-dimensional grid and projecting the corresponding values on the isosurface of electron density. The LDP map of benzene is depicted in the left picture of Figure 5.2. In order to benchmark the results against more sophisticated computational methods, we computed a sparse LDP map of benzene by placing helium atoms at various positions on the surface of benzene and computed the dispersion interaction component between benzene and helium using the SAPT2+(3)δMP2/aTZ level of theory[74] and the corresponding results are depicted in the right picture in Figure 5.2. The distance of the helium atom and benzene was chosen so that the overall BIE is in the attractive regime. Hence, the distance is larger than in the LDP map, for which we use the vdW surface. Therefore, for better comparison, we also computed a sparse LDP map using the distances also used for the benchmark calculations. The corresponding results are shown in the middle picture in Figure 5.2. The results from LDP and SAPT agree very well and show identical positions of local extrema. We benchmarked the LDPs from a library of small molecules including molecules with various heteroatoms and even organometallic compounds against the dispersion energies of noble gas atoms with the molecules using various methods including SAPT and LED. We found that, in general, the LDP maps agree very well with benchmarks from higher level of theory and local extrema are correctly identified.

Group	P_{int}	P_{max}	P_{int} A_{rel}	V_{rel}
	[kcal$^{0.5}$ mol$^{-0.5}$]			
-H	3.6	6.2	1.5	0.29
-F	6.1	7.0	3.5	0.43
-OH	7.8	10.1	5.5	0.62
-NH_2	8.7	12.0	7.4	0.81
-CN	9.6	11.0	10.0	1.03
-Me	9.8	12.5	9.8	1.00
-NC	10.1	12.2	10.9	1.07
-CF_3	10.3	16.5	14.1	1.48
-NO_2	10.4	15.2	12.5	1.24
-OMe	10.4	14.8	13.4	1.36
-COOH	10.9	15.4	14.4	1.41
-CCH	10.9	11.8	13.0	1.24
-$B(OH)_2$	10.9	15.6	16.1	1.62
-$CONH_2$	11.3	15.9	16.3	1.60
-Cl	11.4	14.9	10.6	0.93
-Et	11.5	16.0	17.5	1.72
-Ac	11.6	17.2	18.4	1.80
-$CHCH_2$	11.6	16.5	15.7	1.49
-CF_2CF_3	11.9	19.0	24.5	2.49
-*i*Pr	12.3	18.7	24.4	2.43
-SH	12.3	17.4	13.5	1.18
-*c*Pr	12.4	19.5	21.7	2.07
-OCOMe	12.6	18.5	23.4	2.13
-COOMe	12.6	19.9	23.3	2.12
-*t*Bu	12.9	18.4	30.6	3.13
-$CF(CF_3)_2$	13.0	23.3	34.2	3.45
-Br	13.5	19.5	15.0	1.20
-$CONMe_2$	13.5	21.8	31.7	2.95
-$C(CF_3)_3$	13.9	21.3	43.0	4.35
-SO_3H	14.6	23.3	25.1	2.01
-SeH	14.6	22.2	18.3	1.39
-TMS	14.7	23.4	42.6	3.71
Cy	14.7	20.4	42.5	4.01
-OPO_3H_2	15.1	26.1	31.6	2.51
-Ph	15.3	23.3	38.6	3.30
-OTf	15.3	26.0	38.6	3.19
-TIPS	16.3	26.8	79.7	7.78
-I	16.3	25.1	22.0	1.57
-TPS	19.2	36.4	132.9	10.49

Table 5.2.: Common organic substituents ranked by their P_{int} value.

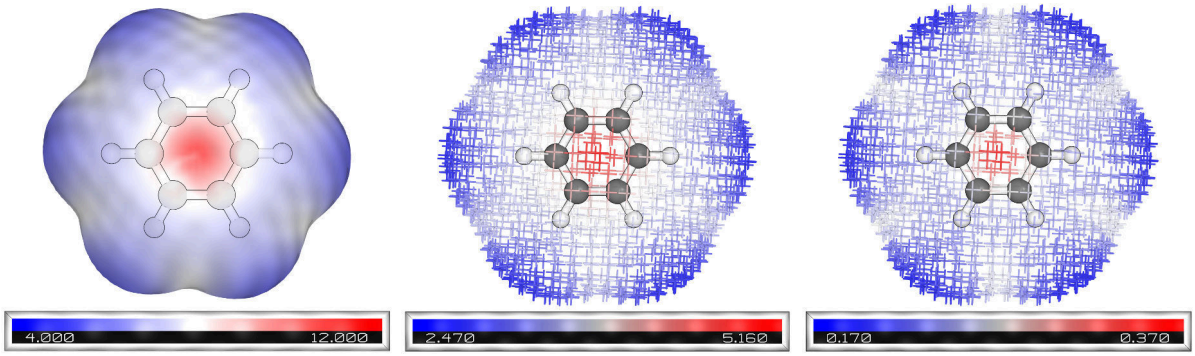


Figure 5.2.: LDP maps of benzene: Left: Full LDP map based on D4 at $1 \cdot 10^{-4}$ e Bohr^{-3}. Middle: Sparse LDP map based on D4 at $5 \cdot 10^{-6}$ e Bohr^{-3}. Right: Sparse map of dispersion energy with He atoms at points of the $5 \cdot 10^{-6}$ e Bohr^{-3} isosurface from SAPT2+(3)δMP2/aTZ.

Additionally, LDP maps can also be used to assess intermolecular interactions. To do that, the P parameter is computed at the Hirshfeld[80,315] or Becke[81] surface between the interacting fragments. As a first demonstration example, the LDP Becke map of the T-shaped dimer of benzene[316] is shown in the left picture of Figure 5.3. We also wanted to compare the LDP Becke map against well-established ways to visualize intermolecular interactions and, hence, we also computed the NCI plot[77,78] and the DID plot,[79] which are shown in the middle and right picture of Figure 5.3. While all three maps highlight the C-H–π contact as the strongest contribution to the intermolecular interaction, they do not depict the same thing. The NCI plot shows that the C-H–π contact is overall attractive but it does not reveal whether the contact is dispersive in nature. The DID plot only depicts dispersion but highlights the atoms contributing strongest to the intermolecular dispersion. The LDP map only shows dispersion and depicts the contact rather than the origin of dispersion. This illustrates that LDP Becke maps are a valuable addition to the existing methods to visualize dispersive interactions.

Case Studies

In this section, we will apply the newly developed tools to several problems in chemistry.

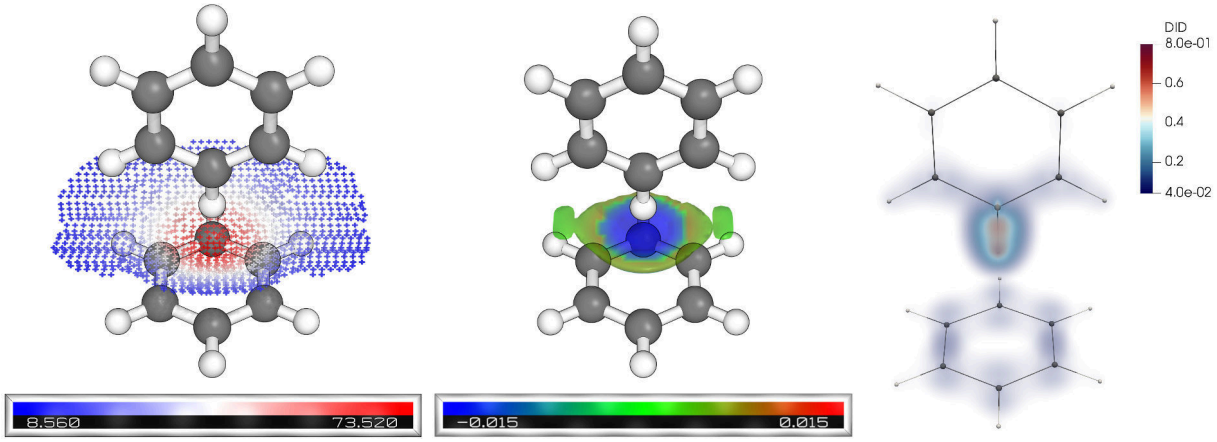


Figure 5.3.: Comparison of visualization methods for intermolecular interactions illustrated using the T-shaped benzene dimer. Left: LDP Becke map. Middle: Intermolecular NCI plot. Right: DID plot.

NCIs in Small Molecules First, we will investigate NCIs in small molecules and a first test case is the dimer of ethene. The intermolecular interaction energy is dominated by dispersion.[317] Intuitively, to maximize dispersion, one would try to bring as many atoms as close together as possible. Therefore, one would probably predict a stacked dimer to be most stable. To minimize also repulsive electrostatic components the stacked dimer could be slightly displaced. The LDP map of ethene is depicted in the left picture in Figure 5.4. It shows that the dispersive maxima are not parallel to the plane of the molecule but rather orthogonal to the plane of the molecule between the H−C=C−H moieties. Hence, based on the LDP map, to a first approximation, one would try to bring the most dispersive regions as close together as possible to predict the most stable geometry. Based on that, the most stable dimer is predicted to have two molecules of ethene with interdigitated H−C=C−H moieities. Indeed, high level WFT computations confirm this structure to be the most preferred one.[317] The most stable dimer is depicted in the right picture of Figure 5.4 together with an LDP Becke map to reveal the intermolecular dispersive contact.

Secondly, we wanted to revisit the intermolecular dispersive interactions of propane and perfluoropropane, that we already discussed in great detail in the previous chapter. We wanted to see whether the new defintion of P still yields the same results regarding the relative dispersive ability of the molecules. Indeed, the P_{int} of perfluoropropane is estimated to be larger than the corresponding va-

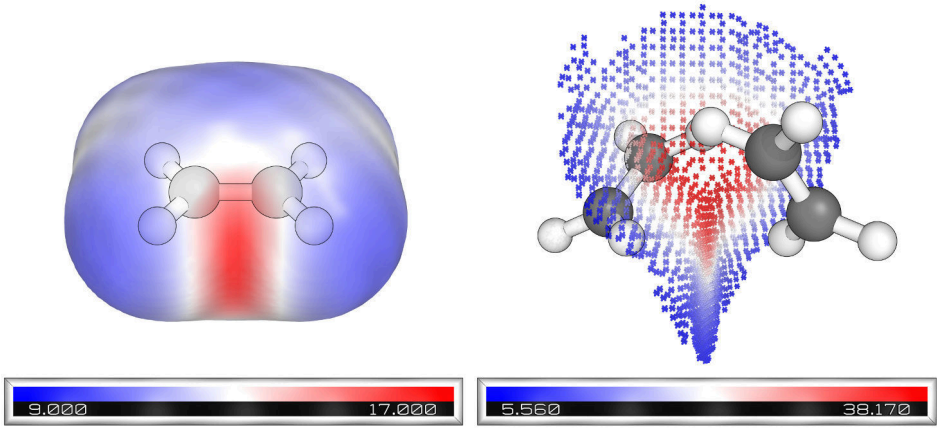


Figure 5.4.: Left: LDP map of ethene from D4 at $1 \cdot 10^{-3}$ e Bohr^{-3}. Right: LDP Becke map of ethene dimer.

lue of propane (ratio of 1.05) which is in good qualitative agreement to the findings of the previous chapter. Furthermore, we also found in the previous chapter that the homodimers of propane and perfluoropropane have distinct structures. We wanted to see whether we can understand the structures based on the LDP maps of these molecules. The corresponding results are shown in the two top pictures in Figure 5.5. The two bottom pictures in Figure 5.5 show the LDP Becke maps of the corresponding homodimers.

The LDP maps reveal that the dispersion maxima of propane are orthogonal to the plane of the carbon atoms facing the middle carbon atom. The dispersive maximum of perfluoropropane lies in the extension of the plane constructed from the three carbon atoms and is close to four fluorine atoms. This difference in the position of the dispersive maxima is also reflected in the preferred interaction geometries of the corresponding dimers. While the dimer of propane aligns, with the carbon chains facing each other, the dimer of perfluoropropane aligns four fluorine atoms of each molecule facing each other. In both cases, the dispersive maxima are aligned towards the respective interaction partner. Another interesting observation can be made from the LDP Becke maps. While the largest value of P on the LDP surface of perfluoropropane is larger than that for propane, the largest value of P on the Becke surface of the dimer of propane is larger than that for the dimer of perfluoropropane. This illustrates the unsuitable geometric arrangement of perfluoropropanes that we established in the previous chapter, as the larger dispersive potential is not reflected in larger dispersive contacts in

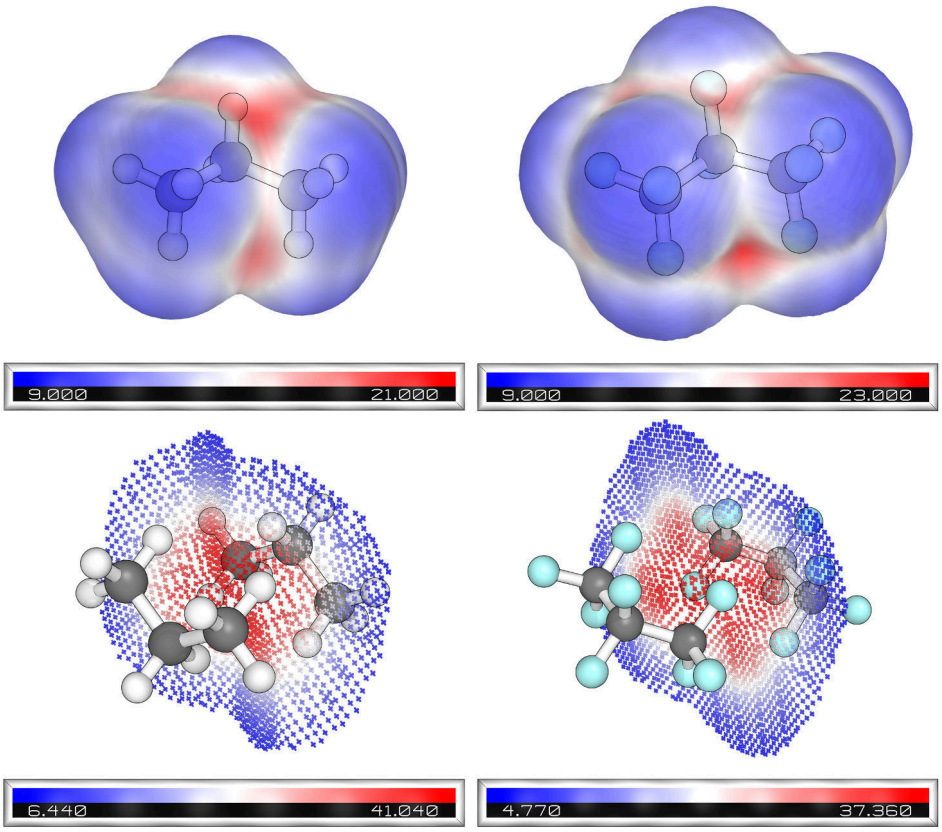
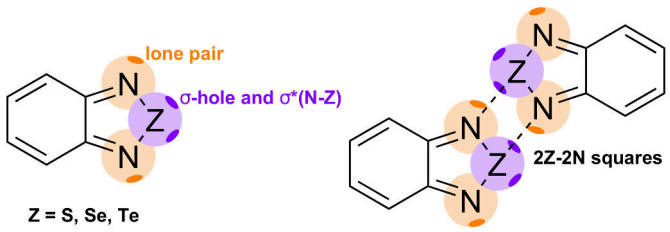


Figure 5.5.: Top left: LDP map of propane from D4 at $1 \cdot 10^{-3}$ e Bohr^{-3}. Top right: LDP map of perfluoropropane from D4 at $1 \cdot 10^{-3}$ e Bohr^{-3}. Bottom left: LDP Becke map of propane dimer. Bottom right: LDP Becke map of perfluoropropane dimer.

the corresponding dimers.

LDP maps can also be applied to the study of chalcogen bonding. There are various different types of chalcogen interactions, and for each type of motif different interaction energy components are most important. In general, electrostatics,[318] orbital interactions[319] and dispersion[318,320] are considered the most important contributors and the relative weight of them can differ significantly between different systems. Notably, electrostatics and dispersion are always attractive in these types of interactions,[321–323] while orbital interactions can be both attractive[319] and repulsive.[318] Various computational methods have been employed to study chalcogen bonding, including NBO[204,324] and EDA-NOCV[323] for orbital interactions and ESP maps[321,322] for electrostatics. However, dispersion is usually studied in less detail, mostly only by energy decomposition schemes.[204,319,323–325] Hence, we decided to use LDP maps to study 2Z-2N squares,[323,326–328] where Z denotes a chalcogen, with respect to the importance of dispersion. A typical molecule showing this type of chalcogen bonding interaction is shown in Scheme 5.1.



Scheme 5.1: Chalcogen interactions in the 2Z-2N motif (Z = S, Se, Te).

We computed the LDP maps of 2,1,3-benzothiadiazole, 2,1,3-benzoselenodiazole and 2,1,3-benzotellurodiazole, they are depicted in Figure 5.6. They illustrate the presence of local dispersion maxima between the chalcogen and the nitrogen atoms. This position is exactly in between the lone pairs of the nitrogen atoms and the σ-holes or the σ*(N-Z) orbitals. Hence, the 2Z-2N squares do not only align the electrostatic and orbital interactions in an ideal way but also the dispersive interactions are arranged to maximize them. The interaction energy of the chalcogen interactions get stronger when going from sulfur via selenium to

tellurium. Notably, this is parallel to the trend of the P_{int} of the molecules, to the trend of the local extrema on the LDP maps and to the P_{int} of the respective chalcogens in their unbound state (*vide supra*). The main contribution in the 2Z-2N squares is electrostatic in nature, but dispersion makes them an even stronger interaction motif. Hence, LDP maps can be used to study the contribution of dispersion in typical systems with important σ-hole interactions to obtain additional insight.

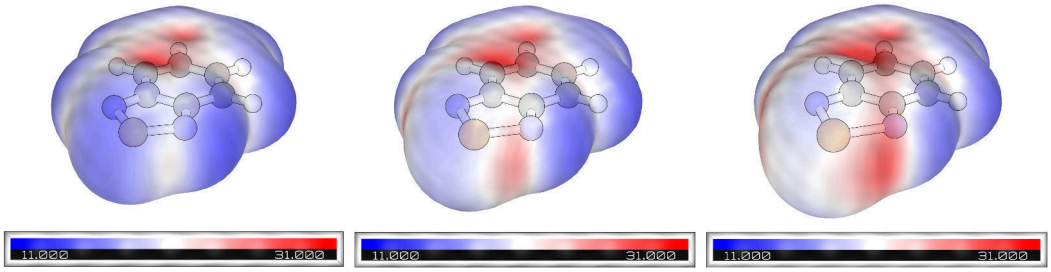
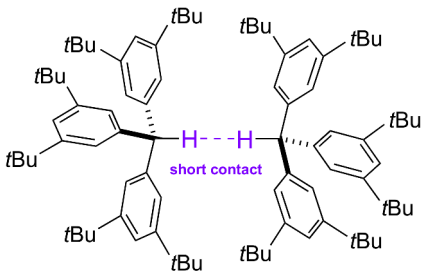


Figure 5.6.: LDP maps of 2,1,3-benzothiadiazole (left), 2,1,3-benzoselenodiazole (middle) and 2,1,3-benzotellurodiazole (right) from D4 at $1 \cdot 10^{-3}$ e Bohr^{-3}.

We were also interested to study the relative strengths of various type of dispersive contacts within one intermolecular complex using the LDP Becke surface. We selected a recently published crystal structure of a very short H-H contact in the crystal structure of tris(3,5-tert-butylphenyl)methane, which is held together by strong dispersion surrounding the short H-H contact.[329] The corresponding structure is illustrated in Scheme 5.2.

To study the relative strength of various types of intermolecular contacts in that structure, we computed the LDP Becke map of the corresponding dimer and compared it to the normalized contact distance (d_{norm})[285] projected onto the Becke surface, to the intermolecular NCI plot and the DID plot. The corresponding maps are shown in Figure 5.7.

The Becke map of the normalized contact distance shows all close contacts and does not distinguish between them by dispersive strength. The LDP Becke map highlights only part of the interactions that are also highlighted by the contact distances. According to the LDP Becke map, the strongest intermolecular



Scheme 5.2: Short H-H contact in the solid state structure of tris(3,5-tert-butylphenyl)methane.

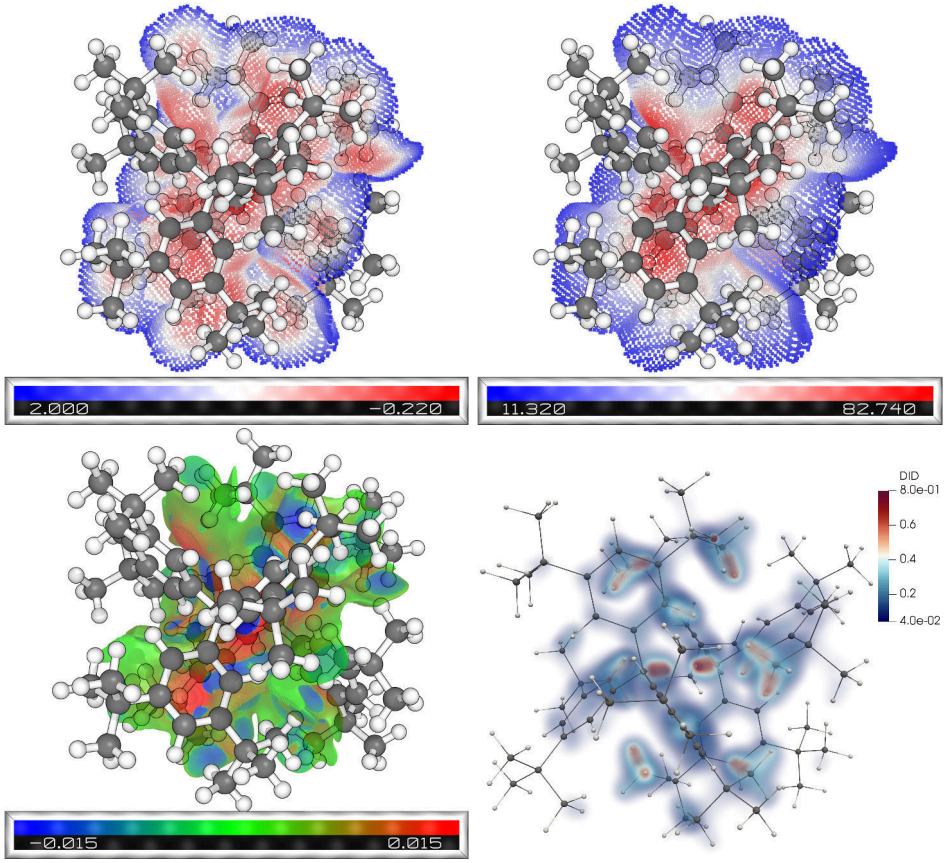


Figure 5.7.: Intermolecular interactions in the dimer of tris(3,5-tert-butylphenyl)methane. Top left: d_{norm} Becke map. Top right: LDP Becke map. Bottom left: Intermolecular NCI plot. Bottom right: DID plot.

dispersive interactions are C-H–π contacts, which are significantly stronger than contacts between two C-H moieties. In contrast, while the NCI plot highlights the same type of interactions, it is harder to assess their relative strength as the C-H–π contacts are attractive on one side of the interaction surface and repulsive on the other. In addition, as mentioned previously, it is not clear whether the particular interaction considered is dispersive in nature or not merely based on the NCI plot. The DID plot confirms that the strongest intermolecular interactions arise from the C-H–π contacts, which can be most easily seen by the highlighted C-H moieities that are in close proximity to aromatic rings. Overall, having investigated various small molecules, we could show that LDP maps are a useful tool to gain spatial understanding of dispersion.

Spectroscopy After having studied NCIs in small molecules, we moved on to a somewhat related topic, namely cryogenic spectroscopy. Typically, very weakly bound intermolecular complexes of the molecules of interest with tag molecules, which are inert and small atoms or molecules like noble gases or dinitrogen,[330–332] are isolated and either directly analyzed[333,334] or analyzed by means of infrared photodissociation,[335–339] which cleaves the weakly bound complex by absorption of IR photons. The latter technique has gained significant interest in recent years due to its ability to study the structure of reactive intermediates.[335,340–343] One of the challenges when applying this technique is the question as to whether the tag really is innocent, or whether it influences the target molecule. It has been shown that it depends strongly on the specific type of system studied.[330–332,335,341,342,344,345] To tackle this challenge, computational studies are typically used, as they can determine the position of the tag and the difference in computed spectra with and without tag can be readily estimated. Recently, the formation of intermolecular complexes of pyrrole with one or more molecules of dinitrogen was studied both experimentally and computationally.[334] In the computations, the position of tag molecules was varied systematically to determine the most stable complexes of pyrrole with different numbers of tag molecules.[334] Based on a large number of calculations including systematic structure exploration and single point calculations, it was found that the preferred binding sites of dinitrogen are the π-system of the ring and the N-H group.[334] Why

these binding sites are favored can be readily understood when both the ESP map and the LDP map of pyrrole are inspected. They are shown in Figure 5.8.

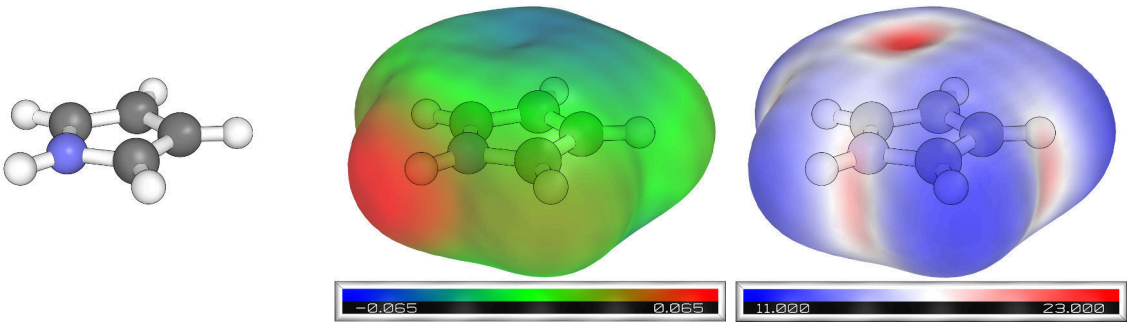


Figure 5.8.: Study of preferred interaction sites of pyrrole (left) using the ESP (middle) and LDP (right) maps at $1 \cdot 10^{-3}$ e Bohr^{-3}.

The region close to the N-H is the most positively polarized region on the vdW surface. Hence, tag molecules will interact either by electrostatic interactions or by induction at this position. The π-system has the two dispersion maxima on both sides of the ring and, hence, tag molecules will interact by dispersion at this position. Furthermore, the π-system also coincides with the most negatively polarized region on the vdW surface, therefore suitable electrostatic interactions can enhance the corresponding binding. Based on computations, it was revealed that the first and second dinitrogen molecules prefer to bind at the two sides of the π-system, so they prefer to bind at the most dispersive sites.[334] Three consecutive dinitrogen molecules bind preferentially at the N-H, which is the most positively polarized site.[334] Therefore, the preferred tag positions of molecules analyzed by cyrogenic spectroscopy can be understood and predicted using a combination of ESP and LDP maps.

Supramolecular Chemistry Another field of application of LDP maps is carbon nanomaterials, for instance the study of hosts for fullerenes and single-walled carbon nanotubes. These types of materials have found numerous applications in recent years.[346,346–353] One of the main challenges is processing[354] and purification.[355] To overcome these problems, hosts have been developed, which bind

selectively to these types of carbon nanomaterials.[356] Concave π-systems have shown particularly promising results in that respect.[356–358] They bind mainly by dispersion with non-negligible electrostatic contributions.[359] One of the main representative motifs is corannulene, which binds selectively to fullerenes.[356] Its concave center is known to be the binding site for C_{60} and C_{70} fullerenes and it is also established that the binding preference cannot be solely explained by the tendency to maximize contact area. Therefore, we generated an LDP map of corannulene to understand the reason for this binding preference. The corresponding results are shown in Figure 5.9.

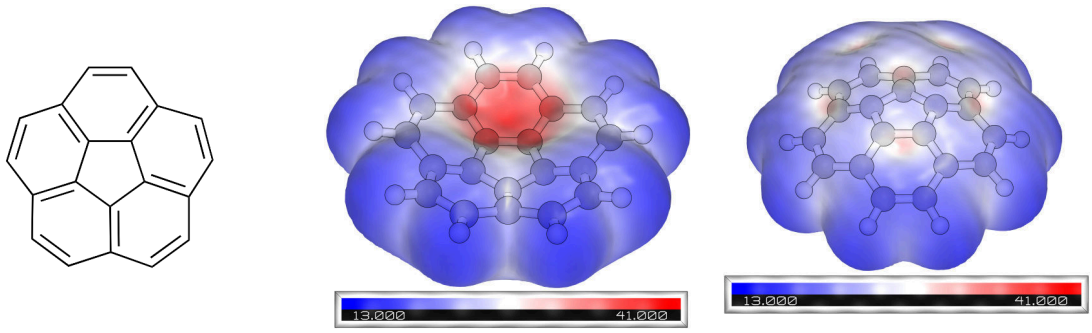


Figure 5.9.: Study of preferred dispersive interaction sites of corannulene (left) based on the LDP map of the concave (middle) and convex (right) faces at $1 \cdot 10^{-3}$ e Bohr^{-3}.

The LDP map reveals that the most dispersive site of corannulene is indeed the concave center of the molecule. The P_{max} on the concave face is 1.37 times as big as P_{max} on the convex face. This demonstrates that solely the concave shape can account for a very significant difference in dispersive binding ability. This can explain why corannulene is very effective in binding curved carbon nanomaterials. It also shows that the LDP is a simple and useful tool to understand hosts in supramolecular chemistry and design even more effective ones.

Organometallic Chemistry The next step was to apply LDP maps to metallophilic interactions.[360] We selected aurophilic interactions[361,362] as one class of d^{10}-d^{10} interactions.[363] ”Metallophilic” was first used by Pyykkö for the attraction

between closed-shell metal cations,[360] for instance Au(I) centers. The interaction was rationalized by attractive intermetallic dispersive interactions.[360 360 364–366] Initially, this stabilizing effect was estimated to reach 12 kcal/mol in stabilization,[367] but more recently it was shown to be in the range of 3 kcal/mol–8 kcal/mol with the ligands also contributing significantly.[368 369] In addition, ionic contributions to the interaction are also significant.[370] In a recent report, metallophilic dimers of Au(I)-carbene complexes observed in solid state structures were studied computationally with WFT methods and DID plots.[371,372] It was found that the interaction between the ligands are actually larger in magnitude.[371,372] Therefore, we decided to also apply our LDP maps to this class of compounds. The corresponding result is depicted Figure 5.10.

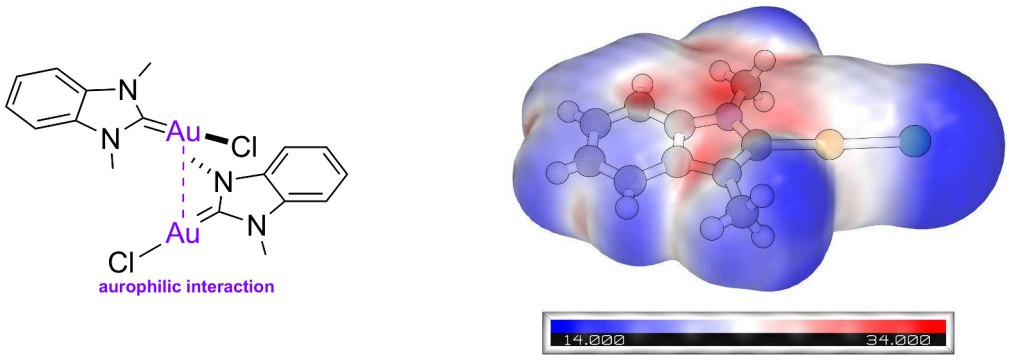
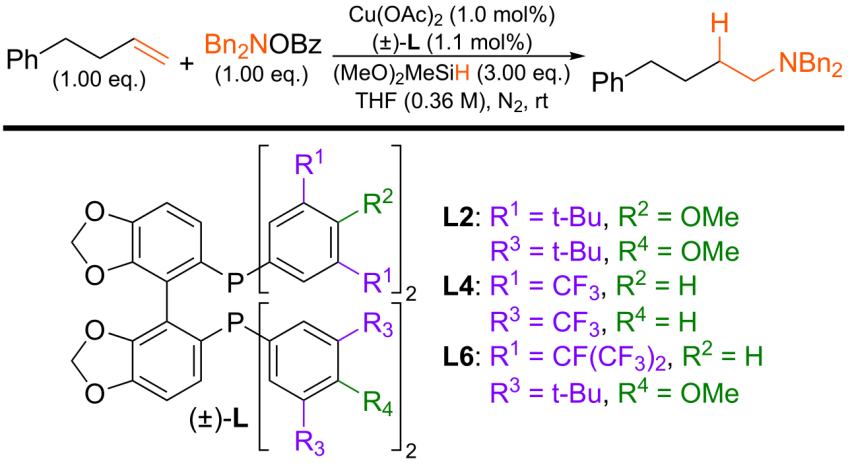


Figure 5.10.: Study of preferred dispersive interaction sites of the Au(I)-carbene complex (left) based on the LDP map (right) at $1 \cdot 10^{-3}$ e Bohr^{-3}.

The LDP reveals that the maximum of dispersion is not on the gold atom but on the pi-system of the N-heterocyclic carbene ligand. This is direct support for the findings in the literature.[371,372] Furthermore, P_{int} can be used to estimate the relative strength of Au-Au, Ag-Ag and Cu-Cu contacts. Hence, we estimate that the relative order in interaction strength is Cu-Cu $\approx$ Ag-Ag $>$ Au-Au, which is in good qualitative agreement with higher level of theory.[369 373] Therefore, the tools we introduced to study dispersion can also be used to understand matallophilic interactions.

Catalysis Last, we apply our new tools to catalysis. Recently, Buchwald and Liu investigated CuH-catalyzed hydroaminations of alkenes.[296,297] The corresponding reaction scheme is depicted in Scheme 5.3.



Scheme 5.3: CuH-catalyzed hydroamination of alkenes with typically employed ligands.

Their computational investigation was facilitated by their newly proposed ligand-substrate interaction model,[296] which is an extension of the distortion/interaction model developed by Bickelhaupt and Houk[374] and decomposes the overall ligand-substrate interaction into through-bond and through-space components. They were able to demonstrate that through-space dispersion between ligand and substrate is important for the stabilization of the hydrocupration transition state (TS), which is probably the turnover-frequency determining TS, by correlating the corresponding TS barrier to through-space interactions of various ligands.[296] Based on that model, they were able to find a new ligand outperforming all the others by improving the through-bond interaction without losing through-space dispersion.[297] Their computational studies involved costly TS optimizations using DFT. Using our newly developed tools, we want to show how this type of study could benefit from using LDP maps and the P_{int} of the ligands for carrying out rational ligand design. The first step was establishing a correlation of the ligand-substrate dispersion in the hydrocupration TS against the P_{int} of the aryl groups on the diphosphine ligand, which is depcted in Figure

5.11.

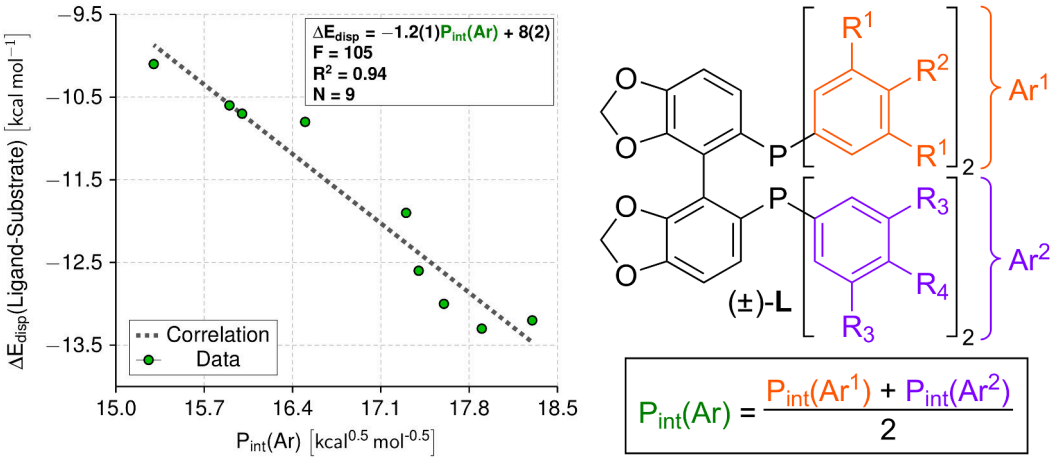


Figure 5.11.: Correlation of dispersion in the CuH-catalyzed hydroamination of alkenes between ligand and substrate in the hydrocupration TS (left) against P_{int} of the aryl groups on the diphosphine ligand (right).

The good correlation between through-space dispersion in the TS and P_{int} of the aryl groups on the diphosphine ligand suggests, that P_{int} could be used directly as a reliable measure for the interactions in the TS. It is important to point out that computing P_{int} for the aryl groups in the ligand can be done within minutes, while finding the transition states combined with EDA at the transition state takes hours to days. The difference in computation time is several orders of magnitude, which makes P_{int} very attractive for screening new potential ligands, which can than be preselected to compute the correspondings TSs for only a subset of the tested ligands. Buchwald and Liu substituted a tert-butyl group with a perfluoro-isopropyl group and found ligand-substrate dispersion to be very similar.[297] This is not very surprising as we established in Table 5.2 that these two groups have roughly the same P_{int}. In addition, the LDP Becke maps can be used to visualize the through-space dispersive interactions in the hydrocupration TS of the various ligands. This can be useful both to understand why the ligands show a strong ligand-substrate dispersion and to also get a feeling for the spatial characteristics of the interactions. The corresponding results for a subset of the TSs studied are provided in Figure 5.12.

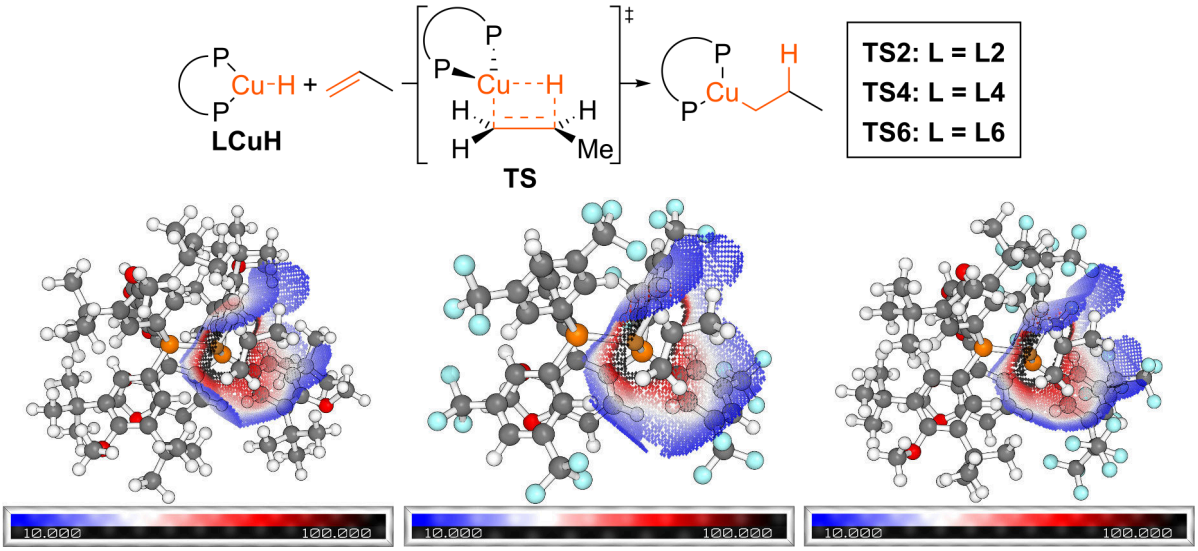
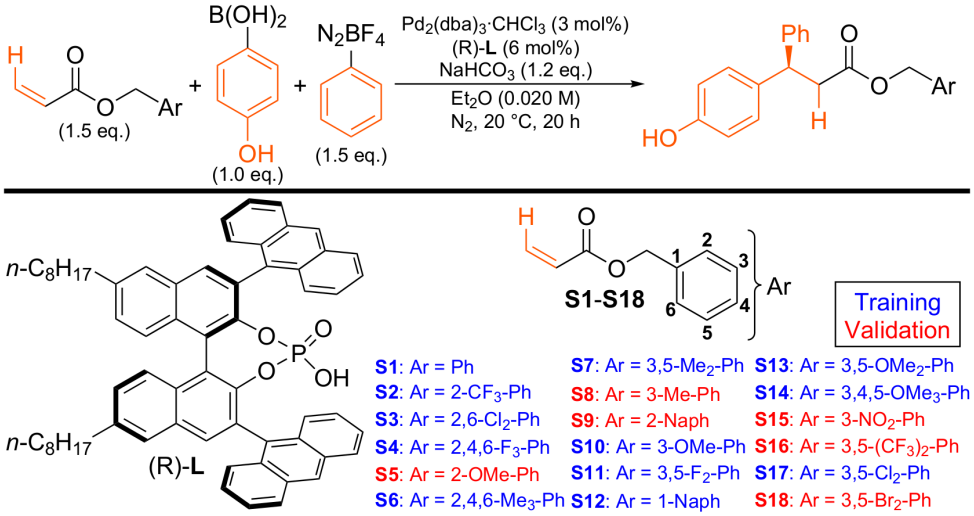


Figure 5.12.: Visualization of ligand-substrate dispersive interactions in the hydrocupration transition states using LDP Becke maps of TS2 (bottom left), TS4 (bottom middle) and TS6 (bottom right).

Figure 5.12 shows that TS4 has significantly lower ligand-substrate dispersion compared to TS2 and TS6 because the substituents on the aromatic ring do not get close enough to the substrate. In addition, it also shows that TS2 and TS6 have very similar ligand-substrate dispersive interactions and also very similar catalytic pockets. These LDP Becke maps can be used to understand the spatial requirements of effects ligands for the hydrocupration. Overall, both P_{int} and LDP Becke maps could be used very effectively for understanding and predicting dispersion in the CuH-catalyzed hydroamination of alkenes.

Finally, we tested whether P_{int} could be used in MLFER studies. Sigman established multiple linear regression for constructing models to predict stereoselectivity in asymmetric catalysis[300,302,375] and gain mechanistic insight and established the use of Sterimol parameters,[298,376] vibrational parameters[299] and atomic charges.[377,378] Lately, he also developed interaction parameter to describe NCIs of aromatic substituents[310] and applied them to study a few exemplary reactions.[310–312] Our idea was to replace the energy parameters, termed Eπ,[310] in the corresponding correlations simply with the P_{int} of the substituents and perform the same type of analysis. Therefore, we reanalyzed the Pd-catalyzed 1,1-diarylation of benzyl acrylates with varying benzyl substituents.[308,311] The corresponding reaction is depicted in Scheme 5.4.



Scheme 5.4: Pd-catalyzed enantioselective 1,1-diarylation of benzyl acrylates and the employed substrate scope for the MLFER model.

We tried replacing ${}^{S}E\pi$ with P_{int} or with the product of P_{int} and A_{rel}. Both parameter substitutions resulted in models of similar quality compared to the original models. The corresponding results are depicted in Figure 5.13. Importantly, using only A_{rel} instead of the interaction energy parameter did not give a model of comparable quality. Notably, in our models with P_{int} the data point of substrate **S4** in the training set was not an outlier. It is also important to point out that the product of NBO_{C1} and P_{int} or $P_{int}A_{rel}$ give good single parameter correlations (R^2 = 0.63 and 0.64, respectively). Hence, overall we demonstrated successfully that P_{int} can be used in MLFER studies.

5.3. Discussion

The following discussion will first deal with methodological details of the dispersive interaction potential P, its definition and relevant theoretical details. The second section will deal with potential applications and comparisons to established models.

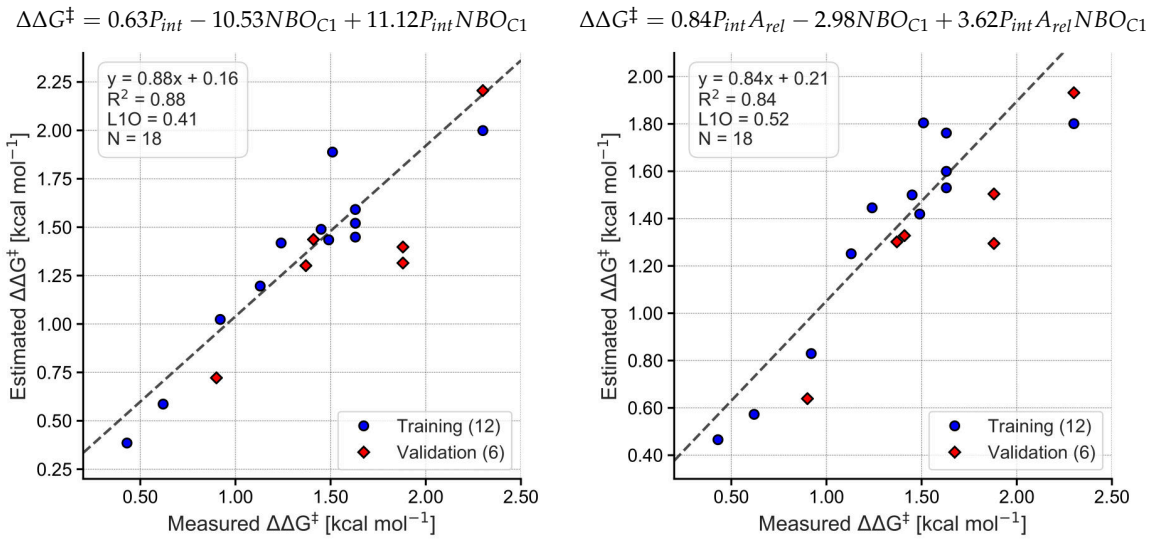


Figure 5.13.: MLFER analysis of Pd-catalyzed 1,1-diarylation of benzyl acrylates with P_{int} (left) and with the product of P_{int} and A_{rel} (right).

5.3.1. Methodology

The newly introduced definition of P is based on dispersion coefficients, which are readily available via modern atom-pairwise dispersion correction schemes. Hence, to compute P parameters, the Tkatchenko-Scheffler[46] dispersion model, the exchange-dipole moment[47–51] (XDM) dispersion model or the D3[52] or D4[55,56] model are all potential approaches that could be chosen. Therefore, our approach is not limited to one particular dispersion correction scheme but is flexible in that respect, as long as the scheme gives access to dispersion coefficients.[3,4] In the context of this work, we tested XDM, D3 and D4 to compute P parameters and found them to yield comparable results in most cases. D4 was our method of choice because it provides the most accurate dispersion coefficients of all the methods named before,[55] which is especially important for perfluoroalkanes, and the best cost benefit ratio. That was the reason it was employed for all our case studies presented in this chapter.

The definition of P is based on several simplifications. The most notable is that we are only using homoatomic dispersion coefficients, as we want to quantify the intrinsic dispersive interaction potential of an atom or molecule without knowing what the interaction partner is. When using P to estimate interaction energies by multiplying the P values of the two interacting moieties, the geometric mean

of the two homoatomic dispersion coefficients are used as an estimate of the corresponding heteromolecular dispersion coefficients. This is the approach that was used in the D2 correction and generally provides a reasonable approximation to exact Casimir-Polder integration.[3,91] Therefore, this approximation can be justified from a theoretical point of view.

In addition, we are using the sum of C_6 and C_8 coefficients to compute P. When the product of two P values is calculated, this does not only lead to C_6 and C_8 terms, but gives rise to a C_7 cross term, which depends on R^{-7}. This cross term artificially overestimates the resulting interaction. However, we also neglect higher order energy terms and do not correct for them by introducing scaling factors, which are commonly used in dispersion corrections.[3] Therefore, the overestimation is partially compensated by the neglected terms.[3] In addition, in molecules, C_7 coefficients are actually larger than zero and can contribute for several percent of the total interaction. This provides some physical justification for an effective C_7 term when P parameters are multiplied.[379] Most importantly, the additional energy contribution of this cross term is only important at short distances as it has a higher order distance dependence compared to the leading term.

It also important to note that at short interaction distances, our approach to compute the P parameter is not very accurate because the dispersion formula employed is exact only in the asymptotic long-range limit. The effect is that short-range dispersive interactions are significantly overestimated by using this formula. Since we neglect higher-order dispersion terms in our approach, the overestimation is somewhat reduced, but still not solved. Modern dispersion correction schemes employ short range damping functions like Becke-Johnson damping[47] to alleviate this problem. However, since the P parameter introduced by us is mainly designed to describe mid- and long-range dispersion quantitatively, the neglect of damping is not a big problem. For the sake of simplicity, we decided against implementing a damping scheme. However, future work could address that shortcoming and introduce damping to the computation of the dispersion interaction potential.

Furthermore, we decided to neglect many-body dispersion in our current definition of P, which removes part of the anisotropy of dispersion. In principle,

the Axilrod-Teller-Muto expression[58,59] could be added to the current definition of P. However, since it was shown that the corresponding contribution is typically only about 5 %[3] and only relevant in large molecules,[44,380] we decided to neglect it in order to avoid additional approximations, which would be necessary. Overall, despite the many approximations inherent in the definition of P, our comparisons to high level computations show that it is a reasonable quantitative descriptor of dispersive interaction potential. In the context of this work, we computed benchmark dispersion maps by placing noble gas atoms on the surface of a set of molecules and determined the corresponding intermolecular dispersion based on SAPT[115,116,237,238] and LED-DLPNO-CCSD(T)[75,239] calculations. For a wide set of molecules, the corresponding results agreed both qualitatively and quantitatively well with P, which provides good theoretical justification for the approximations used.

Another aspect meriting close inspection is the fact that while the mathematical definition of P does not include anisotropy of dispersion, it enters the equation when the electron-density based molecular surface is used for the computation of P or LDP maps. By doing that, anisotropy is introduced due to polar flattening,[381,382] an effect that makes local vdW radii of atoms anisotropic.[383] Polar flattening is a well-known physical effect[381,382] and has been demonstrated to be one of the main reasons for the presence of σ-holes in molecules giving rise to NCIs like halogen bonding.[384–387] Surprisingly, it was shown previously that isotropic atom-centered dispersion correction schemes like D3 can reproduce the anisotropy of intermolecular interactions in Br_2, BrCl and BrF as computed by higher level of theory accurately[383] because polar flattening introduces the anisotropy through the repulsive part of the intermolecular potential.[383,388] To test our approach, we compared the LDP map of Br_2 to the results from higher level of theory and found that the position of the local extrema agree well.[383] However, the maxima surrounding the center of the molecule give roughly the same P value as the maxima at the extension of the covalent bond, which demonstrates that our approach somewhat underestimates the anisotropy in the dispersive interactions of Br_2. They are likely a consequence of the neglect of the Axilrod-Teller-Muto three-body term.[383] However, the general qualitative picture is well reproduced using our approach to compute LDP maps, which shows that using

the electron density to define the vdW surface is a proper means to introduce anisotropy. It also shows that promolecular densities, which are electron densities obtained from the sum of spherically symmetric atomic electron densities of all the constituent atoms of a molecule, are not a good alternative for computing P, as that would remove all the anisotropy and lead to wrong extrema. Hence, LDP maps can be used to get a semi-quantitative picture of the anisotropy of dispersive interactions in molecules.

As a final remark, we would like to discuss the value range of P_{int}. Typical dispersion interaction energies between single atoms at distances to minimize overall interaction energies are around 10^{-1} kcal/mol–10^{1} kcal/mol.[389] However, P_{int} values of atoms and groups of atoms are typically around 10^{1} kcal$^{0.5}$ mol$^{-0.5}$–10^{2} kcal$^{0.5}$ mol$^{-0.5}$. The reason that P_{int} values are rather large compared to the corresponding dispersion interaction energies that that we use the vdW surface for their estimation. However, when two atoms or molecules interact, the interaction distance is not equal to the vdW radius of one interaction partner but it is the sum of the vdW radii of the two interaction partners. To a first approximation, when the size of the two interaction partners is assumed to be similar, the interaction distance is about twice the size of the vdW radius of one interaction partner. This would effectively lead to a correction factor of 8–16 for every P_{int} value when interaction energies are estimated. This correction factor brings the typical value range of P_{int} in good agreement to the typical dispersion interaction energies observed between atoms and groups of atoms.

5.3.2. Applications

The main idea of constructing LDP maps was to have a tool available to study dispersion in the same way ESP maps are used to study electrostatic interactions. Compared to the widely employed NCI plots and to the still relatively rarely used DID plots, LDP maps provide a different type of information. NCI plots cannot distinguish between dispersion and other types of interactions[77–79] and they reveal the portion of the intermolecular or interfragment surface that is overall attractive or repulsive. DID plots only reveal dispersive interactions and they reveal them based on the atoms and molecular fragments contributing most to the respective interactions. They highlight the atoms interacting with each other rather

than highlighting the specific interaction. Furthermore, both NCI plots and DID plots can only be used to study existing interactions. In contrast, the LDP map reveals positions on the surface of molecules, that would provide most or least dispersion in case an interacting atom would be placed at that position. That way, they provide a different kind of information and are a valuable addition to the selection of methods to study and visualize dispersion. Furthermore, LDP Becke maps provide a very similar information as NCI plots with the difference that the former only quantify dispersion while the latter quantify all weak noncovalent interactions.

An alternative method to visualize noncovalent interactions is the projection of the normalized contact distance[285] on Becke surfaces. The question was as to how different the LDP Becke map is to a mere visualization of close contacts because the interaction distance is a very important part of the estimation of LDP. Our results suggest that LDP Becke maps are a more specific indicator of strong interactions as it is not sufficient to have only one atom close to indicate a contact, but it is even more favorable to have several atoms close to have a strong interaction. Therefore, the LDP Becke maps are a simple and insightful tool to study the relative strength of several dispersive contacts in intermolecular complexes. In addition, in the cases we studied with both DID plots and LDP Becke maps, we could see good agreement between the approaches showing that the LDP Becke map also gives good quantitative estimates of the dispersive strength of the contacts. There is one additional aspect specific to LDP Becke maps that needs to be mentioned. Because it uses distances from the Becke surface to atoms rather than atom-atom distances, they are typically too short for a specific contact. When computing P parameters, this increases the weight of higher-order dispersion terms as they have a stronger distance dependence. Hence, the C_8 term is estimated to be more important than the C_6 term relative to their importance in the actual interatomic interaction terms. This makes P values on the Becke surface only semi-quantitative descriptors of the strength of particular interactions.

Furthermore, we developed a very efficient approach to compute P_{int} parameters of molecules based on widely used dispersion corrections. Our computational procedure for computation of P_{int} parameters of molecules or substituents

starts from the SMILES[390–392] code of the molecules, converts them to initial 3D structures using OpenBabel[393] and uses GFN2-xTB[166,167] for extensive exploration of conformer space and structure optimization. As a next step, the electron density is generated. We tested both using promolecular densities using Multiwfn[394] and using DFT densities at the PBE[395]/def2-SVPD[396,397] level of theory. The promolecular density approach is extremely fast but cannot account for local density differences very well. It is still useful to get quick estimates of P_{int} parameters of molecules but is not suitable to obtain good LDP maps. The DFT densities are a compromise between computational cost and theoretical rigor and provide better results compared to the promolecular densities. In addition, the DFT densities used are quite cheap and can be readily carried out for large molecules with hundreds of atoms. The last step is computing P_{int} based on dispersion coefficients from D4[55,56] and the vdW surface obtained from the electron density, which is computationally very cheap. Our approach gives access to P_{int} values of molecules with 100 or atoms within a few minutes to a few hours on a single core. Overall, the bottleneck, especially for large molecules, is the conformer exploration, even though cheap semiempirical methods are employed for that purpose. This approach makes P_{int} available at low computational cost, which is a great benefit for large scale statistical modeling. In that regard, the computational cost is higher than typical QSAR parameters but lower than modern physical organic parameters typically obtained from DFT computations, and P_{int} can be readily used for large structures.[301] Notably, computation of P_{int} is faster by several orders of magnitude compared to the $E\pi$ parameters,[310] which could be replaced by P_{int} in modern MLFER studies of noncovalent interactions, and P_{int} is not limited to aromatic substituents giving access to a much broader scope of applications.

Another notable aspect is that we used both P_{int} and the product of P_{int} and A_{rel} in the MLFER application example to describe the dispersive interaction potential of a particular substituent. It is important to point out that they describe different dispersive properties. P_{int} quantifies the average strength of a single dispersive contact with the molecule. The product of P_{int} and A_{rel} quantifies the average strength of dispersion when the complete surface of the molecule is covered with interaction partners. Hence, the latter provides a measure of how

much dispersion the group could contribute in total while the former provides an estimate of very local dispersive interactions. Table 5.2 shows that there is some correlation between these two, but in general they are not equivalent. Hence, depending on the particular system under study, one or the other parameter is likely to be a better parameter.

We would like to conclude this discussion by discussing potential applications of the newly introduced P parameter that we did not show in the context of this chapter. We envision that LDP could possibly be used to treat dispersive solute-solvent interactions in implicit solvent models. The dispersive ability of the solvent could be described with its P_{int} parameter and the solute-solvent interface could be treated as surface with the same P_{int} value. Every dispersive solute-solvent contact could be treated as interaction between a solvent molecule with its average P_{int} value and the local P value of the solute. With an additional accounting for the number of contacts between solute and solvent, this could enable a more theoretically justified way to account for dispersion compared to the surface area approach currently employed in SMD[6] or COSMO-RS.[5,82,105] However, currently, this is just a first concept idea and additional studies are required to see if it can be useful. In addition, we also think that LDP maps could have applications in fragrance chemistry. Currently, olfaction still suffers from limited understanding[398–400] and the absence of crystal structures of olfactory receptors is also a big challenge.[401] The tool of choice are structure-odor relationships using typical QSAR parameters[398] and machine learning models are also emerging to improve on the available models.[400] LDP maps or even LDP densities could serve as multidimensional descriptors of dispersion which could provide new dimensions to the descriptors and models currently employed.

5.4. Conclusions

In this chapter, we redefined the dispersion interaction parameter P introduced in the previous chapter based on dispersion coefficients in order to make it readily available and easy to compute without losing any accuracy. We established P_{int} as a quantitative descriptor of the average dispersive interaction potential of a molecule and the LDP map as a tool to study which position on the surface

of a molecule provides most dispersion. Based on these tools, we demonstrated numerous applications in the study of London dispersion including the P_{int} of the elements and the P_{int} of common substituents. In addition, we studied dispersion in various classes of molecules in the fields of spectroscopy, supramolecular chemistry, organometallic chemistry and catalysis and we carried out MLFER studies in enantioselective catalysis. The newly introduced tools are simple to compute and are a valuable addition to the computational toolbox to study chemical problems related to noncovalent interactions. We believe that they will find many more applications in the future.

6. Summary and Outlook

In this thesis, we investigated London dispersion in a wide variety of molecular systems. In all of them, it proved to be a fundamentally important energy component contributing significantly to the overall thermodynamic stability. At present, the understanding of the importance of dispersion in the gas phase has advanced to a degree that good quantitative predictions can be made with reasonable effort for most systems. However, discrepancies between experimental gas-phase results established in the context of this work on proton-bound dimers and several computational methods suggest that chemical accuracy cannot be reached yet, especially in large systems. Further work in that regard is necessary to find the origin of the disagreement between experiment and computation in the gas phase in order to reach the ultimate goal of computational methods delivering chemical accuracy for interaction energies in dispersion-dominated systems.

We made a big effort to address the challenging question about the importance of London dispersion in solution. It is clear that compensatory (dispersive) solute-solvent interactions lead to an effective compensation, which makes dispersion less important in solution than in the gas phase in general. Based on a detailed investigation of proton-bound dimer dissociation equilibria, we established that compensation of dispersion in rigid systems in solution is about 60 %–80 % at ambient temperatures and largely invariant with respect to solvent. More important than the choice of solvent is the choice of molecular structure to introduce dispersion as overly flexible systems will lead to increased compensation due to an intrinsically higher gas-phase compensation. In the near future, the established degree of compensation needs to be tested in other systems experimentally and also investigated from a theoretical point of view to understand why it is largely invariant with respect to solvent. Furthermore, the large discrepancies between computed Gibbs free energies of dissociation in solution and the experi-

mental results for the proton-bound dimers reveal very unsatisfactory systematic discrepancies, which could be traced back to the implicit solvent models used. Deviations up to 10 kcal/mol are not uncommon and trends show systematic differences. Hence, new solvent models need to be developed in order to improve on the standard computational methodologies employed in everyday computational chemistry. It needs to be seen if implicit solvent models can be improved with respect to their description of dispersion enough so that the deviations can be decreased to an acceptable level. Alternatively, explicit solvent models based on cheap MM methods or coupled with QM/MM schemes offer an alternative solution for the problem at hand. However, their computational cost is inherently significantly higher due to the high dimension of the problem. In addition, they need to be tested against our experimental benchmark data set to see whether they can provide more reliable results.

We also made a big effort to understand the (dispersive) interactions of perfluoroalkanes and the fluorophobic effect. We revealed that many current computational methods have difficulties to reproduce benchmark interaction energies for perfluoroalkanes and provided a first benchmark set for their noncovalent complexes. Some of the more recent methods were shown to perform better but, in general, interaction parameters need to be updated to properly account for dispersion in fluorocarbons. It also needs to be seen how good new dispersion correction schemes perform for this class of compounds. It is clear that further benchmarking efforts are necessary and current dispersion corrections for DFT need to be updated. In addition, we managed to propose a new model for the interactions of perfluoroalkanes in general, which helps understand their physical properties and the fluorophobic effect. One particular implication of our study is that effective replacements for PFOA could be developed, which do not contain any fluorine atoms. Further work in that direction is necessary to see whether this strategy proves effective.

The final piece of this thesis was the introduction of a simple and general descriptor of dispersive interaction potential. We demonstrated on numerous examples that it provides insight into the dispersive interaction component in a wide variety of chemical systems. It allows to study the interplay of dispersion coefficients and molecular size of atoms in molecules in general. We only carried out

preliminary studies in that regard and further work is necessary to understand the changes in dispersive ability upon bond formation. We believe that, due to its simplicity, LDP maps have the potential to become the method of choice to visualize dispersion in molecular systems. Furthermore, the growing use of dispersive interactions in catalysis can greatly benefit from simple parameters to describe dispersive abilities of molecules like P_{int}. Potential future applications could be implicit solvent models, as the use of London dispersion potential could allow the development of solute-solvent interaction terms to be defined which do not rely on a separate set of parameters but could be based directly on widely employed dispersion correction schemes.

A. Experimental Details

This section provides the experimental details for chapters 2 and 3.

A.1. Compensation of Dispersion in Dichloromethane

A.1.1. General Experimental Methods

In general, unless noted otherwise, chemicals were purchased from commercial suppliers and used without further purification. Diethyl ether (Et_2O) and tetrahydrofuran (THF) were dried by refluxing over Na/benzophenone for several hours followed by distillation prior to use. Toluene was dried by refluxing over Na for several hours followed by distillation prior to use. Molecular sieves 3 Å were purchased from Fluka and dried for 12 h–16 h at 700 °C–800 °C and 10^{-2} mbar overnight prior to use and stored thereafter in a glove-box under N_2. CD_2Cl_2 was dried stepwise over 20 % (m/V) of these highly activated molecular sieves by static drying and degassed by 2 successive freeze-pump-thaw cycles prior to use.

For analytical TLC aluminium-backed silica gel 60 with fluorescence indicator F254 was used. Flash column chromatography was performed on SilicaFlash® P60 from Silicycle (40 µm–63 µm).

High-resolution mass spectrometry (HRMS) for literature-unknown compounds or compounds with no published HRMS measurement results was performed on a Bruker Daltonics maXis by high performance liquid chromatography in combination with high-resolution time-of-flight mass spectrometry, unless noted otherwise.

Chiral HPLC was performed on a Merck Hitachi D-7000 HPLC machine equipped with Daicel Chiralpak IA (4.6 x 250 mm, 5 µm particle size) column. The column was coated with amylose tris(3,5-dimethylphenylcarbamate) immobilized on silica gel.

Isothermal titration calorimetry (ITC) was performed on a MicroCal VP-ITC. The reference cell was filled with dry and freshly distilled (distilled over CaH_2) CH_2Cl_2 and the sample cell was filled with 2 ml of a solution of pyridinium $[BArF_{24}]^-$ salt in dry CH_2Cl_2. The syringe was loaded with 300 µl of a solution of the corresponding pyridine in dry CH_2Cl_2. Injection of the pyridine solution to the pyridinium solution in the sample cell was performed stepwise until adduct formation was complete as judged on the basis of the heat evolved after every injection. The injection sequence, unless noted otherwise, consisted of an initial injection of 2 µl followed by 50 injections of 5 µl with an interval of 75 s between the injections. The initial injection was performed to minimize artefacts from syringe filling and the corresponding data point was not used in the data evaluation.[402] The initial delay before the first injection was 120 s, the stirring rate in the sample cell during measurements was 307 rpm, the reference power was 35.5 µcal s^{-1} and all titrations were performed at 25 °C.

A.1.2. Qualitative ESI-MS Measurements

Qualitative ESI-MS measurements were performed either on a Thermo Finnigan TSQ Quantum triple quadrupole mass spectrometer or on a Thermo Scientific TSQ Quantum Ultra EMR triple quadrupole mass spectrometer. Solutions of compounds were sprayed at a rate of 5.0 µl min^{-1}. CID MS were recorded at 100 µTorr pressure of Ar inside the collision cell.

A.1.3. NMR Measurements

Standard NMR Measurements for Characterization

NMR spectra for characterization purposes were recorded either on a Bruker AVIII 400 MHz, on a Bruker AVIII 300 MHz or on a Bruker Avance-NEO 500 MHz, chemical shifts are reported in ppm, using the residual peak of the deuterated

solvent as internal standard (^{1}H-NMR: $CHCl_3$ at 7.26 ppm, CD_2Cl_2 at 5.32 ppm, CD_3OD at 3.31 ppm; ^{13}C-NMR: $CHCl_3$ at 77.16 ppm, CD_2Cl_2 at 53.84 ppm). For the designation of the observed signal coupling patterns the following abbreviations are used: singlet (s), doublet (d), triplet (t), quartet (q), multiplet (m).

DOSY NMR Measurements

NMR spectra were recorded either on a Bruker AVIII 500 MHz or on a Bruker AVIII 500 MHz HD spectrometer, equipped with a BBFO broadband probe, respectively. NMR measurements were carried out without spinning. The time between midpoints of the gradients (Δ) was chosen as 49.9 ms. The gradient length (δ) was optimized for each sample and it was also optimized for each species present in the sample, when necessary due to big differences in the diffusion constants. Standard DOSY processing, as supplied by the vendor, was used based on peak integrals. The spectrometers were calibrated using the diffusion coefficient of HDO in D_2O obtained from an NMR sample of D_2O.

Samples in CD_2Cl_2 were prepared under dry conditions equivalent to variable-temperature NMR samples (vide infra). Samples in d^4-MeOH were prepared under air and the solvent was used as provided by commercial suppliers without prior drying.

The DOSY NMR measurements were evaluated on the basis of the following equation.[403] The corresponding variables are explained in the table below.

$$ln\left(\frac{I}{I_0}\right) = -(\gamma\delta)^2\, G^2\, \left(\Delta - \frac{\delta}{3}\right)\, D \tag{A.1}$$

G refers to the unit Gauss, a non-SI unit for the magnetic field strength that is still widely used, and also by the NMR software we used. Equation A.1 can be used to linearize experimental data and estimate the diffusion coefficient on the basis of linear regression of the transformed data set. In order to do that, the natural logarithm of the relative intensity is plotted against G^2. This approach was widely used in literature,[403] however, should not be used anymore because experimental data points have different weight on the result of the regression when linearized first. Equation A.1 can be rearranged to give the following equation, which is better suited for nonlinear regression. For nonlinear regression one sim-

Variable	Definition
I	Signal Intensity [a.u.]
I_0	Initial Signal Intensity [a.u.]
$\frac{I}{I_0}$	Relative Signal Attenuation
γ	Gyromagnetic ratio of the NMR nulecus, ^{1}H: $2\pi \cdot 4.258 \cdot 10^3\ G^{-1}s^{-1}$
δ	Gradient length [s]
Δ	Delay between the midpoints of the gradients [s]
G	Gradient strength [G m^{-1}]
D	Diffusion coefficient [m^2 s^{-1}]

a.u. = arbitary unit.

ply uses the values of relative intensity as a function of the gradient strength and equation A.2 as fitting function.

$$\frac{I}{I_0} = e^{-(\gamma\delta)^2\, G^2\, \left(\Delta - \frac{\delta}{3}\right)\, D} \tag{A.2}$$

We evaluated all the DOSY measurements with nonlinear regression and derived the diffusion coefficient as average over all the values obtained from different peaks in the ^{1}H-NMR spectrum. To be able to compare our DOSY results to literature values, we used the diffusion constant of HDO in D_2O as reference. We determined the corresponding value using our experimental setup and derived a correction factor for our experimental diffusion coefficients by comparing our result to a literature value of $D_{HDO} = 1.9 \cdot 10^{-9}\, m^2 s^{-1}$.[404] However, since we found the exact same value, no correction factor was necessary. In order to compare diffusion coefficients for molecules dissolved in different solvents, the hydrodynamic radius r_H of the molecules can be estimated on the basis of the Stokes-Einstein equation (Equation A.3).

$$D = \frac{k_B T}{6\pi\eta r_H} \tag{A.3}$$

The hydrodynamic radius r_H can be determined from the diffusion coefficients D in different media and the corresponding results can be directly compared. By comparing the effective size of molecules in solution in different solvents, the extent of ion pairing in a salt can be estimated.[405]

Variable	Definition
k_B	Boltzmann constant, $1.380\,65 \cdot 10^{-23}\,\mathrm{J\,K^{-1}}$
T	Temperature [K]
π	Number pi, 3.141 592 653 59
η	Dynamic viscosity of the medium [Pa s]
r_H	Hydrodynamic radius (spherical particle) [m]

Variable-Temperature NMR Measurements

Materials made at least partially of plastic and glass syringes were dried at 55 °C at 50 mbar for at least 1 day. All other glassware used in the sample preparation was dried at about 160 °C for at least 1 day. All Samples were prepared inside a glove-box under N_2 atmosphere and transferred into Young NMR tubes. Substances and solvent amounts are weighed with an accuray of 0.01 mg.

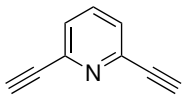
NMR spectra were recorded on a Bruker AVIII 500 MHz equipped with a BBFO broadband probe, chemical shifts are reported in ppm, using the residual peak of the deuterated solvent as internal standard (^{1}H-NMR: CD_2Cl_2 at 5.32 ppm).

Temperatures of samples inside the spectrometer were calibrated using a Pt100 resistance thermometer fixed inside an NMR tube over the whole temperature range investigated. Temperatures from the Pt100 resistance thermometer were validated using the MeOH NMR thermometer. Samples were not spun during measurement. NMR spectra were recorded every 5 K–10 K. Relative amounts of pyridines, acridines, quinolines or amines and the corresponding protonated species were determined by the relative integrals with respect to the ^{1}H signals of the counterion tetrakis[3,5-bis(trifluoromethyl)phenyl]borate ($[BArF_{24}]^-$) using a relaxation delay (D1) of 30 s.

A.1.4. Synthesis of Test Substances

In general, synthetic procedures and therefore especially yields were not optimized and only carried out once in big scale to obtain enough of all the test substances for the subsequent experiments.

2,6-Diethynylpyridine

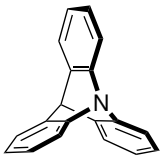


2,6-Diethynylpyridine was prepared according to a literature procedure.[406] The obtained product after flash chromatography was further purified by resublimation.

^{1}H-NMR ($CDCl_3$, 400 MHz): δ (ppm) = 7.65 (dd, *J* = 8.2, 7.4 Hz, 1 H), 7.44 (d, *J* = 7.8 Hz, 2 H), 3.15 (s, 2 H).

^{13}C-NMR ($CDCl_3$, 100 MHz): δ (ppm) = 142.85, 136.69, 127.24, 82.20, 77.92.

Azatrypticene



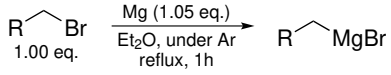
Azatrypticene was prepared according to a literature procedure.[407] Colorless crystals

^{1}H-NMR ($CDCl_3$, 400 MHz): δ (ppm) = 7.56 – 7.44 (m, 3 H), 7.43 – 7.34 (m, 3 H), 7.10 – 6.92 (m, 6 H), 5.31 (s, 1 H).

^{13}C-NMR ($CDCl_3$, 100 MHz): δ (ppm) = 155.06, 149.16, 126.18, 125.24, 123.91, 123.87, 50.70.

HRMS: calculated for $C_{19}H_{14}N^+$ [(M + H)$^+$] 256.1121; found 256.1124.

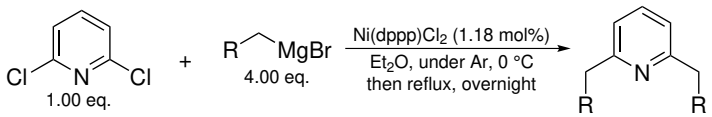
General Procedure for the Preparation of Grignard Reagents



In a typical experiment Mg turnings (689.9 mg, 28.38 mmol, 1.050 eq.) were placed in a two-necked flask equipped with a reflux condensor and the whole apparatus together with the Mg turnings were flame-dried in vacuo directly before the experiment and allowed to cool down to room temperature under Ar

atmosphere. Dry Et_2O (5 ml) was added and the resulting suspension was vigorously stirred at room temperature. About 0.5 ml of a solution of benzyl bromide (4.623 g, 27.03 mmol, 1.000 eq.) in dry Et_2O (9 ml of solvent volume of the solution) was added dropwise via syringe. The resulting mixture was carefully heated until the Grignard reaction started, indicated by an exothermic reaction and the suspension slowly becoming turbid. Afterwards, the remainder of the benzyl bromide solution was added dropwise maintaining gentle reflux. After complete addition, the reaction mixture was refluxed for 1 h and allowed to cool down to room temperature afterwards. The obtained solution was separated from unreacted Mg turnings by decantation and immediately used for the following reaction.

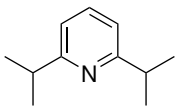
General Procedure for Kumada Cross-coupling of 2,6-Dichloropyridine and 2,7-Dichloroquinoline



A procedure from literature was adapted.[408] In a typical experiment 2,6-dichloropyridine (1.000 g, 6.757 mmol, 1.000 eq.) and [1,3-bis(diphenylphosphino)propane]nickel(II) chloride (43.1 mg, 0.0795 mmol, 1.18 mol%) were placed in a two-necked flask equipped with a reflux condensor under Ar atmosphere and dissolved in dry Et_2O (48 ml). The resulting reaction mixture was cooled down to 0 °C under stirring and a solution of benzylmagnesium bromide in Et_2O (14 ml, 1.9 mol l^{-1}, 27 mmol, 4.0 eq.) was added dropwise. After complete addition the resulting reaction mixture was allowed to warm up to room temperature and subsequently heated under reflux overnight. Reaction progress was controlled by TLC and GC-MS analysis. The reaction was quenched with a mixture of ice and saturated aqueous NH_4Cl (50 ml) and the resulting mixture was stirred vigorously until two clear layers formed. These layers were separated and the aqueous phase was extracted with Et_2O until the corresponding extract did not contain any UV-active substance (checked by spotting one droplet on a TLC plate and viewing it under a UV lamp) anymore. The combined organic phases were dried over $MgSO_4$, the solid was removed again by filtration and

the filtrate was concentrated in vacuo. In general, unless noted otherwise, purification was performed by means of flash chromatography and yielded pure material.

2,6-Diisopropylpyridine

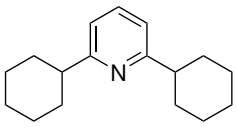


The general Kumada cross-coupling procedure was modified. 2,6-Dichloropyridine (2.000 g, 13.51 mmol, 1.000 eq.) was reacted with [1,3-bis(diphenylphosphino)propane]-nickel(II) chloride (86.2 mg, 0.159 mmol, 1.18 mol%) and isopropylmagnesium chloride (27 ml, 2.0 mol l^{-1} in Et_2O, 54 mmol, 4.0 eq.) in dry Et_2O (48 ml) and the reaction mixture was refluxed for 3 days. Work-up was performed according to the general procedure until extraction of the aqueous phase with Et_2O was finished. The combined organic phases were extracted 3 times with 1 mol l^{-1} aqueous HCl and the combined aqueous extracts were washed three times with CH_2Cl_2. Subsequently, the aqueous phase was basified with concentrated aqueous NH_3 and then extracted 3 times with CH_2Cl_2. The combined organic phases were dried over $MgSO_4$, the solid was removed again by filtration and the filtrate was concentrated in vacuo (until 500 mbar at 40 °C, product is volatile). The crude product was purified by flash chromatography (*n*-pentane : Et_2O 30:1 (v/v)). Yellowish oil (538.0 mg, 24 % of theory).

^{1}H-NMR ($CDCl_3$, 400 MHz): δ (ppm) = 7.52 (t, J = 7.7 Hz, 1 H), 6.96 (d, J = 7.7 Hz, 2 H), 3.03 (m, 2 H), 1.29 (d, J = 6.9 Hz, 12 H).

^{13}C-NMR ($CDCl_3$, 100 MHz): δ (ppm) = 166.58, 136.68, 117.31, 36.58, 22.86.

2,6-Dicyclohexylpyridine



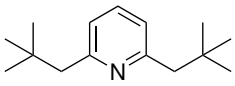
The general Kumada cross-coupling procedure was modified. 2,6-Dichloropyridine (2.000 g, 13.51 mmol, 1.000 eq.) was reacted with [1,3-bis(diphenyl-

phosphino)propane]nickel(II) chloride (86.2 mg, 0.159 mmol, 1.18 mol%) and cyclohexylmagnesium chloride (27 ml, 2.0 mol l^{-1} in Et_2O, 54 mmol, 4.0 eq.) in dry Et_2O (96 ml) and the reaction mixture was refluxed for 64 h. The crude product was purified by flash chromatography (25 %–60 % (v/v) CH_2Cl_2 in *n*-pentane). Yellow oil (1354.6 mg, 41 % of theory).

^{1}H-NMR (CD_2Cl_2, 400 MHz): δ (ppm) = 7.51 (t, *J* = 7.7 Hz, 1 H), 6.94 (d, *J* = 7.7 Hz, 2 H), 2.63 (tt, *J* = 11.6, 3.5 Hz, 2 H), 2.05 – 1.68 (m, 10 H), 1.59 – 1.15 (m, 10 H).

^{13}C-NMR (CD_2Cl_2, 100 MHz): δ (ppm) = 165.97, 136.76, 118.15, 46.96, 33.40, 27.07, 26.64.

2,6-Dineopentylpyridine

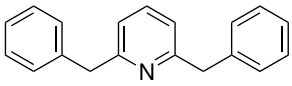


The general Kumada cross-coupling procedure was followed. 2,6-Dichloropyridine (1.000 g, 6.757 mmol, 1.000 eq.) was reacted with [1,3-bis(diphenylphosphino)propane]nickel(II) chloride (43.1 mg, 0.0795 mmol, 1.18 mol%) and freshly prepared neopentylmagnesium bromide (14 ml, 1.9 mol l^{-1} in Et_2O, 27 mmol, 4.0 eq.) in dry Et_2O (50 ml). The crude product was purified by flash chromatography (0.5 %–1.0 % (v/v) MeOH in CH_2Cl_2). Yellowish crystals (1005.7 mg, 67 % of theory). The obtained yellowish solid was purified further by resublimation (50 °C, $1 \cdot 10^{-1}$ mbar–$1 \cdot 10^{-2}$ mbar). Colorless crystals.

^{1}H-NMR (CD_2Cl_2, 400 MHz): δ (ppm) = 7.47 (t, *J* = 7.7 Hz, 1 H), 6.94 (d, *J* = 7.6 Hz, 2 H), 2.64 (s, 4 H), 0.94 (s, 18 H).

^{13}C-NMR (CD_2Cl_2, 100 MHz): δ (ppm) = 159.56, 135.23, 122.06, 52.27, 32.33, 29.88.

2,6-Dibenzylpyridine



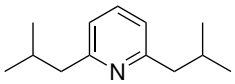
The general Kumada cross-coupling procedure was followed. 2,6-Dichloropyridine (1.000 g, 6.757 mmol, 1.000 eq.) was reacted with [1,3-bis(diphenyl-

phosphino)propane]nickel(II) chloride (43.1 mg, 0.0795 mmol, 1.18 mol%) and freshly prepared benzylmagnesium bromide (14 ml, 1.9 mol l^{-1} in Et_2O, 27 mmol, 4.0 eq.) in dry Et_2O (48 ml). The crude product was purified by flash chromatography (*n*-pentane : EtOAc 9:1 (v/v)). The obtained material was purified further by means of short-path vacuum distillation (140 °C–160 °C, $1 \cdot 10^{-3}$ mbar). Colorless oil (1.418 g, 81 % of theory).

^{1}H-NMR ($CDCl_3$, 400 MHz): δ (ppm) = 7.46 (t, *J* = 7.7 Hz, 1 H), 7.34 – 7.19 (m, 10 H), 6.89 (d, *J* = 7.7 Hz, 2 H), 4.19 (s, 4 H).

^{13}C-NMR ($CDCl_3$, 100 MHz): δ (ppm) = 160.50, 139.60, 137.31, 129.36, 128.66, 126.48, 120.73, 44.64.

2,6-Diisobutylpyridine

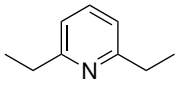


The general Kumada cross-coupling procedure was followed. 2,6-Dichloropyridine (1.000 g, 6.757 mmol, 1.000 eq.) was reacted with [1,3-bis(diphenylphosphino)propane]nickel(II) chloride (43.1 mg, 0.0795 mmol, 1.18 mol%) and isobutylmagnesium chloride (13.5 ml, 2.0 mol l^{-1} in Et_2O, 27 mmol, 4.0 eq.) in dry Et_2O (48 ml). The crude product was purified by flash chromatography (*n*-pentane : Et_2O 20:1 (v/v)). Yellowish oil (1205.8 mg, 93 % of theory).

^{1}H-NMR ($CDCl_3$, 400 MHz): δ (ppm) = 7.49 (t, *J* = 7.6 Hz, 1 H), 6.92 (d, *J* = 7.6 Hz, 2 H), 2.65 (d, *J* = 7.3 Hz, 4 H), 2.08 (m, 2 H), 0.91 (d, *J* = 6.6 Hz, 12 H).

^{13}C-NMR ($CDCl_3$, 100 MHz): δ (ppm) = 161.04, 135.97, 120.66, 47.77, 29.42, 22.52.

2,6-Diethylpyridine



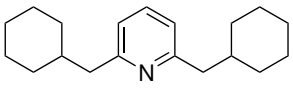
The general Kumada cross-coupling procedure was followed. 2,6-Dichloropyridine (1.000 g, 6.757 mmol, 1.000 eq.) was reacted with [1,3-bis(diphenyl-

phosphino)propane]nickel(II) chloride (43.1 mg, 0.0795 mmol, 1.18 mol%) and ethylmagnesium bromide (9.0 ml, 3.0 mol l^{-1} in Et_2O, 27 mmol, 4.0 eq.) in dry Et_2O (49 ml). The crude product was purified by flash chromatography (*n*-pentane : Et_2O 10:1 (v/v)). Yellowish oil (803.4 mg, 88 % of theory).

^{1}H-NMR ($CDCl_3$, 400 MHz): δ (ppm) = 7.53 (t, *J* = 7.7 Hz, 1 H), 6.98 (d, *J* = 7.7 Hz, 2 H), 2.81 (q, *J* = 7.6 Hz, 4 H), 1.29 (t, *J* = 7.6 Hz, 6 H).

^{13}C-NMR ($CDCl_3$, 100 MHz): δ (ppm) = 162.95, 137.07, 119.15, 31.51, 14.34.

2,6-Bis(cyclohexylmethyl)pyridine



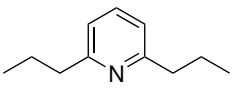
The general Kumada cross-coupling procedure was followed. 2,6-Dichloropyridine (1.000 g, 6.757 mmol, 1.000 eq.) was reacted with [1,3-bis(diphenylphosphino)propane]nickel(II) chloride (43.1 mg, 0.0795 mmol, 1.18 mol%) and freshly prepared (cyclohexylmethyl)magnesium bromide (14.0 ml, 1.9 mol l^{-1} in Et_2O, 27 mmol, 4.0 eq.) in dry Et_2O (50 ml). The crude product was purified by flash chromatography (*n*-pentane : Et_2O 25:1 (v/v)). Yellowish oil (1639.3 mg, 89 % of theory).

^{1}H-NMR ($CDCl_3$, 400 MHz): δ (ppm) = 7.44 (t, *J* = 7.5 Hz, 1 H), 6.88 (d, *J* = 7.9, 2 H), 2.63 (d, *J* = 7.7 Hz, 4 H), 1.82 – 1.53 (m, 13 H), 1.18 (m, 6 H), 0.97 (m, 4 H).

^{13}C-NMR ($CDCl_3$, 100 MHz): δ (ppm) = 160.65, 135.92, 120.76, 46.25, 38.77, 33.21, 26.68, 26.39.

HRMS: calculated for $C_{19}H_{30}N^+$ [(M + H)$^+$] 272.2373; found 272.2378.

2,6-Dipropylpyridine



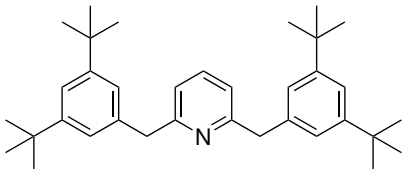
The general Kumada cross-coupling procedure was followed. 2,6-Dichloropyridine (1.000 g, 6.757 mmol, 1.000 eq.) was reacted with [1,3-bis(diphenylphosphino)propane]nickel(II) chloride (43.1 mg, 0.0795 mmol, 1.18 mol%)

and freshly prepared propylmagnesium bromide (14.0 ml, 1.9 mol l^{-1} in Et_2O, 27 mmol, 4.0 eq.) in dry Et_2O (50 ml). The crude product was purified by flash chromatography (*n*-pentane : Et_2O 10:1 (v/v)). Yellowish oil (1025.4 mg, 92 % of theory).

^{1}H-NMR ($CDCl_3$, 400 MHz): δ (ppm) = 7.49 (t, *J* = 7.7 Hz, 1 H), 6.95 (d, *J* = 7.7 Hz, 2 H), 2.79 – 2.70 (m, 4 H), 1.80 – 1.65 (m, 4 H), 0.96 (t, *J* = 7.3 Hz, 6 H).

^{13}C-NMR ($CDCl_3$, 100 MHz): δ (ppm) = 161.73, 136.65, 119.93, 40.55, 23.56, 14.03.

2,6-Bis(3,5-di-tert-butylbenzyl)pyridine



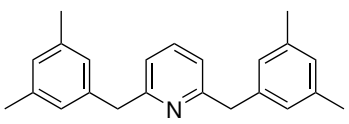
The general Kumada cross-coupling procedure was followed. 2,6-Dichloropyridine (853.3 mg, 5.766 mmol, 1.000 eq.) was reacted with [1,3-bis(diphenylphosphino)propane]nickel(II) chloride (36.8 mg, 0.0678 mmol, 1.18 mol%) and freshly prepared (3,5-di-tert-butylbenzyl)magnesium bromide (11.0 ml, 1.6 mol l^{-1} in Et_2O, 17 mmol, 4.0 eq.) in dry Et_2O (43 ml). The crude product was purified by flash chromatography (33 %–100 % (v/v) CH_2Cl_2 in *n*-pentane). Colorless crystals (1639.3 mg, 89 % of theory).

^{1}H-NMR ($CDCl_3$, 400 MHz): δ (ppm) = 7.44 (t, *J* = 7.7 Hz, 1 H), 7.30 (t, *J* = 1.8 Hz, 2 H), 7.16 (d, *J* = 1.9 Hz, 4 H), 6.89 (d, *J* = 7.7 Hz, 2 H), 4.19 (s, 4 H), 1.32 (s, 36 H).

^{13}C-NMR ($CDCl_3$, 100 MHz): δ (ppm) = 160.93, 150.94, 150.93, 138.64, 136.89, 123.71, 120.42, 120.38, 45.43, 34.95, 31.66, 31.64.

HRMS: calculated for $C_{35}H_{50}N^+$ [(M + H)$^+$] 484.3938; found 484.3941.

2,6-Bis(3,5-dimethylbenzyl)pyridine



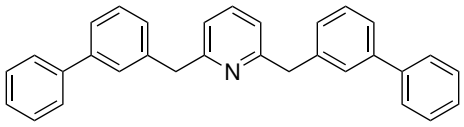
The general Kumada cross-coupling procedure was followed. 2,6-Dichloropyridine (1.000 g, 6.757 mmol, 1.000 eq.) was reacted with [1,3-bis(diphenylphosphino)propane]nickel(II) chloride (43.1 mg, 0.0795 mmol, 1.18 mol%) and freshly prepared (3,5-dimethyl-benzyl)magnesium bromide (14.0 ml, 1.9 mol l^{-1} in Et_2O, 27 mmol, 4.0 eq.) in dry Et_2O (48 ml). The crude product was purified by flash chromatography (*n*-pentane : Et_2O 9:1 (v/v)). Colorless crystals (1575.2 mg, 73 % of theory).

^{1}H-NMR ($CDCl_3$, 400 MHz): δ (ppm) = 7.43 (t, J = 7.7 Hz, 1 H), 6.93 – 6.83 (m, 8 H), 4.10 (s, 4 H), 2.29 (s, 12 H).

^{13}C-NMR ($CDCl_3$, 100 MHz): δ (ppm) = 160.77, 139.65, 138.09, 136.97, 128.05, 127.16, 120.52, 44.78, 21.42.

HRMS: calculated for $C_{23}H_{26}N^+$ [(M + H)$^+$] 316.2060; found 316.2064.

2,6-Bis([1,1'-biphenyl]-3-ylmethyl)pyridine



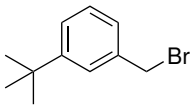
The general Kumada cross-coupling procedure was followed. 2,6-Dichloropyridine (1.000 g, 6.757 mmol, 1.000 eq.) was reacted with [1,3-bis(diphenylphosphino)propane]nickel(II) chloride (43.1 mg, 0.0795 mmol, 1.18 mol%) and freshly prepared ([1,1'-biphenyl]-3-ylmethyl)magnesium bromide (14.0 ml, 1.9 mol l^{-1} in Et_2O, 27 mmol, 4.0 eq.) in dry Et_2O (48 ml). The crude product was purified by flash chromatography (*n*-pentane : Et_2O 9:1 to 4:1 (v/v)). Yellowish oil (2.792 g, 99 % of theory).

^{1}H-NMR ($CDCl_3$, 400 MHz): δ (ppm) = 7.63 – 7.22 (m, 19 H), 6.95 (d, J = 7.7 Hz, 2 H), 4.25 (s, 4 H).

^{13}C-NMR ($CDCl_3$, 100 MHz): δ (ppm) = 160.55, 141.58, 141.31, 140.24, 137.17, 129.06, 128.85, 128.26, 128.21, 127.36, 127.32, 125.33, 120.74, 44.95.

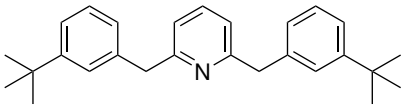
HRMS: calculated for $C_{31}H_{26}N^+$ [(M + H)$^+$] 412.2060; found 412.2058.

1-(Bromomethyl)-3-(tert-butyl)benzene



1-(Bromomethyl)-3-(tert-butyl)benzene was prepared according to a literature procedure.[409] The crude product was purified by distillation in vacuo (104 °C–112 °C at 7 mbar). Colorless oil. ^{1}H-NMR showed the product to be a mixture of the desired monobrominated product and approximately 7 % of bisbrominated product.

2,6-Bis(3-(tert-butyl)benzyl)pyridine



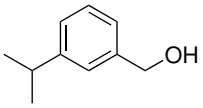
The general Kumada cross-coupling procedure was followed. 2,6-Dichloropyridine (0.500 g, 3.38 mmol, 1.00 eq.) was reacted with [1,3-bis(diphenylphosphino)propane]nickel(II) chloride (21.6 mg, 0.0398 mmol, 1.18 mol%) and freshly prepared (3-(tert-butyl)benzyl)magnesium bromide (8.0 ml, 1.7 mol l^{-1} in Et_2O, 13 mmol, 3.9 eq.) in dry Et_2O (26 ml). The crude product was purified by flash chromatography (*n*-pentane : CH_2Cl_2 1:2 (v/v) to pure CH_2Cl_2). The obtained yellow oil was further purified by flash chromatography (*n*-pentane : EtOAc 25:1 (v/v)). Yellowish oil (826.5 mg, 65 % of theory).

^{1}H-NMR ($CDCl_3$, 400 MHz): δ (ppm) = 7.44 (t, *J* = 7.7 Hz, 1 H), 7.37 – 7.29 (m, 2 H), 7.30 – 7.18 (m, 4 H), 7.07 (dt, *J* = 6.7, 1.7 Hz, 2 H), 6.88 (d, *J* = 7.7 Hz, 2 H), 4.17 (s, 4 H), 1.31 (s, 18 H).

^{13}C-NMR ($CDCl_3$, 100 MHz): δ (ppm) = 160.77, 151.48, 139.33, 136.99, 128.30, 126.47, 126.38, 123.37, 120.55, 45.15, 34.79, 31.53.

HRMS: calculated for $C_{27}H_{34}N^+$ [(M + H)$^+$] 372.2686; found 372.2685.

(3-Isopropylphenyl)methanol

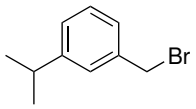


3-Isopropylphenyl)methanol was prepared according to a literature procedure.[410] Yellowish oil.

^{1}H-NMR ($CDCl_3$, 400 MHz): δ (ppm) = 7.30 (t, *J* = 7.5 Hz, 1 H), 7.25 – 7.22 (m, 1 H), 7.21 – 7.14 (m, 2 H), 4.68 (d, *J* = 5.8 Hz, 2 H), 2.99 – 2.85 (m, 1 H), 1.71 (t, *J* = 6.0 Hz, 1 H), 1.27 (d, *J* = 6.9 Hz, 6 H).

^{13}C-NMR ($CDCl_3$, 100 MHz): δ (ppm) = 149.44, 140.97, 128.70, 125.96, 125.28, 124.64, 65.73, 34.26, 24.14.

1-(Bromomethyl)-3-isopropylbenzene

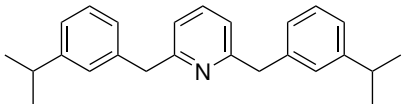


1-(Bromomethyl)-3-isopropylbenzene was prepared according to a literature procedure.[410] Yellowish oil.

^{1}H-NMR ($CDCl_3$, 400 MHz): δ (ppm) = 7.34 – 7.13 (m, 4 H), 4.50 (s, 2 H), 2.91 (m, 1 H), 1.26 (d, *J* = 6.9 Hz, 7 H).

^{13}C-NMR ($CDCl_3$, 100 MHz): δ (ppm) = 149.64, 137.80, 128.90, 127.32, 126.75, 126.65, 34.15, 34.08, 24.05.

2,6-Bis(3-isopropylbenzyl)pyridine



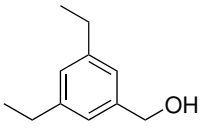
The general Kumada cross-coupling procedure was followed. 2,6-Dichloropyridine (0.500 g, 3.38 mmol, 1.00 eq.) was reacted with [1,3-bis(diphenylphosphino)propane]nickel(II) chloride (21.6 mg, 0.0398 mmol, 1.18 mol%) and freshly prepared (3-isopropyl-benzyl)magnesium bromide (8.0 ml, 1.7 mol l^{-1} in Et_2O, 14 mmol, 4.0 eq.) in dry Et_2O (24 ml). The crude product was purified by flash chromatography (*n*-pentane : CH_2Cl_2 1:2 (v/v) to pure CH_2Cl_2). The obtained yellow oil was further purified by flash chromatography (*n*-pentane : Et_2O 9:1 (v/v)). Yellowish oil (1.1522 g, 98 % of theory).

^{1}H-NMR ($CDCl_3$, 400 MHz): δ (ppm) = 7.44 (t, *J* = 7.7 Hz, 1 H), 7.27 – 7.05 (m, 8 H), 6.89 (d, *J* = 7.7 Hz, 2 H), 4.16 (s, 4 H), 2.88 (m, 2 H), 1.24 (d, *J* = 6.9 Hz, 12 H).

^{13}C-NMR ($CDCl_3$, 100 MHz): δ (ppm) = 160.72, 149.24, 139.66, 137.00, 128.56, 127.63, 126.74, 124.44, 120.56, 44.97, 34.20, 24.16.

HRMS: calculated for $C_{25}H_{30}N^+$ [(M + H)$^+$] 344.2373; found 344.2377.

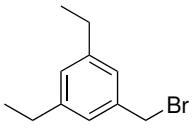
(3,5-Diethylphenyl)methanol



A procedure from literature was adapted.[411] 3,5-Diethylbenzoic acid (3.488 g, 19.57 mmol, 1.00 eq.) was dissolved in dry THF (39 ml) under Ar and the resulting solution cooled down to 0 °C under stirring. A solution of $BH_3 \cdot THF$ (39 ml, 1 mol l^{-1}, $4 \cdot 10^1$ mmol, 2 eq.) was added dropwise and the resulting mixture was allowed to warm up to room temperature and stirred overnight. The reaction was quenched by carefully adding aqueous HCl (100 ml, 2 mol l^{-1}, $2 \cdot 10^2$ mmol) dropwise. CH_2Cl_2 (150 ml) was added and the resulting layers were separated. The aqueous phase was extracted twice with CH_2Cl_2 (50 ml). The combined organic layers were washed once with saturated aqueous $NaHCO_3$ solution and once with brine, dried over $MgSO_4$, the solid removed again by filtration and the filtrate was concentrated in vacuo. The product was found to be reasonably pure and used for the next step without further purification. Yellowish liquid (3.2926 g, quantitative).

^{1}H-NMR ($CDCl_3$, 300 MHz): δ (ppm) = 7.07 – 7.01 (m, 2 H), 7.01 – 6.95 (m, 1 H), 4.66 (d, *J* = 5.2 Hz, 2 H), 2.64 (q, *J* = 7.6 Hz, 4 H), 1.62 (t, *J* = 5.7 Hz, 1 H), 1.24 (t, *J* = 7.6 Hz, 6 H).

1-(Bromomethyl)-3,5-diethylbenzene



A procedure from literature was adapted.[410] 3,5-Diethylphenyl)methanol from the previous step (3.2926 g of the crude product, assumed to correspond to 19.57 mmol, 1.00 eq.) was dissolved in dry Et_2O (35 ml) under Ar and cooled down to 0 °C under stirring. PBr_3 (1.1 ml, 11 mmol, 1.8 eq.) was added dropwise and the resulting mixture was stirred for 1 h at 0 °C. Reaction progress was controlled by TLC. The mixture was allowed to warm up to room temperature and saturated aqueous $NaHCO_3$ solution was added carefully until no gas evolution was observed anymore. The resulting layers were separated and the aqueous phase was extracted twice with CH_2Cl_2. The combined organic extracts were dried over $MgSO_4$, the solid removed again by filtration and the filtrate was concentrated in vacuo. The product was found to be reasonably pure and used for the next step without further purification. Yellowish liquid (4.477 g, 93 % of theory over two steps).

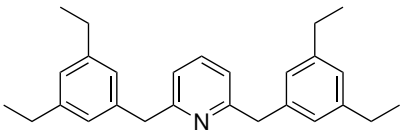
^{1}H-NMR ($CDCl_3$, 400 MHz): δ (ppm) = 7.10 – 7.03 (m, 2 H), 7.01 – 6.95 (m, 1 H), 4.48 (s, 2 H), 2.64 (q, *J* = 7.6 Hz, 4 H), 1.25 (t, *J* = 7.6 Hz, 6 H).

^{13}C-NMR ($CDCl_3$, 100 MHz): δ (ppm) = 144.97, 137.78, 127.93, 126.05, 34.23, 28.81, 15.59.

HRMS: calculated for $C_{11}H_{15}Br^{\cdot+}$ [$M^{\cdot+}$] 226.0352; found 226.0361.

For this compound HRMS analysis was performed on a Waters AutoSpec Ultima (EI-triSector-MS) by gas chromatography in combination with high-resolution sector mass spectrometry.

2,6-Bis(3,5-diethylbenzyl)pyridine



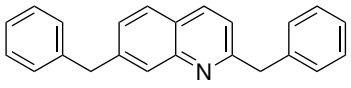
The general Kumada cross-coupling procedure was followed. 2,6-Dichloropyridine (0.400 g, 2.70 mmol, 1.00 eq.) was reacted with [1,3-bis(diphenylphosphino)propane]nickel(II) chloride (17.4 mg, 0.0318 mmol, 1.18 mol%) and freshly prepared (3,5-diethylben-zyl)magnesium bromide (6.4 ml, 1.7 mol l^{-1} in Et_2O, 11 mmol, 4.0 eq.) in dry Et_2O (20 ml). The crude product was purified by flash chromatography (*n*-pentane : Et_2O 14:1 (v/v)). Yellowish oil (433.4 mg, 43 % of theory).

^{1}H-NMR ($CDCl_3$, 400 MHz): δ (ppm) = 7.43 (t, *J* = 7.7 Hz, 1 H), 6.97 – 6.82 (m, 8 H), 4.13 (s, 4 H), 2.59 (q, *J* = 7.6 Hz, 8 H), 1.21 (t, *J* = 7.6 Hz, 12 H).

^{13}C-NMR ($CDCl_3$, 100 MHz): δ (ppm) = 160.81, 144.57, 139.67, 136.96, 126.30, 125.59, 120.52, 44.94, 28.93, 15.75.

HRMS: calculated for $C_{27}H_{34}N^+$ [(M + H)$^+$] 372.2686; found 372.2685.

2,7-Dibenzylquinoline



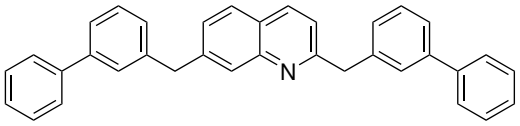
The general Kumada cross-coupling procedure was followed. 2,7-Dichloroquinoline (0.500 g, 2.53 mmol, 1.00 eq.) was reacted with [1,3-bis(diphenylphosphino)propane]nickel(II) chloride (16.2 mg, 0.0297 mmol, 1.18 mol%) and freshly prepared benzylmagnesium bromide (7.0 ml, 1.4 mol l^{-1} in Et_2O, 10 mmol, 4.0 eq.) in dry Et_2O (19 ml). The crude product was purified by flash chromatography (*n*-pentane : Et_2O 5:1 (v/v)). Yellow crystals (731.4 mg, 94 % of theory).

^{1}H-NMR ($CDCl_3$, 400 MHz): δ (ppm) = 7.98 (dd, *J* = 8.5, 0.8 Hz, 1 H), 7.90 (m, 1 H), 7.67 (d, *J* = 8.3 Hz, 1 H), 7.39 – 7.13 (m, 12 H), 4.32 (s, 2 H), 4.18 (s, 2 H).

^{13}C-NMR ($CDCl_3$, 100 MHz): δ (ppm) = 161.47, 148.21, 143.09, 140.55, 139.42, 136.34, 129.36, 129.28, 128.75, 128.76, 128.40, 127.81, 127.68, 126.61, 126.45, 125.38, 121.20, 45.73, 42.43.

HRMS: calculated for $C_{23}H_{20}N^+$ [(M + H)$^+$] 310.1590; found 310.1590.

2,7-Bis([1,1'-biphenyl]-3-ylmethyl)quinoline



The general Kumada cross-coupling procedure was followed. 2,7-Dichloroquinoline (0.500 g, 2.53 mmol, 1.00 eq.) was reacted with [1,3-bis(diphenylphosphino)propane]nickel(II) chloride (16.2 mg, 0.0297 mmol, 1.18 mol%) and freshly prepared ([1,1'-biphenyl]-3-ylmethyl)magnesium bromide (7.0 ml, 1.4 mol l^{-1} in Et_2O, 10 mmol, 4.0 eq.) in dry Et_2O (19 ml). The crude product was

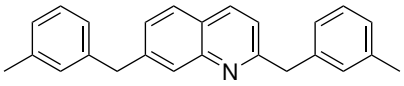
purified by flash chromatography (*n*-pentane : Et_2O 4:1 (v/v)). Orange thick oil (1074.8 mg, 92 % of theory).

^{1}H-NMR ($CDCl_3$, 400 MHz): δ (ppm) = 7.99 (d, *J* = 8.4, 1 H), 7.97 – 7.92 (m, 1 H), 7.68 (d, *J* = 8.3 Hz, 1 H), 7.60 – 7.52 (m, 5 H), 7.51 – 7.18 (m, 15 H), 4.39 (s, 2 H), 4.25 (s, 2 H).

^{13}C-NMR ($CDCl_3$, 100 MHz): δ (ppm) = 161.37, 148.24, 143.02, 141.75, 141.74, 141.29, 141.20, 141.03, 139.90, 136.42, 129.18, 128.85, 128.84, 128.42, 128.32, 128.25, 128.17, 127.83, 127.76, 127.41, 127.38, 127.37, 127.33, 125.52, 125.46, 125.39, 121.24, 45.83, 42.53.

HRMS: calculated for $C_{35}H_{28}N^+$ [(M + H)$^+$] 462.2216; found 462.2216.

2,7-Bis(3-methylbenzyl)quinoline



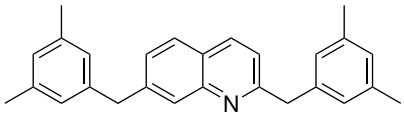
The general Kumada cross-coupling procedure was followed. 2,7-Dichloroquinoline (0.500 g, 2.53 mmol, 1.00 eq.) was reacted with [1,3-bis(diphenylphosphino)propane]nickel(II) chloride (16.2 mg, 0.0297 mmol, 1.18 mol%) and freshly prepared (3-methylbenzyl)-magnesium bromide (7.0 ml, 1.4 mol l^{-1} in Et_2O, 10 mmol, 4.0 eq.) in dry Et_2O (19 ml). The crude product was purified by flash chromatography (*n*-pentane : Et_2O 8:1 (v/v)). Yellow oil (736.6 mg, 86 % of theory).

^{1}H-NMR ($CDCl_3$, 400 MHz): δ (ppm) = 7.97 (d, *J* = 8.4, 1 H), 7.94 – 7.86 (m, 1 H), 7.67 (d, *J* = 8.3 Hz, 1 H), 7.35 (dd, *J* = 8.4, 1.7 Hz, 1 H), 7.23 – 7.14 (m, 3 H), 7.14 – 6.96 (m, 6 H), 4.28 (s, 2 H), 4.14 (s, 2 H), 2.40 – 2.21 (m, 6 H).

^{13}C-NMR ($CDCl_3$, 100 MHz): δ (ppm) = 161.59, 148.20, 143.22, 140.47, 139.33, 138.39, 138.35, 136.31, 130.15, 130.06, 128.63, 128.35, 127.79, 127.65, 127.36, 127.20, 126.40, 126.30, 125.37, 121.20, 45.70, 42.40, 21.56, 21.53.

HRMS: calculated for $C_{25}H_{24}N^+$ [(M + H)$^+$] 338.1903; found 338.1899.

2,7-Bis(3,5-dimethylbenzyl)quinoline



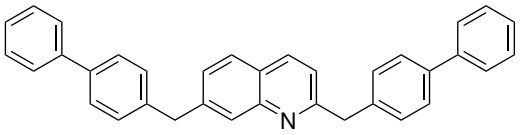
The general Kumada cross-coupling procedure was followed. 2,7-Dichloroquinoline (0.500 g, 2.53 mmol, 1.00 eq.) was reacted with [1,3-bis(diphenylphosphino)propane]nickel(II) chloride (16.2 mg, 0.0297 mmol, 1.18 mol%) and freshly prepared (3,5-dimethylben-zyl)magnesium bromide (7.0 ml, 1.4 mol l^{-1} in Et_2O, 10 mmol, 4.0 eq.) in dry Et_2O (19 ml). The crude product was purified by flash chromatography (*n*-pentane : Et_2O 8:1 (v/v)). Pink crystals (651.0 mg, 71 % of theory).

^{1}H-NMR ($CDCl_3$, 400 MHz): δ (ppm) = 7.97 (d, *J* = 8.4, 1 H), 7.92 – 7.88 (m, 1 H), 7.67 (d, *J* = 8.3 Hz, 1 H), 7.35 (dd, *J* = 8.4, 1.7 Hz, 1 H), 7.17 (d, *J* = 8.4, 1 H), 6.95 – 6.81 (m, 6 H), 4.24 (s, 2 H), 4.10 (s, 2 H), 2.31 – 2.22 (m, 12 H).

^{13}C-NMR ($CDCl_3$, 100 MHz): δ (ppm) = 161.69, 148.19, 143.33, 140.42, 139.27, 138.26, 138.22, 136.27, 128.31, 128.26, 128.10, 127.77, 127.61, 127.20, 127.10, 125.36, 121.20, 45.65, 42.34, 21.42, 21.40.

HRMS: calculated for $C_{27}H_{28}N^+$ [(M + H)$^+$] 366.2216; found 366.2222.

2,7-Bis([1,1'-biphenyl]-4-ylmethyl)quinoline



The general Kumada cross-coupling procedure was followed. 2,7-Dichloroquinoline (0.500 g, 2.53 mmol, 1.00 eq.) was reacted with [1,3-bis(diphenylphosphino)propane]nickel(II) chloride (16.2 mg, 0.0297 mmol, 1.18 mol%) and freshly prepared ([1,1'-biphenyl]-4-ylmethyl)magnesium bromide (7.0 ml, 1.4 mol l^{-1} in Et_2O, 10 mmol, 4.0 eq.) in dry Et_2O (19 ml). Extraction was performed using CH_2Cl_2 instead of Et_2O because at least one compound in the reaction mixture was only sparingly soluble in Et_2O. The crude solid was purified first by trituration with CH_2Cl_2 to extract the desired product. The combined extracts were concentrated in vacuo and the resulting residue was purified by

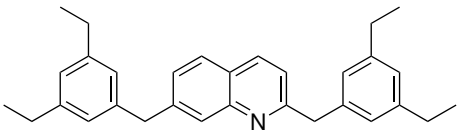
flash chromatography (*n*-pentane : CH_2Cl_2 : Et_2O 8:1:1 (v/v/v)). Yellowish powder (391.6 mg, 33 % of theory).

^{1}H-NMR ($CDCl_3$, 400 MHz): δ (ppm) = 8.07 – 7.89 (m, 2 H), 7.71 (d, *J* = 8.3, 1 H), 7.65 – 7.47 (m, 8 H), 7.47 – 7.26 (m, 12 H), 7.24 (d, *J* = 8.4 Hz, 1 H), 7.60 – 7.52 (m, 5 H), 7.51 – 7.18 (m, 15 H), 4.38 (s, 2 H), 4.24 (s, 2 H).

^{13}C-NMR ($CDCl_3$, 100 MHz): δ (ppm) = 161.36, 143.15, 141.11, 141.02, 139.59, 139.47, 138.42, 136.56, 129.76, 129.69, 128.87, 128.30, 127.92, 127.78, 127.53, 127.51, 127.29, 127.25, 127.19, 127.15, 125.47, 121.29, 45.27, 42.07.

HRMS: calculated for $C_{35}H_{28}N^+$ [(M + H)$^+$] 462.2216; found 462.2213.

2,7-Bis(3,5-diethylbenzyl)quinoline



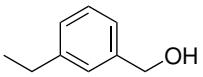
The general Kumada cross-coupling procedure was followed. 2,7-Dichloroquinoline (0.400 g, 2.02 mmol, 1.00 eq.) was reacted with [1,3-bis(diphenylphosphino)propane]nickel(II) chloride (13.0 mg, 0.0238 mmol, 1.18 mol%) and freshly prepared (3,5-diethylben-zyl)magnesium bromide (5.6 ml, 1.4 mol l^{-1} in Et_2O, 8.0 mmol, 4.0 eq.) in dry Et_2O (15 ml). The crude product was purified by flash chromatography (*n*-pentane : Et_2O 14:1 (v/v)). Orange oil (468.9 mg, 55 % of theory).

^{1}H-NMR ($CDCl_3$, 400 MHz): δ (ppm) = 7.97 (d, *J* = 8.4, 1 H), 7.94 – 7.89 (m, 1 H), 7.67 (d, *J* = 8.3 Hz, 1 H), 7.37 (dd, *J* = 8.3, 1.7 Hz, 1 H), 7.18 (d, *J* = 8.4, 1 H), 7.00 – 6.95 (m, 2 H), 6.95 – 6.88 (m, 4 H), 4.28 (s, 2 H), 4.13 (s, 2 H), 2.66 – 2.51 (m, 8 H), 1.30 – 1.13 (m, 12 H).

^{13}C-NMR ($CDCl_3$, 100 MHz): δ (ppm) = 161.74, 148.18, 144.73, 144.68, 143.36, 140.46, 139.29, 136.26, 128.31, 127.79, 127.58, 126.36, 126.22, 125.79, 125.60, 125.35, 121.18, 45.80, 42.49, 28.92, 28.91, 15.74.

HRMS: calculated for $C_{31}H_{36}N^+$ [(M + H)$^+$] 422.2842; found 422.2841.

((3-Ethylphenyl)methanol

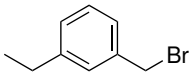


A procedure from literature was adapted.[411] 3-Ethylbenzoic acid (2.628 g, 17.50 mmol, 1.00 eq.) was dissolved in dry THF (35 ml) under Ar and the resulting solution cooled down to 0 °C under stirring. A solution of $BH_3 \cdot THF$ (35 ml, 1 mol l^{-1}, $4 \cdot 10^1$ mmol, 2 eq.) was added dropwise and the resulting mixture was allowed to warm up to room temperature and stirred overnight. The reaction was quenched by carefully adding aqueous HCl (90 ml, 2 mol l^{-1}, $2 \cdot 10^2$ mmol) dropwise. CH_2Cl_2 (150 ml) was added and the resulting layers were separated. The aqueous phase was extracted twice with CH_2Cl_2 (50 ml). The combined organic layers were washed once with saturated aqueous $NaHCO_3$ solution and once with brine, dried over $MgSO_4$, the solid removed again by filtration and the filtrate was concentrated in vacuo. The product was found to be reasonably pure and used for the next step without further purification. Yellowish liquid (2.4840 g, quantitative).

^{1}H-NMR ($CDCl_3$, 400 MHz): δ (ppm) = 7.29 (t, *J* = 7.5 Hz, 1 H), 7.24 – 7.11 (m, 3 H), 4.76 – 4.59 (m, 2 H), 2.67 (q, *J* = 7.6 Hz, 2 H), 1.68 (t, *J* = 5.5 Hz, 1 H), 1.25 (t, *J* = 7.6 Hz, 3 H).

^{13}C-NMR ($CDCl_3$, 100 MHz): δ (ppm) = 144.80, 141.01, 128.70, 127.38, 126.71, 124.46, 65.65, 28.97, 15.73.

1-(Bromomethyl)-3-ethylbenzene



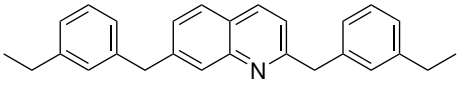
A procedure from literature was adapted.[410] 3-Ethylphenyl)methanol from the previous step (2.4840 g of the crude product, assumed to correspond to 17.50 mmol, 1.00 eq.) was dissolved in dry Et_2O (30 ml) under Ar and cooled down to 0 °C under stirring. PBr_3 (0.95 ml, 8.81 mmol, 1.7 eq.) was added dropwise and the resulting mixture was allowed to warm up to room temperature and stirred for 2 h. Reaction progress was controlled by TLC. Saturated aqueous $NaHCO_3$ solution was added carefully until no gas evolution was observed

anymore. The resulting layers were separated and the aqueous phase was extracted twice with CH_2Cl_2. The combined organic extracts were dried over $MgSO_4$, the solid removed again by filtration and the filtrate was concentrated in vacuo. The product was found to be reasonably pure and used for the next step without further purification. Colorless liquid (3.4605 g, 99 % of theory over two steps).

^{1}H-NMR ($CDCl_3$, 400 MHz): δ (ppm) = 7.30 – 7.19 (m, 3 H), 7.17 – 7.11 (m, 1 H), 4.49 (s, 2 H), 2.65 (q, *J* = 7.6 Hz, 2 H), 1.25 (t, *J* = 7.6 Hz, 3 H).

^{13}C-NMR ($CDCl_3$, 100 MHz): δ (ppm) = 145.01, 137.85, 128.91, 128.70, 128.19, 126.49, 33.98, 28.83, 15.57.

2,7-Bis(3-ethylbenzyl)quinoline



The general Kumada cross-coupling procedure was followed. 2,7-Dichloroquinoline (0.500 g, 2.53 mmol, 1.00 eq.) was reacted with [1,3-bis(diphenylphosphino)propane]nickel(II) chloride (16.2 mg, 0.0297 mmol, 1.18 mol%) and freshly prepared (3-ethylbenzyl)-magnesium bromide (7.0 ml, 1.4 mol l^{-1} in Et_2O, 10 mmol, 4.0 eq.) in dry Et_2O (19 ml). The crude product was purified by flash chromatography (*n*-pentane : Et_2O 8:1 (v/v)). Orange oil (650.8 mg, 70 % of theory).

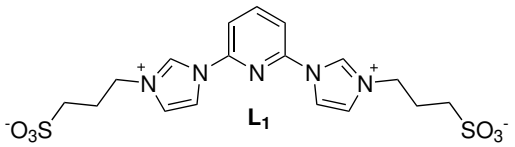
^{1}H-NMR ($CDCl_3$, 400 MHz): δ (ppm) = 8.04 – 7.86 (m, 2 H), 7.67 (d, *J* = 8.3, 1 H), 7.36 (d, *J* = 8.4, 1 H), 7.31 – 6.98 (m, 9 H), 4.30 (s, 2 H), 4.16 (s, 2 H), 2.74 – 2.49 (m, 4 H), 1.33 – 1.10 (m, 6 H).

^{13}C-NMR ($CDCl_3$, 100 MHz): δ (ppm) = 161.61, 148.19, 144.77, 144.72, 143.24, 140.50, 139.35, 136.30, 129.02, 128.90, 128.69, 128.34, 127.80, 127.63, 126.66, 126.57, 126.13, 125.96, 125.36, 121.19, 45.77, 42.47, 28.95, 28.93, 15.71.

HRMS: calculated for $C_{31}H_{36}N^+$ [(M + H)$^+$] 366.2216; found 366.2217.

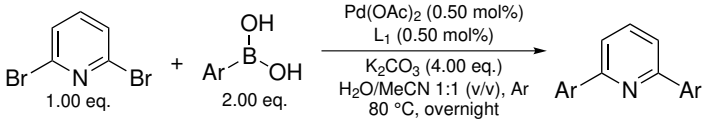
General Procedure for Suzuki Cross-coupling of 2,6-Dibromopyridine

Ligand L_1



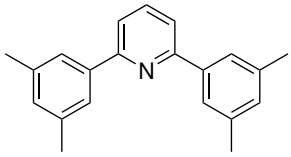
The ligand depicted above was prepared according to a literature procedure.[412] In the following procedure, 3,3'-(Pyridine-2,6-diylbis(1H-imidazole-3-ium-1,3-diyl))bis(propane-1-sulfonate) is abbreviated simply as L_1.

Suzuki Cross-coupling



A procedure from literature was adapted.[412] In a typical experiment $Pd(OAc)_2$ (5.6 mg, 0.025 mmol, 0.50 mol%) and L_1 (11.4 mg, 0.0250 mmol, 0.500 mol%) were placed in a two-necked flask equipped with a reflux condensor under Ar atmosphere and dissolved in H_2O (15 ml) and MeCN (15 ml). The resulting mixture was stirred until a yellow homogeneous solution was formed. 2,6-dibromopyridine (1.184 g, 5.000 mmol, 1.000 eq.) and K_2CO_3 (2.764 g, 20.00 mmol, 4.000 eq.) were added and stirred vigorously for 10 min. Two layers formed in the reaction mixture. Subsequently, (3,5-dimethylphenyl)boronic acid (1.500 g, 10.00 mmol, 2.000 eq.) was added and the resulting biphasic reaction mixture was heated to 80 °C overnight under vigorous stirring. Reaction progress was controlled by TLC and GC-MS analysis. Afterwards, H_2O and Et_2O were added and the resulting two layers were separated. The aqueous phase was extracted twice with Et_2O (no UV-active substance was extracted anymore). The combined organic phases were dried over $MgSO_4$, the solid was removed again by filtration and the filtrate was concentrated in vacuo. In general, unless noted otherwise, purification was performed by means of flash chromatography and yielded pure material.

2,6-Bis(3,5-dimethylphenyl)pyridine

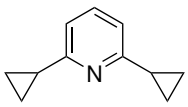


The general Suzuki cross-coupling procedure was followed. 2,6-Dibromopyridine (1.184 g, 6.757 mmol, 1.000 eq.) was reacted with $Pd(OAc)_2$ (5.6 mg, 0.025 mmol, 0.50 mol%), L_1 (11.4 mg, 0.0250 mmol, 0.500 mol%), K_2CO_3 (2.764 g, 20.00 mmol, 4.000 eq.) and (3,5-dimethylphenyl)boronic acid (1.500 g, 10.00 mmol, 2.000 eq.) in H_2O (15 ml) and MeCN (15 ml). The crude product was purified by flash chromatography (*n*-pentane : CH_2Cl_2 2.5:1 (v/v)). Colorless powder (1120.7 mg, 78 % of theory).

^{1}H-NMR ($CDCl_3$, 400 MHz): δ (ppm) = 7.88 – 7.59 (m, 7 H), 7.08 (s, 2 H), 2.43 (s, 12 H).

^{13}C-NMR ($CDCl_3$, 100 MHz): δ (ppm) = 157.43, 139.78, 138.29, 137.35, 130.72, 125.09, 118.94, 21.63.

2,6-Dicyclopropylpyridine



A procedure from literature was adapted.[413] 2,6-Dichloropyridine (1.500 g, 10.14 mmol, 1.000 eq.) and iron(III) acetylacetonate (134.1 mg, 0.3798 mmol, 3.75 mol%) were placed in a two-necked flask under Ar and dissolved in dry THF (66.8 ml) and dry NMP (25.4 ml). The resulting mixture was stirred at room temperature and cyclopropylmagnesium bromide (33.3 ml, 0.7 mol l^{-1} in THF, $2 \cdot 10^1$ mmol, 2 eq.) was added dropwise and the resulting reaction mixture was stirred for 2 h at room temperature. Afterwards, additional iron(III) acetylacetonate (536.4 mg, 1.519 mmol, 14.98 mol%), cyclopropylmagnesium bromide (33.3 ml, 0.7 mol l^{-1} in THF, $2 \cdot 10^1$ mmol, 2 eq.) and dry NMP (16.6 ml) were added. Stirring was continued for 3 h and then additional iron(III) acetylacetonate (134.1 mg, 0.3798 mmol, 3.75 mol%) was added. Stirring was continued for another 1 h. More iron(III) acetylacetonate (134.1 mg, 0.3798 mmol, 3.75 mol%)

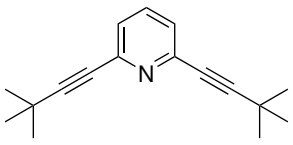
and cyclopropylmagnesium bromide (16.7 ml, 0.7 mol l^{-1} in THF, $1 \cdot 10^1$ mmol, 1 eq.) were added and the resulting mixture was stirred for 4 h at room temperature. Iron(III) acetylacetonate (134.1 mg, 0.3798 mmol, 3.75 mol%) and cyclopropylmagnesium bromide (16.7 ml, 0.7 mol l^{-1} in THF, $1 \cdot 10^1$ mmol, 1 eq.) were added again and stirring was continued overnight. Afterwards, Iron(III) acetylacetonate (134.1 mg, 0.3798 mmol, 3.75 mol%) and cyclopropylmagnesium bromide (33.3 ml, 0.7 mol l^{-1} in THF, $2 \cdot 10^1$ mmol, 2 eq.) were added yet again and stirring was continued for another day. The resulting mixture was quenched with a mixture of ice and saturated aqueous NH_4Cl (200 ml) and the resulting mixture was stirred vigorously until two clear layers formed. These layers were separated and the aqueous phase was extracted with Et_2O until the corresponding extract did only contain NMP (checked by GC-MS analysis). The combined organic phases were washed 3 times with saturated aqueous NH_4Cl. The organic phase was dried over $MgSO_4$, the solid was removed again by filtration and the filtrate was concentrated in vacuo. The residue was taken up in Et_2O (100 ml) and extracted 3 times with diluted aqueous HCl (50 ml, 1 mol l^{-1}). The combined aqueous phases were washed 3 times with CH_2Cl_2 (50 ml). Subsequently, the aqueous phase was basified with concentrated aqueous NH_3 and then extracted 3 times with CH_2Cl_2 (50 ml). The combined phases were dried over $MgSO_4$, the solid was removed again by filtration and the filtrate was concentrated in vacuo. The crude product was purified by flash chromatography (*n*-pentane : Et_2O 49:1 (v/v)). Yellowish oil (441.8 mg, 27 % of theory).

^{1}H-NMR ($CDCl_3$, 400 MHz): δ (ppm) = 7.35 (t, J = 7.7 Hz, 1 H), 6.86 (d, J = 7.7 Hz, 2 H), 1.94 (tt, J = 8.1, 4.8 Hz, 2 H), 1.03 – 0.79 (m, 9 H).

^{13}C-NMR ($CDCl_3$, 100 MHz): δ (ppm) = 161.87, 135.62, 117.91, 17.12, 9.54.

HRMS: calculated for $C_{11}H_{14}N^+$ [(M + H)$^+$] 160.1121; found 160.1123.

2,6-Bis(3,3-dimethylbut-1-yn-1-yl)pyridine



A procedure from literature was adapted.[414] 2,6-Dibromopyridine (1.777 g,

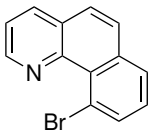
7.500 mmol, 1.000 eq.), tetrakis(triphenylphosphine)palladium(0) (433.4 mg, 0.3750 mmol, 5.00 mol%), copper(I) iodide (71.4 mg, 0.3750 mmol, 5.00 mol%) and 3,3-dimethyl-1-butyne (1.540 g, 18.75 mmol, 2.500 eq.) were placed in a two-necked flask under Ar and diisopropylamine (3.8 ml) and dry THF (15 ml) were added. The resulting reaction mixture was stirred overnight at room temperature. GC-MS showed full conversion of 2,6-dibromopyridine so the mixture was filtered and washed with CH_2Cl_2. The resulting solution was washed once with saturated aqueous NH_4Cl (50 ml) to remove copper salts. The resulting aqueous layer was extracted several times with CH_2Cl_2 until no UV-active aubstance was extracted anymore. The combined organic phases were washed once with brine, dried over $MgSO_4$, the solid was removed again by filtration and the filtrate was concentrated in vacuo. The crude product was purified by flash chromatography (*n*-pentane : Et_2O 25:1 (v/v)). The yellow solid obtained was purified further by resublimation (80 °C–120 °C, $1 \cdot 10^{-2}$ mbar). The colorless crystalline solid obtained was purified again by flash chromatography (*n*-pentane : Et_2O 50:1 (v/v)). Colorless crystals (1616.8 mg, 90 % of theory).

^{1}H-NMR ($CDCl_3$, 400 MHz): δ (ppm) = 7.50 (dd, *J* = 8.1, 7.5 Hz, 1 H), 7.23 (d, *J* = 7.8 Hz, 2 H), 1.31 (s, 18 H).

^{13}C-NMR ($CDCl_3$, 100 MHz): δ (ppm) = 144.25, 136.05, 125.80, 99.00, 78.93, 30.81, 28.02.

HRMS: calculated for $C_{17}H_{22}N^+$ [(M + H)$^+$] 240.1747; found 240.1746.

10-Bromobenzo[h]quinoline

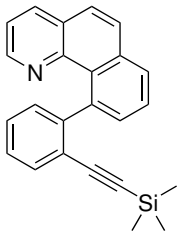


10-Bromobenzo[h]quinoline was prepared according to a literature procedure.[415] The obtained product after flash chromatography was not purified further by recrystallization. Yellowish crystals.

^{1}H-NMR ($CDCl_3$, 400 MHz): δ (ppm) = 9.12 (dd, *J* = 4.3, 1.9 Hz, 1 H), 8.18 (dd, *J* = 8.0, 1.9 Hz, 1 H), 8.11 (dd, *J* = 7.7, 1.3 Hz, 1 H), 7.89 (dd, *J* = 7.9, 1.3 Hz, 1 H), 7.78 (d, *J* = 8.8 Hz, 1 H), 7.71 (d, *J* = 8.7 Hz, 1 H), 7.57 (dd, *J* = 8.0, 4.3 Hz, 1 H), 7.47 (t, *J* = 7.8 Hz, 1 H).

^{13}C-NMR ($CDCl_3$, 100 MHz): δ (ppm) = 147.17, 146.16, 136.41, 135.76, 135.65, 128.51, 128.29, 128.07, 127.50, 126.69, 122.02, 119.62.

10-(2-((Trimethylsilyl)ethynyl)phenyl)benzo[h]quinoline



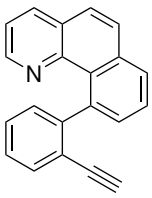
A procedure from literature was adapted.[415] ((2-Bromophenyl)ethynyl)trimethylsilane (4.900 g, 19.35 mmol, 1.200 eq.) was dissolved in dry THF (81 ml) under Ar and the resulting solution was cooled down to −78 °C. A solution of *n*-BuLi (7.7 ml, 2.5 mol l^{-1} in hexanes, 19 mmol, 1.2 eq.) was added dropwise and the resulting solution was stirred for 30 min at −78 °C. CuI (3.993 g, 20.96 mmol, 1.299 eq.) was added all at once and the resulting mixture was stirred for additional 30 min at −78 °C. Afterwards, the resulting suspension was allowed to slowly warm up until it turned into a homogeneous solution. When the solution was clear, 10-bromobenzo[h]quinoline (4.162 g, 16.13 mmol, 1.000 eq.) was added all at once and the resulting solution was allowed to warm up to room temperature. It was stirred at room temperature overnight. GC-MS showed full conversion of 10-bromobenzo[h]quinoline so Et_2O and saturated aqueous NH_3 were added, the resulting layers were separated and the aqueous phase was extracted extensively with Et_2O. The combined organic layers were dried over $MgSO_4$, the solid was removed again by filtration and the filtrate was concentrated in vacuo. The crude product was purified by flash chromatography (*n*-pentane : CH_2Cl_2 4:1 (v/v)). Brown crystals (4.9940 g, 88 % of theory).

^{1}H-NMR ($CDCl_3$, 400 MHz): δ (ppm) = 8.42 (dd, *J* = 4.3, 1.9 Hz, 1 H), 8.08 (dd, *J* = 8.0, 1.9 Hz, 1 H), 7.94 (dd, *J* = 8.0, 1.4 Hz, 1 H), 7.86 (d, *J* = 8.8 Hz, 1 H), 7.73 – 7.65 (m, 2 H), 7.53 – 7.42 (m, 4 H), 7.35 – 7.28 (m, 2 H), −0.47 (s, 9 H).

^{13}C-NMR ($CDCl_3$, 100 MHz): δ (ppm) = 150.26, 147.25, 147.16, 140.09, 135.09, 134.54, 131.08, 131.03, 129.93, 128.48, 128.32, 128.21, 128.07, 127.19, 126.84, 125.87, 125.75, 122.59, 121.08, 105.24, 95.99, −0.59.

HRMS: calculated for $C_{24}H_{22}NSi^+$ [(M + H)$^+$] 352.1516; found 352.1518.

10-(2-Ethynylphenyl)benzo[h]quinoline



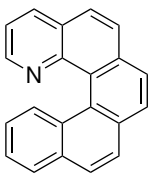
A procedure from literature was adapted.[415] 10-(2-((Trimethylsilyl)ethynyl)-phenyl)benzo-[h]quinoline (4.000 g, 11.38 mmol, 1.000 eq.) was dissolved in dry THF (50 ml) and a solution of TBAF (12.5 ml, 1.0 mol l^{-1} in THF, 13 mmol, 1.1 eq.) was added dropwise. The resulting mixture was stirred for 30 min at room temperature. GC-MS showed full conversion of 10-(2-((trimethylsilyl)ethynyl)phenyl)benzo[h]quinoline so the reaction mixture was concentrated in vacuo. The crude product was purified by flash chromatography (*n*-pentane : CH_2Cl_2 3.5:1 (v/v)). Yellowish viscous oil (3.0955 g, 97 % of theory).

^{1}H-NMR ($CDCl_3$, 400 MHz): δ (ppm) = 8.41 (dd, *J* = 4.3, 1.9 Hz, 1 H), 8.08 (dd, *J* = 8.0, 1.9 Hz, 1 H), 7.96 (dd, *J* = 7.9, 1.4 Hz, 1 H), 7.87 (d, *J* = 8.8 Hz, 1 H), 7.76 – 7.65 (m, 2 H), 7.59 (dd, *J* = 7.7, 1.4 Hz, 1 H), 7.54 (dd, *J* = 7.2, 1.4 Hz, 1 H), 7.43 (td, *J* = 7.5, 1.4 Hz, 1 H), 7.38 – 7.28 (m, 3 H), 2.47 (s, 1 H).

^{13}C-NMR ($CDCl_3$, 100 MHz): δ (ppm) = 149.89, 147.29, 147.00, 139.85, 135.20, 134.65, 132.23, 130.95, 129.69, 128.65, 128.40, 128.37, 128.33, 127.13, 127.02, 125.99, 125.81, 121.30, 121.13, 83.65, 78.82.

HRMS: calculated for $C_{21}H_{14}N^+$ [(M + H)$^+$] 280.1121; found 280.1121.

Phenanthro[3,4-h]quinoline (1-Aza[5]helicene)



A procedure from literature was adapted.[415] To 10-(2-ethynylphenyl)benzo[h]-quinoline (2.900 g, 10.38 mmol, 1.000 eq.) and $PtCl_4$ (349.7 mg, 1.038 mmol,

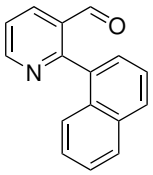
0.1000 eq.) was added dry DCE (52 ml) under Ar and the resulting mixture was heated in a closed pressure flask to 120 °C overnight. The resulting mixture was cooled down to room temperature and GC-MS showed full conversion of 10-(2-ethynylphenyl)benzo[h]quinoline so aqueous NaOH (2 mol l^{-1}) and CH_2Cl_2 were added and the resulting layers were separated. The aqueous phase was extracted 3 times with CH_2Cl_2 and the combined organic layers were dried over $MgSO_4$, the solid was removed again by filtration and the filtrate was concentrated in vacuo. The crude product was purified by flash chromatography (*n*-pentane : CH_2Cl_2 2:3 (v/v)). Orange powder (584.4 mg, 20 % of theory).

^{1}H-NMR ($CDCl_3$, 400 MHz): δ (ppm) = 8.70 (dd, *J* = 4.2, 1.9 Hz, 1 H), 8.28 (dd, *J* = 8.1, 1.9 Hz, 1 H), 8.23 – 8.14 (m, 1 H), 8.02 – 7.83 (m, 7 H), 7.57 – 7.47 (m, 2 H), 7.30 – 7.22 (m, 1 H).

^{13}C-NMR ($CDCl_3$, 100 MHz): δ (ppm) = 146.90, 146.87, 135.75, 134.40, 132.91, 132.88, 131.39, 130.73, 129.06, 128.65, 127.88, 127.79, 127.72, 127.45, 126.78, 126.70, 126.42, 126.26, 125.82, 123.28, 121.96.

Despite the potential for helical chirality of 1-aza[5]helicene and the well-known helical chirality of 1-aza[6]helicene,[415] chiral HPLC analysis of 1-aza[5]helicene dissolved in acetonitrile showed only a single peak indicating its inversion barrier to be low. It should be noted that under these conditions the enantiomers of 1-aza[6]helicene are readily separable (vide infra).

2-(Naphthalen-1-yl)nicotinaldehyde



A procedure from literature was adapted.[416] 2-Bromonicotinaldehyde (4.11 g, 22.1 mmol, 1.00 eq.), naphthalen-1-ylboronic acid (5.70 g, 33.1 mmol, 1.50 eq.), cesium carbonate (7.917 g, 24.30 mmol, 1.10 eq.) and tetrakis(triphenylphosphine)palladium(0)(1.277 g, 1.105 mmol, 0.0500 eq.) were placed in a two-necked round-bottom flask and the atmosphere was exchanged with Ar. Dry toluene (141 ml) and dry ethanol (14 ml) were added and the resulting mixture was heated to 100 °C overnight under Ar. Reaction progress was controlled by

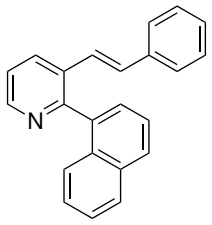
GC-MS analysis. Water and EtOAc were added to the reaction mixture and the resulting layers were separated. The aqueous phase was extracted twice with EtOAc, the combined organic layers were dried over $MgSO_4$, the solid was removed again by filtration and the filtrate was concentrated in vacuo. The crude product was purified by flash chromatography (*n*-pentane : EtOAc 3:1 (v/v)). Yellowish oil (4.445 g, 86 % of theory).

^{1}H-NMR ($CDCl_3$, 400 MHz): δ (ppm) = 9.71 (d, J = 0.9 Hz, 1H), 8.98 (dd, J = 4.8, 1.9 Hz, 1H), 8.41 (dd, J = 7.9, 1.9 Hz, 1H), 7.98 (ddt, J = 20.2, 8.2, 1.0 Hz, 2H), 7.64 – 7.48 (m, 4H), 7.53 – 7.42 (m, 2H).

^{13}C-NMR ($CDCl_3$, 100 MHz): δ (ppm) = 191.39, 162.33, 153.94, 135.42, 134.51, 133.81, 132.27, 131.04, 129.97, 129.01, 128.66, 127.33, 126.53, 125.37, 125.16, 123.23.

HRMS: calculated for $C_{16}H_{12}NO^+$ [(M + H)$^+$] 234.0913; found 234.0919.

2-(Naphthalen-1-yl)-3-styrylpyridine



A procedure from literature was adapted.[417] 2-(Naphthalen-1-yl)nicotinaldehyde (4.364 g, 18.71 mmol, 1.000 eq.), benzyltriphenylphosphonium bromide (21.45 g, 49.50 mmol, 2.646 eq.), potassium carbonate (26.19 g, 189.5 mmol, 10.13 eq.) and 18-crown-6 (49.4 mg, 0.187 mmol, 0.0100 eq.) were placed in a two-necked round-bottom flask and the atmosphere was exchanged with Ar. Dry dichloromethane (180 ml) was added and the resulting suspension was refluxed for 36 h. Reaction progress was controlled by GC-FID analysis. The reaction mixture was cooled down, water was added and the resulting layers were separated. The organic phase was washed once with water, the organic layer was dried over $MgSO_4$, the solid was removed again by filtration and the filtrate was concentrated in vacuo to obtain a yellow solid. The solid was triturated excessively with *n*-pentane until no product was dissolved anymore as indicated by GC-FID analysis of the triturates. The combined triturates were

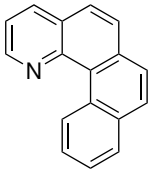
concentrated in vacuo to obtain a yellow solid. The crude product was purified by flash chromatography (*n*-pentane : EtOAc 7:1 (v/v)) and was obtained as mixture of double bond isomers. Colorless powder (5.494 g, 96 % of theory).

^{1}H-NMR ($CDCl_3$, 400 MHz): δ (ppm) = 8.71 – 8.61 (m, 1H), 8.17 (dd, *J* = 8.1, 1.6 Hz, minor diastereomer), 7.91 (d, *J* = 8.2, 2H), 7.75 – 7.36 (m, 6H), 7.30 – 7.11 (m, 6H), 7.07 (d, *J* = 16.3 Hz, minor diastereomer), 6.77 (d, *J* = 16.3 Hz, minor diastereomerz), 6.43 (d, *J* = 12.2 Hz, 1H), 6.16 (d, *J* = 12.2 Hz, 1H).

^{13}C-NMR ($CDCl_3$, 100 MHz): δ (ppm) = 158.29, 148.37, 137.71, 137.63, 136.57, 133.91, 132.95, 131.65, 131.63, 131.08, 128.93, 128.86, 128.81, 128.72, 128.57, 128.43, 128.08, 127.91, 127.59, 127.57, 127.54, 126.75, 126.57, 126.46, 126.06, 125.99, 125.81, 125.41, 125.32, 125.29, 122.86, 122.05.

HRMS: calculated for $C_{23}H_{18}N^+$ [(M + H)$^+$] 308.1434; found 308.1438.

Naphtho[2,1-h]quinoline (1-Aza[4]helicene)



A procedure from literature was followed.[418] Trifluoromethanesulfonic acid (50 ml) was added to 2-(naphthalen-1-yl)-3-styrylpyridine (5.332 g, 17.35 mmol, 1.000 eq.) placed in a round-bottom flask. The resulting mixture was stirred for 12 h. Reaction progress was controlled by GC-MS analysis. The reaction mixture was slowly poured over ice (350 g) and the resulting mixture was basified with aqueous NaOH (10 mol l^{-1}) and an orange precipitate formed in the process. The mixture was transferred to a separatory funnel and CH_2Cl_2 was added. The resulting layers were separated and the aqueous phase was extracted once with CH_2Cl_2. The combined organic phases were washed twice with water and once with brine. The organic layer was dried over $MgSO_4$, the solid was removed again by filtration and the filtrate was concentrated in vacuo to obtain an orange oil. The crude product was purified by flash chromatography (*n*-pentane : EtOAc 8:1 (v/v)). The resulting yellowish solid was purified again by flash chromatography (*n*-pentane : EtOAc 18:1 (v/v)). Yellowish crystalline solid (1.956 g, 49 %

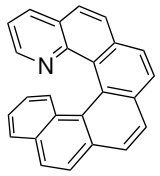
of theory).

^{1}H-NMR ($CDCl_3$, 400 MHz): δ (ppm) = 11.22 (m, 1H), 9.23 (dd, *J* = 4.2, 2.0 Hz, 1H), 8.31 (dd, *J* = 8.1, 2.0 Hz, 1H), 8.07 – 7.99 (m, 2H), 7.97 (d, *J* = 8.5 Hz, 1H), 7.98 – 7.80 (m, 3H), 7.69 (m, 1H), 7.57 (dd, *J* = 8.1, 4.2 Hz, 1H).

^{13}C-NMR ($CDCl_3$, 100 MHz): δ (ppm) = 148.59, 148.45, 136.58, 134.11, 133.76, 131.86, 130.81, 129.95, 128.78, 128.61, 127.99, 127.44, 127.02, 126.73, 126.58, 126.37, 120.72.

HRMS: calculated for $C_{17}H_{12}N^+$ [(M + H)$^+$] 230.0964; found 230.0969.

Benzo[5,6]phenanthro[3,4-h]quinoline (1-Aza[6]helicene)



A synthesis sequence from literature was followed to synthesize 1-aza[6]-helicene.[415]

^{1}H-NMR ($CDCl_3$, 400 MHz): δ (ppm) = 8.12 (dd, *J* = 8.1, 1.7 Hz, 2 H), 8.06 – 7.95 (m, 4 H), 7.95 – 7.89 (m, 3 H), 7.89 – 7.82 (m, 2 H), 7.57 (d, *J* = 8.5 Hz, 1 H), 7.19 (m, 1 H), 7.11 (dd, *J* = 8.0, 4.2 Hz, 1 H), 6.59 (m, 1 H).

^{13}C-NMR ($CDCl_3$, 100 MHz): δ (ppm) = 146.84, 145.92, 135.30, 133.44, 133.24, 131.98, 131.33, 131.15, 129.93, 128.79, 128.76, 128.29, 128.01, 127.65, 127.39, 126.68, 126.49, 126.46, 126.37, 126.21, 126.15, 124.83, 124.36, 124.29, 120.72.

(P)-1-aza[6]helicene and (M)-1-aza[6]helicene were separated using chiral HPLC. The corresponding chromatogram of the racemic mixture is shown in A.1.

Electronic circular dichroism (ECD) spectra of separated (P)-1-aza[6]helicene and (M)-1-aza[6]helicene were measured on a Jasco J-715 spectropolarimeter in a QS cuvette of 1 cm length. The corresponding spectra are shown in A.2.

General Procedures for the Preparation of Protonated Nitrogen Bases

Procedure 1 - Reasonably Water-soluble Species

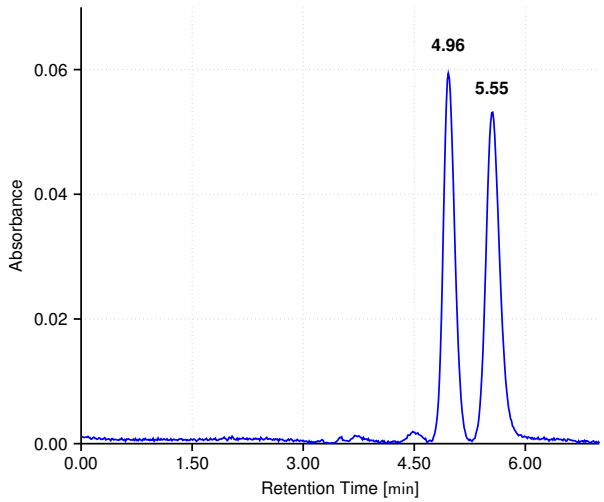


Figure A.1.: Chiral HPLC chromatogram of racemic 1-aza[6]helicene.

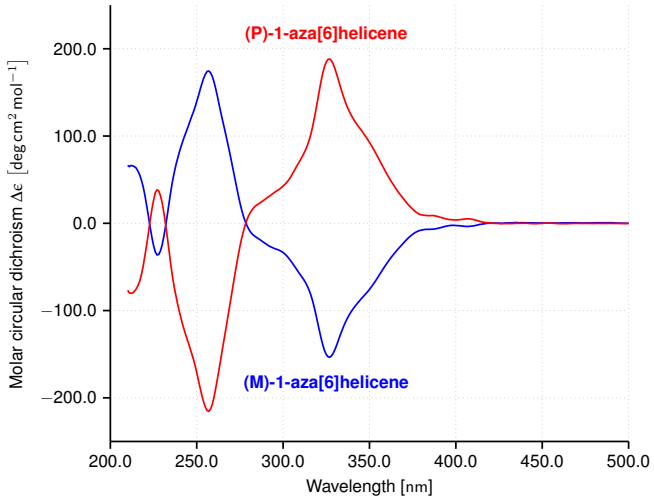
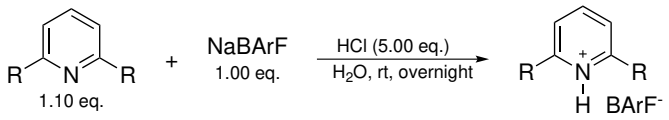
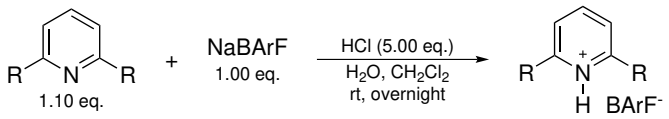


Figure A.2.: ECD spectra of separated (P)-1-aza[6]helicene and (M)-1-aza[6]helicene in MeOH.



In a typical experiment pyridine (18.4 mg, 0.233 mmol, 1.10 eq.) was suspended in deionized H_2O (15 ml) and concentrated aqueous HCl (37 % (m/m), 87.1 µl, 1.06 mmol, 5.00 eq.) was added. The resulting mixture was stirred for 5 min. Subsequently, Na[$BArF_{24}$] (188 mg, 0.212 mmol, 1.00 eq.) was added and the resulting mixture was stirred vigorously overnight. The solid precipitate formed was collected by vacuum filtration and washed extensively with deionized H_2O. The obtained solid was purified by repeated dissolution in CH_2Cl_2 and precipitation in *n*-pentane followed by decantation of the supernatant. Unless noted otherwise, the resulting product was dried at 105 °C at $1 \cdot 10^{-2}$ mbar–$1 \cdot 10^{-3}$ mbar overnight.

Procedure 2 - Water-insoluble Species

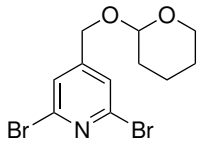


In a typical experiment 2,6-diphenylpyridine (53.9 mg, 0.233 mmol, 1.10 eq.) was dissolved in CH_2Cl_2 (7.5 ml) and deionized H_2O (15 ml) and concentrated aqueous HCl (37 % (m/m), 87.1 µl, 1.06 mmol, 5.00 eq.) were added. The resulting mixture was stirred for 5 min. Subsequently, Na[$BArF_{24}$] (188 mg, 0.212 mmol, 1.00 eq.) was added and the resulting mixture was stirred vigorously overnight. The resulting two layers were separated and the organic layer was washed two additional times with deionized H_2O. The organic phase was dried over $MgSO_4$, the solid was removed again by filtration and the filtrate was concentrated in vacuo. The obtained oil was purified by repeated dissolution in CH_2Cl_2 and precipitation in *n*-pentane followed by decantation of the supernatant. Unless noted otherwise, the resulting product was dried at 105 °C at $1 \cdot 10^{-2}$ mbar–$1 \cdot 10^{-3}$ mbar overnight.

Protonated Pyridines, Quinolines, Acridines and Amines

For pyridine, 2,6-dimethylpyridine, 2,6-diethylpyridine, 2,6-diethynylpyridine, 2,6-dipropyl-pyridine, 2,6-diisopropylpyridine and 2,6-dicyclopropylpyridine procedure 1 for the preparation of the protonated salts was followed. For all other compounds procedure 2 was followed.

2,6-Dibromo-4-(((tetrahydro-2H-pyran-2-yl)oxy)methyl)pyridine

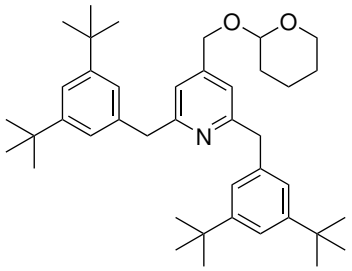


A procedure from literature adapted.[419] (2,6-Dibromopyridin-4-yl)methanol (1.0105 g, 3.786 mmol, 1.00 eq.) and pyridinium p-toluenesulfonate (47.0 mg, 0.187 mmol, 0.049 eq.) were placed in a Schlenk flask and the atmosphere was exchanged with Ar. 3,4-Dihydro-2H-pyran (3.5 ml, 39 mmol, 10 eq.) and dry CH_2Cl_2 (9 ml) were added via syringe through septum and the resulting mixture was stirred at room temperature for 4 h. Reaction progress was controlled by TLC analysis. The reaction mixture was diluted with CH_2Cl_2 and extracted once with saturated aqueous $NaHCO_3$ and H_2O, respectively, the organic layer was dried over $MgSO_4$, the solid was removed again by filtration and the filtrate was concentrated in vacuo to obtain the product in quantitative yield as yellowish oil (1.358 g).

^{1}H-NMR ($CDCl_3$, 400 MHz): δ (ppm) = 7.44 (t, *J* = 0.9 Hz, 2H), 4.78 – 4.63 (m, 2H), 4.45 (dt, *J* = 14.5, 0.9 Hz, 1H), 3.83 (m, 1H), 3.64 – 3.45 (m, 1H), 1.95 – 1.39 (m, 6H).

^{13}C-NMR ($CDCl_3$, 100 MHz): δ (ppm) = 153.39, 140.93, 125.08, 98.63, 65.96, 62.44, 30.43, 25.39, 19.28.

2,6-Bis(3,5-di-tert-butylbenzyl)-4-(((tetrahydro-2H-pyran-2-yl)oxy)-methyl)pyridine



The preparation of organozinc reagent was adapted from two procedures from literature.[420,421] Zinc powder (1.041 g, 15.92 mmol, 1.43 eq.) and LiCl (660 mg, 15.6 mmol, 1.40 eq.) were placed in a 2-necked flask equipped with a reflux condenser and the apparatus was flame-dried in medio vacuo. The atmosphere was backfilled with Ar and the apparatus allowed to cool down. Dry THF (16 ml) was added via syringe through septum. Zinc powder was activated by addition of 1,2-dibromoethane (20 µl) followed by short heating of the reaction mixture under reflux, followed by addition of trimethylsilyl chloride (20 µl), again followed by short heating under reflux. 1-(Bromomethyl)-3,5-di-tert-butylbenzene (3.156 g, 11.14 mmol, 1.00 eq.) was dissolved in dry THF (6 ml) and the resulting solution was added to the reaction mixture. An exothermic reaction was observed. The reaction mixture was heated under reflux for 2 h. Reaction progress was controlled by GC-MS analysis after quenching a sample of the reaction mixture with saturated aqueous NH_4Cl solution and judged by absence of the alkylbromide. The resulting solution (0.5 mol l^{-1}) was decanted from remaining Zinc powder and used immediately.

The procedure for Negishi cross coupling was adapted from literature.[422] 2,6-Dibromo-4-(((tetrahydro-2H-pyran-2-yl)oxy)methyl)pyridine (648.6 mg, 1.848 mmol, 1.00 eq.), $Pd(OAc)_2$ (25.0 mg, 0.111 mmol, 0.0603 eq.) and SPhos (76 mg, 0.19 mmol, 0.10 eq.) were placed in a two-necked flask and the atmosphere was exchanged with Ar. Dry THF (17 ml) was added and the resulting mixture was stirred at room temperature. The alkylzinc solution (20 ml, $1 \cdot 10^1$ mmol, 5 eq.) prepared in the previous step was added dropwise and the resulting mixture was heated to 50 °C for 3 h. Reaction progress was controlled by TLC analysis of the reaction mixture after quenching a sample of the

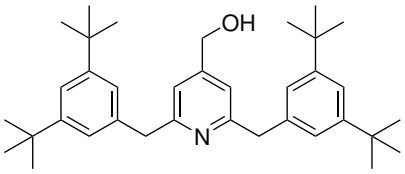
reaction mixture with saturated aqueous NH_4Cl solution. The reaction mixture was quenched with saturated aqueous NH_4Cl solution, THF was removed in vacuo at 50 °C and the resulting mixture was extracted with CH_2Cl_2 twice. The combined organic phases were dried over $MgSO_4$, the solid was removed again by filtration and the filtrate was concentrated in vacuo to obtain a brown oil. The crude product was purified by flash chromatography (*n*-pentane : CH_2Cl_2 1:1 (v/v), to CH_2Cl_2, to CH_2Cl_2 : MeOH 98:2 (v/v)). The resulting brown oil was purified again by flash chromatography (*n*-pentane : EtOAc 14:1 (v/v)). Yellow oil (932.0 mg, 84 % of theory).

^{1}H-NMR ($CDCl_3$, 400 MHz): δ (ppm) = 7.28 (t, J = 1.9 Hz, 2H), 7.15 (d, J = 1.8 Hz, 4H), 6.87 (s, 2H), 4.67 – 4.53 (m, 2H), 4.35 (d, J = 13.5 Hz, 1H), 4.16 (s, 4H), 3.75 (m, 1H), 3.50 – 3.38 (m, 1H), 1.81 – 1.41 (m, 6H), 1.30 (s, 36H).

^{13}C-NMR ($CDCl_3$, 100 MHz): δ (ppm) = 161.01, 150.91, 148.34, 138.61, 123.70, 120.39, 118.71, 98.02, 67.44, 62.08, 45.37, 34.95, 31.65, 30.50, 25.49, 19.25.

HRMS: calculated for $C_{41}H_{60}NO_2^+$ [(M + H)$^+$] 598.46179; found 598.46186.

(2,6-Bis(3,5-di-tert-butylbenzyl)pyridin-4-yl)methanol



Two procedures from literature were adapted.[423,424] 2,6-Bis(3,5-di-tertbutyl-benzyl)-4-(((tetrahydro-2H-pyran-2-yl)oxy)methyl)pyridine (932 mg, 1.56 mmol, 1.00 eq.) and pyridinium p-toluenesulfonate (39.3 mg, 0.156 mmol, 0.100 eq.) were dissolved in absolute EtOH (13 ml) and the resulting solution was heated to 55 °C under ambient atmosphere. After 1 h TLC analysis showed no conversion. Concentrated HCl (5 drops) were added to the reaction mixture and the flask was flushed with Ar. The resulting mixture was heated under reflux for 60 h until full conversion of the substrate. Reaction progress was controlled by TLC analysis. The reaction mixture was concentrated in vacuo and the residue was dissolved in CH_2Cl_2 and extracted once with saturated aqueous $NaHCO_3$. The

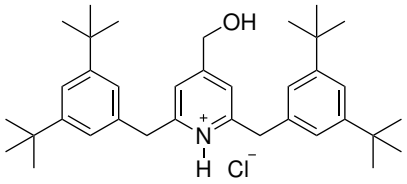
organic layer was dried over $MgSO_4$, the solid was removed again by filtration and the filtrate was concentrated in vacuo to obtain an orange oil which was dried further at 50 °C in medio vacuo. Yellowish solid (761.6 mg, 95 % of theory).

^{1}H-NMR ($CDCl_3$, 400 MHz): δ (ppm) = 7.29 (t, *J* = 1.9 Hz, 2H), 7.16 (d, *J* = 1.9 Hz, 4H), 6.90 (s, 2H), 4.56 (d, *J* = 5.6 Hz, 2H), 4.16 (s, 4H), 1.61 (t, *J* = 6.1 Hz, 1H), 1.30 (s, 36H).

^{13}C-NMR ($CDCl_3$, 100 MHz): δ (ppm) = 161.19, 150.93, 150.52, 138.62, 123.68, 120.44, 118.07, 64.09, 45.42, 34.97, 31.66.

HRMS: calculated for $C_{36}H_{52}NO^+$ [$(M + H)^+$] 514.40416; found 514.40434.

2,6-bis(3,5-di-tert-butylbenzyl)-4-(hydroxymethyl)pyridin-1-ium chloride



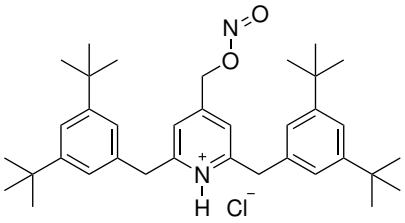
2,6-Bis(3,5-di-tert-butylbenzyl)pyridin-4-yl)methanol (50.0 mg, 0.091 mmol) was dissolved in CH_2Cl_2 and transferred to a separatory funnel filled with a mixture of H_2O (6 ml) and concentrated HCl (87.1 µl). The phases were shaken vigorously and then separated. The organic phase was dried over $MgSO_4$, the solid was removed again by filtration and the filtrate was concentrated in vacuo to obtain a colorless solid. It was triturated once with *n*-pentane and the resulting solid was dried in medio vacuo at 50 °C. Colorless solid.

^{1}H-NMR ($CDCl_3$, 500 MHz): δ (ppm) = 7.36 – 7.28 (m, 4H), 7.23 (m, 4H), 4.75 (s, 2H), 4.59 (s, 4H), 3.72 (broad s, 1H), 1.64 (s, 1H), 1.29 (s, 36H).

^{13}C-NMR ($CDCl_3$, 125 MHz): δ (ppm) = 161.92, 156.60, 151.89, 134.15, 123.92, 121.74, 120.97, 62.66, 39.21, 35.05, 31.60.

HRMS: calculated for $C_{36}H_{52}NO^+$ [$(M)^+$] 514.4042; found 514.4046.

2,6-bis(3,5-di-tert-butylbenzyl)-4-((nitrosooxy)methyl)pyridin-1-ium chloride



A procedure from literature was adapted.[425] 2,6-bis(3,5-di-tert-butylbenzyl)-4-(hydroxy-methyl)pyridin-1-ium chloride (5 mg, 0.009 mmol, 1 eq.) and tert-butyl nitrite (20 mg, 0.19 mmol, 22 eq.) were dissolved in $CDCl_3$ (0.8 ml) and the resulting mixture was directly characterized.

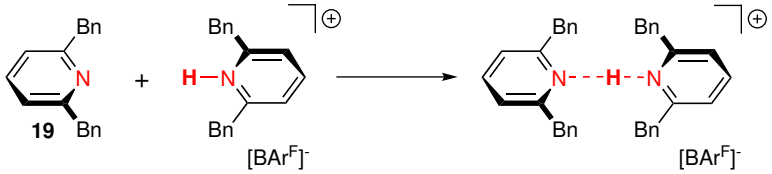
Components: tert-butyl nitrite (*t*BuONO), tert-butyl alcohol (*t*BuOH), 2,6-bis(3,5-di-tert-butylbenzyl)-4-((nitrosooxy)methyl)pyridin-1-ium (ONOPyH), 2,6-bis(3,5-di-tert-butylbenzyl)-4-(hydroxymethyl)pyridin-1-ium (HOPyH).

^{1}H-NMR ($CDCl_3$, 500 MHz): δ (ppm) = 7.37 (t, *J* = 1.7 Hz, 2H, ONOPyH), 7.35 – 7.31 (m, 4H, HOPyH), 7.20 (d, *J* = 1.8 Hz, 4H, HOPyH), 7.18 (d, *J* = 1.8 Hz, 4H, ONOPyH), 7.16 (s, 2H, ONOPyH), 5.73 (s, 2H, ONOPyH), 4.77 – 4.73 (s, 2H, HOPyH), 4.62 (s, 4H, ONOPyH), 4.56 (s, 4H, HOPyH), 1.58 (s, *t*BuONO), 1.30 (s, 36H, ONOPyH), 1.29 (s, 36H, HOPyH), 1.26 (s, *t*BuOH).

^{13}C-NMR ($CDCl_3$, 125 MHz): δ (ppm) = 161.46 (HOPyH), 157.97 (ONOPyH), 156.79 (HOPyH), 155.52 (ONOPyH), 152.18 (ONOPyH), 151.96 (HOPyH), 134.14 (HOPyH), 133.71 (ONOPyH), 123.93 (ONOPyH, HOPyH), 121.98 (ONOPyH), 121.78 (HOPyH), 121.33 (ONOPyH), 120.90 (HOPyH), 82.71 (*t*BuONO), 69.37 (*t*BuOH), 66.43 (ONOPyH), 62.70 (HOPyH), 39.45 (ONOPyH, HOPyH), 35.07 (ONOPyH), 35.05 (HOPyH), 31.58 (HOPyH), 31.55 (ONOPyH), 31.36 (*t*BuOH), 29.16 (*t*BuONO).

HRMS: calculated for $C_{35}H_{51}N_2O_2^+$ [$(M)^+$] 543.3945; found 543.3948.

A.1.5. X-ray Crystal Structure of the Proton-bound Dimer of 2,6-Dibenzylpyridine



2,6-Dibenzylpyridine **19** and the corresponding pyridinium $[BAr^F]^-$ salt were dissolved in equal amounts in CH_2Cl_2 in a vial and the vial was closed. The cap of the vial was pierced with a syringe needle to allow for slow solvent evaporation and the vial left to stand in a glove-box under N_2 for several days. Crystals suitable for X-ray analysis were obtained. Crystals were measured on a Bruker Nonius Kappa Apex2 diffractometer. The crystal was kept at 100.0(2) K during data collection. Using Olex2,[426] the structure was solved with the XT[427] structure solution program using intrinsic phasing and refined with the XL[428] refinement package using least squares minimisation. Analysis details are provided in Table A.1.

A.2. Compensation of Dispersion in Aprotic Solvents

A.2.1. General Experimental Methods

In general, unless noted otherwise, chemicals were purchased from commercial suppliers and used without further purification. Tetrahydrofuran (THF) was dried by refluxing over Na/benzophenone for several hours followed by distillation prior to use. Dry dimethylsulfoxide (DMSO) was purchased over 4 Å molecular sieves and used as received under inert conditions. Dry benzene was purchased from commercial suppliers and used as received under inert conditions. Perfluorohexane was purchased as mixture of isomers with 95 % purity and used as received. All nitrogen bases except for 2,6-dibenzhydrylpyridine, 2,6-di(prop-1-yn-1-yl)pyridine and 2,6-bis((adamantan-1-yl)ethynyl)pyridine (all three prepared

Parameter	Value
Empirical formula	$C_{70}H_{47}BF_{24}N_2$
Formula weight M_r	1382.90
Temperature [K]	100.0(2)
Crystal system	monoclinic
Space group	C2/c
a [Å]	18.8008(12)
b [Å]	17.7797(12)
c [Å]	18.5543(11)
α [°]	90
β [°]	90.787(4)
γ [°]	90
Volume [Å^3]	6201.6(7)
Z	4
ρ_{calc} [g cm^{-3}]	1.481
μ [mm^{-1}]	0.136
F(000)	2808.0
Crystal size [mm^3]	0.16 x 0.1 x 0.08
Radiation	Mo K_α (λ = 0.71073)
2Θ range for data collection [°]	4.582 to 55.062
Index ranges	-24 $\leq$ h $\leq$ 23, -23 $\leq$ k $\leq$ 23, -21 $\leq$ l $\leq$ 24
Reflections collected	32332
Independent reflections	7148 [$R_{int} = 0.0431$, $R_{sigma} = 0.0390$]
Data, restraints, parameters	7148, 324, 523
Goodness-of-fit on F^2	1.049
Final R indexes ($I \geq 2\sigma(I)$)	$R_1 = 0.0423$, $wR_2 = 0.0941$
Final R indexes (all data)	$R_1 = 0.0630$, $wR_2 = 0.1046$
$\Delta\rho_{max}$, $\Delta\rho_{min}$ [e Å^{-3}]	0.31, -0.30

Table A.1.: Crystal data and structure refinement for the proton-bound dimer of 2,6-dibenzylpyridine.

as described below) were prepared according to the description in the previous section.

For analytical thin layer chromatography (TLC), aluminium-backed silica gel 60 with fluorescence indicator F254 was used. Flash chromatography was performed on SilicaFlash® P60 from Silicycle (40 µm–63 µm).

High-resolution mass spectrometry (HRMS) for literature-unknown compounds or compounds with no published HRMS measurement results was performed on a Bruker Daltonics maXis by high performance liquid chromatography in combination with high-resolution time-of-flight mass spectrometry, unless noted otherwise.

Isothermal titration calorimetry (ITC) was performed on a MicroCal VP-ITC. The reference cell was filled with dry and freshly distilled (distilled over CaH_2) CH_2Cl_2 and the sample cell was filled with 2 ml of a solution of pyridinium $[BArF_{24}]^-$ salt in dry CH_2Cl_2. The syringe was loaded with 300 µl of a solution of the corresponding pyridine in dry CH_2Cl_2. Injection of the pyridine solution to the pyridinium solution in the sample cell was performed stepwise until adduct formation was complete as judged based on the heat evolved after every injection. The injection sequence typically consisted of an initial injection of 2 µl followed by 50 injections of 5 µl with an interval of 75 s between the injections. The initial injection was performed to minimize artifacts from syringe filling and the corresponding data point was not used in the data evaluation.[402] The initial delay before the first injection was 120 s, the stirring rate in the sample cell was 307 rpm, the reference power was 35.5 µcal s^{-1} and all titrations were performed at 5 °C, 15 °C or 25 °C.

A.2.2. Isothermal Titration Calorimetry Evaluation

All the data was analyzed using the open-source python package *pytc* (version 1.1.5).[429] To account for the excess of pyridine in most of the isolated pyridinium salts, a new fitting model was added to the available selection. The corresponding code is given here:

```
__description__ = \
"""
```

single_sitem.py

```
Model describing binding of a ligand to a single site with T in initial S.
"""
__author__ = "Robert Pollice"
__date__ = "10-10-2017"

import numpy as np

from .base import ITCModel

class SingleSiteM(ITCModel):
    """
    Binding at a single site with T in initial S solution.
    """

    def param_definition(K=1e6,dH=-4000.0,T_rel_initial=0.05):
        pass

    @property
    def dQ(self):
        """
        Calculate the heats that would be observed across shots for a
        given set of enthalpies and binding constants for each reaction.
        """

        # ---- Determine mole fractions ----
        S_conc_corr = self._S_conc
        b = S_conc_corr + self._T_conc +
            self._S_conc*self.param_values["T_rel_initial"] +
            1/self.param_values["K"]
        ST = (b - np.sqrt((b)**2 -
             4*S_conc_corr*(self._T_conc +
             self._S_conc*self.param_values["T_rel_initial"])))/2

        mol_fx_st = ST/S_conc_corr

        # ---- Relate mole fractions to heat ----
        X = self.param_values["dH"]*(mol_fx_st[1:] - mol_fx_st[:-1])

        to_return = self._cell_volume*S_conc_corr[1:]*X +
                                        self.dilution_heats

        return to_return
```

The following code shows how fitting was carried out for one specific set of data:

```
import pytc
```

```
#-------------------------------------------------------------------------
# Create a global fitting instance
g = pytc.GlobalFit()

#-------------------------------------------------------------------------
# Load in an experimental data set with a single-site binding model.
# Ignore the first shot
a = pytc.ITCExperiment("Py_DCM_5dC_ITC1.DH",
                       pytc.indiv_models.SingleSiteM,
                       shot_start=1,
                       uncertainty=0.1)

# Add the experiment to the fitter
g.add_experiment(a)

# Constraint
g.update_fixed("dilution_heat",0,a)
g.update_fixed("dilution_intercept",0,a)

#-------------------------------------------------------------------------
# Fit the data
g.fit()

#-------------------------------------------------------------------------
# Show the results
fig, ax = g.plot()
fig.savefig('Py_DCM_5dC_ITC1_fit.pdf', bbox_inches='tight')
c = g.corner_plot()
c.savefig('Py_DCM_5dC_ITC1_corner.pdf')
#print(g.fit_as_csv)
f = open('Py_DCM_5dC_ITC1_results.csv', 'a')
f.write(g.fit_as_csv)
f.close()
```

A.2.3. NMR Measurements

Standard NMR Measurements for Characterization

NMR spectra for characterization purposes were recorded either on a Bruker AVIII 400 MHz or on a Bruker AVIII 300 MHz, chemical shifts are reported in ppm, using the residual peak of the deuterated solvent as internal standard (^{1}H-NMR: $CHCl_3$ at 7.26 ppm, CD_2Cl_2 at 5.32 ppm, CD_3OD at 3.31 ppm, d^6-acetone at 2.05 ppm; ^{13}C-NMR: $CHCl_3$ at 77.16 ppm, CD_2Cl_2 at 53.84 ppm). For the designation of the observed signal coupling patterns the following abbreviations

are used: singlet (s), doublet (d), triplet (t), multiplet (m).

DOSY NMR Measurements

NMR spectra were recorded either on a Bruker AVIII 500 MHz HD spectrometer, equipped with a BBFO broadband probe or on a Bruker AVIII 600 MHz HD spectrometer, equipped with a fluorine-selective TBIF probe and a ^{19}F lock system. NMR measurements were carried out without spinning. The time between midpoints of the gradients (Δ) was chosen as 49.9 ms. The gradient length (δ) was optimized for each sample and it was also optimized for each species present in the sample, when necessary due to big differences in the diffusion constants. Standard DOSY processing, as supplied by the vendor, was used based on peak integrals. The spectrometer was calibrated using the diffusion coefficient of HDO in D_2O obtained from an NMR sample of D_2O.

Samples in CD_2Cl_2, $CDCl_3$, d^5-PhBr and 1H-perfluorohexane were prepared under dry conditions equivalent to the variable-temperature NMR samples (*vide infra*). Samples in d^4-MeOH were prepared under air and the solvent was used as provided by commercial suppliers without prior drying.

DOSY Equilibrium Constants

In this section, we explain the determination of ion pairing equilibrium constants from the DOSY NMR results. First, the hydrodynamic radii of all the species in solution were determined using the equations A.2-A.3. Based on the hydrodynamic radii the molar fractions of ion-paired species was determined using the following equations.[180] The variables are explained in the table below.

$$x_{IP} = \frac{1 - \frac{\frac{1}{r_{PyH,CDCl_3}} \cdot f_{corr} - \frac{1}{r_{PyH,IP}}}{\frac{1}{r_{PyH,MeOH}} - \frac{1}{r_{PyH,IP}}}}{1 - x_{Py}} \tag{A.4}$$

$$x_{IP} = 1 - \frac{\frac{1}{r_{BArF_{24},CDCl_3}} \cdot f_{corr} - \frac{1}{r_{BArF_{24},IP}}}{\frac{1}{r_{BArF_{24},MeOH}} - \frac{1}{r_{BArF_{24},IP}}} \tag{A.5}$$

Variable	Definition
x_{IP}	Molar fraction of ion-paired species
$r_{PyH,CDCl_3}$	Hydrodynamic radius of pyridinium in chloroform [Å]
$r_{PyH,MeOH}$	Hydrodynamic radius of pyridinium in methanol [Å]
$r_{PyH,IP}$	Hydrodynamic radius of fully ion-paired pyridinium [Å]
$r_{BArF_{24},CDCl_3}$	Hydrodynamic radius of $BArF_{24}^-$ in chloroform [Å]
$r_{BArF_{24},MeOH}$	Hydrodynamic radius of $BArF_{24}^-$ in methanol [Å]
$r_{BArF_{24},IP}$	Hydrodynamic radius of fully ion-paired $BArF_{24}^-$ [Å]
f_{corr}	Correction factor to account for differences between measurements
x_{Py}	Mole fraction of pyridine in the isolated pyridinium salt

These equations assume that the molar ratio can be determined from the respective diffusion coefficients of the species (hence the inverse of the hydrodynamic radii in the equations). Due to variance between the different measurements, the correction factor is required. It is determined from equations A.4-A.5 by finding the value that results in identical x_{IP}. The equilibrium constant of ion pairing is determined from x_{IP} using the following equation:

$$K_{IP} = \frac{x_{IP}}{(1 - x_{IP})^2 \ (1 - x_{Py}) \ c_{PyH}} \tag{A.6}$$

This procedure was applied to determine the ion pairing equilibrium constants in chloroform for a subset of the pyridinium species.

Variable-Temperature NMR Measurements

Materials made at least partially of plastic and glass syringes were dried at 55 °C at 50 mbar for at least one day. All other glassware used in the sample preparation was dried at about 160 °C for at least one day. All samples were prepared inside a glove-box under N_2 atmosphere and transferred into Young NMR tubes. Substances and solvent amounts were weighed with a precision of 0.01 mg.

NMR spectra were recorded either on a Bruker AVIII 500 MHz equipped with a BBFO broadband probe or on a Bruker AVIII 600 MHz HD spectrometer, equipped with a fluorine-selective TBIF probe and a ^{19}F lock system. Chemical shifts are reported in ppm, using the residual peak of the deuterated solvent as internal standard (^1H-NMR: CD_2Cl_2 at 5.32 ppm), when deuterated solvents were used. For the undeuterated solvents, the reference shifts were determined with tetra-

methylsilane (TMS) as internal standard at room temperature using a single scan measurement. The corresponding reference shifts are provided in Table A.2. We performed a subset of the ^{1}H-NMR measurements in 1H-perfluorohexane both with locking using a ^{19}F-Lock and without locking. The corresponding results were identical within the experimental errors. The measurements using 1-fluorohexane, 1,2-difluorobenzene and bromocyclohexane were carried out without locking.

Solvent	^{1}H-NMR Chemical Shifts [ppm]
1H-Perfluorohexane	5.914
1-Fluorohexane	4.375, 4.280.
1,2-Difluorobenzene	6.978, 6.898.
Bromocyclohexane	4.157, 2.106, 1.775, 1.547, 1.363.

Table A.2.: ^{1}H-NMR reference shifts of four of our solvents with respect to TMS as internal standard at room temperature.

Temperatures of samples inside the spectrometer were calibrated using a Pt100 resistance thermometer fixed inside an NMR tube over the whole temperature range investigated. Temperatures from the Pt100 resistance thermometer were validated using the MeOH NMR thermometer. Samples were not spun during measurement. NMR spectra were recorded every 5 K–10 K. Relative amounts of nitrogen base and the corresponding protonated species were determined by the relative integrals with respect to the ^{1}H signals of the counterions tetrakis[3,5-bis(trifluoromethyl)phenyl]borate ($[BArF_{24}]^{-}$) or the fluorous BAr_{f6}^{-} (*vide infra*) using a relaxation delay (D1) of 30 s.

A.2.4. Variable-Temperature NMR Evaluation

The VT-NMR data of the proton-bound dimers was evaluated using the master equation established in the work of chapter 2:

$$T = \frac{\Delta_R H}{\Delta_R S - R\, ln \left\{ \frac{\frac{\delta_H - \delta(NH)}{\delta(N_2H) - \delta_H}}{c \left[e - \frac{\delta_H - \delta(NH)}{\delta(N_2H) - \delta(NH)} \right]} \right\}} \tag{A.7}$$

In this equation there are three unknown parameters for the systems under investigation which have to be determined by means of nonlinear fitting from the experimental data: $\Delta_R H$, $\Delta_R S$ and $\delta(N_2H)$. Notably, for systems where $\delta(N_2H)$ was known from experiment equation A.7 was still used but with only two unknown parameters. Table C.1.1 explains all the parameters in the master equation.

Variable	Definition
T	Temperature [K]
$\Delta_R H$	Enthalpy of reaction [kcal mol^{-1}]
$\Delta_R S$	Entropy of reaction [kcal mol^{-1}]
R	Universal gas constant
c	Total concentration of the mobile proton [mol l^{-1}]
e	Relative excess of the basic species with respect to the mobile proton ($e \geq 0$)
δ_H	^{1}H-NMR shift of the mobile proton at temperature T
$\delta(NH)$	^{1}H-NMR shift of the protonated nitrogen base (**NH**)
$\delta(N_2H)$	^{1}H-NMR shift of the proton-bound dimer (**N_2H**)

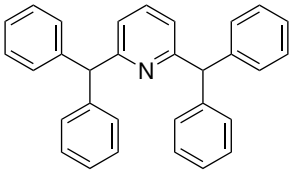
Table A.3.: Explanation of parameters of the master equation for proton-bound dimer equilibria.

Using the master equation, we evaluated all our VT-NMR data.

A.2.5. Synthesis of Test Substances

In general, synthetic procedures and therefore yields were not optimized and only carried out once in big scale to obtain enough of all the test substances for all subsequent experiments.

2,6-Dibenzhydrylpyridine

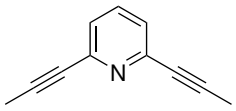


A literature procedure was adapted.[430] 2,6-Pyridinecarboxaldehyde (504 mg, 1.22 mmol, 1.00 eq.) was placed in a round-bottom flask and dissolved in dry benzene (2.5 ml). The resulting mixture was cooled down with an ice bath and triflic acid (7.5 ml) was added dropwise via syringe. The resulting orange solution was stirred overnight. The reaction mixture was poured onto ice and aqueous sodium hydroxide (2 M) was added until the aqueous phase was basic. The resulting aqueous layer was extracted several times with CH_2Cl_2 until no UV-active substance was extracted anymore. The combined organic phases were dried over $MgSO_4$, the solid was removed again by filtration and the filtrate was concentrated *in vacuo*. The crude product was purified by flash chromatography (*n*-pentane : CH_2Cl_2 2:3 (v/v)). Yellowish crystals (1.162 g, 75 % of theory).

^{1}H-NMR (CD_2Cl_2, 400 MHz): δ (ppm) = 7.55 (t, J = 7.7 Hz, 1 H), 7.30 – 7.15 (m, 20 H), 7.03 (d, J = 7.8 Hz, 2 H), 5.54 (s, 2 H).

^{13}C-NMR (CD_2Cl_2, 100 MHz): δ (ppm) = 162.67, 143.66, 137.38, 129.71, 128.54, 126.67, 121.75, 59.43.

2,6-Di(prop-1-yn-1-yl)pyridine



A literature procedure was adapted.[414] 2,6-Dibromopyridine (1.187 g, 5.011 mmol, 1.000 eq.), tetrakis(triphenylphosphine)palladium(0) (289 mg, 0.250 mmol, 4.99 mol%), copper(I) iodide (49.3 mg, 0.259 mmol, 5.17 mol%) and propyne (24.4 ml, 0.5 mol l^{-1} in heptane, $1 \cdot 10^1$ mmol, 3 eq.) were placed in a two-necked flask under Ar and diisopropylamine (3.0 ml) and dry THF (25 ml) were added. The resulting reaction mixture was stirred at room temperature for 6.5 days. GC-MS showed conversion to be incomplete and not

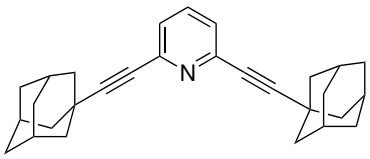
continue anymore. Hence, tetrakis(triphenylphosphine)palladium(0) (146 mg, 0.126 mmol, 2.52 mol%) and propyne (12.2 ml, 0.5 mol l^{-1} in heptane, 6 mmol, 1 eq.) were added and the reaction continued for 2 days. GC-MS showed conversion to be incomplete and not continue anymore. Therefore, tetrakis(triphenylphosphine)palladium(0) (152 mg, 0.131 mmol, 2.62 mol%), copper(I) iodide (48.0 mg, 0.252 mmol, 5.03 mol%) and propyne (12.2 ml, 0.5 mol l^{-1} in heptane, 6 mmol, 1 eq.) were added and the reaction stirred overnight. GC-MS showed full conversion of 2,6-dibromopyridine so the mixture was filtered over Celite® and washed with CH_2Cl_2. The resulting solution was washed once with saturated aqueous NH_4Cl (50 ml) to remove copper salts. The resulting aqueous layer was extracted several times with CH_2Cl_2 until no UV-active aubstance was extracted anymore. The combined organic phases were dried over $MgSO_4$, the solid was removed again by filtration and the filtrate was concentrated in vacuo. The crude product was purified by flash chromatography (*n*-pentane : EtOAc 8:1 (v/v)). The brownish solid obtained was purified further by resublimation (80 °C–105 °C, $1 \cdot 10^{-2}$ mbar). Yellowish crystals (637 mg, 82 % of theory).

^{1}H-NMR ($CDCl_3$, 400 MHz): δ (ppm) = 7.53 (t, *J* = 7.8 Hz, 1 H), 7.23 (d, *J* = 7.8 Hz, 2 H), 2.05 (s, 6 H).

^{13}C-NMR ($CDCl_3$, 100 MHz): δ (ppm) = 144.06, 136.41, 125.26, 87.01, 79.47, 4.43.

HRMS: calculated for $C_{11}H_{10}N^+$ [(M + H)$^+$] 156.0808; found 156.0809.

2,6-Bis((adamantan-1-yl)ethynyl)pyridine



A literature procedure was adapted.[414] 2,6-Dibromopyridine (1.194 g, 5.038 mmol, 1.000 eq.), tetrakis(triphenylphosphine)palladium(0) (289 mg, 0.250 mmol, 4.96 mol%), copper(I) iodide (56.7 mg, 0.298 mmol, 5.91 mol%) and 1-ethynyladamantane (2.021 g, 12.61 mmol, 2.503 eq.) were placed in a two-necked flask under Ar and diisopropylamine (2.5 ml) and dry THF (10 ml) were added. The resulting reaction mixture was stirred at room temperature

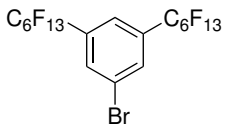
overnight. GC-MS showed full conversion of 2,6-dibromopyridine so the mixture was filtered over Celite® and washed with CH_2Cl_2. The resulting solution was washed once with saturated aqueous NH_4Cl (50 ml) to remove copper salts. The resulting aqueous layer was extracted several times with CH_2Cl_2 until no UV-active aubstance was extracted anymore. The combined organic phases were dried over $MgSO_4$, the solid was removed again by filtration and the filtrate was concentrated in vacuo. The crude solid was triturated with *n*-pentane. The residual solid was purified by flash chromatography (*n*-pentane : EtOAc 50:1 (v/v)). The yellowish solid was dried in medio vacuo at 65 °C overnight. The yellowish residual solid obtained was sufficiently pure. Yellowish crystals (1.809 g, 91 % of theory).

^{1}H-NMR ($CDCl_3$, 400 MHz): δ (ppm) = 7.49 (t, *J* = 7.8 Hz, 1 H), 7.23 (d, *J* = 7.8 Hz, 2 H), 1.97 (s, 12 H), 1.86 (d, *J* = 2.9 Hz, 3 H), 1.70 (t, *J* = 2.7 Hz, 12 H), 1.66 (t, *J* = 3.2 Hz, 3 H).

^{13}C-NMR ($CDCl_3$, 100 MHz): δ (ppm) = 144.31, 136.04, 125.84, 98.79, 79.28, 42.51, 36.47, 30.16, 28.03.

HRMS: calculated for $C_{29}H_{34}N^+$ [(M + H)$^+$] 396.2686; found 396.2684.

1-bromo-3,5-bis(perfluorohexyl)benzene



Two procedures from literature were modified.[431,432] 1,3,5-Tribromobenzene (6.194 g, 19.68 mmol, 1.000 eq.) was dissolved in dry DMSO and copper powder (5.020 g, 79.00 mmol, 4.014 eq.) was added. The atmosphere was exchanged three times with argon and the resulting suspension was heated to 120 °C. At 120 °C, a solution of perfluoro-*n*-hexyliodide (17.500 g, 39.24 mmol, 1.994 eq.) was added dropwise via syringe through septum over the course of 1 hour. The resulting mixture was stirred at 120 °C overnight. The reaction mixture was allowed to cool down to room temperature and water (50 ml) was added. The resulting green precipitate was collected by vacuum filtration using a glass-sintered funnel and the filter cake was washed with Et_2O and dried by letting air pass through for several minutes. The filter cake was continuously extracted with Et_2O

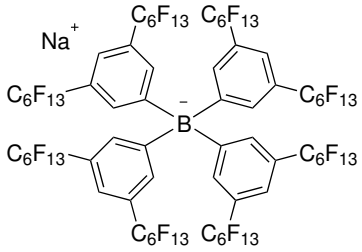
using a Soxhlet apparatus for two hours. The filtrate was extracted several times with Et_2O and combined with the Soxhlet extract. The combined solutions were dried over $MgSO_4$, the insolube solid was removed by filtration and the filtrate was concentrated in vacuo to obtain a yellow oil. The oil was dissolved in perfluorohexane and the fluorous phase was washed twice with acetone. The fluorous phase was concentrated in vacuo to obtain a yellowish oil which solidified to obtain yellowish crystals upon drying in medio vacuo overnight. The crystals were purified by Kugelrohr distillation (125 °C, $1 \cdot 10^{-1}$ mbar). Colorless crystals (11.06 g, 16.3 mol% of 1-3,5-tris(perfluorohexyl)benzene as inseparable impurity on the basis of ^{1}H-NMR analysis, 71 % of theory). The product was used for the subsequent step.

^{1}H-NMR ($CDCl_3$, 300 MHz): δ (ppm) = 8.04 (s, 3 H, B), 7.96 (s, 1 H, A), 7.74 (s, 1 H, A).

^{19}F-NMR ($CDCl_3$, 282 MHz): δ (ppm) = -80.78 (m, 6 F, A+B), -111.10 (m, 4 F, A), -111.35 (m, 4 F, B), -121.37 (m, 4 F, A+B), -121.66 (m, 4 F, A), -121.95 (m, 4 F, B), -122.78 (m, 4 F, A+B), -126.13 (m, 4 F, A+B).

Legend: A = 1-bromo-3,5-bis(perfluorohexyl)benzene
B = 1-3,5-tris(perfluorohexyl)benzene.

Sodium tetrakis(3,5-bis(perfluorohexyl)phenyl)borate ($NaBAr_{f6}$)



A literature procedure was modified.[164] 1-Bromo-3,5-bis(perfluorohexyl)benzene (8.593 g, purity 83.7 mol%, 9.069 mmol, 3.94 eq.) was placed in a flame-dried 2-necked flask under Ar and dissolved in dry Et_2O (300 ml). The resulting solution was cooled down to –78 °C using an acetone/dry ice bath. After 20 min in the cooling bath, *t*-BuLi (10.7 ml, 1.7 mol l^{-1} in *n*-pentane, 18.19 mmol, 7.91 eq.) was added dropwise via syringe through a septum over a period of 15 min. The resulting mixture was stirred for 1 h at –78 °C. Afterwards, BCl_3 (2.3 ml,

1.0 mol l^{-1} in *n*-hexane, 2.3 mmol, 1.0 eq.) was added dropwise via syringe through a septum over a period of 5 min. The resulting mixture was stirred for 2 h at −78 °C. Afterwards, the cooling bath was removed and the reaction mixture was allowed to warm up to room temperature for 3 h. The reaction mixture was poured into H_2O (350 ml) and the resulting mixture was stirred vigorously. Solid NaCl was added until the solution was saturated (i.e. additional solid did not dissolve anymore). The mixture was extracted three times with Et_2O, the combined extracts were dried over $MgSO_4$, the insolube solid was removed by filtration and the filtrate was concentrated in vacuo to obtain a yellow oil. Flash chromatography was performed. The product was dissolved in Et_2O. The column (100 g of silica gel) was prepared in CH_2Cl_2 : Et_2O 1 : 1 and eluted with 1000 ml of CH_2Cl_2 : Et_2O 1 : 1, 250 ml of CH_2Cl_2 : MeOH 98 : 2, 250 ml of CH_2Cl_2 : MeOH 95 : 5 and 2500 ml of CH_2Cl_2 : MeOH 90 : 10. The product was obtained as yellow turbid oil. This oil was dissolved in hot perfluorohexane and filtered to remove insoluble impurities. CH_2Cl_2 was added to the filtrate so that a second layer of CH_2Cl_2 remained on top of the perfluorohexane solution, and the resulting mixture was left to stand on air in a beaker to allow the solvents to evaporate and the product to crystallize. The crystals were collected by filtration and washed with cold (about −10 °C) perfluorohexane. The resulting mother liquor was used for subsequent crystallizations by repeated addition of CH_2Cl_2 in the same way as before and leaving the resulting mixture in a beaker so that the solvents could evaporate. The crystals were collected using gravitation filtration with a glass-sintered funnel in order to avoid evaporation of solvent inside the frit. In total, three cycles of crystallization were carried out in that manner. The obtained crystals were dried in medio vacuo. Colorless crystals (4.114 g, 62 % of theory).

Remark: The product is readily soluble in ethyl acetate or acetone. However, these solvents should be avoided as the product forms stable complexes with these species and residual amounts of both these solvents are very hard to be removed again afterwards.

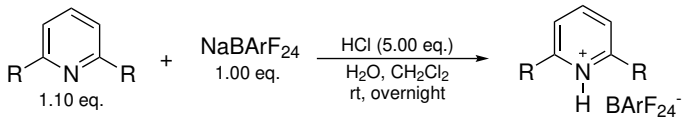
^{1}H-NMR (d^6-acetone, 400 MHz): δ (ppm) = 7.74 – 7.68 (m, 8 H), 7.58 (s, 4 H), 2.83 – 2.78 (m, 3 H).

^{19}F-NMR (d^6-acetone, 282 MHz): δ (ppm) = -81.92 – -82.09 (m, 12 F), -111.60

(t, J = 13.8 Hz, 8 F), -122.01 – -122.47 (m, 8 F), -123.09 – -123.46 (m, 8 F), -123.67 (s, 8 F), -126.94 – -127.27 (m, 8 F).

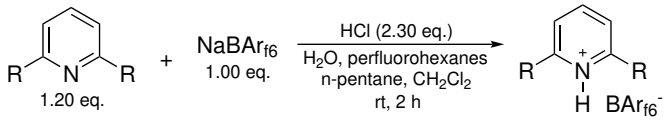
General Procedures for the Preparation of Protonated Nitrogen Bases

Procedure 1 - $BArF_{24}^-$-salts



The procedure established in the work of chapter 2 was modified. In a typical experiment 2,6-di(prop-1-yn-1-yl)pyridine (36.3 mg, 0.234 mmol, 1.10 eq.) was dissolved in CH_2Cl_2 (7.5 ml) and deionized H_2O (15 ml) and concentrated aqueous HCl (37 % (m/m), 87.1 µl, 1.06 mmol, 5.00 eq.) were added. The resulting mixture was stirred for 5 min. Subsequently, Na[$BArF_{24}$] (188 mg, 0.212 mmol, 1.00 eq.) was added and the resulting mixture was stirred vigorously for 2 h. The resulting two layers were separated and the organic layer was washed two additional times with deionized H_2O. The organic phase was dried over $MgSO_4$, the solid was removed again by filtration and the filtrate was concentrated in vacuo. The resulting product was dried at 100 °C at $1 \cdot 10^{-2}$ mbar–$1 \cdot 10^{-3}$ mbar overnight.

Procedure 2 - BAr_{f6}^--salts



In a typical experiment 2,6-diphenylpyridine (14.7 mg, 0.0635 mmol, 1.20 eq.) was dissolved in *n*-pentane(1.9 ml) and CH_2Cl_2 (1.9 ml). Perfluorohexane (1.9 ml), deionized H_2O (3.8 ml) and concentrated aqueous HCl (37 % (m/m), 10.0 µl, 0.122 mmol, 2.30 eq.) were added to the mixture to obtain a triphasic system. The resulting mixture was stirred for 5 min. Subsequently, Na[BAr_{f6}] (153 mg, 0.0530 mmol, 1.00 eq.) was added and the resulting mixture was stirred vigorously for 2 h. From the three layers, the nonaqueous layers were separated and the aqueous phase was extracted three times with a mixture of CH_2Cl_2 and perfluorohexane. The combined non-aqueous phases were dried over $MgSO_4$, the solid was removed again by filtration and the filtrate was concentrated in vacuo. The resulting product was dried at 100 °C at $1 \cdot 10^{-2}$ mbar–$1 \cdot 10^{-3}$ mbar overnight.

Protonated Pyridines, Quinolines, Acridines and Amines

For the synthesis of the $BArF_{24}$-salts of 2,6-bis((adamantan-1-yl)ethynyl)pyridine and 2,6-di(prop-1-yn-1-yl)pyridine, procedure 1 was followed. For the synthesis of the BAr_{f6}^--salts of pyridine and 2,6-dimethylpyridine procedure 2 was followed with modifications, *n*-pentane and CH_2Cl_2 were not used as solvents during the reaction. For the synthesis of the BAr_{f6}^--salts of 2,6-diethylpyridine, 2,6-diisopropylpyridine, 1,2,3,4,5,6,7,8-octahydroacridine, 2,6-dineopentylpyridine, 2,6-dicyclohexylpyridine and 2,6-bis(cyclohexylmethyl)pyridine procedure 2 was followed with modifications, CH_2Cl_2 was not used as solvents during the reaction. For all other nitrogen bases, procedure 2 was followed to synthesize the BAr_{f6}^--salts.

B. Computational Details

This section provides the computational details for chapters 2-5.

B.1. Compensation of Dispersion in Dichloromethane

The Orca suite (Version 3.0.3, unless noted otherwise)[433] was employed for all density functional theory (DFT) calculations, unless noted otherwise, for all frequency calculations, unless noted otherwise, and for all DLPNO-CCSD(T)[30,37] calculations.

The polarization-consistent basis sets[92–94] were employed for DFT calculations, unless noted otherwise. These basis sets are optimized for DFT calculations and both are computationally very efficient and show well-defined and fast convergence to the complete basis set (CBS) limit.[93] In order to improve computational efficiency even further, truncated versions of these basis sets were employed.[95] Analogously to the previously proposed truncation of the f-functions from all atoms and the d-functions from H of the pc-2 basis set to get the so-called pc-2-sp(d) basis set,[95] by removing the g-functions from all atoms and the f-functions from H of the pc-3 basis set we defined the pc-3-spd(f) basis set. It should be noted here, that the pc-2 basis set has triple-zeta quality and the pc-3 hence has quadruple-zeta quality. CBS extrapolation was performed according to the following formula[93] where L denotes the highest angular momentum and n_S denotes the number of s functions included in the basis set, respectively.

$$E = E_\infty + A\,(L+1)\,e^{-B\sqrt{n_S}} \tag{B.1}$$

Unless noted otherwise, we performed two-point CBS extrapolations using the

energies evaluated with the pc-2-sp(d) and pc-3-spd(f) basis sets, respectively, at geometries optimized using the pc-2-sp(d) basis set. In order to do that, we assumed the value of B to be equal to 5.5 as suggested in literature for two-point CBS extrapolations with the parent pc-n basis sets.[93] L was taken to be 2 and 3 for pc-2-sp(d) and pc-3-spd(f), respectively, and n_S was taken as 4 and 6, respectively. However, in general, the resulting bond dissociation energy (BDE) from the pc-3-spd(f) basis-set calculations was reasonably converged to the CBS-limit already. CBS-limit corrections to the BDE typically amounted to less than 0.5 kcal mol^{-1} with respect to energies evaluated using the pc-3-spd(f) basis-set.

Grimme's B97-D3 GGA exchange-correlation functional[434] with the atom-pairwise dispersion correction using the Becke-Johnson damping scheme[52,435] (D3BJ) was used for most calculations, especially for most geometry optimizations. This functional was shown in extensive benchmark studies[83,84] to provide good accuracy of dissociation energies in both H-bonded and dispersion-bound complexes at only moderate computational cost. In addition, Truhlar's meta-GGA exchange-correlation functional M06-L,[43] which was designed for good performance, amongst other areas, in main-group thermochemistry and noncovalent interactions, was also employed. Additional calculations for comparison were done using BP86.[89,90]

Density-fitting (also called Resolution of Identity, RI)[436–438] was employed using the RI-J approximation to further speed up self-consistent field (SCF) calculations using the TZV/J auxiliary basis set in conjunction with the pc-2-sp(d) basis set and the QZVPP/J auxiliary basis set in conjunction with the pc-3-spd(f) basis set. Accompanying errors were shown to be negligible (typically on the order of 0.1 kcal mol^{-1}).

For estimating zero point energies (ZPEs) a computationally more efficient alternative compared to frequency calculations at the level of theory used for geometry optimization, which can become computationally the most expensive step for big molecules, is employing the HF-3c method.[439] This method is based on Hartree-Fock calculations with a small Gaussian AO basis set together with 3 atom-pairwise corrections correcting for London Dispersion, basis set superposition error and short-ranged basis set incompleteness effects.[439] It provides good molecular structures in geometry optimizations and is therefore also well-suited

for frequency calculations. Hence, geometries were reoptimized at the HF-3c level of theory and frequency calculations were performed. All harmonic frequencies were scaled with a factor of 0.86 for HF-3c[439] for ZPE estimation. The obtained results were compared to the results of frequency calculations using the B97-D3/pc-2-sp(d) level of theory.

Zero point energies (ZPE) and enthalpy corrections were determined by analytical frequency calculations at the B97-D3/pc-2-sp(d) level of theory. All harmonic frequencies were used unscaled. For several of the bigger molecules it was not computationally affordable to avoid a few negative frequencies (less than 5) above –30 cm^{-1} which are an artifact of numerical noise in the computations. They would require significant additional computational effort to resolve. For the calculation of the corresponding ZPE-corrections these additional vibrations were ignored since their influence on the overall ZPE-correction is negligible. Entropies were computed via the Q-RRHO approach.[439] Imaginary harmonic frequencies were assumed to have the opposite positive value in order to estimate their contribution towards the overall entropy. Very low harmonic frequencies contribute the most towards the vibrational entropy. Small negative frequencies are indicative of a shallow potential energy surface (PES) so at the true minimum geometry the corresponding positive frequencies would also be very low. Therefore, assuming the negative frequencies to be positive provides better estimates for the entropy rather than neglecting them. It should also be noted that the entropy of mixing for chiral structures was neglected but the entropy of mixing of multiple compounds was accounted for. In addition to that, when BP86 was employed as functional, analytical frequency calculations were performed at the BP86/pc-2-sp(d) level of theory. Analytical frequency calculations with the M06-L functional were performed using the Gaussian 09 (revision D.01)[440] software package using density fitting with an automatically created fitting basis set (keyword ”Auto”).

Domain-based local pair natural orbital coupled cluster theory with single, double and perturbative triple excitations (DLPNO-CCSD(T))[30,37] method was used for single-point energy calculations on the basis of DFT geometries using the default ”NormalPNO” setup ($T_{CutPairs} = 10^{-4}$, $T_{CutPNO} = 3.33 \cdot 10^{-7}$, $T_{CutMKN} = 10^{-3}$).[33] Two-point CBS extrapolation was performed by extrapolating separately the Hartree-Fock (HF) SCF energies and the Coupled-Cluster

(CC) correlation energies using the approach adopted by Neese,[441] as implemented in Orca using the ”Extrapolate” keyword, on the basis of the single point energies employing the cc-pVDZ and cc-pVTZ basis sets. Additionally, the CBS-extrapolated counterpoise-correction (CP)[442] was calculated for the bond dissociation energies (BDEs) by performing the single point energy CBS extrapolations of the two fragments both in the monomer and dimer basis. This procedure is equivalent to calculating the CP-corrected single point energies and extrapolating these CP-corrected cc-pVDZ and cc-pVTZ single point energies to the CBS limit.

The PSI4 electronic structure program (Version 1.0rc)[443] was employed for all density fitting symmetry-adapted perturbation theory (SAPT) calculations.[115,444,445] Calculations were performed with geometries optimized at the B97-D3/pc-2-sp(d) level of theory using the jun-cc-pVDZ basis set[446] and employing the empirically modified sSAPT0 method[74] that benefits from favorable error cancellation between basis set incompleteness, especially in conjunction with the jun-cc-pVDZ (jaDZ), and inappropriate level of theory.[74] Deformation energies of the dissociated molecules at the proton-bound dimer geometries were computed both at the B97-D3/CBS and at the DLPNO-CCSD(T)/CBS level of theory. Total interaction energies from sSAPT0/jun-cc-pVDZ, B97-D3/CBS and DLPNO-CCSD(T)/CBS were compared with each other.

All calculations to estimate the free energy of solvation using the SMD model[6] were carried out using the Gaussian 09 (revision D.01)[440] software package. Free energies of solvation were computed according to their parametrization within the SMD model via thermodynamic cycles. Therefore, geometries were reoptimized both at the M05-2X/6-31g* level of theory and single points energies were calculated at the SMD/M05-2X/6-31g* level of theory. The difference between the respective absolute electronic energies yields the free energy of solvation at 25 °C.[6] It should be noted here, that for the parametrization of SMD[6] geometries were not reoptimized in solution either. Nevertheless, the effect of reoptimization when the solvent was included was also investigated by reoptimization at the SMD/M05-2X/6-31g* level of theory and taking the difference between the respective absolute electronic energies at the M05-2X/6-31g* and the SMD/M05-2X/6-31g* levels of theory.

All calculations to estimate the free energy of solvation using COSMO-RS[5,105] were carried out using ADF2014[447] and the corresponding COSMO-RS module (ADF COSMO-RS 2016).[448] Free energies of solvation were computed according to their parametrization within the COSMO-RS model in the ADF implementation.[448] The ADF combi2005 parameters were used. Single point calculations and COSMO calculations were performed both using B97-D3/pc-2-sp(d) geometries and BP86/TZP geometries at the BP86/TZP level of theory.

Reduced density gradient (RDG) isosurfaces of proton-bound dimers were generated using NCIPLOT[78] and visualized using PyMOL (Version 1.8.4)[449] on the basis of B97-D3/pc-2-sp(d) geometries. Densities were obtained from promolecular estimates. The density cut-offs were chosen on the basis of the reduced density behavior in the 2D-NCI s(ρ) plot for each system individually ($\rho < 0.030$ au for proton-bound dimers of compounds **19** and **28** and $\rho < 0.033$ au for compound **25**). The RDG cut-off was set to s = 0.3. All interactions with at least a fraction of 0.90 of the density from only one of the two fragments of the proton-bound dimer were turned off so that only intermolecular interactions were plotted.

The p-test for the determination of the probability of a particular enthalpy-entropy compensation (EEC) plot being explainable as a consequence of random experimental errors was based on the BASIC program written by Perez-Benito[110] for the analysis of EEC plots obtained from Eyring analysis of the temperature dependence of reaction rates. The code was modified to be applicable to EEC plots obtained from the temperature dependence of equilibrium constants and compiled with the program FreeBASIC (Version 1.05.0 for Windows) using the language version ”qb”. The source code of the p-test is provided in a separate file.

All calculations for the competition between bond dissociation NH-N bond of a proton-bound dimer and the O-NO bond of an alkyl nitrite were performed using unrestricted Kohn-Sham calculations and unrestricted HF and DLPNO-CCSD(T) calculations. The Orca suite (Version 4.0.0)[433] was employed for geometry optimizations at the B97-D3/pc-2-sp(d) level of theory. Calculations were speeded up significantly using the resolution of identity employing an automatically generated fitting basis set with the keyword *AutoAux*. Frequency calculations were

performed at the same level of theory using the Gaussian 09 (revision D.01)[440] software package using an automatically generated fitting basis set with the keyword *Auto*. DFT energies were extrapolated to the CBS limit on the basis of single points at the B97-D3/pc-2 and B97-D3/pc-3 levels of theory again using the resolution of identity in combination with the keyword def2/J as fitting basis set. DLPNO-CCSD(T) calculations were performed using Orca (Version 4.0.0) using the same procedure as described before but using the unrestricted implementation. For the O-NO bond cleavage calculations using DLPNO-CCSD(T) no CP corrections were applied.

B.2. Compensation of Dispersion in Aprotic Solvents

Orca (Version 4.0.1)[433,450] was employed for all density functional theory (DFT) calculations, for all frequency calculations and for all DLPNO-CCSD(T)[30,37] calculations.

The polarization-consistent basis sets[92–94] were employed for all DFT calculations. These basis sets are optimized for DFT calculations and both are computationally very efficient and show well-defined and fast convergence to the complete basis set (CBS) limit[93]. To improve computational efficiency even further, truncated versions of these basis sets were employed[95] for geometry optimizations. It should be noted here, that the pc-2 basis set has triple-zeta quality and, hence, the pc-3 has quadruple-zeta quality. CBS extrapolation was performed according to the following formula[93] where L denotes the highest angular momentum and n_S denotes the number of s functions included in the basis set, respectively.

$$E = E_\infty + A\,(L+1)\,e^{-B\sqrt{n_S}} \tag{B.2}$$

Unless noted otherwise, we performed two-point CBS extrapolations using the energies evaluated with the pc-2 and pc-3 basis sets, respectively, at geometries optimized using the pc-2-sp(d) basis set, using the optimized extrapolation parameters reported in the literature.[93] Notably, the resulting bond dissociation energy (BDE) from the pc-3 calculations was reasonably converged to the CBS-

limit already. CBS-limit corrections to the BDE typically amounted to less than 0.5 kcal mol^{-1} with respect to pc-3 energies.

Grimme's B97-D3 GGA exchange-correlation functional[91] with the atom-pairwise dispersion correction using the Becke-Johnson damping scheme[52,435] (D3BJ) was used for most calculations, especially for most geometry optimizations. This functional was shown in extensive benchmark studies[83,84] to provide good accuracy of dissociation energies in both H-bonded and dispersion-bound complexes at only moderate computational cost. In addition, Truhlar's meta-GGA exchange-correlation functional M06-L,[43] which was designed for good performance, amongst other areas, in main-group thermochemistry and noncovalent interactions, was employed for single-point calculations.

Density-fitting (also resolution of identity, RI)[436–438] was employed using the RI-J approximation to speed up self-consistent field (SCF) calculations using the "def2/J" auxiliary basis set. Accompanying errors were shown to be negligible (*i.e.* on the order of 0.1 kcal mol^{-1}).

Zero point energies (ZPE) and enthalpy corrections were determined by analytical frequency calculations at the B97-D3/pc-2-sp(d) level of theory. All harmonic frequencies were used without scaling. For bigger molecules, it was not computationally affordable to avoid a few imaginary frequencies (less than 5) below 30i cm^{-1}, which are an artifact of numerical noise in the computations. They would require significant additional computational effort to resolve. For the calculation of the corresponding ZPE-corrections these additional vibrations were ignored since their influence on the overall ZPE-correction is negligible. Entropies were computed via the Q-RRHO approach.[177] Imaginary harmonic frequencies were assumed to have the same real value in order to estimate their contribution towards the overall entropy. Very low harmonic frequencies contribute the most towards the vibrational entropy. Small imaginary frequencies are indicative of a shallow potential energy surface (PES) so at the true minimum geometry the corresponding real frequencies would also be very low. Therefore, assuming the imaginary frequencies to be real provides better estimates for the entropy rather than neglecting them. Notably, the entropy of mixing for chiral structures was neglected but the entropy of mixing of multiple compounds was accounted for.

Domain-based local pair natural orbital coupled cluster theory with single,

double and perturbative triple excitations (DLPNO-CCSD(T))[30,37] method was used for single-point energy calculations on the basis of DFT geometries using the default "TightPNO" set of parameters ($T_{CutPairs} = 10^{-5}$, $T_{CutPNO} = 10^{-7}$, $T_{CutMKN} = 10^{-4}$).[33] Two-point CBS extrapolation was performed by extrapolating the Hartree-Fock (HF) SCF energies and the Coupled-Cluster (CC) correlation energies separately using the approach adopted by Neese,[441] as implemented in Orca, using the "Extrapolate" keyword, on the basis of the single point energies employing the cc-pVDZ and cc-pVTZ basis sets. Additionally, the CBS-extrapolated counterpoise-correction (CP)[442] was calculated for the bond dissociation energies (BDEs) by performing the single point energy CBS extrapolations of the two fragments both in the monomer and dimer basis. This procedure is equivalent to calculating the CP-corrected single point energies and extrapolating these CP-corrected cc-pVDZ and cc-pVTZ single point energies to the CBS limit.

All calculations to estimate the free energy of solvation using the SMD model[6] were carried out using the Gaussian 09 (revision D.01)[440] software package. Free energies of solvation were computed according to their parametrization within the SMD model via thermodynamic cycles. Therefore, geometries were re-optimized at the M05-2X/6-31g* level of theory and single points energies were calculated at the SMD/M05-2X/6-31g* level of theory in the respective solvent. The difference between the respective absolute electronic energies yields the free energy of solvation at 25 °C .[6]

All calculations to estimate the free energy of solvation using COSMO-RS[5,105] were carried out using ADF2016[447] and the corresponding COSMO-RS module (ADF COSMO-RS 2016)[448]. Free energies of solvation were computed according to their parametrization within the COSMO-RS model in the ADF implementation.[448] Hence, single point calculations and COSMO calculations were performed using BP86/TZP geometries at the BP86/TZP level of theory. The ADF combi2005 parameters were used.

Conformational analysis was performed using the xtb program[166] (version 5.9) with the GFN1-xTB[166] and GFN2-xTB[451] methods by simulated annealing molecular dynamics ("-siman" option). The simulations were performed in an NVT ensemble and the SHAKE[452] algorithm was used to prevent all covalent bonds

from breaking during the simulations. The mass of hydrogen atoms was increased to 4 amu and the time step used was 4.0 fs. Each simulated annealing run encompassed ten simulations of 50.0 ps, each at successively higher temperatures starting from 298.15 K and going linearly to 1000 K. In each of these simulations, the temperature was kept constant. The conformation was saved every 50 fs for subsequent optimization by either GFN1-xTB or GFN2-xTB. The resulting conformers after optimization were compared based on their rotational constant and all distinct conformers within 20.0 kcal mol^{-1} were considered for Boltzmann averaging. Zero-point energies and thermodynamic corrections were determined by subsequent reoptimization using ”tight” convergence criteria followed by frequency calculations to obtain enthalpies, entropies and Gibbs free energies for all the nitrogen bases, protonated nitrogen bases and proton-bound dimers. Boltzmann averaging was carried out at every temperature separately and the corresponding results were used to estimate Gibbs free energies at 173 K–343 K.

B.3. Immiscibility of Alkanes and Perfluoroalkanes

The following programs were employed for all the computations in chapter 4:

- Orca (versions 3.0.3 and 4.0.1)[433,450]
- PSI4 (versions 1.1 and 1.2)[443,453]
- Molpro (version 2015.1.18)[454]
- Gaussian (version 09, revision D.01)[440]
- ADF (version 2016)[447]
- Q-Chem (version 4.3)[455]
- postg[256,257]
- BootD3[52,271]
- xtb (version 5.8)[166]
- Paraview (version 5.5.0-RC3)[456,457]
- PolaBer (version 0.02)[242]
- AIMAll (version 17.11.14)[458]
- multiwfn (version 3.4.1)[394]
- NCIPLOT (version 3.0)[78,459]

- PyMol (version 1.8.4)[449]
- Chemcraft (version 1.8, build 536b)[460]
- Python (version 3.6.0)[461]

B.3.1. Geometry Optimizations for Benchmarking

Geometry optimizations were performed in Orca (version 3.0.3) using the DSD-PBEP86[233]/def2-TZVP-spd[95,462] level of theory employing the RIJCOSX[463,464] approximation in conjunction with the def2-TZVP/J[465] and def2-TZVP/C[466] fitting basis sets. We performed systematic exploration of the conformational space of the intermolecular complexes by starting DFT geometry optimizations from various initial geometries having distinct spatial arrangements of the interacting molecules. The exploration was performed starting from **1-HH**, **1-HF** and **1-FF** by increasing the carbon chains of the interacting molecules by one and finding the most stable intermolecular complexes for **2-HH**, **2-HF** and **2-FF** by systematically varying the position of the newly introduced groups and the spatial arrangement of the two interacting molecules. In addition, in a second cycle, we started from the most stable intermolecular complexes for **6-HH**, **6-HF** and **6-FF** and systematically reduced the carbon chain length by successively removing distinct end groups from the complexes. That way, the conformational space was explored step-by-step by systematically changing the carbon chain length from one to six and from six to one and performing conformational exploration at every step. The most stable conformers, judged on the basis of DSD-PBEP86/def2-QZVP single point energies, were selected for benchmarking. For **HH** and **FF**, the two lowest energy conformers in each class of complexes was selected for further benchmarking. For **HF**, except for **1-HF** for which we selected only two conformers, the four lowest energy conformers in each class of complexes was selected for benchmarking.

B.3.2. Single Points

Wavefunction Theory (WFT) methods and Density Function Theory (DFT) methods were used for single point calculations. Single point calculations were performed using DSD-PBEP86/def2-TZVP-spd geometries. Two-point complete basis

set (CBS) extrapolations were performed using two single point energies calculated using the either the cc-pVTZ[467] and cc-pVQZ[467] or the aug-cc-pVTZ[468] and aug-cc-pVQZ[468] basis sets and are abbreviated herein as CBS(34) and CBS(aug34), respectively. For DLPNO-CCSD(T) calculations, we also used the following non-standard sets of thresholds together with the TightPNO[33] keyword in Orca:

- VeryTightPNO:[270] TCutPNO = 10^{-8}, TCutPairs = 10^{-6}, TCutMKN = 10^{-4}
- ExtremePNO: TCutPNO = 10^{-9}, TCutPairs = 10^{-7}, TCutMKN = 10^{-5}

Semiempirical Methods

GFN-xTB[166] and GFN2-xTB[167] calculations were performed using xtb (version 5.8).

WFT

Method	Basis set	Software	Comments
CCSD(T)-F12[24,25]	cc-pVTZ-F12[469]	Molpro	–
CCSD(T)-F12[24,25]	cc-pVDZ-F12[469]	Molpro	–
CCSD(T)[470,471]	CBS(34)	Molpro	–
CCSD(T)[470,471]	aug-cc-pVDZ[468]	Molpro	CBS(aug34) extrapolation was performed using Orca (version 4.0.1) with RI-MP2.[235,236]
CCSD(T)[470,471]	aug-cc-pVDZ[468]	Molpro	CBS(aug34) extrapolation was performed using Orca (version 4.0.1) with LPNO-CEPA/1.[86]
DLPNO-CCSD(T)[30,37]	CBS(34)	Orca 4.0.1	The NormalPNO,[33] TightPNO,[33] VeryTightPNO and ExtremePNO sets of threshold parameters were used, respectively. The RI-JK[472] approximation for HF calculations was used.
MP2[234]	CBS(34)	Orca 4.0.1	–

Method	Basis set	Software	Comments
RI-MP2[235,236]	CBS(34)	Orca 4.0.1	The RI-JK[472] approximation for HF calculations was used.
RI-SCS-MP2[21,235,236]	CBS(34)	Orca 4.0.1	The RI-JK[472] approximation for HF calculations was used.
RI-MP3[235,236,473,474]	CBS(34)	Orca 4.0.1	The RI-JK[472] approximation for HF calculations was used.

DFT

Method	Basis set	Software	Comments
B97M-V[64]	def2-QZVP[462]	PSI4 1.2	The RI-J[436–438] approximation with the def2/JK[475] fitting basis set was used.
ωB97M-V[65]	def2-QZVP[462]	PSI4 1.2	The RI-JK[472] approximation with the def2/JK[475] fitting basis set was used.
B3LYP[90,476,477]-XDM[47–51]	aug-cc-pVTZ[468]	Gaussian, postg	–
PBE[395]-XDM[47–51]	aug-cc-pVTZ[468]	Gaussian, postg	–
DSD-PBEP86[233]	def2-QZVP[462]	Orca 3.0.3	The RIJCOSX[463,464] approximation with the def2-TZVP/J[465] and def2-TZVP/C[466] fitting basis sets was used.
BLYP[90,476]-XDM[47–51]	aug-cc-pVTZ[468]	Gaussian, postg	–
M06L[42]	def2-QZVP[462]	Orca 4.0.1	The RI-J[436–438] approximation with the def2/J[475] fitting basis set was used.
B3LYP[90,476,477]-NL[62,63]	def2-QZVP[462]	Orca 4.0.1	The RIJCOSX[463,464] approximation with the def2-TZVP/J[465] fitting basis set was used.

Method	Basis set	Software	Comments
PBE[395]-D3(BJ)[52,435]	def2-QZVP[462]	Orca 4.0.1	The RI-J[436–438] approximation with the def2/J[475] fitting basis set was used.
M06-2X[43]-D3(ZERO)[52,435]	def2-QZVP[462]	Orca 4.0.1	The RIJCOSX[463,464] approximation with the def2-TZVP/J[465] fitting basis set was used.
BLYP[90,476]-NL[62,63]	def2-QZVP[462]	Orca 4.0.1	The RI-J[436–438] approximation with the def2/J[475] fitting basis set was used.
B3LYP[90,476,477]-dDsC[53,54,262]	QZ4P[478]	ADF	Density fitting was performed using the RI approximation[478,479] with the QZ4P[478] fitting basis set.
M06L[42]-D3(ZERO)[52,435]	def2-QZVP[462]	Orca 4.0.1	The RI-J[436–438] approximation with the def2/J[475] fitting basis set was used.
B2PLYP[434]-D3[52,435]	def2-QZVP[462]	Orca 4.0.1	The RIJCOSX[463,464] approximation with the def2-TZVP/J[465] and def2-TZVP/C[466] fitting basis sets were used.
TPSS[480]-D3(BJ)[52,435]	def2-QZVP[462]	Orca 4.0.1	The RI-J[436–438] approximation with the def2/J[475] fitting basis set was used.
B3LYP[90,476,477]-D3(BJ)[52,435]	def2-QZVP[462]	Orca 4.0.1	The RIJCOSX[463,464] approximation with the def2-TZVP/J[465] fitting basis set was used.
ωB97X-D3(BJ)[52,435,481]	def2-QZVP[462]	Orca 4.0.1	The RIJCOSX[463,464] approximation with the def2-TZVP/J[465] fitting basis set was used.
rPW86PW92-vdW-DF10[61]	pc-3[92–94]	Q-Chem	–
PBE[395]-dDsC[53,54,262]	QZ4P[478]	ADF	Density fitting was performed using the RI approximation[479] with the QZ4P[478] fitting basis set.

Method	Basis set	Software	Comments
BLYP[90,476]-D3(BJ)[52,435]	def2-QZVP[462]	Orca 4.0.1	The RI-J[436–438] approximation with the def2/J[475] fitting basis set was used.
M06-2X[43]	def2-QZVP[462]	Orca 4.0.1	The RIJCOSX[463,464] approximation with the def2-TZVP/J[465] fitting basis set was used.
B97-D3(BJ)[52,91,435]	def2-QZVP[462]	Orca 4.0.1	The RI-J[436–438] approximation with the def2/J[475] fitting basis set was used.

B.3.3. Energy Decomposition Schemes

Method	Basis set	Software	Comments
sSAPT0[74]	jun-cc-pVDZ[446,482]	PSI4 1.1	The RI-JK[472] approximation for HF calculations was used with the jun-cc-pVDZ/JK[483] fitting basis set.
SAPT2+[74]	aug-cc-pVDZ[468]	PSI4 1.1	The RI-JK[472] approximation for HF calculations was used with the aug-cc-pVDZ/JK[472] fitting basis set.
SAPT2+(3)δMP2[74]	aug-cc-pVTZ[468]	PSI4 1.1	The RI-JK[472] approximation for HF calculations was used with the aug-cc-pVTZ/JK[472] fitting basis set.
SAPT2+(CCD)δMP2[74]	aug-cc-pVTZ[468]	PSI4 1.1	The RI-JK[472] approximation for HF calculations was used with the aug-cc-pVTZ/JK[472] fitting basis set.
LED[75]-DLPNO-CCSD(T)[30,37]	cc-pVQZ[467]	Orca 4.0.1	The RI-JK[472] approximation for HF calculations was used with the cc-pVQZ/JK[472] fitting basis set. TightPNO[33] thresholds were used.

B.3.4. Visual Analyses

Method	Level of Theory	Software	Comments
DID[79]	SCS-LMP2[484]/aug-cc-pVTZ	Molpro	The cc-pVTZ basis set was used on H atoms. The RI-JK[472] approximation for HF calculations was used with the automatically generated JKFIT basis set. Visualization was performed using Paraview.
PolaBer[242]	MP2[485]/def2-TZVPPD[397,462]	PolaBer, Orca 4.0.1	Calculations were performed once in the absence and in the presence of electric fields of 0.005 a.u. in the positive and negative directions of the cartesian coordinate axes, respectively.
NCIPLOT[78,459]	MP2[485]/def2-TZVPPD[397,462]	NCIPLOT, Orca 4.0.1	The intermolecular cut-off used was 0.9 to neglect intramolecular interactions. The density cut-off was 0.1. The reduced density gradient cut-off was 2.0. The density cut-off in the cube files was 0.007, the RDG cut-off in the cube files was 0.5. Visualization was performed using PyMol.

B.3.5. Additional Analyses

Van der Waals surfaces were calculated with an isovalue of 0.001 using MP2[485]/def2-TZVPPD[397,462] densities in multiwfn. COSMO[105] surfaces were calculated using ADF and the corresponding volumes and surfaces were obtained from the corresponding COSMO-RS module.[448] Additional calculations including the simple interaction models, the Q values, the P values and the $n_{contacts}$ were performed using python.

B.4. Universal Descriptor of Dispersive Potential

In this section, we will describe the computational methods used in chapter 5. The following programs were employed for all the computations:

- Orca (version 4.1.0)[433,450]
- molden2aim (version 4.2.1)[486]
- dftd3 (version 3.1 Rev 1)[52]
- dftd4 (version 2.0)[55,56]
- xtb (version 6.1 beta)[166,167]
- crest (version 2.5.1)[487,488]
- Multiwfn (version 3.6)[394]
- PSI4 (version 1.2)[443,453]
- ADF (version 2016)[447]
- Gaussian (version 09, revision D.01)[440]
- postg[256,257]
- IGMPlot (version 2.4.2)[78,459,489,490]
- Molpro (version 2015.1.18)[454]
- PyMol (version 1.8.4)[449]
- Paraview (version 5.6.0)[456,457]
- Python (version 3.6.8)[461]

Orca (version 4.1.0)[433,450] was employed for all density functional theory (DFT) calculations except for some of the benchmark dispersion map calculations. Density-fitting (also called resolution of identity, RI)[436–438] was employed using the RI-J approximation to speed up self-consistent field (SCF) calculations using the ”def2/J” auxiliary basis set. Unless noted otherwise, ”Grid5” was used in conjunction with ”NoFinalGrid”. Geometry optimizations for small molecules were carried out using the PBE[395]/def2-TZVP[462] with the ”TightOPT” keyword. Other geometries were taken directly from the literature, either from crystal structures or from computations. All geometries are provided in a ZIP archive as additional supporting information material.

P parameters of functional groups were computed as follows. The functional groups were attached to an ethynyl group. The lowest energy conformers of the respective molecules were found using crest (version 2.5.1)[487,488] using

the MTD-GC workflow with default settings at the GFN2-xTB[166,167] level of theory. The electron density was obtained either from PBE/def2-SVPD[397] single point calculations (Orca keywords: RI, def2/J, NormalSCF, Grid4, NoFinalGrid) or from Multiwfn (version 3.6)[394] as promolecular electron density. The van der Waals (vdW) surface was obtained using Multiwfn (version 3.6)[394] based on the respective density with isovalues of 0.00100 (for PBE/def2-SVPD densities) or 0.00174 (for promolecular densities, isovalue optimized to reproduce vdW volume of ethene from PBE/def2-SVPD densities). The spacing between points was 0.25 Bohr. The points of the vdW surface corresponding to the ethynyl group were removed and the dispersion coefficients of the ethynyl group were not included in the computation of the P values. For the remaining points, the P parameter in each point was calculated as defined in the main manuscript. For dispersion coefficients, dftd3 (D3), dftd4 (D4) or postg (XDM) were used. The volume and surface area were determined by taking the total volume and surface area of the molecule and removing the volume and surface area of the ethynyl group, which were determined as half the volume and surface area of buta-1,3-diyne, respectively. All volumes and surface areas were determined as ratios with respect to the methyl group.

LDP maps of molecules were computed as follows. The electron density was obtained from PBE/def2-TZVP single points calculations (Orca keywords: RI, def2/J, Grid4, NoFinalGrid). Density cube files were generated using Multiwfn (version 3.6).[394] P parameters were computed in all points of the cube file. For dispersion coefficients, dftd3 (D3), dftd4 (D4) or postg (XDM) were used. LDP maps were visualized using PyMol (version 1.8.4).[449]

Benchmark dispersion maps were computed as follows. Noble gas atoms were placed systematically across the molecular surface of the respective molecules. The positions were determined based on PBE/def2-TZVP densities using Multiwfn (version 3.6)[394] by generating isosurfaces of electron density with isovalues of 0.000005 for He, 0.000001 for Ne and 0.0000005 for Ar, and with a spacing between points of 1.1 Bohr for He and 1.2 Bohr for Ne and Ar. The isovalues were chosen based on SAPT2+(3)δMP2 calculations in order to have overall attractive interaction energies. Typically, this resulted in 10^2–10^3 points on the molecular surface. The dispersion energy was either determined directly by energy decom-

position analysis methods between the molecules and the respective noble gas atom or as energy difference between the dispersion corrections for the molecules with and without noble gas atom. The following table lists the methods used to determine the dispersion between the molecules and the noble gas atoms together with the corresponding basis sets employed, programs used and additional computational details.

Method	Basis set	Software	Comments
sSAPT0[74]	jun-cc-pVDZ[446,482]	PSI4	The RI-JK[472] approximation for HF calculations was used with the jun-cc-pVDZ/JK[483] fitting basis set.
SAPT2+[74]	aug-cc-pVDZ[468]	PSI4	The RI-JK[472] approximation for HF calculations was used with the aug-cc-pVDZ/JK[472] fitting basis set.
SAPT2+(3)δMP2[74]	aug-cc-pVTZ[468]	PSI4	The RI-JK[472] approximation for HF calculations was used with the aug-cc-pVTZ/JK[472] fitting basis set.
LED[75]-DLPNO-CCSD(T)[30,37]	cc-pVQZ[467]	Orca	The RI-JK[472] approximation for HF calculations was used with the cc-pVQZ/JK[472] fitting basis set. TightPNO[33] thresholds were used.
B3LYP[90,476,477]-D3(BJ,abc)[52]	–	dftd3	–
B3LYP[90,476,477]-D4(BJ,abc)[55,56]	–	dftd4	–
B3LYP[90,476,477]-NL[62,63]	def2-QZVP[462]	Orca	The RIJCOSX[463,464] approximation with the def2/J[465] fitting basis set was used.
B3LYP[90,476,477]-dDsC[53,54,262]	QZ4P[478]	ADF	Density fitting was performed using the RI approximation[478,479] with the QZ4P[478] fitting basis set.

Method	Basis set	Software	Comments
B3LYP[90,476,477]-XDM[47–51]	aug-cc-pVTZ[468]	Gaussian, postg	–

For non-covalent interaction (NCI) plots, reduced density gradient (RDG) isosurfaces (isovalue of 0.75) of intermolecular complexes were generated using IGMPlot[78,459,489,490] and visualized using PyMOL (version 1.8.4).[449] Geometries were taken from the literature. Densities were obtained from promolecular estimates. All interactions with at least a fraction of 0.90 of the density from only one of the two fragments were turned off so that only intermolecular interactions were visualized.

Dispersion interaction density (DID)[79] plots were generated using Molpro (version 2015.1.18)[454] by SCS-LMP2[21,484]/aug'-cc-pVTZ[467,468] calculations making use of the RI-JK[472] approximation for the HF step with the automatically generated JKFIT basis set. Notably, the aug'-cc-pVTZ uses diffuse functions only for non-hydrogen atoms. Localization was carried out using Pipek-Mezey orbitals and visualization was conducted in Paraview.[456,457] Geometries were taken from the literature.

Multiple linear regressions were carried out using Python (version 3.6.8)[461] using the scikit-learn package (also called sklearn, version 0.20.2).[491] Target values and parameters, except for P_{int} and A_{rel}, were taken from the respective literature reports. Both the parameters and target values were normalized before data analysis. Linear regression was carried out using `sklearn.linear_model.LinearRegression`. Leave-one-out cross validation was performed using `sklearn.model_selection.LeaveOneOut`.

Atomic P parameters from hydrides and fluorides of the main group elements were determined as follows. The number of hydrogen and fluorine atoms bound to the central element were chosen to satisfy the octet rule. Geometries were optimized at the TPSS[480]/def2-QZVP level of theory (Orca keywords: RI, def2/J, TightOpt, Grid5, NoFinalGrid). Subsequent single point calculations were carried out at the DKH2[492,493]-TPSS/ANO-RCC-QZP[494] level of theory (Orca keywords: RI, AutoAux, NoFrozenCore, TightSCF) and were used to obtain electron

densities. The vdW surface was obtained using Multiwfn (version 3.6)[394] with isovalues of 0.00100 and a spacing between points of 0.25 Bohr. Consecutively, the P parameter effected solely by the central element was calculated for all the points on the vdW surface. The maximum P value on the vdW surface was used as atomic P parameter.

C. Mathematical Derivations

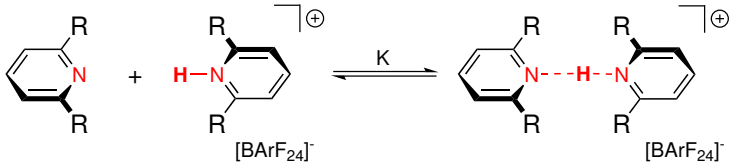
This section provides mathematical derivations for chapters 2 and 4.

C.1. Compensation of Dispersion in Dichloromethane

The standard state assumed in the gas-phase is an ideal gas at a concentration of 1 mol l^{-1} for each species at 298.15 K. The standard state assumed in solution is an ideal solution at a concentration of 1 mol l^{-1} for each species at 298.15 K. In the following calculations all equilibrium constants are unitless. The corresponding necessary divisions of each concentration by the respective standard concentration (1 mol l^{-1}) are omitted for clarity.

C.1.1. Master Equation of Proton-bound Dimer Equilibria

The following equilibrium is investigated in solution.



Therefore, the equilibrium can be formulated as follows.

$$K = \frac{[N_2H]}{[NH]\,[N]} \tag{C.1}$$

Now, the following definitions are made. It should be noted here that the following equations are only valid if the excess e of the basic species is bigger than or equal to 1.

Variable	Definition
c	Total concentration of the mobile proton [mol l^{-1}]
x	Relative mole fraction
e	Relative excess of the basic species with respect to the mobile proton ($e \geq 1$)

Now several basic equations of the equilibrium investigated are formulated.

$$x(NH) = 1 - x(N_2H) \tag{C.2}$$

$$[NH] = c\, x(NH) \tag{C.3}$$

$$[Py_2H] = c\, x(N_2H) \tag{C.4}$$

$$[N] = c\ [e - x(N_2H)] \tag{C.5}$$

Using equations C.2-C.5 and inserting them into equation C.1 results in the following expression.

$$K = \frac{x(N_2H)}{c\ [1 - x(N_2H)]\ [e - x(N_2H)]} \tag{C.6}$$

The following basic equation describes the NMR-shifts observed and their connection to the equilibrium under investigation.

$$\delta_H = x(N_2H)\, \delta(N_2H) + x(NH)\, \delta(NH) \tag{C.7}$$

Inserting the equations above in equation C.7 the following equations are obtained.

$$\delta_H = x(N_2H)\, \delta(N_2H) + [1 - x(N_2H)]\ \delta(NH) \tag{C.8}$$

$$\delta_H = \delta(NH) + [\delta(N_2H) - \delta(NH)]\ x(N_2H) \tag{C.9}$$

In order to obtain an expression for the dependence of the chemical shift of the mobile proton as a function of the equilibrium constant the quadratic equation C.6 in $x(N_2H)$ is solved (cf. equations C.10 – C.12).

$$1 - x(N_2H)\ [e + 1] + x(N_2H)^2 = \frac{x(N_2H)}{K\, c} \tag{C.10}$$

$$x(N_2H)^2 - \left[1 + e + \frac{1}{K\,c}\right] x(N_2H) + 1 = 0 \tag{C.11}$$

$$x(N_2H) = \frac{1+e}{2} + \frac{1}{2\,K\,c} - \sqrt{\left[\frac{1+e}{2} + \frac{1}{2\,K\,c}\right]^2 - 1} \tag{C.12}$$

To connect this equation with the temperature-dependence of the equilibrium constant several basic thermodynamic equations are used (cf. equations C.13-C.15).

$$\Delta_R G = -R\,T\,ln\,K = \Delta_R H - T\,\Delta_R S \tag{C.13}$$

$$ln\,K = -\frac{\Delta_R H}{R\,T} + \frac{\Delta_R S}{R} \tag{C.14}$$

$$K = e^{-\frac{\Delta_R H}{R\,T}}\, e^{\frac{\Delta_R S}{R}} \tag{C.15}$$

Equation C.15 is now substituted into equation C.12 to obtain the following equation.

$$x(N_2H) = \frac{1+e}{2} + \frac{e^{\frac{\Delta_R H}{R\,T}}\, e^{\frac{-\Delta_R S}{R}}}{2\,c} - \sqrt{\left[\frac{1+e}{2} + \frac{e^{\frac{\Delta_R H}{R\,T}}\, e^{\frac{-\Delta_R S}{R}}}{2\,c}\right]^2 - 1} \tag{C.16}$$

This equation is now substituted into equation C.9 to obtain a master equation describing the temperature dependence of the equilibrium under investigation.

$$\delta_H = \delta(NH) + [\delta(N_2H) - \delta(NH)] \left[\frac{1+e}{2} + \frac{e^{\frac{\Delta_R H}{R\,T}}\, e^{\frac{-\Delta_R S}{R}}}{2\,c} - \sqrt{\left[\frac{1+e}{2} + \frac{e^{\frac{\Delta_R H}{R\,T}}\, e^{\frac{-\Delta_R S}{R}}}{2\,c}\right]^2 - 1}\right] \tag{C.17}$$

In this equation there are 3 unknown parameters for the system under investigation which have to be determined by means of nonlinear fitting from the experimental data: $\Delta_R H$, $\Delta_R S$ and $\delta(N_2H)$. However, this equation turned out not to yield very useful results in general because the nonlinear fitting did not converge. Therefore, dependent and independent variable were switched for nonlinear regression. In order to do that, on the basis of equation C.14 the following

rearrangement was carried out.

$$ln\,K - \frac{\Delta_R S}{R} = -\frac{\Delta_R H}{R\,T} \tag{C.18}$$

$$\frac{\Delta_R S - R\,ln\,K}{\Delta_R H} = \frac{1}{T} \tag{C.19}$$

$$T = \frac{\Delta_R H}{\Delta_R S - R\,ln\,K} \tag{C.20}$$

In order to obtain an expression for $ln\,K$ first an expression for $x(N_2H)$ is obtained from equation C.9 by simple rearrangements.

$$x(N_2H) = \frac{\delta_H - \delta(NH)}{\delta(N_2H) - \delta(NH)} \tag{C.21}$$

Equation C.21 is substituted into equation C.6 to obtain an expression for K.

$$K = \frac{\frac{\delta_H - \delta(NH)}{\delta(N_2H) - \delta(NH)}}{c\,\left[1 - \frac{\delta_H - \delta(NH)}{\delta(N_2H) - \delta(NH)})\right]\,\left[e - \frac{\delta_H - \delta(NH)}{\delta(N_2H) - \delta(NH)}\right]} \tag{C.22}$$

$$K = \frac{\frac{\delta_H - \delta(NH)}{\delta(N_2H) - \delta_H}}{c\,\left[e - \frac{\delta_H - \delta(NH)}{\delta(N_2H) - \delta(NH)}\right]} \tag{C.23}$$

Equation C.23 is substituted into equation C.20 to obtain another master equation describing the temperature dependence of the equilibrium under investigation.

$$T = \frac{\Delta_R H}{\Delta_R S - R\,ln\,\left\{\frac{\frac{\delta_H - \delta(NH)}{\delta(N_2H) - \delta_H}}{c\,\left[e - \frac{\delta_H - \delta(NH)}{\delta(N_2H) - \delta(NH)}\right]}\right\}} \tag{C.24}$$

In this equation there are again 3 unknown parameters for the system under investigation which have to be determined by means of nonlinear fitting from the experimental data: $\Delta_R H$, $\Delta_R S$ and $\delta(N_2H)$. This equation turned out to yield reliable results in general. It should also be noted that for systems where $\delta(N_2H)$ was known from experiment equation C.24 was still used but with only 2 unknown parameters.

C.2. Immiscibility of Alkanes and Perfluoroalkanes

In this section mathematical derivations are presented and explained in detail. The corresponding results are applied to analyze and interpret the computational results.

C.2.1. Homomolecular Cuboid Model for Interaction Surfaces

The basic assumption is that the interacting molecules (linear alkanes and perfluoroalkanes) can be approximated as interacting cuboids (cf. Figure C.1).

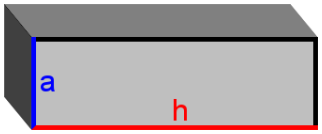


Figure C.1.: Cuboid model.

The total surface (S) of the cuboid can be calculated as follows:

$$S = 2a^2 + 4ah \tag{C.25}$$

The interaction surface corresponds to one of the long faces of the cuboid with the area:

$$S_{int} = ah \tag{C.26}$$

The length of the long face (h) in this model is approximately a linear function of the length of the carbon chain and hence of the number of carbon atoms (n) with u and v being the corresponding slope and intercept, respectively:

$$h \approx u \cdot n + v \tag{C.27}$$

It follows from the combination of equations C.25 and C.27 that the total surface is predicted to be a linear function (with intercept q) of the number of carbon

atoms for each class of compounds (i.e. either **H** or **F**):

$$S \approx 2a^2 + 4av + 4an = 4au \cdot n + q \tag{C.28}$$

Combining equations C.26 and C.27 it can be seen that also the interaction surface is a linear function (with intercept r) of the number of carbon atoms with the slope being a quarter of the corresponding slope of the total surface.

$$S_{int} \approx au \cdot n + av = au \cdot n + r \tag{C.29}$$

Therefore, on the basis of this simple model it can be shown that comparing the slope of the total surface of the isolated molecules to the slope of the interaction surface is a means to estimate whether the interaction surface is ideal or not. Additionally, when comparing the ratio of the slopes of the total surface of **H** and **F** against the ratio of the slopes of the interaction surface area of **H** and **F** provides a means to investigate whether **H** or **F** have a closer to ideal interaction surface.

C.2.2. Heteromolecular Cuboid Model for Interaction Surfaces

The model discussed in the previous section is only valid when the two interacting molecules are identical, i.e. the interacting cuboids have the same size. When they do not, the interaction surface is not equal to the long face of either of the two molecules. This is illustrated in Figure C.2.

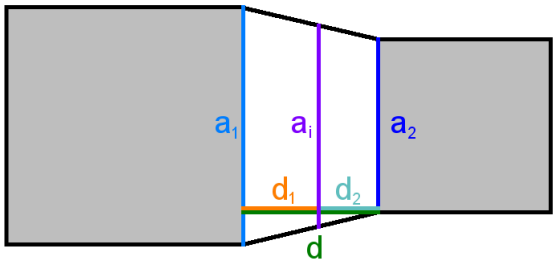


Figure C.2.: Cuboid model for heteromolecular interaction. View from the side on the volume of intercuboidal interaction.

The edge of the interaction surface a_i can determined if we assume the following relationship:

$$\frac{a_1}{a_2} = \frac{d_1}{d_2} \tag{C.30}$$

Using the intercept theorem of elementary geometry we can derive the following relationship:

$$\frac{d_1}{a_1 - a_i} = \frac{d_2}{a_i - a_2} \tag{C.31}$$

Minor rearrangement leads to the following equation:

$$\frac{d_1}{d_2} = \frac{a_1 - a_i}{a_i - a_2} \tag{C.32}$$

Combining equations C.30 and C.32 leads to:

$$\frac{a_1}{a_2} = \frac{a_1 - a_i}{a_i - a_2} \tag{C.33}$$

Solving equation C.33 for a_i yields:

$$a_1 a_i - a_1 a_2 = a_1 a_2 - a_2 a_i \tag{C.34}$$

$$a_i(a_1 + a_2) = 2a_1 a_2 \tag{C.35}$$

$$a_i = \frac{2a_1 a_2}{a_1 + a_2} \tag{C.36}$$

Using equation C.36, a_i can be estimated from both a_1 and a_2. Assuming that a_i corresponds approximately to the slope of the interaction surface in heteromolecular complexes of **H** and **F** with respect n, it can be estimated how close to ideal the interaction geometries in the heteromolecular complexes are.

Bibliography

[1] London, F. Zur Theorie und Systematik der Molekularkräfte. *Zeitschrift für Physik* **1930**, *63*, 245–279.

[2] Eisenschitz, R.; London, F. Über das Verhältnis der van der Waalsschen Kräfte zu den homöopolaren Bindungskräften. *Zeitschrift für Physik* **1930**, *60*, 491–527.

[3] Grimme, S.; Hansen, A.; Brandenburg, J. G.; Bannwarth, C. Dispersion-Corrected Mean-Field Electronic Structure Methods. *Chemical Reviews* **2016**, *116*, 5105–5154.

[4] Hermann, J.; DiStasio, R. A.; Tkatchenko, A. First-Principles Models for van der Waals Interactions in Molecules and Materials: Concepts, Theory, and Applications. *Chemical Reviews* **2017**, *117*, 4714–4758.

[5] Eckert, F.; Klamt, A. Fast Solvent Screening via Quantum Chemistry: COSMO-RS Approach. *AIChE Journal* **2002**, *48*, 369–385.

[6] Marenich, A. V.; Cramer, C. J.; Truhlar, D. G. Universal solvation model based on solute electron density and on a continuum model of the solvent defined by the bulk dielectric constant and atomic surface tensions. *Journal of Physical Chemistry B* **2009**, *113*, 6378–6396.

[7] Schreiner, P. R.; Chernish, L. V.; Gunchenko, P. a.; Tikhonchuk, E. Y.; Hausmann, H.; Serafin, M.; Schlecht, S.; Dahl, J. E. P.; Carlson, R. M. K.; Fokin, A. a. Overcoming lability of extremely long alkane carbon-carbon bonds through dispersion forces. *Nature* **2011**, *477*, 308–311.

[8] Grimme, S.; Huenerbein, R.; Ehrlich, S. On the importance of the dispersion energy for the thermodynamic stability of molecules. *ChemPhysChem* **2011**, *12*, 1258–1261.

[9] Hunter, C. A. van der Waals interactions in non-polar liquids. *Chem. Sci.* **2013**, *4*, 834–848.

[10] Tkatchenko, A. Current Understanding of Van der Waals Effects in Realistic Materials. *Advanced Functional Materials* **2015**, *25*, 2054–2061.

[11] van der Waals, J. D. *Nobel Lectures, Physics 1901-1921*; Elsevier Publishing Company: Amsterdam, 1967; pp 254–265.

[12] Nakamura, K.; Houk, K. N. Theoretical Studies of the Wilcox Molecular Torsion Balance. Is the Edge-to-Face Aromatic Interaction Important? *Organic Letters* **1999**, *1*, 2049–2051.

[13] Hunter, C. A. Quantifying intermolecular interactions: Guidelines for the molecular recognition toolbox. *Angewandte Chemie - International Edition* **2004**, *43*, 5310–5324.

[14] Yang, L.; Adam, C.; Nichol, G. S.; Cockroft, S. L. How much do van der Waals dispersion forces contribute to molecular recognition in solution? *Nature Chemistry* **2013**, *5*, 1006–1010.

[15] Hwang, J.; Dial, B. E.; Li, P.; Kozik, M. E.; Smith, M. D.; Shimizu, K. D. How important are dispersion interactions to the strength of aromatic stacking interactions in solution? *Chem. Sci.* **2015**, *6*, 4358–4364.

[16] London, F. The general theory of molecular forces. *Trans. Faraday Soc.* **1937**, *33*, 8–26.

[17] Koide, A. A new expansion for dispersion forces and its application. *Journal of Physics B: Atomic and Molecular Physics* **1976**, *9*, 3173–3183.

[18] Wagner, J. P.; Schreiner, P. R. London Dispersion in Molecular Chemistry - Reconsidering Steric Effects. *Angew. Chem. Int. Ed.* **2015**, *54*, 12274–12296.

[19] Sinnokrot, M. O.; Sherrill, C. D. High-Accuracy Quantum Mechanical Studies of π–π Interactions in Benzene Dimers. *The Journal of Physical Chemistry A* **2006**, *110*, 10656–10668, PMID: 16970354.

[20] Sherrill, C. D.; Takatani, T.; Hohenstein, E. G. An Assessment of Theoretical Methods for Nonbonded Interactions: Comparison to Complete Basis Set Limit Coupled-Cluster Potential Energy Curves for the Benzene Dimer, the Methane Dimer, Benzene–Methane, and Benzene–H2S. *The Journal of Physical Chemistry A* **2009**, *113*, 10146–10159, PMID: 19689152.

[21] Grimme, S. Improved second-order Møller-Plesset perturbation theory by separate scaling of parallel- and antiparallel-spin pair correlation energies. *Journal of Chemical Physics* **2003**, *118*, 9095–9102.

[22] Antony, J.; Grimme, S. Is Spin-Component Scaled Second-Order Møller–Plesset Perturbation Theory an Appropriate Method for the Study of Noncovalent Interactions in Molecules? *The Journal of Physical Chemistry A* **2007**, *111*, 4862–4868, PMID: 17506533.

[23] Brauer, B.; Kesharwani, M. K.; Kozuch, S.; Martin, J. M. L. The S66x8 benchmark for noncovalent interactions revisited: explicitly correlated ab initio methods and density functional theory. *Phys. Chem. Chem. Phys.* **2016**, *18*, 20905–20925.

[24] Adler, T. B.; Knizia, G.; Werner, H. J. A simple and efficient CCSD(T)-F12 approximation. *Journal of Chemical Physics* **2007**, *127*, 221106.

[25] Knizia, G.; Adler, T. B.; Werner, H. J. Simplified CCSD(T)-F12 methods: Theory and benchmarks. *Journal of Chemical Physics* **2009**, *130*, 054104.

[26] Burns, L. A.; Marshall, M. S.; Sherrill, C. D. Appointing silver and bronze standards for noncovalent interactions: A comparison of spin-component-scaled (SCS), explicitly correlated (F12), and specialized wavefunction approaches. *Journal of Chemical Physics* **2014**, *141*, 234111.

[27] Friedrich, J. Efficient Calculation of Accurate Reaction Energies - Assessment of Different Models in Electronic Structure Theory. *Journal of Chemical Theory and Computation* **2015**, *11*, 3596–3609.

[28] Sirianni, D. A.; Burns, L. A.; Sherrill, C. D. Comparison of explicitly correlated methods for computing high-accuracy benchmark energies for noncovalent interactions. *Journal of Chemical Theory and Computation* **2017**, *13*, 86–99.

[29] Saebo, S. Local Treatment of Electron Correlation. *Annual Review of Physical Chemistry* **2002**, *44*, 213–236.

[30] Riplinger, C.; Neese, F. An efficient and near linear scaling pair natural orbital based local coupled cluster method. *Journal of Chemical Physics* **2013**, *138*, 034106.

[31] Liakos, D. G.; Neese, F. Domain based pair natural orbital coupled cluster studies on linear and folded alkane chains. *Journal of Chemical Theory and Computation* **2015**, *11*, 2137–2143.

[32] Minenkov, Y.; Chermak, E.; Cavallo, L. Accuracy of DLPNO-CCSD (T) method for non- covalent bond dissociation enthalpies from coinage metal cation complexes. *Journal of Chemical Theory and Computation* **2015**, *11*, 1–65.

[33] Liakos, D. G.; Sparta, M.; Kesharwani, M. K.; Martin, J. M. L.; Neese, F. Exploring the accuracy limits of local pair natural orbital coupled-cluster theory. *Journal of Chemical Theory and Computation* **2015**, *11*, 1525–1539.

[34] Liakos, D. G.; Neese, F. Is It Possible to Obtain Coupled Cluster Quality Energies at near Density Functional Theory Cost? Domain-Based Local Pair Natural Orbital Coupled Cluster vs Modern Density Functional Theory. *Journal of Chemical Theory and Computation* **2015**, *11*, 4054–4063.

[35] Minenkov, Y.; Wang, H.; Wang, Z.; Sarathy, S. M.; Cavallo, L. Heats of Formation of Medium-Sized Organic Compounds from Contemporary Electronic Structure Methods. *Journal of Chemical Theory and Computation* **2017**, *13*, 3537–3560.

[36] Paulechka, E.; Kazakov, A. Efficient DLPNO-CCSD(T)-Based Estimation of Formation Enthalpies for C-, H-, O-, and N-Containing Closed-Shell Compounds Validated Against Critically Evaluated Experimental Data. *Journal of Physical Chemistry A* **2017**, *121*, 4379–4387.

[37] Riplinger, C.; Sandhoefer, B.; Hansen, A.; Neese, F. Natural triple excitations in local coupled cluster calculations with pair natural orbitals. *The Journal of Chemical Physics* **2013**, *139*, 134101.

[38] Guo, Y.; Riplinger, C.; Becker, U.; Liakos, D. G.; Minenkov, Y.; Cavallo, L.; Neese, F. Communication: An improved linear scaling perturbative triples correction for the domain based local pair-natural orbital based singles and doubles coupled cluster method [DLPNO-CCSD(T)]. *Journal of Chemical Physics* **2018**, *148*, 011101.

[39] Grimme, S. Density functional theory with London dispersion corrections. *Wiley Interdisciplinary Reviews: Computational Molecular Science* **2011**, *1*, 211–228.

[40] Corminboeuf, C. Minimizing Density Functional Failures for Non-Covalent Interactions Beyond van der Waals Complexes. *Accounts of Chemical Research* **2014**, *47*, 3217–3224, PMID: 24655016.

[41] Klimes, J.; Michaelides, A. Perspective: Advances and challenges in treating van der Waals dispersion forces in density functional theory. *Journal of Chemical Physics* **2012**, *137*.

[42] Zhao, Y.; Truhlar, D. G. A new local density functional for main-group thermochemistry, transition metal bonding, thermochemical kinetics, and noncovalent interactions. *Journal of Chemical Physics* **2006**, *125*, 194101.

[43] Zhao, Y.; Truhlar, D. G. The M06 suite of density functionals for main group thermochemistry, thermochemical kinetics, noncovalent interactions, excited states, and transition elements: Two new functionals and systematic testing of four M06-class functionals and 12 other function. *Theoretical Chemistry Accounts* **2008**, *120*, 215–241.

[44] Risthaus, T.; Grimme, S. Benchmarking of London dispersion-accounting density functional theory methods on very large molecular complexes. *Journal of Chemical Theory and Computation* **2013**, *9*, 1580–1591.

[45] Grimme, S. Accurate description of van der Waals complexes by density functional theory including empirical corrections. *Journal of Computational Chemistry* **2004**, *25*, 1463–1473.

[46] Tkatchenko, A.; Scheffler, M. Accurate molecular van der Waals interactions from ground-state electron density and free-atom reference data. *Physical Review Letters* **2009**, *102*, 6–9.

[47] Johnson, E. R.; Becke, A. D. A post-Hartree-Fock model of intermolecular interactions. *Journal of Chemical Physics* **2005**, *123*, 024101.

[48] Becke, A. D.; Johnson, E. R. A density-functional model of the dispersion interaction. *Journal of Chemical Physics* **2005**, *123*, 154101.

[49] Becke, A. D.; Johnson, E. R. Exchange-hole dipole moment and the dispersion interaction. *Journal of Chemical Physics* **2005**, *122*, 154104.

[50] Johnson, E. R.; Becke, A. D. A post-Hartree-Fock model of intermolecular interactions: Inclusion of higher-order corrections. *Journal of Chemical Physics* **2006**, *124*, 174104.

[51] Becke, A. D.; Johnson, E. R. Exchange-hole dipole moment and the dispersion interaction revisited. *Journal of Chemical Physics* **2007**, *127*, 154104.

[52] Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu. *Journal of Chemical Physics* **2010**, *132*, 154104.

[53] Steinmann, S. N.; Corminboeuf, C. A system-dependent density-based dispersion correction. *Journal of Chemical Theory and Computation* **2010**, *6*, 1990–2001.

[54] Steinmann, S. N.; Corminboeuf, C. A generalized-gradient approximation exchange hole model for dispersion coefficients. *Journal of Chemical Physics* **2011**, *134*, 044117.

[55] Caldeweyher, E.; Bannwarth, C.; Grimme, S. Extension of the D3 dispersion coefficient model. *Journal of Chemical Physics* **2017**, *147*, 034112.

[56] Caldeweyher, E.; Ehlert, S.; Hansen, A.; Neugebauer, H.; Spicher, S.; Bannwarth, C.; Grimme, S. A generally applicable atomic-charge dependent London dispersion correction. *The Journal of Chemical Physics* **2019**, *150*, 154122.

[57] Bursch, M.; Caldeweyher, E.; Hansen, A.; Neugebauer, H.; Ehlert, S.; Grimme, S. Understanding and Quantifying London Dispersion Effects in Organometallic Complexes. *Accounts of Chemical Research* **2018**, *52*, acs.accounts.8b00505.

[58] Axilrod, B. M.; Teller, E. Interaction of the van der Waals Type Between Three Atoms. *The Journal of Chemical Physics* **1943**, *11*, 299–300.

[59] Muto, Y. Force between nonpolar molecules. *Proc. Phys. Soc. Jpn.* **1943**, *17*, 629.

[60] Tkatchenko, A.; DiStasio, R. A.; Car, R.; Scheffler, M. Accurate and Efficient Method for Many-Body van der Waals Interactions. *Phys. Rev. Lett.* **2012**, *108*, 236402.

[61] Lee, K.; Murray, É. D.; Kong, L.; Lundqvist, B. I.; Langreth, D. C. Higher-accuracy van der Waals density functional. *Physical Review B - Condensed Matter and Materials Physics* **2010**, *82*, 3–6.

[62] Vydrov, O. A.; Van Voorhis, T. Nonlocal van der Waals density functional: The simpler the better. *Journal of Chemical Physics* **2010**, *133*, 244103.

[63] Hujo, W.; Grimme, S. Performance of the van der waals density functional VV10 and (hybrid)GGA variants for thermochemistry and noncovalent interactions. *Journal of Chemical Theory and Computation* **2011**, *7*, 3866–3871.

[64] Mardirossian, N.; Head-Gordon, M. Mapping the genome of meta-generalized gradient approximation density functionals: The search for B97M-V. *Journal of Chemical Physics* **2015**, *142*, 074111.

[65] Mardirossian, N.; Head-Gordon, M. ω B97M-V: A combinatorially optimized, range-separated hybrid, meta-GGA density functional with VV10 nonlocal correlation. *Journal of Chemical Physics* **2016**, *144*, 214110.

[66] Acree, W. E., Jr.; Chickos, J. S. In *NIST Chemistry WebBook, NIST Standard Reference Database Number 69*; Linstrom, P. J., Mallard, W. G., Eds.; National Institute of Standards and Technology: Gaithersburg MD, 20899, (retrieved September 17, 2018); Chapter Boiling Point Data.

[67] Acree, W. E., Jr.; Chickos, J. S. In *NIST Chemistry WebBook, NIST Standard Reference Database Number 69*; Linstrom, P. J., Mallard, W. G., Eds.; National Institute of Standards and Technology: Gaithersburg MD, 20899, (retrieved September 17, 2018); Chapter Phase Transition Enthalpy Measurements of Organic and Organometallic Compounds.

[68] McBride, J. The hexaphenylethane riddle. *Tetrahedron* **1974**, *30*, 2009 – 2022.

[69] Kahr, B.; Van Engen, D.; Mislow, K. Length of the ethane bond in hexaphenylethane and its derivatives. *Journal of the American Chemical Society* **1986**, *108*, 8305–8307.

[70] Maier, G.; Pfriem, S.; Schäfer, U.; Matusch, R. Tetra-tert-butyltetrahedrane. *Angewandte Chemie International Edition in English* **1978**, *17*, 520–521.

[71] Autumn, K.; Sitti, M.; Liang, Y. A.; Peattie, A. M.; Hansen, W. R.; Sponberg, S.; Kenny, T. W.; Fearing, R.; Israelachvili, J. N.; Full, R. J. Evidence for van der Waals adhesion in gecko setae. *Proceedings of the National Academy of Sciences* **2002**, *99*, 12252–12256.

[72] Adam, C.; Yang, L.; Cockroft, S. L. Partitioning solvophobic and dispersion forces in alkyl and perfluoroalkyl cohesion. *Angewandte Chemie - International Edition* **2015**, *54*, 1164–1167.

[73] Yang, L.; Brazier, J. B.; Hubbard, T. A.; Rogers, D. M.; Cockroft, S. L. Can dispersion forces govern aromatic stacking in an organic solvent? *Angewandte Chemie - International Edition* **2016**, *55*, 912–916.

[74] Parker, T. M.; Burns, L. A.; Parrish, R. M.; Ryno, A. G.; Sherrill, C. D. Levels of symmetry adapted perturbation theory (SAPT). I. Efficiency and performance for interaction energies. *The Journal of Chemical Physics* **2014**, *140*, 094106.

[75] Schneider, W. B.; Bistoni, G.; Sparta, M.; Saitow, M.; Riplinger, C.; Auer, A. A.; Neese, F. Decomposition of Intermolecular Interaction Energies within the Local Pair Natural Orbital Coupled Cluster Framework. *Journal of Chemical Theory and Computation* **2016**, *12*, 4778–4792.

[76] Cramer, C. J. *Essentials of Computational Chemistry: Theories and Models*, 2nd ed.; Wiley, 2017.

[77] Johnson, E. R.; Keinan, S.; Mori-Sanchez, P.; Contreras-Garcia, J.; Cohen, A. J.; Yang, W. Revealing Noncovalent Interactions. *Journal of the American Chemical Society* **2010**, *132*, 6498–6506.

[78] Contreras-Garcia, J.; Johnson, E. R.; Keinan, S.; Chaudret, R.; Piquemal, J. P.; Beratan, D. N.; Yang, W. NCIPLOT: A program for plotting noncovalent interaction regions. *Journal of Chemical Theory and Computation* **2011**, *7*, 625–632.

[79] Wuttke, A.; Mata, R. A. Visualizing dispersion interactions through the use of local orbital spaces. *Journal of Computational Chemistry* **2017**, *38*, 15–23.

[80] Spackman, M. A.; Jayatilaka, D. Hirshfeld surface analysis. *CrystEngComm* **2009**, *11*, 19–32.

[81] Becke, A. D. A multicenter numerical integration scheme for polyatomic molecules. *The Journal of Chemical Physics* **1988**, *88*, 2547–2553.

[82] Klamt, A.; Schüürmann, G. COSMO: a new approach to dielectric screening in solvents with explicit expressions for the screening energy and its gradient. *J. Chem. Soc., Perkin Trans. 2* **1993**, 799–805.

[83] Burns, L. A.; Mayagoitia, Á. V.; Sumpter, B. G.; Sherrill, C. D. Density-functional approaches to noncovalent interactions: A comparison of dispersion corrections (DFT-D), exchange-hole dipole moment (XDM) theory, and specialized functionals. *The Journal of Chemical Physics* **2011**, *134*, 084107.

[84] Goerigk, L.; Grimme, S. A thorough benchmark of density functional methods for general main group thermochemistry, kinetics, and noncovalent interactions. *Phys. Chem. Chem. Phys.* **2011**, *13*, 6670–6688.

[85] Lee, T. J.; Taylor, P. R. A diagnostic for determining the quality of single-reference electron correlation methods. *International Journal of Quantum Chemistry* **1989**, *36*, 199–207.

[86] Neese, F.; Wennmohs, F.; Hansen, A. Efficient and accurate local approximations to coupled-electron pair approaches: An attempt to revive the pair natural orbital method. *Journal of Chemical Physics* **2009**, *130*, 114108.

[87] Gurinov, A. A.; Lesnichin, S. B.; Limbach, H. H.; Shenderovich, I. G. How short is the strongest hydrogen bond in the proton-bound homodimers of pyridine derivatives? *Journal of Physical Chemistry A* **2014**, *118*, 10804–10812.

[88] Gurinov, A. A.; Denisov, G. S.; Borissova, A. O.; Goloveshkin, A. S.; Greindl, J.; Limbach, H.-H.; Shenderovich, I. G. NMR Study of Solvation Effect on the Geometry of Proton-Bound Homodimers of Increasing Size. *The Journal of Physical Chemistry A* **2017**, 8697–8705.

[89] Perdew, J. P. Density-functional approximation for the correlation energy of the inhomogeneous electron gas. *Physical Review B* **1986**, *33*, 8822–8824.

[90] Becke, A. D. Density-functional exchange-energy approximation with correct asymptotic behavior. *Physical Review A* **1988**, *38*, 3098–3100.

[91] Grimme, S. Semiempirical GGA-type density functional constructed with a long-range dispersion correction. *Journal of Computational Chemistry* **2006**, *27*, 1787–1799.

[92] Jensen, F. Polarization consistent basis sets: Principles. *Journal of Chemical Physics* **2001**, *115*, 9113–9125.

[93] Jensen, F. Polarization consistent basis sets: II. Estimating the Kohn-Sham basis set limit. *Journal of Chemical Physics* **2002**, *116*, 7372–7379.

[94] Jensen, F. Polarization consistent basis sets. III. The importance of diffuse functions. *Journal of Chemical Physics* **2002**, *117*, 9234–9240.

[95] Johnson, E. R.; Otero-De-La-Roza, A.; Dale, S. G.; Dilabio, G. A. Efficient basis sets for non-covalent interactions in XDM-corrected density-functional theory. *Journal of Chemical Physics* **2013**, *139*, 214109.

[96] Bondi, A. van der Waals Volumes and Radii. *The Journal of Physical Chemistry* **1964**, *68*, 441–451.

[97] Alvarez, S. A cartography of the van der Waals territories. *Dalton Transactions* **2013**, *42*, 8617.

[98] Grimme, S.; Steinmetz, M. Effects of London dispersion correction in density functional theory on the structures of organic molecules in the gas phase. *Physical Chemistry Chemical Physics* **2013**, *15*, 16031–16042.

[99] Meot-Ner, M.; Sieck, L. W. The ionic hydrogen bond. 1. Sterically hindered bonds. Solvation and clustering of protonated amines and pyridines. *Journal of the American Chemical Society* **1983**, *105*, 2956–2961.

[100] Steacie, E. W.; Shaw, G. T. The homogeneous unimolecular decomposition of gaseous alkyl nitrites. II. The decomposition of ethyl nitrite. *The Journal of Chemical Physics* **1934**, *2*, 345–348.

[101] Szwarc, M. The determination of bond dissociation energies by pyrolytic methods. *Chemical Reviews* **1950**, *47*, 75–173.

[102] Gray, P. Bond Dissociation Energies in Nitrities and Nitro Compounds and the Reaction of Free Radicals with Nitrogen Dioxide. *Trans. Faraday Soc.* **1955**, *51*, 1367–1374.

[103] Gray, P.; Rathbone, P.; Williams, A. Benzyloxyl radicals in benzyl nitrite pyrolyses. *Journal of the Chemical Society* **1960**, *0*, 3932–3943.

[104] Otero-de-la Roza, A.; Johnson, E. R.; Contreras-García, J. Revealing non-covalent interactions in solids: NCI plots revisited. *Physical Chemistry Chemical Physics* **2012**, *14*, 12165.

[105] Klamt, A. Conductor-like Screening Model for Real Solvents: A New Approach to the Quantitative Calculation of Solvation Phenomena. *Journal of Physical Chemistry* **1995**, *99*, 2224–2235.

[106] Tomasi, J.; Mennucci, B.; Cammi, R. Quantum mechanical continuum solvation models. *Chemical Reviews* **2005**, *105*, 2999–3093.

[107] Aldrich, K. E.; Billow, B. S.; Holmes, D.; Bemowski, R. D.; Odom, A. L. Weakly coordinating yet ion paired: Anion effects on an internal rearrangement. *Organometallics* **2017**, *36*, 1227–1237.

[108] Smid, J. Structure of Ion Pair Solvation Complexes. *Angewandte Chemie International Edition* **1972**, *11*, 112–127.

[109] Liu, L.; Guo, Q.-X. Isokinetic Relationship, Isoequilibrium Relationship, and Enthalpy-Entropy Compensation. *Chemical Reviews* **2001**, *101*, 673–696.

[110] Perez-Benito, J. F.; Mulero-Raichs, M. Enthalpy-Entropy Compensation Effect in Chemical Kinetics and Experimental Errors: A Numerical Simulation Approach. *The Journal of Physical Chemistry A* **2016**, *120*, 7598–7609.

[111] Krug, R. R.; Hunter, W. G.; Grieger, R. A. Enthalpy-Entropy Compensation. 2. Separation of the Chemical from the Statisitcal Effect. *Journal of Physical Chemistry* **1976**, *80*, 2341–2351.

[112] Krug, R. R.; Hunter, W. G.; Grieger, R. A. Statistical Interpretation of Enthalpy-Entropy Compensation. *Nature* **1976**, *261*, 566–567.

[113] Krug, R. R.; Hunter, W. G.; Grieger, R. A. Enthalpy-Entropy Compensation. 1. Some Fundamental Statistical Problems Associated with the Analysis. *Journal of Physical Chemistry* **1976**, *80*, 2335–2341.

[114] Phipps, M. J.; Fox, T.; Tautermann, C. S.; Skylaris, C.-K. Energy decomposition analysis approaches and their evaluation on prototypical protein–drug interaction patterns. *Chem. Soc. Rev.* **2015**, *44*, 3177–3211.

[115] Jeziorski, B.; Jeziorski, B.; Moszynski, R.; Moszynski, R.; Szalewicz, K.; Szalewicz, K. Perturbation Theory Approach to Intermolecular Potential Energy Surfaces of van der Waals Complexes. *Chem. Rev.* **1994**, *94*, 1887–1930.

[116] Szalewicz, K. Symmetry-adapted perturbation theory of intermolecular forces. *Wiley Interdisciplinary Reviews: Computational Molecular Science* **2012**, *2*, 254–272.

[117] Chupka, W. a. Effect of Unimolecular Decay Kinetics on the Interpretation of Appearance Potentials. *The Journal of Chemical Physics* **1959**, *30*, 191.

[118] Rodgers, M.; Armentrout, P. Noncovalent metal-ligand bond energies as studied by threshold collision-induced dissociation. *Mass spectrometry reviews* **2000**, *19*, 215–47.

[119] Narancic, S.; Bach, A.; Chen, P. Simple fitting of energy-resolved reactive cross sections in threshold collision-induced dissociation (T-CID) experiments. *Journal of Physical Chemistry A* **2007**, *111*, 7006–7013.

[120] Baer, T. *Advances in Chemical Physics, Volume 64*; John Wiley & Sons, Inc., 1986; p 111.

[121] Armentrout, P. B.; Baer, T. Gas-Phase Ion Dynamics and Chemistry. *The Journal of Physical Chemistry* **1996**, *100*, 12866–12877.

[122] Chase, M. W. Jr.; Davies, C. A.; Downey, J. R. Jr.; Frurip, D. J.; McDonald, R. A.; Syverud, A. N. JANAF Thermochemical Tables. *J. Phys. Chem. Ref. Data* **1985**, *17*, Suppl. 1.

[123] Lias, S.G.; Bartmess, J.E.; Liebman, J.F.; Holmes, J.L.; Levin, R.D.; Mallard, W. Gas-Phase Ion and Neutral Thermochemistry. *J. Phys. Chem. Ref. Data* **1988**, *17*, Suppl. 1.

[124] Ruscic, B.; Pinzon, R. E.; Morton, M. L.; Von Laszevski, G.; Bittner, S. J.; Nijsure, S. G.; Amin, K. A.; Minkoff, M.; Wagner, A. F. Introduction to active thermochemical tables: Several "Key" enthalpies of formation revisited. *Journal of Physical Chemistry A* **2004**, *108*, 9979–9997.

[125] Stevens, W. R.; Rusic, B.; Baer, T. Heats of Formation of C6H5•,C6H5+, and C6H5NO by Threshold Photoelectron Photoion Coincidence and Active Thermochemical Tables Analysis. *J. Phys. Chem. A* **2010**, *114*, 13134–13145.

[126] Minenkov, Y.; Sliznev, V. V.; Cavallo, L. Accurate Gas Phase Formation Enthalpies of Alloys and Refractories Decomposition Products. *Inorganic Chemistry* **2017**, *56*, 1386–1401.

[127] Calbo, J.; Ortí, E.; Sancho-García, J. C.; Aragó, J. Accurate treatment of large supramolecular complexes by double-hybrid density functionals coupled with nonlocal van der waals corrections. *Journal of Chemical Theory and Computation* **2015**, *11*, 932–939.

[128] Mati, I. K.; Adam, C.; Cockroft, S. L. Seeing through solvent effects using molecular balances. *Chemical Science* **2013**, *4*, 3965–3972.

[129] Yang, L.; Adam, C.; Cockroft, S. L. Quantifying Solvophobic Effects in Nonpolar Cohesive Interactions. *Journal of the American Chemical Society* **2015**, *137*, 10084–10087.

[130] Qian, H.; Hopfied, J. J.; Hopfield, J. J. Entropy-enthalpy compensation: Perturbation and relaxation in thermodynamic systems. *The Journal of Chemical Physics* **1996**, *105*, 9292.

[131] Pan, A.; Biswas, T.; Rakshit, A. K.; Moulik, S. P. Enthalpy-Entropy Compensation (EEC) Effect: A Revisit. *Journal of Physical Chemistry B* **2015**, *119*, 15876–15884.

[132] Pan, A.; Kar, T.; Rakshit, A. K.; Moulik, S. P. Enthalpy–Entropy Compensation (EEC) Effect: Decisive Role of Free Energy. *The Journal of Physical Chemistry B* **2016**, *120*, 10531–10539.

[133] Grunwald, E.; Steel, C. Solvent Reorganization and Thermodynamic Enthalpy-Entropy Compensation. *J. Am. Chem. Soc.* **1995**, *117*, 5687–5692.

[134] Chodera, J. D.; Mobley, D. L. Entropy-enthalpy compensation: role and ramifications in biomolecular ligand recognition and design. *Annual review of biophysics* **2013**, *42*, 121–42.

[135] Starikov, E. B.; Norden, B. Enthalpy-entropy compensation: A phantom or something useful? *Journal of Physical Chemistry B* **2007**, *111*, 14431–14435.

[136] Liu, Y.; Sturtevant, J. M. Significant discrepancies between van't Hoff and calorimetric enthalpies. III. *Biophysical Chemistry* **1997**, *64*, 121–126.

[137] U, D. M.; Bloomfield, V. A. Thermodynamics of the hydrophobic effect . I . Coupling of aggregation and p K a shifts in solutions of aliphatic amines. *Biophysical Chemistry* **2001**, 37–51.

[138] De Lisi, R.; Milioto, S.; Muratore, N. Thermodynamics of surfactants, block copolymers and their mixtures in water: The role of the isothermal calorimetry. *International Journal of Molecular Sciences* **2009**, *10*, 2873–2895.

[139] Klippenstein, S. J.; Pande, V. S.; Truhlar, D. G. Chemical Kinetics and Mechanisms of Complex Systems: A Perspective on Recent Theoretical Advances A Perspective on Recent Theoretical Advances. *J. Am. Chem. Soc.* **2013**, *136*, 528–546.

[140] Cheng, G.-J.; Zhang, X.; Chung, L. W.; Xu, L.; Wu, Y.-D. Computational Organic Chemistry: Bridging Theory and Experiment in Establishing the Mechanisms of Chemical Reactions. *Journal of the American Chemical Society* **2015**, *137*, 1706–1725.

[141] Skyner, R. E.; McDonagh, J. L.; Groom, C. R.; Van, M. T.; Mitchell, J. B. O. A review of methods for the calculation of solution free energies and the modelling of systems in solution. *Phys Chem Chem Phys* **2015**, *17*, 6174–6191.

[142] Decherchi, S.; Masetti, M.; Vyalov, I.; Rocchia, W. Implicit solvent methods for free energy estimation. *European journal of medicinal chemistry* **2015**, *91*, 27–42.

[143] Jensen, J. H. Predicting accurate absolute binding energies in aqueous solution: thermodynamic considerations for electronic structure methods. *Physical chemistry chemical physics : PCCP* **2015**, *17*, 12441–51.

[144] Vidossich, P.; Lledós, A.; Ujaque, G. First-Principles Molecular Dynamics Studies of Organometallic Complexes and Homogeneous Catalytic Processes. *Accounts of Chemical Research* **2016**, *49*, 1271–1278.

[145] Acevedo, O. Simulating chemical reactions in ionic liquids using QM/MM methodology. *Journal of Physical Chemistry A* **2014**, *118*, 11653–11666.

[146] Ratkova, E. L.; Palmer, D. S.; Fedorov, M. V. Solvation Thermodynamics of Organic Molecules by the Molecular Integral Equation Theory: Approaching Chemical Accuracy. *Chemical Reviews* **2015**, *115*, 6312–6356.

[147] Cramer, C. J.; Truhlar, D. G. A universal approach to solvation modeling. *Accounts of chemical research* **2008**, *41*, 760–768.

[148] Klamt, A.; Mennucci, B.; Tomasi, J.; Barone, V.; Curutchet, C.; Orozco, M.; Luque, F. J. On the Performance of Continuum Solvation Methods. A Comment on "Universal Approaches to Solvation Modeling". *Accounts of Chemical Research* **2009**, *42*, 489–492.

[149] Klamt, A.; Eckert, F.; Arlt, W. COSMO-RS - An Alternative to Simulation for Calculating Thermodynamic Properties of Liquid Mixtures. *Annual Review of Chemical and Biomolecular Engineering* **2010**, *1*, 101–122.

[150] Marenich, A. V.; Cramer, C. J.; Truhlar, D. G. Generalized born solvation model SM12. *Journal of Chemical Theory and Computation* **2013**, *9*, 609–620.

[151] Ho, J.; Coote, M. L. A universal approach for continuum solvent pKa calculations: Are we there yet? *Theoretical Chemistry Accounts* **2009**, *125*, 3–21.

[152] Plata, R. E.; Singleton, D. A. A Case Study of the Mechanism of Alcohol-Mediated Morita Baylis-Hillman Reactions. The Importance of Experimental Observations. *J. Am. Chem. Soc.* **2015**, *137*, 3811–3826.

[153] Basdogan, Y.; Keith, J. A. A paramedic treatment for modeling explicitly solvated chemical reaction mechanisms. *Chemical Science* **2018**, *9*, 5341–5346.

[154] Mugridge, J. S.; Bergman, R. G.; Raymond, K. N. Equilibrium Isotope Effects on Noncovalent Interactions in a Supramolecular Host–Guest System. *Journal of the American Chemical Society* **2012**, *134*, 2057–2066.

[155] Pitera, J. W.; Van Gunsteren, W. F. The importance of solute-solvent van der Waals interactions with interior atoms of biopolymers. *Journal of the American Chemical Society* **2001**, *123*, 3163–3164.

[156] Garza, A. J. Solvation Entropy Made Simple. *Journal of Chemical Theory and Computation* **2019**, 10.1021/acs.jctc.9b00214.

[157] Hwang, J.; Li, P.; Smith, M. D.; Shimizu, K. D. Distance-Dependent Attractive and Repulsive Interactions of Bulky Alkyl Groups. *Angewandte Chemie - International Edition* **2016**, *55*, 8086–8089.

[158] Rösel, S.; Balestrieri, C.; Schreiner, P. R. Sizing the role of London dispersion in the dissociation of all-meta tert-butyl hexaphenylethane. *Chemical Science* **2016**, *8*, 405–410.

[159] Rösel, S.; Becker, J.; Allen, W. D.; Schreiner, P. R. Probing the Delicate Balance Between Pauli Repulsion and London Dispersion with Triphenylmethyl Derivatives. *Journal of the American Chemical Society* **2018**, *140*, 14421 – 14432.

[160] Grimme, S.; Schreiner, P. R. Steric crowding can stabilize a labile molecule: Solving the hexaphenylethane riddle. *Angewandte Chemie - International Edition* **2011**, *50*, 12639–12642.

[161] Marquez, C.; Nau, W. M. Polarizabilities inside molecular containers. *Angewandte Chemie - International Edition* **2001**, *40*, 4387–4390.

[162] Van Den Broeke, J.; Deelman, B. J.; Van Koten, G. Tetrakis{3,5-bis(perfluorohexyl)phenyl}borate: A highly fluorous anion. *Tetrahedron Letters* **2001**, *42*, 8085–8087.

[163] Boswell, P. G.; Bühlmann, P. Fluorous bulk membranes for potentiometric sensors with wide selectivity ranges: Observation of exceptionally strong ion pair formation. *Journal of the American Chemical Society* **2005**, *127*, 8958–8959.

[164] Ghosh, S. K.; Ojeda, A. S.; Guerrero-Leal, J.; Bhuvanesh, N.; Gladysz, J. A. New media for classical coordination chemistry: Phase transfer of Werner and related

polycations into highly nonpolar fluorous solvents. *Inorganic Chemistry* **2013**, *52*, 9369–9378.

[165] Attah, I. K.; Platt, S. P.; Meot-Ner, M.; El-Shall, M. S.; Aziz, S. G.; Alyoubi, A. O. Proton-bound dimers of nitrogen heterocyclic molecules: Substituent effects on the structures and binding energies of homodimers of diazine, triazine, and fluoropyridine. *Journal of Chemical Physics* **2014**, *140*.

[166] Grimme, S.; Bannwarth, C.; Shushkov, P. A Robust and Accurate Tight-Binding Quantum Chemical Method for Structures, Vibrational Frequencies, and Noncovalent Interactions of Large Molecular Systems Parametrized for All spd-Block Elements (Z = 1-86). *Journal of Chemical Theory and Computation* **2017**, *13*, 1989–2009.

[167] Bannwarth, C.; Ehlert, S.; Grimme, S. GFN2-xTB – an accurate and broadly quantum chemical method with multipole electrostatics and density-dependent dispersion contributions. *J. Chem. Theory Comput.* **2019**, *15*, 1652–1671.

[168] Product Directory Home. https://www.sigmaaldrich.com/technical-service-home/product-catalog.html (accessed 20th December 2018).

[169] Williams, J. W.; Krchma, I. J. The dielectric constants of binary mixtures. *Journal of the American Chemical Society* **1926**, *48*, 1888–1896.

[170] Gill, J. B. Solute-solute interaction in liquid ammonia solutions: A vibrational spectroscopic view. *Pure and Applied Chemistry* **1981**, *53*, 1365–1381.

[171] Marcus, Y.; Hefter, G. Ion pairing. *Chemical Reviews* **2006**, *106*, 4585–4621.

[172] Eigen, M. Über die Kinetik sehr schnell verlaufender Ionenreaktionen in wässeriger Lösung. *Zeitschrift fur Physikalische Chemie* **1954**, *1*, 176–200.

[173] Fuoss, R. M. Ionic Association. III. The Equilibrium between Ion Pairs and Free Ions. *Journal of the American Chemical Society* **1958**, *80*, 5059–5061.

[174] Otto, S. The role of solvent cohesion in nonpolar solvation. *Chemical Science* **2013**, *4*, 2953–2959.

[175] Hansen, C. M. The Three Dimensional Solubility Parameter and Solvent Diffusion Coefficient. Their Importance in Surface Coating Formulation. *J. Paint Technology* **1967**, 104.

[176] Reilly, A. M.; Tkatchenko, A. van der Waals dispersion interactions in molecular materials: beyond pairwise additivity. *Chem. Sci.* **2015**, *6*, 3289–3301.

[177] Grimme, S. Supramolecular binding thermodynamics by dispersion-corrected density functional theory. *Chem. Eur. J.* **2012**, *18*, 9955–9964.

[178] Bursch, M.; Hansen, A.; Grimme, S. Fast and Reasonable Geometry Optimization of Lanthanoid Complexes with an Extended Tight Binding Quantum Chemical Method. *Inorganic Chemistry* **2017**, *56*, 12485–12491.

[179] Chan, B.; Del Bene, J. E.; Radom, L. What factors determine whether a proton-bound homodimer has a symmetric or an asymmetric hydrogen bond? *Molecular Physics* **2009**, *107*, 1095–1105.

[180] MacChioni, A.; Ciancaleoni, G.; Zuccaccia, C.; Zuccaccia, D. Determining accurate molecular sizes in solution through NMR diffusion spectroscopy. *Chemical Society Reviews* **2008**, *37*, 479–489.

[181] Pregosin, P. S. NMR spectroscopy and ion pairing: Measuring and understanding how ions interact. *Pure and Applied Chemistry* **2009**, *81*, 615–633.

[182] Krossing, I.; Raabe, I. Noncoordinating anions - Fact or fiction? A survey of likely candidates. *Angewandte Chemie - International Edition* **2004**, *43*, 2066–2090.

[183] Lee, B. Solvent reorganization contribution to the transfer thermodynamics of small nonpolar molecules. *Biopolymers* **1991**, *31*, 993–1008.

[184] Petruska, J.; Goodman, M. F.; Boosalis, M. S.; Sowers, L. C.; Cheong, C.; Tinoco, I. Comparison between DNA melting thermodynamics and DNA polymerase fidelity. *Proceedings of the National Academy of Sciences of the United States of America* **1988**, *85*, 6252–6.

[185] Gil, P.; Moghadam, T. T.; Ranjbar, B. Differential Scanning Calorimetry Techniques: Applications in Biology and Nanoscience. **2010**, 167–193.

[186] Houk, K. N.; Leach, A. G.; Kim, S. P.; Zhang, X. Binding Affinities of Host-Guest, Protein-Ligand, and Protein-Transition-State Complexes. *Angewandte Chemie - International Edition* **2003**, *42*, 4872–4897.

[187] Sugihara, G.; Shigematsu, D. S.; Nagadome, S.; Lee, S.; Sasaki, Y.; Igimi, H. Thermodynamic study on the Langmuir adsorption of various bile salts including

taurine and glycine conjugates onto graphite in water. *Langmuir* **2000**, *16*, 1825–1833.

[188] Mills, E. A.; Plotkin, S. S. Protein Transfer Free Energy Obeys Entropy-Enthalpy Compensation. *Journal of Physical Chemistry B* **2015**, *119*, 14130–14144.

[189] Petruska, J.; Goodman, M. F. Enthalpy-entropy compensation in DNA melting thermodynamics. 1995.

[190] Wilfong, E. M.; Kogiso, Y.; Muthukrishnan, S.; Kowatz, T.; Du, Y.; Bowie, A.; Naismith, J. H.; Hadad, C. M.; Toone, E. J.; Gustafson, T. L. A Multidisciplinary Approach to Probing Enthalpy - Entropy Compensation and the Interfacial Mobility Model. *Journal of the American Chemical Society* **2011**, *133*, 11515–11523.

[191] Chen, L.-J.; Lin, S.-Y.; Huang, C.-C. Effect of Hydrophobic Chain Length of Surfactants on Enthalpy–Entropy Compensation of Micellization. *The Journal of Physical Chemistry B* **1998**, *102*, 4350–4356.

[192] Ford, D. M. Enthalpy-entropy compensation is not a general feature of weak association. *Journal of the American Chemical Society* **2005**, *127*, 16167–16170.

[193] Sharp, K. Entropy – enthalpy compensation : Fact or artifact ? *Protein Science* **2001**, *10*, 661–667.

[194] Ryde, U. A fundamental view of enthalpy–entropy compensation. *MedChemComm* **2014**, *5*, 1324–1336.

[195] Exner, O. How to get wrong results from good experimental data: a survey of incorrect applications of regression. *Journal of Physical Organic Chemistry* **1997**, *10*, 797–813.

[196] Eaton, D. F.; Smart, B. E. Are Fluorocarbon Chains “Stiffer” Than Hydrocarbon Chains? Dynamics of End-to-End Cyclization in a C8F16 Segment Monitored by Fluorescence. *Journal of the American Chemical Society* **1990**, *112*, 2821–2823.

[197] Morgado, P.; Garcia, A. R.; Martins, L. F. G.; Ilharco, L. M.; Filipe, E. J. M. Alkane Coiling in Perfluoroalkane Solutions: A New Primitive Solvophobic Effect. *Langmuir* **2017**, *33*, 11429–11435.

[198] Grant, R. F.; Davidson, D. W.; Gray, P. Dielectric Study of Some Liquid Alkyl Nitrites. *The Journal of Chemical Physics* **1960**, *33*, 1713–1718.

[199] Neumann, M. Dipole moment fluctuation formulas in computer simulations of polar systems. *Molecular Physics* **1983**, *50*, 841–858.

[200] Shenderovich, I. G.; Burtsev, A. P.; Denisov, G. S.; Golubev, N. S.; Limbach, H. H. Influence of the temperature-dependent dielectric constant on the H/D isotope effects on the NMR chemical shifts and the hydrogen bond geometry of the collidine-HF complex in CDF3/CDClF2 solution. *Magn Reson Chem* **2001**, *39*, S91–S99.

[201] Riniker, S.; Horta, B. A.; Thijssen, B.; Gupta, S.; Van Gunsteren, W. F.; Hünenberger, P. H. Temperature dependence of the dielectric permittivity of acetic acid, propionic acid and their methyl esters: A molecular dynamics simulation study. *ChemPhysChem* **2012**, *13*, 1182–1190.

[202] Assaf, K. I.; Nau, W. M. Cucurbiturils: from synthesis to high-affinity binding and catalysis. *Chem. Soc. Rev.* **2015**, *44*, 394–418.

[203] Paliwal, S.; Geib, S.; Wilcox, C. S. Molecular Torsion Balance for Weak Molecular Recognition Forces. Effects of "Tilted-T" Edge-to-Face Aromatic Interactions on Conformational Selection and Solid-State Structure. *Journal of the American Chemical Society* **1994**, *116*, 4497–4498.

[204] Pascoe, D. J.; Ling, K. B.; Cockroft, S. L. The Origin of Chalcogen-Bonding Interactions. *Journal of the American Chemical Society* **2017**, *139*, 15160–15167.

[205] Besora, M.; Vidossich, P.; Lledós, A.; Ujaque, G.; Maseras, F. Calculation of Reaction Free Energies in Solution: A Comparison of Current Approaches. *Journal of Physical Chemistry A* **2018**, *122*, 1392–1399.

[206] Ardura, D.; López, R.; Sordo, T. L. Relative Gibbs energies in solution through continuum models: Effect of the loss of translational degrees of freedom in bimolecular reactions on gibbs energy barriers. *Journal of Physical Chemistry B* **2005**, *109*, 23618–23623.

[207] Hermans, J.; Wang, L. Inclusion of loss of translational and rotational freedom in theoretical estimates of free energies of binding. Application to a complex of benzene and mutant T4 lysozyme. *Journal of the American Chemical Society* **1997**, *119*, 2707–2714.

[208] Wertz, D. H. Relationship between the Gas-Phase Entropies of Molecules and Their Entropies of Solvation in Water and 1-Octanol. *Journal of the American Chemical Society* **1980**, *102*, 5316–5322.

[209] Finkelstein, A. V.; Janin, J. The price of lost freedom: Entropy of bimolecular complex formation. *Protein Engineering, Design and Selection* **1989**, *3*, 1–3.

[210] Janin, J. For Guldberg and Waage, with love and cratic entropy. *Proteins: Structure, Function and Genetics* **1996**, *24*, 0–1.

[211] Searle, M. S.; Williams, D. H. The Cost of Conformational Order: Entropy Changes in Molecular Associations. *Journal of the American Chemical Society* **1992**, *114*, 10690–10697.

[212] Jencks, W. P. On the attribution and additivity of binding energies. *Proceedings of the National Academy of Sciences* **1981**, *78*, 4046–4050.

[213] Haselmeier, R.; Holz, M.; Marbach, W.; Weingärtner, H. Water dynamics near a dissolved noble gas. First direct experimental evidence for a retardation effect. *Journal of Physical Chemistry* **1995**, *99*, 2243–2246.

[214] Silverstein, T. P. The Real Reason Why Oil and Water Don't Mix. *Journal of Chemical Education* **1998**, *75*, 116.

[215] Cramer, C. J.; Truhlar, D. G. Implicit Solvation Models: Equilibria, Structure, Spectra, and Dynamics. *Chemical Reviews* **1999**, *99*, 2161–2200.

[216] Zhang, J.; Zhang, H.; Wu, T.; Wang, Q.; van der Spoel, D. Comparison of Implicit and Explicit Solvent Models for the Calculation of Solvation Free Energy in Organic Solvents. *Journal of Chemical Theory and Computation* **2017**, *13*, 1034–1043.

[217] O'Hagan, D. Understanding organofluorine chemistry. An introduction to the C–F bond. *Chem. Soc. Rev.* **2008**, *37*, 308–319.

[218] Horvath, I. T. *Topics in Current Chemistry*; Springer, 2012; Vol. 308.

[219] Dunitz, J. D. Organic fluorine: Odd man out. *ChemBioChem* **2004**, *5*, 614–621.

[220] Smart, B. E. Fluorine substituent effects (on bioactivity). *Journal of Fluorine Chemistry* **2001**, *109*, 3–11.

[221] Lemal, D. M. Perspective on Fluorocarbon Chemistry. *J. Org. Chem.* **2004**, *69*, 1–11.

[222] Schlosser, M.; Michel, D. About the "physiological size" of fluorine substituents: comparison of sensorially active compounds with fluorine and methyl substituted analogs. *Tetrahedron* **1996**, *52*, 99–108, CAN 124:171569.

[223] Cametti, M.; Crousse, B.; Metrangolo, P.; Milani, R.; Resnati, G. The fluorous effect in biomolecular applications. *Chem. Soc. Rev.* **2012**, *41*, 31–42.

[224] Johansson, G.; Percec, V.; Ungar, G.; Smith, K. Fluorophobic Effect Generates a Systematic Approach to the Synthesis of the Simplest Class of Rodlike Liquid Crystals Containing a Single Benzene Unit. *Chemistry of Materials* **1997**, *9*, 164–175.

[225] Johansson, G.; Percec, V.; Ungar, G.; Zhou, J. P. Fluorophobic effect in the self-assembly of polymers and model compounds containing tapered groups into supramolecular columns. *Macromolecules* **1996**, *29*, 646–660.

[226] Percec, V.; Schlueter, D.; Kwon, Y. K.; Blackwell, J.; Möller, M.; Slangen, P. J. Dramatic Stabilization of a Hexagonal Columnar Mesophase Generated from Supramolecular and Macromolecular Columns by the Semifluorination of the Alkyl Groups of Their Tapered Building Blocks. *Macromolecules* **1995**, *28*, 8807–8818.

[227] Grosse, A. V.; Cady, G. H. Properties of Fluorocarbons. *Industrial and Engineering Chemistry* **1947**, *39*, 367–374.

[228] Krafft, M. P.; Riess, J. G. Chemistry, physical chemistry, and uses of molecular fluorocarbon- hydrocarbon diblocks, triblocks, and related compounds-unique ”apolar” components for self-assembled colloid and interface engineering. *Chemical Reviews* **2009**, *109*, 1714–1792.

[229] Chattoraj, J.; Risthaus, T.; Rubner, O.; Heuer, A.; Grimme, S. A multi-scale approach to characterize pure CH4, CF4, and CH4/CF4 mixtures. *The Journal of chemical physics* **2015**, *142*, 164508.

[230] Biller, M. J.; Mecozzi, S. A high level computational study of the CH4 /CF4 dimer: how does it compare with the CH4 /CH4 and CF4 /CF4 dimers? *Molecular Physics* **2012**, *110*, 377–387.

[231] Tsuzuki, S.; Uchimaru, T.; Mikami, M.; Urata, S. Magnitude and orientation dependence of intermolecular interaction between perfluoroalkanes: High level ab initio calculations of CF4 and C2F6 dimers. *The Journal of Chemical Physics* **2002**, *116*, 3309–3315.

[232] Mahlanen, R.; Jalkanen, J.-P.; Pakkanen, T. A. Potential energy surfaces of CF4, CCl4 and CBr4 dimers. *Chemical Physics* **2005**, *313*, 271 – 277.

[233] Kozuch, S.; Martin, J. M. L. Spin-component-scaled double hybrids: An extensive search for the best fifth-rung functionals blending DFT and perturbation theory. *Journal of Computational Chemistry* **2013**, *34*, 2327–2344.

[234] Moller, C.; Plesset, M. S. Note on an Approximation Treatment for Many-Electron Systems. *Phys. Rev.* **1934**, *46*, 618–622.

[235] Feyereisen, M.; Fitzgerald, G.; Komornicki, A. Use of approximate integrals in ab initio theory. An application in MP2 energy calculations. *Chemical Physics Letters* **1993**, *208*, 359–363.

[236] Vahtras, O.; Almlöf, J.; Feyereisen, M. W. Integral approximations for LCAO-SCF calculations. *Chemical Physics Letters* **1993**, *213*, 514–518.

[237] Moszynski, R.; Jeziorski, B.; Szalewicz, K. Many-Body Theory of Exchange Effects in Intermolecular Interactions - 2Nd-Quantization Approach and Comparison With Full Configuration-Interaction Results. *Journal of Chemical Physics* **1994**, *100*, 1312–1325.

[238] Rybak, S.; Jeziorski, B.; Szalewicz, K. Many-body symmetry-adapted perturbation theory of intermolecular interactions. H ₂ O and HF dimers. *The Journal of Chemical Physics* **1991**, *95*, 6576–6601.

[239] Bistoni, G.; Auer, A. A.; Neese, F. Understanding the Role of Dispersion in Frustrated Lewis Pairs and Classical Lewis Adducts: A Domain-Based Local Pair Natural Orbital Coupled Cluster Study. *Chemistry - A European Journal* **2017**, *23*, 865–873.

[240] Hättig, C.; Jansen, G.; Hess, B. A.; Ángyán, J. G. Topologically partitioned dynamic polarizabilities using the theory of atoms in molecules. *Canadian Journal of Chemistry* **1996**, *74*, 976–987.

[241] Keith, T. A. AIMAll (Version 17.11.14).

[242] Krawczuk, A.; Pérez, D.; MacChi, P. PolaBer: A program to calculate and visualize distributed atomic polarizabilities based on electron density partitioning. *Journal of Applied Crystallography* **2014**, *47*, 1452–1458.

[243] Politzer, P.; Jin, P.; Murray, J. S. Atomic polarizability, volume and ionization energy. *Journal of Chemical Physics* **2002**, *117*, 8197–8202.

[244] Schulman, J. M.; Musher, J. I. Hydrogen-Atom Polarizability as a Hartree–Fock Perturbation Expansion: A Geometric Approximation to Atomic Polarizabilities. *The Journal of Chemical Physics* **1968**, *49*, 4845.

[245] Rumble, J. R., Ed. *CRC Handbook of Chemistry and Physics*, 98th ed.; CRC Press/Taylor & Francis: Boca Raton, FL, 2017.

[246] Flory, P. J. Thermodynamics of High Polymer Solutions. *J. Chem. Phys.* **1942**, *10*, 51–61.

[247] Huggins, M. L. Thermodynamic Properties of Solutions of Long-chain Compounds. *Ann. N.Y. Acad. Sci.* **1942**, *43*, 1–32.

[248] Dantzler, E. M.; Knobler, C. M.; Windsor, M. L. Interaction Virial Coefficients in Hydrocarbon Mixtures. *The Journal of Physical Chemistry* **1968**, *72*, 684.

[249] Dantzler, E. M.; Knobler, C. M. Interaction Virial Coefficients in Fluorocarbon Mixtures. *The Journal of Physical Chemistry* **1969**, *73*, 1335–1341.

[250] Dantzler Siebert, E. M.; Knobler, C. M. Interaction virial coefficients in hydrocarbon-fluorocarbon mixtures. *The Journal of Physical Chemistry* **1971**, *75*, 3863–3870.

[251] Hohenstein, E. G.; Sherrill, C. D. Wavefunction methods for noncovalent interactions. *Wiley Interdisciplinary Reviews: Computational Molecular Science* **2012**, *2*, 304–326.

[252] Mardirossian, N.; Head-Gordon, M. Thirty years of density functional theory in computational chemistry: an overview and extensive assessment of 200 density functionals. *Molecular Physics* **2017**, *8976*, 1–58.

[253] Johnson, E. R.; Mackie, I. D.; DiLabio, G. A. Dispersion interactions in density-functional theory. *Journal of Physical Organic Chemistry* **2009**, *22*, 1127–1135.

[254] Goerigk, L.; Kruse, H.; Grimme, S. Benchmarking density functional methods against the S66 and S66x8 datasets for non-covalent interactions. *ChemPhysChem* **2011**, *12*, 3421–3433.

[255] Mardirossian, N.; Head-Gordon, M. How Accurate Are the Minnesota Density Functionals for Noncovalent Interactions, Isomerization Energies, Thermochemistry, and Barrier Heights Involving Molecules Composed of Main-Group Elements? *Journal of Chemical Theory and Computation* **2016**, *12*, 4303–4325.

[256] Otero-de-la Roza, A.; Johnson, E. R. Non-covalent interactions and thermochemistry using XDM-corrected hybrid and range-separated hybrid density functionals Non-covalent interactions and thermochemistry using XDM-corrected hybrid and range-separated hybrid density functionals. *The Journal of Chemical Physics* **2013**, *138*, 204109.

[257] Kannemann, F. O.; Becke, A. D. van der Waals Interactions in Density-Functional Theory: Intermolecular Complexes. *Journal of Chemical Theory and Computation* **2010**, *6*, 1081–1088.

[258] Calbo, J.; Ortí, E.; Sancho-García, J. C.; Aragó, J. The nonlocal correlation density functional VV10: A successful attempt to accurately capture noncovalent interactions. *Annual Reports in Computational Chemistry* **2015**, *11*, 37–102.

[259] Otero-De-La-Roza, A.; Johnson, E. R. Predicting Energetics of Supramolecular Systems Using the XDM Dispersion Model. *Journal of Chemical Theory and Computation* **2015**, *11*, 4033–4040.

[260] Goerigk, L.; Grimme, S. A general database for main group thermochemistry, kinetics, and noncovalent interactions - Assessment of common and reparameterized (meta-)GGA density functionals. *Journal of Chemical Theory and Computation* **2010**, *6*, 107–126.

[261] Goerigk, L.; Grimme, S. Double-hybrid density functionals. *Wiley Interdisciplinary Reviews: Computational Molecular Science* **2014**, *4*, 576–600.

[262] Steinmann, S. N.; Corminboeuf, C. Comprehensive benchmarking of a density-dependent dispersion correction. *Journal of Chemical Theory and Computation* **2011**, *7*, 3567–3577.

[263] Řezáč, J.; Šimová, L.; Hobza, P. CCSD[T] describes noncovalent interactions better than the CCSD(T), CCSD(TQ), and CCSDT methods. *Journal of Chemical Theory and Computation* **2013**, *9*, 364–369.

[264] Burns, L. A.; Marshall, M. S.; Sherrill, C. D. Comparing Counterpoise-Corrected, Uncorrected, and Averaged Binding Energies for Benchmarking Noncovalent Interactions. *Journal of Chemical Theory and Computation* **2014**, *10*, 49–57.

[265] Řezáč, J.; Riley, K. E.; Hobza, P. Benchmark Calculations of Noncovalent Interactions of Halogenated Molecules. *Journal of Chemical Theory and Computation* **2012**, *8*, 4285–4292.

[266] Jurecka, P.; Sponer, J.; Cerný, J.; Hobza, P. Benchmark database of accurate (MP2 and CCSD(T) complete basis set limit) interaction energies of small model complexes, DNA base pairs, and amino acid pairs. *Physical chemistry chemical physics : PCCP* **2006**, *8*, 1985–1993.

[267] Cheng, L.; Gauss, J.; Ruscic, B.; Armentrout, P. B.; Stanton, J. F. Bond Dissociation Energies for Diatomic Molecules Containing 3d Transition Metals: Benchmark Scalar-Relativistic Coupled-Cluster Calculations for 20 Molecules. *Journal of Chemical Theory and Computation* **2017**, *13*, 1044–1056.

[268] Tsuzuki, S.; Uchimaru, T.; Mikami, M.; Urata, S. Magnitude and orientation dependence of intermolecular interaction of perfluoropropane dimer studied by high-level ab initio calculations: Comparison with propane dimer. *Journal of Chemical Physics* **2004**, *121*, 9917–9924.

[269] Liakos, D. G.; Neese, F. Improved correlation energy extrapolation schemes based on local pair natural orbital methods. *Journal of Physical Chemistry A* **2012**, *116*, 4801–4816.

[270] Pavošević, F.; Peng, C.; Pinski, P.; Riplinger, C.; Neese, F.; Valeev, E. F. SparseMaps - A systematic infrastructure for reduced scaling electronic structure methods. V. Linear scaling explicitly correlated coupled-cluster method with pair natural orbitals. *Journal of Chemical Physics* **2017**, *146*, 174108.

[271] Weymuth, T.; Proppe, J.; Reiher, M. Statistical Analysis of Semiclassical Dispersion Corrections. **2018**, 1–34.

[272] Caldeweyher, E.; Ehlert, S.; Hansen, A.; Neugebauer, H.; Spicher, S.; Bannwarth, C.; Grimme, S. A Generally Applicable Atomic-Charge Dependent London Dispersion Correction Scheme. **2019**,

[273] De Silva, P.; Corminboeuf, C. Simultaneous visualization of covalent and noncovalent interactions using regions of density overlap. *Journal of Chemical Theory and Computation* **2014**, *10*, 3745–3756.

[274] Sanderson, R. T. Principles of electronegativity Part II. Applications. *Journal of Chemical Education* **1988**, *65*, 227.

[275] Sanderson, R. T. Principles of electronegativity Part I. General nature. *Journal of Chemical Education* **1988**, *65*, 112.

[276] Parr, R. G.; Donnelly, R. A.; Levy, M.; Palke, W. E. Electronegativity: The density functional viewpoint. *The Journal of Chemical Physics* **1978**, *68*, 3801–3807.

[277] Geerlings, P.; De Proft, F.; Langenaeker, W. Conceptual Density Functional Theory. *Chemical Reviews* **2003**, *103*, 1793–1874.

[278] Proft, F. D.; Ayers, P. W.; Geerlings, P. In *The Conceptual Density Functional Theory Perspective of Bonding. In The Chemical Bond*; Frenking, G., Shaik, S., Eds.; 2014; pp 233–270.

[279] Nagle, J. K. Atomic Polarizability and Electronegativity. *Journal of the American Chemical Society* **1990**, *112*, 4741–4747.

[280] Parr, R. G.; Pearson, R. G. Absolute Hardness: Companion Parameter to Absolute Electronegativity. *Journal of the American Chemical Society* **1983**, *105*, 7512–7516.

[281] Pearson, R. G. Chemical hardness and density functional theory. *Journal of Chemical Sciences* **2005**, *117*, 369–377.

[282] Pearson, R. G. Hard and soft acids and bases, HSAB, part 1: Fundamental principles. *Journal of Chemical Education* **1968**, *45*, 581.

[283] Pearson, R. G. Hard and soft acids and bases, HSAB, part II: Underlying theories. *Journal of Chemical Education* **1968**, *45*, 643.

[284] Byrd, J. N.; Bartlett, R. J.; Montgomery, J. A. At What Chain Length Do Unbranched Alkanes Prefer Folded Conformations? *The Journal of Physical Chemistry A* **2014**, *118*, 1706–1712, PMID: 24524689.

[285] McKinnon, J. J.; Jayatilaka, D.; Spackman, M. A. Towards quantitative analysis of intermolecular interactions with Hirshfeld surfaces. *Chemical Communications* **2007**, 3814–3816.

[286] Simons, J. H.; Dunlap, R. D. The properties of n-pentforane and its mixtures with n-pentane. *The Journal of Chemical Physics* **1950**, *18*, 335–346.

[287] Bedford, R. G.; Dunlap, R. D. Solubilities and Volume Changes Attending Mixing for the System: Perfluoro-n-hexane–n-Hexane. *Journal of the American Chemical Society* **1958**, *80*, 282–285.

[288] Lau, C.; Anitole, K.; Hodes, C.; Lai, D.; Pfahles-Hutchens, A.; Seed, J. Perfluoroalkyl acids: A review of monitoring and toxicological findings. *Toxicological Sciences* **2007**, *99*, 366–394.

[289] L 150. *Official Journal of the European Union* **2017**, *60*, 1–19.

[290] Post, G. B.; Cohn, P. D.; Cooper, K. R. Perfluorooctanoic acid (PFOA), an emerging drinking water contaminant: A critical review of recent literature. *Environmental Research* **2012**, *116*, 93–117.

[291] Prevedouros, K.; Cousins, I. T.; Buck, R. C.; Korzeniowski, S. H. Sources, fate and transport of perfluorocarboxylates. *Environmental Science and Technology* **2006**, *40*, 32–44.

[292] Campos-Vallette, M.; Campos-Vallette, M.; Diaz-Fleming, G. Some physical and thermodynamic properties of n-cnf2n+2 compounds with n = 4-8. 1983.

[293] Morgado, P.; Laginhas, C. M.; Lewis, J. B.; McCabe, C.; Martins, L. F.; Filipe, E. J. Viscosity of liquid perfluoroalkanes and perfluoroalkylalkane surfactants. *Journal of Physical Chemistry B* **2011**, *115*, 9130–9139.

[294] Berkessel, A.; Elfert, S.; Yatham, V. R.; Neudörfl, J. M.; Schlörer, N. E.; Teles, J. H. Umpolung by N-heterocyclic carbenes: Generation and reactivity of the elusive 2,2-diamino enols (breslow intermediates). *Angewandte Chemie - International Edition* **2012**, *51*, 12370–12374.

[295] Berkessel, A.; Yatham, V. R.; Elfert, S.; Neudörfl, J. M. Characterization of the key intermediates of carbene-catalyzed umpolung by nmr spectroscopy and X-Ray diffraction: Breslow intermediates, homoenolates, and azolium enolates. *Angewandte Chemie - International Edition* **2013**, *52*, 11158–11162.

[296] Liu, R. Y.; Buchwald, S. L.; Liu, P.; Yang, Y.; Lu, G.; Lambrecht, D. S.; Fang, C. Ligand–Substrate Dispersion Facilitates the Copper-Catalyzed Hydroamination of Unactivated Olefins. *Journal of the American Chemical Society* **2017**, *139*, 16548–16555.

[297] Thomas, A. A.; Speck, K.; Kevlishvili, I.; Lu, Z.; Liu, P.; Buchwald, S. L. Mechanistically Guided Design of Ligands That Significantly Improve the Efficiency of CuH-Catalyzed Hydroamination Reactions. *Journal of the American Chemical Society* **2018**, *140*, 13976–13984.

[298] Harper, K. C.; Bess, E. N.; Sigman, M. S. Multidimensional steric parameters in the analysis of asymmetric catalytic reactions. *Nature Chemistry* **2012**, *4*, 366–374.

[299] Milo, A.; Bess, E. N.; Sigman, M. S. Interrogating selectivity in catalysis using molecular vibrations. *Nature* **2014**, *507*, 210–214.

[300] Sigman, M. S.; Harper, K. C.; Bess, E. N.; Milo, A. The Development of Multidimensional Analysis Tools for Asymmetric Catalysis and Beyond. *Accounts of Chemical Research* **2016**, *49*, 1292–1301, PMID: 27220055.

[301] Metsänen, T. T.; Lexa, K. W.; Santiago, C. B.; Chung, C. K.; Xu, Y.; Liu, Z.; Humphrey, G. R.; Ruck, R. T.; Sherer, E. C.; Sigman, M. S. Combining traditional 2D and modern physical organic-derived descriptors to predict enhanced enantioselectivity for the key aza-Michael conjugate addition in the synthesis of Prevymis^{TM} (letermovir). *Chemical Science* **2018**, 1–6.

[302] Reid, J. P.; Sigman, M. S. Comparing quantitative prediction methods for the discovery of small-molecule chiral catalysts. *Nature Reviews Chemistry* **2018**,

[303] Zahrt, A. F.; Henle, J. J.; Rose, B. T.; Wang, Y.; Darrow, W. T.; Denmark, S. E. Prediction of higher-selectivity catalysts by computer-driven workflow and machine learning. *Science* **2019**, *363*.

[304] Zahrt, A. F.; Denmark, S. E. Evaluating continuous chirality measure as a 3D descriptor in chemoinformatics applied to asymmetric catalysis. *Tetrahedron* **2019**, *75*, 1841 – 1851.

[305] Ahneman, D. T.; Estrada, J. G.; Lin, S.; Dreher, S. D.; Doyle, A. G. Predicting reaction performance in C–N cross-coupling using machine learning Predicting reaction performance in C–N cross-coupling using machine learning. *Science* **2018**, *360*, 186–190.

[306] Nielsen, M. K.; Ahneman, D. T.; Riera, O.; Doyle, A. G. Deoxyfluorination with sulfonyl fluorides: Navigating reaction space with machine learning. *Journal of the American Chemical Society* **2018**, *140*, 5004–5008.

[307] He, S.; Biedermann, F.; Vankova, N.; Zhechkov, L.; Heine, T.; Hoffman, R. E.; De Simone, A.; Duignan, T. T.; Nau, W. M. Cavitation energies can outperform dispersion interactions. *Nature Chemistry* **2018**, *10*, 5–10.

[308] Chen, Z. M.; Hilton, M. J.; Sigman, M. S. Palladium-Catalyzed Enantioselective Redox-Relay Heck Arylation of 1,1-Disubstituted Homoallylic Alcohols. *Journal of the American Chemical Society* **2016**, *138*, 11461–11464.

[309] Schneider, H. J. Dispersive Interactions in Solution Complexes. *Accounts of Chemical Research* **2015**, *48*, 1815–1822.

[310] Orlandi, M.; Coelho, J. A.; Hilton, M. J.; Toste, F. D.; Sigman, M. S. Parametrization of Non-covalent Interactions for Transition State Interrogation Applied to Asymmetric Catalysis. *Journal of the American Chemical Society* **2017**, *139*, 6803–6806.

[311] Orlandi, M.; Hilton, M. J.; Yamamoto, E.; Toste, F. D.; Sigman, M. S. Mechanistic Investigations of the Pd(0)-Catalyzed Enantioselective 1,1-Diarylation of Benzyl Acrylates. *Journal of the American Chemical Society* **2017**, *139*, 12688–12695.

[312] Coelho, J. A. S.; Matsumoto, A.; Orlandi, M.; Hilton, M. J.; Sigman, M. S.; Toste, F. D. Enantioselective fluorination of homoallylic alcohols enabled by the tuning of non-covalent interactions. *Chemical Science* **2018**, 7153–7158.

[313] Bader, R. F.; Carroll, M. T.; Cheeseman, J. R.; Chang, C. Properties of Atoms in Molecules: Atomic Volumes. *Journal of the American Chemical Society* **1987**, *109*, 7968–7979.

[314] Rahm, M.; Hoffmann, R.; Ashcroft, N. W. Atomic and Ionic Radii of Elements 1–96. *Chemistry - A European Journal* **2016**, *22*, 14625–14632.

[315] Spackman, M. A.; Byrom, P. G. A novel definition of a molecule in a crystal. *Chemical Physics Letters* **1997**, *267*, 215–220.

[316] Sinnokrot, M. O.; Sherrill, C. D. Highly accurate coupled cluster potential energy curves for the benzene dimer: Sandwich, T-shaped, and parallel-displaced configurations. *Journal of Physical Chemistry A* **2004**, *108*, 10200–10207.

[317] Kalugina, Y. N.; Cherepanov, V. N.; Buldakov, M. A.; Zvereva-Loîte, N.; Boudon, V. Theoretical investigation of the ethylene dimer: Interaction energy and dipole moment. *Journal of Computational Chemistry* **2012**, *33*, 319–330.

[318] Tsuzuki, S.; Sato, N. Origin of attraction in chalgogen-nitrogen interaction of 1,2,5-chalcogenadiazole dimers. *Journal of Physical Chemistry B* **2013**, *117*, 6849–6855.

[319] Cozzolino, A. F.; Vargas-Baca, I.; Mansour, S.; Mahmoudkhani, A. H. The nature of the supramolecular association of 1,2,5-chalcogenadiazoles. *Journal of the American Chemical Society* **2005**, *127*, 3184–3190.

[320] Bleiholder, C.; Werz, D. B.; Köppel, H.; Gleiter, R. Theoretical investigations on chalcogen-chalcogen interactions: What makes these nonbonded interactions bonding? *Journal of the American Chemical Society* **2006**, *128*, 2666–2674.

[321] Murray, J. S.; Lane, P.; Clark, T.; Politzer, P. σ-hole bonding: Molecules containing group VI atoms. *Journal of Molecular Modeling* **2007**, *13*, 1033–1038.

[322] Murray, J. S.; Lane, P.; Politzer, P. Simultaneous σ-Hole and Hydrogen Bonding by Sulfur- and Selenium- Containing Heterocycles. *International Journal of Quantum Chemistry* **2008**, *108*, 2770–2781.

[323] De Vleeschouwer, F.; Denayer, M.; Pinter, B.; Geerlings, P.; De Proft, F. Characterization of chalcogen bonding interactions via an in-depth conceptual quantum chemical analysis. *Journal of Computational Chemistry* **2018**, *39*, 557–572.

[324] Wang, W.; Ji, B.; Zhang, Y. Chalcogen bond: A sister noncovalent bond to halogen bond. *Journal of Physical Chemistry A* **2009**, *113*, 8132–8135.

[325] Seferos, D. S.; Taylor, M. S.; Gibson, G. L.; Straus, R. N.; Garrett, G. E. Chalcogen Bonding in Solution: Interactions of Benzotelluradiazoles with Anionic and Uncharged Lewis Bases. *Journal of the American Chemical Society* **2015**, *137*, 4126–4133.

[326] Cozzolino, A. F.; Whitfield, P. S.; Vargas-Baca, I. Supramolecular chromotropism of the crystalline phases of 4,5,6,7-tetrafluorobenzo-2,1,3-telluradiazole. *Journal of the American Chemical Society* **2010**, *132*, 17265–17270.

[327] Ams, M. R.; Trapp, N.; Schwab, A.; Milić, J. V.; Diederich, F. Chalcogen Bonding "2S–2N Squares" versus Competing Interactions: Exploring the Recognition Properties of Sulfur. *Chemistry – A European Journal* **2018**, chem.201804261.

[328] Riwar, L.-J.; Trapp, N.; Root, K.; Zenobi, R.; Diederich, F. Supramolecular Capsules: Strong versus Weak Chalcogen Bonding. *Angewandte Chemie International Edition* **2018**, 17259–17264.

[329] Rösel, S.; Quanz, H.; Logemann, C.; Becker, J.; Mossou, E.; Cañadillas-Delgado, L.; Caldeweyher, E.; Grimme, S.; Schreiner, P. R. London Dispersion

Enables the Shortest Intermolecular Hydrocarbon H···H Contact. *Journal of the American Chemical Society* **2017**, *139*, 7428–7431.

[330] Dopfer, O.; Nizkorodov, S. A.; Meuwly, M.; Bieske, E. J.; Maier, J. P. The $ nu$$_3 $ infrared spectrum of the He-NH$_4$$^+$ complex. *Chemical Physics Letters* **1996**, *260*, 545–550.

[331] Yoshioka, K.; Raston, P. L.; Anderson, D. T. Infrared spectroscopy of chemically doped solid parahydrogen. *International Reviews in Physical Chemistry* **2006**, *25*, 469–496.

[332] Coussan, S.; Bouteiller, Y.; Perchard, J. P.; Zheng, W. Q. Rotational isomerism of ethanol and matrix isolation infrared spectroscopy. *Journal of Physical Chemistry A* **1998**, *102*, 5789–5793.

[333] Häber, T.; Schmitt, U.; Suhm, M. A. FTIR-spectroscopy of molecular clusters in pulsed supersonic slit-jet expansions. *Physical Chemistry Chemical Physics* **1999**, *1*, 5573–5582.

[334] Oswald, S.; Suhm, M. A.; Coussan, S. Incremental NH stretching downshift through stepwise nitrogen complexation of pyrrole: a combined jet expansion and matrix isolation study. *Physical Chemistry Chemical Physics* **2019**, *21*, 1277–1284.

[335] Fritsche, L.; Bach, A.; Miloglyadova, L.; Tsybizova, A.; Chen, P. A 4 K FT-ICR cell for infrared ion spectroscopy. *Review of Scientific Instruments* **2018**, *89*, 063119.

[336] Eyler, J. R. INFRARED MULTIPLE PHOTON DISSOCIATION SPECTROSCOPY OF IONS IN PENNING TRAPS. *Mass Spectrometry Reviews* **2009**, *28*, 448–467.

[337] MacAleese, L.; Maitre, P. INFRARED SPECTROSCOPY OF ORGANOMETALLIC IONS IN THE GAS PHASE: FROM MODEL TO REAL WORLD COMPLEXES Luke. *Mass Spectrometry Reviews* **2007**, *26*, 583–605.

[338] Fridgen, T. D. INFRARED CONSEQUENCE SPECTROSCOPY OF GASEOUS PROTONATED AND METAL ION CATIONIZED COMPLEXES. *Mass Spectrometry Reviews* **2009**, *28*, 586–607.

[339] Polfer, N. C.; Oomens, J. VIBRATIONAL SPECTROSCOPY OF BARE AND SOLVATED IONIC COMPLEXES OF BIOLOGICAL RELEVANCE. *Mass Spectrometry Reviews* **2009**, *28*, 468–494.

[340] Roithová, J. Characterization of reaction intermediates by ion spectroscopy. *Chemical Society Reviews* **2012**, *41*, 547–559.

[341] Jašík, J.; Žabka, J.; Roithová, J.; Gerlich, D. Infrared spectroscopy of trapped molecular dications below 4 K Dedicated to the memory of Detlef Schröder. *International Journal of Mass Spectrometry* **2013**, *354-355*, 204–210.

[342] Johnson, C. J.; Wolk, A. B.; Fournier, J. A.; Sullivan, E. N.; Weddle, G. H.; Johnson, M. A. Communication: He-tagged vibrational spectra of the SarGlyH+ and H+(H2O)2,3 ions: Quantifying tag effects in cryogenic ion vibrational predissociation (CIVP) spectroscopy. *Journal of Chemical Physics* **2014**, *140*.

[343] Dillinger, S.; Mohrbach, J.; Hewer, J.; Gaffga, M.; Niedner-Schatteburg, G. Infrared spectroscopy of N2 adsorption on size selected cobalt cluster cations in isolation. *Physical Chemistry Chemical Physics* **2015**, *17*, 10358–10362.

[344] Pirim, C.; Krim, L. A neon-matrix isolation study of the reaction of non-energetic H-atoms with CO molecules at 3 K. *Physical Chemistry Chemical Physics* **2011**, *13*, 19454–19459.

[345] Tsiouris, M.; Pollack, I. B.; Lester, M. I. Infrared action spectroscopy and inelastic recoil dynamics of the CH4-OD reactant complex. *Journal of Physical Chemistry A* **2002**, *106*, 7722–7727.

[346] Thompson, B. C.; Fréchet, J. M. Polymer-fullerene composite solar cells. *Angewandte Chemie - International Edition* **2008**, *47*, 58–77.

[347] Imahori, H. Porphyrin – fullerene linked systems as arti fi cial photosynthetic mimics. *Org. Biomol. Chem.* **2004**, *2*, 1425–1433.

[348] D'Souza, F.; Ito, O. Supramolecular donor-acceptor hybrids of porphyrins/phthalocyanines with fullerenes/carbon nanotubes: Electron transfer, sensing, switching, and catalytic applications. *Chemical Communications* **2009**, 4913–4928.

[349] Schnorr, J. M.; Swager, T. M. Emerging applications of carbon nanotubes. *Chemistry of Materials* **2011**, *23*, 646–657.

[350] Li, J.; Pomerantseva, E.; Maleski, K.; Gogotsi, Y.; Wang, G.; Byles, B. W.; Ren, C. E.; Xie, X.; Anasori, B.; Zhao, M.-Q. Porous heterostructured MXene/-carbon nanotube composite paper with high volumetric capacity for sodium-based energy storage devices. *Nano Energy* **2016**, *26*, 513–523.

[351] Peng, C.; Zhang, S.; Jewell, D.; Chen, G. Z. Carbon nanotube and conducting polymer composites for supercapacitors. *Progress in Natural Science* **2008**, *18*, 777–788.

[352] Avouris, P.; Freitag, M.; Perebeinos, V. Carbon-nanotube photonics and optoelectronics. *Nature Photonics* **2008**, *2*, 341–350.

[353] Cao, Q.; Kim, H. S.; Pimparkar, N.; Kulkarni, J. P.; Wang, C.; Shim, M.; Roy, K.; Alam, M. A.; Rogers, J. A. Medium-scale carbon nanotube thin-film integrated circuits on flexible plastic substrates. *Nature* **2008**, *454*, 495–500.

[354] Niyogi, S.; Hamon, M. A.; Hu, H.; Zhao, B.; Bhowmik, P.; Sen, R.; Itkis, M. E.; Haddon, R. C. Chemistry of single-walled carbon nanotubes. *Acc. Chem. Res.* **2002**, *35*, 1105–1113.

[355] Zheng, M.; Jagota, A.; Strano, M. S.; Santos, Adelina, P.; Barone, P.; Chou, S. G.; Diner, B. A.; Dresselhaus, M. S.; Mclean, R. S.; Onoa, G. B.; Samsonidze, G. G.; Semke, E. D.; Usrey, M.; Walls, D. J. Structure-Based Carbon Nanotube Sorting by Sequence-Dependent DNA Assembly Ming Zheng,. *Science* **2010**, *302*, 1545–1549.

[356] Selmani, S.; Schipper, D. π-Concave Hosts for Curved Carbon Nanomaterials. *Chemistry - A European Journal* **2019**, *25*, 1–21.

[357] Tashiro, K.; Aida, T. Metalloporphyrin hosts for supramolecular chemistry of fullerenes. *Chemical Society Reviews* **2007**, *36*, 189–197.

[358] Pérez, E. M.; Martín, N. Curves ahead: Molecular receptors for fullerenes based on concave-convex complementarity. *Chemical Society Reviews* **2008**, *37*, 1512–1519.

[359] Hermann, J.; Alfè, D.; Tkatchenko, A. Nanoscale π-π Stacked molecules are bound by collective charge fluctuations. *Nature Communications* **2017**, *8*.

[360] Pyykkö, P.; Li, J.; Runeberg, N. Predicted ligand dependence of the Au(I)...Au(I) attraction in (XAuPH3)2. *Chemical Physics Letters* **1994**, *218*, 133–138.

[361] Scherbaum, F.; Grohmann, A.; Huber, B.; Krüger, C.; Schmidbaur, H. “Aurophilicity” as a Consequence of Relativistic Effects: The Hexakis(triphenylphosphaneaurio)methane Dication [(Ph3PAu)6C]2+. *Angewandte Chemie International Edition in English* **1988**, *27*, 1544–1546.

[362] Schmidbaur, H. The fascinating implications of new results in gold chemistry. *Gold Bulletin* **1990**, *23*, 11–21.

[363] Pyykkö, P. Strong closed-shell interactions in inorganic chemistry. *Chemical Reviews* **1997**, *97*, 597–636.

[364] Pyykkö, P.; Zhao, Y. Ab initio Calculations on the (ClAuPH3)2 Dimer with Relativistic Pseudopotential: Is the “Aurophilic Attraction” a Correlation Effect? *Angewandte Chemie International Edition in English* **1991**, *30*, 604–605.

[365] Cl, X.; Ph, L.; Pyykko, P.; Mendizabal, F. Theory of the d10 -d10 Closed-Shell Attraction : 2. Long-Distance Behaviour and Nonadditive Effects in Dimers and Trimers of Type [(X-Au-L)n] (n = 2, 3; X = Cl, I, H; L = PH3, PMe3, -N=CH). *Chem. Eur. J.* **1997**, *3*, 1458–1465.

[366] Pyykkö, P.; Runeberg, N.; Mendizabal, F. Theory of the d10-d10 Closed-Shell Attraction: 1. Dimers Near Equilibrium. *Chemistry - A European Journal* **1997**, *3*, 1451–1457.

[367] Mendizabal, F.; Pyykkö, P. Aurophilic attraction in binuclear complexes with Au(I) and Au(III). A theoretical stuy. *Phys. Chem. Chem. Phys.* **2004**, *6*, 900–905.

[368] Muñiz, J.; Wang, C.; Pyykkö, P. Aurophilicity: The effect of the neutral ligand l on [{ClAuL}2] systems. *Chemistry - A European Journal* **2011**, *17*, 368–377.

[369] Otero-De-La-Roza, A.; Mallory, J. D.; Johnson, E. R. Metallophilic interactions from dispersion-corrected density-functional theory. *Journal of Chemical Physics* **2014**, *140*.

[370] Runeberg, N.; Schütz, M.; Werner, H. J. The aurophilic attraction as interpreted by local correlation methods. *Journal of Chemical Physics* **1999**, *110*, 7210–7215.

[371] Andrejić, M.; Mata, R. A. Study of ligand effects in aurophilic interactions using local correlation methods. *Physical Chemistry Chemical Physics* **2013**, *15*, 18115–18122.

[372] Wuttke, A.; Feldt, M.; Mata, R. A. All That Binds Is Not Gold-The Relative Weight of Aurophilic Interactions in Complex Formation. *Journal of Physical Chemistry A* **2018**, *122*, 6918–6925.

[373] O'Grady, E.; Kaltsoyannis, N. Does metallophilicity increase or decrease down group 11? Computational investigations of [Cl-M-PH3]2(M = Cu, Ag, Au, [111]). *Physical Chemistry Chemical Physics* **2004**, *2*, 680–687.

[374] Bickelhaupt, F. M.; Houk, K. N. Analyzing Reaction Rates with the Distortion/Interaction-Activation Strain Model. *Angewandte Chemie - International Edition* **2017**, *56*, 10070–10086.

[375] Santiago, C. B.; Guo, J. Y.; Sigman, M. S. Predictive and mechanistic multivariate linear regression models for reaction development. *Chemical Science* **2018**, *9*, 2398–2412.

[376] Verloop, A.; Hoogenstraaten, W.; Tipker, J. *Drug Design: Volume Vii*; ACADEMIC PRESS, INC., 1976; Vol. 1962; pp 165–207.

[377] Foster, J. P.; Weinhold, F. Natural Hybrid Orbitals. *Journal of the American Chemical Society* **1980**, *102*, 7211–7218.

[378] Zhang, C.; Santiago, C. B.; Crawford, J. M.; Sigman, M. S. Enantioselective Dehydrogenative Heck Arylations of Trisubstituted Alkenes with Indoles to Construct Quaternary Stereocenters. *Journal of the American Chemical Society* **2015**, *137*, 15668–15671.

[379] Krishtal, A.; Vannomeslaeghe, K.; Geldof, D.; Van Alsenoy, C.; Geerlings, P. Importance of anisotropy in the evaluation of dispersion interactions. *Phys. Rev. A* **2011**, *83*, 024501.

[380] Sure, R.; Grimme, S. Comprehensive Benchmark of Association (Free) Energies of Realistic Host-Guest Complexes. *Journal of Chemical Theory and Computation* **2015**, *11*, 3785–3801.

[381] Stevens, E. Experimental electron density distribution of molecular chlorine. *Molecular Physics* **1979**, *37*, 27–45.

[382] Nyburg, S. C. 'Polar flattening': non-spherical effective shapes of atoms in crystals. *Acta Crystallographica Section A* **1979**, *35*, 641–645.

[383] El Kerdawy, A.; Murray, J. S.; Politzer, P.; Bleiziffer, P.; Heßelmann, A.; Görling, A.; Clark, T. Directional noncovalent interactions: Repulsion and dispersion. *Journal of Chemical Theory and Computation* **2013**, *9*, 2264–2275.

[384] Awwadi, F. F.; Willett, R. D.; Peterson, K. A.; Twamley, B. The Nature of Halogen–Halogen Synthons: Crystallographic and Theoretical Studies. *Chemistry – A European Journal* **2006**, *12*, 8952–8960.

[385] Politzer, P.; Murray, J. S. Halogen Bonding: An Interim Discussion. *ChemPhysChem* **2013**, *14*, 278–294.

[386] Politzer, P.; Murray, J. S.; Clark, T. Halogen bonding and other σ-hole interactions: a perspective. *Phys. Chem. Chem. Phys.* **2013**, *15*, 11178–11189.

[387] Clark, T. σ-Holes. *Wiley Interdisciplinary Reviews: Computational Molecular Science* **2013**, *3*, 13–20.

[388] Williams, G. J.; Stone, A. J. Distributed dispersion: A new approach. *The Journal of Chemical Physics* **2003**, *119*, 4620–4628.

[389] Ogilvie, J.; Wang, F. Y. Potential-energy functions of diatomic molecules of the noble gases. *Journal of Molecular Structure* **1992**, *273*, 277–290.

[390] Weininger, D. SMILES, a Chemical Language and Information System: 1: Introduction to Methodology and Encoding Rules. *Journal of Chemical Information and Computer Sciences* **1988**, *28*, 31–36.

[391] Weininger, D.; Weininger, A.; Weininger, J. L. SMILES. 2. Algorithm for Generation of Unique SMILES Notation. *Journal of Chemical Information and Computer Sciences* **1989**, *29*, 97–101.

[392] Weininger, D. Smiles. 3. Depict. Graphical Depiction of Chemical Structures. *Journal of Chemical Information and Computer Sciences* **1990**, *30*, 237–243.

[393] Boyle, N. M. O.; Banck, M.; James, C. A.; Morley, C.; Vandermeersch, T.; Hutchison, G. R. Open Babel: An open chemical toolbox. *Journal of Cheminformatics* **2011**, *3*, 33.

[394] Lu, T.; Chen, F. Atomic Dipole Moment Corrected Hirshfeld Population Method. *Journal of Theoretical and Computational Chemistry* **2012**, *11*, 163–183.

[395] Perdew, J. P.; Burke, K.; Ernzerhof, M. Generalized gradient approximation made simple. *Physical Review Letters* **1996**, *77*, 3865–3868.

[396] Weigend, F.; Ahlrichs, R. {B}alanced Basis Sets of Split Valence, Triple Zeta Valence and Quadruple Zeta Valence Quality for {H} to {R}n: {D}esign and Assessment of Accuracy. *Phys. Chem. Chem. Phys.* **2005**, *7*, 3297–3305.

[397] Rappoport, D.; Furche, F. Property-optimized Gaussian basis sets for molecular response calculations. *Journal of Chemical Physics* **2010**, *133*, 1–11.

[398] Rossiter, K. J. Structure-odor relationships. *Chemical Reviews* **1996**, *96*, 3201–3240.

[399] Sell, C. S. On the unpredictability of odor. *Angewandte Chemie - International Edition* **2006**, *45*, 6254–6261.

[400] Keller, A. et al. Predicting human olfactory perception from chemical features of odor molecules. *Science* **2017**, *355*, 820–826.

[401] Don, C. G.; Riniker, S. Scents and sense: in silico perspectives on olfactory receptors. *Journal of computational chemistry* **2014**, *35*, 2279–2287.

[402] Freyer, M. W.; Lewis, E. A. Isothermal Titration Calorimetry: Experimental Design, Data Analysis, and Probing Macromolecule/Ligand Binding and Kinetic Interactions. 2008.

[403] Valentini, M.; Rüegger, H.; Pregosin, P. S. Applications of pulsed-gradient spin-echo (PGSE) diffusion measurements in organometallic chemistry. *Helvetica Chimica Acta* **2001**, *84*, 2833–2853.

[404] Tyrrell, H. J. V.; Harris, K. R. *Diffusion in Liquids*, 1st ed.; 1984.

[405] Martìnez-viviente, E.; Pregosin, P. S. Lew Temperature PGSE Diffusion Measurements. Applications to Cationic Alcohol Complexes. *Helvetica Chimica Acta* **2003**, *86*, 2364–2378.

[406] Allampally, N. K.; Daniliuc, C. G.; Strassert, C. A.; De Cola, L. Tuning the structural and photophysical properties of cationic Pt(II) complexes bearing neutral bis(triazolyl)pyridine ligands. *Inorganic Chemistry* **2015**, *54*, 1588–1596.

[407] Wittig, G.; Steinhoff, G. Azatriptycen. *Justus Liebigs Annalen der Chemie* **1964**, *676*, 21–31.

[408] Gonzalez, J.; Jewell, T. M.; Li, H.; Linton, A.; Tatlock, J. H. Inhibitors of Hepatitis C Virus RNA-dependent RNA Polymerase, and Compositions and Treatments using the same. WO2006018725 (A1). February 23, 2006.

[409] Chang, H. K.; Oh, Y. S.; Jang, Y. J. Caspase Inhibitors Based on Pyridazinone Scaffold. WO2008016239 (A1). February 7, 2008.

[410] Surendra, K.; Rajendar, G.; Corey, E. J. Useful catalytic enantioselective cationic double annulation reactions initiated at an internal π-bond: Method and applications. *Journal of the American Chemical Society* **2014**, *136*, 642–645.

[411] Gibson, S. E.; Mainolfi, N.; Kalindjian, S. B.; Wright, P. T.; White, A. J. P. A new class of non-racemic chiral macrocycles: A conformational and synthetic study. *Chemistry - A European Journal* **2005**, *11*, 69–80.

[412] Prajapati, D.; Schulzke, C.; Kindermann, M. K.; Kapdi, a. R. Selective palladium-catalysed arylation of 2,6-dibromopyridine using N-heterocyclic carbene ligands. *RSC Adv.* **2015**, *5*, 53073–53085.

[413] Malhotra, S.; Seng, P. S.; Koenig, S. G.; Deese, A. J.; Ford, K. A. Chemoselective sp 2-sp3 cross-couplings: Iron-catalyzed alkyl transfer to dihaloaromatics. *Organic Letters* **2013**, *15*, 3698–3701.

[414] Li, A.; Muddana, H. S.; Gilson, M. K. Quantum Mechanical Calculation of Noncovalent Interactions: A Large-Scale Evaluation of PMx, DFT, and SAPT Approaches. *Journal of chemical theory and computation* **2014**, *10*, 1563–1575.

[415] Weimar, M.; Correa Da Costa, R.; Lee, F. H.; Fuchter, M. J. A scalable and expedient route to 1-Aza[6]helicene derivatives and its subsequent application to a chiral-relay asymmetric strategy. *Organic Letters* **2013**, *15*, 1706–1709.

[416] Nuñez, A.; Abarca, B.; Cuadro, A. M.; Alvarez-Builla, J.; Vaquero, J. J. Ring-Closing Metathesis Approach to Heteroaromatic Cations: Synthesis of Benzo[a]quinolizinium Salts. *European Journal of Organic Chemistry* **2011**, *2011*, 1280–1290.

[417] Voets, M.; Antes, I.; Scherer, C.; Biemel, K.; Barassin, C.; Marchais-Oberwinkler, S.; Hartmann, R. W. Synthesis and evaluation of heteroaryl substituted dihydronaphthalenes and indenes: potent and selective inhibitors of aldosterone synthase ({CYP11B2}) for the treatment of congestive heart failure and myocardial fibrosis. *Journal of Medicinal Chemistry* **2005**, *48*, 1563–1575.

[418] Li, A.; Gilbert, T. M.; Klumpp, D. A. Preparation of aza-polycyclic aromatic compounds via superelectrophilic cyclizations. *Journal of Organic Chemistry* **2008**, *73*, 3654–3657.

[419] Palmer, B. D.; Sutherland, H. S.; Blaser, A.; Kmentova, I.; Franzblau, S. G.; Wan, B.; Wang, Y.; Ma, Z.; Denny, W. A.; Thompson, A. M. Synthesis and Structure-Activity Relationships for Extended Side Chain Analogues of the Antitubercular Drug (6 S)-2-Nitro-6-{[4-(trifluoromethoxy)benzyl]oxy}-6,7-dihydro-5 H -imidazo[2,1- b][1,3]oxazine (PA-824). *Journal of Medicinal Chemistry* **2015**, *58*, 3036–3059.

[420] Fleckenstein, J. E.; Koszinowski, K. Lithium organozincate complexes LiRZnX2: Common species in organozinc chemistry. *Organometallics* **2011**, *30*, 5018–5026.

[421] Coutant, E. P.; Janin, Y. L. A study of Negishi cross-coupling reactions with benzylzinc halides to prepare original 3-ethoxypyrazoles. *Synthesis (Germany)* **2015**, *47*, 511–516.

[422] Blachut, D.; Szawkalo, J.; Czarnocki, Z. Suzuki-Miyaura and Negishi approaches to a series of forensically relevant pyridines and pyrimidines. *Synthesis* **2011**, 3496–3506.

[423] Elhaik, J.; Pask, C. M.; Kilner, C. A.; Halcrow, M. A. Synthesis of 2,6-di(pyrazol-1-yl)-4-bromomethylpyridine, and its conversion to other 2,6-di(pyrazol-1-yl)pyridines substituted at the pyridine ring. *Tetrahedron* **2007**, *63*, 291–298.

[424] Bogdan, N. D.; Matache, M.; Meier, V. M.; Dobrotä, C.; Dumitru, I.; Roiban, G. D.; Funeriu, D. P. Protein-inorganic array construction: Design and synthesis of the building blocks. *Chemistry - A European Journal* **2010**, *16*, 2170–2180.

[425] Glover, S. A.; Goosen, A.; Mccleland, C. W.; Vogel, F. R. Alcoholysis of t-butyl nitrite . Synthesis of alkyl nitrites i :. *South African Journal of Chemistry* **1981**, *34*, 96–100.

[426] Dolomanov, O. V.; Bourhis, L. J.; Gildea, R. J.; Howard, J. A. K.; Puschmann, H. OLEX2: A complete structure solution, refinement and analysis program. *Journal of Applied Crystallography* **2009**, *42*, 339–341.

[427] Sheldrick, G. M. SHELXT - Integrated space-group and crystal-structure determination. *Acta Crystallographica Section A: Foundations of Crystallography* **2015**, *71*, 3–8.

[428] Sheldrick, G. M. A short history of SHELX. *Acta Crystallographica Section A: Foundations of Crystallography* **2007**, *64*, 112–122.

[429] Duvvuri, H.; Wheeler, L. C.; Harms, M. J. Pytc: Open-Source Python Software for Global Analyses of Isothermal Titration Calorimetry Data. *Biochemistry* **2018**, *57*, 2578–2583.

[430] Klumpp, D. A.; Zhang, Y.; Kindelin, P. J.; Lau, S. Superacid-catalyzed reactions of pyridinecarboxaldehydes. *Tetrahedron* **2006**, *62*, 5915–5921.

[431] Cavazzini, M.; Manfredi, A.; Montanari, F.; Quici, S.; Pozzi, G. Asymmetric Epoxidation of Alkenes in Fluorinated Media, Catalyzed by Second-Generation Fluorous Chiral (Salen)manganese Complexes. *European Journal of Organic Chemistry* **2001**, *2001*, 4639.

[432] Adams, D. J.; Bennett, J. a.; Cole-Hamilton, D. J.; Hope, E. G.; Hopewell, J.; Kight, J.; Pogorzelec, P.; Stuart, A. M. Rhodium catalysed hydroformylation of alkenes using highly fluorophilic phosphines. *Dalton transactions (Cambridge, England : 2003)* **2005**, 3862–7.

[433] Neese, F. The ORCA program system. *Wiley Interdisciplinary Reviews: Computational Molecular Science* **2012**, *2*, 73–78.

[434] Grimme, S. Semiempirical hybrid density functional with perturbative second-order correlation. *Journal of Chemical Physics* **2006**, *124*, 034108.

[435] Grimme, S.; Huenerbein, R.; Ehrlich, S. On the importance of the dispersion energy for the thermodynamic stability of molecules. *ChemPhysChem* **2011**, *12*, 1258–1261.

[436] Newton, M. D.; Ostlund, N. S.; Pople, J. A. Projection of Diatomic Differential Overlap: Least-Squares Projection of Two-Center Distributions onto One-Center Functions. *The Journal of Chemical Physics* **1968**, *49*, 5192–5194.

[437] Baerends, E. J.; Ellis, D. E.; Ros, P. Self-consistent molecular Hartree-Fock-Slater calculations I. The computational procedure. *Chemical Physics* **1973**, *2*, 41–51.

[438] Skylaris, C. K.; Gagliardi, L.; Handy, N. C.; Ioannou, A. G.; Spencer, S.; Willetts, A. On the resolution of identity Coulomb energy approximation in density functional theory. *Journal of Molecular Structure: THEOCHEM* **2000**, *501-502*, 229–239.

[439] Sure, R.; Grimme, S. Corrected small basis set Hartree-Fock method for large systems. *Journal of Computational Chemistry* **2013**, *34*, 1672–1685.

[440] Frisch, M. J. et al. Gaussian09 Revision D.01.

[441] Neese, F.; Valeev, E. F. Revisiting the atomic natural orbital approach for basis sets: Robust systematic basis sets for explicitly correlated and conventional correlated ab initio methods? *Journal of Chemical Theory and Computation* **2011**, *7*, 33–43.

[442] Boys, S. S.; Bernardi, F. The calculation of small molecular interactions by the differences of separate total energies. Some procedures with reduced errors. *Molecular Physics* **1970**, *19*, 553–566.

[443] Turney, J. M. et al. Psi4: An open-source ab initio electronic structure program. *Wiley Interdisciplinary Reviews: Computational Molecular Science* **2012**, *2*, 556–565.

[444] Hohenstein, E. G.; Sherrill, C. D. Density fitting and Cholesky decomposition approximations in symmetry-adapted perturbation theory: Implementation and application to probe the nature of $\pi - \pi$ interactions in linear acenes. *Journal of Chemical Physics* **2010**, *132*.

[445] Hohenstein, E. G.; Parrish, R. M.; Sherrill, C. D.; Turney, J. M.; Schaefer, H. F. Large-scale symmetry-adapted perturbation theory computations via density fitting and Laplace transformation techniques: Investigating the fundamental forces of DNA-intercalator interactions. *Journal of Chemical Physics* **2011**, *135*.

[446] Papajak, E.; Zheng, J.; Xu, X.; Leverentz, H. R.; Truhlar, D. G. Perspectives on basis sets beautiful: Seasonal plantings of diffuse basis functions. *Journal of Chemical Theory and Computation* **2011**, *7*, 3027–3034.

[447] Baerends, E. J. et al. ADF2017.

[448] Pye, C. C.; Ziegler, T.; van Lenthe, E.; Louwen, J. N. An implementation of the conductor-like screening model of solvation within the Amsterdam density functional package — Part II. COSMO for real solvents 1. *Canadian Journal of Chemistry* **2009**, *87*, 790–797.

[449] Schrödinger, L. The PyMOL Molecular Graphics System, Version 1.8. 2015.

[450] Neese, F. Software update: The ORCA program system, version 4.0. *Wiley Interdisciplinary Reviews: Computational Molecular Science* **2017**, 1–6.

[451] Bannwarth, C.; Ehlert, S.; Grimme, S. GFN2-xTB – an accurate and broadly quantum chemical method with multipole electrostatics and density-dependent dispersion contributions. *ChemRxiv* Preprint.

[452] Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *Journal of Computational Physics* **1977**, *23*, 327–341.

[453] Parrish, R. M. et al. Psi4 1.1: An Open-Source Electronic Structure Program Emphasizing Automation, Advanced Libraries, and Interoperability. *Journal of Chemical Theory and Computation* **2017**, *13*, 3185–3197.

[454] Werner, H. J.; Knowles, P. J.; Knizia, G.; Manby, F. R.; Schütz, M. Molpro: A general-purpose quantum chemistry program package. *Wiley Interdisciplinary Reviews: Computational Molecular Science* **2012**, *2*, 242–253.

[455] Shao, Y. et al. Advances in molecular quantum chemistry contained in the Q-Chem 4 program package. *Molecular Physics* **2015**, *113*, 184–215.

[456] Ahrens, J.; Geveci, B.; Law, C. In *ParaView: An End-User Tool for Large Data Visualization*; Handbook, V., Ed.; Elsevier, 2005.

[457] Ayachit, U. *The ParaView Guide: A Parallel Visualization Application*; Kitware, 2015.

[458] Keith, T. A. *AIMAll (Version 17.11.14)*; TK Gristmill Software, 2017 (aim.tkgristmill.com).

[459] Johnson, S.; Drowns, M.; Tatlock, J.; Linton, A.; Gonzalez, J.; Hoffman, R.; Jewell, T.; Patel, L.; Blazel, J.; Tang, M.; Li, H. Synthetic route optimization of PF-00868554, an HCV polymerase inhibitor in clinical evaluation. *Synlett* **2010**, 796–800.

[460] *Chemcraft - graphical software for visualization of quantum chemistry computations.*; https://www.chemcraftprog.com.

[461] Foundation, P. S. *Python Language Reference, version 3.6*; Available at http://www.python.org.

[462] Weigend, F.; Ahlrichs, R. Balanced basis sets of split valence, triple zeta valence and quadruple zeta valence quality for H to Rn: Design and assessment of accuracy. *Physical Chemistry Chemical Physics* **2005**, *7*, 3297–305.

[463] Neese, F.; Wennmohs, F.; Hansen, A.; Becker, U. Efficient, approximate and parallel Hartree-Fock and hybrid DFT calculations. A 'chain-of-spheres' algorithm for the Hartree-Fock exchange. *Chemical Physics* **2009**, *356*, 98–109.

[464] Kossmann, S.; Neese, F. Efficient structure optimization with second-order many-body perturbation theory: The RIJCOSX-MP2 method. *Journal of Chemical Theory and Computation* **2010**, *6*, 2325–2338.

[465] Weigend, F. Accurate Coulomb-fitting basis sets for H to Rn. *Physical Chemistry Chemical Physics* **2006**, *8*, 1057–1065.

[466] Hellweg, A.; Hättig, C.; Höfener, S.; Klopper, W. Optimized accurate auxiliary basis sets for RI-MP2 and RI-CC2 calculations for the atoms Rb to Rn. *Theoretical Chemistry Accounts* **2007**, *117*, 587–597.

[467] Dunning, T. H. Gaussian basis sets for use in correlated molecular calculations. I. The atoms boron through neon and hydrogen. *The Journal of Chemical Physics* **1989**, *90*, 1007–1023.

[468] Kendall, R. A.; Dunning, T. H.; Harrison, R. J. Electron affinities of the first-row atoms revisited. Systematic basis sets and wave functions. *The Journal of Chemical Physics* **1992**, *96*, 6796–6806.

[469] Peterson, K. A.; Adler, T. B.; Werner, H. J. Systematically convergent basis sets for explicitly correlated wavefunctions: The atoms H, He, B-Ne, and Al-Ar. *Journal of Chemical Physics* **2008**, *128*, 084102.

[470] Sinanoğlu, O. Many-electron theory of atoms and molecules. I. Shells, electron pairs vs many-electron correlations. *The Journal of Chemical Physics* **1962**, *36*, 706–717.

[471] Cizek, J. On the Correlation Problem in Atomic and Molecular Systems. Calculation of Wavefunction Components in Ursell-Type Expansion Using Quantum-Field Theoretical Methods. *J. Chem. Phys.* **1966**, *45*, 4256.

[472] Weigend, F. A fully direct RI-HF algorithm: Implementation, optimised auxiliary basis sets, demonstration of accuracy and efficiency. *Physical Chemistry Chemical Physics* **2002**, *4*, 4285–4291.

[473] Pople, J. A.; Binkley, J. S.; Seeger, R. Theoretical models incorporating electron correlation. *International Journal of Quantum Chemistry* **1976**, *10*, 1–19.

[474] Pople, J. A.; Seeger, R.; Krishnan, R. Variational configuration interaction methods and comparison with perturbation theory. *International Journal of Quantum Chemistry* **1977**, *11*, 149–163.

[475] Weigend, F. Hartree-fock exchange fitting basis sets for H to Rn. *Journal of Computational Chemistry* **2008**, *29*, 167–175.

[476] Lee, C.; Yang, W.; Parr, R. G. Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. *Physical Review B* **1988**, *37*, 785–789.

[477] Becke, A. D. Density-functional thermochemistry. III. The role of exact exchange. *The Journal of Chemical Physics* **1993**, *98*, 5648–5652.

[478] Van Lenthe, E.; Baerends, E. J. Optimized Slater-type basis sets for the elements 1-118. *Journal of Computational Chemistry* **2003**, *24*, 1142–1156.

[479] Franchini, M.; Philipsen, P. H. T.; Van Lenthe, E.; Visscher, L. Accurate Coulomb potentials for periodic and molecular systems through density fitting. *Journal of Chemical Theory and Computation* **2014**, *10*, 1994–2004.

[480] Tao, J.; Perdew, J. P.; Staroverov, V. N.; Scuseria, G. E. Climbing the density functional ladder: Nonempirical meta–generalized gradient approximation designed for molecules and solids. *Physical Review Letters* **2003**, *91*, 3–6.

[481] Lin, Y. S.; Li, G. D.; Mao, S. P.; Chai, J. D. Long-range corrected hybrid density functionals with improved dispersion corrections. *Journal of Chemical Theory and Computation* **2013**, *9*, 263–272.

[482] Papajak, E.; Truhlar, D. G. Convergent Partially Augmented Basis Sets for Post-Hartree - Fock Calculations of Molecular Properties and Reaction Barrier Heights. *Journal of Chemical Theory and Computation* **2011**, 10–18.

[483] Gonthier, J. F.; Sherrill, C. D. Density-fitted open-shell symmetry-adapted perturbation theory and application to π-stacking in benzene dimer cation and ionized DNA base pair steps. *Journal of Chemical Physics* **2016**, *145*, 134106.

[484] Werner, H. J.; Manby, F. R.; Knowles, P. J. Fast linear scaling second-order Møller-Plesset perturbation theory (MP2) using local and density fitting approximations. *Journal of Chemical Physics* **2003**, *118*, 8149–8160.

[485] Head-Gordon, M.; Pople, J. A.; Frisch, M. J. MP2 energy evaluation by direct methods. *Chemical Physics Letters* **1988**, *153*, 503–506.

[486] Zou, W. Molden2AIM. 2019; `https://github.com/zorkzou/Molden2AIM`.

[487] Hansen, A.; Pracht, P.; Bannwarth, C.; Pisarek, J.; Seibert, J.; Dohm, S.; Neese, F.; Grimme, S. Fully Automated Quantum-Chemistry-Based Computation of Spin-Spin-Coupled Nuclear Magnetic Resonance Spectra. *Angew. Chemie Int. Ed.* **2017**, *56*, 14763–14769.

[488] Grimme, S. Exploration of Chemical Compound, Conformer, and Reaction Space with Meta-Dynamics Simulations Based on Tight-Binding Quantum Chemical Calculations. *Journal of Chemical Theory and Computation* **2019**,

[489] Khartabil, H.; Contreras-García, J.; Boisson, J.-C.; Lefebvre, C.; Hénon, E.; Rubez, G. Accurately extracting the signature of intermolecular interactions present in the NCI plot of the reduced density gradient versus electron density. *Phys. Chem. Chem. Phys.* **2017**, *19*, 17928–17936.

[490] Lefebvre, C.; Khartabil, H.; Boisson, J. C.; Contreras-García, J.; Piquemal, J. P.; Hénon, E. The Independent Gradient Model: A New Approach for Probing Strong and Weak Interactions in Molecules from Wave Function Calculations. *ChemPhysChem* **2018**, 724–735.

[491] Pedregosa, F. et al. Scikit-learn: Machine Learning in Python . *Journal of Machine Learning Research* **2011**, *12*, 2825–2830.

[492] Reiher, M.; Wolf, A. Exact decoupling of the Dirac Hamiltonian. I. General theory. *J. Chem. Phys.* **2004**, *121*, 2037–2047.

[493] Reiher, M. Douglas-Kroll-Hess theory: A relativistic electrons-only theory for chemistry. *Theor. Chem. Acc.* **2006**, *116*, 241–252.

[494] Roos, B. O.; Lindh, R.; Malmqvist, P. Å.; Veryazov, V.; Widmark, P. O. Main Group Atoms and Dimers Studied with a New Relativistic ANO Basis Set. *J. Phys. Chem. A* **2004**, *108*, 2851–2858.

D. Curriculum Vitae

Name	Robert Pollice
Date of Birth	28th March 1992
Place of Birth	Neunkirchen, Austria
Citizenship	Austria

Education

2015 – 2019	Ph.D. thesis in the group of Prof. Dr. Peter Chen, ETH Zurich, Switzerland Doctoral thesis: London Dispersion in Molecular Systems
2013 – 2015	M.Sc. Technical Chemistry, TU Wien, Austria Master thesis: Mechanistic Investigations of C–H Activation Reactions
2010 – 2013	B.Sc. Technical Chemistry, TU Wien, Austria Bachelor thesis: Mechanistic & substrate scope investigation of direct alkylation reactions by C–H activation

Experience

2015 – 2019	Ph.D. research with Prof. Dr. Peter Chen
2018 – 2019	Supervision of two undergraduate research projects and one master thesis, ETH Zurich, Switzerland
2018	Teaching assistant for Organic Chemistry IV: Physical Organic Chemistry, ETH Zurich, Switzerland
2015 – 2016	Teaching assistant for Practical Course Organic Chemistry, ETH Zurich, Switzerland
2013 – 2015	M.Sc. research with Prof. Dr. Michael Schnürch
2014	Teaching assistant for Synthesis Laboratory Course, TU Wien, Austria