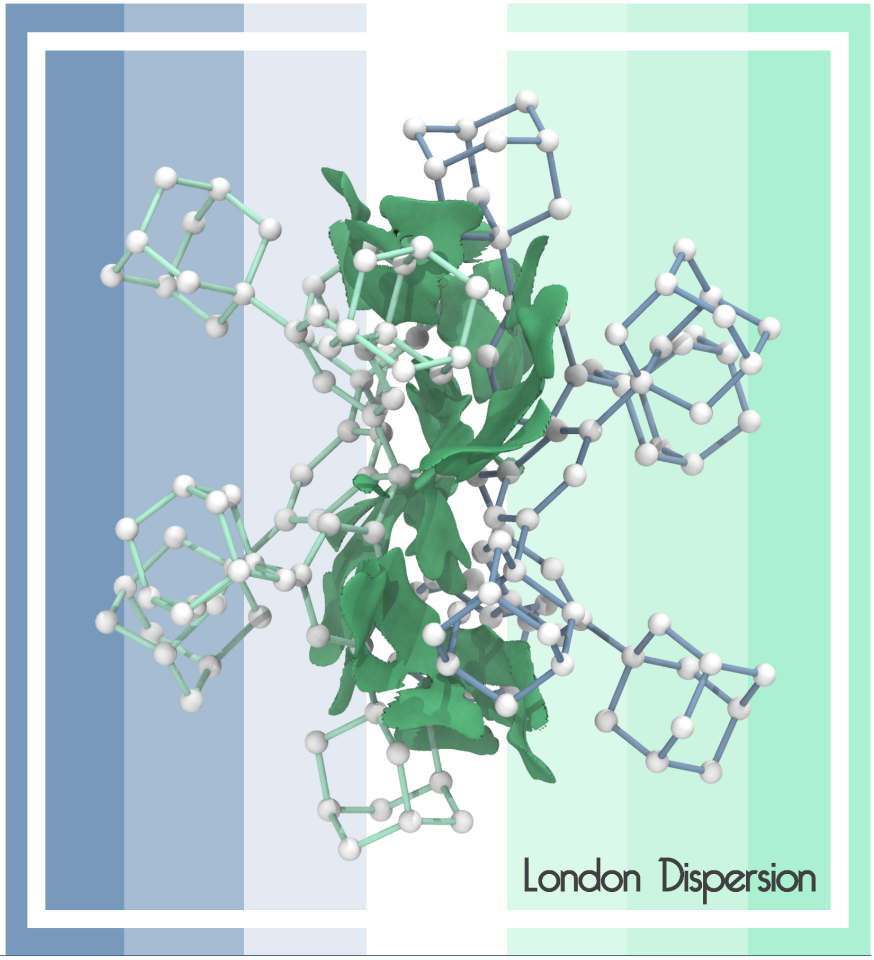
BEILSTEIN JOURNAL OF ORGANIC CHEMISTRY

Dispersion interactions

Edited by Peter R. Schreiner



Imprint

Beilstein Journal of Organic Chemistry
www.bjoc.org
ISSN 1860-5397
Email: journals-support@beilstein-institut.de

The *Beilstein Journal of Organic Chemistry* is published by the Beilstein-Institut zur Förderung der Chemischen Wissenschaften.

Beilstein-Institut zur Förderung der Chemischen Wissenschaften
Trakehner Straße 7–9
60487 Frankfurt am Main
Germany
www.beilstein-institut.de

BEILSTEIN JOURNAL OF ORGANIC CHEMISTRY

Dispersion interactions

Peter R. Schreiner

Editorial

Open Access

Address:
Institute of Organic Chemistry, Justus-Liebig University, Heinrich-Buff-Ring 17, 35392 Giessen, Germany

Email:
Peter R. Schreiner - prs@uni-giessen.de

Keywords:
London dispersion; van-der-Waals potential

Beilstein J. Org. Chem. **2018,** *14,* 3076–3077.
doi:10.3762/bjoc.14.286

Received: 20 November 2018
Accepted: 07 December 2018
Published: 18 December 2018

This article is part of the Thematic Series "Dispersion interactions".

Guest Editor: P. R. Schreiner

London dispersion (LD) [1-3], the attractive part of the van-der-Waals [4] (vdW) potential (Figure 1), has long been recognized as an important binding interaction, just not so much in molecular organic, or for that matter, also inorganic chemistry. The repulsive part of the vdW potential has been well appreciated and is synonymous with the notion of "steric repulsion" [5,6]. The dilemma is that without a balanced description of attractive and repulsive forces, there cannot be a full understanding of structure and reactivity in chemistry. That is why we refer to "equilibrium structures" in spectroscopy and theory because there is perfect (time-averaged) equilibrium of all forces at work: they sum up to zero. The neglect of the attractive part of the vdW potential probably derives from the fact that there is no classic analogue as in the case of repulsion for which the hard sphere atom model works well.

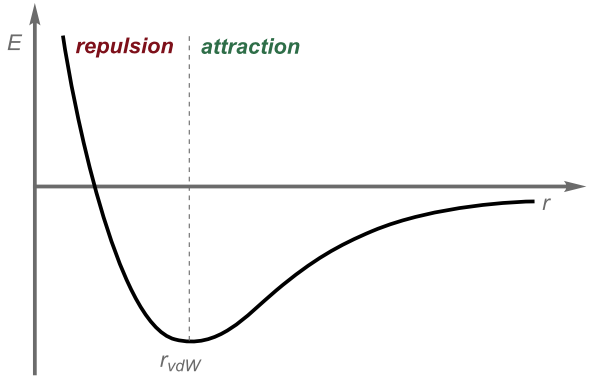


Figure 1: Dispersion = attractive part of the van-der-Waals potential.

LD is a purely quantum mechanical effect due to electron correlation. It is present for all matter starting from atoms (e.g., the condensation of helium) over molecules (e.g., aggregation) to materials (e.g., adhesion). It is also difficult to include in theoretical approaches and for the longest time it had been left virtually unnoticed that many density functional theory (DFT) approaches largely lacked the inclusion of LD. The often very good results of standard DFT implementations in the description of chemical reactions can in part be traced back to the compensating effects of neglecting both dispersion and solvation. This is not to say that dispersion vanishes [7] in solution but it is certainly attenuated [8]. We are still in the process of understanding just by how much.

LD is a driving force for molecular aggregation that plays a key role in the thermodynamic stability, molecular recognition, chemical selectivity through transition-state stabilization, pro-

tein folding, enzyme catalysis, and much more. Hence, this thematic issue covers selected aspects of the role LD plays for structures and reactivity. Naturally, it addresses diverse topics for which LD is particularly apparent.

Peter R. Schreiner

Giessen, November 2018

References

1. Eisenschitz, R.; London, F. *Z. Phys.* **1930,** *60,* 491–527. doi:10.1007/bf01341258
2. London, F. *Z. Phys.* **1930,** *63,* 245–279. doi:10.1007/bf01421741
3. London, F. *Trans. Faraday Soc.* **1937,** *33,* 8b. doi:10.1039/tf937330008b
4. van der Waals, J. D. Over de Continuiteit van den Gas - en Vloeistoftoestand. Ph.D. Thesis, Leiden University, Leiden, Netherlands, 1873.
5. Wagner, J. P.; Schreiner, P. R. *Angew. Chem., Int. Ed.* **2015,** *54,* 12274–12296. doi:10.1002/anie.201503476
6. Liptrot, D. J.; Power, P. P. *Nat. Rev. Chem.* **2017,** *1,* No. 0004. doi:10.1038/s41570-016-0004
7. Yang, L.; Adam, C.; Nichol, G. S.; Cockroft, S. L. *Nat. Chem.* **2013,** *5,* 1006–1010. doi:10.1038/nchem.1779
8. Pollice, R.; Bot, M.; Kobylianskii, I. J.; Shenderovich, I.; Chen, P. *J. Am. Chem. Soc.* **2017,** *139,* 13126–13140. doi:10.1021/jacs.7b06997

License and Terms

BEILSTEIN JOURNAL OF ORGANIC CHEMISTRY

Terahertz spectroscopy of 2,4,6-trinitrotoluene molecular solids from first principles

Ido Azuri[1], Anna Hirsch[1], Anthony M. Reilly[2], Alexandre Tkatchenko[3], Shai Kendler[4], Oded Hod[5] and Leeor Kronik[*1]

Full Research Paper

Open Access

Address:
[1]Department of Materials and Interfaces, Weizmann Institute of Science, Rehovoth 76100, Israel, [2]School of Chemical Sciences, Dublin City University, Glasnevin, Dublin 9, Ireland, [3]Physics and Materials Research Unit, University of Luxembourg, L-1511 Luxembourg, [4]Israel Institute for Biological Research, Ness Ziona 74100, Israel and [5]Department of Physical Chemistry, School of Chemistry, The Raymond and Beverly Sackler Faculty of Exact Sciences and The Sackler Center for Computational Molecular and Materials Science, Tel Aviv University, Tel Aviv 6997801, Israel

Email:
Leeor Kronik[*] - leeor.kronik@weizmann.ac.il

* Corresponding author

Keywords:
density functional theory; THz spectroscopy

Beilstein J. Org. Chem. **2018,** *14,* 381–388.
doi:10.3762/bjoc.14.26

Received: 29 October 2017
Accepted: 22 January 2018
Published: 09 February 2018

This article is part of the Thematic Series "Dispersion interactions".

Guest Editor: P. Schreiner

Abstract

We present a computational analysis of the terahertz spectra of the monoclinic and the orthorhombic polymorphs of 2,4,6-trinitrotoluene. Very good agreement with experimental data is found when using density functional theory that includes Tkatchenko–Scheffler pair-wise dispersion interactions. Furthermore, we show that for these polymorphs the theoretical results are only weakly affected by many-body dispersion contributions. The absence of dispersion interactions, however, causes sizable shifts in vibrational frequencies and directly affects the spatial character of the vibrational modes. Mode assignment allows for a distinction between the contributions of the monoclinic and orthorhombic polymorphs and shows that modes in the range from 0 to ca. 3.3 THz comprise both inter- and intramolecular vibrations, with the former dominating below ca. 1.5 THz. We also find that intramolecular contributions primarily involve the nitro and methyl groups. Finally, we present a prediction for the terahertz spectrum of 1,3,5-trinitrobenzene, showing that a modest chemical change leads to a markedly different terahertz spectrum.

Introduction

The transparency of many non-conductive materials in the terahertz range (0.1 to 10 THz) of the electromagnetic spectrum has led to the development of several potential applications of terahertz radiation [1]. Terahertz spectroscopy of molecular solids, in particular, has gained significant attention as it offers the possibility of distinguishing between different solid forms based on the signature of intermolecular vibrations [1-8]. For example, terahertz spectroscopy has been used to distinguish between dif-

ferent polymorphs of molecular solids used for pharmaceutical purposes [4,5,9-11], to detect contamination in food [12], in genetic [13] and skin-cancer [14] diagnosis, and in other applications [1,2,4-7].

One intriguing application of terahertz spectroscopy is the detection of energetic materials, which is of obvious importance for defense purposes [15]. Indeed, terahertz spectra were measured and analyzed for a variety of typical energetic materials, such as octahydro-1,3,5,7-tetranitro-1,3,5,7-tetraazacyclooctane (HMX), 2,4-dinitrotoluene (DNT), cyclotrimethylenetrinitramine (RDX), and pentaerythritol tetranitrate (PETN) [8,16-18]. Specifically for 2,4,6-trinitrotoluene (TNT), a very well-known energetic material, terahertz spectra were measured by several groups [16-19]. In particular, Melinger et al. [16] have obtained a high-resolution, low-temperature (12 K) spectrum at frequencies up to approx. 3.5 THz. The analysis of these spectra, especially with respect to the assignment of vibrational modes, is complicated by the fact that TNT samples typically contain at least two co-existing polymorphs, a monoclinic one (majority) and an orthorhombic one (minority). Furthermore, the relative contribution of inter- and intramolecular vibrational components remains under debate.

The above difficulties in analysis can be overcome by employing first principles calculations, in which each pure polymorph can be studied individually [20-22]. A leading first principles approach that can yield reliable simulated spectra for complex materials is density functional theory (DFT) [23]. A significant complication, however, is that conventionally used exchange–correlation energy functionals in DFT do not describe the intermolecular dispersion interactions well. Therefore, early calculations employing them were not always able to achieve satisfactory agreement with experiment, as cautioned in [24]. Recent years have seen major improvements in DFT augmented by pair-wise dispersion interaction terms, to the point where they can be used regularly to predict properties of molecular solids [21,25,26]. Indeed, simulated terahertz spectra based on dispersion-inclusive DFT were recently reported for several molecular crytals [20,27].

Here, we employ dispersion-inclusive DFT calculations based on the Tkatchenko–Scheffler (TS) approach [28] to study the terahertz spectra of TNT. These calculations are used to assign modes in both TNT polymorphs studied. They reveal modes contributed by the different polymorphs and distinguish modes dominated by intermolecular motion (at low frequencies) from modes dominated by a combination of inter- and intramolecular movement (at higher frequencies). These results are further validated by comparing them to uncorrected DFT calculations on the one hand and to more sophisticated many-body dispersion DFT calculations on the other hand. The same methodology is then used to predict the terahertz spectra of the related 1,3,5-trinitrobenzene (TNB) molecular solid, demonstrating that the elimination of the methyl group changes significant fingerprints in the terahertz spectrum.

Computational Approach

All calculations were performed based on the generalized-gradient approximation exchange–correlation functional of Perdew, Burke and Ernzerhof (PBE) [29], with or without Tkatchenko–Scheffler–van der Waals (TS-vdW) interactions [28]. In this approach, the vdW energy is added as a pair-wise interaction and has only one semi-empirical parameter. This parameter determines the onset of the pair-wise interaction and is fitted, once and for all per a given functional, against the S22 data set of weakly bounded complexes [30]. For going beyond pair-wise interactions, we apply the many-body dispersion (MBD) method [31,32]. Within this approach, one first evaluates the TS-vdW dispersion parameters. Then, the atomic response functions are mapped onto a set of quantum harmonic oscillators that are coupled through dipole–dipole interactions to obtain self-consistent screened polarizabilities. The latter are used to calculate the correlation energy of the interacting oscillator model system, within the random-phase approximation.

Most calculations presented here were performed using VASP, a projector-augmented planewave code [33], using an energy planewave cutoff of 950 eV. Comparison to MBD calculations was performed within the CASTEP code [34], with an energy planewave cutoff of 800 eV. For both above-mentioned polymorphs of TNT, the Brillouin zone of the crystallographic unit cell was sampled using a Monkhorst–Pack *k*-grid [35] of $2 \times 4 \times 1$. We have additionally computed the orthorhombic polymorph of TNB, using a Monkhorst–Pack *k*-point grid of $2 \times 1 \times 2$ along the three reciprocal lattice vectors. The self-consistent cycle was converged to better than 10^{-7} eV for the total energy, to allow for numerically stable derivatives. Complete relaxation of all forces and stress components was performed prior to the calculation of vibrational frequencies. All forces in the optimized structures were smaller than 5×10^{-3} eV/Å, and all stress components were smaller than 0.02 GPa in the VASP and CASTEP codes.

Vibrational frequencies, f_n, for each structure were calculated from $f_n = (1/2\pi)\varepsilon_n^{1/2}$, where ε_n are the eigenvalues of the mass-weighted Hessian matrix, W. The matrix elements of W are given by:

$$W_{\alpha i,\beta j} = \frac{H_{\alpha i,\beta j}}{\sqrt{M_\alpha M_\beta}}, \quad (1)$$

where M_α and M_β are the masses of atoms α and β, and $H_{\alpha i,\beta j}$ are matrix elements of the Hessian matrix H, given by:

$$H_{\alpha i,\beta j} = \frac{\partial^2 E}{\partial r_{\alpha i} \partial r_{\beta j}}, \quad (2)$$

where i, j denote the Cartesian directions x,y, or z.

The Hessian matrix elements were determined numerically from the forces acting on the atoms for a given displacement (as computed via analytical derivatives), using the form:

$$H_{\alpha i,\beta j} = \frac{1}{2}\left(\frac{F_{\beta j}^{+\delta r_{\alpha i}} - F_{\beta j}^{-\delta r_{\alpha i}}}{2\delta r} + \frac{F_{\alpha i}^{+\delta r_{\beta j}} - F_{\alpha i}^{-\delta r_{\beta j}}}{2\delta r} \right), \quad (3)$$

where F is the force acting on atom β/α in direction j/i (see subscript of F), as a result of atom α/β being displaced by ±δr in the direction i/j (see superscript of F). The displacement amplitude for constructing the Hessian was chosen as 0.01 Å for TNT and 0.015 Å for TNB. These displacement amplitudes were chosen to be large enough to minimize numerical noise but small enough to minimize anharmonic contributions. We found the calculated frequencies to be numerically stable to within 0.05 THz at most and typically less than that.

The normalized eigenvectors, $v_{\alpha i}^n$, of W are factorized by $M_\alpha^{-(1/2)}$ to yield the normal mode displacement eigenvectors:

$$u_{\alpha i}^n = \frac{1}{\sqrt{M_\alpha}} v_{\alpha i}^n. \quad (4)$$

Finally, the absorption intensity I_n of each mode is given by [36-41]:

$$I_n \propto d_n \sum_j \overline{Z}_{n,j}{}^2, \quad (5)$$

where

$$\overline{Z}_{n,i} = \sum_{\alpha j} u_{\alpha j}^n Z_{\alpha,ij}^*, \quad (6)$$

d_n is the degeneracy of the mode, and

$$Z_{\alpha,ij}^* = \frac{1}{e}\frac{\partial F_j}{\partial E_i}$$

are the Born effective charge tensor elements of each atom, with F the force, E an external electric field, and e the electron charge.

To facilitate comparison with experiment, Lorentzian functions of the type

$$F\left(f, f_n, \gamma, I_n\right) = \frac{I_n}{1 + \left(\frac{f - f_n}{\gamma}\right)^2}$$

were used to broaden peaks at frequencies f_n and intensities I_n. We used a broadening parameter of γ = 0.0075 THz, which is similar to the measured experimental widths. The computed spectra have been scaled with respect to the highest-intensity peak observed in experiment, for the range of frequencies studied here.

Results and Discussion

Structural analysis

Before discussing vibrational properties, we first ascertain that our computational approach is sufficiently accurate for obtaining reliable structural predictions. Crystallographic coordinates for orthorhombic and monoclinic 2,4,6-TNT, which crystallize in the space groups $Pca2_1$ and $P2_1/a$, respectively, were obtained from [42,43]. These coordinates correspond to measurements at room temperature and 100 K, respectively, and were used as the starting point for computational structural relaxation. For 1,3,5-TNB, crystallographic data of a solid with the space group symmetry of *Pbca*, measured at room temperature, were taken from [44] and subsequently relaxed. The structures are shown in Figure 1, with a comparison between the measured and computed lattice parameters given in Table 1.

Clearly, excellent agreement with experiment is obtained when the PBE+TS-vdW method is used, with residual differences between theory and experiment of the order of 1%. Notably, agreement is much less satisfactory if such interactions are not included, underscoring their importance. In the absence of vdW interactions, lattice parameters are in general too large (owing to the lack of van der Waals attraction) and errors with respect to experiment are of the order of 5–10%. These observations are fully consistent with previous studies that have compared PBE and PBE+TS-vdW predictions for geometries [25,45].

Terahertz spectra

The excellent agreement between the experimental and computed lattice parameters serves as the foundation for computing terahertz spectra. The latter, however, require not only the calculation of reliable equilibrium structures but also accurate

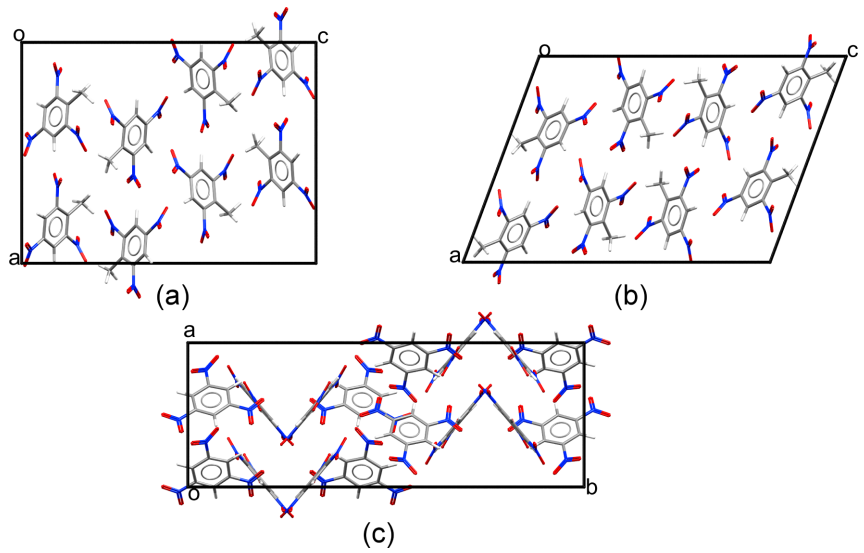


Figure 1: Structures of: (a) orthorhombic TNT, (b) monoclinic TNT, (c) orthorhombic TNB. For each structure, the direction of two of the three lattice vectors (*a*, *b*, *c*) is shown in the figure and the third one points inwards at the origin (*o*).

Table 1: Measured and computed lattice parameters for orthorhombic TNT, monoclinic TNT, and orthorhombic TNB. *a*, *b*, and *c* are lattice parameters (in Å) and β is the angle between *a* and *c*, in degrees, for the monoclinic polymorph. Numbers in parentheses indicate the relative error with respect to experiment.

		a	*b*	*c*	β
TNT orthorhombic	experimental [42]	14.99	6.08	20.02	
	PBE+TS-vdW	15.11 (0.8%)	6.06 (−0.33%)	19.94 (−0.4%)	
	PBE	15.65 (4.4%)	6.32 (3.95%)	21.87 (9.24%)	
TNT monoclinic	experimental [43]	14.91	6.03	20.88	110.37
	PBE+TS-vdW	15.13 (1.48%)	6.07 (0.66%)	21.17 (1.39%)	110.26 (−0.1%)
	PBE	15.66 (5.03%)	6.33 (4.98%)	23.27 (11.45%)	110.11 (−0.24%)
TNB orthorhombic	experimental [44]	9.78	26.94	12.82	
	PBE+TS-vdW	9.61 (−1.74%)	27.14 (0.74%)	12.82 (0.0%)	
	PBE	10.71 (9.51%)	27.73 (2.93%)	13.38 (4.37%)	

potential energy surface curvatures, placing a more severe challenge for the TS-vdW scheme used here. The computed spectra for the two TNT polymorphs, in the range from 0 to 3.3 THz, is given in Figure 2. Several interesting observations can be drawn from the figure. First, it is readily observed that inclusion of TS-vdW interactions improves agreement with experiment dramatically, also for terahertz spectroscopy. Some discrepancies remain, e.g., the theoretical group of peaks denoted in gray, starting at ca. 2.5 THz, is slightly shifted to lower frequencies compared to experiment. Nevertheless, as shown in the lower two panels of the figure, without TS-vdW interactions various vibrational modes are strongly shifted to much lower frequencies (consistent with the missing treatment of van der Waals interactions) and, furthermore, the overall spectral shape is different.

Second, the calculation indicates that the experimental spectrum contains contributions from both the monoclinic and the orthorhombic polymorphs. Perhaps the clearest examples are the peaks at ca. 1.4 THz and ca. 2.0 THz (underlined by purple and pink, respectively, in the figure), which arise from the orthorhombic polymorph. Other important peaks, for example a major feature at ca. 2.3 THz (underlined by black in the figure), arise from both polymorphs.

Next, we analyze the nature of the vibrational modes, illustrated qualitatively for selected modes in Figure 3. Our analysis shows that modes below ca. 1.5 THz are mostly dominated by intermolecular vibration, whereas modes above ca. 2.0 THz possess both inter- and intramolecular vibrational components. Specifically, the dominant intramolecular motion involves

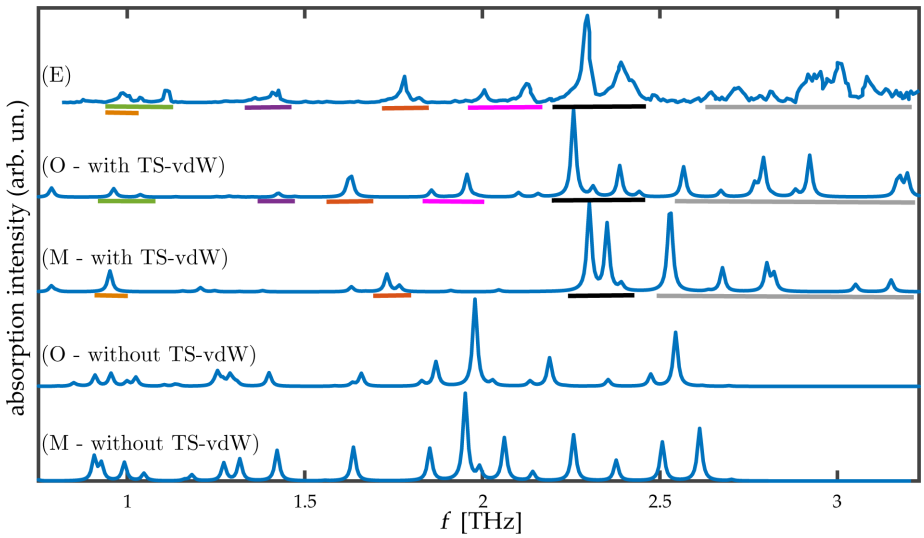


Figure 2: Absorption intensity (in arbitrary units) as a function of the frequency for the two TNT polymorphs studied in this work. (E) Experimental spectrum for the mixed-polymorph TNT crystal, taken from [16], as captured by a data analysis program. All other spectra are computed for: (O) the orthorhombic structure and (M) the monoclinic structure, calculated with and without TS-vdW interactions. Colored horizontal lines appearing below various experimental and computed peaks denote peak assignment.

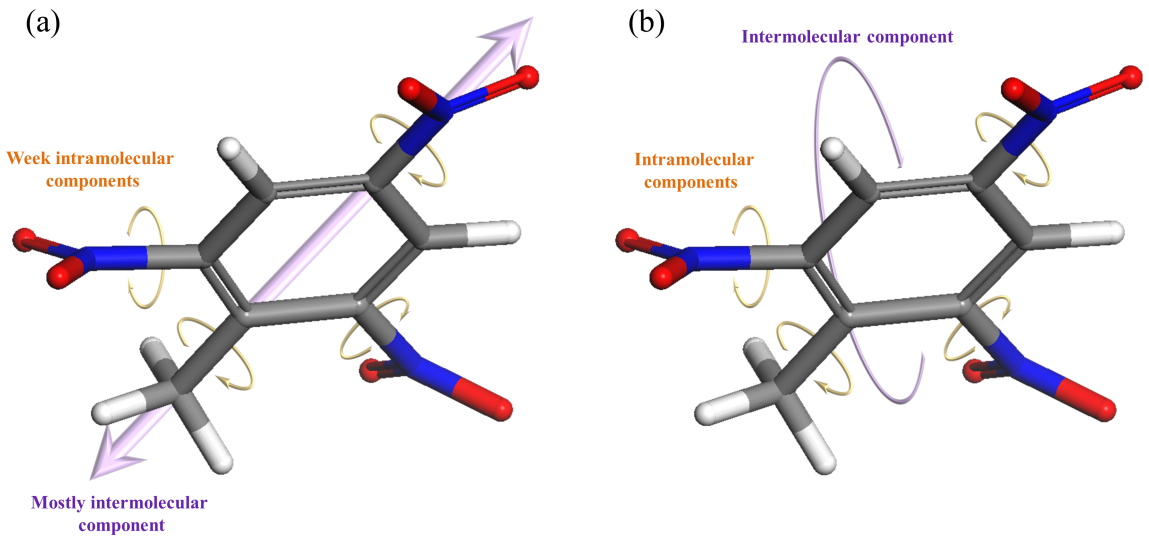


Figure 3: Schematic representation of typical TNT vibrational modes, demonstrated using modes of the orthorhombic polymorph at frequencies of (a) ca. 1.0 THz, (b) ca. 2.3 THz. Purple and orange arrows denote the direction of primary inter- and intramolecular displacement components.

torsion of nitro and methyl groups, marked by orange arrows in the figure. This is in agreement with the assignment assumed in [16], as well as with force-field calculations reported in [46]. Here, however, this assignment is obtained from first principles.

The effect of the pair-wise dispersion interactions on the vibrational modes themselves (i.e., beyond just a shift in their frequencies) can be assessed by considering the (absolute value of the) scalar product of eigenvectors of the mass-weighted Hessian matrix, with and without TS-vdW interactions. For perfectly identical modes, the multiplication should be equal to one, with the (absolute value of the) product decreasing owing to differences, down to zero for perfectly orthogonal modes. Therefore, the matrix constructed from all scalar products between modes obtained with and without TS-vdW interactions should be the identity matrix if the modes are identical. The computed matrix, arranged by mode number (in order of increasing mode frequency) for both the orthorhombic and monoclinic polymorphs, is given in Figure 4. Clearly, in the high-frequency regime (above mode 125, ca. 8 THz) the matrix ele-

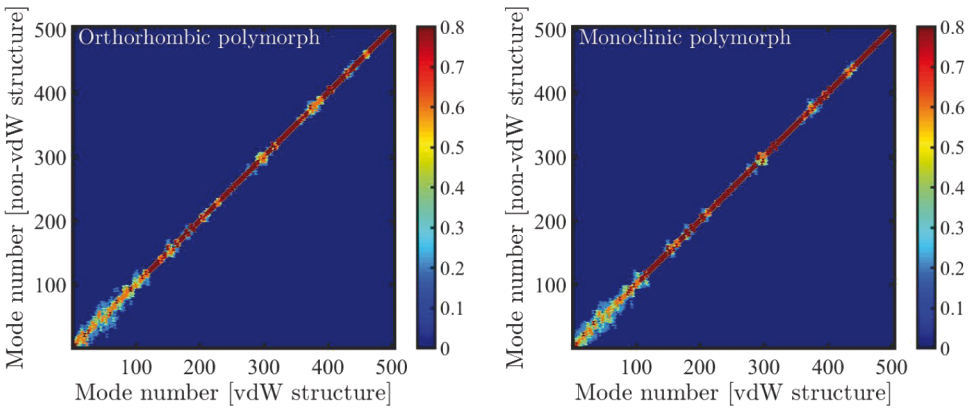


Figure 4: Absolute value of the scalar products between eigenvectors of the mass-weighted Hessian matrix, as obtained with and without TS-vdW pair-wise interactions for both the orthorhombic and monoclinic polymorphs of 2,4,6-trinitrotoluene. Mode no. 125, above which the eigenvectors are very similar with and without pair-wise interactions, is at ca. 8.1 THz for the dispersion-inclusive computation.

ments are close to those of the identity matrix, indicating that dispersion interactions have very little direct effect on the vibrational mode. This is because these modes are mostly dominated by intramolecular motions. However, in the lower-terahertz regime (below mode 125, ca. 8 THz), which is of relevance here, significant deviations from the identity matrix arise. As discussed above, in this regime there is a non-negligible (and sometimes dominant) component of intermolecular movement. We therefore conclude that dispersion interactions are important not only for determining the vibrational frequencies but also for understanding the spatial character of the vibrational modes.

As demonstrated above, semi-local DFT approximations, augmented by TS-vdW dispersion interactions, provide good agreement with experiment. Nevertheless, it would be instructive to estimate the performance of more advanced treatments going beyond the pair-wise approximation, especially as these have been shown to affect vibrational spectra in some cases [10,26,47]. We explore this issue by comparing TS-vdW with MBD calculations for both TNT polymorphs. Owing to the use of a different code for this comparison (see section "Computational Approach"), in which tight convergence is more expensive, some difference in the results is encountered already at the TS-vdW level of theory. Nevertheless, Figure 5 clearly establishes that, all other computational details being equal, the difference between TS-vdW and MBD results for the TNT crystal is relatively small, with an average vibrational frequency shift of only 0.15 THz and 0.18 THz for the orthorhombic and monoclinic polymorphs, respectively. Having ruled out

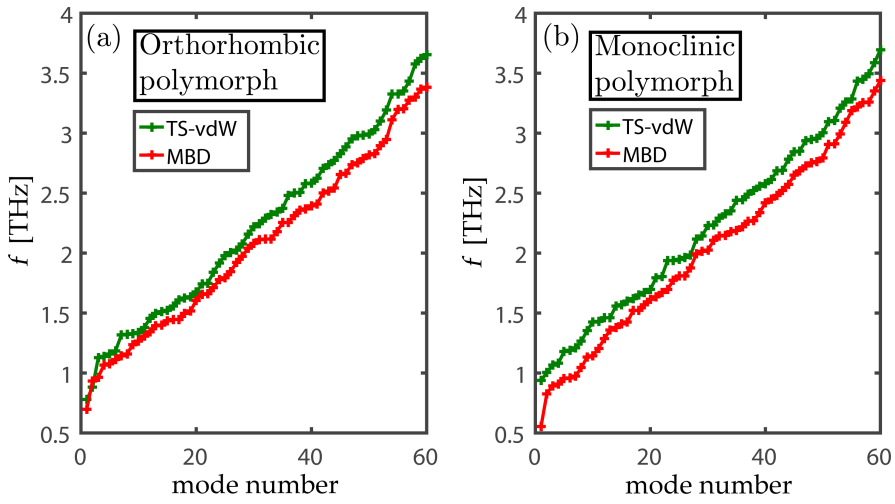


Figure 5: THz vibrational frequencies, as a function of mode number, obtained for the orthorhombic (a) and monoclinic (b) polymorphs of TNT using the TS-vdW and MBD approaches.

MBD interactions as a major issue, one can assume that the remaining theoretical limitations may arise from the underlying exchange–correlation functional itself and/or from anharmonic effects [26,48]. These issues are subject for further research.

Finally, having gained confidence in the predictive power of our approach, we consider 1,3,5-TNB (see structure in Figure 1), for which we are unaware of experimental data in the relevant terahertz range. The TNB molecule differs from TNT merely by the removal of a methyl side group. A computed terahertz spectrum is given in Figure 6, where it is compared to that of orthorombic TNT. The relatively modest chemical modification leaves clear fingerprints in the terahertz spectrum. This demonstrates the significant selectivity of terahertz spectroscopy and the importance of the ability to predict such spectra using advanced computational tools.

Conclusion

In this article, we have calculated terahertz spectra for the monoclinic and orthorhombic polymorphs of 2,4,6-TNT, using DFT both with and without Tkatchenko–Scheffler pair-wise dispersion interactions. We obtained very good agreement with experimental data upon inclusion of dispersion interactions, whereas lack of dispersion interaction causes sizable shifts in vibrational frequencies and directly affects the spatial character of the vibrational modes. The agreement between theory and experiment allowed us to distinguish between contributions of the two polymorphs to the observed spectrum. Furthermore, we could show that modes in the range from 0 to ca. 3.3 THz bear contributions from both inter- and intramolecular vibrations, with the former dominating below ca. 1.5 THz and the latter primarily involving nitro and methyl groups. Finally, we showed that the theoretical results are little affected by the inclusion many-body dispersion terms for this system, allowing us to present a prediction for the terahertz spectrum of 1,3,5-TNB and showing that a modest chemical modification may result in a markedly different terahertz spectrum.

ORCID® iDs

Ido Azuri - https://orcid.org/0000-0002-2891-1514
Anthony M. Reilly - https://orcid.org/0000-0002-1090-1086
Shai Kendler - https://orcid.org/0000-0002-3722-8522
Oded Hod - https://orcid.org/0000-0003-3790-8613
Leeor Kronik - https://orcid.org/0000-0001-6791-8658

References

1. Tonouchi, M. *Nat. Photonics* **2007,** *1,* 97–105. doi:10.1038/nphoton.2007.3
2. Jepsen, P. U.; Cooke, D. G.; Koch, M. *Laser Photonics Rev.* **2011,** *5,* 124–166. doi:10.1002/lpor.201000011
3. Kemp, M. C.; Taday, P. F.; Cole, B. E.; Cluff, J. A.; Fitzgerald, A. J.; Tribe, W. R. *Proc. SPIE* **2003,** *5070*. doi:10.1117/12.500491
4. Choi, M. K.; Bettermann, A.; van der Weide, D. W. *Philos. Trans. R. Soc., A* **2004,** *362,* 337–349. doi:10.1098/rsta.2003.1319
5. Taday, P. F. *Philos. Trans. R. Soc., A* **2004,** *362,* 351–364. doi:10.1098/rsta.2003.1321
6. Baxter, J. B.; Guglietta, G. W. *Anal. Chem.* **2011,** *83,* 4342–4368. doi:10.1021/ac200907z
7. McIntosh, A. I.; Yang, B.; Goldup, S. M.; Watkinson, M.; Donnan, R. S. *Chem. Soc. Rev.* **2012,** *41,* 2072–2082. doi:10.1039/C1CS15277G
8. Melinger, J. S.; Harsha, S. S.; Laman, N.; Grischkowsky, D. *J. Opt. Soc. Am. B* **2009,** *26,* A79–A89. doi:10.1364/JOSAB.26.000A79

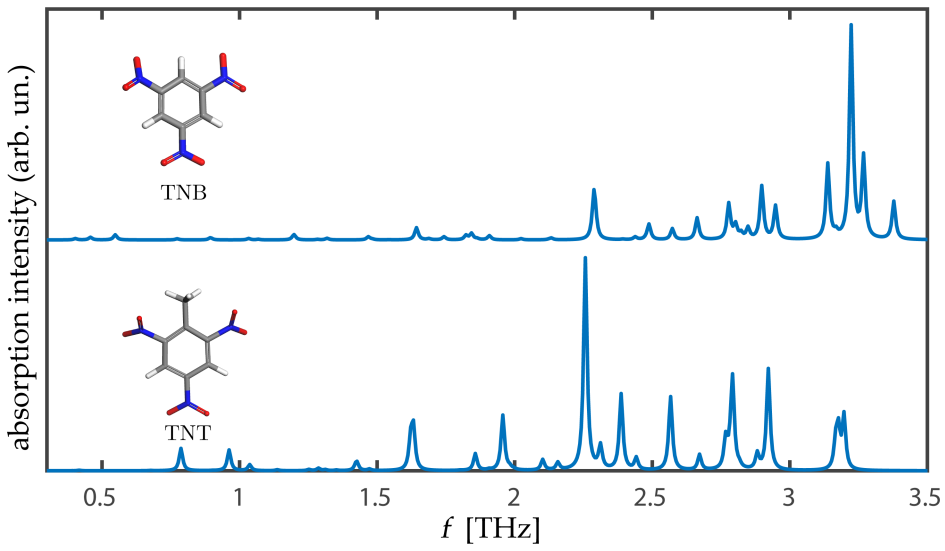


Figure 6: Comparison of computed terahertz spectra, as computed with the PBE+TS-vdW approximation, for the orthorhombic TNB and TNT crystals. Schematic views of the TNT and TNB molecules are given as insets.

9. Walther, M.; Fischer, B.; Schall, M.; Helm, H.; Jepsen, P. U. *Chem. Phys. Lett.* **2000,** *332,* 389–395. doi:10.1016/S0009-2614(00)01271-9
10. Reilly, A. M.; Tkatchenko, A. *Phys. Rev. Lett.* **2014,** *113,* 055701. doi:10.1103/PhysRevLett.113.055701
11. Taday, P. F.; Bradley, I. V.; Arnone, D. D.; Pepper, M. *J. Pharm. Sci.* **2003,** *92,* 831–838. doi:10.1002/jps.10358
12. Jördens, C.; Koch, M. *Opt. Eng. (Bellingham, WA, U. S.)* **2008,** *47,* 037003. doi:10.1117/1.2896597
13. Nagel, M.; Bolivar, P. H.; Brucherseifer, M.; Kurz, H.; Bosserhoff, A.; Büttner, R. *Appl. Phys. Lett.* **2002,** *80,* 154–156. doi:10.1063/1.1428619
14. Woodward, R. M.; Cole, B. E.; Wallace, V. P.; Pye, R. J.; Arnone, D. D.; Linfield, E. H.; Pepper, M. *Phys. Med. Biol.* **2002,** *47,* 3853. doi:10.1088/0031-9155/47/21/325
15. Hübsch, D.; Sprenger, T. Method and device for examination of mail items with regard to dangerous content. U.S. Patent US2013021416, Aug 23, 2013.
16. Melinger, J. S.; Laman, N.; Grischkowsky, D. *Appl. Phys. Lett.* **2008,** *93,* 011102. doi:10.1063/1.2949068
17. Leahy-Hoppa, M. R.; Fitch, M. J.; Zheng, X.; Hayden, L. M.; Osiander, R. *Chem. Phys. Lett.* **2007,** *434,* 227–230. doi:10.1016/j.cplett.2006.12.015
18. Hu, Y.; Huang, P.; Guo, L.; Wang, X.; Zhang, C. *Phys. Lett. A* **2006,** *359,* 728–732. doi:10.1016/j.physleta.2006.07.046
19. Chen, Y.; Liu, H.; Deng, Y.; Veksler, D. B.; Shur, M. S.; Zhang, X.-C.; Schauki, D.; Fitch, M. J.; Osiander, R.; Dodson, C.; Spicer, J. B. *Proc. SPIE* **2004,** *5411*. doi:10.1117/12.540945
20. Ruggiero, M. T.; Sibik, J.; Zeitler, J. A.; Korter, T. M. *J. Phys. Chem. A* **2016,** *120,* 7490–7495. doi:10.1021/acs.jpca.6b05702
21. Beran, G. J. O. *Chem. Rev.* **2016,** *116,* 5567–5613. doi:10.1021/acs.chemrev.5b00648
22. Bedoya-Martínez, N.; Schrode, B.; Jones, A. O. F.; Salzillo, T.; Ruzié, C.; Demitri, N.; Geerts, Y. H.; Venuti, E.; Della Valle, R. G.; Zojer, E.; Resel, R. *J. Phys. Chem. Lett.* **2017,** *8,* 3690–3695. doi:10.1021/acs.jpclett.7b01634
23. Sholl, D. S.; Steckel, J. A. *Density Functional Theory: A Practical Introduction;* Wiley: Hoboken, New Jersey, 2009.
24. Byrd, E. F. C.; Scuseria, G. E.; Chabalowski, C. F. *J. Phys. Chem. B* **2004,** *108,* 13100–13106. doi:10.1021/jp0486797
25. Kronik, L.; Tkatchenko, A. *Acc. Chem. Res.* **2014,** *47,* 3208–3216. doi:10.1021/ar500144s
26. Hoja, J.; Reilly, A. M.; Tkatchenko, A. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2017,** *7,* e1294. doi:10.1002/wcms.1294
27. Zhang, F.; Hayashi, M.; Wang, H.-W.; Tominaga, K.; Kambara, O.; Nishizawa, J.-i.; Sasaki, T. *J. Chem. Phys.* **2014,** *140,* 174509. doi:10.1063/1.4873421
28. Tkatchenko, A.; Scheffler, M. *Phys. Rev. Lett.* **2009,** *102,* 073005. doi:10.1103/PhysRevLett.102.073005
29. Perdew, J. P.; Burke, K.; Ernzerhof, M. *Phys. Rev. Lett.* **1996,** *77,* 3865–3868. doi:10.1103/PhysRevLett.77.3865
30. Jurečka, P.; Šponer, J.; Černy, J.; Hobza, P. *Phys. Chem. Chem. Phys.* **2006,** *8,* 1985–1993. doi:10.1039/B600027D
31. Tkatchenko, A.; DiStasio, R. A., Jr.; Car, R.; Scheffler, M. *Phys. Rev. Lett.* **2012,** *108,* 236402. doi:10.1103/PhysRevLett.108.236402
32. Ambrosetti, A.; Reilly, A. M.; DiStasio, R. A., Jr.; Tkatchenko, A. *J. Chem. Phys.* **2014,** *140,* 18A508. doi:10.1063/1.4865104
33. Kresse, G.; Furthmüller, J. *Comput. Mater. Sci.* **1996,** *6,* 15–50. doi:10.1016/0927-0256(96)00008-0
34. Clark, S. J.; Segall, M. D.; Pickard, C. J.; Hasnip, P. J.; Probert, M. I. J.; Refson, K.; Payne, M. C. *Z. Kristallogr.* **2005,** *220,* 567–570. doi:10.1524/zkri.220.5.567.65075
35. Monkhorst, H. J.; Pack, J. D. *Phys. Rev. B* **1976,** *13,* 5188–5192. doi:10.1103/PhysRevB.13.5188
36. Porezag, D.; Pederson, M. R. *Phys. Rev. B* **1996,** *54,* 7830–7836. doi:10.1103/PhysRevB.54.7830
37. Gonze, X.; Lee, C. *Phys. Rev. B* **1997,** *55,* 10355–10368. doi:10.1103/PhysRevB.55.10355
38. Zicovich-Wilson, C. M.; Torres, F. J.; Pascale, F.; Valenzano, L.; Orlando, R.; Dovesi, R. *J. Comput. Chem.* **2008,** *29,* 2268–2278. doi:10.1002/jcc.20993
39. Fernández-Torre, D.; Escribano, R.; Archer, T.; Pruneda, J. M.; Artacho, E. *J. Phys. Chem. A* **2004,** *108,* 10535–10541. doi:10.1021/jp047249d
40. Valenzano, L.; Noël, Y.; Orlando, R.; Zicovich-Wilson, C. M.; Ferrero, M.; Dovesi, R. *Theor. Chem. Acc.* **2007,** *117,* 991–1000. doi:10.1007/s00214-006-0213-2
41. Zhen-Long, L.; Jing-Han, Y.; Yuan-Yuan, Z.; Hui, W. *Commun. Theor. Phys.* **2011,** *55,* 513. doi:10.1088/0253-6102/55/3/26
42. Carper, W. R.; Davis, L. P.; Extine, M. W. *J. Phys. Chem.* **1982,** *86,* 459–462. doi:10.1021/j100393a009
43. Vrcelj, R. M.; Sherwood, J. N.; Kennedy, A. R.; Gallagher, H. G.; Gelbrich, T. *Cryst. Growth Des.* **2003,** *3,* 1027–1032. doi:10.1021/cg0340704
44. Choi, C. S.; Abel, J. E. *Acta Crystallogr., Sect. B* **1972,** *28,* 193–201. doi:10.1107/S0567740872002067
45. Al-Saidi, W. A.; Voora, V. K.; Jordan, K. D. *J. Chem. Theory Comput.* **2012,** *8,* 1503–1513. doi:10.1021/ct200618b
46. Shi, L.; Duan, X.-H.; Zhu, L.-G.; Liu, X.; Pei, C.-H. *J. Phys. Chem. A* **2016,** *120,* 1160–1167. doi:10.1021/acs.jpca.5b10782
47. Hoja, J.; Neumann, M.; Tkatchenko, A. to be published.
48. Katz, G.; Zybin, S.; Goddard, W. A., III; Zeiri, Y.; Kosloff, R. *J. Phys. Chem. Lett.* **2014,** *5,* 772–776. doi:10.1021/jz402801m

License and Terms

BEILSTEIN JOURNAL OF ORGANIC CHEMISTRY

Local energy decomposition analysis of hydrogen-bonded dimers within a domain-based pair natural orbital coupled cluster study

Ahmet Altun[1,2], Frank Neese[*1,2] and Giovanni Bistoni[*1,2]

Full Research Paper

Open Access

Address:
[1]Max Planck Institute for Chemical Energy Conversion, Stifstrasse 34-36, D-45470 Mülheim an der Ruhr, Germany and [2]Max Planck Institute for Coal Research, Kaiser-Wilhelm-Platz 1, D-45470 Mülheim an der Ruhr, Germany

Email:
Frank Neese[*] - Frank.Neese@kofo.mpg.de; Giovanni Bistoni[*] - giovanni.bistoni@kofo.mpg.de

* Corresponding author

Keywords:
DLPNO-CCSD(T); hydrogen-bond interaction; interaction energy; local energy decomposition; London dispersion

Beilstein J. Org. Chem. **2018,** *14,* 919–929.
doi:10.3762/bjoc.14.79

Received: 19 January 2018
Accepted: 06 April 2018
Published: 25 April 2018

This article is part of the Thematic Series "Dispersion interactions" and is dedicated to the memory of Keiji Morokuma, who passed away in December 2017, in recognition of his pioneering contributions to energy decomposition analysis schemes.

Guest Editor: P. Schreiner

Abstract

The local energy decomposition (LED) analysis allows for a decomposition of the accurate domain-based local pair natural orbital CCSD(T) [DLPNO-CCSD(T)] energy into physically meaningful contributions including geometric and electronic preparation, electrostatic interaction, interfragment exchange, dynamic charge polarization, and London dispersion terms. Herein, this technique is employed in the study of hydrogen-bonding interactions in a series of conformers of water and hydrogen fluoride dimers. Initially, DLPNO-CCSD(T) dissociation energies for the most stable conformers are computed and compared with available experimental data. Afterwards, the decay of the LED terms with the intermolecular distance (r) is discussed and results are compared with the ones obtained from the popular symmetry adapted perturbation theory (SAPT). It is found that, as expected, electrostatic contributions slowly decay for increasing r and dominate the interaction energies in the long range. London dispersion contributions decay as expected, as r^{-6}. They significantly affect the depths of the potential wells. The interfragment exchange provides a further stabilizing contribution that decays exponentially with the intermolecular distance. This information is used to rationalize the trend of stability of various conformers of the water and hydrogen fluoride dimers.

Introduction

Hydrogen bonds are of fundamental importance for regulating molecular properties like polarizability [1] and in various biochemical processes, including protein folding [2] and stability [3], replication of DNA and RNA [4], enzyme catalysis [5], proton relay mechanism [6], and drug delivery [7].

Energy decomposition analysis (EDA) schemes have been instrumental in providing insights into the nature of these interactions, by partitioning the total interaction energy of two (or more) interacting fragments into several chemically meaningful contributions [8-10]. EDA methods are mainly based on an

early variational study of Morokuma [11]. They are typically carried out at the Hartree–Fock (*HF*) or density functional theory (DFT) level. In these schemes, the interacting system is treated as a supermolecule and the overall interaction energy is decomposed into various terms such as electrostatic interaction, charge transfer, polarization, and the so-called Pauli or exchange-repulsion terms [12-16].

Instead of decomposing DFT or *HF* interaction energies, the widely used symmetry-adapted perturbation theory (SAPT) [17] provides a perturbative expansion of the interaction energy based on the wave functions of the monomers. For weakly interacting monomers, this approach permits to obtain accurate interaction energies as well as their constituting electrostatic, induction, dispersion, and exchange-repulsion terms [9,10].

Although these schemes provide different quantitative estimates for the important components of the interaction, they also provide useful interpretative frameworks in which to discuss experimental observables. For instance, they can be used for discussing trends of dissociation energies [8,9] or the relative stability of conformers [11,15,18,19]. However, two fundamental aspects must be considered when an EDA scheme is applied to a specific chemical problem. The chosen approach must provide: (i) a sufficiently accurate estimate for the observables of interest, which are typically relative energies; (ii) a useful decomposition of the observable into a series of chemically meaningful terms representing the correct physics in the asymptotic region.

In order to address the first issue, the coupled-cluster method with single, double, and perturbative treatment of triple excitations [CCSD(T)] has proven its reliability in a wide range of contexts. This method typically allows for the calculation of relative energies with chemical accuracy (1 kcal/mol) [20-22]. Moreover, our group has recently developed the domain-based local pair natural orbital CCSD(T) method [DLPNO-CCSD(T)] [23-30], which scales linearly with system size and typically provides around 99.9% of the canonical CCSD(T) correlation energy if TightPNO settings are used [31,32]. Thus, DLPNO-CCSD(T) single-point energies can now be obtained for systems with hundreds of atoms and thousands of basis functions while essentially retaining the accuracy and reliability of canonical CCSD(T).

However, the CCSD wave function is a highly complex object that is nonlinear in its parameters (cluster amplitudes). Hence, its direct physical interpretation is not immediately apparent. In order to facilitate the interpretation of DLPNO-CCSD(T) results, we have thus recently introduced the local energy decomposition (LED) analysis scheme, which decomposes the DLPNO-CCSD(T) interaction energy of two or more molecules in terms of electronic and geometric preparation, electrostatic interaction, interfragment exchange, dynamic charge polarization, and London dispersion terms [33].

Herein, the DLPNO-CCSD(T)/LED methodology is applied to the study of H-bond interactions in a series of conformers of water (H_2O) and hydrogen fluoride (HF) dimers, which are shown in Figure 1.

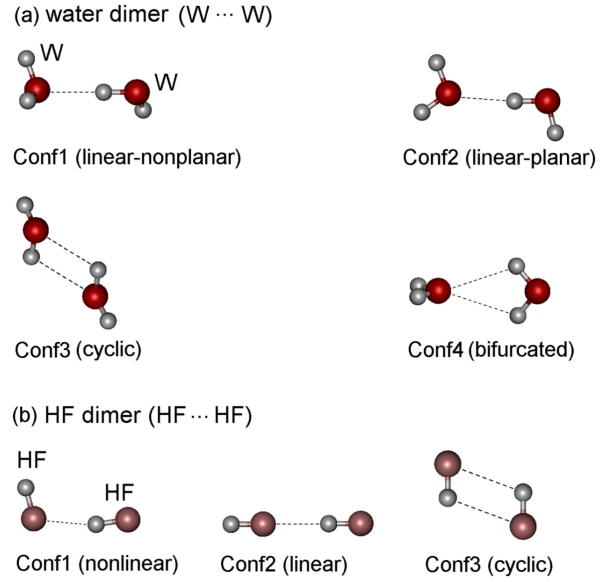


Figure 1: The conformers of (a) water dimer and (b) HF dimer.

These systems are representative examples of H-bond interactions and are often used as model systems for newly developed methods, including EDA schemes [18,34-41]. Although these dimers have been studied extensively, the principal mechanisms of interaction between their constituting monomers are still under debate. The debate concerns the magnitude of individual terms and the importance of London dispersion, charge transfer and polarization effects compared with the dominating electrostatic interaction [15,16,18,19]. Herein, particular emphasis is given in discussing the role played by London dispersion, which constitutes the attractive part of the van der Waals potential and has long been considered a weak effect compared to the other components of the interaction. However, in recent years, several studies have demonstrated that this component of the interaction plays a fundamental role in controlling the stability and reactivity of a wide range of systems [42,43].

This paper is organized as follows. Following a description of the computational details, computed geometries and dissociation energies are compared with previously published experi-

mental and computational data. In the following section, the decay of LED terms with the intermolecular distance between the monomers is discussed for the water dimer case, and results are compared with those obtained from SAPT. This information is then used to rationalize the trends in stability of various conformers of the water and hydrogen fluoride dimers. The last section is devoted to the discussion of the results and concluding remarks.

Computational Details and Theoretical Aspects

Computational details

All DLPNO-CCSD(T) calculations and geometry optimizations were performed with a development version of the ORCA 4.0 suite of programs [44,45].

Geometry optimizations and relaxed PES scans constraining only the reaction coordinates were carried out at the RI-MP2 level, employing aug-cc-pVTZ basis set with matching auxiliary basis sets [46-49]. The RIJK approach applying RI approximation for both Coulomb *J* and exchange *K* parts was used [50,51]. Harmonic vibrational frequencies and zero-point energy (ZPE) corrections were computed with the same level of theory used for the geometry optimizations. All valence electrons were included in the correlation treatment (only the core 1s orbitals of oxygen and fluorine atoms were frozen).

Single point DLPNO-CCSD(T) energies and LED calculations employed the Foster–Boys scheme [52] for the localization of the occupied orbitals. All valence electrons were included in the correlation treatment. "TightPNO" settings were used [31,32]. All electron pairs were included in the coupled cluster treatment. The RIJK approximation was used in the *HF* part. The Pipek–Mezey [53] orbital localization scheme was applied for the localization of the PNOs in the LED scheme. In all cases, augmented correlation consistent basis sets of triple-ζ (aug-cc-pVTZ) and quadruple-ζ (aug-cc-pVQZ) qualities were used in conjunction with matching auxiliary basis sets [46-49]. DLPNO-CCSD(T) energies were first corrected for the basis set superposition error (BSSE) [54] and then extrapolated to the complete basis set (CBS) limit using a two-point scheme [55] based on Equation 1 and Equation 2.

$$E_{HF}^{\mathrm{CBS}} = \frac{E_{HF}^{(n)} \cdot e^{-\alpha\sqrt{m}} - E_{HF}^{(m)} \cdot e^{-\alpha\sqrt{n}}}{e^{-\alpha\sqrt{m}} - e^{-\alpha\sqrt{n}}} \quad (1)$$

$$E_{\mathrm{C}}^{\mathrm{CBS}} = \frac{n^{\beta} E_{\mathrm{corr}}^{(n)} - m^{\beta} E_{\mathrm{corr}}^{(m)}}{n^{\beta} - m^{\beta}} \quad (2)$$

where $E^{(n)}$ and $E^{(m)}$ are the energies obtained with a basis set of n–ζ and m–ζ cardinality (here n = 3 and m = 4), respectively. The previously calibrated values [56] of the constants (α = 5.46, and β = 3.05) for the 3/4 extrapolation were used. The individual LED contributions were also extrapolated.

It is worth mentioning that BSSE-corrected and -uncorrected interaction energies converge to the same value upon extrapolation within 0.1 kcal/mol. For completeness, all energies are reported in Supporting Information File 1.

SAPT calculations were carried out with the MOLPRO [57] program package (version 2012.1) using RI-MP2 geometries. The nondispersive terms of density functional-based SAPT (DFT-SAPT) converge quickly with the basis set size and do not require CBS extrapolation [58,59]. Thus, only the dispersion terms of the DFT-SAPT energies have been extrapolated to CBS limit through Equation 2 using β = 3.05.

DFT-SAPT calculations were carried out with the asymptotically-corrected exchange-correlation functional PBE0AC [60,61]. This functional is a modified PBE0 hybrid functional in which the long-range tail contains 75% of LB94 exchange. The shift parameter applied for the bulk potential within this correction was calculated as the sum of the ionization potential and highest occupied molecular orbital (HOMO) energy of each fragment optimized in the gas phase. The experimentally determined ionization potential of an isolated water molecule was used (0.4638 E_{h} [62,63]).

LED analysis in the DLPNO-CCSD(T) framework

The theory and implementation of the DLPNO-CCSD(T) method and of the LED scheme have been described in detail in a series of recent publications [23-33]. We thus only recall here the main features of this technique.

Within a supramolecular approach, the energy of a molecular adduct *XY* relative to the total energies of noninteracting fragments *X* and *Y*, i.e, dimerization energy (Δ*E*), can be written as:

$$\Delta E = \Delta E_{\mathrm{geo-prep}} + \Delta E_{\mathrm{int}} \quad (3)$$

where $\Delta E_{\mathrm{geo-prep}}$ is the geometric preparation energy needed to distort the fragments *X* and *Y* from their structures at infinite separation to their in-adduct geometry. ΔE_{int} is the interaction energy of the fragments *X* and *Y* at a given geometry of the adduct *XY*.

ΔE_{int} can be decomposed into an *HF* contribution $\Delta E_{\mathrm{int}}^{HF}$ and a correlation contribution $\Delta E_{\mathrm{int}}^{\mathrm{C}}$:

$$\Delta E_{\text{int}} = \Delta E_{\text{int}}^{HF} + \Delta E_{\text{int}}^{\text{C}} \quad (4)$$

By exploiting the localization of the occupied orbitals in the DLPNO-CCSD(T) framework, the $\Delta E_{\text{int}}^{HF}$ is then decomposed into three contributions [33,64]:

$$\Delta E_{\text{int}}^{HF} = \Delta E_{\text{el-prep}}^{HF} + E_{\text{elstat}} + E_{\text{exch}} \quad (5)$$

The electronic preparation $\Delta E_{\text{el-prep}}^{HF}$ is positive and thus repulsive. It corresponds to the energy needed to bring the electronic structures of the isolated fragments into the one optimal for the interaction. E_{elstat} and E_{exch} are the electrostatic and exchange interactions, respectively, between the interacting fragments. It is worth noting here that the intermolecular exchange describes a stabilizing component of the interaction, lowering the repulsion between electrons of the same spin.

The DLPNO-CCSD(T) correlation energy (E_{C}) can be written as a sum of electron-pair correlation energy (ε_{ij}, where i and j denote the localized orbitals) contributions plus a perturbative triples correction ($E_{\text{C-(T)}}$). Local second-order many-body perturbation theory is used to divide the ε_{ij} terms into "weak pairs", with expected negligible contribution to the correlation energy, and "strong pairs". The contribution coming from the weak pairs is kept at the second-order level, whereas the strong pairs are treated at the coupled cluster level. Hence, the overall correlation energy reads [30]:

$$E_{\text{C}} = E_{\text{C-SP}} + E_{\text{C-WP}} + E_{\text{C-(T)}} \quad (6)$$

where $E_{\text{C-SP}}$ and $E_{\text{C-WP}}$ are the strong-pairs and weak-pairs components of the correlation energy, respectively. The correlation contribution to the interaction energy $\Delta E_{\text{int}}^{\text{C}}$ can thus be expressed as a sum of three contributions:

$$\Delta E_{\text{int}}^{\text{C}} = \Delta E_{\text{int}}^{\text{C-SP}} + \Delta E_{\text{int}}^{\text{C-WP}} + \Delta E_{\text{int}}^{\text{C-(T)}} \quad (7)$$

in which $\Delta E_{\text{int}}^{\text{C-SP}}$, $\Delta E_{\text{int}}^{\text{C-WP}}$, and $\Delta E_{\text{int}}^{\text{C-(T)}}$ are the strong pairs, weak pairs and triples correction components of the correlation contribution to the interaction energy, respectively.

The $\Delta E_{\text{int}}^{\text{C-WP}}$ and $\Delta E_{\text{int}}^{\text{C-(T)}}$ terms can be further divided into electronic preparation and interfragment interaction based on the localization of the occupied orbitals [30]. However, these terms are very small for the systems studied in this work and thus are not decomposed herein.

For the dominant strong pairs contribution $\Delta E_{\text{int}}^{\text{C-SP}}$, the decomposition exploits the localization of both the occupied and the virtual orbitals in the DLPNO-CCSD(T) framework. Hence, the $\Delta E_{\text{int}}^{\text{C-SP}}$ term is divided into three contributions: the electronic preparation energy $\Delta E_{\text{el-prep}}^{\text{C-SP}}$, the charge transfer or charge polarization contribution ($\Delta E_{\text{CT}}^{\text{C-SP}}$), and London dispersion ($\Delta E_{\text{DISP}}^{\text{C-SP}}$).

$$\Delta E_{\text{int}}^{\text{C-SP}} = \Delta E_{\text{el-prep}}^{\text{C-SP}} + E_{\text{CT}}^{\text{C-SP}} + E_{\text{DISP}}^{\text{C-SP}} \quad (8)$$

The relevant pair excitation contributions constituting these terms are shown pictorially in Figure 2.

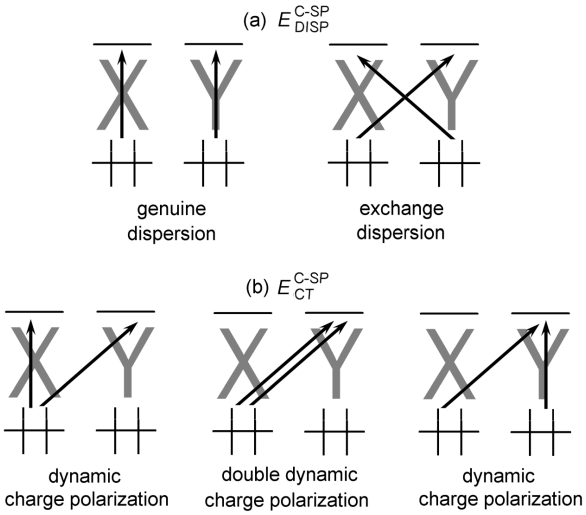


Figure 2: Schematic representation of strong pair excitations in the framework of the DLPNO-CCSD(T) method. Electronic preparation arises from excitations occurring within the same fragment, which are not shown. Only the charge transfer excitations from *X* to *Y* are shown. Analogous charge transfer excitations also exist from *Y* to *X*.

It may be useful to combine several terms depending on the molecular system of interest. For example, $\Delta E_{\text{el-prep}}^{\text{C-SP}}$ and $E_{\text{CT}}^{\text{C-SP}}$ have opposite signs and typically compensate each other [33,64]. Hence, these two terms can be combined to give the SP contribution to the interaction energy excluding dispersion contribution ($E_{\text{res}}^{\text{C-SP}}$):

$$E_{\text{int}}^{\text{C-SP}} = \left(\Delta E_{\text{el-prep}}^{\text{C-SP}} + E_{\text{CT}}^{\text{C-SP}}\right) + E_{\text{DISP}}^{\text{C-SP}} = E_{\text{res}}^{\text{C-SP}} + E_{\text{DISP}}^{\text{C-SP}} \quad (9)$$

As a final remark, it is worth underscoring that one of the aims of this paper is to discuss the decay of the different components of $\Delta E_{\text{int}}^{\text{C-SP}}$ with the distance between the interacting fragments. Hence, we decided to include all electron pairs in the coupled cluster treatment. In this case, the weak-pair contribution only

corrects for the pair natural orbital (PNO) truncation and only marginally affects the overall correlation energy. The latter is thus dominated by the strong pairs irrespective of the distance between the fragments.

SAPT analysis

Symmetry-adapted perturbation theory (SAPT) is a well-established method for the calculation of interaction energies in the context of weak intermolecular interactions [65,66]. It expresses the interaction energy in various terms in the perturbation series that are physically meaningful. In this work, the terms of SAPT are compared with those from our LED scheme.

In SAPT, the non-dispersive interaction energy includes the first order perturbative terms polarization (E_{pol}) and exchange-repulsion (E_{exch}), and the second order terms induction (E_{ind}) and exchange-induction ($E_{exch-ind}$). In the DFT variant of SAPT (DFT-SAPT), the difference between the *HF* interaction energy and the sum of the non-dispersive HF-SAPT terms (denoted as δ(*HF*)) is also included in the nondispersive interaction energy to approximately account for the effect of the higher order terms.

$$E_{no-disp} = E_{pol} + E_{exch} + E_{ind} + E_{exch-ind} + \delta(HF) \quad (10)$$

The dispersive energy (E_{disp}) of the DFT-SAPT includes both the genuine dispersion and its exchange correction, which are calculated both at the second order: the sum of the nondispersive ($E_{no-disp}$) and dispersive (E_{disp}) terms is the DFT-SAPT interaction energy.

$$\Delta E_{int} = E_{no-disp} + E_{disp} \quad (11)$$

Results and Discussion

Geometries and dissociation energies: comparison with experiment

In this section, the computed geometries and dissociation energies for the water and HF dimers in their global minimum (Conf1 of Figure 1) are compared with available experimental data and previously published computational predictions.

For the water dimer in its global minimum, experiments estimate an r_0(O···O) distance between 2.946 and 2.976 Å [67]. The CCSD(T)-based best estimate of r_e(O···O) reported in literature is 2.912 ± 0.005 Å [68]. The r_e(O···O) distance calculated in this work at the RI-MP2/aug-cc-pVTZ level (2.908 Å) is very close to the CCSD(T) prediction. The effect of excitations beyond CCSD(T) has been shown to be negligible by means of CCSDTQ calculations [69].

For the HF dimer, the r_e(F···F) distance of Conf1 calculated at the RI-MP2/aug-cc-pVTZ level (2.746 Å) agrees reasonably well with its previous best estimate (2.735 ± 0.010 Å) on a theoretical potential energy surface refined by comparing multidimensional nuclear quantum energy levels with the corresponding experimental data [70] and the CCSD(T) result of 2.737 Å with a quintuple−ζ basis [71].

The equilibrium ΔE_e and zero-point corrected ΔE_0 dimerization energies of water and HF dimers are given in Table 1. These correspond to the equilibrium D_e and zero-point D_0 dissociation energies with opposite sign, respectively.

The ΔE_e value of the water dimer calculated previously at the CCSD(T)/CBS level (−5.01 kcal/mol [68]) agrees remarkably well with the present DLPNO-CCSD(T) result (ΔE_e = −4.95 kcal/mol, Table 1). The accurate calculation of ZPE correction of H-bonded systems requires larger basis sets and the inclusion of anharmonic effects [68,71]. Thus, the present RI-MP2/aug-cc-pVTZ harmonic ZPE contribution (2.13 kcal/mol) is slightly larger than the experimental value of 1.72 kcal/mol [71]. Using the experimentally determined ZPE contribution for correcting the DLPNO-CCSD(T) ΔE_e value, one obtains a ΔE_0 value of −3.23 kcal/mol, which is very close to the experimental value of −3.16 ± 0.03 kcal/mol [72,73].

For the HF dimer, the present DLPNO-CCSD(T) values of ΔE_e (−4.511 kcal/mol) and harmonic ΔE_e (−2.694 kcal/mol) are consistent with the previously calculated ΔE_e (−4.580 ± 0.004 kcal/mol) and harmonic ΔE_0 (−2.775 ± 0.024 kcal/mol) values at the CCSD(T)/CBS level [74]. It was shown [74] that the effects of quadruple excitations Q (−0.008 ± 0.004 kcal/mol), relativity (0.016 ± 0.001 kcal/mol), and the diagonal Born–Oppenheimer correction (−0.012 ± 0.000 kcal/mol) to the dimerization energy of the HF dimer are negligible while the anharmonic ZPE contribution (−0.185 ± 0.019 kcal/mol) is significant. Adding these corrections to the calculated ΔE_e energies, the best fully theoretical estimates of ΔE_0 become −2.964 ± 0.047 and −2.883 kcal/mol at the CCSD(T)/CBS [74] and the present DLPNO-CCSD(T) levels, respectively. These results are similar to that calculated on a potential [70] refined by using experimental data (−3.036 ± 0.003 kcal/mol).

The consistency of the present DLPNO-CCSD(T) and the previous experimental or CCSD(T) dimerization energies of the water and HF dimers indicates that the present computational level can be reliably applied to investigation of dissociation and interaction energies of other H-bonded molecules.

Table 1: The DLPNO-CCSD(T) dimerization energies (kcal/mol) of the conformers of water and HF dimers together with the individual LED terms.

	water dimer				HF dimer		
	Conf1	Conf2	Conf3	Conf4	Conf1	Conf2	Conf3
ΔE_e	−4.95	−4.38	−4.15	−3.16	−4.51	−3.56	−3.52
ΔE_o	−2.82[a]	−2.86	−2.22	−1.80	−2.69[b]	−2.45	−2.07
decomposition of ΔE							
$\Delta E_{geo-prep}$	0.07	0.05	0.05	0.15	0.11	0.04	0.09
ΔE_{int}	−5.01	−4.43	−4.20	−3.31	−4.62	−3.60	−3.61
decomposition of ΔE_{int}^{HF}							
ΔE_{int}^{HF}	−3.67	−3.30	−2.70	−2.51	−3.89	−3.33	−2.74
$\Delta E_{el-prep}^{HF}$	22.91	18.33	16.52	8.74	20.52	10.99	13.94
E_{elstat}	−22.83	−18.60	−16.43	−9.75	−21.22	−12.56	−14.47
E_{exch}	−3.76	−3.03	−2.79	−1.50	−3.19	−2.22	−2.22
decomposition of ΔE_{int}^{C-SP}							
ΔE_{res}^{C-SP}	0.19	0.23	0.16	0.26	0.42	0.64	0.31
ΔE_{DISP}^{C-SP}	−1.24	−1.10	−1.36	−0.89	−0.94	−0.80	−0.96
WP and triple corrections							
ΔE_{int}^{C-WP}	−0.08	−0.08	−0.08	−0.06	−0.08	−0.07	−0.08
$\Delta E_{int}^{C-(T)}$	−0.22	−0.17	−0.22	−0.11	−0.13	−0.05	−0.13

[a]Experiment: −3.16 ± 0.03 kcal/mol [72,73]. When the experimentally determined ZPE (1.72 kcal/mol [71]) is used, the resulting value (−3.23 kcal/mol) is very close to the experiment. [b]When the effect of the anharmonicity of the vibrational energy levels estimated to be −0.185 ± 0.019 kcal/mol [74] is included, the resulting value (−2.88 kcal/mol) agrees reasonably well with the value found on an empirical potential (−3.036 ± 0.003 kcal/mol) [70].

Decay of LED terms with the intermolecular distance

In this section, the decay of the LED terms with the intermolecular distance (r) is discussed for the water dimer. However, the derived conclusions are rather general and thus hold true for the HF dimer as well, as shown in Supporting Information File 1. From now on, we use the term "short-range" to indicate the region where $r_{O\cdots H} \leq 3.5$ Å, and "long-range" for the region where $r_{O\cdots H} > 3.5$ Å.

Let us start with the analysis of the DLPNO-CCSD(T) energy profile (Figure 3) for the dissociation of the water dimer as a function of the H-bond distance $r_{O\cdots H}$. The corresponding *HF* and DLPNO-CCSD profiles are also reported for comparison.

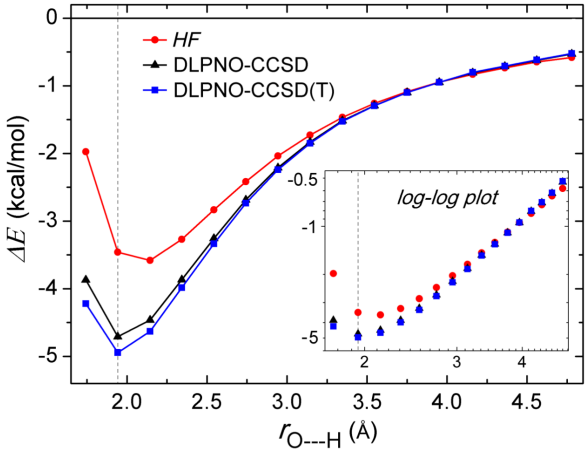


Figure 3: Dissociation curve of Conf1 of water dimer as a function of the H-bond distance. Its nearly linear relation in the log–log scale for the long range is shown as insert on the graph. The black dotted vertical line at 1.943 Å corresponds to the equilibrium r_e(O---H) distance.

In the long range, the *HF* and coupled cluster energies show smooth polynomial decays, which are evident from their linear log–log relation shown in the insert of Figure 3. In this range, the correlation contribution to the interaction energy is small and positive and the overall interaction is dominated by the *HF* term. Conversely, in the short range, the correlation energy becomes a significant stabilizing component of the interaction. At the equilibrium position, correlation contributes to the interaction energy of the water dimer by −1.34 kcal/mol. Interestingly, the effect of the perturbative triples (T) is small for all distances.

A deeper insight into the nature of the water dimer interaction comes by decomposing both the *HF* and correlation component of the dissociation energy into their LED components. Let us start by discussing the *HF* contributions, reported in the upper panel of Figure 4.

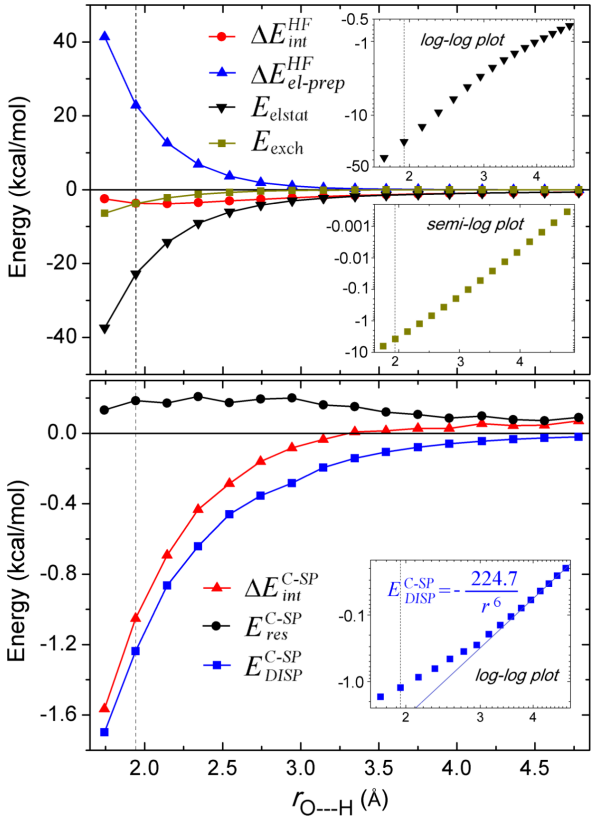


Figure 4: Decomposed *HF* energy terms (top), and correlation energy terms (bottom) of Conf1 of water dimer as a function of the H-bond distance. The nearly linear relation of the long range electrostatic and London dispersion energy terms in the log–log scale as well as the exchange energies in the semi-log scale are given as inserts on the graphs. The black dotted vertical line at 1.943 Å corresponds to the equilibrium r_e(O---H) distance.

In the long range, the only significant LED term is the electrostatic energy, which shows a slow polynomial decay with the distance (indicated by the linear relation in the log–log plot shown in the insert of Figure 4 top). This is not surprising considering the strong dipole of water. However, in the short range, the repulsive electronic preparation arising from distortion of the electronic clouds of the interacting monomers assumes large values and almost entirely counteracts the electrostatic contribution at the equilibrium position. In this position, the remaining *HF* term, i.e., the attractive exchange interaction, amounts to −3.76 kcal/mol, which is very close to the overall *HF* contribution to the interaction energy (−3.67 kcal/mol), and thus provides a fundamental stabilizing component. As expected, the exchange term decays exponentially with intermolecular distance, which is indicated by the linear relation of the semi-log plot shown in the insert of Figure 4 top.

In the correlation part, the weak pair correction is very small in the whole distance range and amounts to −0.08 kcal/mol at the equilibrium position. Hence, only the dominant strong pair contribution is decomposed in the following for the sake of simplicity (lower panel of Figure 4). In the long range, the sum of the dynamic charge polarization and electronic preparation energies (E_{res}^{C-SP}) is always positive and provides a small correction to the electrostatics computed at the *HF* level, which is known to overestimate the dipole of water [75]. The remaining correlation term, i.e., London dispersion, is always attractive and decays with r^{-6} in the long range, as shown in the log–log insert in the lower panel of Figure 4.

The E_{res}^{C-SP} term is dominant in the long range, which makes the overall correlation contribution to the interaction energy positive. However, it reaches just 0.19 kcal/mol at the equilibrium, being much smaller than the corresponding dispersion term (−1.24 kcal/mol). Therefore, the short-range correlation behavior is largely dominated by the London dispersion. The slight fluctuations of the correlation energy terms (Figure 4, bottom panel) arise mainly from difficulties in localizing the PNOs [33].

The r^{-6} behavior of the LED estimate of the London dispersion contribution deserves to be discussed in more detail. In order to do that, it is useful to look at the expression for the strong pair correlation energy in the DLPNO-CCSD(T) method [33]:

$$E_{C-SP} = \sum_{i \geq j} \sum_{\tilde{a}_{ij}\tilde{b}_{ij}} \frac{4\tau_{\tilde{a}_{ij}\tilde{b}_{ij}}^{ij} - 2\tau_{\tilde{a}_{ij}\tilde{b}_{ij}}^{ji}}{1+\delta_{ij}} (i\tilde{a}_{ij} \mid j\tilde{b}_{ij}) \qquad (12)$$

in which $\tilde{a}_{ij}$ and $\tilde{b}_{ij}$ are PNOs that belong to pair ij, $(i\tilde{a}_{ij} \mid j\tilde{b}_{ij})$ represents the two electron integrals in Mulliken notation, and $\tau_{\tilde{a}_{ij}\tilde{b}_{ij}}^{ij}$ is defined as

$$\tau_{\tilde{a}_{ij}\tilde{b}_{ij}}^{ij} = t_{\tilde{a}_{ij}}^{i} t_{\tilde{b}_{ij}}^{j} + t_{\tilde{a}_{ij}\tilde{b}_{ij}}^{ij} \qquad (13)$$

in which $t_{\tilde{a}_{ij}}^{i}$ and $t_{\tilde{b}_{ij}}^{j}$ are the singles amplitudes and $t_{\tilde{a}_{ij}\tilde{b}_{ij}}^{ij}$ are the corresponding doubles amplitudes. From a multipole expansion of the integrals, it follows that the $(i\tilde{a}_{ij} \mid j\tilde{b}_{ij})$ terms decay with r_{ij}^{-3}, where r_{ij} is the separation between the charge centroids of the local occupied orbitals ϕ_i and ϕ_j. The $t_{\tilde{a}_{ij}\tilde{b}_{ij}}^{ij}$ amplitudes also

decay with r_{ij}^{-3}, whilst the $t^{i}_{\tilde{a}_{ij}}\, t^{j}_{\tilde{b}_{ij}}$ terms have no explicit dependence on the distance between the centroids of ϕ_i and ϕ_j. In the LED definition of London dispersion, only the terms of Equation 12, in which ϕ_i and ϕ_j are assigned to different fragments, are included (see Figure 2). Hence, the overall London dispersion consists of two terms, one decaying with r^{-3} (due to singles) and the other with r^{-6} (due to doubles). Although one could argue that these two terms have different physical meanings, the contribution of the singles to the London dispersion is typically negligible. At the equilibrium distance, it amounts to the 0.34% of the overall London dispersion contribution (see Supporting Information File 1).

Comparison with DFT-SAPT

DFT-SAPT treats the interaction energy as a perturbation on the isolated fragments. Hence, the terms of the DFT-SAPT interaction energy are difficult to compare with the ones from energy decomposition schemes based on a supramolecular approach, e.g., the LED. Despite these differences, it is still interesting to compare whether both approaches lead to a similar partitioning of dispersion and electrostatics, in order to draw connections between different interpretative frameworks. A comparison of total interaction, electrostatic, and London dispersion energies calculated with DLPNO-CCSD(T)/LED and DFT-SAPT for the water dimer at various intermolecular distances is reported in Figure 5 (see Supporting Information File 1 for the individual data).

Let us start by discussing the behavior of the total interaction energy (Figure 5 upper panel). At the equilibrium geometry, DFT-SAPT underestimates the interaction energy by 0.42 kcal/mol, whilst the DLPNO-CCSD(T) reproduces the experimental interaction energy within 0.1 kcal/mol (see above). However, the difference between the DFT-SAPT and DLPNO-CCSD(T) total interaction energies decreases with increasing intermolecular distance.

A comparison of electrostatic interactions estimated by LED and DFT-SAPT is shown in the central panel of Figure 5. At the equilibrium geometry, the sum of E_{elstat} and $E_{\text{res}}^{\text{C-SP}}$ (the $E_{\text{res}}^{\text{C-SP}}$ only provides a small correction, see above) is about two times larger than the sum of first-order polarization and second-order induction terms of DFT-SAPT, providing similar results to those recently found using the ALMO-EDA [15] decomposition. Note that in the DFT-SAPT and ALMO-EDA schemes electrostatics and induction are given as separate terms, whilst in LED both effects are included in E_{elstat}. At least part of the difference between DFT-SAPT and LED/ALMO-EDA arises from the fact that the latter schemes rely on orthogonal orbitals, whilst orbitals belonging to different fragments are not orthogonal in DFT-SAPT. In fact, LED and DFT-SAPT values converge to similar results in the long range, where the overlap between the orbitals is negligible.

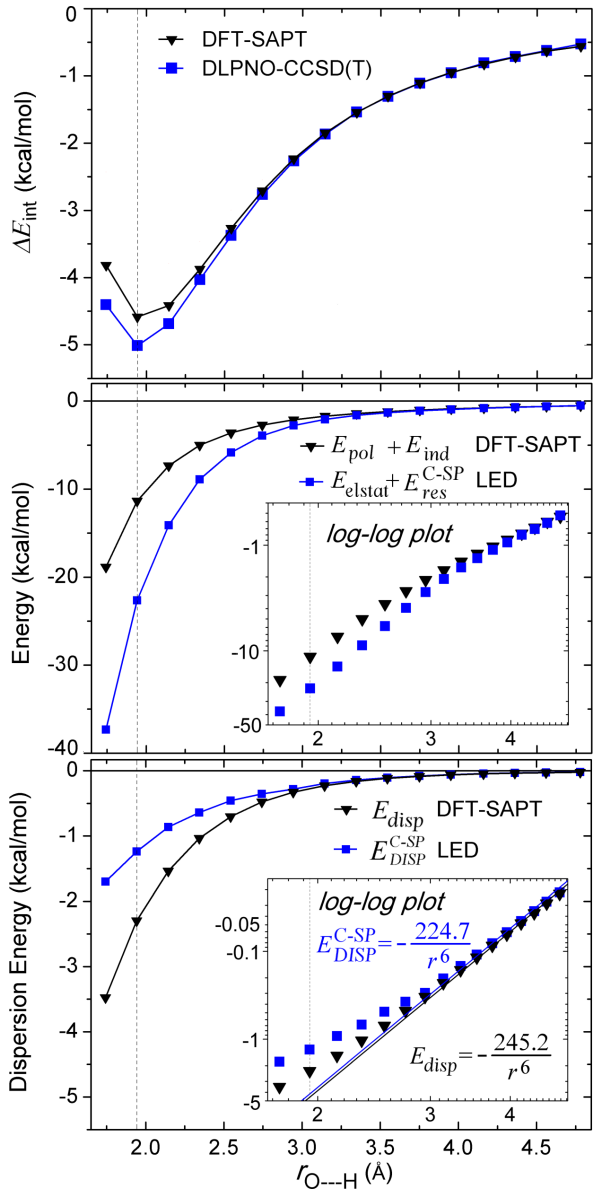


Figure 5: Comparison of total interaction, electrostatic interaction, and London dispersion energies calculated with DLPNO-CCSD(T)/LED and DFT-SAPT for Conf1 of water dimer. The black dotted vertical line at 1.943 Å corresponds to the equilibrium r_e(O---H) distance.

Finally, the comparison of London dispersion extracted from LED and DFT-SAPT is shown in the lower panel of Figure 5. At the equilibrium geometry of the water dimer, the DFT-SAPT London dispersion is −1.06 kcal/mol larger than the present DLPNO-CCSD(T)/LED result (Figure 5). It is worth mentioning that the difference in the calculated dispersion energy reduces to about half when a coupled cluster variant of SAPT is

used [76]. Again, the difference in the present LED and DFT-SAPT dispersion energies diminishes as the fragments move apart. In the long range, both definitions of dispersion decay with r^{-6} dependence with C_6 coefficients differing only by 8.4%.

These results demonstrate that, despite the non-uniqueness of the definition of the interaction energy terms, both schemes represent the correct physics in the asymptotic region.

Relative stabilities of the conformers

The four conformers of water dimer (Conf1: nonplanar with a nearly linear OH···O bond; Conf2: planar-Conf1; Conf3: cyclic; and Conf4: bifurcated) and the three conformers of HF dimer (Conf1: nonlinear; Conf2: linear; Conf3: cyclic) investigated are shown in Figure 1. In all cases, Conf1 corresponds to the absolute minimum whilst the other conformers are not stable intermediates and present at least one imaginary frequency.

The dimerization energies and their constituting LED contributions for the conformers of water and HF dimers are shown in Table 1. In all cases, the dimerization energies of the conformers are quite similar. For the water conformers, they range from −4.95 kcal/mol (Conf1) to −3.16 kcal/mol (Conf4). The situation is similar for the HF conformers, for which the dimerization energies range from −4.51 kcal/mol (Conf1) to −3.52 kcal/mol (Conf3). Therefore, the conformers of both dimers lie within 2 kcal/mol. The inclusion of the ZPE correction to relative energies, which amounts up to 0.7 kcal/mol, makes the energetic separation between the conformers even smaller.

These results already suggest that subtle differences in the various terms of the interaction determine the trend in the interaction energies. Deeper insight into this aspect can be obtained by looking at the individual contributions from the LED decomposition. Consistent with what was discussed in the previous sections, the geometric preparation, weak pairs, and perturbative triples do not contribute significantly to the relative stabilities of the different conformers. In all cases, the largest LED terms are electronic preparation and electrostatic interactions at the *HF* level. This is not surprising, considering the strong dipole moments of water and HF and in light of the fact that the electrostatic interaction is well described at the HF level, as shown in the previous section. Interestingly, E_{elstat} and ΔE_{e} show similar trends, thus highlighting the importance of classical electrostatic interactions in determining the relative stabilities of different conformers. However, as E_{elstat} and $\Delta E^{HF}_{\text{el-prep}}$ largely cancel each other, the other contributions of the interaction also play an important role. In particular, the interfragment exchange energy provides a fundamental stabilizing component for all conformers and is typically of the same order of the overall $\Delta E^{HF}_{\text{int}}$.

Electron correlation also affects the energetic separation of various conformers. For example *HF* predicts a large energetic separation between Conf2 and Conf3 for both water and HF dimers (about 0.6 kcal/mol) whilst the inclusion of electron correlation makes them virtually degenerate. The LED decomposition of the strong pairs shows that $\Delta E^{\text{C-SP}}_{\text{res}}$, i.e., the sum of the counteracting dynamic electronic preparation and dynamic charge polarization, is positive and ranges from 0.16 to 0.64 kcal/mol. Hence, the major correlation contribution to ΔE_{e} in all cases arises from the London dispersion $E^{\text{C-SP}}_{\text{DISP}}$, which ranges from −0.9 to 1.4 kcal/mol. However, it is worth underscoring that London dispersion and $E^{\text{C-SP}}_{\text{res}}$ show similar variations among the various conformers. This picture holds true for all conformers of water and HF dimers. In brief, the LED analyses show that the energetic ordering for the conformers of the water and HF dimers arises from a balance of the stabilizing electrostatic (which is dominated by the *HF* contribution), interfragment exchange, and dispersion terms, which are partially counteracted by the positive electronic preparation.

Conclusion

The recently developed LED scheme in the DLPNO-CCSD(T) framework is a useful and affordable tool to accurately quantify interaction energies and provides their decomposition into physically meaningful terms. In this work, this scheme was applied to the study of H-bond interactions on a series of prototype molecular systems, i.e., a series of conformers of water and HF dimers. For the water dimer, results are compared to the ones obtained from the popular DFT-SAPT approach.

The dissociation energy of water and HF dimers in their equilibrium structure was computed at the DLPNO-CCSD(T) level and results were found to be in perfect agreement with available experimental and previously available CCSD(T) data. On the other hand, the DFT-SAPT was found to underestimate the interaction energy in the water dimer by 0.42 kcal/mol.

For the water dimer, the decay of the different LED components with the intermolecular distance was studied. It was found that, when the water dimer is in its equilibrium structure, the electrostatic interaction estimated via the LED scheme is about twice as large as that obtained from DFT-SAPT. This difference mainly arises from the fact that LED uses orthogonal orbitals whilst the orbitals of different fragments are non-orthogonal in DFT-SAPT. However, both schemes converge to the same asymptotic value. The London dispersion interaction calculated by DFT-SAPT and LED schemes differ by

1 kcal/mol in the equilibrium position, but also converge to the same values in the long range, showing in both cases the expected r^{-6} decay. The LED analysis demonstrates the presence of another stabilizing contribution in the short range, i.e., the interfragment exchange. This component of the interaction decays exponentially and acts by lowering the repulsion of electrons with the same spin.

In the last part of the paper, the DLPNO-CCSD(T)/LED scheme is used to rationalize the trend of stability of a series of conformers of water and HF dimers. It was found that the energetic separation introduced by different H-bond networks arise from a balance between many terms.

Supporting Information

Supporting Information File 1

The Cartesian coordinates of the optimized structures; the individual and total DLPNO-CCSD(T)/LED energies computed with aug-cc-pVTZ and aug-cc-pVQZ; and HF-SAPT and DFT-SAPT energies computed with aug-cc-pVTZ and aug-cc-pVQZ. The energetics includes CBS, BSSE, and BSSE-followed CBS corrected values. Additional data.
[https://www.beilstein-journals.org/bjoc/content/supplementary/1860-5397-14-79-S1.pdf]

Acknowledgements

We gratefully acknowledge the Priority Program “Control of Dispersion Interactions in Molecular Chemistry” (SPP 1807) of the Deutsche Forschungsgemeinschaft for financial support.

ORCID® iDs

Ahmet Altun - https://orcid.org/0000-0001-8818-9925
Frank Neese - https://orcid.org/0000-0003-4691-0547
Giovanni Bistoni - https://orcid.org/0000-0003-4849-1323

References

1. Yang, F.; Wang, X.; Yang, M.; Krishtal, A.; van Alsenoy, C.; Delarue, P.; Senet, P. *Phys. Chem. Chem. Phys.* **2010,** *12,* 9239–9248. doi:10.1039/c001007c
2. Bowie, J. U. *Curr. Opin. Struct. Biol.* **2011,** *21,* 42–49. doi:10.1016/J.SBI.2010.10.003
3. Altun, A.; Morokuma, K.; Yokoyama, S. *ACS Chem. Biol.* **2011,** *6,* 775–780. doi:10.1021/cb200100f
4. Goodman, M. F. *Proc. Natl. Acad. Sci. U. S. A.* **1997,** *94,* 10493–10495. doi:10.1073/PNAS.94.20.10493
5. Altun, A.; Guallar, V.; Friesner, R. A.; Shaik, S.; Thiel, W. *J. Am. Chem. Soc.* **2006,** *128,* 3924–3925. doi:10.1021/JA058196W
6. Ishikita, H.; Saito, K. *J. R. Soc., Interface* **2014,** *11,* 20130518. doi:10.1098/rsif.2013.0518
7. Ok, S.; Altun, A.; Kasimogullari, R.; Sen, E. *J. Chem. Eng. Data* **2013,** *58,* 3521–3527. doi:10.1021/je400806h
8. von Hopffgarten, M.; Frenking, G. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2012,** *2,* 43–62. doi:10.1002/wcms.71
9. Phipps, M. J. S.; Fox, T.; Tautermann, C. S.; Skylaris, C.-K. *Chem. Soc. Rev.* **2015,** *44,* 3177–3211. doi:10.1039/C4CS00375F
10. Pastorczak, E.; Corminboeuf, C. *J. Chem. Phys.* **2017,** *146,* 120901. doi:10.1063/1.4978951
11. Morokuma, K. *J. Chem. Phys.* **1971,** *55,* 1236–1244. doi:10.1063/1.1676210
12. Kitaura, K.; Morokuma, K. *Int. J. Quantum Chem.* **1976,** *10,* 325–340. doi:10.1002/qua.560100211
13. Morokuma, K. *Acc. Chem. Res.* **1977,** *10,* 294–300. doi:10.1021/ar50116a004
14. Umeyama, H.; Morokuma, K. *J. Am. Chem. Soc.* **1977,** *99,* 1316–1332. doi:10.1021/ja00447a007
15. Mao, Y.; Horn, P. R.; Head-Gordon, M. *Phys. Chem. Chem. Phys.* **2017,** *19,* 5944–5958. doi:10.1039/C6CP08039A
16. Mo, Y.; Gao, J.; Peyerimhoff, S. D. *J. Chem. Phys.* **2000,** *112,* 5530–5538. doi:10.1063/1.481185
17. Jeziorski, B.; Moszynski, R.; Szalewicz, K. *Chem. Rev.* **1994,** *94,* 1887–1930. doi:10.1021/cr00031a008
18. Tafipolsky, M. *J. Phys. Chem. A* **2016,** *120,* 4550–4559. doi:10.1021/acs.jpca.6b04861
19. Wang, B.; Jiang, W.; Dai, X.; Gao, Y.; Wang, Z.; Zhang, R.-Q. *Sci. Rep.* **2016,** *6,* No. 22099. doi:10.1038/srep22099
20. Ramabhadran, R. O.; Raghavachari, K. *J. Chem. Theory Comput.* **2013,** *9,* 3986–3994. doi:10.1021/ct400465q
21. Řezáč, J.; Hobza, P. *J. Chem. Theory Comput.* **2013,** *9,* 2151–2155. doi:10.1021/ct400057w
22. Mata, R. A.; Suhm, M. A. *Angew. Chem., Int. Ed.* **2017,** *56,* 11011–11018. doi:10.1002/anie.201611308
23. Neese, F.; Hansen, A.; Liakos, D. G. *J. Chem. Phys.* **2009,** *131,* 64103. doi:10.1063/1.3173827
24. Neese, F.; Hansen, A.; Wennmohs, F.; Grimme, S. *Acc. Chem. Res.* **2009,** *42,* 641–648. doi:10.1021/ar800241t
25. Neese, F.; Wennmohs, F.; Hansen, A. *J. Chem. Phys.* **2009,** *130,* 114108. doi:10.1063/1.3086717
26. Hansen, A.; Liakos, D. G.; Neese, F. *J. Chem. Phys.* **2011,** *135,* 214102. doi:10.1063/1.3663855
27. Liakos, D. G.; Hansen, A.; Neese, F. *J. Chem. Theory Comput.* **2011,** *7,* 76–87. doi:10.1021/ct100445s
28. Huntington, L. M. J.; Hansen, A.; Neese, F.; Nooijen, M. *J. Chem. Phys.* **2012,** *136,* 64101. doi:10.1063/1.3682325
29. Riplinger, C.; Neese, F. *J. Chem. Phys.* **2013,** *138,* 34106. doi:10.1063/1.4773581
30. Riplinger, C.; Sandhoefer, B.; Hansen, A.; Neese, F. *J. Chem. Phys.* **2013,** *139,* 134101. doi:10.1063/1.4821834
31. Riplinger, C.; Pinski, P.; Becker, U.; Valeev, E. F.; Neese, F. *J. Chem. Phys.* **2016,** *144,* 24109. doi:10.1063/1.4939030
32. Liakos, D. G.; Sparta, M.; Kesharwani, M. K.; Martin, J. M. L.; Neese, F. *J. Chem. Theory Comput.* **2015,** *11,* 1525–1539. doi:10.1021/ct501129s
33. Schneider, W. B.; Bistoni, G.; Sparta, M.; Saitow, M.; Riplinger, C.; Auer, A. A.; Neese, F. *J. Chem. Theory Comput.* **2016,** *12,* 4778–4792. doi:10.1021/acs.jctc.6b00523
34. Fedorov, D. G.; Kitaura, K. *J. Comput. Chem.* **2007,** *28,* 222–237. doi:10.1002/jcc.20496

35. Glendening, E. D. *J. Phys. Chem. A* **2005,** *109,* 11936–11940. doi:10.1021/JP058209S
36. Glendening, E. D.; Streitwieser, A. *J. Chem. Phys.* **1994,** *100,* 2900–2909. doi:10.1063/1.466432
37. Chen, W.; Gordon, M. S. *J. Phys. Chem.* **1996,** *100,* 14316–14328. doi:10.1021/jp960694r
38. Schenter, G. K.; Glendening, E. D. *J. Phys. Chem.* **1996,** *100,* 17152–17156. doi:10.1021/JP9612994
39. Khaliullin, R. Z.; Cobar, E. A.; Lochan, R. C.; Bell, A. T.; Head-Gordon, M. *J. Phys. Chem. A* **2007,** *111,* 8753–8765. doi:10.1021/jp073685z
40. Røeggen, I. *Mol. Phys.* **1990,** *70,* 353–375. doi:10.1080/00268979000101051
41. Hartke, B.; Schütz, M.; Werner, H.-J. *J. Chem. Phys.* **1988,** *239,* 561–572. doi:10.1016/S0301-0104(98)00322-X
42. Wagner, J. P.; Schreiner, P. R. *Angew. Chem., Int. Ed.* **2015,** *54,* 12274–12296. doi:10.1002/anie.201503476
43. Liptrot, D. J.; Power, P. P. *Nat. Rev. Chem.* **2017,** *1,* No. 4. doi:10.1038/s41570-016-0004
44. Neese, F. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2012,** *2,* 73–78. doi:10.1002/wcms.81
45. Neese, F. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2018,** *8,* e1327. doi:10.1002/wcms.1327
46. Dunning, T. H., Jr. *J. Chem. Phys.* **1989,** *90,* 1007–1023. doi:10.1063/1.456153
47. Balabanov, N. B.; Peterson, K. A. *J. Chem. Phys.* **2005,** *123,* 64107. doi:10.1063/1.1998907
48. Peterson, K. A.; Dunning, T. H., Jr. *J. Chem. Phys.* **2002,** *117,* 10548–10560. doi:10.1063/1.1520138
49. Woon, D. E.; Dunning, T. H., Jr. *J. Chem. Phys.* **1994,** *100,* 2975–2988. doi:10.1063/1.466439
50. Weigend, F.; Häser, M.; Patzelt, H.; Ahlrichs, R. *Chem. Phys. Lett.* **1998,** *294,* 143–152. doi:10.1016/S0009-2614(98)00862-8
51. Eichkorn, K.; Treutler, O.; Öhm, H.; Häser, M.; Ahlrichs, R. *Chem. Phys. Lett.* **1995,** *240,* 283–290. doi:10.1016/0009-2614(95)00621-A
52. Boys, S. F. *Rev. Mod. Phys.* **1960,** *32,* 296–299. doi:10.1103/RevModPhys.32.296
53. Pipek, J.; Mezey, P. G. *J. Chem. Phys.* **1989,** *90,* 4916–4926. doi:10.1063/1.456588
54. Boys, S. F.; Bernardi, F. *Mol. Phys.* **1970,** *19,* 553–566. doi:10.1080/00268977000101561
55. Halkier, A.; Helgaker, T.; Jørgensen, P.; Klopper, W.; Koch, H.; Olsen, J.; Wilson, A. K. *Chem. Phys. Lett.* **1998,** *286,* 243–252. doi:10.1016/S0009-2614(98)00111-0
56. Neese, F.; Valeev, E. F. *J. Chem. Theory Comput.* **2011,** *7,* 33–43. doi:10.1021/ct100396y
57. Werner, H.-J.; Knowles, P. J.; Knizia, G.; Manby, F. R.; Schütz, M. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2012,** *2,* 242–253. doi:10.1002/wcms.82
58. Řezáč, J.; Hobza, P. E. *J. Chem. Theory Comput.* **2011,** *7,* 685–689. doi:10.1021/ct200005p
59. Halkier, A.; Helgaker, T.; Jørgensen, P.; Klopper, W.; Olsen, J. *Chem. Phys. Lett.* **1999,** *302,* 437–446. doi:10.1016/S0009-2614(99)00179-7
60. Heßelmann, A.; Jansen, G. *Chem. Phys. Lett.* **2002,** *357,* 464–470. doi:10.1016/S0009-2614(02)00538-9
61. Jansen, G. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2014,** *4,* 127–144. doi:10.1002/wcms.1164
62. Reutt, J. E.; Wang, L. S.; Lee, Y. T.; Shirley, D. A. *J. Chem. Phys.* **1986,** *85,* 6928–6939. doi:10.1063/1.451379
63. Page, R. H.; Larkin, R. J.; Shen, Y. R.; Lee, Y. T. *J. Chem. Phys.* **1988,** *88,* 2249–2263. doi:10.1063/1.454058
64. Bistoni, G.; Auer, A. A.; Neese, F. *Chem. – Eur. J.* **2017,** *23,* 865–873. doi:10.1002/chem.201604127
65. Szalewicz, K. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2012,** *2,* 254–272. doi:10.1002/wcms.86
66. Hohenstein, E. G.; Sherrill, C. D. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2012,** *2,* 304–326. doi:10.1002/wcms.84
67. Odutola, J. A.; Dyke, T. R. *J. Chem. Phys.* **1980,** *72,* 5062–5070. doi:10.1063/1.439795
68. Klopper, W.; van Duijneveldt-van de Rijdt, J. G. C. M.; van Duijneveldt, F. B. C. *Phys. Chem. Chem. Phys.* **2000,** *2,* 2227–2234. doi:10.1039/a910312k
69. Lane, J. R. *J. Chem. Theory Comput.* **2013,** *9,* 316–323. doi:10.1021/ct300832f
70. Klopper, W.; Quack, M.; Suhm, M. A. *J. Chem. Phys.* **1998,** *108,* 10096–10115. doi:10.1063/1.476470
71. Howard, J. C.; Gray, J. L.; Hardwick, A. J.; Nguyen, L. T.; Tschumper, G. S. *J. Chem. Theory Comput.* **2014,** *10,* 5426–5435. doi:10.1021/ct500860v
72. Samanta, A. K.; Czakó, G.; Wang, Y.; Mancini, J. S.; Bowman, J. M.; Reisler, H. *Acc. Chem. Res.* **2014,** *47,* 2700–2709. doi:10.1021/ar500213q
73. Rocher-Casterline, B. E.; Ch'ng, L. C.; Mollner, A. K.; Reisler, H. *J. Chem. Phys.* **2011,** *134,* 211101. doi:10.1063/1.3598339
74. Řezáč, J.; Hobza, P. *J. Chem. Theory Comput.* **2014,** *10,* 3066–3073. doi:10.1021/ct500047x
75. Thirman, J.; Head-Gordon, M. *J. Phys. Chem. Lett.* **2014,** *5,* 1380–1385. doi:10.1021/jz500165u
76. Korona, T. *J. Chem. Theory Comput.* **2009,** *5,* 2663–2678. doi:10.1021/ct900232j

License and Terms

BEILSTEIN JOURNAL OF ORGANIC CHEMISTRY

Correlation effects and many-body interactions in water clusters

Andreas Heßelmann

Full Research Paper

Open Access

Address:
Lehrstuhl für Theoretische Chemie, Universität Erlangen-Nürnberg, Egerlandstr. 3, 91058 Erlangen, Germany

Email:
Andreas Heßelmann - andreas.hesselmann@fau.de

Keywords:
dispersion; many-body effects; water

Beilstein J. Org. Chem. **2018,** *14,* 979–991.
doi:10.3762/bjoc.14.83

Received: 15 February 2018
Accepted: 11 April 2018
Published: 02 May 2018

This article is part of the Thematic Series "Dispersion interactions".

Guest Editor: P. Schreiner

Abstract

Background: The quantum-chemical description of the interactions in water clusters is an essential basis for deriving accurate and physically sound models of the interaction potential for water to be used in molecular simulations. In particular, the role of many-body interactions beyond the two-body interactions, which are often not explicitly taken into account by empirical force fields, can be accurately described by quantum chemistry methods on an adequate level, e.g., random-phase approximation electron correlation methods. The relative magnitudes of the different interaction energy contributions obtained by accurate ab initio calculations can therefore provide useful insights that can be exploited to develop enhanced force field methods.

Results: In line with earlier theoretical studies of the interactions in water clusters, it has been found that the main contribution to the many-body interactions in clusters with a size of up to $N = 13$ molecules are higher-order polarisation interaction terms. Compared to this, many-body dispersion interactions are practically negligible for all studied sytems. The two-body dispersion interaction, however, plays a significant role in the formation of the structures of the water clusters and their stability, since it leads to a distinct compression of the cluster sizes compared to the structures optimized on an uncorrelated level. Overall, the many-body interactions amount to about 13% of the total interaction energy, irrespective of the cluster size. The electron correlation contribution to these, however, amounts to only about 30% to the total many-body interactions for the largest clusters studied and is repulsive for all structures considered in this work.

Conclusion: While this shows that three- and higher-body interactions can not be neglected in the description of water complexes, the electron correlation contributions to these are much smaller in comparison to the two-body electron correlation effects. Efficient quantum chemistry approaches for describing intermolecular interactions between water molecules may therefore describe higher-body interactions on an uncorrelated Hartree–Fock level without a serious loss in accuracy.

Introduction

The description of the intermolecular interactions between water molecules is essential for an understanding of the structures and properties of water through the different stages of assemblies, from the dimer over the liquid phase to the bulk phase. Moreover, many chemical processes are explicitly or implicitly influenced by a water environment. An example for this is the hydrogen-bond cooperativity effect that can have a significant impact on the properties of the bare solute molecules [1]. In order to describe such phenomena, computer simulations have become an indispensable tool, since they enable a description of water on a molecular level that often can provide further insights than are accessible from spectroscopic measurements.

The basis for such simulations are the so-called force fields that describe both the covalent as well as the noncovalent interactions within the system. These commonly depend on a number of empirical parameters that are determined either by a fit to experimentally known liquid or bulk properties, or by fitting to energies from ab initio quantum chemistry methods. The most popular potentials for water are the TIP3P [2], TIP4P [2,3] and TIP5P [4,5] force fields, which are based on a modeling of the water pair potential using an electrostatic contribution described by interacting point charges and a van der Waals interaction contribution using Lennard-Jones potentials. In more advanced ab initio water pair potentials the force field is fitted to high-level quantum chemistry results for the water dimer. Force fields belonging to this category are, e.g., the TTM3-F and TTM4-F models [6,7], the AMOEBA force field [8,9], the DPP2 model [10] and various force fields derived by Szalewicz and co-workers [11-13]. There also exist a number of pair potentials that go beyond the point charge approximation [14].

A comparison with high-level coupled-cluster energies for a large number of water dimers and tetramers has revealed, however, that polarisation effects, which are not accounted for in the classical point charge potentials, are essential to describe the structures of $(H_2O)_2$ and $(H_2O)_3$ in many different conformations [15]. A rather good correlation between the coupled-cluster energies and the force-field energies is found for the polarisable AMOEBA2003 [9,16] and TTM4-F [17] potentials both for the dimer and for the trimer. Both these methods are based on the induced dipole scheme in which polarisable point dipoles, which are assigned to the molecules, interact with the surrounding electric field and are computed in a self-consistent manner. While these force fields, too, rely on a certain degree of empiricism, a number of other force field exist that aim at a more physically sound decomposition of the interaction energy into distinct contributions. Examples for such force fields are the sum of interaction between fragments (SIBFA) [18-20] and the effective fragment potential (EFP) [21] method. The most recent version of the latter, EFP2, can describe both (long-range) polarisation as well as charge-transfer interactions. The latter was found to yield a significant contribution to the interaction energy of the water dimer [22].

However, one of the most significant results of [22] was that the dipole–quadrupole polarisability term of the multipole expansion of the dispersion interaction, which is usually neglected in force fields or dispersion corrected DFT methods, is a quite large positive (for the clusters considered) and anisotropic contribution to the interaction energy of small water clusters [22]. It was found to be almost half of the magnitude of the leading-order dipole–dipole term. It was therefore concluded by Guidez et al. that this term should not be neglected in the description of the interactions in water.

Another challenge for water models is the description of nonadditive many-body terms to the interaction energy [23-25]. It has been found that these contribute 15% to the total interactions in the condensed phase [26] and this amount even increases to 17–30% for small water clusters [26-28]. Explicit evaluations of three-body interaction energy terms using symmetry-adapted perturbation theory (SAPT) for the water trimer have revealed that the strongest contribution to the three-body energy originates from the polarisation (induction) energy while the three-body dispersion interaction is rather small [28,29]. Moreover, many-body exchange effects, including exchange–induction and exchange–dispersion, are relatively large yet cancel each other due to opposite signs at the uud minimum configuration [29] ("uud" indicates that two "free" hydrogen atoms point above the plane formed by the three oxygen atoms (u→up) and one below it (d→down) [30]). In the study by Hodges et al. it has been shown for several structures of the water tetramer that total four-body interactions are in most cases much smaller or even negligible compared to the three-body interactions [27]. The only exception to this was observed for the squared geometry in which the hydrogen bonds act cooperatively to enhance the induction energy. In a recent work by Hapka et al. it was shown that also standard density functional theory methods are able to describe nonadditive effects to the interaction energy quite well for hydrogen-bonded clusters, yet, fail to do so for dispersion bound complexes [31]. The common conclusion from the quantum chemistry studies of small water clusters was that damped classical polarization models should be able to accurately capture the nonadditive effects for larger clusters of water, because of the fact that short-range contributions, including many-body exchange effects, grow less strongly with the size of the system than induction, dispersion or electrostatic interaction energies [32].

In this work we will study the impact of electron correlation effects and many-body interactions on the structures and energies of water clusters $(H_2O)_n$ with cluster sizes ranging between $n = 2$ and $n = 13$. In doing so, we focus on the importance of dispersion energy contributions to the interaction energy, including two-body and many-body dispersion effects. It will be shown that overall electron correlation effects have only a minor impact on the orientation of the water molecules in the various minimum structures of the clusters. Dispersion interactions, however, make up a significant contribution of about 60% to the total interaction energies in larger water clusters and lead to a compression of the cluster sizes on average. The total magnitude of the electron correlation contribution to the interaction energy is, however, only about half the size of the sum of the two-, three- and four-body dispersion interactions, i.e., the large dispersion interaction contribution is strongly quenched by further repulsive correlation contributions. Noting that the correlation effect to the molecular dipole moment of the water molecule reduces the dipole moment from the Hartree–Fock method by about 0.13 Debye [33], it can be assessed that a fraction of this repulsive correlation contribution originates from the reduction of the electrostatic interaction energy [34]. Empirical models for water that are based on a fitting to ab initio results therefore have to take the influences of the different correlation effects carefully into account.

Many-Body Expansion of the Interaction Energy

Consider a cluster system containing N molecules. The total interaction energy of this system is given by

$$\begin{aligned} E_N &= E_{\text{int}}(123\ldots N) \\ &= E(123\ldots N) - E(1) - E(2) - \ldots - E(N), \end{aligned} \tag{1}$$

with $E(123\ldots N)$ denoting the total energy of the system and $E(A)$ (with $A = 1,2,\ldots$) denoting the one-body clusters (monomers). The idea of the many-body expansion is to decompose the total interaction energy into terms arising from the two-body, three-body, four-body, etc. interactions, so that the total energy is given by

$$\begin{aligned} E_N = &\sum_{A=1}^{N} E(A) + \sum_{A=1}^{N-1}\sum_{B>A}^{N} \Delta^2 E(AB) \\ &+ \sum_{A=1}^{N-2}\sum_{B>A}^{N-1}\sum_{C>B}^{N} \Delta^3 E(ABC) + \ldots \\ &+ \sum_{A=1}^{N-3}\sum_{B>A}^{N-2}\sum_{C>B}^{N-1}\sum_{D>C}^{N} \Delta^4 E(ABCD) + \ldots \\ &+ \Delta^N E(1234\ldots N), \end{aligned} \tag{2}$$

with Δ^M denoting the M-th body interaction contribution. These describe the change of the total energy of the system due to the interactions of the M-th body clusters. The first three terms are defined as

$$\begin{aligned} \Delta^2 E(AB) &= E(AB) - [E(A) + E(B)] \\ \Delta^3 E(ABC) &= E(ABC) - [E(A) + E(B) + E(C)] \\ &\quad - [\Delta^2(AB) + \Delta^2(AC) + \Delta^2(BC)] \\ \Delta^4 E(ABCD) &= E(ABCD) - [E(A) + E(B) + E(C) + E(D)] \\ &\quad - [\Delta^2(AB) + \Delta^2(AC) + \Delta^2(AD) \\ &\quad + \Delta^2(BC) + \Delta^2(BD) + \Delta^2(CD)] \\ &\quad - [\Delta^3(ABC) + \Delta^3(ABD) \\ &\quad + \Delta^3(ACD) + \Delta^3(BCD)]. \end{aligned} \tag{3}$$

As shown in Figure 1, the number of three- and four-body clusters increases much more rapidly with the size of the system than the number of two-body clusters. In spite of this, however, a truncation of the expansion of Equation 2 after the two-body term will usually capture 90% and more of the total interaction energy. This originates in part from the fact that for larger many-body systems many three-body and four-body clusters possess structures with far distant molecules. Moreover, the individual interaction energy contributions to the three- and higher-body interaction energy tend to be much smaller than the sum of the two-body interaction energies, as will be shown in this work.

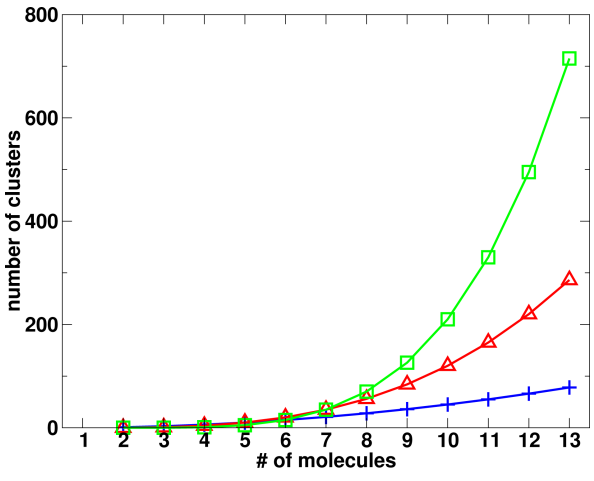


Figure 1: Number of two-body, three-body and four-body clusters for systems with up to 13 molecules.

Method

The calculations in this work have been performed using the EXX-RPA (exact-exchange random-phase approximation) electron correlation method [35-37], which is based on an exact-

exchange Kohn–Sham reference determinant [38-43]. The total energy of the EXX-RPA method is given by

$$E^{\text{EXX-RPA}} = E^{\text{EXX}} + E_{\text{corr}}^{\text{EXX-RPA}}, \tag{4}$$

where E^{EXX} denotes the energy of the EXX reference determinant and $E_{\text{corr}}^{\text{EXX-RPA}}$ is the EXX-RPA correlation energy. Like the Hartree–Fock method, the EXX method lacks the description of electron correlation effects, i.e., electron–electron interactions beyond zeroth and first order in the Coulomb interaction. Thus, any inter- and intramolecular correlation effects of the EXX-RPA method are described by the correlation term $E_{\text{corr}}^{\text{EXX-RPA}}$ of Equation 4.

In order to analyze individual interaction energy contributions of the two-body clusters, the DFT-SAPT method [44-52] method has been used. With this method, the interaction energy of a dimer system *AB* is given by

$$\begin{aligned}\Delta^2 E(AB)[\text{DFT-SAPT}] = {} & E_{\text{elst}}^{(1)} + E_{\text{exch}}^{(1)} + E_{\text{ind}}^{(2)} \\ & + E_{\text{exch-ind}}^{(2)} + E_{\text{disp}}^{(2)} + E_{\text{exch-disp}}^{(2)}.\end{aligned} \tag{5}$$

The interaction energy terms in Equation 5 are: $E_{\text{elst}}^{(1)}$ electrostatic interaction energy, $E_{\text{exch}}^{(1)}$ first-order exchange interaction energy, $E_{\text{ind}}^{(2)}$ induction energy, $E_{\text{exch-ind}}^{(2)}$ exchange-induction energy, $E_{\text{disp}}^{(2)}$ (two-body) dispersion energy and $E_{\text{exch-disp}}^{(2)}$ exchange-dispersion energy. The exchange interaction energy terms in Equation 5 are short-range contributions to the interaction energy and stem from a tunneling of the electrons between the two monomers. They quickly decay exponentially with the distance of the interacting systems. The superscripts (1) and (2) denote the order of the individual terms with respect to the interaction energy operator.

In this work all terms in Equation 5 have been computed using EXX monomer wave functions. Moreover, a time-dependent EXX (TDEXX) response approach was used to compute the second-order interaction energy contributions [53-59]. Due to this choice, the subtotal

$$E_{\text{elst}}^{(1)} + E_{\text{exch}}^{(1)} + E_{\text{ind}}^{(2)} + E_{\text{exch-ind}}^{(2)} \subset \Delta^2 E^{\text{EXX}}(AB) \tag{6}$$

approximates the EXX interaction energy between the monomers *A* and *B*. The difference between $\Delta^2 E^{\text{EXX}}(AB)$ and these interaction energy terms can be interpreted as higher-order exchange–induction interaction terms not accounted for by the DFT-SAPT method (when truncated at second order).

Moreover, the sum of the dispersion and exchange–dispersion energy

$$E_{\text{disp}}^{(2)} + E_{\text{exch-disp}}^{(2)} \subset \Delta^2 E_{\text{corr}}^{\text{EXX-RPA}}(AB) \tag{7}$$

is a fraction of the EXX-RPA correlation energy contribution to the intermolecular interaction energy, since both the DFT-SAPT terms as well as the EXX-RPA correlation energy are computed with the exact-exchange response kernel [53,54]. Here, the main difference of the quantities on the left-hand and right-hand side of Equation 7 stems from the correlation energy contributions to the electrostatic, induction and their exchange interaction energy counterparts. Note that an analogous decomposition of the supermolecular interaction energy into distinct terms is also possible for the second-order Møller-Plesset perturbation theory method [60].

Finally, we have also computed three-body and four-body dispersion interaction energies that contribute to the three-body and four-body interaction energy terms:

$$E_{\text{disp}}^{(3)}(ABC) \subset \Delta^3 E(ABC), \tag{8}$$

$$E_{\text{disp}}^{(4)}(ABCD) \subset \Delta^4 E(ABCD). \tag{9}$$

Thus, it will be possible to evaluate the importance of dispersion interactions both for the total many-body interaction energy as well as for its correlation interaction contribution.

Computational Details

The structures of the water clusters have been taken from the work of Maheshwary et al., see Figure 2. This set of structures contains both the prism form as well as the cage form of the water hexamer. These are almost isoenergetic and form the first noncyclic global minimum structures of $(H_2O)_x$. In addition, the cyclic-chair conformation of the hexamer is considered, too, in this work, since it is known that this structure is characterised by strong many-body interactions [61-63]. For consistency, the cyclic-chair geometry has been optimized on the same level that was used in [23], namely Hartree–Fock employing the 6-31G(d,p) basis set [64]. The resulting structure is shown in Figure 3.

All calculations of this work were performed using the aug-cc-pVTZ basis set by Dunning and co-workers [65]. In order to correct the basis set error of the correlation energy terms, these were extrapolated using a double-ζ to triple-ζ two-point extrap-

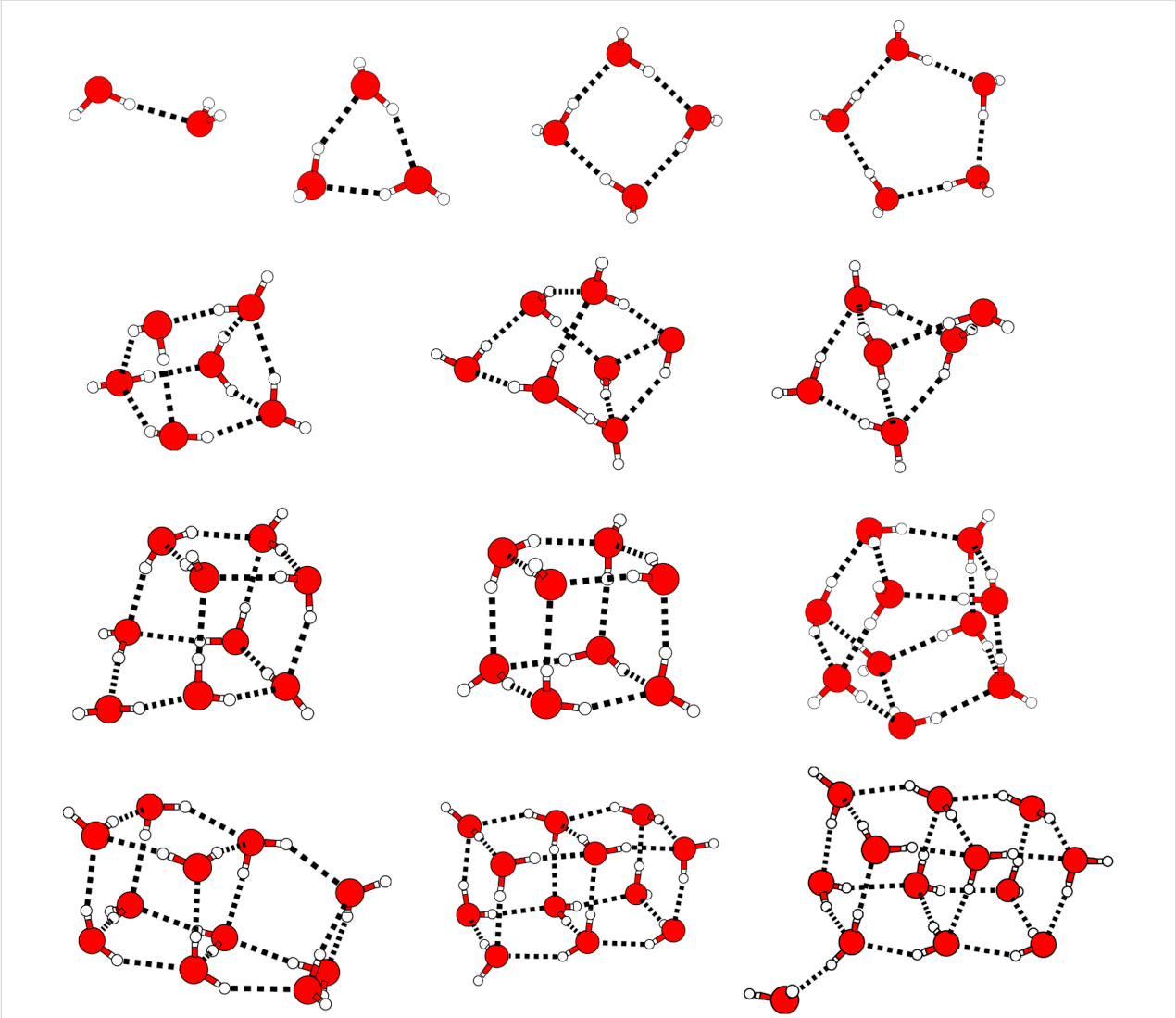


Figure 2: Water clusters from [23] studied in this work.

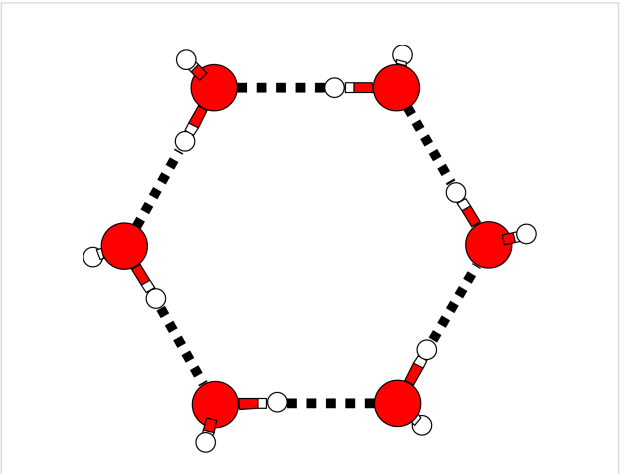


Figure 3: Cyclic-chair structure of the water hexamer.

olation using the formula from Bak and co-workers [66]. Core electrons have been kept frozen in these calculations.

The counterpoise correction of Boys and Bernadi [67] has been employed in all sub-cluster calculations to reduce the basis set superposition error. All calculations have been performed using a developers version of the Molpro quantum chemistry package [68,69].

Results and Discussion

Influence of correlation effects on the structures

To investigate the influence of electron correlation effects (including intermolecular dispersion interactions) on the structure of the water clusters, we have reoptimized the structures from [23] on the Hartree–Fock (HF) level using the def2-TZVP basis

set [70]. These modified HF geometries were then further optimized with the MP2 (second-order Møller-Plesset perturbation theory) method, which takes electron correlation effects into account at the second order. The resulting structures for the dimer up to the heptamer are shown in Figure 4. Here, for better visibility, the HF structures are colored in blue and the MP2 structures in red. They are superpositioned in such a way that the average distances of the atoms of the respective geometries are minimised.

As can be seen in Figure 4, the electron correlation effects described by the MP2 method hardly change the global structures of the various clusters. This indicates that they mainly originate from electrostatic and induction interactions that can already be described reasonably well with the HF method. However, except for the trimer, where HF and MP2 geometries hardly differ from each other, one can observe that the correlation effects lead to a compression of the structures relative to the ones obtained with HF. This can be attributed to dispersion interactions between the molecules which is an additional attractive interaction energy contribution not accounted for by the HF method, see also below.

The dependence of the electrostatic and dispersion energy on the structure is highlighted in Figure 5 for three different conformations of the water dimer. These three structures have in common the distance of the oxygen atoms (2.98 Å in this example), but they possess different orientations of the hydrogen atoms. The first structure in Figure 5 corresponds to the equilibrium. As can be seen in the figure, the electrostatic interaction energy is strongly influenced by the orientation of the molecules and changes by almost +10 kcal/mol from the hydrogen-bonded structure to the second one that is characterised by parallel dipoles of the two water molecules. Compared to this, the dispersion interaction hardly changes upon a disordering of the hydrogen atoms. While it has a minimum, too, at the equilibrium structure, for the other two structures it lies only +0.7 kcal/mol higher in energy. In line with the structure changes displayed in Figure 4, one can thus conclude that dispersion interactions almost act isotropically and therefore

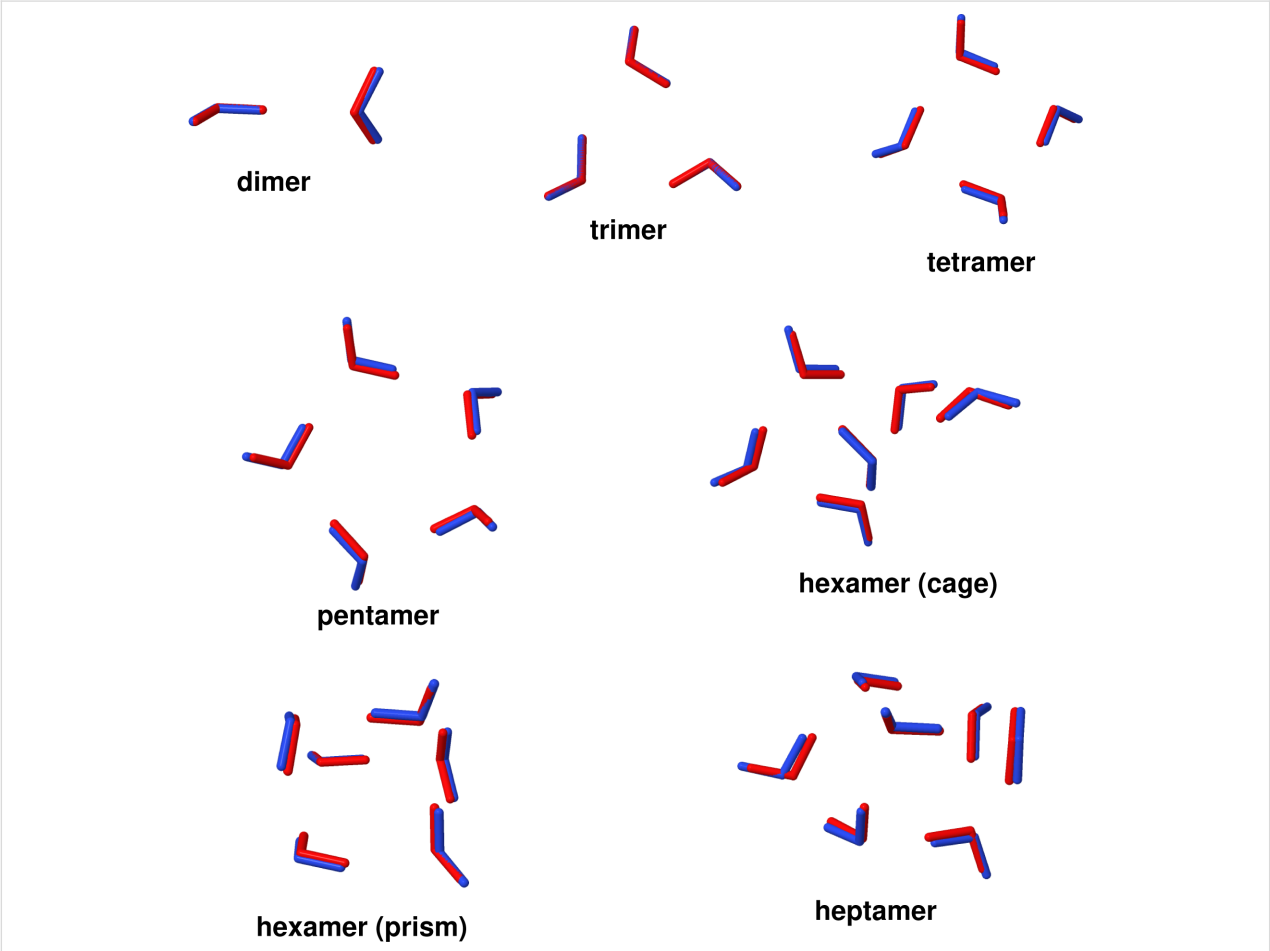


Figure 4: Comparison of the structures of the first seven water clusters optimized on the Hartree–Fock and MP2 levels of theory. Blue: HF, red: MP2 (def2-TZVP basis set).

will generally try to maximise the contacts of the interacting sites.

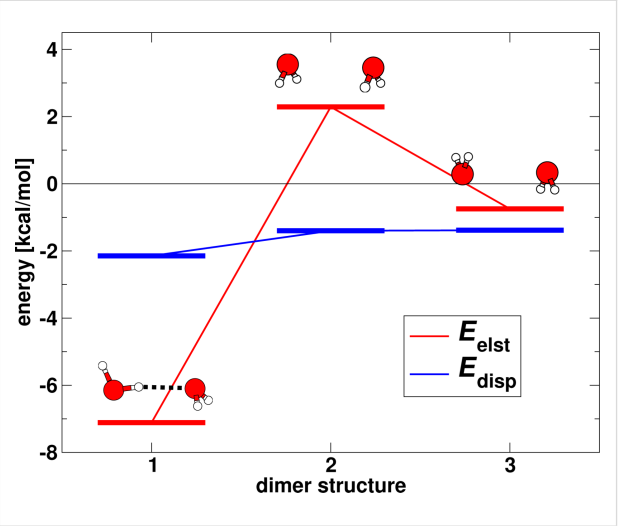


Figure 5: Electrostatic and dispersion interaction energy for three different structures of the water dimer. In all structures the oxygen–oxygen distance is equal to 2.98 Å. The first structure corresponds to the equilibrium (aug-cc-pVTZ basis set).

This is also illustrated in Figure 6, which shows the two-body and three-body dispersion energies for three different structures of the water trimer. Here, again structure 1 corresponds to the equilibrium, which is characterised by an equilateral triangle formed by the three oxygen atoms. Compared to this, in structure 2 one of the sides of the triangle is extended, fixing however the other two at a length of 2.87 Å as in the first one. Finally, structure 3 is a linear shaped structure, see Figure 6. As can be seen in the diagram, when the trimer transforms from the most compact equilibrium structure to the linear form, the two-body dispersion interaction strongly reduces by +6 kcal/mol, yet remains attractive having a magnitude of about −6 kcal/mol for the structures 2 and 3. The blue horizontal bars in Figure 6 show the energy levels of the corresponding three-body dispersion energy for the three conformations. As can be seen, it possesses just the opposite dependency on the structure as the two-body dispersion energy. That is, it is repulsive at the triangularly shaped equilibrium structure but turns into an attractive contribution when the structure changes to the linearly shaped form (see the scale on the right-hand side of the diagram in the figure). One can readily describe this anisotropic behavior of the three-body dispersion energy by the simplified Axilrod–Teller form of the interaction energy between three atoms [71]:

$$E_{\text{disp}}(ABC) = 3C_9^{ABC}\frac{1+3\cos\theta_A\cos\theta_B\cos\theta C}{R_{AB}^3R_{AC}^3R_{BC}^3}, \quad (10)$$

where R_{ij} and θ_i denote the sides and the angles of the ABC triangle and $C_9^{ABC} > 0$ is a constant coefficient. Equation 10 shows that when ABC is in a linear configuration, the three-body dispersion energy is negative (an attractive contribution), while the equilateral triangular configuration leads to a repulsive interaction. Figure 7 shows the number of three-body subclusters that possess a stabilising three-body dispersion energy contribution (determined from the results of the calculations performed in this work). It can be seen that this number grows much less strongly with the cluster size than the number of subclusters with a destabilising three-body dispersion interaction. This qualitatively explains that the total three-body dispersion energy for the respective water clusters considered in this work is always repulsive, too (see below). Furthermore, a comparison of the total magnitudes of the two-body and three-body dispersion energies shows that the three-body dispersion interaction is much weaker than the two-body dispersion interaction. Therefore, its effect on the shape and energies of the larger water clusters can almost be neglected, see also below.

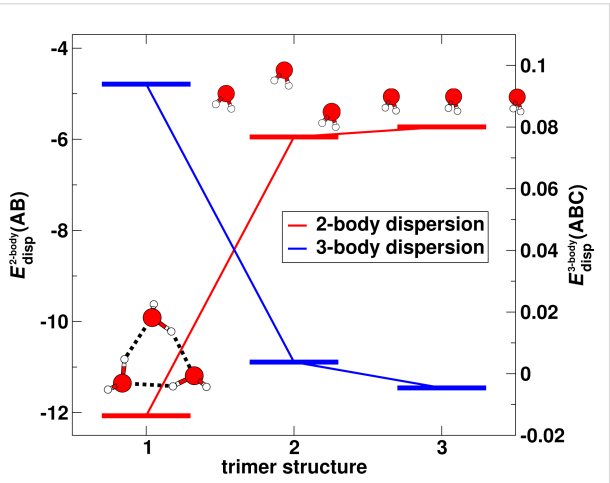


Figure 6: Two- and three-body dispersion energies for three structures of the water trimer. In all conformations the minimum oxygen–oxygen distance amounts to 2.87 Å. The first structure corresponds to the equilibrium. Note that an alternate ordinate scale is used for the two- and three-body energies (aug-cc-pVTZ basis set).

Many-body interactions in the water trimer and tetramer structures

In this section we analyse the three- and four-body interactions in the cyclic water trimer and tetramer structures, see Figure 2. More precisely, we here want to identify the main interaction energy contributions to the many-body interaction in these two cases. Note, though, that the interaction energy terms of the DFT-SAPT method, see Equation (Equation 5), can only describe the two-body interactions between two subsystems A and B. While three-body contributions to the DFT-SAPT method have been developed by Podeszwa and Szalewicz [29]

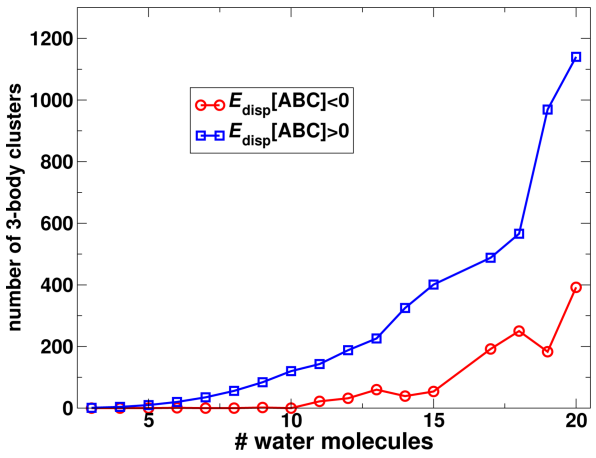


Figure 7: Number of three-body subclusters for which the three-body dispersion energy is attractive (red) or repulsive (blue).

many-body effects can, however, be also described by the two-body DFT-SAPT terms with the aid of the pseudodimer technique. For this, recall that the interaction energy of a trimer ABC can be approximated by the sum of all mutual two-body interactions:

$$E_{\text{int}}(ABC) \approx \Delta^2(AB) + \Delta^2(AC) + \Delta^2(BC). \quad (11)$$

The two-body interaction terms in Equation 11 describe the interactions between two isolated monomers neglecting, however, the perturbation by the third one. Alternatively, two-body interactions can be determined by combining two monomers to one single (pseudo-)monomer and calculating the interaction with the remaining one. In case of the trimer three different possibilities exist:

$$\Delta^2(AB-C),\ \Delta^2(AC-B) \text{ and } \Delta^2(BC-A), \quad (12)$$

where the term $\Delta^2(AB - C)$ now denotes the interaction of the combined system AB with monomer C. As can be easily understood, the term $\Delta^2(AB - C)$ contains the two two-body interactions $\Delta^2(AC)$ and $\Delta^2(BC)$ and a remainder that describes the change of these two interactions due to the perturbation of A by B and vice versa. This precisely is the contribution that is not described by Equation 11 and, thus, is a contribution to the three-body interaction energy $\Delta^3(ABC)$. According to the three different possible pseudodimers, see Equation 12, one can extract the following three terms:

$$\Delta^3(AB-C) = \Delta^2(AB-C) - \Delta^2(AC) - \Delta^2(BC), \quad (13)$$

$$\Delta^3(AC-B) = \Delta^2(AC-B) - \Delta^2(AB) - \Delta^2(BC), \quad (14)$$

$$\Delta^3(BC-A) = \Delta^2(BC-A) - \Delta^2(AB) - \Delta^2(AC). \quad (15)$$

Recall that using the supermolecule method the term $\Delta^3(AB - C)$ is given by $\Delta^3(AB - C) = E(ABC) - E(AB) - E(C)$. An insertion of the corresponding total energy expressions into the other terms in Equation 13–Equation 15 shows that

$$\begin{aligned}&\Delta^3(AB-C) + \Delta^3(AC-B) + \Delta^3(BC-A)\\ &= 3\left[E(ABC) - E(A) - E(B) - E(C)\right]\\ &= 3\Delta^3(ABC),\end{aligned} \quad (16)$$

and therefore the sum of the terms $\Delta^3(AB - C)$, $\Delta^3(AC - B)$ and $\Delta^3(BC - A)$ scaled by a factor of one third can be identified as three-body interaction energy.

We have calculated these terms using the DFT-SAPT method, which allows us to analyse the contribution of the different interaction energy terms of Equation 5 to the three-body interaction energy. The results for the water trimer are shown in Figure 8. In addition, the diagram also contains the three-body interaction energy of the EXX-RPA method as well as its correlation contribution. We have also compared the sum of the terms $\bar{E}^{(1)}_{\text{elst}} + \bar{E}^{(1)}_{\text{exch}} + \bar{E}^{(2)}_{\text{ind}} + \bar{E}^{(2)}_{\text{exch-ind}}$ (the overlines are used here to distinguish the terms from the two-body SAPT interaction terms) to the total EXX three-body interaction energy and found a good agreement (DFT-SAPT: −1.63 kcal/mol, EXX: −1.70 kcal/mol). This shows that higher-order interaction contributions to the three-body interaction energy of the EXX method are small.

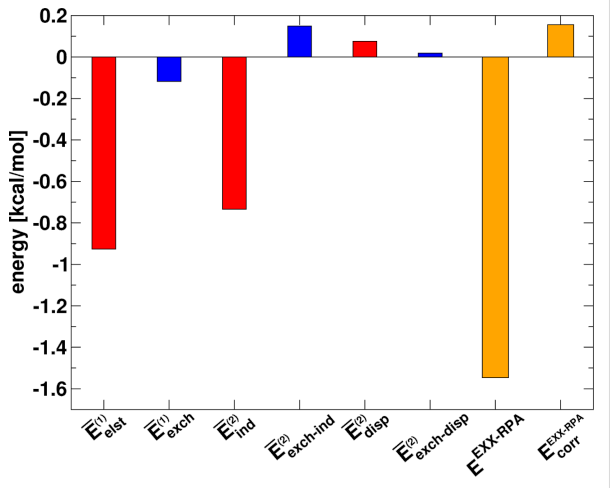


Figure 8: DFT-SAPT energy decomposition of the three-body interaction energy of the water trimer calculated using the pseudo-dimer technique.

The diagram in Figure 8 shows that the main contributions to the three-body interaction energy stem from the electrostatic and induction energies. These should not be confused with the standard two-body interaction energy counterparts but should rather be interpreted as the changes of the electrostatic and induction interactions due to the perturbations by another monomer in the trimer. Accordingly, one may interpret the electrostatic contribution in Figure 8 as an induction effect and the induction contribution as a higher-order polarisation effect contained in the three-body interaction energy. The exchange energy counterparts to $\overline{E}^{(1)}_{\text{elst}}$ and $\overline{E}^{(2)}_{\text{ind}}$ are quite small and strongly cancel each other, making up only a marginal net contribution to the three-body energy. This holds true also for the correlation contribution to the three-body interaction, including the dispersion energy, which is found to be repulsive. Therefore, in line with the findings of the previous section, the many-body correlation effects are negligible in the sum of the different three-body terms, see also Figure 6.

The pseudodimer scheme can also be applied to the tetramer. While this can be done in various ways, in this work we used the following terms:

$$\Delta^4(AB-CD) = \Delta^2(AB-CD) - \Delta^2(AC) - \Delta^2(BC) - \Delta^2(AD) - \Delta^2(BD), \quad (17)$$

$$\Delta^4(AC-BD) = \Delta^2(AC-BD) - \Delta^2(AB) - \Delta^2(AD) - \Delta^2(BC) - \Delta^2(CD), \quad (18)$$

$$\Delta^4(AD-BC) = \Delta^2(AD-BC) - \Delta^2(AB) - \Delta^2(AC) - \Delta^2(BD) - \Delta^2(CD). \quad (19)$$

One finds

$$\Delta^4(AB-CD) + \Delta^4(AC-BD) + \Delta^4(AD-BC) = 3\left[\Delta^4(ABCD) + \Delta^3(ABCD)\right], \quad (20)$$

and therefore, analogously to the procedure described above for the trimer, the sum has to be scaled by a factor of one third to reproduce the sum of the three-body and four-body interactions. The DFT-SAPT interaction energy decomposition of the many-body interactions of the water tetramer is shown in Figure 9. While compared to the trimer case the magnitudes of the individual components are distinctly larger, qualitatively the situation is similar to the trimer case. Namely, the main contributions to the many-body interactions in the tetramer stem from the (changes in the) electrostatic and induction energy while again correlation effects are comparably small. The difference between the total EXX many-body interaction energy and the sum of the first and second order energies (excluding the (exchange–)dispersion energy) amounts to −0.15 kcal/mol, which is slightly larger than in case of the trimer. This indicates an increasing importance of higher order interaction energy terms to the many-body interaction energy for larger cluster sizes.

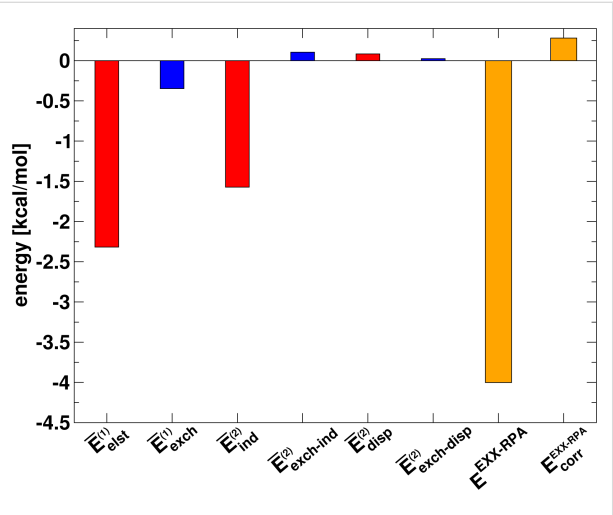


Figure 9: DFT-SAPT energy decomposition of the sum of the three- and four-body interaction energy of the water tetramer calculated using the pseudo-dimer technique.

Dependence of energy contributions on the cluster size

Various contributions to the interaction energy of the water clusters are presented in Table 1 up to the water tridecamer. The second to fourth column show the two-body, three-body and four-body dispersion energies, the fifth column contains the sum of all two-body interactions ($\sum_{A=1}^{N-1}\sum_{B>A}^{N}\Delta^2 E(AB)$) and the seventh column the total (all-body) interaction energy, i.e., $E_N = E(123\ldots N) - \sum_{A=1}^{N} E(A)$. In addition, the sixth and the eighth column display the correlation contributions to the two-body and all-body interaction energies, respectively.

As can be seen in Table 1, even for large cluster sizes the three- and four-body dispersion energies are fairly small compared to the two-body dispersion interaction. Moreover, since they possess alternate signs, they also partially cancel such that their sum amounts to only 1% of the two-body dispersion at a cluster size of $N = 13$. One can therefore conclude that many-body dispersion effects are negligible for the description of water clusters. This is also illustrated in the diagram in Figure 10 in which the three- and four-body dispersion interaction is plotted along with the total many-body correlation interaction $\Delta^{3-N}_{\text{corr}}$, defined by the difference of the $E^{\text{all-body}}_{\text{corr}}$ and $E^{\text{two-body}}_{\text{corr}}$ terms

Table 1: Two-, three- and four-body dispersion energies and two-body and total (all-body) interaction energies of the water clusters. All energies are in kcal/mol.

$(H_2O)_x$	dispersion			two-body		all-body	
	$E_{disp}^{two\text{-}body}$	$E_{disp}^{three\text{-}body}$	$E_{disp}^{four\text{-}body}$	$E_{total}^{two\text{-}body}$	$E_{corr}^{two\text{-}body}$	$E_{total}^{all\text{-}body}$	$E_{corr}^{all\text{-}body}$
2	−2.28			−4.47	−1.24	−4.47	−1.24
3	−8.02	0.06		−12.22	−4.47	−13.76	−4.31
4	−13.21	0.07	−0.01	−20.27	−6.88	−24.27	−6.61
5	−16.89	0.04	−0.03	−25.62	−8.33	−32.05	−8.17
6 (cage)	−24.09	0.31	−0.04	−35.12	−13.79	−40.35	−12.78
6 (prism)	−24.69	0.33	−0.05	−35.53	−14.25	−40.72	−13.17
6 (ring)	−20.33	0.01	−0.01	−31.63	−10.06	−39.72	−9.60
7	−29.86	0.36	−0.06	−43.64	−17.00	−50.71	−15.74
8	−37.72	0.48	−0.09	−55.15	−21.64	−64.06	−19.99
9	−41.59	0.44	−0.08	−61.60	−23.53	−72.36	−21.72
10	−47.41	0.49	−0.12	−70.27	−27.06	−82.32	−24.82
11	−53.43	0.79	−0.12	−76.72	−31.97	−87.45	−28.86
12	−61.96	0.91	−0.18	−89.79	−37.14	−102.94	−33.33
13	−64.94	0.92	−0.19	−94.26	−38.88	−107.82	−34.93

from Table 1. Here one can see that the total many-body correlation interaction, like the many-body dispersion energy, is repulsive and is a significant contribution to the total many-body interaction energy at larger cluster sizes. For instance, for $N = 13$ the correlation contribution reduces the overall attractive many-body interaction energy by 22%.

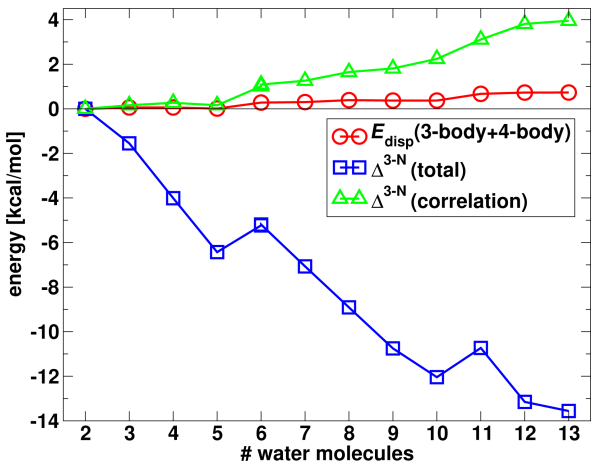


Figure 10: Sum of three- and four-body dispersion interactions compared to the total many-body interactions Δ^{3-N} in the water clusters.

Yet, how significant are many-body interactions compared to the sum of the two-body interactions in the water clusters? From the results shown in Table 1 one finds that the many-body interactions lead to a lowering of the interaction energy of about 13% for all clusters on average. The only exception to this is found for the cyclic structures of the water tetramer, pentamer and hexamer for which the many-body interaction contributes even 20–25% to the total interaction energy. This strong increase of the many-body interactions in the cyclic structures of water clusters is well known [61]. More recently, Bates et al. [72], Hincapie et al. [73] and Chen et al. [74] have performed high-level coupled-cluster calculations for various isomers of $(H_2O)_6$. These more recent investigations support the findings of earlier studies of the water hexamer [62,63] that the cyclic structures are less favorable than the prism and cage forms in spite of the strong many-body interaction contribution. The results for the three hexamer structures studied in this work indicate why the ring form is less stable than the other two structures. We find a considerably lower stabilisation of the ring-hexamer due to two-body dispersion and (thereby) total two-body interactions by about 4 kcal/mol compared to the cage and prism isomers, see Table 1. This result agrees well with the local molecular orbital energy decomposition analysis of the MP2 interaction energies for the corresponding hexamer structures by Chen and co-workers [74]. Yet, the total interaction energies for the three structures are within a range of 1 kcal/mol (note that compared to this Bates et al. found that the cyclic-chair structure is more unstable by +1.83 kcal/mol than the prism structure using the CCSD(T) method [72]). Thus, the decrease of the two-body interaction energy and the increase of the many-body interaction when switching from the cage/prism form to the ring form of $(H_2O)_6$ almost cancel. Apparently, the water hexamer is the first cluster where the two-body interactions start to dominate the global shape of the cluster geometry,

favoring more compact structures than noncompact ring structures.

The above analyis shows that while the many-body interactions in the water clusters are a significant contribution to the total interaction their contribution is scalable and could be modeled, e.g., by scaling the sum of the pair interactions to describe the interactions in the system effectively. This is an approach that is used also in many force field parametrisations for water.

In Figure 11 the ratio of the two-body dispersion energy over the two-body and total interaction energies (including the two-body energies) is plotted. One can see that for all cluster sizes the two-body dispersion amounts to about 60% to the total interaction energy and even almost 70% to the total two-body interaction. This clearly demonstrates the significance of dispersion interactions for the stabilisation of the water clusters. The almost constant dependence of the ratio on the cluster size again demonstrates that the increase of the dispersion interaction energy is very similar to the increase of the total interaction energy. Again, this shows that the magnitude of the total interaction energy of the water clusters (in their equilibrium) could be approximated well by a pair-interaction model.

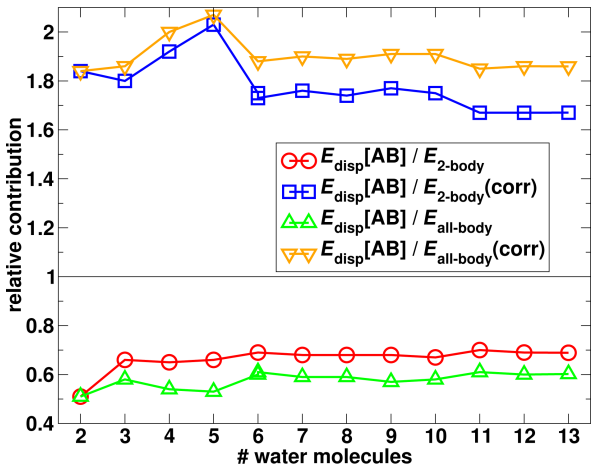


Figure 11: Relative contribution of the two-body dispersion interaction energy to the complete two-body and total (all-body) interaction energies.

Conclusion

Different interaction energy contributions to the total interaction energy of water clusters have been analyzed in this work. It has been shown that the main orientation of the water clusters, characterised by the formation of hydrogen bonds, can be well reproduced already on an uncorrelated level using the Hartree–Fock method. However, electron correlation effects to the interaction energy, including the two-body dispersion interaction, lead to a compression of the cluster sizes relative to the structures optimized with the Hartree–Fock method. This global effect originates from the almost isotropic character of the two-body dispersion energy.

Compared to this, the three-body dispersion interaction energy is more strongly dependent on the orientation of the water molecules. However, its was found that dispersion interactions beyond the two-body level are negligible for the description of the stability of water clusters.

The main contributions to the many-body interactions (beyond the two-body level) are described by higher order polarisation interactions, in line with previous studies of the interactions in water clusters [28,29]. This was found through a decomposition of the interaction energies using the DFT-SAPT method with the aid of the pseudodimer technique. Overall, many-body interactions are quite significant and contribute about 13% to the total interaction energy of the water clusters. This amount was found to be almost independent on the cluster size. Because of this, many-body interactions in water should be accurately reproduceable by properly scaled two-body interaction energy terms. Since, however, the many-body polarisation interactions, like their two-body counterparts, may be very anisotropic, this approach needs to be carefully tested also for nonequilibrium structures not considered in this work.

It can generally be concluded that electron correlation effects, including dispersion interactions, are crucial for the description of the two-body interactions in water clusters, yet they yield comparably smaller contributions to the many-body interactions. Efficient computational approaches that are based on the many-body expansion to describe the interactions in water may therefore restrict the description of electron correlation to the two-body level without a severe loss in accuracy.

Acknowledgements

Financial support of this work through the DFG (Deutsche Forschungsgemeinschaft) priority Program No. SPP1807 “Control of London dispersion interactions in molecular chemistry” is gratefully acknowledged.

References

1. Ludwig, R. *Phys. Chem. Chem. Phys.* **2002,** *4,* 5481. doi:10.1039/B207000F
2. Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983,** *79,* 926. doi:10.1063/1.445869
3. Horn, H. W.; Swope, W. C.; Pitera, J. W.; Madura, J. D.; Dick, T. J.; Hura, G. L.; Head-Gordon, T. *J. Chem. Phys.* **2004,** *120,* 9665. doi:10.1063/1.1683075
4. Mahoney, M. W.; Jorgensen, W. L. *J. Chem. Phys.* **2000,** *112,* 8910. doi:10.1063/1.481505

5. Rick, S. W. *J. Chem. Phys.* **2004,** *120,* 6085. doi:10.1063/1.1652434
6. Fanourgakis, G. S.; Xantheas, S. S. *J. Chem. Phys.* **2008,** *128,* 074506. doi:10.1063/1.2837299
7. Burnham, C. J.; Anick, D. J.; Mankoo, P. K.; Reiter, G. F. *J. Chem. Phys.* **2008,** *128,* 154519. doi:10.1063/1.2895750
8. Ren, P. Y.; Ponder, J. W. *J. Phys. Chem. B* **2004,** *108,* 13427. doi:10.1021/jp0484332
9. Ponder, J. W.; Wu, C.; Ren, P.; Pande, V. S.; Chodera, J. D.; Schnieders, M. J.; Haque, I.; Mobley, D. L.; Lambrecht, D. S.; DiStasio, R. A., Jr.; Head-Gordon, M.; Clark, G. N. I.; Johnson, M. E.; Head-Gordon, T. *J. Phys. Chem. B* **2010,** *114,* 2549. doi:10.1021/jp910674d
10. Kumar, R.; Wang, F.-F.; Jenness, G. R.; Jordan, K. D. *J. Chem. Phys.* **2010,** *132,* 014309. doi:10.1063/1.3276460
11. Mas, E. M.; Szalewicz, K.; Bukowski, R.; Jeziorski, B. *J. Chem. Phys.* **1997,** *107,* 4207. doi:10.1063/1.474795
12. Mas, E. M.; Bukowski, R.; Szalewicz, K.; Groenenboom, G. C.; Wormer, P. E. S.; van der Avoird, A. *J. Chem. Phys.* **2000,** *113,* 6687. doi:10.1063/1.1311289
13. Cencek, W.; Szalewicz, K.; Leforestier, C.; van Harrevelt, R.; van der Avoird, A. *Phys. Chem. Chem. Phys.* **2008,** *10,* 4716. doi:10.1039/b809435g
14. Torheyden, M.; Jansen, G. *Mol. Phys.* **2006,** *104,* 2101. doi:10.1080/00268970600679188
15. Cisneros, G. A.; Wikfeldt, K. T.; Ojamäe, L.; Lu, J.; Xu, Y.; Torabifard, H.; Bartok, A. P.; Csányi, G.; Molinero, V.; Paesani, F. *Chem. Rev.* **2016,** *116,* 7501. doi:10.1021/acs.chemrev.5b00644
16. Ren, P.; Ponder, J. W. *J. Comput. Chem.* **2002,** *23,* 1497. doi:10.1002/jcc.10127
17. Burnham, C. J.; Li, J.; Xantheas, S. S.; Leslie, M. *J. Chem. Phys.* **1999,** *110,* 4566. doi:10.1063/1.478797
18. Gresh, N.; Claverie, P.; Pullman, A. *Theor. Chim. Acta* **1984,** *66,* 1. doi:10.1007/BF00577135
19. Piquemal, J.-P.; Cisneros, G. A.; Reinhardt, P.; Gresh, N.; Darden, T. A. *J. Chem. Phys.* **2006,** *124,* 104101. doi:10.1063/1.2173256
20. Piquemal, J.-P.; Chevreau, H.; Gresh, N. *J. Chem. Theory Comput.* **2007,** *3,* 824. doi:10.1021/ct7000182
21. Day, P. N.; Jensen, J. H.; Gordon, M. S.; Webb, S. P.; Stevens, W. J.; Krauss, M.; Garmer, D.; Basch, H.; Cohen, D. *J. Chem. Phys.* **1996,** *105,* 1968. doi:10.1063/1.472045
22. Guidez, E. B.; Gordon, M. S. *J. Phys. Chem. A* **2017,** *121,* 3736. doi:10.1021/acs.jpca.6b11403
23. Maheshwary, S.; Patel, N.; Sathyamurthy, N.; Kulkarni, A. D.; Gadre, S. R. *J. Phys. Chem. A* **2001,** *105,* 10525. doi:10.1021/jp013141b
24. Kulkarni, A. D.; Ganesh, V.; Gadre, S. R. *J. Chem. Phys.* **2004,** *121,* 5043. doi:10.1063/1.1780156
25. Stone, A. *The Theory of Intermolecular Forces,* 2nd ed.; Oxford University Press, 2013.
26. Mas, E. M.; Bukowski, R.; Szalewicz, K. *J. Chem. Phys.* **2003,** *118,* 4404. doi:10.1063/1.1542872
27. Hodges, M. P.; Stone, A. J.; Xantheas, S. S. *J. Phys. Chem. A* **1997,** *101,* 9163. doi:10.1021/jp9716851
28. Milet, A.; Moszynski, R.; Wormer, P. E. S.; van der Avoird, A. *J. Phys. Chem. A* **1999,** *103,* 6811. doi:10.1021/jp990773d
29. Podeszwa, R.; Szalewicz, K. *J. Chem. Phys.* **2007,** *126,* 194101. doi:10.1063/1.2733648
30. Schütz, M.; Bürgi, T.; Leutwyler, S.; Bürgi, H. B. *J. Chem. Phys.* **1993,** *99,* 5228. doi:10.1063/1.465991
31. Hapka, M.; Rajchel, Ł.; Modrzejewski, M.; Schäffer, R.; Chałasiński, G.; Szczęśniak, M. M. *J. Chem. Phys.* **2017,** *147,* 084106. doi:10.1063/1.4986291
32. Misquitta, A. J.; Stone, A. J. *J. Chem. Theory Comput.* **2008,** *4,* 7. doi:10.1021/ct700104t
33. Heßelmann, A.; Jansen, G. *Chem. Phys. Lett.* **1999,** *315,* 248. doi:10.1016/S0009-2614(99)01251-8
34. Heßelmann, A.; Jansen, G. *J. Chem. Phys.* **2000,** *112,* 6949. doi:10.1063/1.481309
35. Heßelmann, A.; Görling, A. *Mol. Phys.* **2010,** *108,* 359. doi:10.1080/00268970903476662
36. Heßelmann, A.; Görling, A. *Mol. Phys.* **2011,** *109,* 2473. doi:10.1080/00268976.2011.614282
37. Bleiziffer, P.; Heßelmann, A.; Görling, A. *J. Chem. Phys.* **2012,** *136,* 134102. doi:10.1063/1.3697845
38. Talman, J. D.; Shadwick, W. F. *Phys. Rev. A* **1976,** *14,* 36. doi:10.1103/PhysRevA.14.36
39. Görling, A. *Phys. Rev. Lett.* **1999,** *83,* 5459. doi:10.1103/PhysRevLett.83.5459
40. Hirata, S.; Ivanov, S.; Grabowski, I.; Bartlett, R. J.; Burke, K.; Talman, J. D. *J. Chem. Phys.* **2001,** *115,* 1635. doi:10.1063/1.1381013
41. Holas, A.; Cinal, M. *Phys. Rev. A* **2005,** *72,* 032504. doi:10.1103/PhysRevA.72.032504
42. Heaton-Burgess, T.; Bulat, F. A.; Yang, W. *Phys. Rev. Lett.* **2007,** *98,* 256401. doi:10.1103/PhysRevLett.98.256401
43. Heßelmann, A.; Götz, A. W.; Della Sala, F.; Görling, A. *J. Chem. Phys.* **2007,** *127,* 054102. doi:10.1063/1.2751159
44. Heßelmann, A.; Jansen, G. *Chem. Phys. Lett.* **2002,** *357,* 464. doi:10.1016/S0009-2614(02)00538-9
45. Heßelmann, A.; Jansen, G. *Chem. Phys. Lett.* **2002,** *362,* 319. doi:10.1016/S0009-2614(02)01097-7
46. Misquitta, A.; Szalewicz, K. *Chem. Phys. Lett.* **2002,** *357,* 301. doi:10.1016/S0009-2614(02)00533-X
47. Heßelmann, A.; Jansen, G. *Chem. Phys. Lett.* **2003,** *367,* 778. doi:10.1016/S0009-2614(02)01796-7
48. Misquitta, A. J.; Jeziorski, B.; Szalewicz, K. *Phys. Rev. Lett.* **2003,** *91,* 033201. doi:10.1103/PhysRevLett.91.033201
49. Heßelmann, A.; Jansen, G. *Phys. Chem. Chem. Phys.* **2003,** *5,* 5010. doi:10.1039/b310529f
50. Heßelmann, A.; Jansen, G.; Schütz, M. *J. Chem. Phys.* **2005,** *122,* 014103. doi:10.1063/1.1824898
51. Misquitta, A. J.; Podeszwa, R.; Jeziorski, B.; Szalewicz, K. *J. Chem. Phys.* **2005,** *123,* 214103. doi:10.1063/1.2135288
52. Jansen, G. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2013,** *4,* 127. doi:10.1002/wcms.1164
53. Görling, A. *Int. J. Quantum Chem.* **1998,** *69,* 265. doi:10.1002/(SICI)1097-461X(1998)69:3<265::AID-QUA6>3.0.CO;2-T
54. Görling, A. *Phys. Rev. A* **1998,** *57,* 3433. doi:10.1103/PhysRevA.57.3433
55. Hirata, S.; Ivanov, S.; Grabowski, I.; Bartlett, R. J. *J. Chem. Phys.* **2002,** *116,* 6468. doi:10.1063/1.1460869
56. Shigeta, Y.; Hirao, K.; Hirata, S. *Phys. Rev. A* **2006,** *73,* 010502. doi:10.1103/PhysRevA.73.010502
57. Heßelmann, A.; Ipatov, A.; Görling, A. *Phys. Rev. A* **2009,** *80,* 012507. doi:10.1103/PhysRevA.80.012507
58. Ipatov, A.; Heßelmann, A.; Görling, A. *Int. J. Quantum Chem.* **2010,** *110,* 2202. doi:10.1002/qua.22561
59. Heßelmann, A.; Görling, A. *J. Chem. Phys.* **2011,** *134,* 034120. doi:10.1063/1.3517312

60. Chałasiński, G.; Szcześniak, M. M. *Mol. Phys.* **1988,** *63,* 205. doi:10.1080/00268978800100171
61. Xantheas, S. S. *J. Chem. Phys.* **1994,** *100,* 7523. doi:10.1063/1.466846
62. Mhin, B. J.; Kim, J.; Lee, S.; Lee, J. Y.; Kim, K. S. *J. Chem. Phys.* **1994,** *100,* 4484. doi:10.1063/1.466279
63. Gregory, J. K.; Clary, D. C. *J. Phys. Chem. C* **1996,** *100,* 18014. doi:10.1021/jp9616019
64. Curtiss, L. A.; Raghavachari, K.; Redfern, P. C.; Rassolov, V.; Pople, J. A. *J. Chem. Phys.* **1998,** *109,* 7764. doi:10.1063/1.477422
65. Dunning, T. H., Jr. *J. Chem. Phys.* **1989,** *90,* 1007. doi:10.1063/1.456153
66. Bak, K. L.; Jørgensen, P.; Olsen, J.; Helgaker, T.; Klopper, W. *J. Chem. Phys.* **2000,** *112,* 9229. doi:10.1063/1.481544
67. Boys, S. F.; Bernardi, F. *Mol. Phys.* **1970,** *19,* 553. doi:10.1080/00268977000101561
68. *MOLPRO, a package of ab initio programs*, 2015.1;TTI - Technologie-Transfer-Initiative GmbH an der Universität Stuttgart: Stuttgart, Germany, 2015. see http://www.molpro.net
69. Werner, H.-J.; Knowles, P. J.; Knizia, G.; Manby, F. R.; Schütz, M. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2012,** *2,* 242. doi:10.1002/wcms.82
70. Weigend, F.; Ahlrichs, R. *Phys. Chem. Chem. Phys.* **2005,** *7,* 3297. doi:10.1039/b508541a
71. Axilrod, B. M.; Teller, E. *J. Chem. Phys.* **1943,** *11,* 299. doi:10.1063/1.1723844
72. Bates, D. M.; Tschumper, G. S. *J. Phys. Chem. A* **2009,** *113,* 3555. doi:10.1021/jp8105919
73. Hincapie, G.; Acelas, N.; Castaño, M.; David, J.; Restrepo, A. *J. Phys. Chem. A* **2010,** *114,* 7809. doi:10.1021/jp103683m
74. Chen, Y.; Li, H. *J. Phys. Chem. A* **2010,** *114,* 11719. doi:10.1021/jp104822e

License and Terms

BEILSTEIN JOURNAL OF ORGANIC CHEMISTRY

Are dispersion corrections accurate outside equilibrium? A case study on benzene

Tim Gould*[1], Erin R. Johnson[2] and Sherif Abdulkader Tawfik[3,4]

Full Research Paper

Open Access

Address:
[1]Queensland Micro- and Nanotechnology Centre, Griffith University, Nathan, Queensland 4111, Australia, [2]Department of Chemistry, Dalhousie University, Halifax, Nova Scotia, B3H 4R2, Canada, [3]School of Mathematical and Physical Sciences, University of Technology Sydney, Ultimo, New South Wales 2007, Australia and [4]Institute for Biomedical Materials and Devices (IBMD), Faculty of Science, University of Technology Sydney, Sydney, NSW, Australia

Email:
Tim Gould* - t.gould@griffith.edu.au

* Corresponding author

Keywords:
benzene; DFT; dispersion; van der Waals

Beilstein J. Org. Chem. **2018,** *14,* 1181–1191.
doi:10.3762/bjoc.14.99

Received: 22 February 2018
Accepted: 30 April 2018
Published: 23 May 2018

This article is part of the Thematic Series "Dispersion interactions".

Guest Editor: P. Schreiner

Abstract

Modern approaches to modelling dispersion forces are becoming increasingly accurate, and can predict accurate binding distances and energies. However, it is possible that these successes reflect a fortuitous cancellation of errors at equilibrium. Thus, in this work we investigate whether a selection of modern dispersion methods agree with benchmark calculations across several potential-energy curves of the benzene dimer to determine if they are capable of describing forces and energies outside equilibrium. We find the exchange-hole dipole moment (XDM) model describes most cases with the highest overall agreement with reference data for energies and forces, with many-body dispersion (MBD) and its fractionally ionic (FI) variant performing essentially as well. Popular approaches, such as Grimme-D and van der Waals density functional approximations (vdW-DFAs) underperform on our tests. The meta-GGA M06-L is surprisingly good for a method without explicit dispersion corrections. Some problems with SCAN+rVV10 are uncovered and briefly discussed.

Introduction

The past decade has seen an increasing awareness of the role played by van der Waals dispersion forces in chemistry and materials science [1-6]. It has consequently become firmly established that including dispersion forces can be vital for understanding and predicting the behaviour and structure of molecules, materials and surfaces [7-12].

The increased attention being paid to dispersion forces has paralleled, and been driven by, an increased interest in how to accurately model them. Multiple families of approaches for including dispersion forces in quantum chemical simulations now exist, mostly based around the principle of improving density functional theory (DFT) calculations (see, e.g., some key and recent summaries [3,5,13,14]) through a dispersion correction. The latest variants of these approaches have been highly successful in predicting key properties of a wide range of molecules and materials, such as binding energies and molecular/material structures [4,5,15]. Methods

are increasingly converging towards accurate prediction of these properties [16].

What is less well known, however, is how well methods predict the properties of systems outside of internal equilibrium, i.e., whether they can predict energies and forces when a system has not relaxed to its lowest energy geometry. This question is important as it is feasible that methods benefit from a cancellation of errors at equilibrium, which may give false expectations about their general accuracy. We therefore need to understand the limitations of approaches in dealing with dispersion forces generally, and not just for systems at their equilibrium geometries. This is especially important for predicting how a system (or sub-system) behaves when subject to external forces, or when dispersion forces compete with other weak forces within molecules or structures. It is particularly relevant as recent work has shown that modern approaches can often provide accurate binding distances or binding energies in layered materials [17], but not both, suggesting limits to their accuracy.

The work by Řezáč et al. goes some way to resolving this question, by providing benchmark values (the S66x8 set) for eight equilibrium and non-equilibrium geometries of different molecular pairs [18]. This set has been used to test various dispersion methods [19]. However, while S66x8 certainly improves on tests only at optimized geometries, it may still fail to expose issues with forces or other energetic differences.

In this work we thus test the accuracy of modern dispersion approaches in reproducing the energetics of the benzene dimer, an important model system, in different geometries. The coupled-cluster with singles, doubles and perturbative triples [CCSD(T)] benchmark set of Sinnokrot et al. [20] is used, with the aim of establishing which approximations can best reproduce the full potential energy curves. This simple test is not designed to be comprehensive, but rather to interrogate the predictive ability of different approaches. Note that the benchmark set, while slightly inaccurate by modern standards, is predicted to be within 0.1 kcal/mol of more recent benchmarks [21], which is similar to methodological errors caused by using modern pseudopotential methods [22]. We thus feel that its range more than makes up for any limitations it may have.

Moreover, interactions between ring structures feature widely across organic chemistry [23-25]. Recently, there has been increasing interest in utilizing non-covalent π-stacking for synthetic catalysis – and it is notable that most structures shown in a recent review on the topic feature rings that interact at distances greater than the potential minimum [26]. Benzene dimers also feature in the S22 benchmark set [21,27] that is often used to semi-empirically optimize dispersion corrections, and is almost always used as a test of such methods. They are thus an excellent test of the quality of dispersion models on a system where failures may have chemical relevance.

Results and Discussion

The origin of dispersion forces

Dispersion forces are a semi-classical effect coming from quantum fluctuations. Most simply, they can be viewed from the perspective of pairs of interacting fluctuating dipoles, in which a temporary dipole in one (sub)system induces a dipole in the other, and consequently lowers the total energy. Since the field from each temporary dipole falls off as the inverse cube of the distance D between the systems, and the contributions come in pairs, this leads to an interaction with the asymptotic form $U_{\mathrm{vdW}} = -C_6 D^{-6}$, where C_6 is a coefficient that depends on the properties of the independent systems. A similar semi-classical analysis can also be applied to more general multipoles, such as quadropoles, which give rise to terms proportional to D^{-8} (quadropole with dipole), D^{-10} (quadrupole–quadropole) etc.

In addition to direct coupling between pairs of multipoles, more general forms of coupling can also lead to higher order contributions, including 3rd-order effects between three multipoles, 4th-order etc., as detailed in, e.g., Dobson and Gould [2]. In certain cases, including graphene [28-31], this can lead to fundamental deviations from the simple model outlined above [18,31-36]. However, the most extreme deviations from the pairwise model do not affect the benzene dimer system.

Note, the importance of higher-order (many-body) dispersion terms in "typical" systems has been the subject of some debate. It is critical to differentiate between non-additive many-body electronic interactions [34,35] and non-additive C_9 or Axilrod-Teller-Muto (ATM) dispersion interactions here. The former cause large differences in the effective pairwise C_6 and higher-order dispersion coefficients, relative to corresponding values for free atoms [33,37,38] (these are known as Type-B non-additivity effects in the classification scheme of Dobson [33]). This is a particularly significant effect for metals and can alter the C_6 coefficients by more than an order of magnitude in some cases [39,40]. In contrast, the 3-body C_9 contribution to the dispersion energy is typically smaller in magnitude than the pairwise C_{10} contribution [41] and consequently is negligible for most applications compared to 3-body electronic effects [42-44].

Mathematically, the ATM treatment is most applicable when the energy contribution from 3rd and higher order terms converge rapidly as a function of inverse distance and may thus be truncated after the 2nd-order or 3rd-order contributions. Many-body effects arise from a divergence or slow convergence in the same series due to Dobson-B or -C effects, so that

the contributions must be treated as a formal power series and rewritten as an explicit function of the polarisability tensor.

Summary of modern dispersion methods

Over the last decade, a number of new approaches have been developed that explicitly introduce dispersion forces into electronic structure theory methods – typically density functional theory (DFT). These approaches seek to overcome the fundamental lack of dispersion forces in DFT and Hartree–Fock theory by introducing an explicit long-range correction, giving a total energy $E = E_{DFT} + \Delta E_{vdW}$ for the system. Typically, E_{DFT} is taken from a standard density functional approximation (DFA), such as PBE [45] or B3LYP [37], while ΔE_{vdW} is one of a range of dispersion correction models. Note, this is different to seamless approaches like MP2, RPA or other quantum chemistry methods which include dispersion forces automatically.

Common van der Waals corrections can be broadly divided into three categories, as will be detailed below. Substantial effort has seen steady improvements in the quality of approaches in all three categories. In this paper we focus only on recent (or older, but still very popular) iterations within each category, to reflect how the methods are designed to be used in practice. The three classes of approaches considered are:

1) Purely empirical corrections based only on semi-classical models of the nuclei, and their neighbours, without drawing from the electronic density [15,46-50]. Of these we include Grimme's D2, D3 and D3-BJ functionals, as corrections on PBE and, in the case of D3-BJ, on B3LYP. Here and henceforth "on X" means that the correction is taken on top of the X (hybrid) density functional approximation, which we also denote as X-Y (e.g., PBE-D3-BJ);

2) atomic-dipole with density methods, which correct first-principles or empirical models of atomic dipoles (and sometimes multipoles) using properties of the electronic density. Of these we include XDM [51,52] (on various DFAs), TS [53] (on PBE), TS-MBD@rsRSC (MBD for short, on PBE) and FI-MBD@rsRSC (FI for short, on PBE). Both MBD [54,55] and FI [56,57] include dispersion contributions to all orders using the many-body dispersion method of Tkatchenko et al. [54], but involve different treatment of polarisabilities and screening; and

3) first principles density functionals, in which dispersion forces depend only on the density in a totally seamless fashion [58-60] and in which the base DFA forms part of the functional itself. We include the functional of Dion et al. [59], vdW-DF2 [61], optPBEvdW [62] and optB88vdW [62]. We also include SCAN+rVV10, based on the strongly-constrained and normalised (SCAN) meta-GGA, which has been shown to be very successful in initial testing [63]. We refer collectively to these as vdW-DFAs.

In addition to the van der Waals functionals, we also show energies from the generalized gradient approximations (GGAs) PBE and B3LYP, and the meta-GGAs M06L [64] and SCAN [65] without dispersion corrections. Although both GGAs are known to neglect dispersion physics, the meta-GGAs M06-L and SCAN are expected to capture some dispersion-binding contributions through the large-gradient behaviour of their exchange functionals. Note, the interplay between exchange-correlation functionals and dispersion corrections has been the topic of some discussion [66,67]. Finally, we note that we include only general functionals, and avoid approaches that are designed to address one type of system (e.g., molecules, bulks or layered structures) only.

Calculation details

We performed most calculations using VASP 5.4 [68,69] where the valence electrons are separated from the core by use of projector-augmented wave pseudopotentials (PAW) [70]. The energy tolerance for the electronic structure determinations was set at 10^{-7} eV. Calculations used only the Γ $\boldsymbol{k}$-point. ENCUT was set to 400 eV in all calculations, which were carried out in a 15 × 15 × 25 Å^3 supercell. SCAN(+rVV10) required ENCUT = 700 eV as results showed significant noise with the standard energy cutoff, which led us to reduce the box dimensions to 12 × 12 × 20 Å^3. We will discuss issues with SCAN later. Both MBD approaches (TS-MBD and FI-MBD) used the reciprocal space implementation [71], the latter in a custom version of VASP 5.4.1 [57]. The vdW-DFs use the implementation of Klimeš [72] of the Román-Pérez and Soler [73] method.

Some methods are not implemented in VASP and in these cases, the calculations were performed using other codes. XDM results were obtained using Gaussian09 (PBE, B3LYP, and LC-wPBE) or Psi4 [74] (B86bPBE) and the postg application. We include XDM results on several base functionals due to its broad success. M06-L results were calculated with Gaussian16 using the aug-cc-pVTZ basis set due to convergence difficulties with the plane-waves/pseudopotential approach.

To put these settings in context, we purposefully employed the methods as they are intended to be used, i.e., using more-or-less standard convergence parameters and recommended settings.

Results

Now that we have established the background methodology, let us summarise the shared features of Figures 1–3 to aid in

detailed assessment. Each figure is composed of sub-figures showing results for selected groupings of methods. Each sub-figure shows as solid lines the benchmark potential energy curve U_{bench} and the potential energy curves from the selected methods U_{method}. They also show the benchmark force F_{bench}, and the forces for different methods F_{method}, all in dashes. All energies and forces are reported as functions of distance, either between the centre of dimers (Figure 1 and Figure 2) or the sliding distance (Figure 3).

We adopt some steps to ensure all energies and forces are calculated in the same way, so as to reduce uncontrolled errors from, e.g., basis set superpositions or pseudopotentials. Firstly, we use the electronic structure codes to calculate energies $E(R)$ directly. We then fit $E(R) = E_\infty - C_6/R^6$ to the last five points of the parallel configuration data to find E_∞, the extrapolated energy of two monomers, which lets us determine interaction energies $U(R) = E(R) - E_\infty$. We plot $U(R)$ in Figure 1 and Figure 2, and use the minimum-energy values directly in Table 1 – Figure 3 shows $U(R) = E(R) - E(0)$. Secondly, we obtain all forces by fitting cubic splines through the energy data, and taking the derivative of the splines.

Figure 1 shows the interaction energy for the parallel configuration of the benzene dimer (labelled P – with D_{6h} symmetry). Despite having a minimum as a function of distance between the two centres, this arrangement is unstable as the dimers wish to slide apart sideways (see later discussion on Figure 3) to reduce electrostatic effects, such as overlap of the densities of the monomers and static quadrupole–quadropole interactions, which make metastable AA graphite ≈0.23 kcal/mol/C less energetically favourable compared to AB graphite [75]. This configuration thus involves competition between dispersion forces, repulsive electrostatic forces, and other exchange and correlation effects, making it a good test of dispersion corrections.

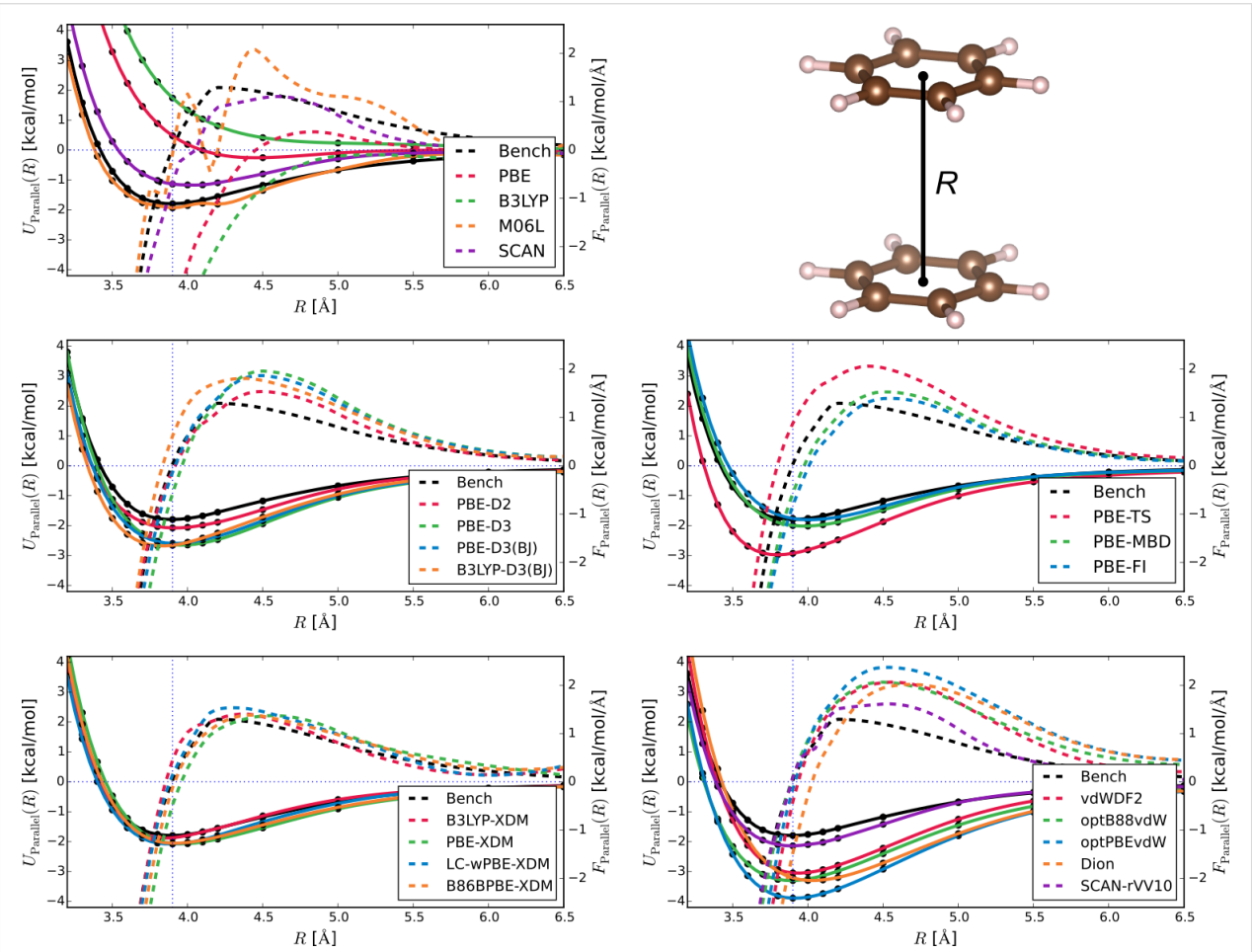


Figure 1: Interaction energies (solid lines) and forces (dashed lines) for the parallel configuration of the benzene dimer. Each panel groups a different family of computational approach. The top row (from left to right) shows GGAs and meta-GGAs without dispersion corrections, and the dimer geometry. The second row reports Grimme-D variants (l) and TS/MBD variants (r). The bottom row shows XDM on different DFAs (l), and vdW-DFAs (r). The benchmark data is always shown in black.

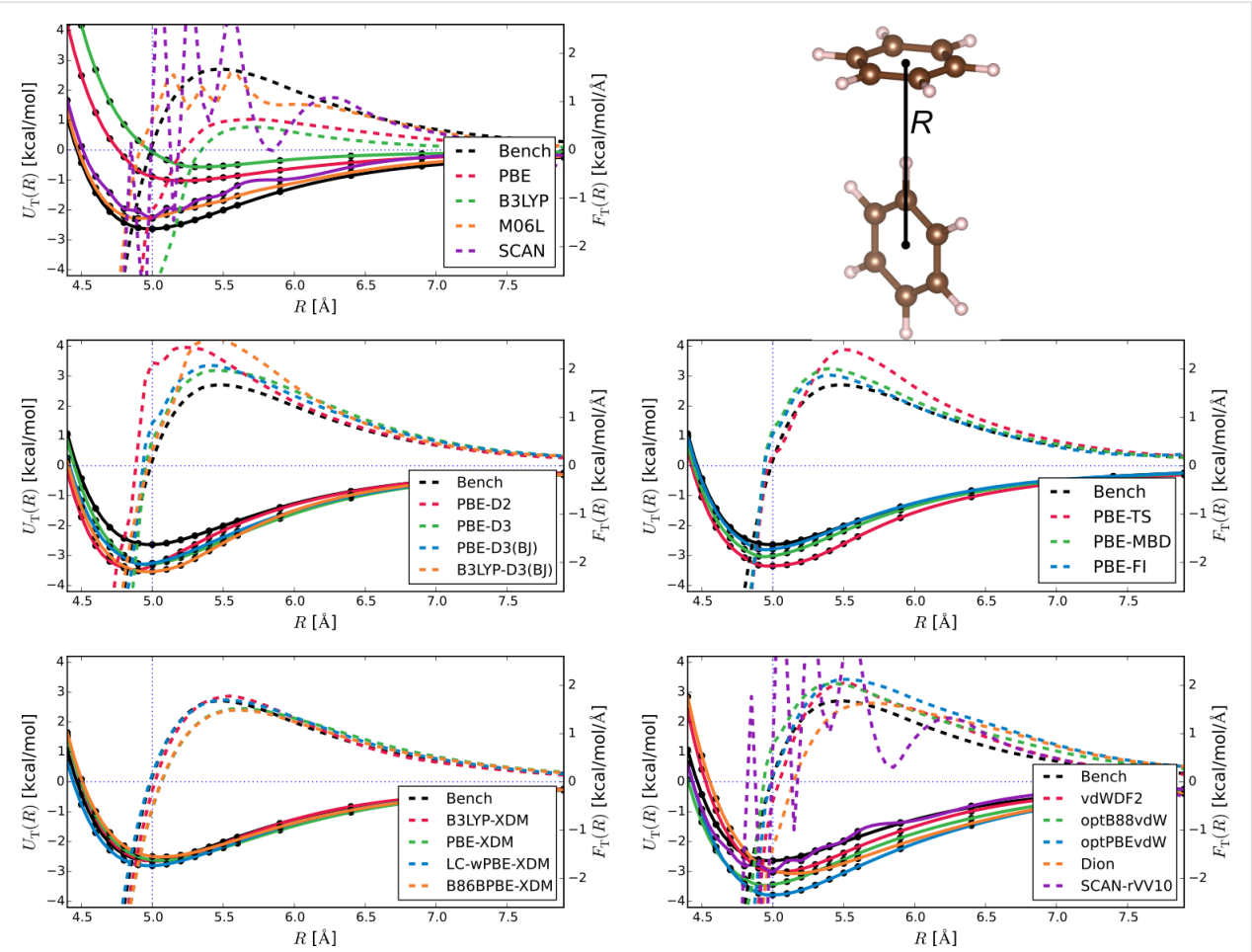


Figure 2: Interaction energies (solid lines) and forces (dashed lines) for the T configuration of the benzene dimer. Panels are the same as in Figure 1.

It is clear from the figure that D2, XDM (all variants), MBD and FI all give reliable energies across the entire curve. Their forces are slightly worse, but still within 0.5 kcal/mol/Å of the reference data at reasonable intermolecular distances. The more modern Grimme variants fare worse than their older cousin, and none of the two-point vdW-DFAs work very well at all, for energies or forces, except near the minima. Indeed, most of the tested vdW-DFAs give force errors outside equilibrium that are similar in magnitude to the force itself. A notable exception is SCAN+rVV10 which is broadly on a par with XDM and TS/FI-MBD. Somewhat surprisingly, the semi-local meta-GGA M06L gives an energy curve which is also in broad agreement with the benchmark, but which fluctuates [76] making the spline-derived forces less reliable. Other dispersion-free methods are less successful, as expected.

Figure 2 then reports the energies for T configuration as a function of distance (T – with C_{2v} symmetry), which includes the global minimum for benzene dimer interactions, or at least a local minimum that is energetically very close to it. Here the balance of energetic contributions is more strongly skewed to dispersion, and it is expected that vdW dispersion corrections should work better than for the parallel configuration shown in Figure 1.

Indeed, the successful methods for the parallel geometry (XDM, MBD, FI) seem to work very well here, reproducing the reference energies and forces to within methodological error of ≈0.1 kcal/mol. The vdW-DFAs perform slightly better than in the parallel case, as one would hope. D2, however, is quite poor despite its success in the parallel case and its more modern cousins are conversely much better. Again, M06L works well. Here, however, we notice that SCAN shows significant oscillations around the true curve, which SCAN+rVV10 inherits (to be discussed later).

Next, Figure 3 reports the potential energy curves for sliding of parallel benzene molecules relative to one another at a fixed inter-planar distance D, known as the slipped-parallel configuration (SP – with C_{2h} symmetry). We show results for

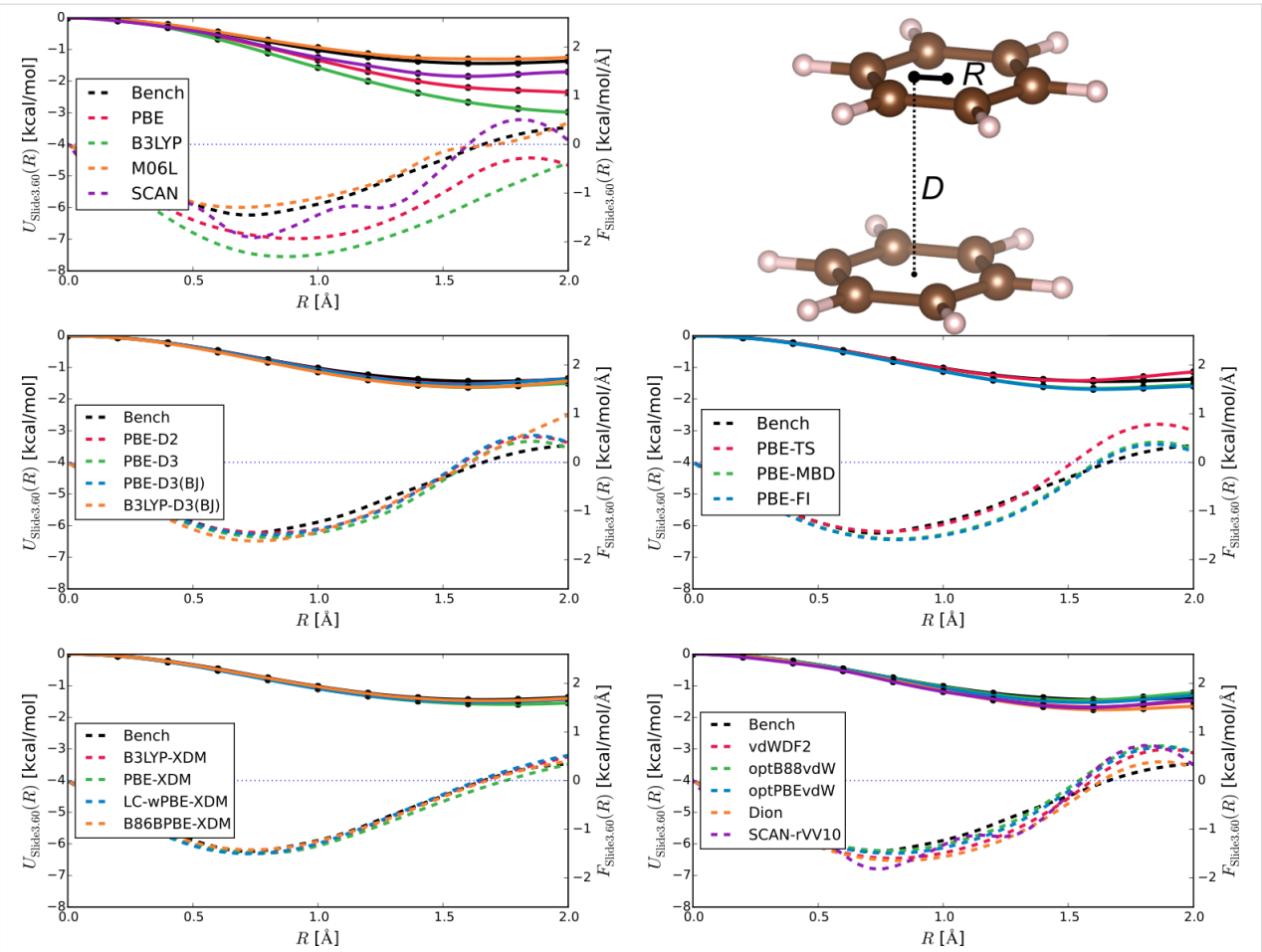


Figure 3: Interaction energies $U(R;D) = E(R;D) - E(0;D)$ (solid lines) and forces (dashed lines) for the slipped-parallel configuration of the benzene dimer at D = 3.6 Å. Panels are the same as in Figure 1.

D = 3.6 Å, shown relative to their energy in the perfectly parallel configuration [i.e., $U(R;D) = E(R;D) - E(0;D)$]. Here XDM is a stand-out, giving almost perfect agreement with the benchmarks, thus indicating its ability to simultaneously capture competing energy contributions. All other methods are much more successful here than in the previous tests, reflecting their consistency in reproducing electrostatic effects compared to dispersion interactions which are more-or-less constant across the curves. These results are replicated in other tests (not shown) at D = 3.2, 3.4 and 3.8 Å. Again, the SCAN and SCAN+rVV10 curves display oscillations.

The strange behaviour of SCAN and SCAN+rVV10 warrants special attention. Previous tests of meta-GGAs using Gaussian-type orbital codes suggest this issue might be related to the density of the real-space grid [76]. In Figure 4 we thus show results for SCAN and SCAN+rVV10 for all four intermolecular distances (D = 3.2, 3.4, 3.6 and 3.8 Å) and for both the large energy cutoff 700 eV used in previous calculations, and also a smaller cutoff of 450 eV.

It is obvious from these curves that the oscillations seem to be smallest when the overlap between the orbitals is largest, as in the parallel case and the two closest (D = 3.2, 3.4) sliding cases, versus the T case and the more distant sliding cases. Furthermore, the oscillations seem to hold consistently for the smaller and larger energy cutoffs, a result we find somewhat perplexing as, if they were sensitive to the grid, we would expect them to decrease with a larger cutoff (and consequently finer grid).

Finally, in Table 1 we quantify how the different methods perform in prediction of the relative energies of the various local minima, U_0(T), U_0(P) and U_0(SP), for the T, parallel (P) and slipped-parallel (SP) configurations, respectively. We thus show the energy differences, ΔU(P/SP) = U_0(P/SP) − U_0(T), between the local minima for the parallel and slipped-parallel configurations, and the (presumed) global minimum for the T configuration. In all cases we fit quadratic curves to data to obtain a value as close to the true minimum as possible. We also take advantage of revised benchmark values from Takatani et al.

Table 1: Relative energy differences, $\Delta U(\mathrm{P/SP}) = U_0(\mathrm{P/SP}) - U_0(\mathrm{T})$ [in kcal/mol], between lowest energies $U_0(\mathrm{T/P/SP})$ for the T, parallel (P) and slipped-parallel (SP) configurations of the benzene dimer, with respect to the minimum-energy T configuration. Here we use the revised benchmarks from Takatani et al. [21] for references, and to quantify the error in our main source of benchmark data [20]. Solid lines separate the different groupings of functionals used in this paper, which are ranked within each section according to |Error|.
$|\mathrm{Error}| = 1/2[|\Delta U(\mathrm{P})_{\mathrm{method}} - \Delta U(\mathrm{P})_{\mathrm{revbench}}| + |\Delta U(\mathrm{SP})_{\mathrm{method}} - \Delta U(\mathrm{SP})_{\mathrm{revbench}}|]$.

	ΔU(P)	ΔU (SP)	|Error|
revBench[a]	0.86	0.11	–
Bench[b]	0.91	−0.01	0.09
SCAN	1.33	0.27	0.32
PBE	0.95	0.82	0.40
M06L	0.51	−0.38	0.42
B3LYP	0.65	1.08	0.59
PBE-D3(BJ)	0.77	−0.16	0.18
PBE-D3	0.71	−0.11	0.19
B3LYP-D3(BJ)	0.80	−0.51	0.34
PBE-D2	1.50	0.28	0.41
PBE-FI	0.93	0.11	0.04
PBE-MBD	0.93	0.03	0.08
PBE-TS	0.43	−0.89	0.71
LC-wPBE-XDM	0.83	−0.14	0.14
B3LYP-XDM	0.83	−0.19	0.17
PBE-XDM	0.59	−0.15	0.27
B86BPBE-XDM	0.48	−0.31	0.40
SCAN-rVV10	1.03	−0.21	0.24
optB88vdW	0.18	−0.72	0.76
vdWDF2	0.06	−0.69	0.80
Dion	−0.27	−0.73	0.98
optPBEvdW	−0.12	−0.90	1.00

[a]From Takatani et al. [21], [b]from Sinnokrot et al. [20].

[21] to establish errors in the main reference data used for the full potential curves.

Here, FI and MBD are the best-performing methods, with absolute errors smaller than those for the older benchmark data set. Variants of XDM (LC-wPBE-XDM, B3LYP-XDM) and some Grimme methods are a little poorer, but are still very good. The vdW-DFAs and PBE-TS method can be quite poor, however, further reflecting their poorer treatment of dispersion energies and forces away from equilibrium.

Conclusion

In this work we used benchmark results for several configurations of the benzene dimer to test the ability of dispersion-corrected density functional theory to obtain accurate energies and forces away from equilibrium, and thus to understand their predictive capabilities. All the methods tested here are backed by previously reported successes on a wide range of chemical and/or materials systems. We have shown that many of them do not match these successes at equilibrium by guaranteed success outside it. In the worst cases, some methods have errors in the predicted forces as large as the force itself.

The exchange-hole dipole moment (XDM) model, which we tested with several DFAs, performs very well in general. PBE-MBD and PBE-FI (which incorporates an improved treatment of polarisabilities into an MBD-like calculation) both perform similarly well. We suspect any of these methods can be reliably trusted for predictions in systems involving benzene ring structures

Grimme's various methods, TS theory, and various two-point van der Waals density functionals are less successful in our tests, however. M06L is surprisingly accurate for a meta-GGA without explicit dispersion corrections, but is numerically noisy [76]. We thus advise caution when using any of these methods

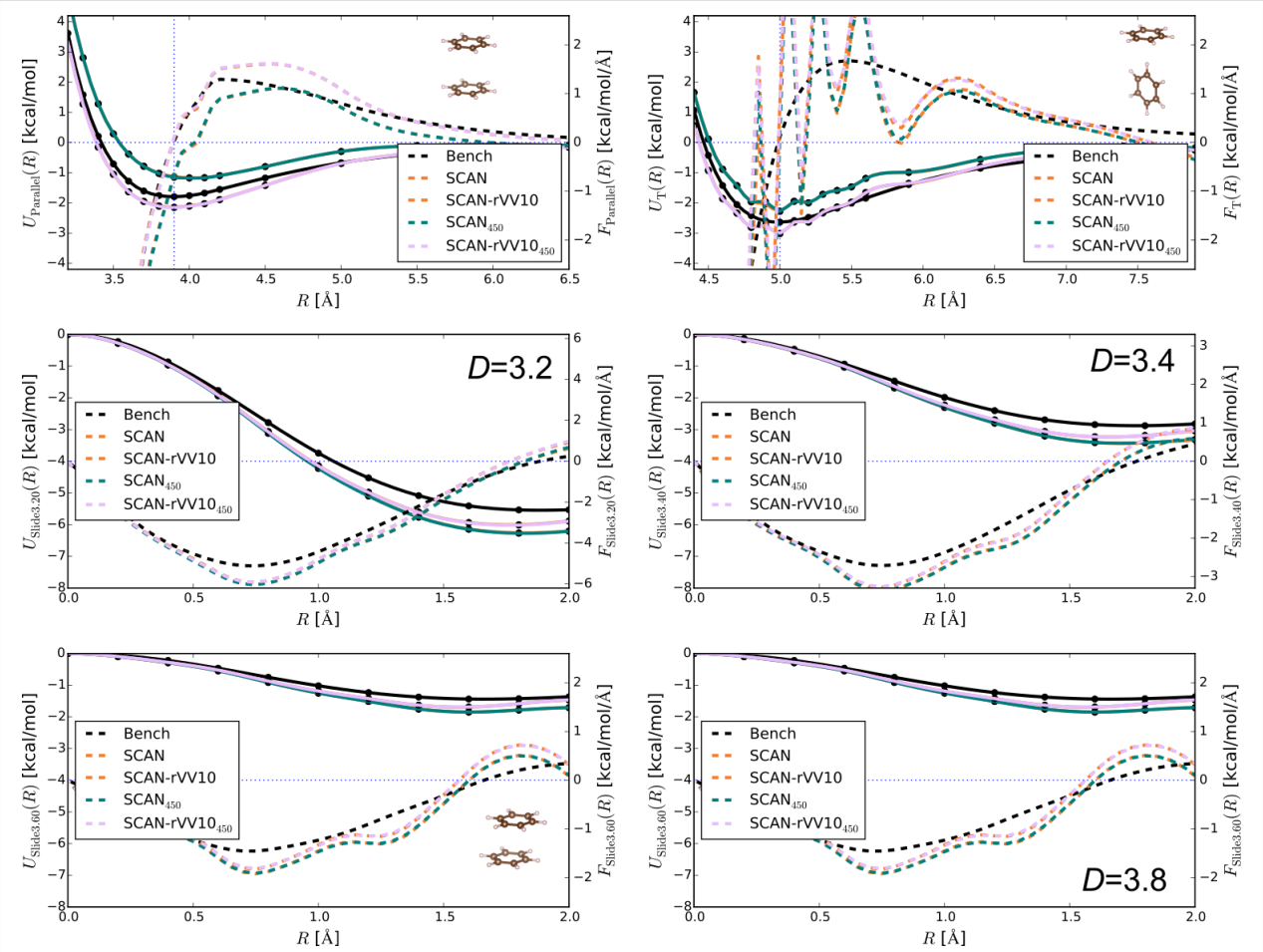


Figure 4: SCAN and SCAN+rVV10 results for the Parallel (top left) and T (top right) configurations of the benzene dimer, as well as for the slipped-parallel configurations at all four intermolecular distances. Results are shown for two different energy cutoffs to test convergence (the default 700 eV and a smaller cutoff of 450 eV). The panels show interaction energies (solid lines) and forces (dashed lines), as in Figure 1.

for systems where interactions between ring structures might be important.

The results for SCAN-rVV10 are troubling. We suspect that the oscillations in the potential energy curves reflect previously reported problems with the integration grid [76]. We could, however, not remove them even with a large energy cutoff of 700 eV, just shy of the value used by its developers for rare gas solids [63]. Also, the results were very similar when a smaller energy cutoff was employed, hinting at a deeper underlying problem. This convergence issue is certainly something that should be investigated before dispersion-corrected SCAN is applied widely to weak-bonding problems.

The results reported here also strongly support the importance of using good polarizabilities (dipole and higher) in dispersion models. XDM, MBD and its FI variant include contributions from both the local density and geometry, and thus can capture type-A and -B non-additivity (the latter semi-locally in the case of XDM), in the classification scheme of Dobson [33]. By contrast, the other methods tested involve more simplistic treatment of environmental contributions to polarisabilities and dispersion coefficients.

Finally, we note that we have only tested one type of molecular dimer which means our conclusions are necessarily limited, despite the benzene dimer being an important and difficult example involving competing contributions to the interaction energies and forces. The results uncovered here are interesting enough, we feel, to establish an impetus to carry out further testing of dispersion forces away from equilibrium and to establish the role of effects beyond dipole pairs (including Axilrod–Teller–Muto terms, quadropole and higher multipole terms, and full many-body terms). We hope this work will drive development and use of new benchmarking sets, along these lines, which can test a wider range of physics and chemistry at and outside of equilibrium, and be used to improve future dispersion models.

Supporting Information

We provide all data generated through this project to allow other researchers to carry out their own analyses. This file contains all reference data used for this paper, stored in comma separated variables format with "#" used to indicate comments.

Supporting Information File 1

All reference data.
[https://www.beilstein-journals.org/bjoc/content/supplementary/1860-5397-14-99-S1.csv]

Acknowledgements

TG acknowledges support of the Griffith University Gowonda HPC Cluster. SAT acknowledges funding by the Australian Government through the Australian Research Council (ARC DP160101301). Theoretical calculations by TG and SAT were undertaken with resources provided by the National Computational Infrastructure (NCI) supported by the Australian Government and, for SAT, by the Pawsey Supercomputing Centre funded by the Australian Government and the Government of Western Australia. ERJ thanks the Natural Sciences and Engineering Research Council (NSERC) of Canada for financial support and the Atlantic Computational Excellence Network (ACEnet) for computing resources.

ORCID® iDs

Tim Gould - https://orcid.org/0000-0002-7191-9124
Erin R. Johnson - https://orcid.org/0000-0002-5651-468X
Sherif Abdulkader Tawfik - https://orcid.org/0000-0003-3592-1419

References

1. Grimme, S. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2011,** *1,* 211–228. doi:10.1002/wcms.30
2. Dobson, J. F.; Gould, T. *J. Phys.: Condens. Matter* **2012,** *24,* 073201. doi:10.1088/0953-8984/24/7/073201
3. Grimme, S.; Hansen, A.; Brandenburg, J. G.; Bannwarth, C. *Chem. Rev.* **2016,** *116,* 5105–5154. doi:10.1021/acs.chemrev.5b00533
4. Woods, L. M.; Dalvit, D. A. R.; Tkatchenko, A.; Rodriguez-Lopez, P.; Rodriguez, A. W.; Podgornik, R. *Rev. Mod. Phys.* **2016,** *88,* 045003. doi:10.1103/RevModPhys.88.045003
5. Hermann, J.; DiStasio, R. A., Jr.; Tkatchenko, A. *Chem. Rev.* **2017,** *117,* 4714–4758. doi:10.1021/acs.chemrev.6b00446
6. Klimeš, J.; Michaelides, A. *J. Chem. Phys.* **2012,** *137,* 120901. doi:10.1063/1.4754130
7. Geim, A. K.; Grigorieva, I. V. *Nature* **2013,** *499,* 419–425. doi:10.1038/nature12385
8. Rösel, S.; Quanz, H.; Logemann, C.; Becker, J.; Mossou, E.; Cañadillas-Delgado, L.; Caldeweyher, E.; Grimme, S.; Schreiner, P. R. *J. Am. Chem. Soc.* **2017,** *139,* 7428–7431. doi:10.1021/jacs.7b01879
9. Fabrizio, A.; Corminboeuf, C. *J. Phys. Chem. Lett.* **2018,** *9,* 464–470. doi:10.1021/acs.jpclett.7b03316
10. Johnson, E. R.; Clarkin, O. J.; Dale, S. G.; DiLabio, G. A. *J. Phys. Chem. A* **2015,** *119,* 5883–5888. doi:10.1021/acs.jpca.5b03251
11. Reimers, J. R.; Li, M.; Wan, D.; Gould, T.; Ford, M. J. *Non-Covalent Interactions in Quantum Chemistry and Physics;* Elsevier, 2017; pp 387–416. doi:10.1016/B978-0-12-809835-6.00015-3
12. Gould, T.; Lebègue, S.; Björkman, T.; Dobson, J. F. In *Semiconductors and Semimetals;* Iacopi, F.; Boeckl, J. J.; Jagadish, C., Eds.; 2D Materials, Vol. 95; Elsevier, 2016; pp 1–33.
13. Goerigk, L. *Non-Covalent Interactions in Quantum Chemistry and Physics;* Elsevier, 2017; pp 195–219. doi:10.1016/B978-0-12-809835-6.00007-4
14. Schröder, E.; Cooper, V. R.; Berland, K.; Lundqvist, B. I.; Hyldgaard, P.; Thonhauser, T. *Non-Covalent Interactions in Quantum Chemistry and Physics;* Elsevier, 2017; pp 241–274. doi:10.1016/B978-0-12-809835-6.00009-8
15. Grimme, S.; Ehrlich, S.; Goerigk, L. *J. Comput. Chem.* **2011,** *32,* 1456–1465. doi:10.1002/jcc.21759
16. Goerigk, L.; Hansen, A.; Bauer, C.; Ehrlich, S.; Najibi, A.; Grimme, S. *Phys. Chem. Chem. Phys.* **2017,** *19,* 32184–32215. doi:10.1039/C7CP04913G
17. Tawfik, S. A.; Gould, T.; Stampfl, C.; Ford, M. J. *Phys. Rev. Materials* **2018,** *2,* 034005. doi:10.1103/PhysRevMaterials.2.034005
18. Řezáč, J.; Riley, K. E.; Hobza, P. *J. Chem. Theory Comput.* **2011,** *7,* 2427–2438. doi:10.1021/ct2002946
19. Goerigk, L.; Kruse, H.; Grimme, S. *ChemPhysChem* **2011,** *12,* 3421–3433. doi:10.1002/cphc.201100826
20. Sinnokrot, M. O.; Sherrill, C. D. *J. Phys. Chem. A* **2004,** *108,* 10200–10207. doi:10.1021/jp0469517
21. Takatani, T.; Hohenstein, E. G.; Malagoli, M.; Marshall, M. S.; Sherrill, C. D. *J. Chem. Phys.* **2010,** *132,* 144104. doi:10.1063/1.3378024
22. Lejaeghere, K.; Bihlmayer, G.; Björkman, T.; Blaha, P.; Blügel, S.; Blum, V.; Caliste, D.; Castelli, I. E.; Clark, S. J.; Corso, A. D.; de Gironcoli, S.; Deutsch, T.; Dewhurst, J. K.; Di Marco, I.; Draxl, C.; Dułak, M.; Eriksson, O.; Flores-Livas, J. A.; Garrity, K. F.; Genovese, L.; Giannozzi, P.; Giantomassi, M.; Goedecker, S.; Gonze, X.; Grånäs, O.; Gross, E. K. U.; Gulans, A.; Gygi, F.; Hamann, D. R.; Hasnip, P. J.; Holzwarth, N. A. W.; Iuşan, D.; Jochym, D. B.; Jolle, F.; Jones, D.; Kresse, G.; Koepernik, K.; Küçükbenli, E.; Kvashnin, Y. O.; Locht, I. L. M.; Lubeck, S.; Marsman, M.; Marzari, N.; Nitzsche, U.; Nordström, L.; Ozaki, T.; Paulatto, L.; Pickard, C. J.; Poelmans, W.; Probert, M. I. J.; Refson, K.; Richter, M.; Rignanese, G.-M.; Saha, S.; Scheffler, M.; Schlipf, M.; Schwarz, K.; Sharma, S.; Tavazza, F.; Thunström, P.; Tkatchenko, A.; Torrent, M.; Vanderbilt, D.; van Setten, M. J.; Van Speybroeck, V.; Wills, J. M.; Yates, J. R.; Zhang, G.-X.; Cottenier, S. *Science* **2016,** *351,* aad3000. doi:10.1126/science.aad3000
23. Hermann, J.; Alfè, D.; Tkatchenko, A. *Nat. Commun.* **2017,** *8,* No. 14052. doi:10.1038/ncomms14052
24. Sutton, C.; Risko, C.; Brédas, J.-L. *Chem. Mater.* **2016,** *28,* 3–16. doi:10.1021/acs.chemmater.5b03266
25. Zang, L.; Che, Y.; Moore, J. S. *Acc. Chem. Res.* **2008,** *41,* 1596–1608. doi:10.1021/ar800030w
26. Neel, A. J.; Hilton, M. J.; Sigman, M. S.; Toste, F. D. *Nature* **2017,** *543,* 637–646. doi:10.1038/nature21701
27. Jurečka, P.; Šponer, J.; Černý, J.; Hobza, P. *Phys. Chem. Chem. Phys.* **2006,** *8,* 1985–1993. doi:10.1039/B600027D

28. Lebègue, S.; Harl, J.; Gould, T.; Ángyán, J. G.; Kresse, G.; Dobson, J. F. *Phys. Rev. Lett.* **2010,** *105,* 196401. doi:10.1103/PhysRevLett.105.196401
29. Gould, T.; Gray, E.; Dobson, J. F. *Phys. Rev. B* **2009,** *79,* 113402. doi:10.1103/PhysRevB.79.113402
30. Gould, T.; Simpkins, K.; Dobson, J. F. *Phys. Rev. B* **2008,** *77,* 165134. doi:10.1103/PhysRevB.77.165134
31. Dobson, J. F.; Gould, T.; Vignale, G. *Phys. Rev. X* **2014,** *4,* No. 021040. doi:10.1103/PhysRevX.4.021040
32. Misquitta, A. J.; Spencer, J.; Stone, A. J.; Alavi, A. *Phys. Rev. B* **2010,** *82,* 075312. doi:10.1103/PhysRevB.82.075312
33. Dobson, J. F. *Int. J. Quantum Chem.* **2014,** *114,* 1157–1161. doi:10.1002/qua.24635
34. DiStasio, R. A., Jr.; von Lilienfeld, O. A.; Tkatchenko, A. *Proc. Natl. Acad. Sci. U. S. A.* **2012,** *109,* 14791–14795. doi:10.1073/pnas.1208121109
35. Ambrosetti, A.; Ferri, N.; DiStasio, R. A., Jr.; Tkatchenko, A. *Science* **2016,** *351,* 1171–1176. doi:10.1126/science.aae0509
36. Gobre, V. V.; Tkatchenko, A. *Nat. Commun.* **2013,** *4,* No. 2341. doi:10.1038/ncomms3341
37. Kim, K.; Jordan, K. D. *J. Phys. Chem.* **1994,** *98,* 10089–10094. doi:10.1021/j100091a024
38. Johnson, E. R. *J. Chem. Phys.* **2011,** *135,* 234109. doi:10.1063/1.3670015
39. Maurer, R. J.; Ruiz, V. G.; Tkatchenko, A. *J. Chem. Phys.* **2015,** *143,* 102808. doi:10.1063/1.4922688
40. Christian, M. S.; Otero-de-la-Roza, A.; Johnson, E. R. *J. Chem. Theory Comput.* **2016,** *12,* 3305–3315. doi:10.1021/acs.jctc.6b00222
41. Otero-de-la-Roza, A.; Johnson, E. R. *J. Chem. Phys.* **2013,** *138,* 054103. doi:10.1063/1.4789421
42. Otero-de-la-Roza, A.; DiLabio, G. A.; Johnson, E. R. *J. Chem. Theory Comput.* **2016,** *12,* 3160–3175. doi:10.1021/acs.jctc.6b00298
43. Huang, Y.; Beran, G. J. O. *J. Chem. Phys.* **2015,** *143,* 044113. doi:10.1063/1.4927304
44. Řezáč, J.; Huang, Y.; Hobza, P.; Beran, G. J. O. *J. Chem. Theory Comput.* **2015,** *11,* 3065–3079. doi:10.1021/acs.jctc.5b00281
45. Perdew, J. P.; Burke, K.; Ernzerhof, M. *Phys. Rev. Lett.* **1996,** *77,* 3865–3868. doi:10.1103/PhysRevLett.77.3865
46. Wu, Q.; Yang, W. *J. Chem. Phys.* **2002,** *116,* 515–524. doi:10.1063/1.1424928
47. Wu, X.; Vargas, M. C.; Nayak, S.; Lotrich, V.; Scoles, G. *J. Chem. Phys.* **2001,** *115,* 8748–8757. doi:10.1063/1.1412004
48. Grimme, S. *J. Comput. Chem.* **2006,** *27,* 1787–1799. doi:10.1002/jcc.20495
49. Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. *J. Chem. Phys.* **2010,** *132,* 154104. doi:10.1063/1.3382344
50. Elstner, M.; Hobza, P.; Frauenheim, T.; Suhai, S.; Kaxiras, E. *J. Chem. Phys.* **2001,** *114,* 5149–5155. doi:10.1063/1.1329889
51. Becke, A. D.; Johnson, E. R. *J. Chem. Phys.* **2007,** *127,* 154108. doi:10.1063/1.2795701
52. Johnson, E. R. *Non-Covalent Interactions in Quantum Chemistry and Physics;* Elsevier, 2017; pp 169–194. doi:10.1016/B978-0-12-809835-6.00006-2
53. Tkatchenko, A.; Scheffler, M. *Phys. Rev. Lett.* **2009,** *102,* 073005. doi:10.1103/PhysRevLett.102.073005
54. Tkatchenko, A.; DiStasio, R. A., Jr.; Car, R.; Scheffler, M. *Phys. Rev. Lett.* **2012,** *108,* 236402. doi:10.1103/PhysRevLett.108.236402
55. Ambrosetti, A.; Reilly, A. M.; DiStasio, R. A., Jr.; Tkatchenko, A. *J. Chem. Phys.* **2014,** *140,* 18A508. doi:10.1063/1.4865104
56. Gould, T.; Bučko, T. *J. Chem. Theory Comput.* **2016,** *12,* 3603–3613. doi:10.1021/acs.jctc.6b00361
57. Gould, T.; Lebègue, S.; Ángyán, J. G.; Bučko, T. *J. Chem. Theory Comput.* **2016,** *12,* 5920–5930. doi:10.1021/acs.jctc.6b00925
58. Dobson, J. F.; Dinte, B. P.; Wang, J.; Gould, T. *Aust. J. Phys.* **2000,** *53,* 575–596.
59. Dion, M.; Rydberg, H.; Schröder, E.; Langreth, D. C.; Lundqvist, B. I. *Phys. Rev. Lett.* **2004,** *92,* 246401. doi:10.1103/PhysRevLett.92.246401
60. Vydrov, O. A.; Van Voorhis, T. *J. Chem. Phys.* **2011,** *133,* 244103. doi:10.1063/1.3521275
61. Lee, K.; Murray, É. D.; Kong, L.; Lundqvist, B. I.; Langreth, D. C. *Phys. Rev. B* **2010,** *82,* 081101. doi:10.1103/PhysRevB.82.081101
62. Klimeš, J.; Bowler, D. R.; Michaelides, A. *J. Phys.: Condens. Matter* **2009,** *22,* 022201. doi:10.1088/0953-8984/22/2/022201
63. Peng, H.; Yang, Z.-H.; Perdew, J. P.; Sun, J. *Phys. Rev. X* **2016,** *6,* 041005. doi:10.1103/PhysRevX.6.041005
64. Zhao, Y.; Truhlar, D. G. *Theor. Chem. Acc.* **2008,** *120,* 215–241. doi:10.1007/s00214-007-0310-x
65. Sun, J.; Ruzsinszky, A.; Perdew, J. P. *Phys. Rev. Lett.* **2015,** *115,* 036402. doi:10.1103/PhysRevLett.115.036402
66. Pernal, K.; Podeszwa, R.; Patkowski, K.; Szalewicz, K. *Phys. Rev. Lett.* **2009,** *103,* 263201. doi:10.1103/PhysRevLett.103.263201
67. Hermann, J.; Tkatchenko, A. *J. Chem. Theory Comput.* **2018,** *14,* 1361–1369. doi:10.1021/acs.jctc.7b01172
68. Kresse, G.; Furthmüller, J. *Phys. Rev. B* **1996,** *54,* 11169–11186. doi:10.1103/PhysRevB.54.11169
69. Kresse, G.; Joubert, D. *Phys. Rev. B* **1999,** *59,* 1758–1775. doi:10.1103/PhysRevB.59.1758
70. Blochl, P. E. *Phys. Rev. B* **1994,** *50,* 17953–17979. doi:10.1103/PhysRevB.50.17953
71. Bučko, T.; Lebègue, S.; Gould, T.; Ángyán, J. G. *J. Phys.: Condens. Matter* **2016,** *28,* 045201. doi:10.1088/0953-8984/28/4/045201
72. Klimeš, J.; Bowler, D. R.; Michaelides, A. *Phys. Rev. B* **2011,** *83,* 195131. doi:10.1103/PhysRevB.83.195131
73. Román-Pérez, G.; Soler, J. M. *Phys. Rev. Lett.* **2009,** *103,* 096102. doi:10.1103/PhysRevLett.103.096102
74. Parrish, R. M.; Burns, L. A.; Smith, D. G. A.; Simmonett, A. C.; DePrince, A. E., III; Hohenstein, E. G.; Bozkaya, U.; Sokolov, A. Y.; Di Remigio, R.; Richard, R. M.; Gonthier, J. F.; James, A. M.; McAlexander, H. R.; Kumar, A.; Saitow, M.; Wang, X.; Pritchard, B. P.; Verma, P.; Schaefer, H. F., III; Patkowski, K.; King, R. A.; Valeev, E. F.; Evangelista, F. A.; Turney, J. M.; Crawford, T. D.; Sherrill, C. D. *J. Chem. Theory Comput.* **2017,** *13,* 3185–3197. doi:10.1021/acs.jctc.7b00174
75. Leconte, N.; Jung, J.; Lebègue, S.; Gould, T. *Phys. Rev. B* **2017,** *96,* 195431. doi:10.1103/PhysRevB.96.195431
76. Johnson, E. R.; Becke, A. D.; Sherrill, C. D.; DiLabio, G. A. *J. Chem. Phys.* **2009,** *131,* 034111. doi:10.1063/1.3177061

License and Terms

BEILSTEIN JOURNAL OF ORGANIC CHEMISTRY

London dispersion as important factor for the stabilization of (*Z*)-azobenzenes in the presence of hydrogen bonding

Andreas H. Heindl, Raffael C. Wende and Hermann A. Wegner*

Full Research Paper

Open Access

Address:
Institut für Organische Chemie, Justus-Liebig-Universität Gießen, Heinrich-Buff-Ring 17, 35392 Gießen, Germany

Email:
Hermann A. Wegner* - hermann.a.wegner@org.chemie.uni-giessen.de

* Corresponding author

Keywords:
azobenzene; hydrogen bonding; London dispersion; molecular switches

Beilstein J. Org. Chem. **2018,** *14,* 1238–1243.
doi:10.3762/bjoc.14.106

Received: 27 February 2018
Accepted: 11 May 2018
Published: 29 May 2018

This article is part of the Thematic Series "Dispersion interactions".

Guest Editor: P. Schreiner

Abstract

The understanding and control of the light-induced isomerization of azobenzenes as one of the most important classes of molecular switches is crucial for the design of light-responsive materials using this entity. Herein, we present the stabilization of metastable (*Z*)-azobenzenes by London dispersion interactions, even in the presence of comparably stronger hydrogen bonds in various solvents. The *Z*→*E* isomerization rates of several *N*-substituted 4,4′-bis(4-aminobenzyl)azobenzenes were measured. An intramolecular stabilization was observed and explained by the interplay of intramolecular amide and carbamate hydrogen bonds as well as London dispersion interactions. Whereas in toluene, 1,4-dioxane and *tert*-butyl methyl ether the hydrogen bonds dominate, the variation in stabilization of the different substituted azobenzenes in dimethyl sulfoxide can be rationalized by London dispersion interactions. These findings were supported by conformational analysis and DFT computations and reveal low-energy London dispersion forces to be a significant factor, even in the presence of hydrogen bonds.

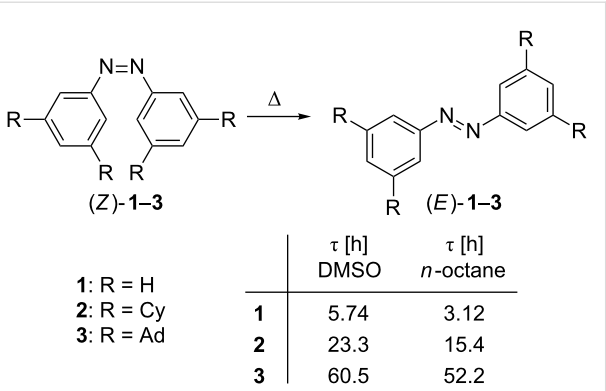
Introduction

The photo-controlled *E*→*Z* isomerization of azobenzene has been known for decades [1] and has originated a wide field of applications in recent years. This molecular switch has been utilized inter alia in the rising field of photopharmacology [2,3], the manipulation of biomolecular processes [4-6] as well as in molecular machinery [7,8] and materials science [9-11]. Azobenzenes are highly stable, easily synthesized [12] and show reversible isomerization from the thermally stable *E*- to the *Z*-isomer upon irradiation with UV light. The metastable *Z*-azobenzene re-isomerizes to the *E*-conformer either thermally or upon irradiation with visible light [13,14]. Interestingly, the thermal stability of azobenzene isomers can be reversed by the incorporation of azobenzene units in macrocyclic arrangements [15]. For example, the groups of Tamaoki [16] and Herges [17] reported azobenzophanes that isomerize thermally to their energetically lower *Z*-conformations from

their corresponding higher energy *E*-isomers. Moreover, our group presented a highly strained bisazobenzophane that was found to be stable exclusively in its (*Z,Z*)-form. Consequently, no isomerization, neither photochemically nor thermally, to its (*E,Z*)- or (*E,E*)-state occurred on the observed time scale [18].

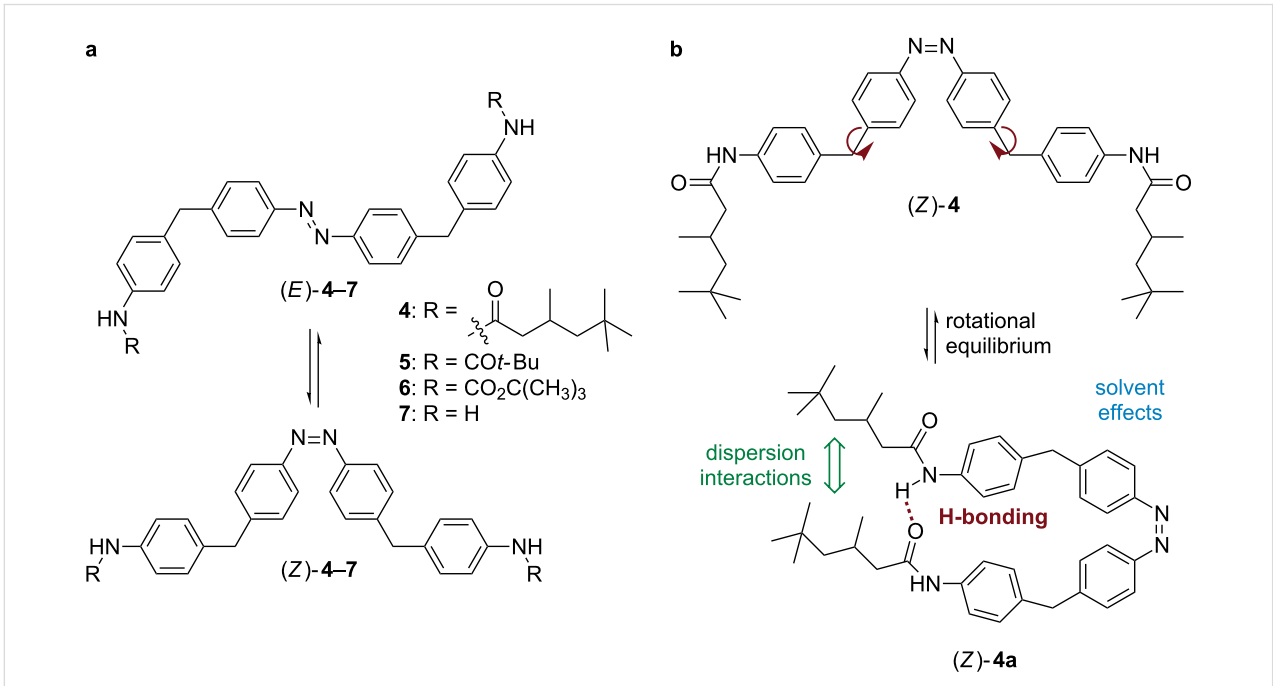
The thermal half-lives of acyclic (*Z*)-azobenzenes can be prolonged dramatically by *ortho*-fluorine substitution, resulting in half-lives of up to two years at room temperature [19]. These highly thermally stable (*Z*)-*ortho*-fluoroazobenzenes can be re-isomerized almost instantaneously in an electrocatalytic fashion [20]. Furthermore, (*Z*)-azobenzenes can also be stabilized without changing their electronic configuration by attractive London dispersion (LD) interactions [21,22]. In a recent study, we reported that all-*meta*-alkyl-substituted (*Z*)-azobenzenes increase in stability with increasingly larger substituents (Scheme 1) [23]. Supported by density functional theory (DFT) computations, attractive LD forces were identified as the origin of this stability trend. Based on this study, LD interactions represent a valuable tool for the design of novel azobenzene photoswitches [24]. Herein we provide further evidence for the importance of LD as unneglectable stabilizing element in controlling interactions in functionalized molecules.

During studies on cyclobisazodiphenylmethane [18] we noticed that the bulky (*Z*)-4,4'-bis[4-(3,5,5-trimethylhexanoylamino)benzyl]azobenzene (**4**, Scheme 2a) showed an unexpectedly increased thermal half-life compared to other azobenzenes



Scheme 1: Half-lives for the thermal *Z*→*E* isomerization of all-*meta*-alkyl-substituted azobenzenes **1**–**3** in DMSO and *n*-octane at 53.2 °C (examples taken from ref. [23], Cy = cyclohexyl, Ad = adamantyl).

5–**7** in *tert*-butyl methyl ether (TBME) as solvent. This observation is surprising, since azobenzenes **4**–**7** are electronically very similar as the methylene linker prevents conjugation of the *N*-aryl and the azobenzene moieties. At first thought, the large spatial separation of the *N*-substituted moieties in (*Z*)-**4**–**7** make intramolecular stabilizing interactions unlikely. Nevertheless, it can be envisioned that the freely rotatable Ar–CH_2–Ar units should allow the formation of close proximity conformers of (*Z*)-**4a** in solution, in which attractive interactions, such as hydrogen bonding and London dispersion, may indeed become possible (Scheme 2b) [25]. To further investigate this phenomenon, the rates of the thermal *Z*→*E* isomerization of azoben-



Scheme 2: Isomerization of *N*-substituted *Z*-azobenzenes (a). Rotational equilibrium of (*Z*)-**4** allowing intramolecular interactions (b).

zenes **4**–**7** were determined in different solvents and at different temperatures (for syntheses, see ref. [18] and experimental section).

Results and Discussion

To further investigate the prior observations, the thermal *Z*→*E* isomerization rates of azobenzenes **4**–**7** were determined by UV–vis spectroscopy in several solvents at 25 °C and 35 °C, respectively. The lowest thermal isomerization rates were found for the crowded 4,4'-bis[4-(3,5,5,-trimethylhexanoylamino)benzyl]azobenzene (**4**, Table 1, entry 1), expressed in the longest half-lives τ. However, in the highly polar solvent dimethyl sulfoxide (DMSO) all azoamides (**4**, **5**, Table 1, entries 1 and 2) as well as the azocarbamate **6** (Table 1, entry 3) showed very similar isomerization rates. Additionally, nearly the same half-lives were observed in all solvents for the *Z*→*E* isomerizations of azoamine **7** (Table 1, entry 4) and azocarbamate **6** (Table 1, entry 3). The isomerization rates of bis(*tert*-butylcarbonylaminobenzyl)azobenzene **5** could not be measured in toluene and TBME due to the insolubility of the compound in these solvents. In 1,4-dioxane and DMSO the isomerization depends on the temperature. While the rates are in comparison rather long at 25 ºC, these processes are considerably faster at 35 ºC.

As outlined before, several stabilizing interactions such as LD, hydrogen bonding and solvation effects are possible for the stabilization of *Z*-azobenzenes **4**–**7**. To estimate the influences of those effects, a conformer distribution analysis was performed to identify low-lying conformations of the corresponding (*Z*)-azobenzenes. The energetically favored conformers found (within 1.5 kcal mol^{-1} for **4** and 5 kcal mol^{-1} for **5**–**7**, respectively, relative to the lowest energy conformer) were then re-optimized at the B3LYP/6-31G** [26-29] level of theory with and without D3(BJ) [30,31] dispersion correction (gas phase) (conformations of one enantiomer of each diastereomer of **4** were analyzed. The maximum stabilization was found for (*R*,*S*)-**4**. For the other diastereomer, see Supporting Information File 1).

As it can be seen in Table 2, the computations reproduced the stabilization of the tight conformations relative to their open forms, which is in agreement with the experimentally observed kinetics of azobenzenes **4**–**7**. Comparing the free energies of azobenzenes **4**–**7** relative to their sterically less crowded conformations, where the diphenylmethane units point away from each other, azodiamide **4** was found to feature the highest stabilization, whereas the *tert*-butylcarbonylamino compound **5** and Boc-protected derivative **6** are almost equally stabilized.

Table 1: Thermal *Z*→*E* isomerization half-lives τ (standard deviations in parentheses) of azobenzenes **4**–**7** in various solvents.

entry	compound	$\tau_{Z\rightarrow E}$ [h]							
		toluene		1,4-dioxane		DMSO		TBME	
		25 °C	35 °C	25 °C	35 °C	25 °C	35 °C	25 °C	35 °C
1	**4** R = TMH[a]	97 (6)	32.2 (0.8)	45 (7)	16.7 (0.4)	50.1 (0.2)	14.7 (0.002)	65 (13)	26 (1)
2	**5** R = CO*t*-Bu	–	–	46.3 (0.6)	14.0 (0.1)	50.0 (0.2)	14.4 (0.2)	–	–
3	**6** R = Boc	48 (2)	14.0 (0.6)	39.7 (0.2)	14.9 (0.1)	51.1 (0.3)	14.4 (0.2)	36 (10)	14 (1)
4	**7** R = H	35.4 (0.2)	11.5 (0.5)	38 (2)	12.0 (0.3)	45.7 (0.01)	12.8 (0.3)	37.6 (0.7)	14.3 (0.3)

[a]TMH = 3,5,5-trimethylhexanoyl.

Table 2: Computational results for (*Z*)-azobenzenes **4**–**7**. Δ*G* is the free energy of the most stable *Z*-conformer relative to the corresponding open (*Z*)-conformation (method in parentheses, see Figure S4 in Supporting Information File 1 for graphical representations of the compared conformers).

Compound	**4** R = TMH	**5** R = CO*t*-Bu	**6** R = Boc	**7** R = H
Δ*G* relative to open conformations (B3LYP-D3(BJ)[a]) [kcal mol^{-1}]	−16.9	−10.5	−10.7	−3.7
Δ*G* relative to open conformations (B3LYP[a]) [kcal mol^{-1}]	1.2	−0.7	0.6	0.5
$d_{NH\cdots X}$ [Å][a,b]	1.87	2.14	1.90	2.11
closest $d_{H\cdots H}$ [Å]	2.38	2.53	2.39	–

[a]Basis set: 6-31G**, [b]X = O for **4**, **5** and **6**; X = N for **7**.

Furthermore, azodiamine **7** shows the lowest relative stabilization and thus isomerizes most rapidly. Additionally, all compounds form intramolecular hydrogen bonds between the *N*-substituted moieties. However, the higher relative free energies computed with D3(BJ) dispersion correction, as well as the close H–H contacts in all compounds of about 2.4 Å to 3.0 Å support the stabilizing effect of LD interactions in (*Z*)-azobenzenes **4**–**7**. Figure 1 visualizes the computational findings for the compared conformers of (*Z*)-**4**. A noncovalent interaction (NCI) analysis [32,33] of (*R*,*S*)-**4** in its tightly folded conformation revealed multiple attractive H–H as well as H–π interactions between the amide residues and the aryl groups to be responsible for the overall stabilization (green and blue isosurfaces in Figure 1b).

Apparently, hydrogen bonds are much stronger than LD interactions and are mainly responsible for the observed stabilization effects, which is also represented as a blue surface in the NCI plot of (*R*,*S*)-**4** c1. Furthermore, it is known that amides form stronger hydrogen bonds than carbamates and the strength also depends on the steric bulk of the amide [34]. This explains the highest stabilization of the trimethylhexanoylamide **4** (experimentally and computationally), followed by the *tert*-butylamide **5** and *tert*-butyl carbamate **6**, which are almost equally stabilized. Nevertheless, hydrogen bonds are strongly dependent on the solvent system and become weaker with increasing dielectric constant of the solvent [35]. This fact becomes obvious when comparing the half-lives of azobenzene **4** in non-polar toluene or TBME with polar DMSO. In contrast to toluene, the half-life of **4** decreased almost by 50% in DMSO due to the weakening of intramolecular hydrogen bonding. As a result, all azobenzenes despite **7** showed comparable isomerization half-lives in DMSO. Accordingly, LD interactions can be responsible for the slightly but significantly prolonged half-lives compared to azodiamine **7**. Analyzing the isomerization in 1,4-dioxane, weak hydrogen bonds as well as dispersion interactions are operative, which is expressed by the same isomerization half-life trend as in toluene, yet with lower absolute values. These results show, that LD interactions may indeed contribute to the overall stabilization of complex molecules even in the presence of stronger interactions, such as hydrogen bonds [36,37].

Conclusion

In conclusion, unexpected variations in the isomerization rates of azobenzenes with different remote nitrogen substituents were observed. The experimental and computational investigations reveal a subtle interplay of hydrogen bonding, LD interactions and solvent effects. In general, intramolecular hydrogen bonds were found to have the strongest influence on the observed thermal *Z*→*E* isomerization half-lives. However, LD becomes the decisive factor in polar solvents in which hydrogen bonding plays a minor role. This study demonstrates the importance of even small energy interactions, such as LD, and provides new insights for the application of LD as design element in complex systems in general.

Experimental

Synthesis of bis(*tert*-butylcarbonylamino)azobenzene 5: To a solution of 4,4′-bis(4-aminobenzyl)azobenzene (**7**) [18] (78 mg, 0.20 mmol, 1.0 equiv) and NEt_3 (65 µL, 0.44 mmol, 2.2 equiv) in THF (2 mL), pivaloyl chloride (54.4 µL, 0.437 mmol, 2.20 equiv) in THF (0.5 mL) was added dropwise at 0 °C. Then, the reaction was allowed to warm to rt while stirring for 2 h. After quenching with sat. aq. NH_4Cl solution (5 mL), the aqueous phase was extracted with THF (5 mL). The organic

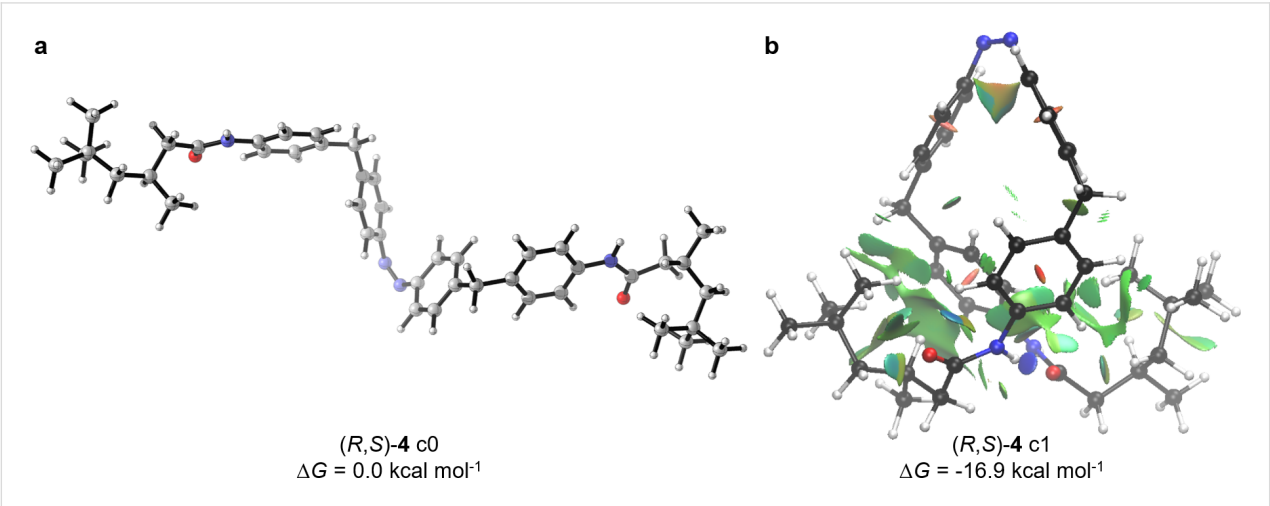


Figure 1: a) Optimized geometry of open conformer (*R*,*S*)-**4** c0. b) NCI plot of the most stable conformer of (*R*,*S*)-**4** (green and blue isosurfaces indicate attractive noncovalent interactions).

phase was washed with sat. aq. $NaHCO_3$ (2 × 10 mL) and was dried over $MgSO_4$, filtered and concentrated. The residue was suspended in hot EtOH and was filtered while hot. After evaporation of the filtrate, the residue was washed with H_2O and was dried under high vacuum to yield the product as a yellow solid (39 mg, 35%). ^{1}H NMR (400 MHz, DMSO-d_6) δ ppm 9.14 (s, 2H), 7.79 (d, J = 8.0 Hz, 4H), 7.56 (d, J = 8.2 Hz, 4H), 7.40 (d, J = 8.0 Hz, 4H), 7.18 (d, J = 8.1 Hz, 4H), 3.98 (s, 4H), 1.20 (s, 18H); ^{13}C NMR (101 MHz, DMSO-d_6) δ ppm 176.3, 150.3, 145.3, 137.5, 135.3, 129.6, 128.7, 122.6, 120.5, 40.7*, 39.5*, 27.2; HRMS (ESI) m/z: $[M + Na]^+$ calcd, 583.3043; found, 583.3043. *identified by HSQC and HMBC spectroscopy.

Supporting Information

Supporting Information File 1

NMR spectra of azobenzene **5**, UV–vis data and detailed procedures, details for conformational analysis and DFT computations.
[https://www.beilstein-journals.org/bjoc/content/supplementary/1860-5397-14-106-S1.pdf]

Acknowledgements

Financial support by the Deutsche Forschungsgemeinschaft is gratefully acknowledged. Lea E. Schäfer is gratefully acknowledged for help with the TOC figure.

ORCID® iDs

Andreas H. Heindl - https://orcid.org/0000-0002-5403-2177
Raffael C. Wende - https://orcid.org/0000-0002-2242-4723
Hermann A. Wegner - https://orcid.org/0000-0001-7260-6018

References

1. Hartley, G. S. *Nature* **1937,** *140,* 281. doi:10.1038/140281a0
2. Velema, W. A.; Szymanski, W.; Feringa, B. L. *J. Am. Chem. Soc.* **2014,** *136,* 2178–2191. doi:10.1021/ja413063e
3. Broichhagen, J.; Trauner, D. *Curr. Opin. Chem. Biol.* **2014,** *21,* 121–127. doi:10.1016/j.cbpa.2014.07.008
4. Van der Berg, J. P.; Velema, W. A.; Szymanski, W.; Driessen, A. J. M.; Feringa, B. L. *Chem. Sci.* **2015,** *6,* 3593–3598. doi:10.1039/C5SC00215J
5. Bellotto, S.; Chen, S.; Rentero Rebollo, I.; Wegner, H. A.; Heinis, C. *J. Am. Chem. Soc.* **2014,** *136,* 5880–5883. doi:10.1021/ja501861m
6. Schweighauser, L.; Wegner, H. A. *ChemBioChem* **2015,** *16,* 1709–1711. doi:10.1002/cbic.201500276
7. Merino, E.; Ribagorda, M. *Beilstein J. Org. Chem.* **2012,** *8,* 1071–1090. doi:10.3762/bjoc.8.119
8. Heitmann, G.; Schütt, C.; Herges, R. *Eur. J. Org. Chem.* **2016,** 3817–3823. doi:10.1002/ejoc.201600548
9. Pieraccini, S.; Gottarelli, G.; Labruto, R.; Masiero, S.; Pandoli, O.; Spada, G. P. *Chem. – Eur. J.* **2004,** *10,* 5632–5639. doi:10.1002/chem.200400463
10. Ikeda, T.; Tsutsumi, O. *Science* **1995,** *268,* 1873–1875. doi:10.1126/science.268.5219.1873
11. Lennartson, A.; Roffey, A.; Moth-Poulsen, K. *Tetrahedron Lett.* **2015,** *56,* 1457–1465. doi:10.1016/j.tetlet.2015.01.187
12. Merino, E. *Chem. Soc. Rev.* **2011,** *40,* 3835–3853. doi:10.1039/C0CS00183J
13. Griffiths, J. *Chem. Soc. Rev.* **1972,** *1,* 481–493. doi:10.1039/CS9720100481
14. Bandara, H. M. D.; Burdette, S. C. *Chem. Soc. Rev.* **2012,** *41,* 1809–1825. doi:10.1039/c1cs15179g
15. Reuter, R.; Wegner, H. A. *Chem. Commun.* **2011,** *47,* 12267–12276. doi:10.1039/c1cc13773e
16. Norikane, Y.; Katoh, R.; Tamaoki, N. *Chem. Commun.* **2008,** 1898–1900. doi:10.1039/B718813G
17. Siewertsen, R.; Neumann, H.; Buchheim-Stehn, B.; Herges, R.; Näther, C.; Renth, F.; Temps, F. *J. Am. Chem. Soc.* **2009,** *131,* 15594–15595. doi:10.1021/ja906547d
18. Heindl, A. H.; Schweighauser, L.; Logemann, C.; Wegner, H. A. *Synthesis* **2017,** *49,* 2632–2639. doi:10.1055/s-0036-1589006
19. Knie, C.; Utecht, M.; Zhao, F.; Kulla, H.; Kovalenko, S.; Brouwer, A. M.; Saalfrank, P.; Hecht, S.; Bléger, D. *Chem. – Eur. J.* **2014,** *20,* 16492–16501. doi:10.1002/chem.201404649
20. Goulet-Hanssens, A.; Utecht, M.; Mutruc, D.; Titov, E.; Schwarz, J.; Grubert, L.; Bléger, D.; Saalfrank, P.; Hecht, S. *J. Am. Chem. Soc.* **2017,** *139,* 335–341. doi:10.1021/jacs.6b10822
21. Wagner, J. P.; Schreiner, P. R. *Angew. Chem., Int. Ed.* **2015,** *54,* 12274–12296. doi:10.1002/anie.201503476
22. Rösel, S.; Balestrieri, C.; Schreiner, P. R. *Chem. Sci.* **2017,** *8,* 405–410. doi:10.1039/C6SC02727J
23. Schweighauser, L.; Strauss, M. A.; Bellotto, S.; Wegner, H. A. *Angew. Chem., Int. Ed.* **2015,** *54,* 13436–13439. doi:10.1002/anie.201506126
24. Grimme, S.; Huenerbein, R.; Ehrlich, S. *ChemPhysChem* **2011,** *12,* 1258–1261. doi:10.1002/cphc.201100127
25. Schaefer, T.; Niemczura, W.; Danchura, W.; Wildman, T. A. *Can. J. Chem.* **1979,** *57,* 1881–1886. doi:10.1139/v79-298
26. *Gaussian 16,* Revision A.03; Gaussian, Inc.: Wallingford, CT, 2016.
27. Becke, A. D. *J. Chem. Phys.* **1993,** *98,* 1372–1377. doi:10.1063/1.464304
28. Krishnan, R.; Binkley, J. S.; Seeger, R.; Pople, J. A. *J. Chem. Phys.* **1980,** *72,* 650–654. doi:10.1063/1.438955
29. Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1988,** *37,* 785–789. doi:10.1103/PhysRevB.37.785
30. Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. *J. Chem. Phys.* **2010,** *132,* 154104. doi:10.1063/1.3382344
31. Grimme, S.; Ehrlich, S.; Goerigk, L. *J. Comput. Chem.* **2011,** *32,* 1456–1465. doi:10.1002/jcc.21759
32. Johnson, E. R.; Keinan, S.; Mori-Sánchez, P.; Contreras-García, J.; Cohen, A. J.; Yang, W. *J. Am. Chem. Soc.* **2010,** *132,* 6498–6506. doi:10.1021/ja100936w
33. Contreras-García, J.; Johnson, E. R.; Keinan, S.; Chaudret, R.; Piquemal, J.-P.; Beratan, D. N.; Yang, W. *J. Chem. Theory Comput.* **2011,** *7,* 625–632. doi:10.1021/ct100641a
34. Le Questel, J.-Y.; Laurence, C.; Lachkar, A.; Helbert, M.; Berthelot, M. *J. Chem. Soc., Perkin Trans. 2* **1992,** 2091–2094. doi:10.1039/P29920002091
35. Rospenk, M.; Zeegers-Huyskens, T. *J. Phys. Chem.* **1987,** *91,* 3974–3977. doi:10.1021/j100299a011

36. Wende, R. C.; Seitz, A.; Niedek, D.; Schuler, S. M. M.; Hofmann, C.; Becker, J.; Schreiner, P. R. *Angew. Chem., Int. Ed.* **2016,** *55,* 2719–2723. doi:10.1002/anie.201509863
37. Müller, C. E.; Zell, D.; Hrdina, R.; Wende, R. C.; Wanka, L.; Schuler, S. M. M.; Schreiner, P. R. *J. Org. Chem.* **2013,** *78,* 8465–8484. doi:10.1021/jo401195c

License and Terms

BEILSTEIN JOURNAL OF ORGANIC CHEMISTRY

Steric “attraction”: not by dispersion alone

Ganna Gryn’ova and Clémence Corminboeuf*

Full Research Paper

Open Access

Address:
Institut des Sciences et Ingénierie Chimiques, École polytechnique fédérale de Lausanne, CH-1015 Lausanne, Switzerland

Email:
Clémence Corminboeuf* - clemence.corminboeuf@epfl.ch

* Corresponding author

Keywords:
charge penetration; dispersion; hydrocarbon; non-covalent interactions; steric attraction

Beilstein J. Org. Chem. **2018,** *14,* 1482–1490.
doi:10.3762/bjoc.14.125

Received: 26 February 2018
Accepted: 02 June 2018
Published: 19 June 2018

This article is part of the Thematic Series "Dispersion interactions".

Guest Editor: P. Schreiner

Abstract

Non-covalent interactions between neutral, sterically hindered organic molecules generally involve a strong stabilizing contribution from dispersion forces that in many systems turns the ‘steric repulsion’ into a ‘steric attraction’. In addition to London dispersion, such systems benefit from electrostatic stabilization, which arises from a short-range effect of charge penetration and gets bigger with increasing steric bulk. In the present work, we quantify this contribution for a diverse set of molecular cores, ranging from unsubstituted benzene and cyclohexane to their derivatives carrying *tert*-butyl, phenyl, cyclohexyl and adamantyl substituents. While the importance of electrostatic interactions in the dimers of sp^2-rich (e.g., π-conjugated) cores is well appreciated, less polarizable assemblies of sp^3-rich systems with multiple short-range CH···HC contacts between the bulky cyclohexyl and adamantyl moieties are also significantly influenced by electrostatics. Charge penetration is drastically larger in absolute terms for the sp^2-rich cores, but still has a non-negligible effect on the sp^3-rich dimers, investigated herein, both in terms of their energetics and equilibrium interaction distances. These results emphasize the importance of this electrostatic effect, which has so far been less recognized in aliphatic systems compared to London dispersion, and are therefore likely to have implications for the development of force fields and methods for crystal structure prediction.

Introduction

In the recent years, perception of the vaguely defined ‘steric’ interactions as categorically repulsive has shifted towards recognizing the crucial role of attractive dispersion in the bulky systems [1]. London dispersion was shown to be capable of bending σ-bonded acene dimers (**2** in Figure 1A) [2] and stabilizing extremely crowded systems, hexaphenylethane (**3**) being the mascot of this concept. As elegantly illustrated by Schreiner and Grimme [3,4], while bare hexaphenylethane (**3**) is thermodynamically unstable due to significant Pauli repulsion between the phenyl rings, its analogue **4** carrying all-*meta*-*tert*-butyl substituents (termed ‘dispersion donors’) can be synthesized and characterized thanks to sufficient stabilization by dispersion (Figure 1B). This realization sparked a race towards the longest covalent C–C bonds [5,6]: already impressive

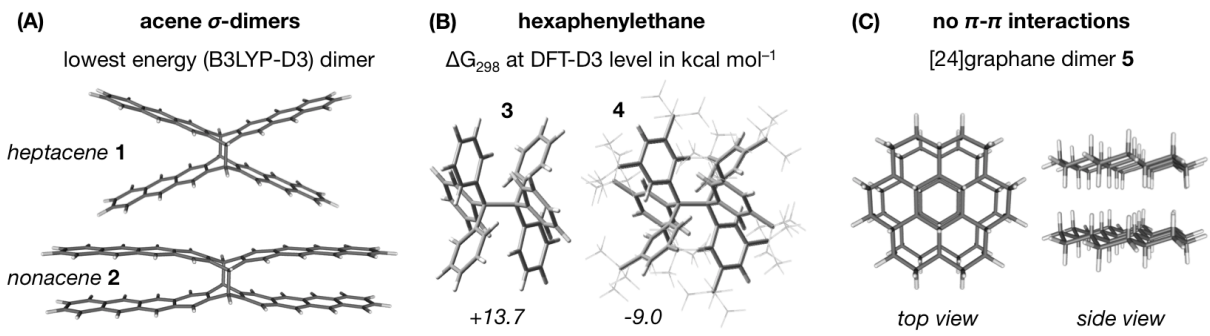


Figure 1: (A) Dispersion is insufficient to bend the heptacene σ-dimer, but becomes sizable enough in nonacene [2]. (B) Bare hexaphenylethane is not thermodynamically stable, but its all-*meta*-*tert*-butyl derivative is [3,4]. (C) Thermodynamically stable (at B97D/6-31G(d,p) level) bilayer of a fully saturated hydrocarbon, [24]-graphane [15].

1.67 Å in hexakis(3,5-di-*tert*-butylphenyl)ethane is not even a limit and stable diamondoid dimer with a central C–C bond as long as 1.71 Å has been achieved [7,8]. Bulky alkyl groups assist not only in achieving the longest C–C bonds, but also the shortest intermolecular H···H contacts [9], which are otherwise tackled by squeezing them inside the cages [10]. Intermolecular interactions in hydrocarbons are also subject to significant dispersion contribution. In the unsaturated systems, from benzene dimer to higher acenes and, ultimately, graphenes, dispersion is increasingly the key force behind the π–π stacking interactions [11]. Large and flat π-conjugated moieties (e.g., ligands) are even referred to as 'sticky pancakes' in homage to strong attractive interactions between them [12-14]. Less intuitively, similarly strong attractive forces are found in extended saturated systems, e.g., the double sheet graphanes and [*n*]ladderane dimers, where the interaction occurs via the CH···HC and CH···C contacts (Figure 1C) [15-24].

While London dispersion is deservedly paraded as the champion of 'steric attraction' in bulky hydrocarbons, several studies have recently pointed to the somewhat less expected electrostatic contribution to it [25]. For example, "dispersion dominates and electrostatics commands" is the 'punch line' of the 2017 computational study on the σ–σ, σ–π and π–π stacking interactions between benzene, cyclohexane and some of their fluorinated derivatives [26]. The authors show that while electrostatics is not the largest stabilizing energetic contribution, it is nonetheless the one that defines the trend in the total interaction energy for a range of investigated dimers. Electrostatic stabilization in graphane and graphene dimers has been attributed to the charge transfer ($\sigma_{CH} \rightarrow \sigma_{HC*}$ hyperconjugative interaction) [17], and a similar argument was used to suggest the possibility of manipulating the band gap of patterned hydrogenated graphene C_4H bilayer by an external electric field [27]. Furthermore, Schreiner et al. showed that approx. 10% of the total interaction energy in the tris(3,5-di-*tert*-butylphenyl)methane dimer (the system mentioned above for its shortest intermolecular H···H contacts) comes from stabilizing electrostatics [9]. Similarly, the interaction energy difference between the all-meta-*tert*-butyl-hexaphenylethane and the bare hexaphenylethane features ≈14% electrostatic contribution at the ISAPT0/jun-cc-pVDZ level [28]. These studies have identified the penetration energy as the dominant component of the electrostatic interaction energy. While at long range electrostatics is virtually entirely due to interactions between the permanent multipoles of the interacting species, at small interaction distances it is instead strongly influenced and in some systems even dominated by charge penetration [29]. The latter is an outcome of the overlapping diffuse electron clouds of interacting molecules. The resulting attraction between the nuclei of one molecule to the electron density of the other is greater than the electron–electron and nuclei–nuclei repulsion. The crucial role of charge penetration has been demonstrated for a diverse range of chemical systems, including the saturated [9,28] and unsaturated hydrocarbons [30], nucleic acids [31], metal ions interacting with proteins [32], heteroaromatic cores that are the common building blocks for organic semiconductors [33], cyclophanes [34], Wilcox torsion balance systems [35], etc.

The recognized importance of charge penetration in various chemical problems is paralleled by myriad developments aimed at accurately describing intermolecular interactions. Effective fragment potential (EFP) methods estimate it by adding a damping term to their classical multipolar expansion [36,37]. To account for this effect, quantum mechanically derived force fields (FFs) are fitted to semi-empirical [38], dispersion-corrected density functional theory [39], post-Hartree–Fock [40,41], symmetry adapted perturbation theory (SAPT) [42-46] data or to a combination of the latter two (e.g., the monomer electron density force field, MEDFF) [47]. The latter approach has been subsequently exploited in the machine learning parameterization of physics-based potentials [48]. Explicit correc-

tions for the missing penetration term in standard FFs were also introduced based on SAPT [49,50], a Gaussian electrostatic model (GEM), which uses density fitting to afford continuous description of molecular charge [51,52], a charge-distribution model based on a promolecule augmented with point charges [53] and a screened charge model with a molecular mechanics outer density screening algorithm [54]. In the context of hydrocarbon chemistry, the need to include charge penetration in FFs when modeling π–π and CH···π interactions in unsaturated hydrocarbons has been emphasized by Sherrill et al. in 2009 [55]. Accordingly, several potentials with accurate electrostatics treatment have been developed and successfully applied to describe the intermolecular interactions of anthracene [56], polycyclic aromatic hydrocarbons [57-60] and fullerene with graphite [61]. However, the importance of introducing the penetration effects in the molecular mechanics united-atom and all-atom force fields, commonly employed to describe the aliphatic systems [62,63], including such industrially relevant representatives as graphane [64] and polyethylene [65], is far less – if at all – recognized.

In the present work, we quantify the penetration energy in a diverse range of hydrocarbon dimers, including π-conjugated moieties and bulky aliphatic substituents. A direct one-to-one quantitative comparison between the fairly polarizable sp^2-rich (π-conjugated) and the much less polarizable sp^3-rich (aliphatic) systems demonstrates that charge penetration is important in both. While the energetic and structural consequences of neglecting this term are more drastic in the former, the resulting errors in the aliphatic dimer systems are nonetheless significant, i.e., ≈50% in interaction energy and 0.3 Å in interaction range. We discuss the implications of these results for the modeling of intermolecular interactions involving extended alkyl side chains, graphanes and various aliphatic systems in general.

Results and Discussion

Here, we investigate the nature of non-covalent interactions for a range of hydrocarbon dimers featuring both aromatic and aliphatic skeletons and bearing substituents, from methyl all the way to bulky adamantyl (Figure 2). First, we consider the dimers, constructed from the optimized monomers that are kept fixed (frozen) in terms of all geometry parameters, except for the intermonomer distance, *d*. Configurations, corresponding to the lowest total interaction energy, E_{tot}, in these constricted energy profiles (see Figures S1–S3 in Supporting Information File 1) are called 'frozen dimers' and are used in this work to compare the various systems on equal grounds. Second, to go beyond this somewhat constrained insight into the non-covalent interactions in the hydrocarbons, we relaxed the geometries of the frozen dimers. The resulting optimized dimers are, in general, structurally similar to the frozen counterparts albeit feature shorter interaction distances and in some cases undergo pronounced changes (e.g., lateral shifts and tilts) upon relaxation (see Figures S4 and S6 in Supporting Information File 1).

We start by considering the electrostatic (E_{elst}) and non-electrostatic ($E_{non\text{-}elst}$) contributions to the total SAPT0/jun-cc-pVDZ intermolecular interaction energies (E_{tot}) of the investigated dimers. The electrostatic part of E_{tot} consists of the distributed multipole and charge penetration terms ($E_{elst} = E_{DMA} + E_{Cpen}$), while the non-electrostatic contribution includes exchange, dispersion and induction terms ($E_{non\text{-}elst} = E_{exch} + E_{disp} + E_{ind}$). This breakdown allows us to discriminate between systems, driven by the non-electrostatic ($E_{elst} < E_{non\text{-}elst}$), and those, driven by the electrostatic ($E_{elst} > E_{non\text{-}elst}$) terms. Comparison between the two types of dimers – frozen and optimized – reveals the following three classes of hydrocarbons (Figure 3):

1. Non-substituted and substituted by comparatively non-bulky Me and *t*-Bu groups benzene and cyclohexane cores, **6a–c** and **9a–c**. These systems are associated with relatively small interaction energies, which do not change appreciably upon geometry relaxation. However, in their frozen dimers E_{tot} is dominated by the non-electrostatic term, while in their optimized assemblies electrostatics takes over. The intermonomer distances *d* are approx. 0.1–0.6 Å shorter in the optimized dimers, in which the monomers have the freedom to shift, e.g., laterally, compared to the frozen ones. The electrostatic term is thus almost entirely due to charge penetration (see Figure S7 in Supporting Information File 1), which increases exponentially at shorter range. The non-electrostatic term grows at a slower pace with shorter *d* since it depends both on exchange, which increases exponentially, and dispersion, which increases slower, i.e., as $1/d^6$ [66]. Similar behavior of E_{Cpen} and E_{exch} is rooted in their dependence on the extent of density overlap [67] and the following expression connecting them has been suggested as far back as 1970 [68]: $E_{exch} = -E_{Cpen}(a + b \times d)$, where *a* and *b* are empirical parameters. The extent to which this linearity holds depends on the rank of multipolar extension, used to compute the E_{Cpen} from E_{elst}, as well as the geometric features of the molecular core and its dimer [69].

2. Extended sp^2-rich cores **7**, **8**, **6f** and **9f**, which have the freedom to shift and get significantly closer upon geometry relaxation (see Figure S8 in Supporting Information File 1). Notably shorter interaction range in their optimized dimers (by 0.3–1.3 Å) compared to the frozen ones is associated with stabilizing effects, similar to those in class (I) above and thus dominated by enhanced charge penetration, albeit significantly amplified by the extended system size (Figure 4).

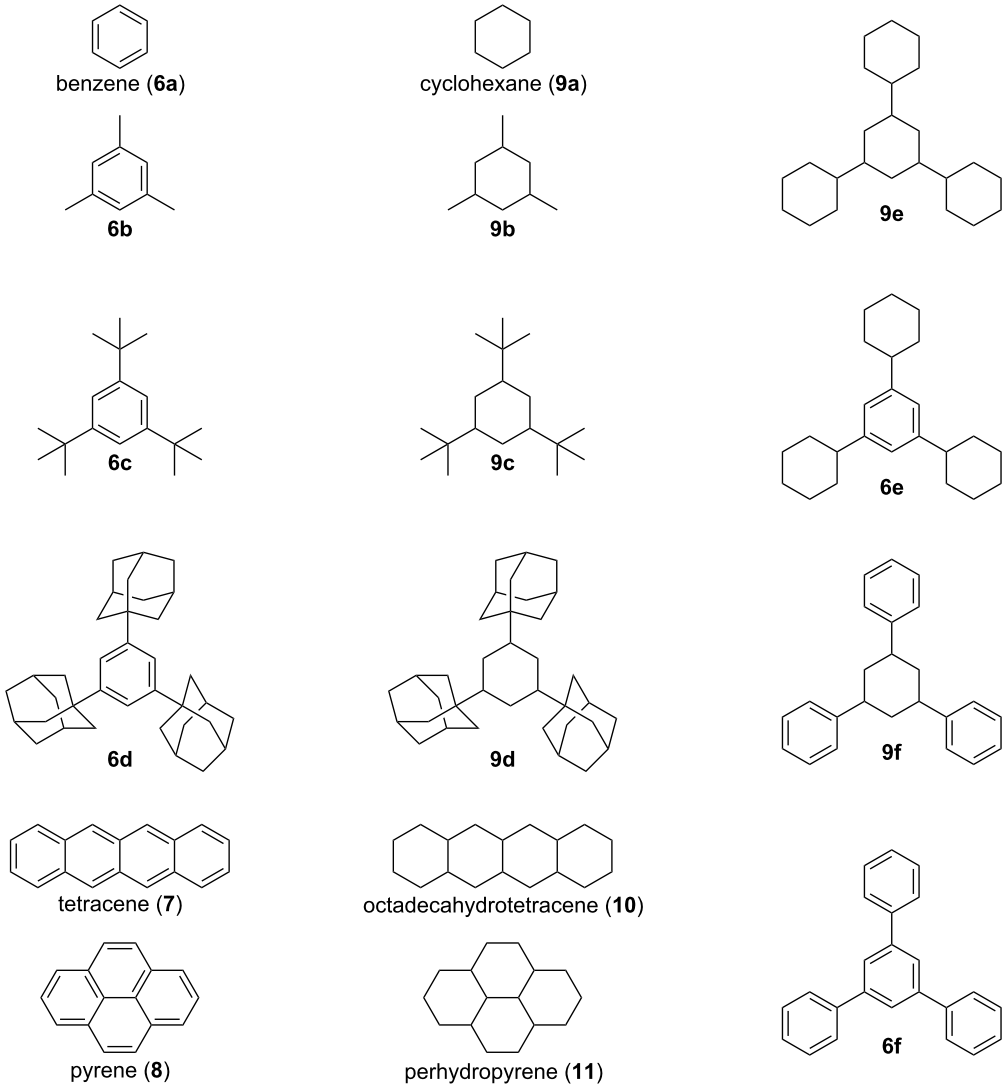


Figure 2: Studied monomer cores and their abbreviations, adopted here.

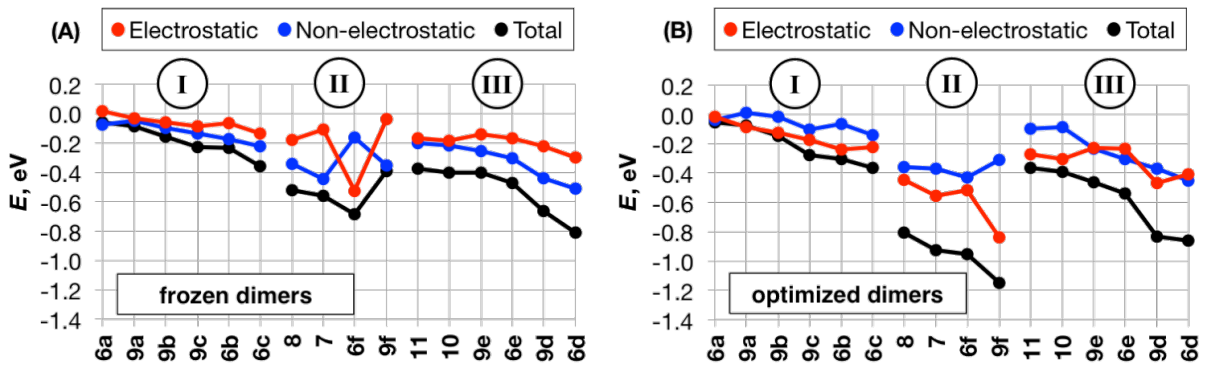


Figure 3: Breakdown of the SAPT0/jun-cc-pVDZ total interaction energies into electrostatic and non-electrostatic contributions in frozen (A) and optimized (B) dimers.

3. Bulky, sterically congested sp^3-rich cores **10**, **11**, **6d**,**e** and **9d**,**e**, which do not have the space to move considerably upon optimization. This results in a moderate decrease in the intermonomer distances and interaction energies upon geometry relaxation (by approx. 0.1–0.7 Å) due to the competition between destabilizing exchange and stabilizing dispersion, with the latter becoming increasingly dominant as the bulk of the substituents increases [3,4,28]. However, the associated shortening of the multiple CH···HC contacts between the bulky cyclohexyl and adamantyl units brings about appreciable – even dominant – electrostatic stabilization (Figure 4). The extent of charge penetration increases with the increasing number of close-range CH···HC contacts, e.g., from **9e** (1,3,5-tricyclohexylcyclohexane) to **10** (less bulky perhydrotetracene) to **9d** (1,3,5-triadamantylcyclohexane, see Figure S8 in Supporting Information File 1).

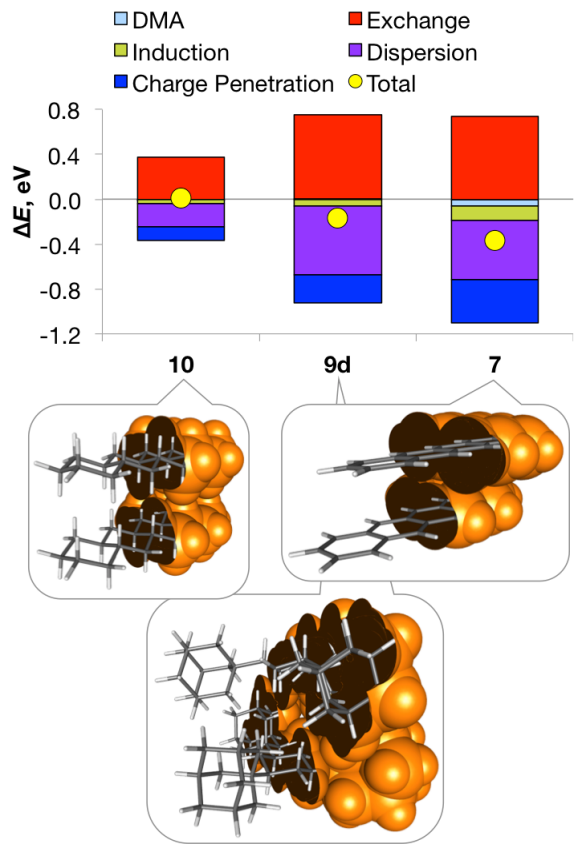


Figure 4: Decomposition of the SAPT0/jun-cc-pVDZ energy difference between the optimized and frozen dimers (*i.e.*, $\Delta E = E_{\text{total or component}}[\text{optimized}] - E_{\text{total or component}}[\text{frozen}]$) for the representative cores from classes (II) and (III), as well as the M06-2X/def2-SVP geometries of their optimized dimers (clipped van der Waals surfaces are shown in orange).

The crucial role of electrostatics is well appreciated in the interactions of polarizable π-conjugated cores [9,30]. Bulky sp^3-rich systems, despite being significantly less polarizable and generally featuring smaller (in absolute terms) stabilization, also involve appreciable electrostatic contribution from charge penetration both in their intra- [28] and intermolecular (Figure 4) assemblies.

The demonstrated quantitative significance of E_{Cpen} nonetheless does not directly reflect its qualitative importance, nor does it reveal the implications of this term for the chemical and physical properties of the bulky π-conjugated and saturated hydrocarbons. To address this question, we have compared the SAPT0/jun-cc-pVDZ energy profiles of the dimer interplanar separation for representative class (II) and class (III) systems: tetracene (**7**) and its fully saturated analogue, perhydrotetracene (**10**, Figure 5A and B). For each monomer, four types of dimers were compared – perfectly stacked (i.e., the frozen dimer) and shifted transversally, laterally and in both directions by approx. half the ring (Figure S11 in Supporting Information File 1). The main purpose of this exercise is to mimic the results of geometrical relaxation with and without the charge penetration for a diverse sample of dimer arrangements, i.e., beyond the model frozen and optimized geometries. For tetracene, the shifted dimer is energetically favored, while for perhydrotetracene the stacked dimer is preferred. Our results in Figure 5C comparing selected – shifted and stacked – dimers (see all four dimers comparison in Figure S11 in Supporting Information File 1) illustrate that in tetracene (**7**) neglecting the penetration effects would result in a drastic underestimation of the energy difference between the stacked and shifted dimer geometries; in perhydrotetracene (**10**), even though this term accounts for almost half the difference in E_{tot}, the relative error would be less significant. In terms of geometries of the energetically preferred dimers (Figure 5D), excluding E_{Cpen} leads to a longer interaction range both for **7** (by 0.5 Å) and, to a lesser extent, **10** (by 0.3 Å).

From the methodological viewpoint, these results have relevant implications on the use of existing and for the development of improved force fields (FFs) and other methods for the modeling and crystal structure prediction of hydrocarbons. In the case of π-conjugated complexes or assemblies, neglecting E_{Cpen} would flatten the potential energy surface dramatically. As shown in Figure 5, the penetration energy contribution strongly discriminates even between the slightly different (perfectly π-stacked and shifted by ≈1.2 Å in long and short axis) geometries. This contribution is therefore vital when exploring the free-energy landscape. While the energetic consequences are less pronounced for the sp^3-rich systems, the absence of charge penetration would lead to elongated intermolecular distances. This might potentially be one of the reasons why the molecular mechanics force fields, commonly applied to aliphatics (see

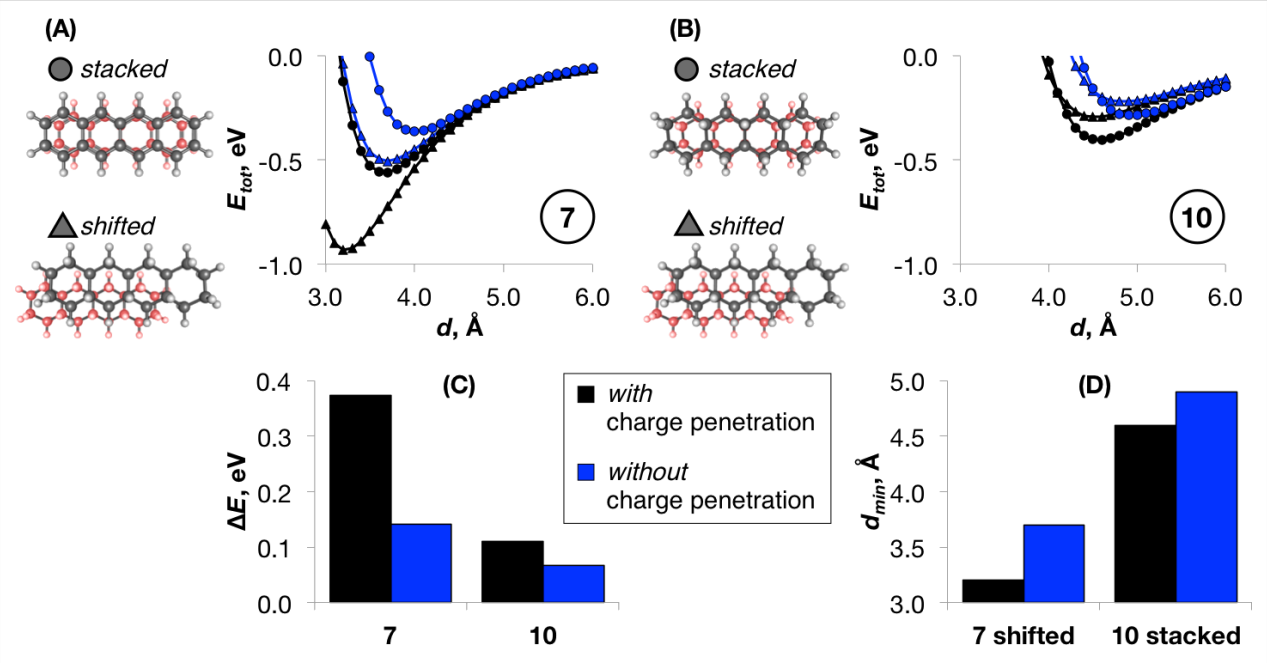


Figure 5: Structures and SAPT0/jun-cc-pVDZ interaction energy profiles with and without the charge penetration contribution vs varying intermonomer distance *d* for stacked and shifted dimers of tetracene (**7**, A) and perhydrotetracene (**10**, B). (C) The absolute SAPT0/jun-cc-pVDZ energetic difference between the energy minima on the shifted and stacked dimer profiles with and without the E_{Cpen}. (D) The interplanar separation, corresponding to the minimum energy dimer for **7** and **10** with and without the E_{Cpen}.

Introduction), significantly underestimate the liquid density and vapor pressure for long chain linear alkanes [70] and branched alkanes [71], fail to accurately reproduce the chain length dependence of the tilt and twist angles in alkanethiol self-assembled monolayers [72] and increasingly deviate (by as much as 15%) from experimental data for the hydrophobic solvation free energies of alkanes in alkanes with the increasing chain lengths [73].

Conclusion

The significance of stabilizing dispersion and electrostatic effects within sterically hindered hydrocarbons is well recognized. In such systems, electrostatic contributions are generally dominated by charge penetration, which increases with system size (bulk) and shorter interaction distances. In the present work, we have performed a direct comparison between the sp^2 and sp^3-rich hydrocarbons and quantified the E_{Cpen} term of their intermolecular interactions. The electrostatic effects are, not surprisingly, important in systems with strong π–π interactions. Our results illustrate that the less polarizable saturated hydrocarbon dimers with increasingly more and shorter CH···HC contacts can also be significantly influenced by electrostatics. In absolute terms, the penetration energy is greater in the π-conjugated systems and is thus crucial for the correct modeling of the energetic and structural properties of their bulk assemblies. In the bulky aliphatic systems, this contribution still constitutes a significant portion of the total interaction energy and accounts for approx. 0.3 Å difference in the interaction range. This re-emphasizes the importance of accounting for these effects even when modeling saturated hydrocarbons and provides the context for the underperformance of the molecular mechanics force fields, commonly applied to aliphatics.

Computational Details

Geometries of the isolated monomer cores were optimized at the M06-2X/def2-SVP level using Gaussian 09 software package [74]. The molar volume of each monomer was computed using the Monte-Carlo integration inside a contour of 0.001 electrons/$Bohr^3$ density in conjunction with M06-2X/def2-SVP density. The dimers were constructed from optimized monomers by translating one monomer with respect to another along the perpendicular axis and, in some systems, rotating it around this axis by 60° to achieved a staggered arrangement (for details, see Figure S3 in Supporting Information File 1). A range of intermonomer distances, *d*, was screened (3.0–6.0 Å or 4.0–7.0 Å depending on the system) with a 0.1 Å step size (see Figures S1and S2 in Supporting Information File 1). For each of these 'frozen' dimer geometries, the total interaction energy was evaluated using the method, considered a bronze standard for non-covalent interactions [75] – the zeroth-order symmetry-adapted perturbation theory (SAPT0) with jun-cc-pVDZ basis [76], which allows decomposing the total interaction energy E_{tot} into the exchange E_{exch}, electrostatic E_{elst}, dispersion E_{disp} and induction E_{ind}

components. SAPT0 computations were performed using the Psi4 code [77] and employed the density-fitting algorithm (DF-SAPT) [78,79]. For the distributed multipole analyses (DMA) [80] computations, the atom-centered multipoles up to the 8th-order were generated using Molpro [81] at the HF/6-311G** level (see also Figure S5A in Supporting Information File 1 regarding the different basis sets in SAPT0 and DMA computations). The multipole–multipole interaction energies were computed up to 32-poles (i.e., including all R^{-n} terms, where $n \leq 6$) using an in-house program of the Sherrill research group [31]. Charge penetration E_{Cpen} was evaluated as the difference between the electrostatic energy term of the SAPT0 total interaction energy, E_{elst}, and the DMA electrostatic term E_{DMA}. Furthermore, for each system the dimer with the lowest E_{tot} (called here the 'frozen dimer') was then used as a starting point for geometry relaxation at M06-2X/def2-SVP and PBE0-dDsC/def2-SVP levels, producing the 'optimized dimer'. The two methods were used to allow comparison of different dispersion treatments and yielded very similar results (for details, see Figure S5B–D in Supporting Information File 1); for consistency, the M06-2X results are discussed in the manuscript. Energy decomposition analyses for the optimized dimers were performed in the same way as for the frozen dimers.

Supporting Information

Supporting Information File 1

Additional figures, complete set of computed data and geometries of the studied monomers and dimers.
[https://www.beilstein-journals.org/bjoc/content/supplementary/1860-5397-14-125-S1.pdf]

Acknowledgements

This work has received support from EPFL and funding from the European Union's Horizon 2020 research and innovation program under the Marie Skłodowska-Curie grant agreement No. 701885 "Single Molecule Junctions with Non-Conventional Architectures, Crafted *in silico*" and from the European Research Council (ERC Grant 306528 COMPOREL). The authors also thank Prof. Konrad Patkowski (Auburn University) and Mr. Kun-Han Lin (EPFL) for helpful comments.

ORCID® iDs

Ganna Gryn'ova - https://orcid.org/0000-0003-4229-939X
Clémence Corminboeuf - https://orcid.org/0000-0001-7993-2879

References

1. Wagner, J. P.; Schreiner, P. R. *Angew. Chem., Int. Ed.* **2015,** *54,* 12274–12296. doi:10.1002/anie.201503476
2. Ehrlich, S.; Bettinger, H. F.; Grimme, S. *Angew. Chem., Int. Ed.* **2013,** *52,* 10892–10895. doi:10.1002/anie.201304674
3. Grimme, S.; Schreiner, P. R. *Angew. Chem., Int. Ed.* **2011,** *50,* 12639–12642. doi:10.1002/anie.201103615
4. Rösel, S.; Balestrieri, S.; Schreiner, P. R. *Chem. Sci.* **2017,** *8,* 405–410. doi:10.1039/C6SC02727J
5. 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. *Nature* **2011,** *477,* 308–311. doi:10.1038/nature10367
6. Fokin, A. A.; Zhuk, T. S.; Blomeyer, S.; Pérez, C.; Chernish, L. V.; Pashenko, A. E.; Antony, J.; Vishnevskiy, Y. V.; Berger, R. J. F.; Grimme, S.; Logemann, C.; Schnell, M.; Mitzel, N. W.; Schreiner, P. R. *J. Am. Chem. Soc.* **2017,** *139,* 16696–16707. doi:10.1021/jacs.7b07884
7. Fokin, A. A.; Chernish, L. V.; Gunchenko, P. A.; Tikhonchuk, E. Y.; Hausmann, H.; Serafin, M.; Dahl, J. E. P.; Carlson, R. M. K.; Schreiner, P. R. *J. Am. Chem. Soc.* **2012,** *134,* 13641–13650. doi:10.1021/ja302258q
8. Allinger, N. L.; Lii, J.-H.; Schaefer, H. F., III. *J. Chem. Theory Comput.* **2016,** *12,* 2774–2778. doi:10.1021/acs.jctc.5b00926
9. Rösel, S.; Quanz, H.; Logemann, C.; Becker, J.; Mossou, E.; Cañadillas-Delgado, L.; Caldeweyher, E.; Grimme, S.; Schreiner, P. R. *J. Am. Chem. Soc.* **2017,** *139,* 7428–7431. doi:10.1021/jacs.7b01879
10. Firouzi, R.; Shahbazian, S. *ChemPhysChem* **2016,** *17,* 51–54. doi:10.1002/cphc.201501002
11. Grimme, S. *Angew. Chem., Int. Ed.* **2008,** *47,* 3430–3434. doi:10.1002/anie.200705157
12. Preuss, K. E. *Polyhedron* **2014,** *79,* 1–15. doi:10.1016/j.poly.2014.04.005
13. Wolters, L. P.; Koekkoek, R.; Bickelhaupt, F. M. *ACS Catal.* **2015,** *5,* 5766–5775. doi:10.1021/acscatal.5b01354
14. Cui, Z.; Lischka, H.; Beneberu, H. Z.; Kertesz, M. *J. Am. Chem. Soc.* **2014,** *136,* 12958–12965. doi:10.1021/ja505624y
15. Fokin, A. A.; Gerbig, D.; Schreiner, P. R. *J. Am. Chem. Soc.* **2011,** *133,* 20036–20039. doi:10.1021/ja206992j
16. Wagner, J. P.; Schreiner, P. R. *J. Chem. Theory Comput.* **2014,** *10,* 1353–1358. doi:10.1021/ct5000499
17. Wang, C.; Mo, Y.; Wagner, J. P.; Schreiner, P. R.; Jemmis, E. D.; Danovich, D.; Shaik, S. *J. Chem. Theory Comput.* **2015,** *11,* 1621–1630. doi:10.1021/acs.jctc.5b00075
18. Echeverría, J.; Aullón, G.; Danovich, D.; Shaik, S.; Alvarez, S. *Nat. Chem.* **2013,** *3,* 323–330. doi:10.1038/nchem.1004
19. Danovich, D.; Shaik, S.; Neese, F.; Echeverria, J.; Aullón, G.; Alvarez, S. *J. Chem. Theory Comput.* **2013,** *9,* 1977–1991. doi:10.1021/ct400070j
20. Janowski, T.; Pulay, P. *J. Am. Chem. Soc.* **2012,** *134,* 17520–17525. doi:10.1021/ja303676q
21. Alonso, M.; Woller, T.; Martín-Martínez, F. J.; Contreras-García, J.; Geerlings, P.; De Proft, F. *Chem. – Eur. J.* **2014,** *20,* 4931–4941. doi:10.1002/chem.201400107
22. Li, Y.; Chen, Z. *J. Phys. Chem. Lett.* **2013,** *4,* 269–275. doi:10.1021/jz301821n
23. Pastorczak, E.; Prlj, A.; Gonthier, J. F.; Corminboeuf, C. *J. Chem. Phys.* **2015,** *143,* 224107. doi:10.1063/1.4936830
24. Ninković, D. B.; Vojislavljević-Vasilev, D. Z.; Medaković, V. B.; Hall, M. B.; Brothers, E. N.; Zarić, S. D. *Phys. Chem. Chem. Phys.* **2016,** *18,* 25791–25795. doi:10.1039/C6CP03734H
25. Podeszwa, R. *J. Chem. Phys.* **2010,** *132,* 044704. doi:10.1063/1.3300064

26. Cabaleiro-Lago, E. M.; Rodríguez-Otero, J. *ChemistrySelect* **2017,** *2,* 5157–5166. doi:10.1002/slct.201700671
27. Li, F.; Li, Y. *Theor. Chem. Acc.* **2016,** *135,* 90. doi:10.1007/s00214-016-1851-7
28. Parrish, R. M.; Gonthier, J. M.; Corminbœuf, C.; Sherrill, C. D. *J. Chem. Phys.* **2015,** *143,* 051103. doi:10.1063/1.4927575
29. Sherrill, C. D. *Acc. Chem. Res.* **2013,** *46,* 1020–1028. doi:10.1021/ar3001124
30. Hohenstein, E. G.; Duan, J.; Sherrill, C. D. *J. Am. Chem. Soc.* **2011,** *133,* 13244–13247. doi:10.1021/ja204294q
31. Parker, T. M.; Hohenstein, E. G.; Parrish, R. M.; Hud, N. V.; Sherrill, C. D. *J. Am. Chem. Soc.* **2013,** *135,* 1306–1316. doi:10.1021/ja3063309
32. Jing, Z.; Qi, R.; Liu, C.; Ren, P. *J. Chem. Phys.* **2017,** *147,* 161733. doi:10.1063/1.4985921
33. Gryn'ova, G.; Corminboeuf, C. *J. Phys. Chem. Lett.* **2016,** *7,* 5198–5204. doi:10.1021/acs.jpclett.6b02585
34. Xia, J. L.; Liu, S. H.; Cozzi, F.; Mancinelli, M.; Mazzanti, A. *Chem. – Eur. J.* **2012,** *18,* 3611–3620. doi:10.1002/chem.201103639
35. Sherman, M. C.; Ams, M. R.; Jordan, K. D. *J. Phys. Chem. A* **2016,** *120,* 9292–9298. doi:10.1021/acs.jpca.6b09193
36. Slipchenko, L.; Gordon, M. S. *J. Comput. Chem.* **2006,** *28,* 276–291. doi:10.1002/jcc.20520
37. Smith, Q. A.; Gordon, M. S.; Slipchenko, L. V. *J. Phys. Chem. A* **2011,** *115,* 4598–4609. doi:10.1021/jp201039b
38. Grimme, S.; Bannwarth, C.; Caldeweyher, E.; Pisarek, J.; Hansen, A. *J. Chem. Phys.* **2017,** *147,* 161708. doi:10.1063/1.4991798
39. Grimme, S. *J. Chem. Theory Comput.* **2014,** *10,* 4497–4514. doi:10.1021/ct500573f
40. Bukowski, R.; Szalewicz, K.; Groenenboom, G. C.; van der Avoird, A. *Science* **2007,** *315,* 1249–1252. doi:10.1126/science.1136371
41. Donchev, A. G.; Galkin, N. G.; Illarionov, A. A.; Khoruzhii, O. V.; Olevanov, M. A.; Ozrin, V. D.; Pereyaslavets, L. B.; Tarasov, V. I. *J. Comput. Chem.* **2008,** *29,* 1242–1249. doi:10.1002/jcc.20884
42. Taylor, D. E.; Rob, F.; Rice, B. M.; Podeszwa, R.; Szalewicz, K. *Phys. Chem. Chem. Phys.* **2011,** *13,* 16629–16636. doi:10.1039/c1cp21342c
43. McDaniel, J. G.; Schmidt, J. R. *J. Phys. Chem. A* **2013,** *117,* 2053–2066. doi:10.1021/jp3108182
44. Szalewicz, K. *Acc. Chem. Res.* **2014,** *47,* 3266–3274. doi:10.1021/ar500275m
45. Van Vleet, M. J.; Misquitta, A. J.; Stone, A. J.; Schmidt, J. R. *J. Chem. Theory Comput.* **2016,** *12,* 3851–3870. doi:10.1021/acs.jctc.6b00209
46. Aina, A. A.; Misquitta, A. J.; Price, S. L. *J. Chem. Phys.* **2017,** *147,* 161722. doi:10.1063/1.4999789
47. Vandebrande, S.; Waroquier, M.; Van Speybroeck, V.; Verstraelen, T. *J. Chem. Theory Comput.* **2017,** *13,* 161–179. doi:10.1021/acs.jctc.6b00969
48. Bereau, T.; DiStasio, R. A., Jr.; Tkachenko, A.; von Lilienfeld, O. A. *J. Chem. Phys.* **2018,** *148,* 241706. doi:10.1063/1.5009502
49. Wang, Q.; Rackers, J. A.; He, C.; Qi, R.; Narth, C.; Lagardere, L.; Gresh, N.; Ponder, J. W.; Piquemal, J.-P.; Ren, P. *J. Chem. Theory Comput.* **2015,** *11,* 2609–2618. doi:10.1021/acs.jctc.5b00267
50. Rackers, J. A.; Wang, Q.; Liu, C.; Piquemal, J.-P.; Ren, P.; Ponder, J. W. *Phys. Chem. Chem. Phys.* **2017,** *19,* 276–291. doi:10.1039/C6CP06017J
51. Cisneros, G. A. *J. Chem. Theory Comput.* **2012,** *8,* 5072–5080. doi:10.1021/ct300630u
52. Chaudret, R.; Gresh, N.; Narth, C.; Lagardère, L.; Darden, T. A.; Cisneros, G. A.; Piquemal, J.-P. *J. Phys. Chem. A* **2014,** *118,* 7598–7612. doi:10.1021/jp5051657
53. Bojarowski, S. A.; Kumar, P.; Dominiak, P. M. *ChemPhysChem* **2016,** *17,* 2455–2460. doi:10.1002/cphc.201600390
54. Wang, B.; Truhlar, D. G. *J. Chem. Theory Comput.* **2014,** *10,* 4480–4487. doi:10.1021/ct5005142
55. Sherrill, C. D.; Sumpter, B. G.; Sinnokrot, M. O.; Marshall, M. S.; Hohenstein, E. G.; Walker, R. C.; Gould, I. R. *J. Comput. Chem.* **2009,** *30,* 2187–2193. doi:10.1002/jcc.21226
56. Grančič, P.; Bylsma, R.; Meekes, H.; Cuppen, H. M. *Cryst. Growth Des.* **2015,** *15,* 1625–1633. doi:10.1021/cg5013507
57. Totton, T. S.; Misquitta, A. J.; Kraft, M. *J. Chem. Theory Comput.* **2010,** *6,* 683–695. doi:10.1021/ct9004883
58. Totton, T. S.; Misquitta, A. J.; Kraft, M. *Phys. Chem. Chem. Phys.* **2012,** *14,* 4081–4094. doi:10.1039/c2cp23008a
59. Tafipolsky, M.; Engels, B. *J. Chem. Theory Comput.* **2011,** *7,* 1791–1803. doi:10.1021/ct200185h
60. Pascazio, L.; Sirignano, M.; D'Anna, A. *Combust. Flame* **2017,** *185,* 53–62. doi:10.1016/j.combustflame.2017.07.003
61. Donchev, A. G. *Phys. Rev. B* **2006,** *74,* 235401. doi:10.1103/PhysRevB.74.235401
62. Chilukoti, H. K.; Kikugawa, G.; Ohara, T. *J. Phys. Chem. B* **2016,** *120,* 7207–7216. doi:10.1021/acs.jpcb.6b05332
63. Janeček, J.; Paricaud, P. *Fluid Phase Equilib.* **2016,** *429,* 27–36. doi:10.1016/j.fluid.2016.08.023
64. Artyukhov, V. I.; Chernozatonskii, L. A. *J. Phys. Chem. A* **2010,** *114,* 5389–5396. doi:10.1021/jp1003566
65. Kumar, V.; Locker, C. R.; in 't Veld, P. J.; Rutledge, G. C. *Macromolecules* **2017,** *50,* 1206–1214. doi:10.1021/acs.macromol.6b02458
66. Ryno, S. M.; Risko, C.; Brédas, J.-L. *Chem. Mater.* **2016,** *28,* 3990–4000. doi:10.1021/acs.chemmater.6b01340
67. Freitag, M. A.; Gordon, M. S.; Jensen, J. H.; Stevens, W. J. *J. Chem. Phys.* **2000,** *112,* 7300–7306. doi:10.1063/1.481370
68. Murrell, J. N.; Teixeira-Dias, J. J. C. *Mol. Phys.* **1970,** *19,* 521–531. doi:10.1080/00268977000101531
69. Misquitta, A. J.; Stone, A. J.; Fazeli, F. *J. Chem. Theory Comput.* **2014,** *10,* 5405–5418. doi:10.1021/ct5008444
70. Müller, E. A.; Mejia, A. *J. Phys. Chem. B* **2011,** *115,* 12822–12834. doi:10.1021/jp203236q
71. Mick, J. R.; Barhaghi, M. S.; Jackman, B.; Schwiebert, L.; Potoff, J. J. *J. Chem. Eng. Data* **2017,** *62,* 1806–1818. doi:10.1021/acs.jced.6b01036
72. Bhadra, P.; Siu, S. W. I. *J. Phys. Chem. C* **2017,** *121,* 26340–26349. doi:10.1021/acs.jpcc.7b08092
73. Jorge, M.; Nuno, M. G.; Simões, C. J. V.; Silva, C. G.; Brito, R. M. M. *J. Comput. Chem.* **2017,** *38,* 346–358. doi:10.1002/jcc.24690
74. *Gaussian 09,* Revision D.01; Gaussian, Inc.: Wallingford CT, 2009.
75. Parker, T. M.; Burns, L. A.; Parrish, R. M.; Ryno, A. G.; Sherrill, C. D. *J. Chem. Phys.* **2014,** *140,* 094106. doi:10.1063/1.4867135
76. Jeziorski, B.; Moszynski, R.; Szalewicz, K. *Chem. Rev.* **1994,** *94,* 1887–1930. doi:10.1021/cr00031a008
77. Turney, J. M.; Simmonett, A. C.; Parrish, R. M.; Hohenstein, E. G.; Evangelista, F. A.; Fermann, J. T.; Mintz, B. J.; Burns, L. A.; Wilke, J. J.; Abrams, M. L.; Russ, N. J.; Leininger, M. L.; Janssen, C. L.; Seidl, E. T.; Allen, W. D.; Schaefer, H. F.; King, R. A.; Valeev, E. F.; Sherrill, C. D.; Crawford, T. D. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2012,** *2,* 556–565. doi:10.1002/wcms.93

78. Hohenstein, E. G.; Parrish, R. M.; Sherrill, C. D.; Turney, J. M.; Schaefer, H. F., III. *J. Chem. Phys.* **2011,** *135,* 174017. doi:10.1063/1.3656681
79. Hohenstein, E. G.; Sherrill, C. D. *J. Chem. Phys.* **2010,** *132,* 184111. doi:10.1063/1.3426316
80. Stone, A. J. *Chem. Phys. Lett.* **1981,** *83,* 233–239. doi:10.1016/0009-2614(81)85452-8
81. *MOLPRO,* Version 2012.1; a package of ab initio programs, http://www.molpro.net.

License and Terms

BEILSTEIN JOURNAL OF ORGANIC CHEMISTRY

Cobalt-catalyzed C–H cyanations: Insights into the reaction mechanism and the role of London dispersion

Eric Detmar[1], Valentin Müller[2], Daniel Zell[2], Lutz Ackermann[*2] and Martin Breugst[*1]

Full Research Paper

Open Access

Address:
[1]Department für Chemie, Universität zu Köln, Greinstraße 4, 50939 Köln, Germany and [2]Institut für Organische und Biomolekulare Chemie, Georg-August-Universität Göttingen, Tammannstraße 2, 37077 Göttingen, Germany

Email:
Lutz Ackermann[*] - Lutz.Ackermann@chemie.uni-goettingen.de;
Martin Breugst[*] - mbreugst@uni-koeln.de.

* Corresponding author

Keywords:
catalysis; C–H activation; density functional theory; London dispersion; reaction mechanisms

Beilstein J. Org. Chem. **2018,** *14,* 1537–1545.
doi:10.3762/bjoc.14.130

Received: 27 February 2018
Accepted: 02 June 2018
Published: 25 June 2018

This article is part of the Thematic Series "Dispersion interactions".

Guest Editor: P. Schreiner

Abstract

Carboxylate-assisted cobalt(III)-catalyzed C–H cyanations are highly efficient processes for the synthesis of (hetero)aromatic nitriles. We have now analyzed the cyanation of differently substituted 2-phenylpyridines in detail computationally by density functional theory and also experimentally. Based on our investigations, we propose a plausible reaction mechanism for this transformation that is in line with the experimental observations. Additional calculations, including NCIPLOT, dispersion interaction densities, and local energy decomposition analysis, for the model cyanation of 2-phenylpyridine furthermore highlight that London dispersion is an important factor that enables this challenging C–H transformation. Nonbonding interactions between the Cp* ligand and aromatic and C–H-rich fragments of other ligands at the cobalt center significantly contribute to a stabilization of cobalt intermediates and transition states.

Introduction

For a long time, large and bulky substituents have intuitively been considered to act through unfavorable steric interactions, although London dispersion – the attractive part of the van-der-Waals interaction – is known for more than 100 years [1,2]. The stabilizing nature of C–H···H–C interactions and their importance for organic transformations has only been fully realized within the last decades [3]. Among others, these interactions explain the hexaarylethane riddle [4] and are responsible for the high stability of singly bonded diamondoid dimers resulting in very long C–C bonds [5,6], or very short H···H contacts in

tris(3,5-di-*tert*-butylphenyl)methane [7]. Besides a remarkable effect on organic structures, dispersion can also affect the outcome of chemical transformations. Presumably due to attractive dispersive interactions between two adamantyl groups in the transition state of a [4 + 2] cycloaddition of benzynes (Scheme 1), the seemingly sterically more hindered product is formed preferentially [8].

Similar to other noncovalent interactions [9-11], London dispersion can also play a crucial role in different transition-metal-catalyzed reactions [12-17]. The C–H-rich di-1-adamantylphosphine oxide – a typical dispersion element – was experimentally found to be an excellent preligand for ruthenium- and palladium-catalyzed C–H functionalizations [18-23]. Similarly, computational studies revealed the importance of dispersion effects in palladium-catalyzed cross-coupling reactions [24-27]. For example, the contribution of London dispersion (up to 37 kcal mol^{-1}) has a huge influence on the ligand dissociation process within the $Pd(PPh_3)_4$ system [25]. Furthermore, only the results obtained from dispersion-corrected density functional theory [28,29] were in agreement with the experimental observations and dispersion reduces the activation free energies by up to 30 kcal mol^{-1} [27].

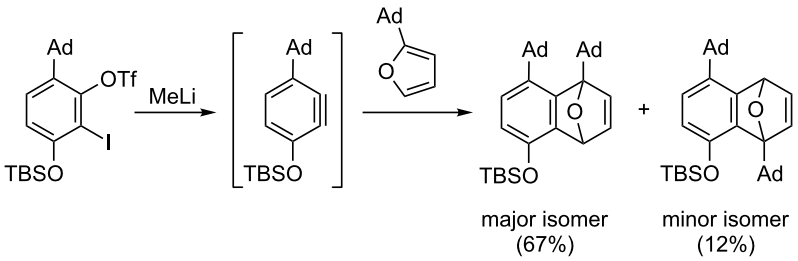
Currently, the strategic application of London dispersion in catalysis is still very difficult to achieve and, as a consequence, detailed insights in how dispersion influences organic reactions continue to be in high demand. Therefore, we have computationally analyzed the recently developed cobalt-catalyzed C–H cyanation of arenes (Scheme 2) [30-34]. Dispersion effects can be envisioned to be highly important in this system, as the relatively C–H-rich ligand Cp* can interact with both substrates within the cobalt complexes. In 2015, Li and Ackermann have proposed the catalytic cycle (C–H cobaltation, ligand coordination, insertion) shown in Scheme 2 which served as the starting point of this investigation [30]. We now report on our computational findings supported by novel kinetic investigations to establish the reaction mechanism of this synthetically useful C–H activation and to elucidate the role of London dispersion in these transformations.

Results and Discussion

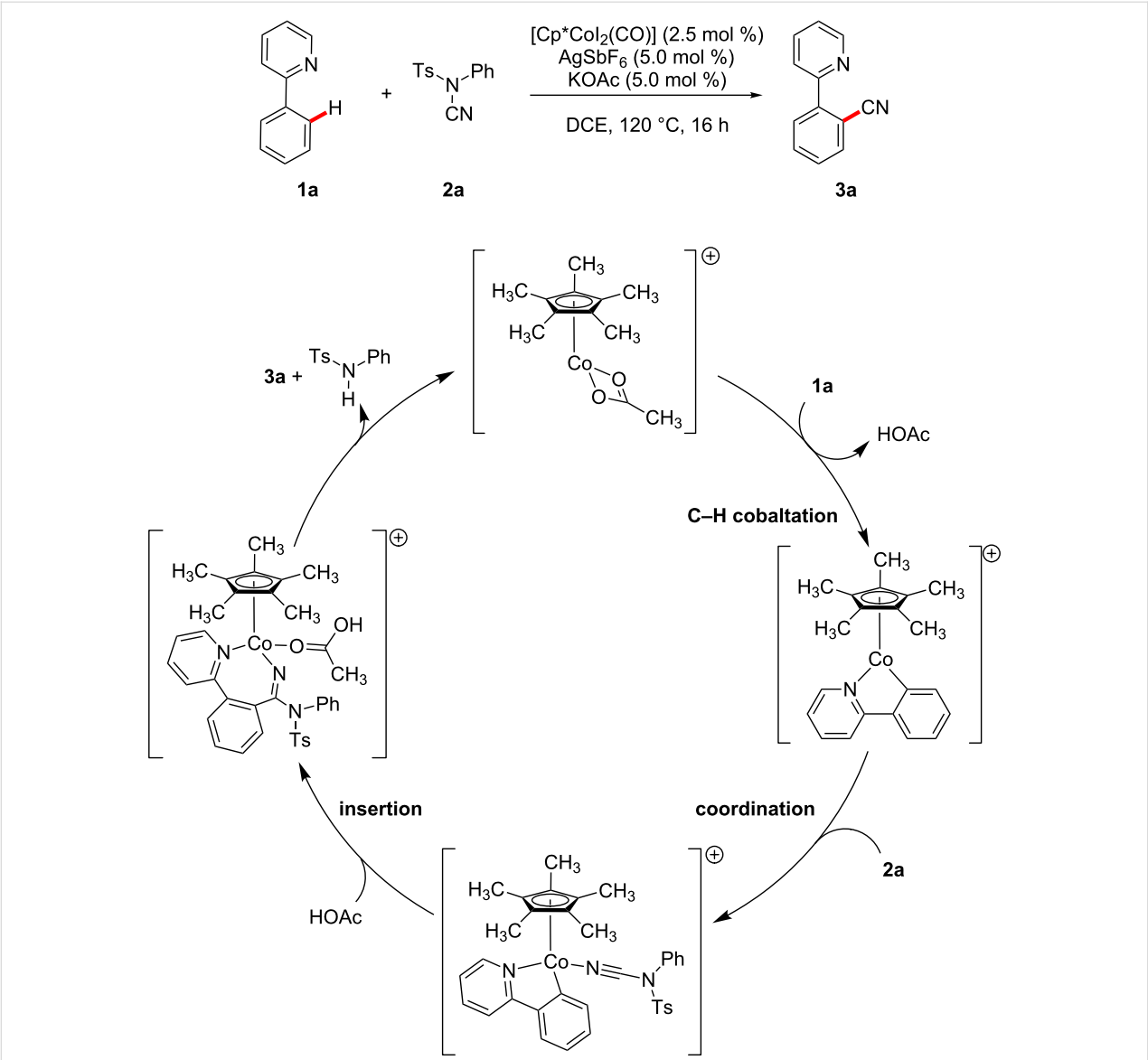
Analysis of the reaction mechanism

To unravel the importance of London dispersion on the cobalt-catalyzed C–H cyanation of 2-phenylpyridine (**1a**), the underlying catalyst's mode of action has to be fully understood. The available experimental data indicated a reversible C–H metalation, which led to the suggested catalytic cycle of Scheme 2 [30]. As computational investigations also allow the study of intermediates that are too unstable to be observed under the experimental conditions, we have analyzed the underlying reaction mechanism in more detail employing density functional theory. A complete free energy profile on the B3LYP-D3BJ/def2-QZVP/COSMO//B3LYP-D3BJ/def2-TZVP potential energy surface is depicted in Figure 1 (black line), while the free-energy profile on the M06-L surface is summarized in Supporting Information File 1. Selected intermediates and transition states are shown in Figure 2.

The computational analysis starts with the catalytically active cobalt(III) acetate complex **4** which is generated in situ from the precatalyst [Cp*CoI_2(CO)], $AgSbF_6$, and KOAc. While the iodine ions are captured by Ag^+, carbon monoxide dissociates and leaves the reaction mixture as a gas. Although the SbF_6^- counter ion to the cationic cobalt complexes is considered to be weakly coordinating [37], specific interactions cannot be completely ruled out. We have assumed that all of the positively charged cobalt complexes on the reaction path are similarly affected by ion pairing and therefore, we base the following investigation mainly on the reactions of the cobalt complexes and do not include ion pairing in our analysis. Coordination of 2-phenylpyridine (**1a**) to this 16-electron species leads to the intermediate **5a** (Figure 2) in a highly exergonic reaction step (ΔG = −18.8 kcal mol^{-1}), which also is the resting state of the catalytic cycle. This intermediate could therefore be amenable to spectroscopic characterization. Based on our computational analysis, the subsequent C–H cobaltation (**5a** → **7a**) is endergonic ($\Delta\Delta G$ = +12 kcal mol^{-1}) and proceeds in a step-wise fashion. A similar mechanism has previously been described by McMullin, Williams, and Frost [38], as well as by Ackermann [39,40] for ruthenium-catalyzed C–H alkenylations. In the first



Scheme 1: Cycloaddition reaction of in situ generated benzynes resulting in the sterically more hindered adduct (Ad = 1-adamantyl) [8].



Scheme 2: Recently developed cobalt-catalyzed C–H cyanation [30].

transition state (**TS1a**, Figure 2), the κ^2-coordination of the acetate ligand changes to a κ^1-coordination. The resulting intermediate **6a** is stabilized by an agostic interaction between the C–H bond and the metal atom as well as by an additional weak hydrogen bond between the C–H bond and the acetate oxygen (O···H distance 2.26 Å). A natural population analysis of structure **6a** further confirms the stabilizing nature of these interactions. In the second transition state **TS2a** (Figure 2), the C–H bond is broken and the proton is transferred to the acetate which results in the formation of the cobaltacycle **7a**.

Acetic acid dissociates, and *N*-cyano-*N*-phenyl-*p*-toluenesulfonamide (**2a**) coordinates to the 16-electron intermediate **8a** yielding **9a**. Next, the insertion of the cyanating agent **2a** into the cobalt–carbon bond takes place through **TS3a**. Within the four-membered transition state (Figure 2), the C–C bond to be formed is still rather long (C–C distance 1.92 Å), while the C–N distance is already significantly elongated (1.15 Å in **9a**, 1.22 Å in **TS3a**, 1.26 Å in **10a**). Furthermore, a significant reorganization has to take place during this step: the former almost linear N–C–N fragment (179.2°) changes to 137.9° in **TS3a** and 124.4° in **10a**, which results in a high barrier for this step. Subsequent coordination of acetic acid leads to intermediate **11a**. No transition states could be obtained for the following β-elimination and proto-demetalation resulting in product **3a**, the cobalt(III) acetate complex **4**, and *N*-phenyl-*p*-toluenesulfonamide. All attempts starting from different potential transition state structures resulted in barrierless reactions

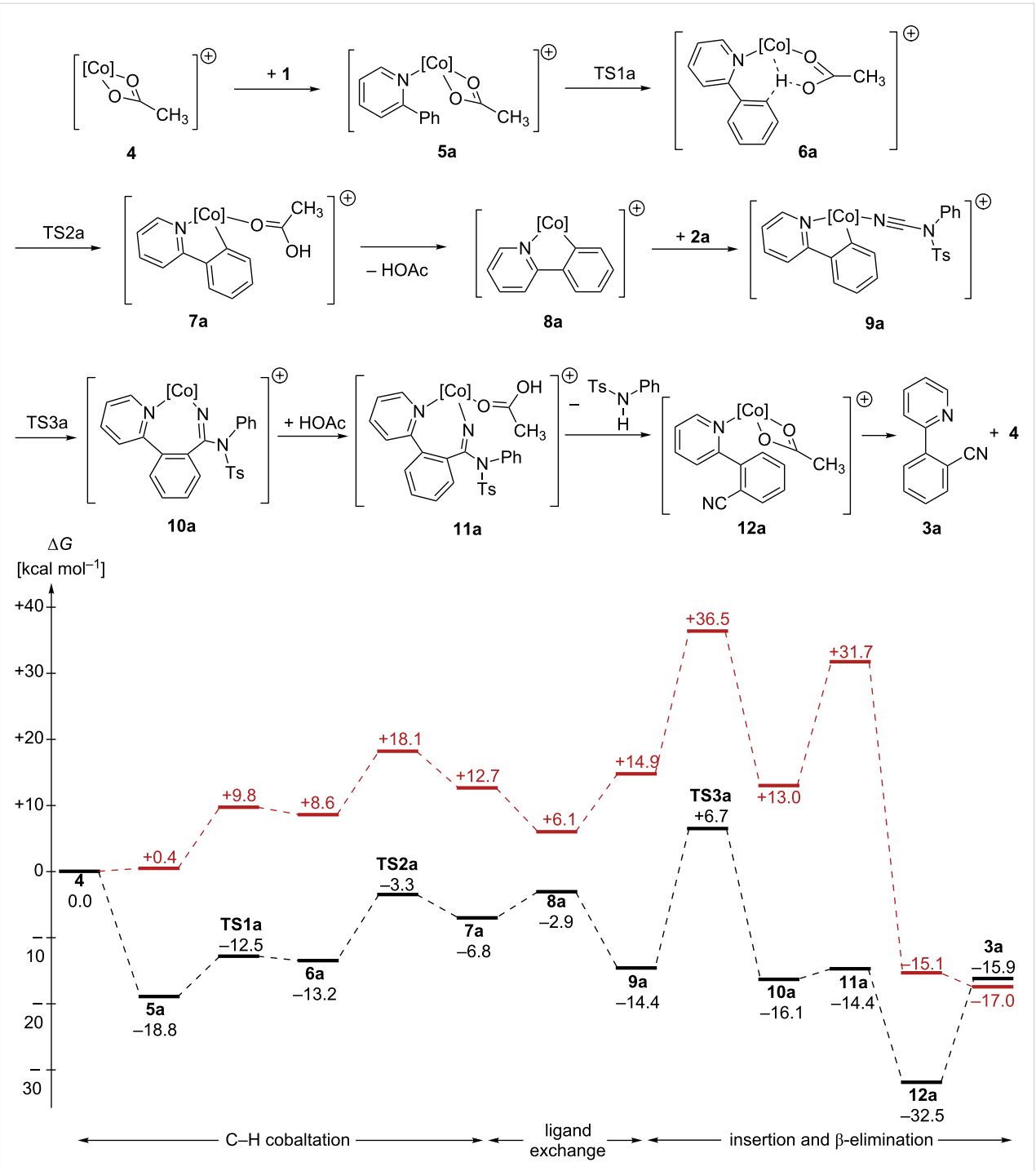


Figure 1: Calculated free-energy profile for the cobalt-catalyzed C–H cyanation of 2-phenylpyridine (**1a**) [in kcal mol^{-1}, [Co] = Cp*Co, black lines indicate that dispersion (D3 correction with Becke–Johnson damping) [35,36] was included in the calculations while red lines indicate that dispersion was not included].

when a proton approaches the amidine substructure (→ **12a**). As the cyanated 2-phenylpyridine **3a** is less Lewis-basic compared to the starting material **1a**, **12a** could also react with **1a** in a thermodynamically favorable ligand exchange reaction (ΔG = −2.2 kcal mol^{-1}) to yield complex **5a**.

In contrast to previous computational studies on manganese(I)-catalyzed fluoro-allylation reactions where β-fluoride and HF eliminations played an important role [41], similar reactions involving amine eliminations seem to be not relevant in this reaction. Furthermore, a comparison with previous computational

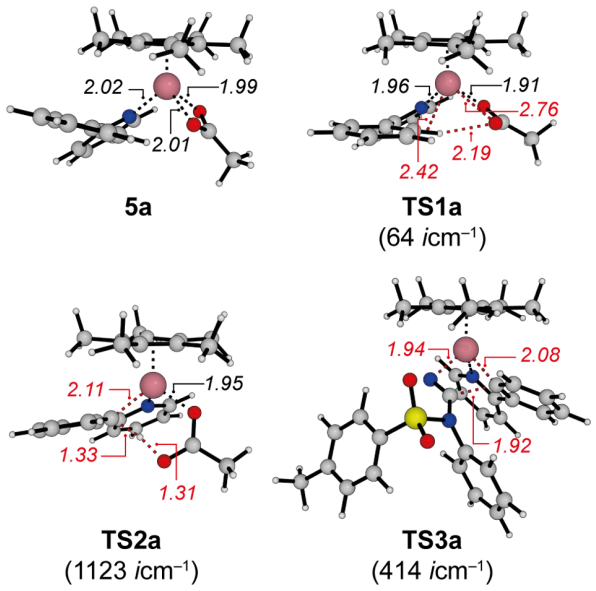
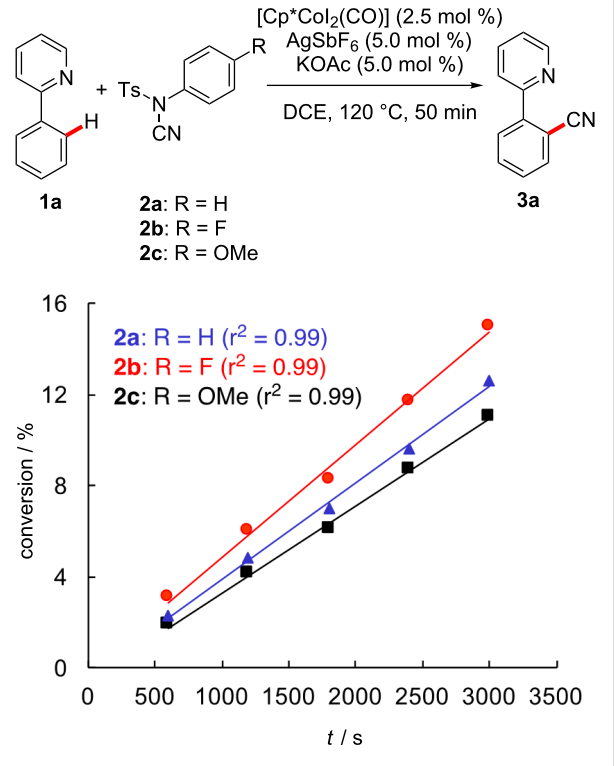


Figure 2: Calculated structures, selected bond lengths (in Å), and imaginary frequencies for representative intermediates and transition states for the cobalt-catalyzed C–H cyanation of 2-phenylpyridine (**1a**).



Scheme 3: Kinetic profile of the cobalt-catalyzed C–H cyanation with differently substituted cyanating agents **2**.

investigations on copper-catalyzed *ortho* C–H cyanations of vinylarenes revealed that those reactions take place via a completely different mechanism involving two distinct catalytic cycles (copper-catalyzed electrophilic cyanative dearomatization and base-catalyzed hydrogen transposition) [42,43].

Inspired by this computational analysis, we experimentally probed the effect of differently substituted cyanation agents **2** on the kinetics of the cobalt(III)-catalyzed C–H cyanation (Scheme 3). Thus, we observed that electron-withdrawing groups significantly facilitated the desired transformation. As the calculated rate-limiting transition state **TS3** benefits from a stabilization of the developing negative charge on the sulfonamide, the relative rates of Scheme 3 provide further support for the migratory insertion representing the rate determining step [44].

As differently substituted 2-phenylpyridines **1** have been employed experimentally, we included five representative substrates (R = H, CH_3, F, $C(O)CH_3$, CN) into the computational analysis as well. For these calculations, only one functional (B3LYP-D3BJ) and a smaller basis set (def2-SVP for non-metals and def2-TZVP for Co) were employed during the optimization to reduce the computational cost. These results are summarized in Table 1.

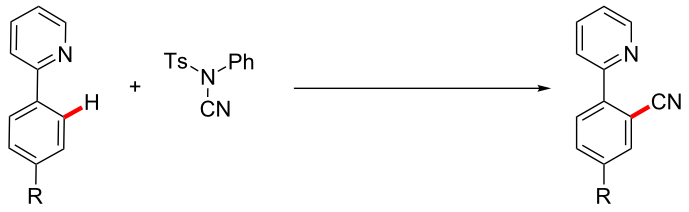
For the unsubstituted 2-phenylpyridine (**1a**), both computational methods (Figure 1 and Table 1) and the optimized structures are generally rather similar to one another. Based on the computational analysis depicted in Table 1, the turnover-limiting step for all substrates **1** is represented by the insertion of the cyanating agent **2a** into the cobalt–carbon bond, which can also be concluded based on the kinetic data of Scheme 3.

Based on the computational analysis of Figure 1 and the experimental data depicted in Scheme 2, the turnover-limiting step for this transformation is the insertion of **2a** with an overall barrier of 25.5 kcal mol^{-1}. The initial C–H cobaltation occurs with a smaller activation free energy of 15.5 kcal mol^{-1}. These values are also in good qualitative agreement with the experimental findings: The calculated high barriers match the prolonged reaction times and high temperature required in the experimental studies and the reversible C–H metalation [30].

Influence of London dispersion

In recent years, London dispersion, the attractive part of the van-der-Waals force, has been repeatedly identified as key to stabilizing organic structures and facilitating novel reactivities [3]. As the Cp* ligand is a C–H-rich molecule, we envisioned that dispersive interactions should be important for this transformation as well. As a consequence, we have analyzed this reaction additionally with B3LYP without dispersion correction and the dispersion-corrected M06-L functional under otherwise identical conditions as a first starting point. Independent of the

Table 1: Calculated free energies for the reaction mechanism involving differently substituted 2-phenylpyridines **1a–e** [B3LYP-D3BJ/def2-QZVP/COSMO//B3LYP-D3BJ/def2-SVP, def2-TZVP for Co].



1a–e **2a** **3a–e**

	4	5	TS1	6	TS2	7	8	9	TS3	10	11	12	3
R = H (**1a**)	0.0	−20.1	−12.4	−14.1	−5.4	−8.4	−4.2	−15.5	+3.4	−18.3	−16.0	−34.5	−18.0
R = CH_3 (**1b**)	0.0	−21.0	−14.2	−15.7	−7.1	−9.6	−4.5	−18.5	+0.6	−18.8	−16.1	−35.5	−18.9
R = F (**1c**)	0.0	−19.8	−12.0	−12.4	−4.8	−9.7	−4.4	−17.5	+4.5	−17.3	−14.9	−32.3	−16.8
R = $C(O)CH_3$ (**1d**)	0.0	−18.1	−9.6	−11.2	−3.4	−7.4	−2.1	−17.7	+3.4	−17.5	−13.2	−30.9	−16.4
R = CN (**1e**)	0.0	−18.0	−9.9	−9.9	−3.1	−8.9	−3.5	−16.8	+5.4	−14.9	−12.4	−29.4	−15.5

computational method, the overall reaction free energy for the transformation of Scheme 2 is almost identical [−15.9 (B3LYP-D3BJ), −17.0 (B3LYP), and −15.9 (M06-L) kcal mol^{-1}] indicating that dispersion is less important for the overall thermodynamics of this reaction. In contrast, a strong effect of the functional was observed for the complete energy profile. While the dispersion-corrected functional M06-L (see the Supporting Information File 1 for details) resulted in a comparable profile to that obtained with B3LYP-D3BJ (black lines in Figure 1), a significant deviation was observed when the latter was used without any dispersion correction (red lines in Figure 1). All cobalt complexes are substantially stabilized by dispersive interactions resulting in a significant net reduction of the activation free energy by 11 kcal mol^{-1}. Comparable contributions of London dispersion have also been calculated with other functionals (TPSS [45] and PBE [46,47]). As expected, complexes with more nonbonding contacts (e.g., **10a**) are better stabilized than complexes where the Cp* ligand is located farther away from other ligands (e.g., **8a**). In comparison to computational investigations of Pd-catalyzed reactions [27], similar dispersive stabilizations of individual complexes have been calculated here.

A closer qualitative analysis of the intramolecular interactions in these complexes employing the NCIPLOT program [48,49] furthermore confirms these noncovalent interactions. While all plots are shown in Supporting Information File 1, Figure 3 summarizes those for selected intermediates and transition states. For all structures, significant interactions can be found between the Cp* ligand and the various phenyl groups of the reagents. In addition, the presence of additional stabilizing interactions such as further hydrogen bonds can also be confirmed by this analysis (e.g., in **TS3**, see also the Supporting Information File 1).

To further probe the dispersive interaction of the Cp* ligand and the other ligands, we have additionally calculated the dispersion interaction densities (DID) [50] for all intermediates and transition states at the SCS-LMP2/def2-TZVPP level of theory. The DID plots of Figure 4 reveal that medium to strong

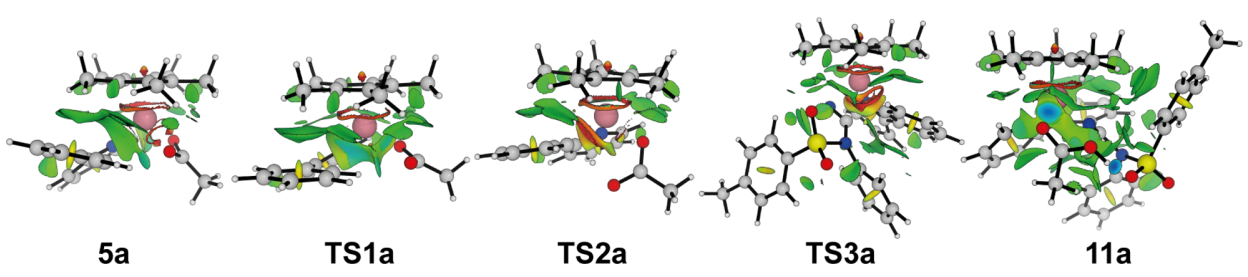


Figure 3: Noncovalent interaction (NCI) analysis for selected intermediates and transition states. The gradient isosurfaces (s = 0.5 au) are colored according to the sign of $(\lambda_2)\rho$ over the range of −0.05 (blue) to +0.05 (red).

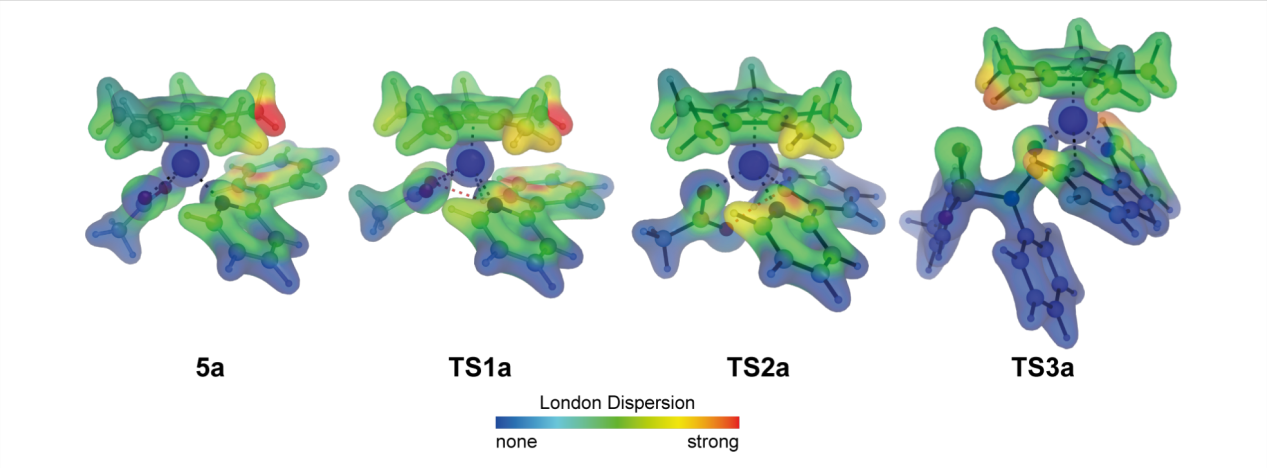


Figure 4: Projected dispersion interaction density (DID) plots for selected intermediates and transition states. The molecular density isosurfaces (0.1 *e*/Bohr^3) are colored from zero interaction energy (blue) to the strongest dispersion interaction (red).

dispersive interactions can be found between the Cp* ligand and the aromatic and C–H-rich fragments in its proximity. In line with the analyses presented in Figure 3 and Figure 4, a local energy decomposition (LED) analysis [51] using DLPNO-CCSD(T)/cc-pVDZ also confirmed medium to strong dispersive interactions up to 12 kcal mol^{-1} between the Cp* ligand and the other ligands. Based on the computational analysis, London dispersion is not only highly beneficial for the synthetically important cobalt-catalyzed C−H cyanation reaction, but it also emphasises that the Cp* ligand does not exclusively act as a sterically demanding ligand in transition-metal-catalyzed reactions.

Conclusion

We have analyzed the cobalt(III)-catalyzed C–H cyanation of differently substituted 2-phenylpyridines with *N*-cyano-*N*-aryl-*p*-toluenesulfonamide using density functional theory. On the basis of our computational and experimental data, we can propose a reaction mechanism for this transformation. After an initial and reversible C–H cobaltation, the subsequent insertion of the cyanating agents is the rate-limiting step. In addition, our calculations unravel that all the cobalt intermediates are considerably affected by London dispersion, which also results in a significant stabilization of the rate-limiting transition state.

Computational Details

For all structures, geometry optimizations were performed with three different functionals using the def2-TZVP (def2-TZVPP for M06-L) basis set [52] and the m4 numerical quadrature grid in the gas phase. The hybrid functional B3LYP [53,54] with and without Grimme's dispersion correction D3 (Becke–Johnson damping) [35,36] as well as Truhlar's dispersion-corrected M06-L [55] functional were employed in this investigation. For the latter, the density fitting RI-J approach was used to accelerate the calculations [56,57]. For the analysis of the substituent effect, the B3LYP functional with Grimme's dispersion correction D3 (Becke–Johnson damping) was employed together with the def2-SVP basis set for all non-metals and the def2-TZVP basis set for Co. Vibrational analysis verified that each structure was a minimum or transition state ($i\omega$ < 30 cm^{-1} were tolerated). Thermal corrections were calculated from unscaled harmonic vibrational frequencies at the same levels of theory and refer to a standard state of 298.15 K and 1 mol L^{-1}. Entropic contributions to the reported free energies were obtained from partition functions evaluated with Truhlar's quasi-harmonic approximation [58]. This method uses the same approximations as the usual harmonic oscillator approximation except that all vibrational frequencies lower than 100 cm^{-1} are set equal to 100 cm^{-1}. Energies were subsequently derived from single-point calculations employing the functionals described above, the quadruple-ζ basis set def2-QZVP [52] and the COSMO solvation model [59] for dichloroethane (ε = 10.125). The dispersion interaction densities (DID) [50] were calculated at the SCS-LMP2/def2-TZVPP level of theory using MOLPRO 2015 [60,61]. The local energy decomposition analysis [51] was performed employing Neese's domain-based local pair-natural orbital (DLPNO) approach to the CCSD(T) method [DLPNO-CCSD(T)] [62-64] with tightPNO settings and the double-ζ cc-pVDZ basis set as implemented in ORCA 4 [65]. All DFT calculations were performed with Turbomole 7.1 [66,67] and the NCIPLOT code was employed for the visualization noncovalent interactions [48,49].

Experimental Details

General remarks: Catalytic reactions were carried out in Schlenk flasks under nitrogen atmosphere using predried glassware. 1,2-Dichlorethane (DCE) was dried and distilled over CaH_2 under N_2. *N*-Cyano-*N*-phenyl-*p*-toluenesulfonamide (**2a**)

[68] and Cp*Co(CO)I_2 [69] were synthesized according to previously described methods. Other chemicals were obtained from commercial sources and were used without further purification.

Kinetic experiments of the cobalt(III)-catalyzed C–H cyanation: A suspension of **1** (78 mg, 0.50 mmol), **2** (0.75 mmol), [Cp*Co(CO)I_2] (6.0 mg, 2.5 mol %), $AgSbF_6$ (8.6 mg, 5.0 mol %) and KOAc (2.5 mg, 5.0 mol %) in DCE (2.0 mL) was heated at 120 °C. Aliquots up to ca 15% conversion (25 µL; 10, 20, 30, 40, 50 min) were periodically removed by a syringe and directly analyzed by GC using *n*-dodecane (30 µL) as internal standard.

2-(Pyridin-2-yl)benzonitrile (3a): ^{1}H NMR ($CDCl_3$, 400 MHz) δ 8.73–8.70 (ddd, *J* = 4.7, 1.8, 0.9 Hz, 1H), 7.82–7.71 (m, 4H), 7.64 (dd, *J* = 7.6, 1.4 Hz, 1H), 7.49 (dd, *J* = 7.6, 1.2 Hz, 1H), 7.31 (ddd, *J* = 7.4, 4.7, 1.2 Hz, 1H); ^{13}C NMR ($CDCl_3$, 125 MHz) δ 155.1 (C_q), 149.8 (CH), 143.4 (C_q), 136.7 (CH), 134.0 (CH), 132.7 (CH), 129.9 (CH), 128.6 (CH), 123.2 (CH), 123.1 (CH), 118.6 (C_q), 111.0 (C_q); IR (ATR): 3350, 2224, 1560, 1464, 758, 509 cm^{-1}. EIMS *m*/*z* (relative intensity): 180 (100) [M^+], 154 (5), 140 (5), 126 (5), 102 (5), 75 (5); HRMS (EI) *m*/*z*: [M^+] calcd. for $C_{12}H_8N_2$, 180.0687; found, 180.0684. The analytical data are in accordance with those reported in literature [30].

Supporting Information

Supporting Information File 1

Cartesian coordinates, energies of all calculated structures, and details of computational methods.
[https://www.beilstein-journals.org/bjoc/content/supplementary/1860-5397-14-130-S1.pdf]

Acknowledgements

Financial support from the Fonds der chemischen Industrie (Liebig scholarship to M.B.), the DFG (SPP 1807), and from the University of Cologne within the excellence initiative is gratefully acknowledged. We are very grateful to Professor Ricardo Mata and Axel Wuttke for help with the calculation of dispersion interaction densities. We thank the Regional Computing Center of the University of Cologne (RRZK) for providing computing time on the DFG-funded High Performance Computing (HPC) system CHEOPS as well as for their support.

References

1. London, F. *Z. Phys.* **1930,** *63,* 245–279. doi:10.1007/BF01421741
2. London, F. *Trans. Faraday Soc.* **1937,** *33,* 8b–26. doi:10.1039/tf937330008b
3. Wagner, J. P.; Schreiner, P. R. *Angew. Chem., Int. Ed.* **2015,** *54,* 12274–12296. doi:10.1002/anie.201503476
4. Grimme, S.; Schreiner, P. R. *Angew. Chem., Int. Ed.* **2011,** *50,* 12639–12642. doi:10.1002/anie.201103615
5. 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. *Nature* **2011,** *477,* 308–311. doi:10.1038/nature10367
6. Fokin, A. A.; Chernish, L. V.; Gunchenko, P. A.; Tikhonchuk, E. Yu.; Hausmann, H.; Serafin, M.; Dahl, J. E. P.; Carlson, R. M. K.; Schreiner, P. R. *J. Am. Chem. Soc.* **2012,** *134,* 13641–13650. doi:10.1021/ja302258q
7. Rösel, S.; Quanz, H.; Logemann, C.; Becker, J.; Mossou, E.; Cañadillas-Delgado, L.; Caldeweyher, E.; Grimme, S.; Schreiner, P. R. *J. Am. Chem. Soc.* **2017,** *139,* 7428–7431. doi:10.1021/jacs.7b01879
8. Aikawa, H.; Takahira, Y.; Yamaguchi, M. *Chem. Commun.* **2011,** *47,* 1479–1481. doi:10.1039/C0CC03025B
9. Hobza, P.; Müller-Dethlefs, K. *Non-covalent Interactions;* The Royal Society of Chemistry: Cambridge, 2009. doi:10.1039/9781847559906
10. Wheeler, S. E.; Seguin, T. J.; Guan, Y.; Doney, A. C. *Acc. Chem. Res.* **2016,** *49,* 1061–1069. doi:10.1021/acs.accounts.6b00096
11. Walden, D. M.; Ogba, O. M.; Johnston, R. C.; Cheong, P. H.-Y. *Acc. Chem. Res.* **2016,** *49,* 1279–1291. doi:10.1021/acs.accounts.6b00204
12. Daugulis, O.; Roane, J.; Tran, L. D. *Acc. Chem. Res.* **2015,** *48,* 1053–1064. doi:10.1021/ar5004626
13. Wencel-Delord, J.; Glorius, F. *Nat. Chem.* **2013,** *5,* 369–375. doi:10.1038/nchem.1607
14. Yeung, C. S.; Dong, V. M. *Chem. Rev.* **2011,** *111,* 1215–1292. doi:10.1021/cr100280d
15. Ackermann, L. *Chem. Rev.* **2011,** *111,* 1315–1345. doi:10.1021/cr100412j
16. Chen, X.; Engle, K. M.; Wang, D.-H.; Yu, J.-Q. *Angew. Chem., Int. Ed.* **2009,** *48,* 5094–5115. doi:10.1002/anie.200806273
17. Bergman, R. G. *Nature* **2007,** *446,* 391–393. doi:10.1038/446391a
18. Ghorai, D.; Müller, V.; Keil, H.; Stalke, D.; Zanoni, G.; Tkachenko, B. A.; Schreiner, P.; Ackermann, L. *Adv. Synth. Catal.* **2017,** *359,* 3137–3141. doi:10.1002/adsc.201700663
19. Graux, L. V.; Giorgi, M.; Buono, G.; Clavier, H. *Dalton Trans.* **2016,** *45,* 6491–6502. doi:10.1039/C5DT04683A
20. Zell, D.; Warratz, S.; Gelman, D.; Garden, S. J.; Ackermann, L. *Chem. – Eur. J.* **2016,** *22,* 1248–1252. doi:10.1002/chem.201504851
21. Ackermann, L.; Barfüsser, S.; Kornhaass, C.; Kapdi, A. R. *Org. Lett.* **2011,** *13,* 3082–3085. doi:10.1021/ol200986x
22. Ackermann, L.; Althammer, A.; Born, R. *Angew. Chem., Int. Ed.* **2006,** *45,* 2619–2622. doi:10.1002/anie.200504450
23. Ackermann, L. *Org. Lett.* **2005,** *7,* 3123–3125. doi:10.1021/ol051216e
24. Sperger, T.; Sanhueza, I. A.; Schoenebeck, F. *Acc. Chem. Res.* **2016,** *49,* 1311–1319. doi:10.1021/acs.accounts.6b00068
25. Ahlquist, M. S. G.; Norrby, P.-O. *Angew. Chem., Int. Ed.* **2011,** *50,* 11794–11797. doi:10.1002/anie.201105928
26. Hansen, A.; Bannwarth, C.; Grimme, S.; Petrović, P.; Werlé, C.; Djukic, J.-P. *ChemistryOpen* **2014,** *3,* 177–189. doi:10.1002/open.201402017
27. Lyngvi, E.; Sanhueza, I. A.; Schoenebeck, F. *Organometallics* **2015,** *34,* 805–812. doi:10.1021/om501199t
28. Ehrlich, S.; Moellmann, J.; Grimme, S. *Acc. Chem. Res.* **2013,** *46,* 916–926. doi:10.1021/ar3000844
29. Grimme, S. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2011,** *1,* 211–228. doi:10.1002/wcms.30

30. Li, J.; Ackermann, L. *Angew. Chem., Int. Ed.* **2015,** *54,* 3635–3638. doi:10.1002/anie.201409247
31. Yu, D.-G.; Gensch, T.; de Azambuja, F.; Vásquez-Céspedes, S.; Glorius, F. *J. Am. Chem. Soc.* **2014,** *136,* 17722–17725. doi:10.1021/ja511011m
32. Pawar, A. B.; Chang, S. *Org. Lett.* **2015,** *17,* 660–663. doi:10.1021/ol503680d
33. Liu, W.; Ackermann, L. *Chem. Commun.* **2014,** *50,* 1878–1881. doi:10.1039/c3cc49502g
34. Cui, J.; Song, J.; Liu, Q.; Liu, H.; Dong, Y. *Chem. – Asian J.* **2018,** *13,* 482–495. doi:10.1002/asia.201701611
35. Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. *J. Chem. Phys.* **2010,** *132,* 154104. doi:10.1063/1.3382344
36. Grimme, S.; Ehrlich, S.; Goerigk, L. *J. Comput. Chem.* **2011,** *32,* 1456–1465. doi:10.1002/jcc.21759
37. Krossing, I.; Raabe, I. *Angew. Chem., Int. Ed.* **2004,** *43,* 2066–2090. doi:10.1002/anie.200300620
38. Leitch, J. A.; Wilson, P. B.; McMullin, C. L.; Mahon, M. F.; Bhonoah, Y.; Williams, I. H.; Frost, C. G. *ACS Catal.* **2016,** *6,* 5520–5529. doi:10.1021/acscatal.6b01370
39. Kumar, N. Y. P.; Rogge, T.; Yetra, S. R.; Bechtoldt, A.; Clot, E.; Ackermann, L. *Chem. – Eur. J.* **2017,** *23,* 17449–17453. doi:10.1002/chem.201703680
40. Bu, Q.; Rogge, T.; Kotek, V.; Ackermann, L. *Angew. Chem., Int. Ed.* **2018,** *57,* 765–768. doi:10.1002/anie.201711108
41. Zell, D.; Dhawa, U.; Müller, V.; Bursch, M.; Grimme, S.; Ackermann, L. *ACS Catal.* **2017,** *7,* 4209–4213. doi:10.1021/acscatal.7b01208
42. Yang, Y.; Buchwald, S. L. *Angew. Chem., Int. Ed.* **2014,** *53,* 8677–8681. doi:10.1002/anie.201402449
43. Yang, Y.; Liu, P. *ACS Catal.* **2015,** *5,* 2944–2951. doi:10.1021/acscatal.5b00443
44. Orientating calculations using **2b** and **2c** as the cyanating agent parallel the trend observed in the kinetic data ($\Delta G^{‡}$(**2b**) < $\Delta G^{‡}$(**2b**) < $\Delta G^{‡}$(**2c**)).
45. Tao, J.; Perdew, J. P.; Staroverov, V. N.; Scuseria, G. E. *Phys. Rev. Lett.* **2003,** *91,* 146401. doi:10.1103/PhysRevLett.91.146401
46. Perdew, J. P.; Burke, K.; Ernzerhof, M. *Phys. Rev. Lett.* **1996,** *77,* 3865–3868. doi:10.1103/PhysRevLett.77.3865
47. Perdew, J. P.; Burke, K.; Ernzerhof, M. *Phys. Rev. Lett.* **1997,** *78,* 1396. doi:10.1103/PhysRevLett.78.1396
48. Johnson, E. R.; Keinan, S.; Mori-Sánchez, P.; Contreras-García, J.; Cohen, A. J.; Yang, W. *J. Am. Chem. Soc.* **2010,** *132,* 6498–6506. doi:10.1021/ja100936w
49. Contreras-García, J.; Johnson, E. R.; Keinan, S.; Chaudret, R.; Piquemal, J.-P.; Beratan, D. N.; Yang, W. *J. Chem. Theory Comput.* **2011,** *7,* 625–632. doi:10.1021/ct100641a
50. Wuttke, A.; Mata, R. A. *J. Comput. Chem.* **2017,** *38,* 15–23. doi:10.1002/jcc.24508
51. Schneider, W. B.; Bistoni, G.; Sparta, M.; Saitow, M.; Riplinger, C.; Auer, A. A.; Neese, F. *J. Chem. Theory Comput.* **2016,** *12,* 4778–4792. doi:10.1021/acs.jctc.6b00523
52. Weigend, F.; Ahlrichs, R. *Phys. Chem. Chem. Phys.* **2005,** *7,* 3297–3305. doi:10.1039/b508541a
53. Becke, A. D. *J. Chem. Phys.* **1993,** *98,* 5648–5652. doi:10.1063/1.464913
54. Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988,** *37,* 785–789. doi:10.1103/PhysRevB.37.785
55. Zhao, Y.; Truhlar, D. G. *J. Chem. Phys.* **2006,** *125,* 194101. doi:10.1063/1.2370993
56. Eichkorn, K.; Weigend, F.; Treutler, O.; Ahlrichs, R. *Theor. Chem. Acc.* **1997,** *97,* 119–124. doi:10.1007/s002140050244
57. Deglmann, P.; May, K.; Furche, F.; Ahlrichs, R. *Chem. Phys. Lett.* **2004,** *384,* 103–107. doi:10.1016/j.cplett.2003.11.080
58. Ribeiro, R. F.; Marenich, A. V.; Cramer, C. J.; Truhlar, D. G. *J. Phys. Chem. B* **2011,** *115,* 14556–14562. doi:10.1021/jp205508z
59. Klamt, A.; Schüürmann, G. *J. Chem. Soc., Perkin Trans. 2* **1993,** 799–805. doi:10.1039/P29930000799
60. Werner, H.-J.; Knowles, P. J.; Knizia, G.; Manby, F. R.; Schütz, M. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2012,** *2,* 242–253. doi:10.1002/wcms.82
61. *MOLPRO,* version 2015.1; a package of ab initio programs, http://www.molpro.net.
62. Riplinger, C.; Sandhoefer, B.; Hansen, A.; Neese, F. *J. Chem. Phys.* **2013,** *139,* 134101. doi:10.1063/1.4821834
63. Riplinger, C.; Neese, F. *J. Chem. Phys.* **2013,** *138,* 034106. doi:10.1063/1.4773581
64. Liakos, D. G.; Neese, F. *J. Chem. Theory Comput.* **2015,** *11,* 4054–4063. doi:10.1021/acs.jctc.5b00359
65. Neese, F. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2012,** *2,* 73–78. doi:10.1002/wcms.81
66. *TURBOMOLE,* V7.1, 2015; TURBOMOLE GmbH, since 2007, http://www.turbomole.com.
a development of University of Karlsruhe and Forschungszentrum Karlsruhe GmbH, 1989–2007.
67. Furche, F.; Ahlrichs, R.; Hättig, C.; Klopper, W.; Sierka, M.; Weigend, F. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2014,** *4,* 91–100. doi:10.1002/wcms.1162
68. Liu, W.; Richter, S. C.; Mei, R.; Feldt, M.; Ackermann, L. *Chem. – Eur. J.* **2016,** *22,* 17958–17961. doi:10.1002/chem.201604621
69. Sun, B.; Yoshino, T.; Matsunaga, S.; Kanai, M. *Adv. Synth. Catal.* **2014,** *356,* 1491–1495. doi:10.1002/adsc.201301110

License and Terms

BEILSTEIN JOURNAL OF ORGANIC CHEMISTRY

The phenyl vinyl ether–methanol complex: a model system for quantum chemistry benchmarking

Dominic Bernhard[1], Fabian Dietrich[1], Mariyam Fatima[2,3], Cristóbal Pérez[2,3], Hannes C. Gottschalk[4], Axel Wuttke[4], Ricardo A. Mata[*4], Martin A. Suhm[*4], Melanie Schnell[*2,3,5] and Markus Gerhards[*1]

Full Research Paper

Open Access

Address:
[1]Fachbereich Chemie & Research Center Optimas, Technische Universität Kaiserslautern, Erwin-Schrödinger-Str. 52, D-67663 Kaiserslautern, Germany, [2]Max Planck Institute for the Structure and Dynamics of Matter, Luruper Chaussee 149, D-22761 Hamburg, Germany, [3]Deutsches Elektronen Synchrotron (DESY), Notkestrasse 85, D-22607 Hamburg, Germany, [4]Institut für Physikalische Chemie, Georg-August-Universität Göttingen, Tammannstrasse 6, D-37077 Göttingen, Germany and [5]Institute of Physical Chemistry, Christian-Albrechts-Universität zu Kiel, Max-Eyth-Strasse 1, D-24118 Kiel, Germany

Email:
Ricardo A. Mata[*] - rmata@gwdg.de;
Martin A. Suhm[*] - msuhm@gwdg.de;
Melanie Schnell[*] - melanie.schnell@desy.de;
Markus Gerhards[*] - gerhards@chemie.uni-kl.de

* Corresponding author

Keywords:
dispersion interactions; IR spectroscopy; quantum-chemical calculations; rotational spectroscopy; structure determination; weak hydrogen bonds

Beilstein J. Org. Chem. **2018,** *14,* 1642–1654.
doi:10.3762/bjoc.14.140

Received: 16 March 2018
Accepted: 02 June 2018
Published: 02 July 2018

This article is part of the Thematic Series "Dispersion interactions".

Guest Editor: P. Schreiner

Abstract

The structure of the isolated aggregate of phenyl vinyl ether and methanol is studied by combining a multi-spectroscopic approach and quantum-chemical calculations in order to investigate the delicate interplay of noncovalent interactions. The complementary results of vibrational and rotational spectroscopy applied in molecular beam experiments reveal the preference of a hydrogen bond of the methanol towards the ether oxygen (OH···O) over the π-docking motifs via the phenyl and vinyl moieties, with an additional less populated OH···P(phenyl)-bound isomer detected only by microwave spectroscopy. The correct prediction of the energetic order of the isomers using quantum-chemical calculations turns out to be challenging and succeeds with a sophisticated local coupled cluster method. The latter also yields a quantification as well as a visualization of London dispersion, which prove to be valuable tools for understanding the role of dispersion on the docking preferences. Beyond the structural analysis of the electronic ground state (S_0), the electronically excited (S_1) state is analyzed, in which a destabilization of the OH···O structure compared to the S_0 state is observed experimentally and theoretically.

Introduction

The balance of different noncovalent interactions is crucial for chemical and biochemical processes as it controls molecular recognition and aggregation [1-6]. In order to gain a deeper understanding of these processes, knowledge on exact structural arrangements and the respective role of different intermolecular forces such as electrostatic, dispersion and induction forces is needed. Thus, experimental examination as well as the precise prediction of a preferred molecular docking site for different molecules is of crucial importance. Despite the remarkable progress made in experiments and theory/computational chemistry, there is still a need for improvement and benchmarking [7].

Many aromatic solute–solvent complexes have been studied in the gas phase (cf. [8-10] and references therein). Studied systems involving methanol as attached solvent molecule include the works on benzene–methanol clusters by the Zwier group [11] and on fluorobenzene–methanol clusters by the Brutschy group [12], to mention only two examples. Complexes of aromatic ethers with polar solvent molecules are of special interest due to the presence of different competing hydrogen bond acceptor sites. An extensive study on diphenoxyethane–water clusters was performed by the Zwier group [13-15] including studies in the excited S_1 and S_2 states. Concerning aggregates of aromatic ethers with alcohols, there is a work of Pietraperzia et al. [16] on the anisole–phenol complex in which an OH···O structure was identified. In a systematic study by the Suhm group on complexes of anisole derivatives with methanol, a balance between OH···O and OH···π structures being very sensitive to the substitution pattern at the anisole moiety was identified [17,18]. In previous multi-spectroscopic studies by the Schnell, Suhm and Gerhards groups on diphenyl ether (DPE)–solvent complexes [19-22], the influence of different attached solvent molecules on the structural preference was compared. It could be shown that the balance between OH···π- and OH···O-bound structures is very sensitive to the size of the attached alcohol. Torsional balances in solution have been used to probe aromatic OH···π interactions and to show that these interactions remain important at room temperature [23].

In such aromatic solute–solvent systems, one frequently encounters hydrogen bonds formed towards oxygen or nitrogen lone pairs, or R–H···π binding motifs (R = O, N, C, S,…) involving aromatic π systems. Less often, R–H···π bound complexes are found involving nonconjugated, localized C=C double bonds. Exceptions include the ethene–methanol complex [24] as well as bulky olefin–*tert*-butyl alcohol complexes [25] investigated by jet FTIR spectroscopy. The observed OH stretching red-shifts compared to the free alcohols are small, indicating a comparatively weak hydrogen bond, which is also reflected in calculated binding energies [24,25].

With the herein presented work, we now extend our overall multi-spectroscopic study to mixed aromatic olefinic ethers: in the case of phenyl vinyl ether (PVE), there is an ethenyl moiety replacing one of the phenyl rings compared to DPE. This introduces a localized π system along with the delocalized phenyl π system as hydrogen bond acceptor sites. Thereby, the complexity of the system is increased, as now three qualitatively different basic binding motifs have to be regarded instead of only two for DPE. This also provides an enhanced challenge for theory, with no clear preference for one of the motifs to be expected. As shown, e.g., in the case of DPE–*t*-BuOH [20], there is a need for benchmarking systems in order to improve and develop better theoretical approaches especially for non-covalently bound complexes. The study on PVE–MeOH is meant to present a further benchmark system, probably even more challenging than DPE–*t*-BuOH.

For an experimental elucidation of structural arrangements and energetic preferences, investigations on a molecular level are required on isolated molecular aggregates, allowing for an ideal comparison with gas phase calculations. This can be achieved by molecular beam experiments, which can be combined with a variety of spectroscopic techniques. For our multi-spectroscopic studies, we utilize FTIR spectroscopy, mass- and isomer-selective IR/UV techniques (IR/R2PI, for methodical developments, cf., e.g., [8,26-29] and UV/IR/UV spectroscopy, cf., e.g., [30-43]) and chirped-pulse Fourier transform microwave (CP-FTMW) spectroscopy. Comparing spectroscopic results with quantum-chemical calculations is often mandatory for the interpretation of experiments. Furthermore, such comparison enables a critical evaluation of the approximations used, comparing the relative stability of different binding motifs.

In this paper, the first structural investigation on the complex of phenyl vinyl ether with methanol is presented. An established multi-spectroscopic approach [19,20] is used, coupling FTIR, IR/UV and microwave spectroscopy with theoretical treatments including dispersion-corrected density functional theory (DFT-D3) [44,45], spin-component-scaled approximated coupled cluster-singles-doubles (SCS-CC2) [46] as well as explicitly correlated local coupled cluster theory (LCCSD(T0)-F12) [47] calculations, the latter allowing for a quantification and visualization of London dispersion interactions [48]. The aim of the presented study is the unambiguous experimental identification of the preferred binding site of a first methanol solvent molecule to the multivalent hydrogen bond scaffold of phenyl vinyl ether, followed by a classification of theoretical methods in

terms of success or failure to predict this preference. Visualization of possible reasons for the subtle preference is a valuable additional asset.

Experimental Setup

FTIR setup

For the FTIR experiments, the so-called filet-jet setup, as described in detail in [49], was used. In this setup, the scans of a Bruker IFS 66 v/s spectrometer (80 kHz, resolution 2 cm^{-1}) are synchronized to a pulsed supersonic expansion through a 600 × 0.2 mm^2 slit nozzle. Using two separate cooled saturators, low concentrations (<0.1%) of PVE (Sigma-Aldrich, 97%, used as purchased) and methanol (Sigma-Aldrich, ≥99.8%, used as purchased) were added to the carrier gas helium (Linde, 99.996%) and premixed at a pressure of 0.75 bar in a 67 L reservoir before being expanded through the slit nozzle. The pulsed operation with waiting times of 30–90 s between 150 ms long pulses combined with a buffer volume of 12–23 m^3 and a pumping capacity of 500–2500 m^3/h resulted in background pressures of less than 0.1 mbar before expansions. This facilitated measurements of clusters of methanol and PVE in the zone of silence of the expansion at an average distance of 10 mm to the nozzle. A calcium fluoride beam splitter, lenses and windows were used in combination with a 150 W tungsten filament and an optical filter (4200–2450 cm^{-1}) to maximize the signal-to-noise ratio in the OH stretching range of the vibrational spectra. For the final spectra, 150 to 775 pulses were co-added to further improve signal-to-noise.

IR/UV setup

The experimental setup for the IR/UV experiments is described in detail elsewhere [29,50], thus only a brief description is given here. All experiments were carried out in a molecular beam apparatus consisting of a differentially pumped linear time-of-flight (TOF) mass spectrometer with a pulsed valve (Series 9 and pulse driver Iota One, General Valve, 500 µm orifice) for skimmed jet expansion. PVE was synthesized according to the procedure reported in [51] (cf. Supporting Information File 1 for details). MeOH (Sigma-Aldrich, ≥99.7%) and PVE were both supplied via separate cooled reservoirs (approx. −8 °C and −13 °C, respectively) and co-expanded with the carrier gas neon (2.5–3.0 bar).

For the one- and two-color R2PI, the IR/R2PI and the UV/IR/UV experiments up to three tunable nanosecond laser systems were necessary, including two independent UV laser systems and one IR laser system. The UV laser radiation is obtained via second harmonic generation in a BBO crystal using the output of a dye laser (Cobra-Stretch and PrecisionScan, Sirah). They are pumped by the second harmonic (532 nm) of a Nd:YAG laser (SpitLight 600 and SpitLight 1000, Innolas). The IR laser radiation in the range of 3520–3750 cm^{-1} is generated by difference frequency mixing (DFM) in a $LiNbO_3$ crystal using the fundamental (1064 nm) of a seeded Nd:YAG laser (Quanta-Ray Pro-230, Spectra-Physics) and the output of a further dye laser (PrecisionScan, Sirah), which is pumped by the second harmonic (532 nm) of the same Nd:YAG laser. Amplification of the resulting IR radiation is obtained by an optical parametric amplification (OPA) process in a further $LiNbO_3$ crystal using the output of the DFM process and the fundamental (1064 nm) of the Nd:YAG laser.

For the IR/R2PI spectra, the IR laser was fired 50 ns prior to the UV excitation laser, whereas for the UV/IR/UV spectra the IR laser was fired 2.0–3.0 ns after the UV excitation laser. The time delay between UV excitation and ionizing laser was 4.0–4.5 ns.

CP-FTMW setup

The rotational spectroscopy measurements were performed with the Hamburg chirped-pulse Fourier transform microwave (CP-FTMW) spectrometer COMPACT covering the 2–8 GHz frequency range, which has been described in detail in [52]. The molecules were seeded into a supersonic expansion with neon as the carrier gas by using a pulse nozzle (Parker General Valve, Series 9, 0.9 mm diameter orifice) equipped with a heatable reservoir close to the valve orifice, operating at 8 Hz. PVE was synthesized as described above and used without further purification.

The liquid sample was held in the reservoir at room temperature, which resulted in sufficient vapor pressure (standard boiling point of about 155 °C) for recording the rotational spectrum. MeOH was kept in a separate reservoir. PVE–MeOH clusters were generated by first flowing the carrier gas (neon) through the reservoir containing methanol that was external to the chamber, followed by picking up PVE vapor. After supersonic expansion into vacuum using neon at 3 bar, the molecular jet was polarized with a 4 µs chirp spanning 2–8 GHz. The chirp was generated with an arbitrary waveform generator, amplified to 300 W with a traveling wave tube amplifier, and transmitted into the vacuum chamber via a horn antenna. Following excitation, 40 µs of the free induction decay (FID) of the macroscopic ensemble of polarized molecules was recorded. The fast frame capability [53] of the Tektronix DPO 71254C was used in which eight consecutive excitation chirps, each followed by 40 µs during which the FID could be collected, were recorded and averaged. This resulted in an effective repetition rate of 64 Hz.

For the spectrum of the PVE–MeOH dimer, 3 million FIDs were co-added. A resulting signal-to-noise ratio of about 500:1

to 600:1 for the stronger transitions of the dominant complex allowed us to determine the positions of the carbon atoms with respect to the center of mass of the overall complex (see below) exploiting the presence of ^{13}C isotopologues in natural abundance and using the Kraitchman approach [54]. Fourier transformation of the averaged time domain FID, recorded at point spacings of 10 ps, resulted in a frequency domain rotational spectrum with frequency resolution of 25 kHz.

The assignment was performed with the program JB95 [55], then the fits to an asymmetric-rotor Hamiltonian were performed using SPFIT/SPCAT. The experimental results were complemented by and compared with the results of electronic structure calculations. B3LYP-D3/aug-cc-pVTZ calculations were performed using the Gaussian 09, rev. D.01 program suite [56] to guide the assignment.

Computational Methods

Various input structures for the PVE–MeOH complex were generated by using the MMFF94s force field [57] as implemented in Avogadro [58]. Afterwards, geometry optimizations were performed by applying the Berny optimization algorithm of Gaussian 09 [56] with energies and gradients obtained from Turbomole 7.0 [59]. The DFT functional B3LYP with Grimme's two-body D3 corrections and Becke–Johnson damping [45] was used in combination with the def2-TZVP basis set based on the documented performance of this level of theory for the similar diphenyl ether–methanol system [19]. Furthermore, the obtained structures were re-optimized with the SCS-CC2 method using the def2-TZVP basis set, both in the electronic ground (S_0) and first excited state (S_1). The ricc2 module in Turbomole 7.0 requires an auxiliary Coulomb fitting basis set (cbas) for the resolution-of-identity approximation (RI) for which def2-TZVP-cbas was chosen [60]. All obtained geometries were confirmed as minima by harmonic frequency calculations.

In order to evaluate the relative stability of the different conformers found on the potential hypersurface, density fitted explicitly correlated local coupled cluster with singles and doubles excitations and perturbative triples (DF-LCCSD(T0)-F12) calculations were carried out [47]. In order to converge the energies relative to the one particle basis, the VTZ-F12 and VQZ-F12 basis sets [61,62] were used together with a Schwenke style basis set extrapolation, as proposed in [63]. The orbitals were Pipek–Mezey [64] localized and orbital domains determined by natural population analysis with a threshold of TNPA = 0.03 [65]. Defaults were used for the pair classification, with all pairs included in the F12 treatment. Furthermore, the intermolecular pairs were classified as strong (meaning that they were treated at the highest level of theory). The latter method will be denoted as LCCSD(T0)-F12/CBS[T:Q]. In all correlated calculations the 1s electrons were removed from the treatment (frozen-core approximation). Furthermore, we analyzed the relative impact of dispersion interactions in the different complexes through a local orbital analysis of the CCSD (connected) doubles energy terms. The latter discussion is complemented with dispersion interaction density (DID) plots [48]. The coupled cluster calculations were carried out with Molpro 2015.1 [66].

Results and Discussion

Theoretical results

In contrast to the already studied diphenyl ether–alcohol clusters [19,20,22], phenyl vinyl ether offers three different binding sites for possible interactions with small solvent molecules: the ether oxygen, the phenyl ring and the vinyl moiety. Since both the phenyl ring and the vinyl moiety interact with the solvent via a π cloud, preferred binding sites are indicated using the following nomenclature: P (phenyl) and E (ethenyl), respectively. The optimizations using B3LYP-D3(BJ)/def2-TZVP yield six different structures, representing each binding motif with two isomers (cf. Figure 1).

In order to verify the structures of the complexes, a second level of theory was applied, namely SCS-CC2/def2-TZVP. Similar minima were found in the latter calculations, confirming the rich variety of binding motifs. However, distinct differences were found between the two methods: While B3LYP-D3 predicted the OH–E conformer as the most stable complex SCS-CC2 gave OH–P as the lowest minimum (cf. Table S1 in Supporting Information File 1). This is in contrast to our results of diphenyl ether–alcohol clusters [19,20,22], where both computational levels predicted the same energetic order of the isomers. The two structures correspond to quite different docking positions, reflecting well the demanding test this system imposes on quantum chemical methods. Several minima are separated by energy differences of 1 kJ/mol or less. The complete energetic analysis at both levels of theory is presented in Supporting Information File 1 (cf. Tables S1 and S2). The reasons behind the discrepancies are manifold, ranging from the method to the small basis set used. In order to obtain a more reliable theoretical prediction, LCCSD(T0)-F12/CBS[T:Q] calculations were carried out on top of the DFT-optimized geometries. The results are presented in Table 1, with and without zero-point vibrational energy (ZPVE) corrections.

The coupled cluster results show a clear energetic preference for the OH–O and OH–O' isomers. Observing the intermolecular contacts, which may or may not be designated as weak hydrogen bonds but are expected to stabilize the complexes, the main difference between the two structures is a phenyl vs

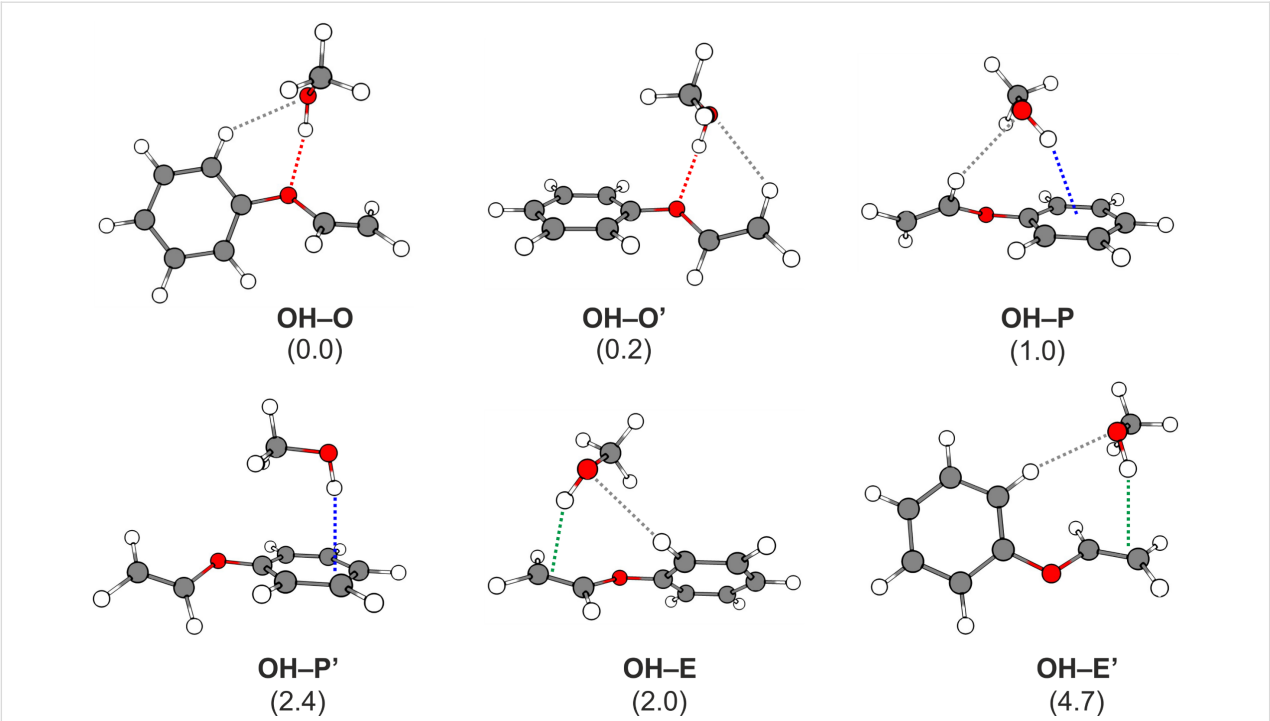


Figure 1: Minimum structures of the most stable PVE–MeOH dimers obtained at the B3LYP-D3(BJ)/def2-TZVP level; dashed colored lines indicate the different primary docking motifs, dashed gray lines illustrate secondary CH–O contacts; values in parentheses correspond to the relative, zero-point-corrected energies $E_{0,rel}$ with respect to the OH–O isomer, calculated at the LCCSD(T0)-F12/CBS[T:Q]//B3LYP-D3/def2-TZVP level of theory (cf. Table 1).

ethenyl CH to methanol O contact (cf. dashed gray lines in Figure 1). Both are separated by only a few tenths of a kJ/mol, which is within the error of the method used (considering that the coupled cluster expansion is truncated at triples excitations and the neglect of core-valence correlation effects, which should be the largest sources of error along with the harmonic B3LYP ZPVE error). It also confirmed the subtle difference between the six conformers, with an energy span of approximately 4–5 kJ/mol (≈1 kcal/mol, the commonly accepted definition of chemical accuracy) among all structures.

Also featured in Table 1 are the computed O–H stretch fundamentals together with the IR intensity at the two different levels of theory used in the optimizations. The frequencies were scaled according to the experimental value of the OH–π isomer of DPE–MeOH [19]. Based on the computational results, the vibrational spectral signals of the OH–O and OH–O’ isomers will be extremely hard to distinguish, as they lie less than 3 cm^{-1} apart, with very similar intensities. The same can be asserted for the less stable OH–E and OH–E’ structures. This is not surprising, given the similarities of the OH binding pattern for both sets of structures.

Table 1: Comparison of different structures for PVE–MeOH dimers in the S_0 state with LCCSD(T0)-F12/CBS[T:Q]//B3LYP-D3/def2-TZVP electronic energies E_{rel} and B3LYP zero-point corrected energies $E_{0,rel}$ relative to the minimum OH–O structure. The scaled wavenumbers $\tilde{\nu}$ of the OH-stretching vibration together with the respective IR intensity I are presented for two levels of theory: B3LYP-D3 (scaling factor: 0.9600) and SCS-CC2/def2-TZVP (scaling factor: 0.9635).

	E_{rel} [kJ/mol]	$E_{0,rel}$ [kJ/mol]	B3LYP-D3		SCS-CC2	
			$\tilde{\nu}$ [cm^{-1}]	I [km/mol]	$\tilde{\nu}$ [cm^{-1}]	I [km/mol]
OH–O	0.0	0.0	3597	219	3619	160
OH–O’	−0.3	0.2	3600	193	3621	144
OH–P	1.4	1.0	3619	112	3631	67
OH–P’	3.9	2.4	3631	127	3636	110
OH–E	1.5	2.0	3567	187	3607	121
OH–E’	4.8	4.7	3567	197	3606	128

In order to gain further insight into the energetic order of the different isomers, we conducted an analysis of the dispersion interactions present in the system by decomposing the CCSD energy terms obtained with the largest basis set (VQZ-F12). The latter procedure is based on the classification of the intermolecular excitation classes as detailed in [48,67]. The results are shown in Table 2. Beyond the total dispersion contributions, we also made use of the local analysis to separate the contribution of different molecular moieties in the PVE molecule (phenyl, ether oxygen and ethenyl). Shared orbitals are split up according to their NPA (natural population analysis) charges as described in [68].

The dispersion interaction energies show an interesting pattern. Although all structures are significantly stabilized by dispersion, with a maximum energy difference of 2.6 kJ/mol when summed all together, the relative weight of the different molecular fragments varies quite significantly. The moiety with the largest potential as dispersion energy donor (DED) is the phenyl ring. This results in the strongest stabilization for the two conformers whereby the methanol is closest to the ring (OH–P and OH–P'). The other conformers have much more spread out contributions. What is surprising is that even for the ethenyl binding complexes the contribution of the phenyl ring is sizeable. Geometrically, this seems unlikely, given that the methanol moiety is not oriented favorably relative to the ring. The effect can, however, be understood by inspecting the respective dispersion interaction densities (DIDs, cf. Figure 2), which allow for an even finer-grained analysis. There, one can observe that the major contributor is not the π-system of the phenyl ring,

Table 2: Comparison of different structures for PVE–MeOH dimers, with dispersion energies calculated at the LCCSD/VQZ-F12 level of theory (the parentheses contain the percentage of the fragment's dispersion relative to the total dispersion energy).

	ΔE_{disp}(total) [kJ/mol]	ΔE_{disp}(phenyl) [kJ/mol]	ΔE_{disp}(O) [kJ/mol]	ΔE_{disp}(ethenyl) [kJ/mol]
OH–O	−14.3	−6.2 (43.2)	−5.1 (35.8)	−3.0 (21.1)
OH–O'	−15.7	−6.1 (38.7)	−4.9 (31.5)	−4.7 (29.8)
OH–P	−16.9	−11.9 (70.4)	−1.5 (9.2)	−3.5 (20.5)
OH–P'	−16.0	−14.0 (87.6)	−1.3 (8.0)	−0.7 (4.4)
OH–E	−15.6	−6.9 (44.3)	−2.1 (13.3)	−6.6 (42.4)
OH–E'	−13.1	−4.5 (34.0)	−0.9 (6.7)	−7.8 (59.3)

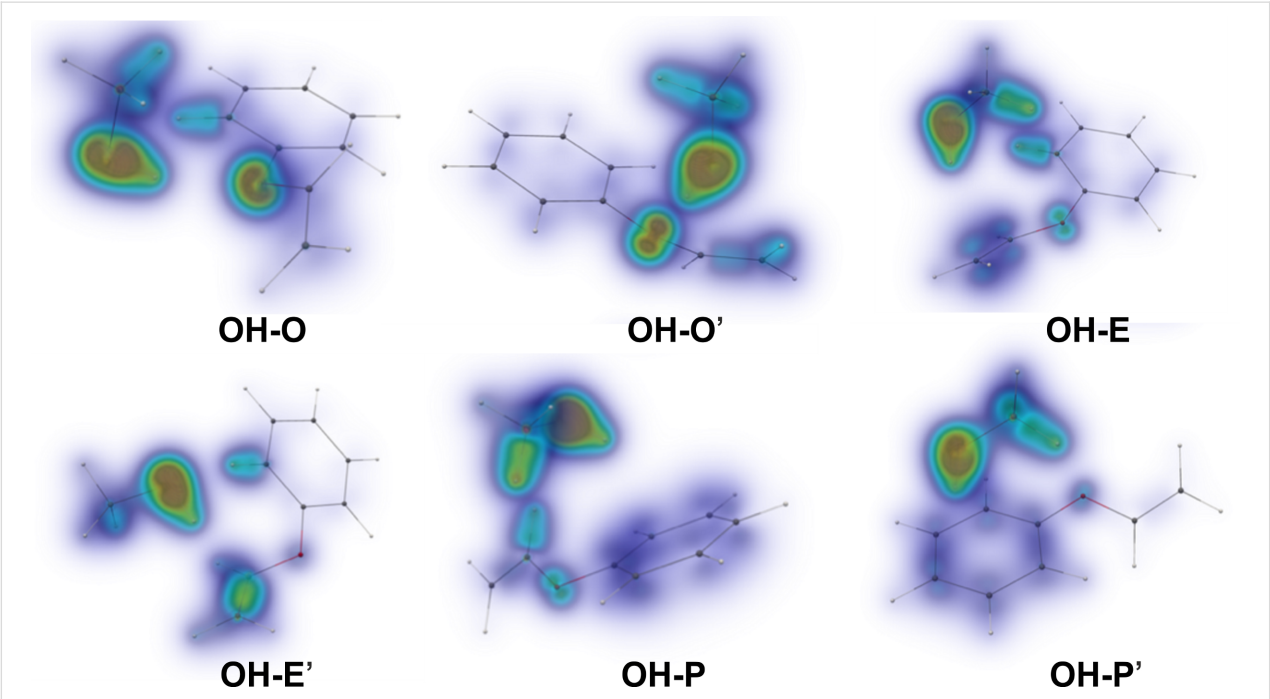


Figure 2: Dispersion interaction density (DID) plots calculated at the LCCSD/VQZ-F12 level. The brown zones indicate regions of electron density in a monomer which interact strongly by dispersion interactions with the other molecule. Blue stands for weaker/diffuse contributions. For example, in the top left figure one can observe that the OH group of methanol interacts strongly with the ether oxygen, with some dispersion energy coming as well from a CH orbital in the phenyl close to the methanol.

but a C–H contact to the methanol (a similar effect had already been observed in diphenyl ether–methanol complexes [19]). This contact is reminiscent of stabilization effects observed in coupled diamondoids [69] or supramolecular complexes [70], where such interactions can be found in large numbers.

Electronic ground state spectra

FTIR spectroscopy

The results of an FTIR exploration of the conformational diversity of this system are shown in Figure 3. Besides methanol monomer, methanol dimer and a signal clearly attributed to a larger cluster, only a single, reasonably narrow absorption at 3625 cm^{-1} is observed. It can be attributed to mixed dimers of MeOH with PVE and allows for a single rigorous conclusion, due to the linearity of the technique and the comparable IR absorption cross section of all predicted dimer conformations (cf. Table 1 and Table S1, Supporting Information File 1): the global minimum structure and any other, higher lying isomers which are initially formed and impeded from relaxation to the global minimum due to broad or high interconversion barriers must have their OH stretching fundamental at 3625 ± 5 cm^{-1} or be significantly less abundant.

If one were to trust the relative harmonic wavenumber predictions from the preceding subsection (cf. Table 1), this would imply a single docking motif, as different docking motifs are predicted to lead to larger spectral separations. However, different extents of anharmonicity do not allow to completely ruling out overlapping docking motifs. Therefore, conformationally selective methods are desirable to investigate this possibility. Finally, the actual docking site has to be identified by structural or electronic excitation spectroscopy.

IR/R2PI spectroscopy

Additional insight can be gained by using the mass- and isomer-selective IR/R2PI technique. This method requires knowledge on electronic excitation energies of the PVE–MeOH complex. For this reason, one-color R2PI spectra were recorded in the range of 36100–37600 cm^{-1} (cf. Figure S1 in Supporting Information File 1). While the R2PI spectrum of the PVE monomer shows well-resolved vibrational progressions (cf. Figure S1a, Supporting Information File 1), the spectrum of the solvent aggregate is broadened and affected by ionization-induced fragmentation of larger clusters (cf. Figure S1b, Supporting Information File 1). This is also reflected in the recorded IR/R2PI

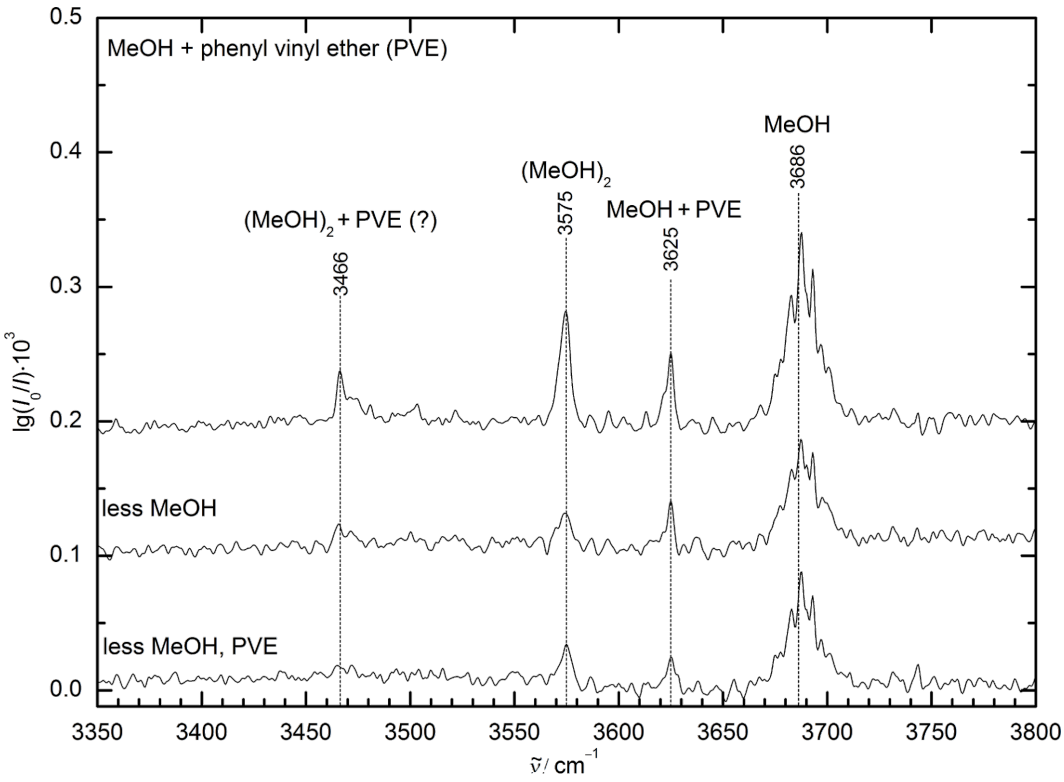


Figure 3: FTIR spectra of the supersonic expansion of methanol (MeOH) and phenyl vinyl ether (PVE) at different concentrations in helium. The spectra are spread out along the ordinate to improve visualization. Only one dominant mixed dimer band is visible in the spectra, lying at 3625 cm^{-1} (marked MeOH + PVE) between the methanol monomer at 3686 cm^{-1} (MeOH) and methanol dimer at 3575 cm^{-1} ($(MeOH)_2$). By comparing the spectrum at the top with the other two spectra recorded at reduced concentrations of methanol (middle) or methanol and PVE (bottom), the further down-shifted band at 3466 cm^{-1} can be attributed to a higher cluster, probably a methanol-rich mixed trimer ($(MeOH)_2$ + PVE (?)), due to its scaling with the variation of the concentrations.

spectra (cf. Figure S2, Supporting Information File 1), yielding solely the spectrum shown in Figure 4 via the excitation energy of 36885 cm^{-1} containing an OH stretching vibration of a PVE–MeOH dimer.

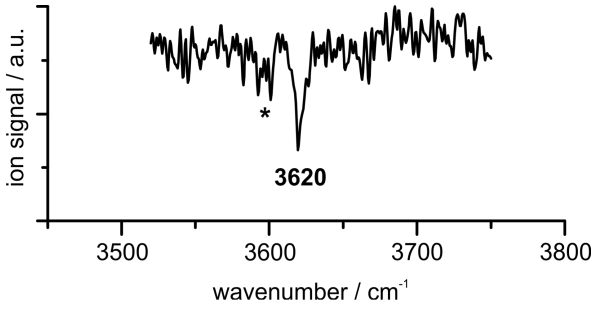


Figure 4: The IR/R2PI spectrum in the range of 3520–3750 cm^{-1} was obtained via the excitation energy of 36885 cm^{-1} using the carrier gas neon; the asterisk (*) indicates ionization-induced fragmentation from larger clusters (cf. Figure S2 in Supporting Information File 1).

Comparing the calculated OH stretching frequencies for the different isomers obtained at the DFT-D3 and SCS-CC2 levels (cf. Table 1) to the experiment does not allow for a clear structural assignment: the DFT-D3 calculations show the best agreement for the OH–P structure (3619 cm^{-1}, scaling factor 0.9600). Regarding the relative electronic energies, the OH–O structure is somewhat favored, with OH–O' coming as a close second. The latter frequencies are 3597 and 3600 cm^{-1}, respectively, at the same level of theory. On the other hand, the computed SCS-CC2 frequencies would provide a coincident assignment, as both O-docking isomers would have the closest fundamentals compared to the measured frequency (3619 and 3621 cm^{-1}). The assignment, however, would be tentative at best with this information alone. The OH–E isomers on the other hand can be excluded due to their lower OH stretching frequencies as well as the energetic disadvantage at the LCCSD(T0)-F12/CBS[T:Q] level (cf. Table 1).

In order to elucidate this problem, the electronic excitation energies can serve as a further indication for the binding motif, as shown for DPE–alcohol clusters before [20-22]. Comparing the vertical excitation energies for the different isomers with the experimental excitation energy of 36885 cm^{-1} yields the best agreement for the OH–O or OH–O' isomer, which also show a significantly blue-shifted $S_1 \leftarrow S_0$ transition compared to the PVE monomer (adiabatic excitation energies of 38291 and 38164 cm^{-1}, respectively, compared to 38034 cm^{-1} for the PVE monomer, cf. Table S2, Supporting Information File 1), as observed experimentally. In contrast to that, a red-shifted $S_1 \leftarrow S_0$ transition compared to PVE is predicted for the OH–P isomer (37907 cm^{-1}), which would coincide with the fragmentation-dominated region of the R2PI spectrum, where, however, only signatures of larger clusters could be identified. These considerations strengthen the arguments for the presence of an OH···O structure laid before, on the basis of the computed coupled cluster energies and the SCS-CC2 fundamental stretch frequencies. Additional experimental insight will be gained from the UV/IR/UV spectrum of the S_1 state as well as the microwave investigations in the following section.

Chirp pulse Fourier transform microwave (CP-FTMW) spectroscopy

From the broadband CP-FTMW spectra obtained with neon as a carrier gas, we assigned two PVE–MeOH complexes with significantly different intensities. Complex 1 is about ten times more intense than complex 2. The experimental rotational constants (Table 3) for the two isomers agree the best with the values calculated for the OH–O' isomer (as also indicated in the FTIR and the IR–UV investigations, which are, however, unable to distinguish OH–O from OH–O') and the OH–P isomer, respectively. The identification of the two complexes to the OH–O' and the OH–P isomers is guided by the absolute and relative values of the B and C rotational constants. Generally, the rotational constants calculated at the SCS-CC2/def2-TZVP level of theory agree somewhat better with the experimental values than the B3LYP-D3(BJ)/def2-TZVP values (note that we compare experimental B_0 rotational constants with theoretical B_e rotational constants here). For the OH–P complex, however, we find that the B3LYP-D3(BJ)/def2-TZVP level of theory provides a better prediction of the magnitudes of the dipole-moment components. Experimentally, we only observe a-type transitions for this complex, which points to rather low values for μ_b and μ_c. SCS-CC2 calculations predict all three dipole-moment components to be of comparable magnitude. At the B3LYP-D3(BJ) level, μ_a is predicted to be significantly stronger than μ_b and μ_c. This change in magnitude for the dipole-moment components for different levels of calculation is more often observed for weakly bound complexes because the exact arrangement of the two monomers with respect to each other can have a major influence on the dipole-moment components. Also note that in none of the spectroscopic experiments, we observe the OH–E isomer that is also predicted to be of relatively low energy (cf. Table 1).

The rotational spectra of the two isomers are qualitatively different. For the OH–O' isomer (complex 1), we observe a characteristic line splitting into so-called A and E components (cf. Figure 5) arising from internal rotation of the methyl group of methanol, similar to the case of the DPE–MeOH complex. For the OH–P isomer (complex 2), no line splitting due to internal rotation was observed. This is consistent with the higher barrier

Table 3: Experimental rotational constants of the two observed complexes, using neon as carrier gas, that are assigned to the OH–O' and the OH–P isomers, respectively. The experimental rotational parameters for the OH–O' isomer (called Exp 1) are the results of a fit to a rigid-rotor asymmetric Hamiltonian including solely the A lines of the internal rotation splitting. Rotational parameters of a global fit (XIAM) including both A and E levels due to internal rotation for the OH–O' isomer are presented in the Table S9 of Supporting Information File 1.

	Complex 1 (OH–O' isomer)		Complex 2 (OH–P isomer)	
	Exp 1	SCS-CC2/def2-TZVP	Exp 1	SCS-CC2/def2-TZVP
A [MHz]	1466.59120(26)	1501.94	1275.7623(49)	1297.89
B [MHz]	697.48965(11)	697.58	818.45271(73)	818.01
C [MHz]	572.109900(95)	589.94	640.2184(11)	646.81
Δ_J [kHz]	0.72697(62)		0.070(14)	
Δ_{JK} [kHz]	−0.6669(26)		2.19(10)	
Δ_K [kHz]	5.6217(62)		–	
δ_J [kHz]	0.15121(11)		–	
δ_K [kHz]	2.5783(29)		–	
A state transition	213 (49/104/60)		20(20/0/0)	
Dipole moment (D) ($\mu_a/\mu_b/\mu_c$)		2.2/1.9/1.2		0.8/0.4/0.8
σ [kHz]	6.7		7.9	

for this motion due to the secondary interactions of the methyl group with PVE (cf. Figure 2).

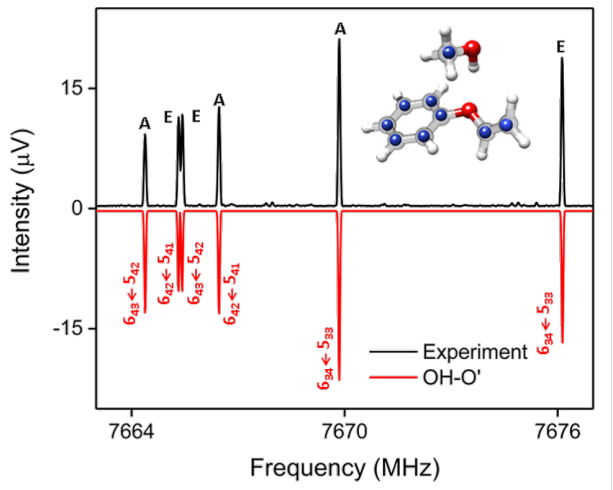


Figure 5: A section of the experimental 2–8 GHz spectrum using a mixture of PVE and MeOH (3 million acquisitions). The upper experimental trace in black is compared with simulations, based on fitted parameters that can be assigned to the OH–O' isomer (complex 1, red) for the PVE–MeOHcomplex. The observed complex has a clear splitting pattern due to the internal rotation of the methyl group of methanol, labeled with A and E. The experimental ^{13}C positions (blue atoms) (r_s substitution structure) deduced from a Kraitchman analysis are compared to the calculated structure at the SCS-CC2/def2-TZVP level of theory and further confirm the observation of the OH–O' isomer.

Two different ways of analyzing the rotational spectrum of the OH–O' isomer (complex 1) were performed. In Table 3, the results from a fit to an asymmetric-top Hamiltonian of only the A state species of the internal rotation splitting pair, which is often a good approximation, is summarized (Exp 1). In addition, we used the program XIAM to perform a global fit including both A and E lines. This global fit does not only provide the rotational constants, but also parameters of the internal rotor, in this case the methyl group. This includes the barrier height for internal rotation as well as the geometrical arrangement of the rotor with respect to the overall rotating molecule, as also discussed for the DPE–MeOH complex [19]. For PVE–MeOH, the barrier height was determined to be 261 cm^{-1}, as summarized in Table S9 of Supporting Information File 1. This value is in agreement to barrier heights observed for other complexes with methanol [19]. It is somewhat lower than in the case of free methanol (373 cm^{-1}) and also lower than the calculated barrier height of 341 cm^{-1} (cf. Table S9, Supporting Information File 1). This somewhat lower methyl group internal rotation barrier for the OH–O' isomer could point to a softening of the C–O bond of methanol due to the hydrogen bond. The DID plots in Figure 2 also indicate that the methyl group is basically free from other interactions, so that no additional hindering is expected.

Furthermore, the transition intensities for the OH–O' isomer are strong enough (with a signal-to-noise (SNR) of about 500:1 to 600:1 for the stronger transitions) to assign rotational transitions arising from all nine singly substituted ^{13}C isotopologues in natural abundance (about 1%, cf. Figure S3, Supporting Information File 1). The additional data sets of rotational constants are summarized in Supporting Information File 1 (Table S12) together with line lists of the main isotopologues (Tables S10–S11) and the ^{13}C isotopologues (Tables S13–S21). They allow us, using Kraitchman's equations, to determine the

carbon substitution structure, r_s, of the complexes, which are the positions of the respective substituted carbon atoms with respect to the center of mass of the complex and thus the carbon backbone structure. The obtained r_s structure for complex 1 (cf. Figure 5) further confirms the assignment of complex 1 as the OH–O' isomer, where the methyl group of the methanol moiety points towards the phenyl ring.

As mentioned, the OH–O' isomer is about ten times more intense than the OH–P isomer. The intensity observed in CP-FTMW spectroscopy directly depends on the number of molecules, i.e., the population of the respective isomers, as well as the square of the transition dipole moments. Since the μ_a values for the two isomers differ by a factor of two (cf. Table S11 in Supporting Information File 1), the OH–P isomer can be considered to be about 2.5 times less populated than the OH–O' isomer, as an upper estimate. Taking the predicted energy difference of 0.8 kJ/mol for granted, this ratio would correspond to a plausible [18] conformational freezing temperature of 100 K. A three-fold lower or three-fold higher conformational temperature appears unlikely, and thus a tentative experimental energy penalty for OH–P relative to OH–O' ranges from 0.3 to 2 kJ/mol. This contradicts both inexpensive approaches (B3LYP-D3 and SCS-CC2 with def2-TZVP) and suggests that these methods somewhat underestimate the stability of OH···O contacts.

Electronically excited state spectrum

For the investigation of the electronically excited state by using the UV/IR/UV technique, a two-color R2PI signal is required. For this reason, the one-color R2PI signal was suppressed by attenuating the laser power of the excitation laser. On the other hand, higher pulse energies were used for the ionizing laser. The latter was set to 31847 cm^{-1} for the UV/IR/UV experiment in order to yield the best two-color R2PI signal. Figure 6 shows the recorded UV/IR/UV spectrum for the PVE–MeOH mass trace.

Due to temporally overlapping laser pulses, the spectrum contains transitions from the S_0 state, the electronically excited (S_1, at 3637 cm^{-1}) and also the ionic D_0 state (at 3667 cm^{-1}). This could not be avoided, as the lifetime of the excited state, which is estimated to be in the order of 5–7 ns, is shorter than the laser pulse-widths of 7–10 ns. The OH stretching vibration at 3637 cm^{-1} originating from the electronically excited state of the PVE–MeOH complex is blue-shifted compared to the ground state, which indicates a decrease of the hydrogen bond strength in the S_1 compared to the S_0 state. A comparison with SCS-CC2 calculations shows a good agreement of a blue-shifted OH stretching frequency at 3642 cm^{-1} (cf. Table S2, Supporting Information File 1, scaled by 0.9635) compared to

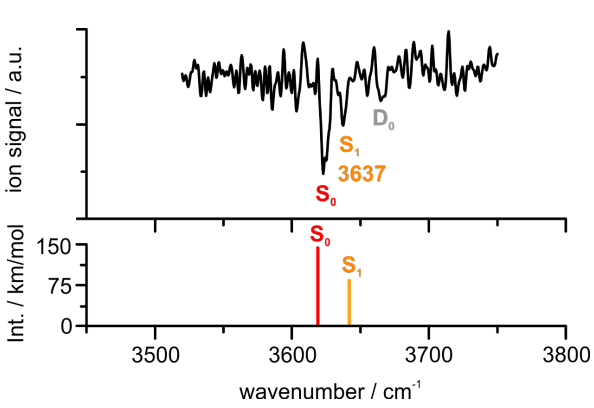


Figure 6: UV/IR/UV spectrum of PVE–MeOH in the range of 3520–3750 cm^{-1}; excitation laser: 36741 cm^{-1}, ionizing laser: 31847 cm^{-1}, carrier gas helium; the lower trace shows the calculated OH stretching frequencies at the SCS-CC2/def2-TZVP level for the optimized S_0 and S_1 structure of the OH–O' isomer scaled by 0.9635.

the ground state at 3619 cm^{-1} (cf. Table 1) for the OH–O' isomer, which is also reflected in an increase of the H···O hydrogen bond distance from 2.068 to 2.168 Å from S_0 to S_1 state geometry. This destabilization of the OH···O hydrogen bond is further reflected in the calculated binding energies of the PVE–MeOH complex obtained at the SCS-CC2/def2-TZVP level, which are reduced by 0.9 kJ/mol regarding D_0 and 1.6 kJ/mol regarding D_e in the S_1 state compared to the S_0 state for OH–O' (cf. Table S3, Supporting Information File 1). The spectral shift can be explained by regarding the HOMO and LUMO orbitals involved in the $S_1 \leftarrow S_0$ transition, which is predicted to be mainly a π–π^* transition with a small charge transfer contribution from the ether oxygen to the phenyl ring. The latter leads to a slightly decreased electron density at the binding site for the methanol molecule and therefore weakens the hydrogen bond. These findings are in line with observations in previous studies on diphenyl ether–alcohol complexes [20,21].

In principle, as the OH–O' isomer has been identified in the S_0 state, the observation of a respective OH···O-bound structure can be expected in the S_1 state as well. However, the OH–P isomers are predicted to be significantly stabilized in the S_1 state (cf. Table S2, Supporting Information File 1). Nevertheless, due to the predicted red-shifts of the OH stretching frequencies of the OH–P isomers (indicating an increased hydrogen bond strength compared to the S_0 state), their presence, i.e., by a rearrangement reaction from the OH–O' isomer, can be excluded. By exciting the electronic origin of the OH–O' isomer the formation of OH–E isomers can also be excluded as their expected excitation energies are higher than the one for OH–O' and in addition they are energetically less stable (cf. Table S2, Supporting Information File 1).

Conclusion

In this paper, the first spectroscopic and theoretical investigation on the isolated phenyl vinyl ether–methanol complex is presented. From the FTIR spectra, the existence of one isomer is concluded, which is confirmed by IR/UV spectroscopy in the electronic ground state (S_0). The combined vibrational and electronic spectroscopic investigations, including a comparison of vibrational frequencies and electronic excitation energies, allow for an assignment of an OH···O-bound structure. Broadband rotational (CP-FTMW) spectroscopy ultimately identifies OH–O' as the observed isomer, ruling out the presence of the nearly isoenergetic OH–O. One explanation for its elusiveness would be a low interconversion barrier. However, rotational spectroscopy further reveals the presence of the OH–P isomer as a second isomer, being less populated, which is not observed with the less sensitive FTIR technique and might be superimposed by fragmentation of larger clusters in the usually more sensitive IR/UV experiments or it is even not populated due to different expansion conditions. No evidence was found for an OH···ethenyl-bound structure, which is in agreement with the more pronounced energetic discrimination of OH–E isomers compared to the other binding motifs predicted at the LCCSD(T0)-F12/CBS[T:Q] level of theory.

In the electronically excited state (S_1), the OH stretching vibration of the attached methanol undergoes a blue-shift compared to the S_0 state. This indicates a weakening of the OH···O bond upon electronic excitation compared to the ground state and is in good agreement with the calculated frequency shift for the S_0 and S_1 state structures obtained at the SCS-CC2/def2-TZVP level and is furthermore in line with findings for similar diphenyl ether–alcohol complexes from previous investigations [20,21].

In summary, we present a multi-spectroscopic analysis on a molecular complex with a very delicate balance between, for the first time, three different binding motifs. This provides an excellent benchmark system for theory, since DFT-D3 as well as SCS-CC2 methods fail in predicting the correct energetic order, whereas LCCSD(T0)-F12 succeeds in the preferred docking motif. These differences are in the range of only 2 kJ/mol, when considering relative electronic energies, but that is already enough to tip the scales in the wrong direction. Comparing VTZ-F12 and VQZ-F12 results, we observe that the electronic energies are well converged for the smaller basis (Table S6, Supporting Information File 1). This would place the main accuracy bottleneck in the electronic structure method (i.e., functional, correlation truncation) chosen.

Finally, regarding the docking preference in comparison to the previously investigated diphenyl ether complex with methanol, a conclusion might be that methanol needs the interaction with a second phenyl ring in order to prefer the OH···π motif over OH···O, as observed for diphenyl ether. The secondary interaction of methanol with a smaller ethenyl moiety being present in phenyl vinyl ether instead of a phenyl ring seems to be insufficient to favor the phenyl docking site.

Supporting Information

Supporting Information File 1

Additional computational and experimental data.
[https://www.beilstein-journals.org/bjoc/content/supplementary/1860-5397-14-140-S1.pdf]

Acknowledgements

The authors thank the Deutsche Forschungsgemeinschaft (DFG, Ge 961/9-1, Schn 1280/4-1, Su 121/5-1, and Ma 5063/3-1) for financial support in context of the priority program SPP 1807 on the Control of London dispersion interactions in molecular chemistry. M.F. acknowledges support from the Hamburg International Max Planck Research School UFAST. This work is part of the Ph.D. theses of D.B., M.F., H.G. and A.W.

ORCID® iDs

Dominic Bernhard - https://orcid.org/0000-0002-6148-0445
Fabian Dietrich - https://orcid.org/0000-0001-5813-5248
Cristóbal Pérez - https://orcid.org/0000-0001-5248-5212
Hannes C. Gottschalk - https://orcid.org/0000-0003-4343-9542
Ricardo A. Mata - https://orcid.org/0000-0002-2720-3364
Martin A. Suhm - https://orcid.org/0000-0001-8841-7705
Melanie Schnell - https://orcid.org/0000-0001-7801-7134
Markus Gerhards - https://orcid.org/0000-0002-8748-2940

References

1. van der Waals, J. D. Ph.D. Thesis, Leiden, 1873.
2. London, F. *Z. Phys.* **1930,** *63,* 245–279. doi:10.1007/BF01421741
3. Müller-Dethlefs, K.; Hobza, P. *Chem. Rev.* **2000,** *100,* 143–168. doi:10.1021/cr9900331
4. Lehn, J.-M. *Angew. Chem., Int. Ed. Engl.* **1988,** *27,* 89–112. doi:10.1002/anie.198800891
5. Černý, J.; Hobza, P. *Phys. Chem. Chem. Phys.* **2007,** *9,* 5291–5303. doi:10.1039/b704781a
6. Meyer, E. A.; Castellano, R. K.; Diederich, F. *Angew. Chem., Int. Ed.* **2003,** *42,* 1210–1250. doi:10.1002/anie.200390319
7. Mata, R. A.; Suhm, M. A. *Angew. Chem., Int. Ed.* **2017,** *56,* 11011–11018. doi:10.1002/anie.201611308
8. Zwier, T. S. *Annu. Rev. Phys. Chem.* **1996,** *47,* 205–241. doi:10.1146/annurev.physchem.47.1.205
9. Ebata, T.; Fujii, A.; Mikami, N. *Int. Rev. Phys. Chem.* **1998,** *17,* 331–361. doi:10.1080/014423598230081
10. Brutschy, B. *Chem. Rev.* **2000,** *100,* 3891–3920. doi:10.1021/cr990055n
11. Pribble, R. N.; Hagemeister, F. C.; Zwier, T. S. *J. Chem. Phys.* **1997,** *106,* 2145. doi:10.1063/1.473784

12. Djafari, S.; Barth, H.-D.; Buchhold, K.; Brutschy, B. *J. Chem. Phys.* **1997,** *107,* 10573–10581. doi:10.1063/1.474221
13. Buchanan, E. G.; Gord, J. R.; Zwier, T. S. *J. Phys. Chem. Lett.* **2013,** *4,* 1644–1648. doi:10.1021/jz400641p
14. Walsh, P. S.; Buchanan, E. G.; Gord, J. R.; Zwier, T. S. *J. Chem. Phys.* **2015,** *142,* 154304. doi:10.1063/1.4917307
15. Walsh, P. S.; Buchanan, E. G.; Gord, J. R.; Zwier, T. S. *J. Chem. Phys.* **2015,** *142,* 154303. doi:10.1063/1.4917305
16. Pietraperzia, G.; Pasquini, M.; Mazzoni, F.; Piani, G.; Becucci, M.; Biczysko, M.; Michalski, D.; Bloino, J.; Barone, V. *J. Phys. Chem. A* **2011,** *115,* 9603–9611. doi:10.1021/jp200444a
17. Gottschalk, H. C.; Altnöder, J.; Heger, M.; Suhm, M. A. *Angew. Chem., Int. Ed.* **2016,** *55,* 1921–1924. doi:10.1002/anie.201508481
18. Poblotzki, A.; Gottschalk, H. C.; Suhm, M. A. *J. Phys. Chem. Lett.* **2017,** *8,* 5656–5665. doi:10.1021/acs.jpclett.7b02337
19. Medcraft, C.; Zinn, S.; Schnell, M.; Poblotzki, A.; Altnöder, J.; Heger, M.; Suhm, M. A.; Bernhard, D.; Stamm, A.; Dietrich, F.; Gerhards, M. *Phys. Chem. Chem. Phys.* **2016,** *18,* 25975–25983. doi:10.1039/C6CP03557D
20. Bernhard, D.; Dietrich, F.; Fatima, M.; Perez, C.; Poblotzki, A.; Jansen, G.; Suhm, M. A.; Schnell, M.; Gerhards, M. *Phys. Chem. Chem. Phys.* **2017,** *19,* 18076–18088. doi:10.1039/C7CP02967E
21. Bernhard, D.; Holzer, C.; Dietrich, F.; Stamm, A.; Klopper, W.; Gerhards, M. *ChemPhysChem* **2017,** *18,* 3634–3641. doi:10.1002/cphc.201700722
22. Dietrich, F.; Bernhard, D.; Fatima, M.; Perez, C.; Schnell, M.; Gerhards, M. *Angew. Chem., Int. Ed.* **2018,** in press. doi:10.1002/anie.201801842
23. Maier, J. M.; Li, P.; Vik, E. C.; Yehl, C. J.; Strickland, S. M. S.; Shimizu, K. D. *J. Am. Chem. Soc.* **2017,** *139,* 6550–6553. doi:10.1021/jacs.7b02349
24. Heger, M.; Mata, R. A.; Suhm, M. A. *Chem. Sci.* **2015,** *6,* 3738–3745. doi:10.1039/C5SC01002K
25. Medel, R.; Heger, M.; Suhm, M. A. *J. Phys. Chem. A* **2015,** *119,* 1723–1730. doi:10.1021/jp508424p
26. Mons, M.; Robertson, E. G.; Simons, J. P. *J. Phys. Chem. A* **2000,** *104,* 1430–1437. doi:10.1021/jp993178k
27. Page, R. H.; Shen, Y. R.; Lee, Y. T. *J. Chem. Phys.* **1988,** *88,* 4621. doi:10.1063/1.453775
28. Riehn, C.; Lahmann, C.; Wassermann, B.; Brutschy, B. *Chem. Phys. Lett.* **1992,** *197,* 443–450. doi:10.1016/0009-2614(92)85798-F
29. Gerhards, M.; Unterberg, C. *Phys. Chem. Chem. Phys.* **2002,** *4,* 1760–1765. doi:10.1039/b110029g
30. Loquais, Y.; Gloaguen, E.; Habka, S.; Vaquero-Vara, V.; Brenner, V.; Tardivel, B.; Mons, M. *J. Phys. Chem. A* **2015,** *119,* 5932–5941. doi:10.1021/jp509494c
31. Ebata, T.; Fujii, A.; Mikami, N. *Int. J. Mass Spectrom. Ion Processes* **1996,** *159,* 111–124. doi:10.1016/S0168-1176(96)04445-X
32. Gerhards, M.; Schumm, S.; Unterberg, C.; Kleinermanns, K. *Chem. Phys. Lett.* **1998,** *294,* 65–70. doi:10.1016/S0009-2614(98)00823-9
33. Palmer, P. M.; Chen, Y.; Topp, M. R. *Chem. Phys. Lett.* **2000,** *318,* 440–447. doi:10.1016/S0009-2614(00)00036-1
34. Robertson, E. G. *Chem. Phys. Lett.* **2000,** *325,* 299–307. doi:10.1016/S0009-2614(00)00661-8
35. Chen, Y.; Palmer, P. M.; Topp, M. R. *Int. J. Mass Spectrom.* **2002,** *220,* 231–251. doi:10.1016/S1387-3806(02)00772-8
36. Chen, Y.; Topp, M. R. *Chem. Phys.* **2002,** *283,* 249–268. doi:10.1016/S0301-0104(02)00613-4
37. Ishiuchi, S.-i.; Daigoku, K.; Saeki, M.; Sakai, M.; Hashimoto, K.; Fujii, M. *J. Chem. Phys.* **2002,** *117,* 7077–7082. doi:10.1063/1.1508103
38. Stearns, J. A.; Das, A.; Zwier, T. S. *Phys. Chem. Chem. Phys.* **2004,** *6,* 2605–2610. doi:10.1039/b313831c
39. Fricke, H.; Bartl, K.; Funk, A.; Gerlach, A.; Gerhards, M. *ChemPhysChem* **2008,** *9,* 2592–2600. doi:10.1002/cphc.200800499
40. Bartl, K.; Funk, A.; Gerhards, M. *J. Chem. Phys.* **2008,** *129,* 234306. doi:10.1063/1.3037023
41. Bartl, K.; Funk, A.; Gerhards, M. *ChemPhysChem* **2009,** *10,* 1882–1886. doi:10.1002/cphc.200900097
42. Stamm, A.; Weiler, M.; Brächer, A.; Schwing, K.; Gerhards, M. *Phys. Chem. Chem. Phys.* **2014,** *16,* 21795–21803. doi:10.1039/C4CP02546F
43. Miyazaki, M.; Fujii, M. *Phys. Chem. Chem. Phys.* **2017,** *19,* 22759–22776. doi:10.1039/C7CP03480F
44. Grimme, S. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2011,** *1,* 211–228. doi:10.1002/wcms.30
45. Grimme, S.; Ehrlich, S.; Goerigk, L. *J. Comput. Chem.* **2011,** *32,* 1456–1465. doi:10.1002/jcc.21759
46. Hellweg, A.; Grün, S. A.; Hättig, C. *Phys. Chem. Chem. Phys.* **2008,** *10,* 4119–4127. doi:10.1039/b803727b
47. Adler, T. B.; Werner, H.-J. *J. Chem. Phys.* **2009,** *130,* 241101. doi:10.1063/1.3160675
48. Wuttke, A.; Mata, R. A. *J. Comput. Chem.* **2017,** *38,* 15–23. doi:10.1002/jcc.24508
49. Suhm, M. A.; Kollipost, F. *Phys. Chem. Chem. Phys.* **2013,** *15,* 10702–10721. doi:10.1039/c3cp51515j
50. Unterberg, C.; Jansen, A.; Gerhards, M. *J. Chem. Phys.* **2000,** *113,* 7945. doi:10.1063/1.1315610
51. Iwasaki, T.; Miyata, Y.; Akimoto, R.; Fujii, Y.; Kuniyasu, H.; Kambe, N. *J. Am. Chem. Soc.* **2014,** *136,* 9260–9263. doi:10.1021/ja5043534
52. Schmitz, D.; Alvin Shubert, V.; Betz, T.; Schnell, M. *J. Mol. Spectrosc.* **2012,** *280,* 77–84. doi:10.1016/j.jms.2012.08.001
53. Pérez, C.; Lobsiger, S.; Seifert, N. A.; Zaleski, D. P.; Temelso, B.; Shields, G. C.; Kisiel, Z.; Pate, B. H. *Chem. Phys. Lett.* **2013,** *571,* 1–15. doi:10.1016/j.cplett.2013.04.014
54. Kraitchman, J. *Am. J. Phys.* **1953,** *21,* 17–24. doi:10.1119/1.1933338
55. *JB95 Spectral fitting program;* NIST. Plusquellic, D.
56. *Gaussian 09,* revision D.01; Gaussian, Inc.: Wallingford, CT, USA, 2009.
57. Halgren, T. A. *J. Comput. Chem.* **1996,** *17,* 490–519. doi:10.1002/(SICI)1096-987X(199604)17:5/6<490::AID-JCC1>3.0.CO;2-P
58. Hanwell, M. D.; Curtis, D. E.; Lonie, D. C.; Vandermeersch, T.; Zurek, E.; Hutchison, G. R. *J. Cheminf.* **2012,** *4,* 17. doi:10.1186/1758-2946-4-17
59. Furche, F.; Ahlrichs, R.; Hättig, C.; Klopper, W.; Sierka, M.; Weigend, F. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2014,** *4,* 91–100. doi:10.1002/wcms.1162
60. Weigend, F.; Köhn, A.; Hättig, C. *J. Chem. Phys.* **2002,** *116,* 3175. doi:10.1063/1.1445115
61. Peterson, K. A.; Adler, T. B.; Werner, H.-J. *J. Chem. Phys.* **2008,** *128,* 84102. doi:10.1063/1.2831537
62. Yousaf, K. E.; Peterson, K. A. *J. Chem. Phys.* **2008,** *129,* 184108. doi:10.1063/1.3009271

63. Hill, J. G.; Peterson, K. A.; Knizia, G.; Werner, H.-J. *J. Chem. Phys.* **2009,** *131,* 194105. doi:10.1063/1.3265857
64. Pipek, J.; Mezey, P. G. *J. Chem. Phys.* **1989,** *90,* 4916–4926. doi:10.1063/1.456588
65. Reed, A. E.; Weinstock, R. B.; Weinhold, F. *J. Chem. Phys.* **1985,** *83,* 735–746. doi:10.1063/1.449486
66. *MOLPRO,* version 2015.1; a package of ab initio programs, http://www.molpro.net.
67. Schütz, M.; Rauhut, G.; Werner, H.-J. *J. Phys. Chem. A* **1998,** *102,* 5997–6003. doi:10.1021/jp981168y
68. Andrejić, M.; Mata, R. A. *Phys. Chem. Chem. Phys.* **2013,** *15,* 18115–18122. doi:10.1039/c3cp52931b
69. 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. *Nature* **2011,** *477,* 308–311. doi:10.1038/nature10367
70. Löffler, S.; Wuttke, A.; Zhang, B.; Holstein, J. J.; Mata, R. A.; Clever, G. H. *Chem. Commun.* **2017,** *53,* 11933–11936. doi:10.1039/C7CC04855F

License and Terms

BEILSTEIN JOURNAL OF ORGANIC CHEMISTRY

Evaluation of dispersion type metal···π arene interaction in arylbismuth compounds – an experimental and theoretical study

Ana-Maria Preda[1], Małgorzata Krasowska[2], Lydia Wrobel[1], Philipp Kitschke[1], Phil C. Andrews[3], Jonathan G. MacLellan[3], Lutz Mertens[1], Marcus Korb[4], Tobias Rüffer[4], Heinrich Lang[4], Alexander A. Auer[*2] and Michael Mehring[*1]

Full Research Paper

Open Access

Address:
[1]Technische Universität Chemnitz, Fakultät für Naturwissenschaften, Institut für Chemie, Professur Koordinationschemie, 09107 Chemnitz, Germany, [2]Max-Planck-Institut für Kohlenforschung, Kaiser-Wilhelm-Platz 1, 45470 Mülheim an der Ruhr, Germany, [3]School of Chemistry, Monash University, Clayton, Melbourne, VIC 3800, Australia and [4]Technische Universität Chemnitz, Fakultät für Naturwissenschaften, Institut für Chemie, Professur Anorganische Chemie, 09107 Chemnitz, Germany

Email:
Alexander A. Auer[*] - alexander.auer@kofo.mpg.de;
Michael Mehring[*] - michael.mehring@chemie.tu-chemnitz.de

* Corresponding author

Keywords:
arylbismuth compounds; DFT-D; dispersion type Bi···π arene interaction; DLPNO-CCSD(T); electronic structure calculations; polymorphism; single crystal X-ray structure

Beilstein J. Org. Chem. **2018,** *14,* 2125–2145.
doi:10.3762/bjoc.14.187

Received: 19 March 2018
Accepted: 12 July 2018
Published: 15 August 2018

This article is part of the Thematic Series "Dispersion interactions".

Guest Editor: P. Schreiner

Abstract

The dispersion type Bi···π arene interaction is one of the important structural features in the assembly process of arylbismuth compounds. Several triarylbismuth compounds and polymorphs are discussed and compared based on the analysis of single crystal X-ray diffraction data and computational studies. First, the crystal structures of polymorphs of Ph_3Bi (**1**) are described emphasizing on the description of London dispersion type bismuth···π arene interactions and other van der Waals interactions in the solid state and the effect of it on polymorphism. For comparison we have chosen the substituted arylbismuth compounds $(C_6H_4\text{-}CH{=}CH_2\text{-}4)_3Bi$ (**2**), $(C_6H_4\text{-}OMe\text{-}4)_3Bi$ (**3**), $(C_6H_3\text{-}t\text{-}Bu_2\text{-}3,5)_3Bi$ (**4**) and $(C_6H_3\text{-}t\text{-}Bu_2\text{-}3,5)_2BiCl$ (**5**). The structural analyses revealed that only two of them show London dispersion type bismuth···π arene interactions. One of them is the styryl derivative **2**, for which two polymorphs were isolated. Polymorph **2a** crystallizes in the orthorhombic space group $P2_12_12_1$, while polymorph **2b** exhibits the monoclinic space group $P2_1/c$. The general structure of **2a** is similar to the monoclinic $C2/c$ modification of Ph_3Bi (**1a**), which leads to the formation of zig-zag Bi–$arene_{centroid}$ ribbons formed as a result of bismuth···π arene interactions and π···π intermolecular contacts. In the crystal structures of the polymorph **2b** as well as for **4** bismuth···π arene interactions are not observed, but

both compounds revealed C–H_{Ph}···π intermolecular contacts, as likewise observed in all of the three described polymorphs of Ph_3Bi. For compound **3** intermolecular contacts as a result of coordination of the methoxy group to neighboring bismuth atoms are observed overruling Bi···π arene contacts. Compound **5** shows a combination of donor acceptor Bi···Cl and Bi···π arene interactions, resulting in an intermolecular pincer-type coordination at the bismuth atom. A detailed analysis of three polymorphs of Ph_3Bi (**1**), which were chosen as model systems, at the DFT-D level of theory supported by DLPNO-CCSD(T) calculations reveals how van der Waals interactions between different structural features balance in order to stabilize molecular arrangements present in the crystal structure. Furthermore, the computational results allow to group this class of compounds into the range of heavy main group element compounds which have been characterized as dispersion energy donors in previous work.

Introduction

Although known for more than a century, the interest on metal···π arene interaction of main group metals has increased significantly, both experimentally and theoretically in the past decade [1-5]. Especially the development of novel computational tools demonstrated the importance of London dispersion type interactions for structures and functions of molecules [6-8]. With regard to this the high relevance of London dispersion type interactions in molecular organometallic chemistry was recently summarized by Liptrot and Power [9]. It should be noted that in this context and more generally organometallic bismuth compounds are witnessing growing attention since applications in the field of supramolecular chemistry [10-12] and pharmacology are of interest [13-15].

Lately, several studies regarding the metal···π interactions in organometallic compounds of antimony and bismuth [16-19] have been reported including intramolecular [20-22] and intermolecular coordination [23,24]. Special attention was given to bismuth···π arene interaction by us including the formation of dimers and networks [1,25-28], and recently we reported a study on the effect of intermolecular dispersion type interaction on polymorphism and phase transition of compounds of the type Ar_3Bi (Ar = C_4H_3NMe, C_4H_3O, C_4H_3S, C_4H_3Se) [28,29].

Other state of the art examples on the formation of supramolecular assemblies via dispersion type metal···π arene intermolecular interactions [10,11] were summarized by Caracelli et al., and recently Tiekink classified this type of interaction as one of the emerging intermolecular interactions that are of particular interest to coordination chemists with regard to supramolecular chemistry [12]. However, most reports on main group metal···π interactions are based on the description of the single crystal structures and lack a profound description of the theoretical background so far. Rare examples on theoretical work about the pnictogen···π interaction were given by Frontera et al. [30,31]. While analysis of structural parameters like interatomic distances allows to assess the plausibility of certain interactions, this is exceedingly difficult and sometimes misleading for weak intermolecular interactions. Here, the accurate quantification that is possible using computational methods allows to gain a deeper understanding of which interactions are dominating. This way, a given crystal structure can be rationalized, for example, as consisting of strongly interacting dimers which themselves interact weakly with their surroundings based on the actual interaction energies. Elucidation of this is already possible at the DFT-D level of theory, if functionals with established accuracy are used, or at the DLPNO-CCSD(T) level of theory, which yields near-quantitative accuracy from first principles and can be applied to fairly large systems [32-38].

Herein, we report on intermolecular interactions with focus on bismuth···π arene interactions for the crystal structures of three polymorphs of Ph_3Bi (**1**). For comparison the crystal structures of substituted arylbismuth compounds of the type Ar_3Bi [Ar = C_6H_4-CH═CH_2-4 (**2a**, **2b**), C_6H_4-OMe-4 (**3**)], Ar'_3Bi (**4**) and Ar'_2BiCl (**5**, Ar' = C_6H_3-*t*-Bu_2-3,5) were analyzed with regard to their packing in the solid state. Electronic structure calculations were carried out on Ph_3Bi···C_6H_6 and selected polymorphs of Ph_3Bi (**1**). For this purpose, a series of electronic structure methods are applied for a model compound in order to assess the performance of different methods and to conceptually investigate and quantify the heavy main group element···π interaction present in these type of compounds. In the second part, DFT-D and DLPNO-CCSD(T) calculations are carried out for a series of molecular structures, dimers, trimers and tetramers that have been taken from the crystal structures of three selected polymorphs of compound **1**. This allows to quantify and to rationalize the balance of dispersion type interactions between bismuth and aromatic ligands as well as between the aromatic ligands itself.

Results and Discussion

Synthesis

So far, four polymorphs of Ph_3Bi (**1**) have been reported in the literature [39-45], but none of these reports contains an analysis of dispersion type interactions including bismuth···π interaction in the solid state. This prompted us to have a closer look at these simple organometallic compounds. Noteworthy, the first

report on the synthesis of Ph_3Bi dates back to 1887, which was based on the reaction of sodium alloy and bromobenzene [46,47]. A more convenient synthetic route makes use of the Grignard reagent phenylmagnesium bromide and its reaction with bismuth trichloride [48]. Following this approach with slight modifications provides Ph_3Bi with a yield of more than 80%. Crystallization from EtOH gave single crystals of the monoclinic *C*2/*c* polymorph **1a**, which was already subject of several studies including the description of its crystal structure [39-43]. Therefore, it is somewhat surprising that the bismuth···π interaction was not noted so far. We obtained polymorph **1a** upon crystallization from solution, but we isolated another polymorph **1b** by crystallization from the melt. Polymorph **1b** was obtained starting from **1a** in a temperature-dependent PXRD experiment (see Supporting Information File 1, Figure S1). The polymorph **1b** was obtained as a microcrystalline material, but Andrews and MacLellan did obtain single crystals on this orthorhombic form **1b** prior to this study [45]. Noteworthy, a phase transition of **1a** to **1b** does not occur before melting.

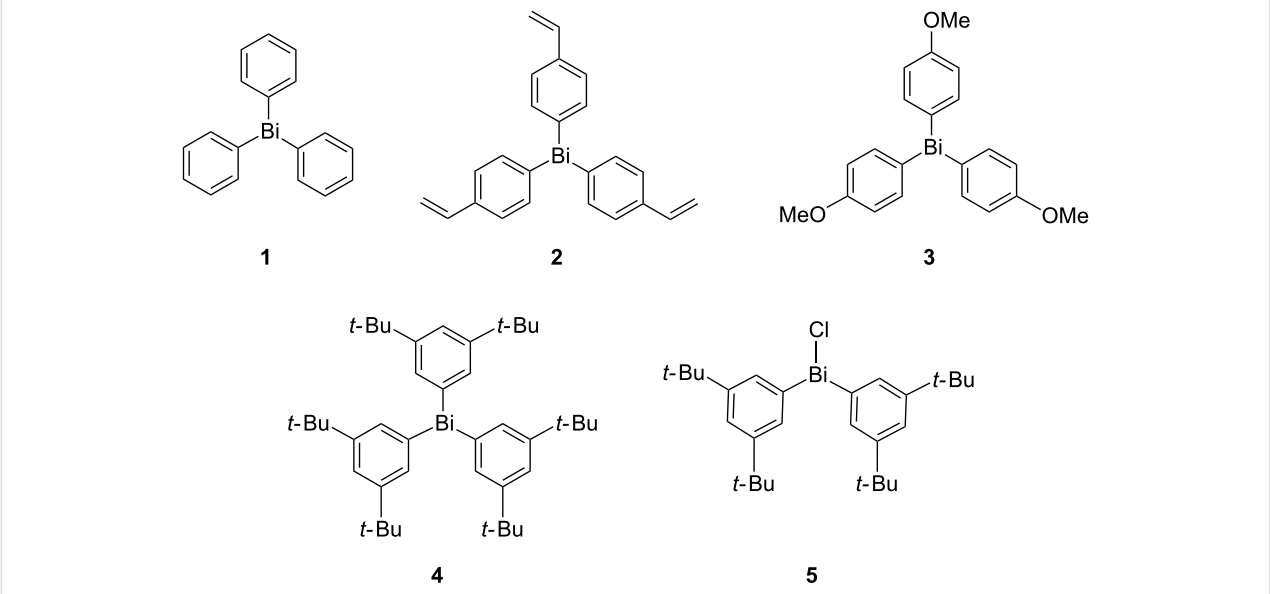
The latest report on a polymorph of Ph_3Bi was made by Stammler and Neumann, which submitted the crystallographic data of a monoclinic $P2_1/c$ (**1c**) polymorph to the Cambridge Crystallographic Data Base [44]. In addition a monoclinic polymorph **1d** was mentioned in a brief report of Wetzel as early as 1942, but the atomic parameters were not given [39].

Following the Grignard route we were able to develop a straightforward synthetic protocol for $(C_6H_4\text{-}CH{=}CH_2\text{-}4)_3Bi$ (**2**) starting from 4-bromostyrene and isolated compound **2** with 84% yield. The synthesis of **2** with very low yield is mentioned in a patent from 1964 [49], but **2** was neither fully characterized, nor was its crystal structure determined. We were able to crystallize two polymorphs of **2**, an orthorhombic form **2a** and a monoclinic form **2b**, both were obtained from iPrOH solution.

In order to develop a better understanding with regard to the effects of substituents, $(C_6H_4\text{-OMe-}4)_3Bi$ (**3**) [50,51] was prepared starting from $BiCl_3$ and the corresponding organolithium reagent following a general method as reported by Wang et al. [52]. Compound **3** was obtained as colorless block-shaped crystals in yields of 83%. While our work was in progress, a crystal structure of **3** was reported by Gagnon et al. The authors confirmed the formation of **3** from the corresponding Grignard reagent and $BiCl_3$ upon crystallization at 20 °C by diffusion of *n*-hexane into CH_2Cl_2 solution [53], but only gave a very brief description of the molecular structure.

The Ar_3Bi compound $(C_6H_3\text{-}t\text{-}Bu_2\text{-}3,5)_3Bi$ (**4**) was prepared with a yield of 73% following the Grignard route, with the intention to study the influence of very bulky substituents. Finally its chloro derivative $(C_6H_3\text{-}t\text{-}Bu_2\text{-}3,5)_2BiCl$ (**5**) was synthesized in 11% yield using the organolithium derivative $(C_6H_3\text{-}t\text{-}Bu_2\text{-}3,5)Li$ and $BiCl_3$.

This series of compounds and polymorphs (Scheme 1) allows to deduce some general trends regarding dispersion type interactions including bismuth···π, π···π and C–H···π interactions in organobismuth compounds and therefore the crystal structures are described and discussed in the following chapter. Please



Scheme 1: Triarylbismuth compounds, that serve as examples for the investigation of bismuth···π interactions in the solid state.

note that the term C–H···π is used as a structure descriptor rather than to describe a special type of bonding. Thus we follow the criticism given by Grimme [54] and Iverson et al. [55] on the unreflected use of terms such as C–H···π, or π···π stacking previously. In most cases, these interactions rely on London dispersion forces rather than special types of bonding due to the π system.

Crystal structures

In all of the presented compounds, the arrangement at the bismuth atom is best described as a slightly distorted trigonal pyramid, with the C–Bi–C angles significantly smaller than the tetrahedral angle, indicating that the lone pair is of mainly 6s character [41]. The Bi–C distances and C–Bi–C angles correspond to bond lengths and angles as observed for the various modifications of Ph_3Bi [1,39,41-44] and other Ar_3Bi compounds (Ar = Mes [56], *p*-Tolyl [57]). The molecular structures of **2a**, **2b**, **4** and **5** are illustrated in Figures S1–S4 (Supporting Information File 1), the selected bond lengths and angles are listed in the corresponding figure captions. Here, we focus mainly on the description of the supramolecular arrangements of these compounds in the solid state.

In the literature several reports exist on the monoclinic polymorph of Ph_3Bi (**1a**), which crystallizes in the space group *C*2/*c* [39-43]. The Bi···π arene interactions range from 3.727–3.856 Å, leading to the formation of 1D ribbons in the solid state due to Bi···π arene interactions (see zig-zag $(Bi–arene_{centroid})_\infty$ chain in Figure 1a). These chains are further connected via $C–H_{Ph}···π$ ($arene_{centroid}$) intermolecular contacts with $C33–H33_{Ph}···π$ ($arene_{centroid}$) distances of 3.030 Å (blue dashed line), γ = 10.9°, (two parallelograms connected via one edge in Figure 1b). Furthermore, two $C–H_{Ph}···π$ ($arene_{centroid}$) intermolecular contacts are observed with $C14–H14_{Ph}···π$ ($arene_{centroid}$) distances of 3.042 Å (green dashed line, γ = 19.5°) and $C15–H15_{Ph}···π$ ($arene_{centroid}$) distances of 2.760 Å (black dashed line, γ = 6.4°) to give a 2D network (Figure 1c). Other additional $C–H_{Ph}···π$ ($arene_{centroid}$) intermolecular contacts with $C36–H36_{Ph}···π$ ($arene_{centroid}$) distances of

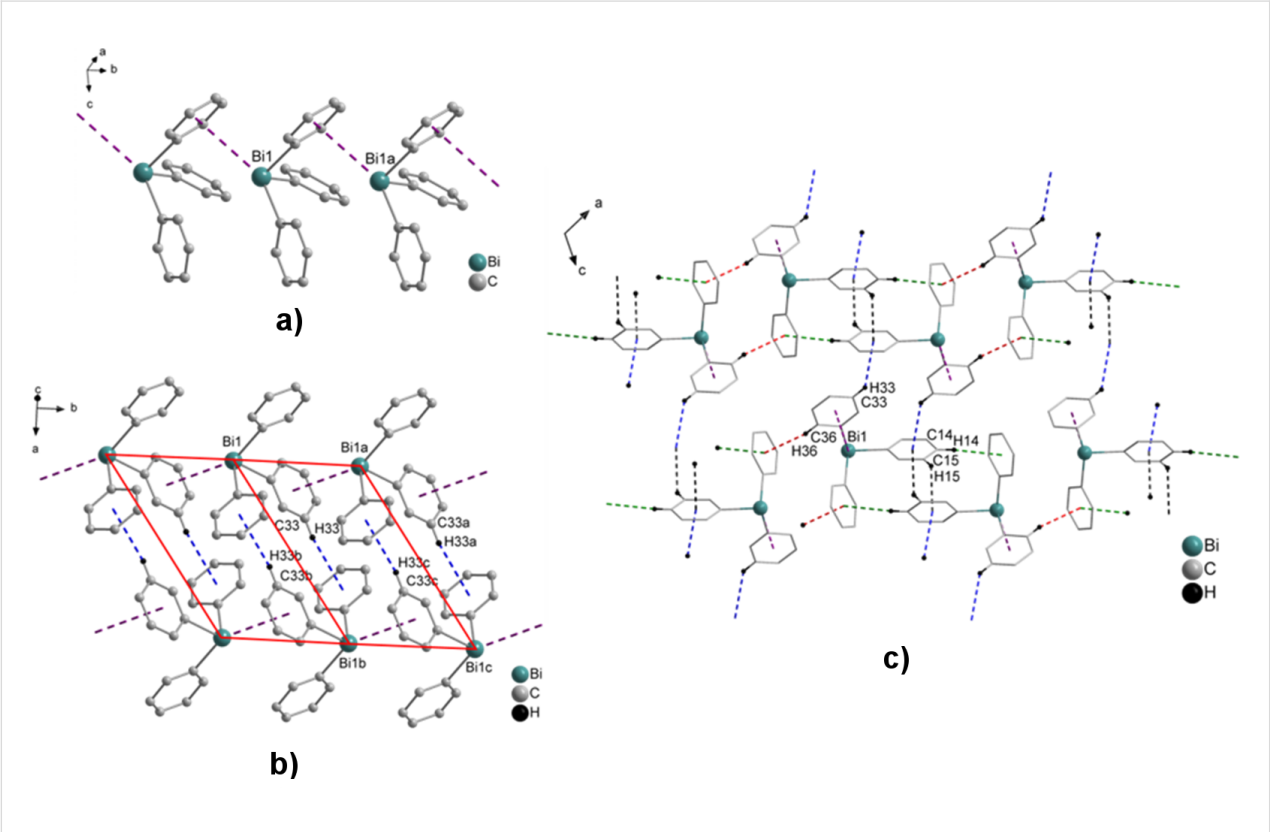


Figure 1: Ball and stick model of a fragment of a) the zig-zag chain of a 1D arrangement of Ph_3Bi (**1a**). Hydrogen atoms were omitted for clarity. Selected distances [Å]: Bi1–$arene_{centroid}$ 3.763 (violet dash line) [39-43]; b) formation of the two parallelograms connected via one edge, formed via two $C–H_{Ph}···π$ ($arene_{centroid}$) intermolecular contacts with $C33–H33_{Ph}···π$ ($arene_{centroid}$) 3.030 Å (blue dashed line, γ = 10.9°); c) wire and stick model of 2D and 3D networks formed via $C–H_{Ph}···π$ ($arene_{centroid}$) intermolecular contacts $C14–H14_{Ph}···π$ ($arene_{centroid}$) distances of 3.042 Å (green dashed line, γ = 19.5°), $C15–H15_{Ph}···π$ ($arene_{centroid}$) distances of 2.760 Å (black dashed line, γ = 6.4°) and $C36–H36_{Ph}···π$ ($arene_{centroid}$) distances of 2.740 Å (red dashed line, γ = 11.2°), respectively (only hydrogen atoms involved in $C–H_{Ph}···π$ ($arene_{centroid}$) contacts are shown). Symmetry transformation: a = *x*, 1 + *y*, *z*; b = ½ – *x*, ³⁄₂ – *y*, –*z*; c = ½ – *x*, ⁵⁄₂ – *y*, –*z*.

2.740 Å (red dashed line, γ = 11.2°) lead to the formation of a 3D network in the solid state (Figure 1c). The C–H_{Ph}···$arene_{centroid}$ contacts are shorter than 3.1 Å with an angle γ between the normal to the arene ring and the line defined by the H atom and the $arene_{centroid}$ smaller than 30° [58,59].

The polymorph **1b** crystallizes in the orthorhombic space group $Pna2_1$ [45]. The crystal structure of polymorph **1b** shows two different bismuth atoms in the unit cell, each of them being involved in Bi···π arene intermolecular interactions, with Bi2–$arene_{centroid}$ 3.468 Å (grey dashed line in Figure 2) and Bi1–$arene_{centroid}$ 3.561 Å (blue dashed line in Figure 2), thus resulting in zig-zag type 1D ribbons. In addition C–H_{Ph}···π ($arene_{centroid}$) intermolecular contacts with C17–$H17_{Ph}$···π ($arene_{centroid}$) 3.083 Å (γ = 9.5°) and C29–$H29_{Ph}$···π ($arene_{centroid}$) 3.097 Å, (γ = 16.4°, green and purple dashed line in Figure 2a, respectively) complement the structure. The ribbons are connected via two additional C–H_{Ph}···π ($arene_{centroid}$) intermolecular contacts with C15–$H15_{Ph}$···π ($arene_{centroid}$) 3.034 Å (black dashed line, γ = 8.4°) and C28–$H28_{Ph}$···π ($arene_{centroid}$) 2.890 Å (brown dashed line, γ = 13.5°) and lead to the formation of a 2D network (Figure 2b).

Another polymorph of Ph_3Bi (**1c**) was reported by Neumann and co-workers in a CSD communication. The polymorph **1c** crystallizes in the monoclinic space group $P2_1/c$ [1,44]. The crystal structure of polymorph **1c** reveals the formation of non-centrosymmetric dimers in the solid state, which are formed via two Bi···π arene intermolecular contacts (Figure 3a). These distances amount to Bi1–$arene_{centroid}$ 3.787 Å (green dashed line) and Bi2–$arene_{centroid}$ 3.939 Å (lime dashed line). A closer look at the crystal structure of polymorph **1c** reveals that the dimeric units self-assemble via C–H_{Ph}···$arene_{centroid}$ contacts, which leads to the formation of centrosymmetric units, based on C–H_{Ph}···π ($arene_{centroid}$) intermolecular contacts (four-membered ring in Figure 3b), with C20–$H20_{Ph}$···π ($arene_{centroid}$) 2.801 Å (dark red dashed line, γ = 12.3°) and C27–$H27_{Ph}$···π ($arene_{centroid}$) 2.763 Å (teal dashed line, γ = 12.6°), (1D layers in Figure 3b), respectively. Additionally, the 1D layers are connected via two C–H_{Ph}···π ($arene_{centroid}$) intermolecular contacts, with C3–$H3_{Ph}$···π ($arene_{centroid}$) 3.037 Å (blue dashed line, γ = 7.9°) to give a 3D network in the solid state (Figure 3c).

It is worth to note that Wetzel has reported another crystal structure of Ph_3Bi (**1d**) in 1942, which crystallizes in the triclinic space group $P\bar{1}$ [39]. Unfortunately, the crystal structure of the latter could not be analyzed by us due to the lack of atomic parameters.

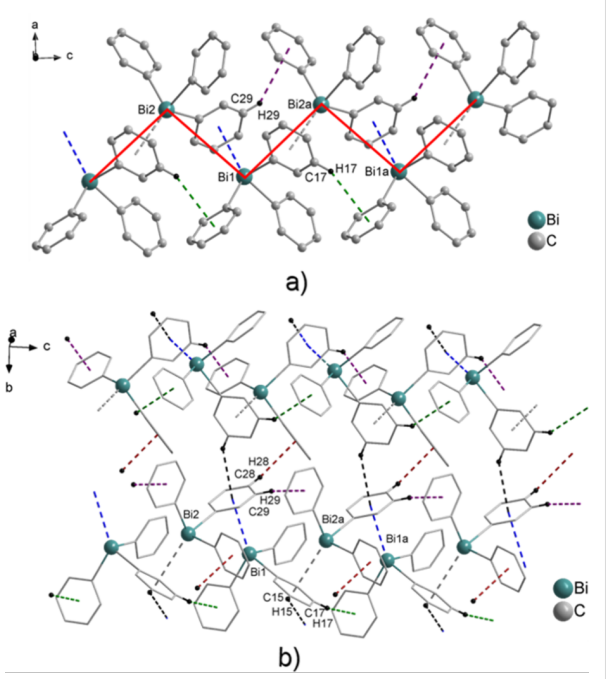


Figure 2: Ball and stick model of a fragment of the zig-zag type arrangement of Ph_3Bi (**1b**) [45], view along the *b* axis. Hydrogen atoms were omitted for clarity. Selected distances [Å]: Bi1–$arene_{centroid}$ 3.561 (blue dashed line), Bi2–$arene_{centroid}$ 3.468 (grey dash line); a) the formation of dimers via two C–H_{Ph}···π ($arene_{centroid}$) intermolecular contacts with C17–$H17_{Ph}$···π ($arene_{centroid}$) 3.083 Å (green dashed line, γ = 9.5°) and C29–$H29_{Ph}$···π ($arene_{centroid}$) 3.097 Å (γ = 16.4°); b) wire and stick model of a 2D network build via additional C–H_{Ph}···π ($arene_{centroid}$) intermolecular contacts with C15–$H15_{Ph}$···π ($arene_{centroid}$) 3.034 Å (black dashed line), (γ = 8.4°) and C28–$H28_{Ph}$···π ($arene_{centroid}$) 2.890 Å (brown dashed line, γ = 13.5°), (only hydrogen atoms involved in C–H_{Ph}···π ($arene_{centroid}$) contacts are shown). Symmetry transformations: a = *x*, *y*, 1 + *z*.

$(C_6H_4$-CH=CH_2-4$)_3Bi$ (**2**)

Crystallization of $(C_6H_4$-CH=CH_2-4$)_3Bi$ (**2**) from iPrOH solution gave pale yellow crystals, which either form needles or rarely a more compact morphology. Both types of crystals of **2** were suitable for single crystal X-ray diffraction analysis and revealed the formation of two polymorphs **2a** (colorless acicular crystals) and **2b** (light yellow block-shaped crystals) in the solid state. Polymorph **2a** crystallizes in the orthorhombic space group $P2_12_12_1$ (Figure 4), while the crystal structure analysis of polymorph **2b** revealed the monoclinic space group $P2_1/c$ (Figure 5).

For **2a** Bi···π arene interactions between the bismuth atom and the aryl ring of the neighboring molecule are deduced, which leads to the formation of zig-zag Bi–$arene_{centroid}$ chains along the crystallographic axis (1D ribbons in Figure 4a). The Bi–$arene_{centroid}$ distance amounts at 3.835 Å (Σ_{vdW} (Bi–C) = 3.77–4.31 Å), which corresponds to the distances of 3.47 to 3.96 Å, as reported for the polymorphs of Ph_3Bi [1,39,41-44]. The overall crystal structure of **2a** is very similar to the mono-

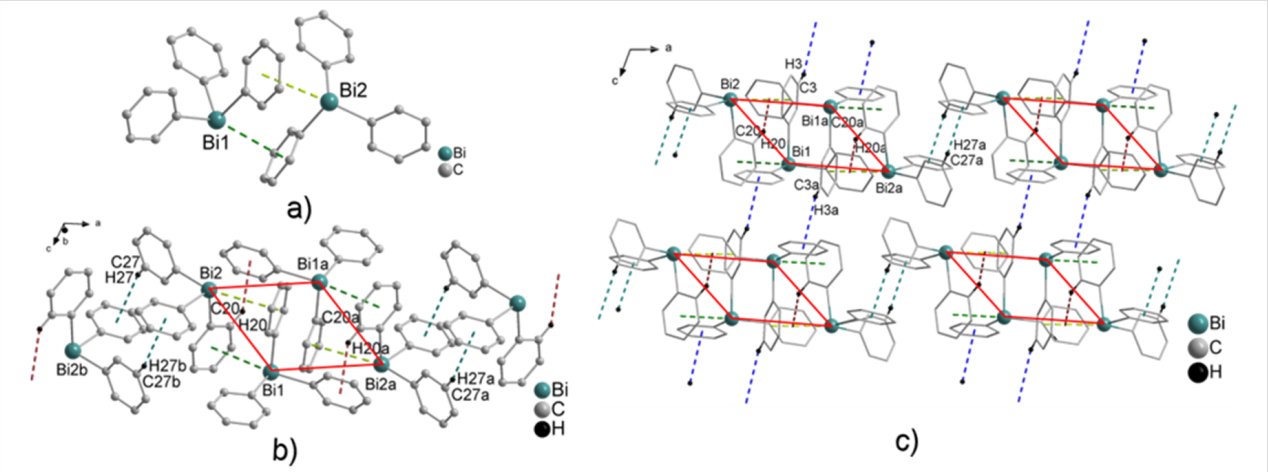


Figure 3: Ball and stick model of Ph_3Bi (**1c**) showing: a) non-centrosymmetric dimers formed via two Bi···π arene intermolecular contacts. Hydrogen atoms were omitted for clarity. Selected distances [Å]: Bi1–$arene_{centroid}$ 3.787 (green dashed line), Bi2–$arene_{centroid}$ 3.939 (lime dash line) [1,44]; b) the formation of a four-membered ring of a 1D layer build via two C–H_{Ph}···π ($arene_{centroid}$) intermolecular contacts with C20–$H20_{Ph}$···π ($arene_{centroid}$) 2.801 Å (dark red dashed line, γ = 12.3°) and C27–$H27_{Ph}$···π ($arene_{centroid}$) 2.763 Å (teal dashed line, γ = 12.6°); c) the formation of a 3D network build via additional C–H_{Ph}···π ($arene_{centroid}$) intermolecular contacts with C3–$H3_{Ph}$···π ($arene_{centroid}$) 3.037 Å (blue dashed line, γ = 7.9°), (only hydrogen atoms involved in C–H_{Ph}···π ($arene_{centroid}$) contacts are shown). Symmetry transformation: a = 2 – *x*, –*y*, 1 – *z*; b = 1 – *x*, –*y*, 1 – *z*.

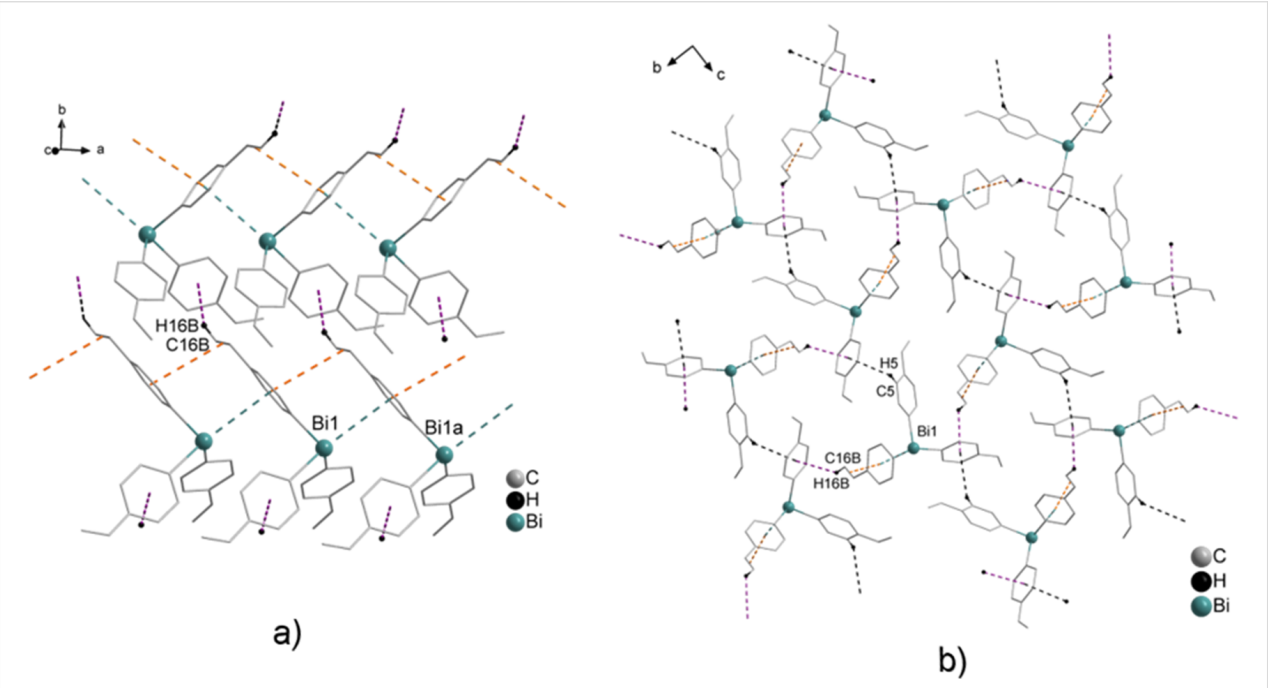


Figure 4: Wire and stick representation of $(C_6H_4$-CH=CH_2-4$)_3$Bi (**2a**) showing: a) zig-zag chains of 1D ribbons formed via Bi···π arene intermolecular contacts (Bi1–$arene_{centroid}$ 3.835 Å), accompanied by π_{Ph}···π_{vinyl} contacts of 3.798 Å and the formation of 2D network via C16B–$H16B_{vinyl}$···π ($arene_{centroid}$) 2.980 Å (violet dashed line, γ = 14.5°); b) the formation of the 3D network via C5–H5···π ($arene_{centroid}$) 3.094 Å (black dashed line, γ = 26.0°, only hydrogen atoms involved in C–H_{Ph}···π ($arene_{centroid}$) contacts are shown). Symmetry transformations: a = 1 + *x*, *y*, *z*; b = –1 + *x*, *y*, *z*.

clinic *C*2/*c* modification of Ph_3Bi (**1a**) [41-43]. Additionally, short π···π distances were observed between one of the vinyl groups and the aryl ligand, with a distance from the centroid of the aromatic ring to the midpoint of the C=C double bond of 3.798 Å. The angle to the plane through the aryl ligand of the neighboring molecule amounts at 13.1° and thus a nearly linear arrangement between a bismuth atom, an aryl ligand, and the vinyl group with an angle of 171.8° is observed (Figure 4a).

In **2a** each Bi···π arene contact is accompanied by a π_{Ph}···π_{vinyl} contact (orange dashed line in Figure 4a). The 1D chains are connected via C–H_{vinyl}···π ($arene_{centroid}$) intermolecular

contacts with C16B–H16B$_{vinyl}$···π (arene$_{centroid}$) 2.980 Å (violet dashed line, γ = 14.5°) to form a 2D network (Figure 4a). Additional C–H$_{Ph}$···π (arene$_{centroid}$) intermolecular contacts with C5–H5···π (arene$_{centroid}$) of 3.094 Å (black dashed line, γ = 26.0°) lead to the formation of a 3D network in the solid state (Figure 4b). By contrast, the crystal structure of **2b** did not show any Bi···π arene interaction, and it reveals only the presence of C–H$_{Ph}$···arene$_{centroid}$ contacts. Two sorts of C–H$_{Ph}$···π (arene$_{centroid}$) intermolecular contacts, with C11–H11$_{Ph}$···π (arene$_{centroid}$) 2.817 Å (cyan dashed line, γ = 8.5°) and C18–H18$_{Ph}$···π (arene$_{centroid}$) 2.940 Å (dark blue dashed line, γ = 10.0°) are observed (Figure 5a). Furthermore, additional C–H$_{vinyl}$···π (arene$_{centroid}$) intermolecular contacts with C16Bb–H16Bb$_{vinyl}$···π (arene$_{centroid}$) 2.960 Å (green dashed line, γ = 25.0°) lead to the formation of a 2D network, while other intermolecular contacts with C24Ac–H24Ac$_{vinyl}$···π (arene$_{centroid}$) of 2.904 Å (brown dashed line, γ = 7.0°) result in the formation of a 3D network (Figure 5b). The Bi···Bi contacts are considerably shorter than the sum of the van der Waals radii of bismuth atoms (Bi···Bi contacts of 4.046 Å; Σr_{vdW}(Bi, Bi) 4.08–5.14 Å)) [60-62] and are in good agreement with the ones reported recently in a theoretical study by Jansen and co-workers who discussed dispersion type Bi···Bi interactions in the context of structure formation in R_3Bi compounds (Bi···Bi contacts vary between 4.015 and 4.059 Å) [63].

$(C_6H_4$-OMe-4$)_3$Bi (**3**)

$(C_6H_4$-OMe-4$)_3$Bi (**3**) crystallized from $CHCl_3$ in the form of colorless cube-shaped or block-shaped crystals which were suitable for single crystal X-ray structure analysis. Compound **3** crystallizes in the trigonal space group $R\bar{3}$ (Figure 6). While our work was in progress, Gagnon and co-workers reported the crystal structure of **3**, which exhibits very similar lattice parameters [64]. However, the packing structure has not been discussed in detail and thus its description is given here.

A closer look at the bismuth environment reveals that for the molecular structure of **3** the bismuth atom might be described as six-coordinated being surrounded by six 4-methoxyphenyl groups. Three of them are bonded covalently to bismuth with Bi–C of 2.248(3) Å and three units are bonded via weaker Bi···O interactions, which have identical values (Bi···O 3.781 Å), finally leading to a [3 + 3] coordination (van der Waals radii Σ$_{vdW}$ (Bi,O) = 3.57–4.09 Å) [60-62]. The 4-methoxyphenyl moieties are pointing via the oxygen atoms to the bismuth atom, which actually hinders the formation of bismuth Bi···π arene interactions. A similar coordination environment was observed in the case of the two polymorphs of $(2\text{-}C_4H_3S)_3Bi$, where three thienyl molecules of the neighboring molecules interact with Ar_3Bi [29]. The oxygen atoms of the methoxy groups each interact with the bismuth atom of a neighboring molecule in a way that each bismuth atom inter-

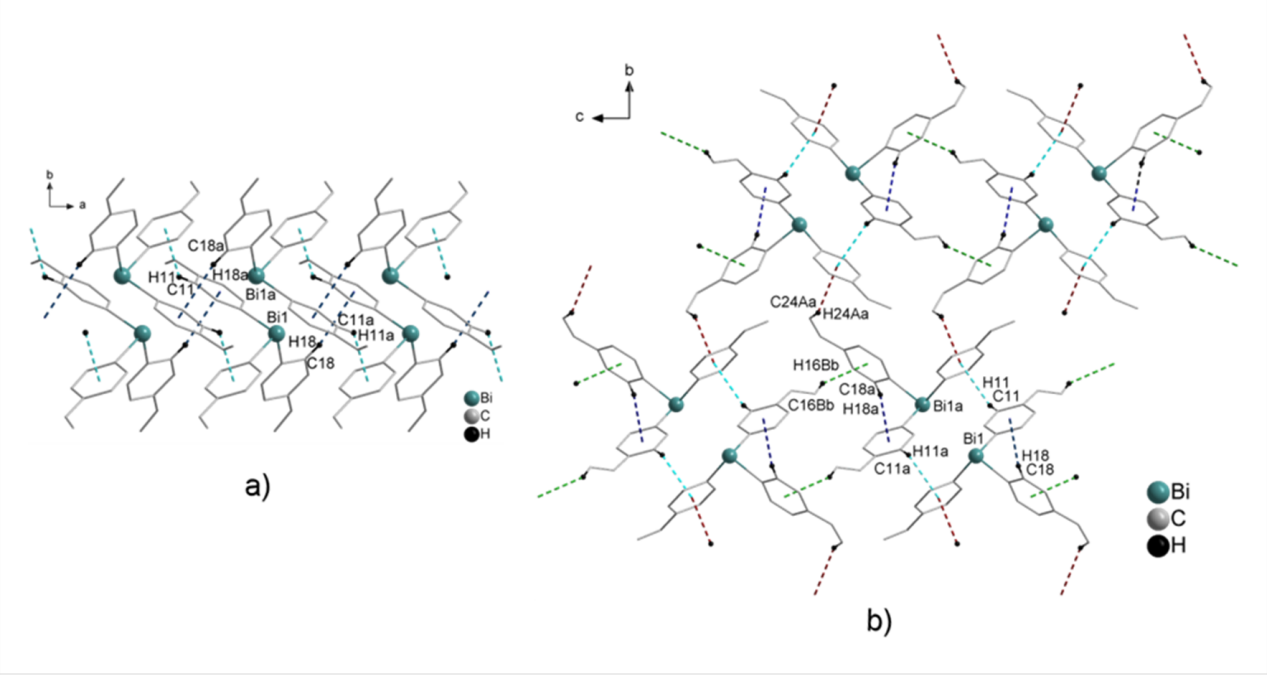


Figure 5: Wire and stick model of $(C_6H_4\text{-}CH{=}CH_2\text{-}4)_3Bi$ (**2b**) showing: a) the formation of 1D ribbons build via two C–H$_{Ph}$···π (arene$_{centroid}$) intermolecular contacts with C11–H11$_{Ph}$···π (arene$_{centroid}$) 2.817 Å (cyan dashed line, γ = 8.5°) and C18–H18$_{Ph}$···π (arene$_{centroid}$) 2.940 Å (dark blue dashed line, γ = 10.0°); b) 2D network formed via C16Bb–H16Bb$_{vinyl}$···π (arene$_{centroid}$) 2.980 Å (γ = 25.0°) and a 3D network formed via C24Ac–H24Ac$_{vinyl}$···π (arene$_{centroid}$) of 2.904 Å (γ = 7.0°), (only hydrogen atoms involved in C–H$_{Ph}$···π (arene$_{centroid}$) contacts are shown). Symmetry transformation: a = 1 – *x*, 1 – *y*, 1 – *z*; b = 1 + *x*, *y*, 1 + *z*; Bi···Bi of 4.046 Å.

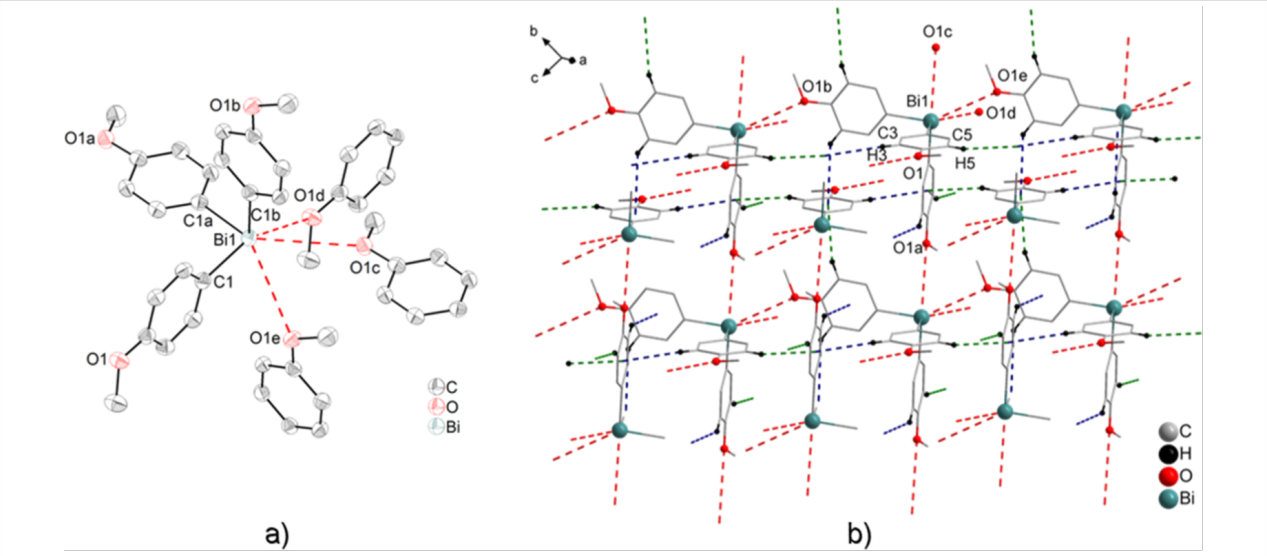


Figure 6: Molecular structure of $(C_6H_4\text{-OMe-}4)_3Bi$ (**3**) showing: a) Thermal ellipsoids that are set at 50% probability level. Hydrogen atoms were omitted for clarity. Symmetry transformations: a = $-x + y + 2, -x + 1, z$; b = $-y + 1, x - y - 1, z$; c = $^7/_3 - x + y, ^5/_3 - x, -^1/_3 + z$; d = $^2/_3 + x, -^1/_3 + y, -^1/_3 + z$; e = $^4/_3 + y, -^4/_3 + x - y, -^1/_3 + z$. Selected bond lengths [Å]: Bi1–C1 2.248(3), Bi1–C1a 2.248(3), Bi1–C1b 2.248(3), Bi1–O1c 3.781, Bi1–O1d 3.781, Bi1–O1e 3.781. Selected bond angles [°]: C1–Bi1–C1a 93.7(11), C1–Bi1–C1b 93.7(11), C1a–Bi1–C1b 93.7(11), O1c–Bi1–O1d 69.6, O1c–Bi1–O1e 69.6, O1d–Bi1–O1e 69.6; b) wire and stick representation of a 3D network built via Bi···O intermolecular interactions and via two C–H_{Phyl}···π ($arene_{centroid}$) intermolecular contacts C3–$H3_{Ph}$···π ($arene_{centroid}$) 3.005 Å (blue dashed line, γ = 14.7°) and C5–$H5_{Ph}$···π ($arene_{centroid}$) 2.845 Å (green dashed line, γ = 11.6°).

acts with three oxygen atoms of different neighbors. In the resulting three-dimensional structure, one molecule of **3** interacts with six other molecules (Figure 6b). Similar Bi···O interactions are also found in tris(2-methoxyphenyl)- and tris(2,4-dimethoxyphenyl)bismuthine [65]. However, the coordination sphere of the bismuth atoms in these compounds is complemented intramolecularly through the methoxy groups in the *ortho* position (Bi···O of 3.05 Å and 3.15 Å) [65].

The crystal structure of **3** did not show any Bi···π arene interaction, but reveals the presence of C–H_{Ph}···$arene_{centroid}$ contacts. These intermolecular C–H_{Ph}···π ($arene_{centroid}$) contacts amount to C3–$H3_{Ph}$···π ($arene_{centroid}$) 3.005 Å (blue dashed line, γ = 14.7°) and C5–$H5_{Ph}$···π ($arene_{centroid}$) 2.845 Å (green dashed line, γ = 11.6°, Figure 6b).

$(C_6H_3\text{-}t\text{-}Bu_2\text{-}3,5)_3Bi$ (**4**) and $(C_6H_3\text{-}t\text{-}Bu_2\text{-}3,5)_2BiCl$ (**5**)

Colorless single crystals suitable for X-ray crystallography were isolated upon crystallization from a CH_2Cl_2 solution at ambient temperature (for **4**) and at −28 °C (for **5**). Compounds **4** and **5**·$2CH_2Cl_2$ crystallize in the hexagonal space group $P6_3$ and orthorhombic space group $Pna2_1$, respectively. The crystal structure of compound **4** does not exhibit any Bi···π arene interactions, but shows four C–$H_{t\text{-}Bu}$···π ($arene_{centroid}$) intermolecular contacts, with C14–H14C1$a_{t\text{-}Bu}$···π ($arene_{centroid}$) 2.877 Å (black dashed line, γ = 10.7°), C27–H27C$b_{t\text{-}Bu}$···π ($arene_{centroid}$) 2.884 Å (blue dashed line, γ = 7.8°), C37–H37C$a_{t\text{-}Bu}$···π ($arene_{centroid}$) 3.012 Å (brown dashed line, γ = 11.5°) and C95–H95C$c_{t\text{-}Bu}$···π ($arene_{centroid}$) 3.002 Å (green dashed line, γ = 6.5°) in Figure 7, respectively. The presence of bismuth···π arene interactions could not be observed most probably due to the bulky *t*-Bu groups attached to the aryl ligands, which hinder the interactions of the bismuth atom with the aryl ligands of the neighboring molecules (Figure 7).

The crystal structure analysis of **5** revealed intermolecular donor acceptor Bi···Cl interactions of Bi1–Cl1b 2.805(7) Å, which are accompanied by Bi···π arene contacts of Bi1–$arene_{centroid}$ 3.725 Å. This arrangement results in a sort of intermolecular pincer-type coordination of the bismuth atom, and thus in the formation of a 1D chain in the solid state (Figure 8). Due to the Bi···π arene interactions, the local geometry of the bismuth atom becomes distorted square pyramidal with one carbon atom of the $(C_6H_3\text{-}t\text{-}Bu_2\text{-}3,5)$ ligand in the axial positions and the two chlorine atoms, another carbon atom of the aryl ligand and the $arene_{centroid}$ placed in the equatorial positions, describing the basal plane. This is reflected in the bond angles of C15–Bi1–$arene_{centroid}$ 89.7°, C15–Bi1–Cl1 90.9(10)°, C15–Bi1–Cl1b 92.7(9)°, and C1–Bi1–C15 92.4(10)°. Besides these contacts, the crystal structure of **5** revealed short C–$H_{t\text{-}Bu}$···π ($arene_{centroid}$) contacts for C9–H9A$_{t\text{-}Bu}$···π ($arene_{centroid}$) 2.662 Å (brown dashed line, γ = 7.4°).

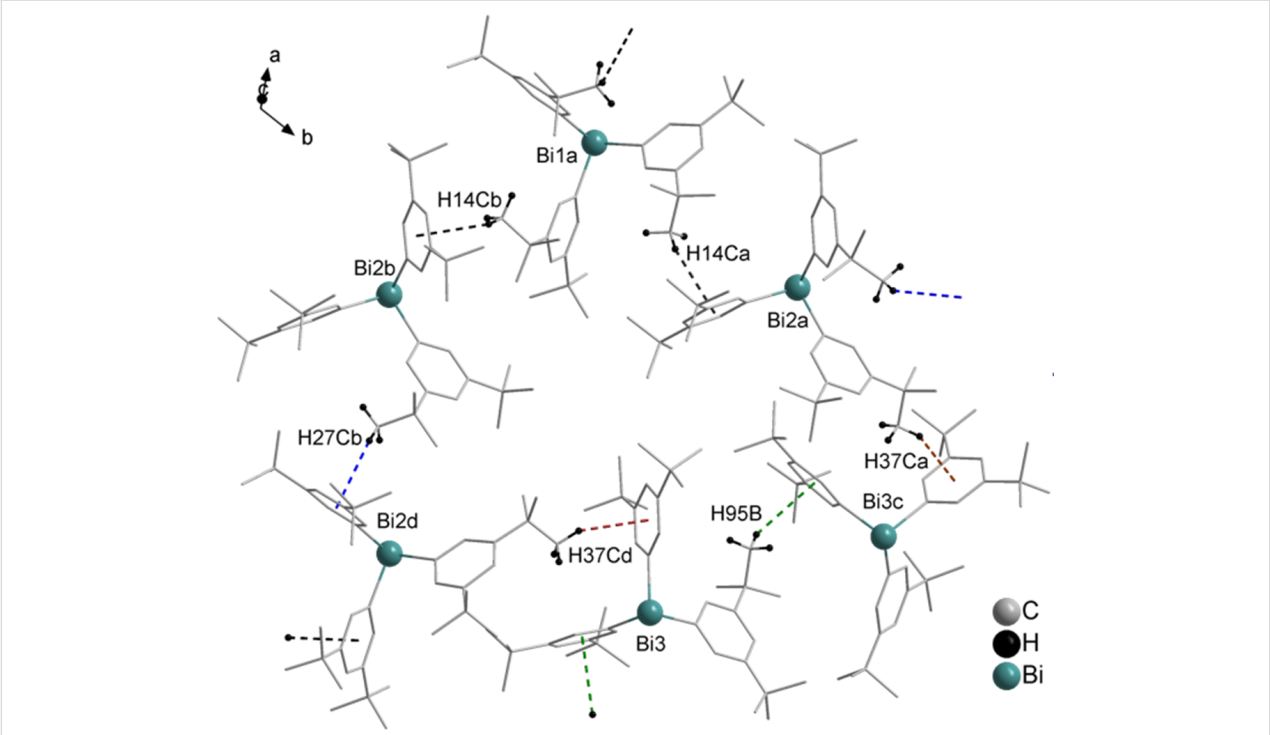


Figure 7: Wire and stick model of $(C_6H_3$-*t*-Bu_2-3,5)$_3$Bi (**4**) showing a 3D network build via four C–H$_{t\text{-}Bu}$···π (arene$_{centroid}$) intermolecular contacts with C14–H14Ca$_{t\text{-}Bu}$···π (arene$_{centroid}$) 2.877 Å (black dashed line, γ = 10.7°), C27–H27Cb$_{t\text{-}Bu}$···π (arene$_{centroid}$) 2.884 Å (blue dashed line, γ = 7.8°), C37–H37Ca$_{t\text{-}Bu}$···π (arene$_{centroid}$) 3.012 Å (brown dashed line, γ = 11.5°) and C95–H95Bc$_{t\text{-}Bu}$···π (arene$_{centroid}$) 3.002 Å (green dashed line, γ = 6.5°), (only hydrogen atoms involved in C–H$_{t\text{-}Bu}$···π (arene$_{centroid}$) contacts are shown). Symmetry transformation: a = *x*, *y*, 1 + *z*; b = 1 – *y*, 1 + *x* – *y*, 1 + *z*; c = 1 + *x* – *y*, 1 + *x*, ½ + *z*; d = –1 + *y*, –*x* + *y*, ½ + *z*.

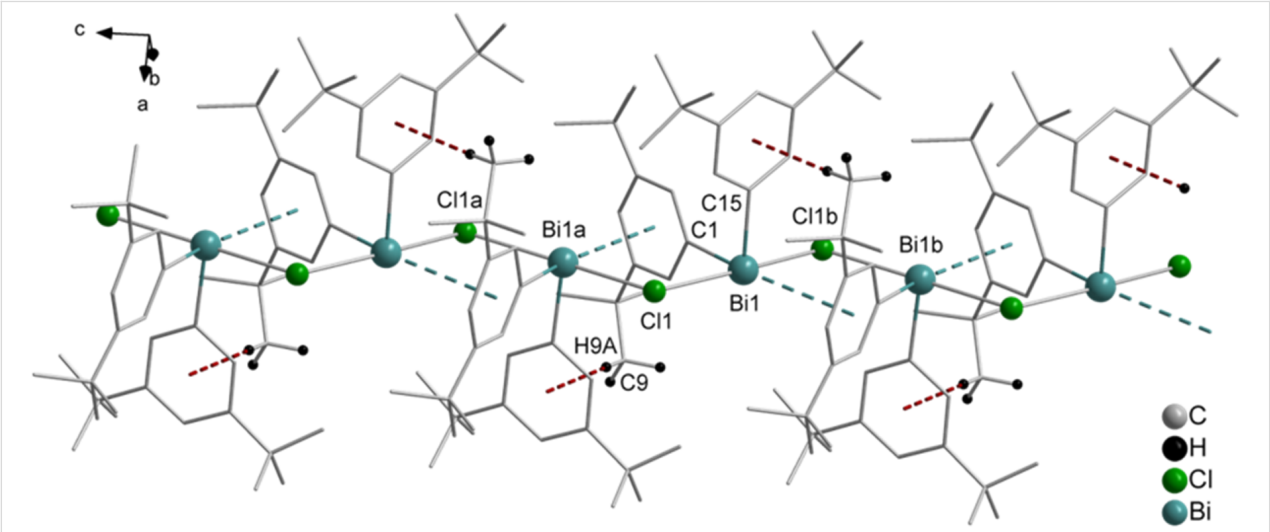


Figure 8: Wire and stick model of $(C_6H_3$-*t*-Bu_2-3,5)$_2$BiCl (**5**) showing a fragment of the 1D arrangement, view along the *b* axis. Symmetry transformations: a = 2 – *x*, –*y*, ½ + *z*; b = 2 – *x*, –*y*, –½ + *z*. Selected bond lengths and distances [Å]: Bi1–Cl1 2.811(7), Bi1–Cl1b 2.805(7), Bi1a–Cl1 2.805(7), Bi1–arene$_{centroid}$ 3.725. Selected bond angles [°]: C15–Bi1–Cl1 90.9(10), C15–Bi1–Cl1b 92.7(9), C(15)–Bi(1)–arene$_{centroid}$ 89.7, Cl1–Bi1–Cl1b 171.5(17), Bi1–Cl1–Bi1a 112.6(19). C–H$_{t\text{-}Bu}$···π (arene$_{centroid}$) intermolecular contacts, C9–H9A$_{t\text{-}Bu}$···π (arene$_{centroid}$) 2.662 Å (brown dashed line), (γ = 7.4°).

As shown in this section, the crystal structures of **1**–**5** described above revealed the presence of London dispersion type interactions in the solid state, with bismuth acting as dispersion energy donor (DED) only in some cases. In the absence of strong donor acceptor type interactions a competition between the different types of dispersion interactions (Bi···π, π···π or C–H···π) is ob-

served and thus leads to different structural features in the solid state (Table 1). Here the question arises how important and how large these interactions are and whether any type of interaction is dominating. For this reason computational studies have been performed with the focus on the crystal structures of three polymorphs of Ph_3Bi (**1**), which revealed either the formation of non-centrosymmetric dimers as basic building block, or the formation of 1D ribbons (i.e., zig-zag type). Both are based on bismuth···π arene interactions: the formation of 2D networks is built up via $C–H_{Ph}$···π intermolecular contacts of T-shape. Noteworthy, polymorph **2a** showed bismuth···π and π···π interactions leading to 1D ribbons in the solid state, while **2b** did not reveal Bi···π interactions. Thus, it is concluded that Bi···π, π···π and C–H···π interactions must be of similar strength. Similar to the situation of **2b**, compound **4** did not show any bismuth···π arene interactions, but also exhibits $C–H_{t\text{-}Bu}$···π intermolecular contacts. In compounds **3** and **5** intermolecular Bi···O and Bi···Cl bonds are dominating.

Electronic structure calculations on selected polymorphs of Ph_3Bi

In order to assess the role of dispersion interactions for the existence of structural features in compounds including Bi···π interactions, we focus our study on the wealth of structural information for $BiPh_3$ (compound **1**). Note that various structural motifs present in the polymorphs of compound **1** can be also found in polymorphs of compound **2**.

We will proceed as follows: First, an idealized model compound is studied such that the basic Bi···π interaction can be classified in comparison to other systems previously studied. Then, we will turn to the crystal structures of polymorph **1a**, **1b**, and **1c** and investigate each polymorph in terms of intermolecular interactions to assess which influences are decisive for structure formation. For this purpose, several tetrameric units have been extracted from the crystal structure for each polymorph. This way, all relevant intermolecular interactions in the solid state can be studied based on monomer distortion, intermolecular interactions of representative dimers and the interaction of one molecule with several of its neighbors.

Distance scan for the idealized $BiPh_3$···benzene complex

In a previous study on the nature of Bi···π arene interactions in various benzene complexes with BiR_3 (with R = Me, OMe, and Cl), we found that interaction energies for this type of compounds range from −10 kJ mol^{-1} to −40 kJ mol^{-1}. The character of the interaction varies from purely dispersive for $BiMe_3$ to dispersive with pronounced donor–acceptor character in case of $Bi(OMe)_3$ and $BiCl_3$ complexes [66,67]. In order to assess

Table 1: Various intermolecular distances observed in the crystal structures of arylbismuth compounds.

	Bi···π intermolecular distances	π···π intermolecular distances	C–H···π intermolecular distances	Bi···O/Cl intermolecular distances	structural features
Ph_3Bi (**1**)					
polymorph **1a**	3.763 Å		3.030 Å 3.042 Å 2.760 Å 2.740 Å		3D network
polymorph **1b**	3.468 Å 3.561 Å		3.083 Å 3.097 Å 3.034 Å 2.890 Å		2D network
polymorph **1c**	3.787 Å 3.939 Å		2.801 Å 2.763 Å 3.037 Å		3D network
$(C_6H_4\text{-}CH{=}CH_2\text{-}4)_3Bi$ (**2**)					
polymorph **2a**	3.835 Å	3.798 Å			1D ribbons
polymorph **2b**			2.817 Å 2.940 Å		1D ribbons
$(C_6H_4\text{-}OMe\text{-}4)_3Bi$ (**3**)			3.005 Å 2.845 Å	3.781 Å	3D network
$(C_6H_3\text{-}t\text{-}Bu_2\text{-}3,5)_3Bi$ (**4**)			2.877 Å 2.884 Å 3.012 Å 3.002 Å		3D network
$(C_6H_3\text{-}t\text{-}Bu_2\text{-}3,5)_2BiCl$ (**5**)	3.725 Å		2.662 Å	2.811 Å 2.805 Å	1D ribbons

the nature of the interaction in the $BiPh_3$···π complexes (compound **1**), rigid potential energy surface scans for the idealized $BiPh_3$–benzene complex were performed at the PBE-D3 and DLPNO-CCSD(T) level of theory. The idealized structure was constructed in order to disentangle pure Bi···π arene interaction from the influence of substituents. The interaction potential curve is shown in Figure 9. The distance scans obtained at the PBE-D3 level of theory are in good agreement with the DLPNO-CCSD(T) results which shows that using the PBE-D3 functional is a cost efficient alternative to the DLPNO-CCSD(T) method. The minimum on the DLPNO-CCSD(T) potential energy curve estimated by interpolation corresponds to a distance of 3.66 Å and to −17 kJ mol^{-1}. Note that the Bi···π arene contact minimum distance is shorter than for **1a**, **1c**, and **2** but slightly longer than for **1b** (see Table 1). The curve for the interaction energy without dispersion contribution (Figure 9, *E*(int-disp)) is slightly attractive.

The interaction energy of the $BiPh_3$ complex is higher than the interaction energy obtained for $BiMe_3$ but smaller than for $Bi(OMe)_3$ (see Figure 10a), however, the dispersion contribution to the interaction energy in the $BiPh_3$ complex (see Figure 10b) is comparable to the dispersion contributions in other BiR_3–benzene complexes. This implies that the character

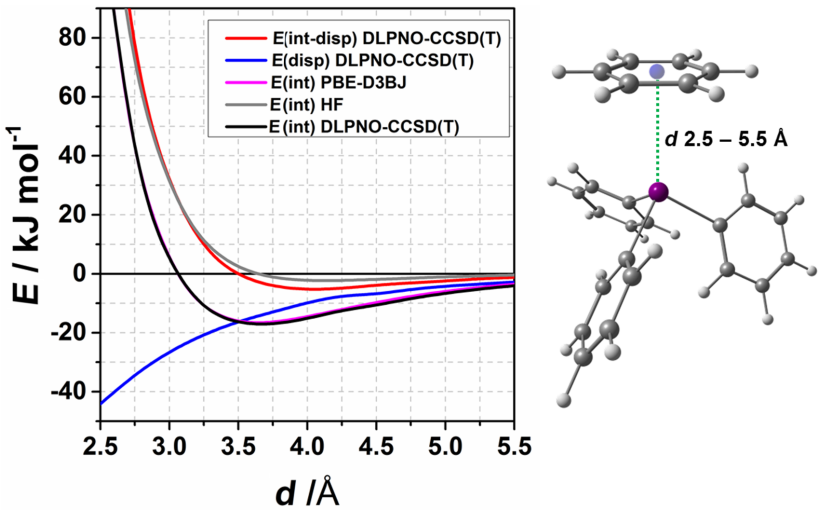


Figure 9: Computed interaction potentials of the distance scan for the idealized $BiPh_3$–benzene complex. *E*(int) denotes the interaction energy obtained at the particular level of theory, *E*(disp) denotes the dispersion energy, *E*(int-disp) – the interaction energy without the dispersion contribution.

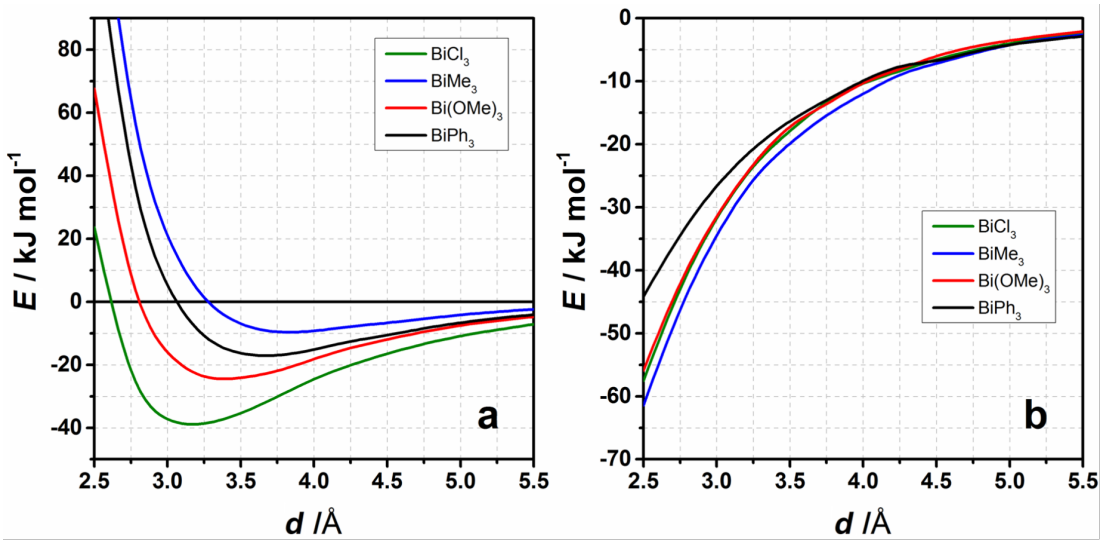


Figure 10: a) The $BiPh_3$ potential energy curve for idealized interaction structures compared to the interaction energy curves for the series of compounds BiR_3···C_6H_6 with R = Me, OMe, Cl obtained at the DLPNO-CCSD(T) level of theory. b) Dispersion energy contributions according to LED (DLPNO-CCSD(T)) for the distance scans shown in Figure 9.

of the interaction in the $BiPh_3$–benzene complex is closer to that of $BiMe_3$ rather than to $Bi(OMe)_3$ and that the interaction is dominated by dispersion with minor contribution of donor–acceptor character.

Computational description of the polymorphs of $BiPh_3$

In this section we address the question which factors influence the structure in solid state for $BiPh_3$ and what possible intermolecular interactions within the polymorphs of compound **1** are. As mentioned already earlier several tetrameric units of polymorphs of compound **1** were chosen in order to obtain a simplified description of the crystal structure. These tetramers contain the information on the different intermolecular interactions present in the solid state of $BiPh_3$. The structures of the studied tetramers are shown in Figures 11, 13, and 15. Subsequently, tetramers were divided into dimeric units that represent specific intermonomer interactions, and can be considered as building blocks of the bulk. Monomer distortion energies (geometry preparation; denoted as E_{prep} in Figures 11, 13, and 15) were computed in order to gain knowledge on crystal packing effects. The distortion energy is obtained as the difference between the energy of a single relaxed molecule and the energy of an unrelaxed molecule in the crystal structure geometry. Important information on the intermolecular interaction strength within the tetramer is obtained when one monomer is removed from the system. This energy (depicted as E_{remove} in Figures 11, 13, and 15) contains the interactions with neighboring molecules and also possible long-range interactions within the tetramer. This energy is determined as a difference between the interaction energies of a tetramer and trimer formed after removing the appropriate monomer.

The interaction energies of all tetramers and dimers were computed at the PBE-D3/def2-TZVP level of theory and are depicted in Figures 11, 13, and 15 as $E_{tetramer}$ and E_{dim}, respectively. A color code is introduced in Figures 11, 13, and 15 to facilitate the understanding of the construction of tetramers and dimers. A different color is ascribed to each monomer within the tetramer. The interaction energies were computed with reference to the sum of the energies of all unrelaxed monomers (crystal structure geometry) included in the tetramer or dimer. Interaction energies of tetrameric, trimeric, and dimeric structures can yield information about the additivity of intermolecular interactions. Additionally, interaction energies of all Bi···π arene type dimers and selected π-stacking dimers were computed at the DLPNO-CCSD(T)/cc-pVQZ (cc-pwCVQZ-PP for Bi) level of theory with TightPNO settings (see Figures 12, 14, and 16). Local energy decomposition analysis was performed in order to obtain the dispersion energy contributions to the interaction energies. The dispersion energies of the specific dimers were then visualized as DED plots and are shown in Figures 12, 14, and 16. The structures and interaction energies of all studied π-stacking dimers are given in Supporting Information File 1 (see Figures S12–S14 and Table S4 in Supporting Information File 1). Please note that the positions of the hydrogen atoms in the tetrameric and dimeric structures were optimized at the PBE-D3/def2-TZVP level of theory. Hence, the intermolecular distances involving C–H groups may vary from the crystallographic data given in the previous sections.

Polymorph **1a**

In case of polymorph **1a** three different tetramers were chosen that are shown in Figure 11. The simplest tetramer **1a-1** consists of linear chains of $BiPh_3$ molecules belonging to one layer. It is built from three equivalent Bi···phenyl dimers with an interaction energy of −46 kJ mol^{-1} (computed at the PBE-D3/def2-TZVP level of theory, depicted in Figure 11 as E_{dim3_2}). Tetramers **1a-2** and **1a-3** are constructed from two Bi···π arene dimers of two different layers of $BiPh_3$ molecules. Within these two tetramers not only Bi···π arene interactions are present but also π-stacking contacts of monomers between two layers. In tetramer **1a-2** three C–H_{Ph}···π stacking interactions can be found. Two equal interactions between monomers 2 (green; numbering of the monomers is the same as the numbering of the Bi atoms within a specific tetramer as depicted in Figure 11) and 3 (red), and monomers 1 (grey) and 4 (blue). There is also a C–H_{Ph}···π type interaction between monomers 2 (green) and 4 (blue). All of these dimers have interaction energies of −29 kJ mol^{-1} (depicted in Figure 11 as E_{dim2_3} and E_{dim2_4}) indicating that they are interacting fairly strongly. Similarly, in tetramer **1a-3** three π-stacking interactions between monomers can be found. Two of them (monomers 1 and 4, and 2 and 3) are equivalent and their interaction energy amounts to −42 kJ mol^{-1} indicating strong interactions between the layers. The next neighbor interaction between monomers 2 (green) and 4 (blue) in tetramer **1a-3** is much weaker and amounts to −15 kJ mol^{-1}.

The distortion energies (E_{prep}) of the monomers in polymorph **1a** are quite large and amount to almost 17 kJ mol^{-1}. This indicates that, although the interactions between specific monomers to form dimeric structures by Bi···π arene are strong, the crystal packing effects are significant in this case.

Another factor that is useful in describing the energetics within the tetramers is the energy required to remove one of the monomers from the corresponding tetramer. This quantity is depicted in Figure 11 as E_{remove}. By removing one of the molecules from the tetramer the interactions with this specific monomer are broken, which can involve contacts between neighboring molecules or long-range interactions. In tetramer **1a-1** the energy needed to remove one of the outer monomers (3 and 4) is

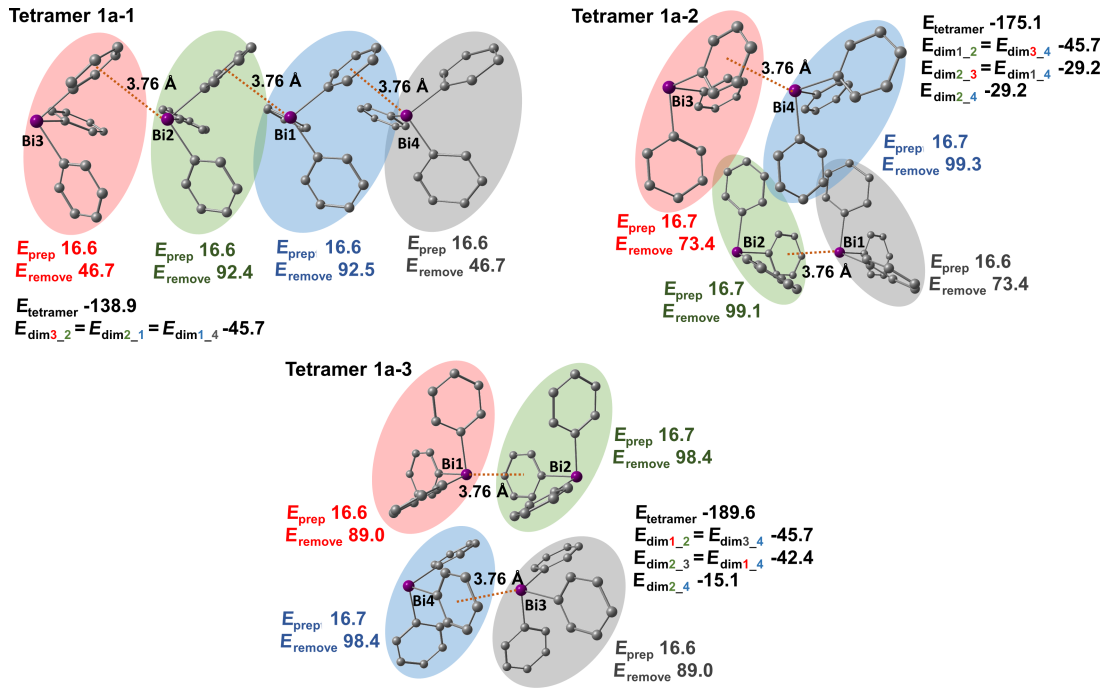


Figure 11: Structures of studied $BiPh_3$ tetramers extracted from the crystal structure of polymorph **1a**. $E_{tetramer}$ indicates the interaction energy computed for a given tetramer; E_{dim} is the interaction energy computed for a specific dimer; E_{prep} is the distortion energy of a monomer in the crystal structure; E_{remove} denotes the energy required to remove a monomer from the tetramer. All energies were computed at the PBE-D3/def2-TZVP and are given in kJ mol^{-1}. Numbering of monomers is the same as numbering of bismuth atoms within the specific tetramer. Hydrogen atoms in tetramers were omitted for the clarity.

almost equal to the interaction energy of one Bi···π arene dimer. This indicates that only one dimer breaks. To remove one of the inner monomers (2 or 1) an energy of 92 kJ mol^{-1} is required. This energy is very close to the sum of the interaction energies of two dimers. For example, E_{remove} of monomer 2 in tetramer **1a-1** is roughly the sum of the interaction energies of two Bi···π arene dimers (91.4 kJ mol^{-1}). This simple example shows the additivity of the dimeric intermolecular interactions within the tetramer. The situation is more complicated for tetramers **1a-2** and **1a-3**. In general, the energies to remove one of the molecules from these tetramers are higher than for tetramer **1a-1** as they contain interactions between the chains and each of the monomers has more contacts on average within the tetramer. The energy needed to remove one of the monomers from tetramer **1a-2** or **1a-3** is roughly the sum of the interaction energies involving this monomer (deviating by at most 5 kJ mol^{-1}). Another important aspect is that the interaction energy computed for tetramers **1a-1**–**1a-3** can be also expressed as a sum of energies of particular dimeric interactions present in the specific tetramer. For instance, the sum of the interaction energies of specific dimers in tetramer **1a-2** amounts to −179 kJ mol^{-1} which is higher by 4 kJ mol^{-1} than the interaction energy of the whole tetramer (−175.1 kJ mol^{-1}). For tetramer **1a-3** this sum is −191.3 kJ mol^{-1} which is very similar to the computed value (−189.6 kJ mol^{-1}). This shows that the interactions in the crystal structure of polymorph **1a** are pairwise neighbor interactions. Long-range interactions are probably mostly weak dipolar interactions that do not contribute significantly. The analysis of specific dimeric interactions in tetramers **1a** shows that not only Bi···π arene interactions are important structure building factors but also several C–H_{Ph}···π arene contacts play a crucial role in structure formation. The strength of the π-stacking interactions depends on the number of contacts and distances between interacting molecules.

Figure 12 depicts structures of Bi···π arene dimers and one of the (strongest) C–H_{Ph}···π arene dimers in polymorph **1a**. The C–H_{Ph}···π arene dimer (depicted as **1a-3-1** in Figure 12) has two very short (2.6 Å) contacts between a C–H group and a phenyl ring. The interaction and dispersion energies given in Figure 12 were obtained at the DLPNO-CCSD(T)/cc-pVQZ (cc-pwCVQZ-PP for Bi and TightPNO settings) level of theory. Note that PBE-D3 and DLPNO-CCSD(T) give very similar results. Inspection of dispersion energies obtained from the LED analysis reveals that both types of dimers (Bi···π arene and π-stacking) are exclusively dispersive. Figure 12 depicts two

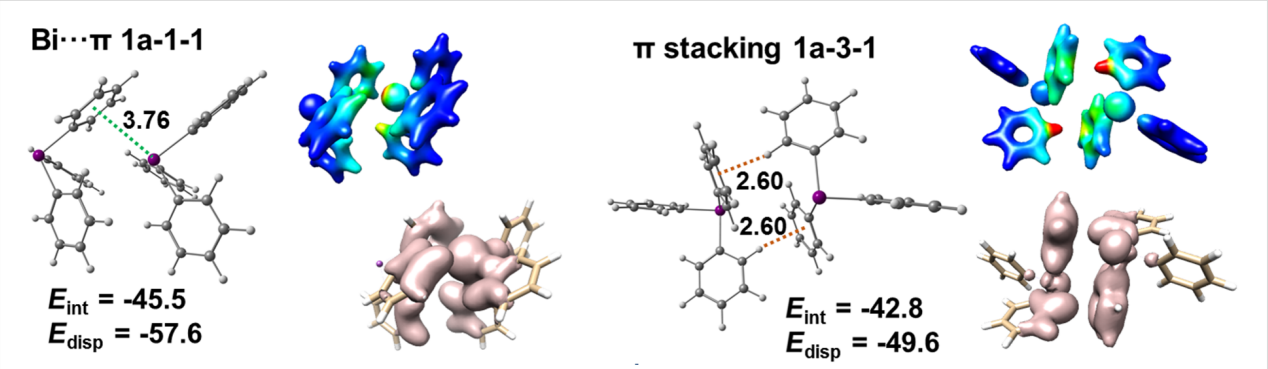


Figure 12: Detailed structures of selected dimers extracted from the crystal structure of polymorph **1a** and dispersion energy plots. E_{int} denotes the interaction energy of a given dimer and E_{disp} is the dispersion energy obtained from LED analysis computed at the DLPNO-CCSD(T)/cc-pVQZ (cc-pwCVQZ-PP for Bi, TightPNO settings) level of theory. The distances are given in Å and energies are in kJ mol^{-1}.

complementary graphical interpretations of the dispersion energy density that can be plotted either as an isosurface (coral plots) or mapped on an isodensity surface (color gradient) of the electron density. Both plots display the regions with the highest contributions to the dispersion interaction present in the complex.

In summary, polymorph **1a** is formed by one dimensional chains consisting of strong Bi···π arene contacts with interaction energies of −46 kJ mol^{-1} (see tetramer **1a-1** in Figure 11). These 1D chains are bound strongly by the π-type interactions present between $BiPh_3$ molecules belonging to different layers. The energies of such contacts range from −15 to −40 kJ mol^{-1}. The distortion energies of monomers (17 kJ mol^{-1}) suggest that crystal packing effects are large in polymorph **1a**.

Polymorph 1b

In case of polymorph **1b** four tetrameric units were identified and are depicted in Figure 13. Tetramers **1b-1** and **1b-2** consist of zig-zag chains formed by Bi···π arene interactions between

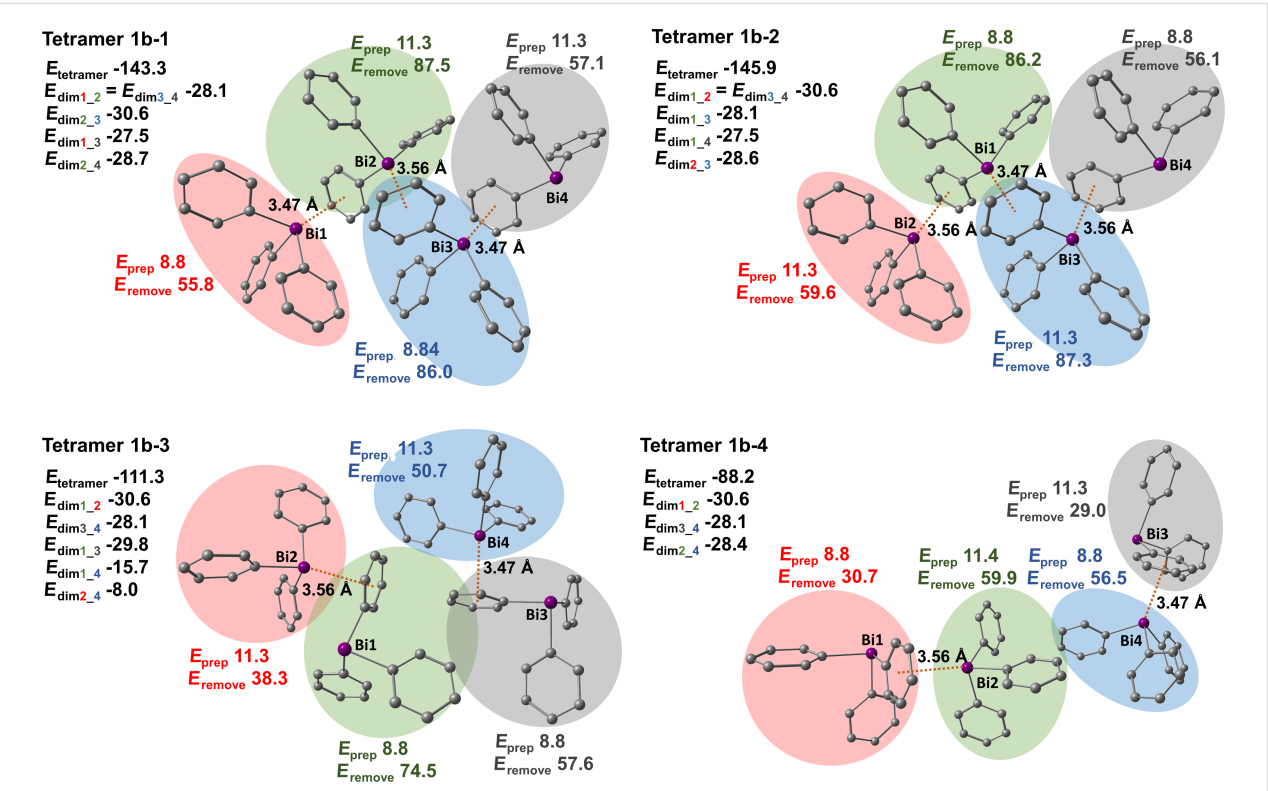


Figure 13: Structures of studied $BiPh_3$ tetramers extracted from the crystal structure of polymorph **1b**. See Figure 11 for details.

one layer of $BiPh_3$ molecules. The difference between these tetramers is that tetramer **1b-1** contains two Bi···π arene contacts with a distance of 3.47 Å between the Bi atom and the phenyl ring centroid and one Bi···π arene contact with a distance of 3.56 Å. In tetramer **1b-2** there are two 3.56 Å Bi···π arene contacts and one 3.47 Å contact. Both types of Bi···π arene dimers have very similar interaction energies of −28 kJ mol^{-1} (E_{dimer1_2} of tetramer **1b-1**) and −31 kJ mol^{-1} (E_{dimer2_3} of tetramer **1b-1**). In tetramers **1b-1** and **1b-2** π-stacking interactions are also present. We only discuss the interactions in tetramer **1b-1** as they are the same as in tetramer **1b-2**. For example, the C–H_{Ph}···π arene-type interactions between monomers 1 (red) and 3 (blue), and monomers 2 (green) and 4 (grey) are as strong as the Bi···π arene interactions in polymorph **1b** and amount to −28 kJ mol^{-1} and −29 kJ mol^{-1}, respectively. Tetramers **1b-3** and **1b-4** on the other hand, exhibit interactions between neighboring zig-zag chains. In tetramer **1b-3** two Bi···π arene interactions and three other stacking interactions are present that vary in energy and contact area between the $BiPh_3$ molecules. For example, the strongest C–H_{Ph}···π arene dimer of tetramer **1b-3** with an interaction energy of −30 kJ mol^{-1} is formed between monomers 1 (green) and 3 (grey). The other two stacking interactions between monomers 1 (green) and 4 (blue), and monomers 2 (red) and 4 (blue) are much weaker due to the smaller contact area between the molecules. Their interaction energies amount to −16 kJ mol^{-1} and −8 kJ mol^{-1}, respectively. In tetramer **1b-4** only one additional π-stacking interaction between molecules 2 (green) and 4 (blue) is present but is quite strong (E_{dim2_4} = −28 kJ mol^{-1}).

The distortion energies of the monomers (E_{prep}) are much lower than that for polymorph **1a** and range from 9 to 11 kJ mol^{-1}. This implies that crystal packing effects are less pronounced in this polymorph.

Energies required to remove one of the monomers from the tetramer in case of polymorph **1b** also depend on the number of interactions and the energy value is roughly the sum of the energies of these interactions. For instance, one of the lowest energies needed to remove a molecule is found for monomer 2 (red) in tetramer **1b-3** that amounts to 38 kJ mol^{-1}. This energy is simply the sum of the already discussed dimer energies (E_{dim1_2} and E_{dim2_4}). Another example is when monomer 3 (blue) is removed from tetramer **1b-1** which requires an energy of 86 kJ mol^{-1}. This energy is a sum of the dimer energies that are formed including this monomer (E_{dim1_3}, E_{dim2_3}, and E_{dim3_4}). This again demonstrates the additivity of intermolecular interactions and that the intermolecular interactions in the bulk can be described as a sum of dimer interactions. Adding up energies of all the dimeric units of tetramer **1b-3** results in an energy of −112.2 kJ mol^{-1} which is roughly equal to the computed interaction energy of this tetramer (−111.3 kJ mol^{-1}).

Figure 14 depicts Bi···π arene dimers and selected π-stacking dimers that are present in the structures of tetramers **1b-1**–**1b-4** and their dispersion energy density plots. It is concluded that the dispersion energies are a few kJ mol^{-1} higher for Bi···π arene dimers than for π-stacking dimers but in general all of these interactions are purely dispersive. Note that typically the dispersion contribution is larger than the overall interaction energy as it compensates the monomer preparation.

Summarizing, polymorph **1b** is built from zig-zag chains of $BiPh_3$ molecules consisting of Bi···π arene contacts with interaction energies of about −30 kJ mol^{-1}. The preparation energy is notably smaller than in polymorph **1a**. The contacts between zig-zag chains are of π-stacking type and their interaction energies amount also to about −30 kJ mol^{-1}.

Polymorph **1c**

For polymorph **1c** we identified three tetramers that are depicted in Figure 15. Each of them contains two Bi···π dimers that consist of two Bi···π arene contacts. The interaction energy of these dimers is high and amounts to −47 kJ mol^{-1}. Tetramer **1c-1** represents Bi···π arene dimers belonging to the same layer. There are two identical C–H_{Ph}···π arene dimers in tetramer **1c-1** with an interaction energy of −38 kJ mol^{-1}. These π-stacking dimers are formed by monomers 2 (green) and 4 (grey), and monomers 1 (red) and 3 (blue). The detailed structure of this dimer is shown in Figure 16 (depicted as **1c-1-1**). Monomers 2 (green) and 3 (blue) also interact to form a dimer that is based on CH···CH and CH···phenyl interactions (see Figure 16, **1c-1-2**). The interaction energy of this dimeric unit is moderate and amounts to −20 kJ mol^{-1}. Tetramers **1c-2** and **1c-3** include the Bi···π arene dimers from two different molecular layers. In tetramer **1c-2** one very strongly bound π-stacking dimer is formed between monomers 1 (green) and 3 (blue). Its interaction energy amounts to −64 kJ mol^{-1} and is the highest among all studied dimers, including Bi···π arene dimers. Its structure resembles the structure of a cube with multiple short C–H_{Ph}···π arene contacts (see Figure 16, dimer **1c-2-1**). In case of tetramer **1c-3** the π-stacking interaction between the molecules are weaker than in the other two tetramers. The interaction between monomers 1 (green) and 3 (blue) is of moderate strength and the interaction energy is −20 kJ mol^{-1}. The second π-stacking interaction present in the tetramer involves monomers 2 (red) and 3 (blue) and is a rather weak interaction with −11 kJ mol^{-1}.

The distortion energies (E_{prep}) of the monomers of polymorph **1c** are very low and range from 5 to 7 kJ mol^{-1}. This means that

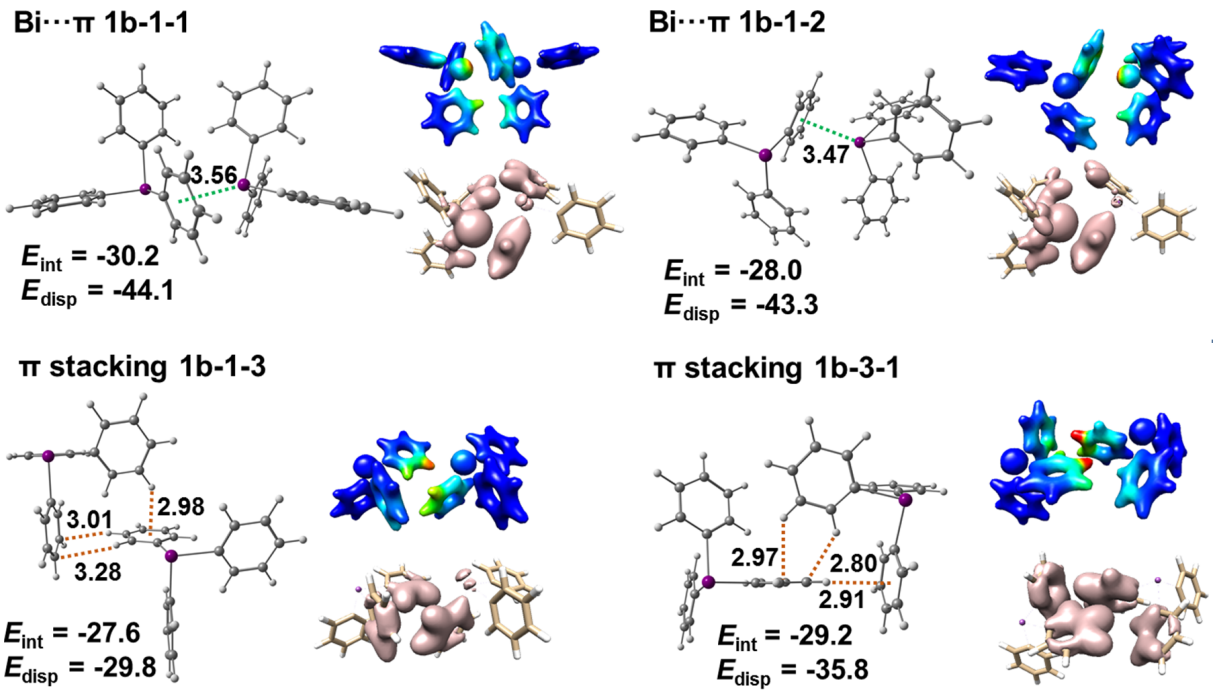


Figure 14: Detailed structures of selected dimers extracted from the crystal structure of polymorph **1b** and dispersion energy plots. See Figure 12 for details.

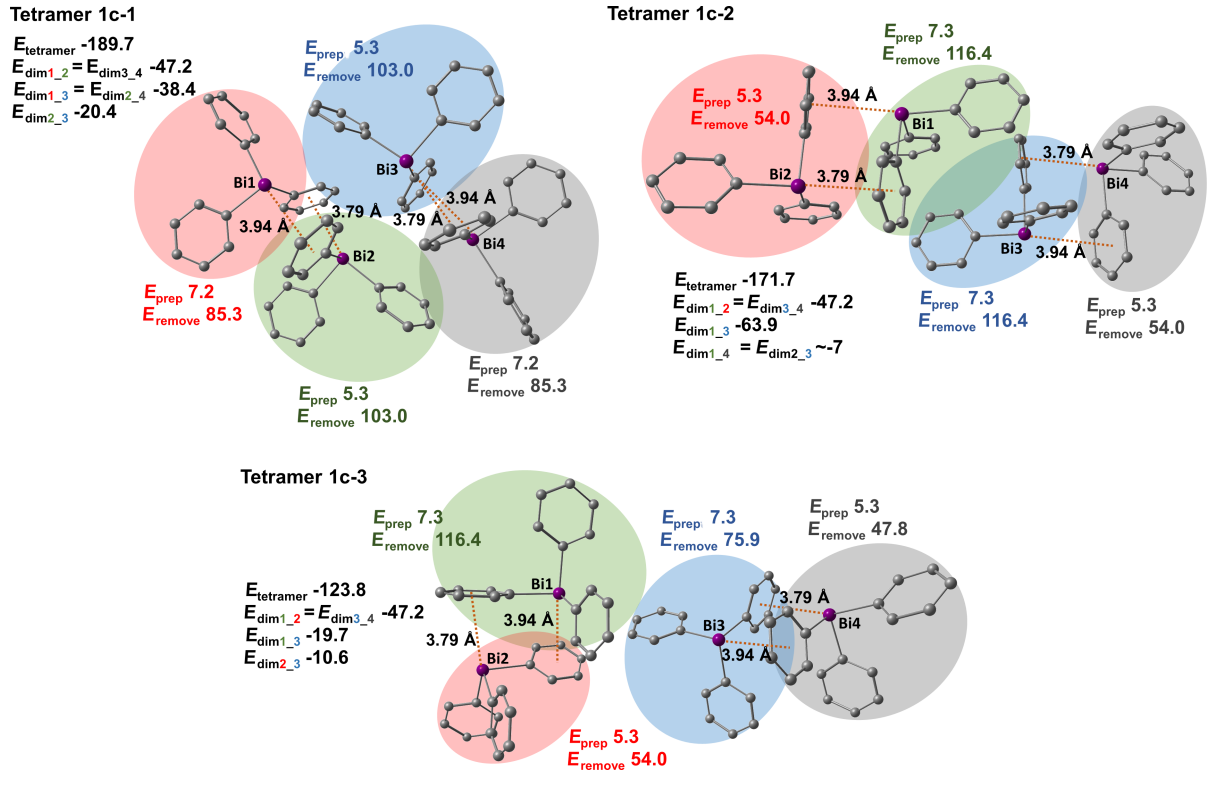


Figure 15: Structures of studied $BiPh_3$ tetramers extracted from the crystal structure of polymorph **1c**. See Figure 11 for details.

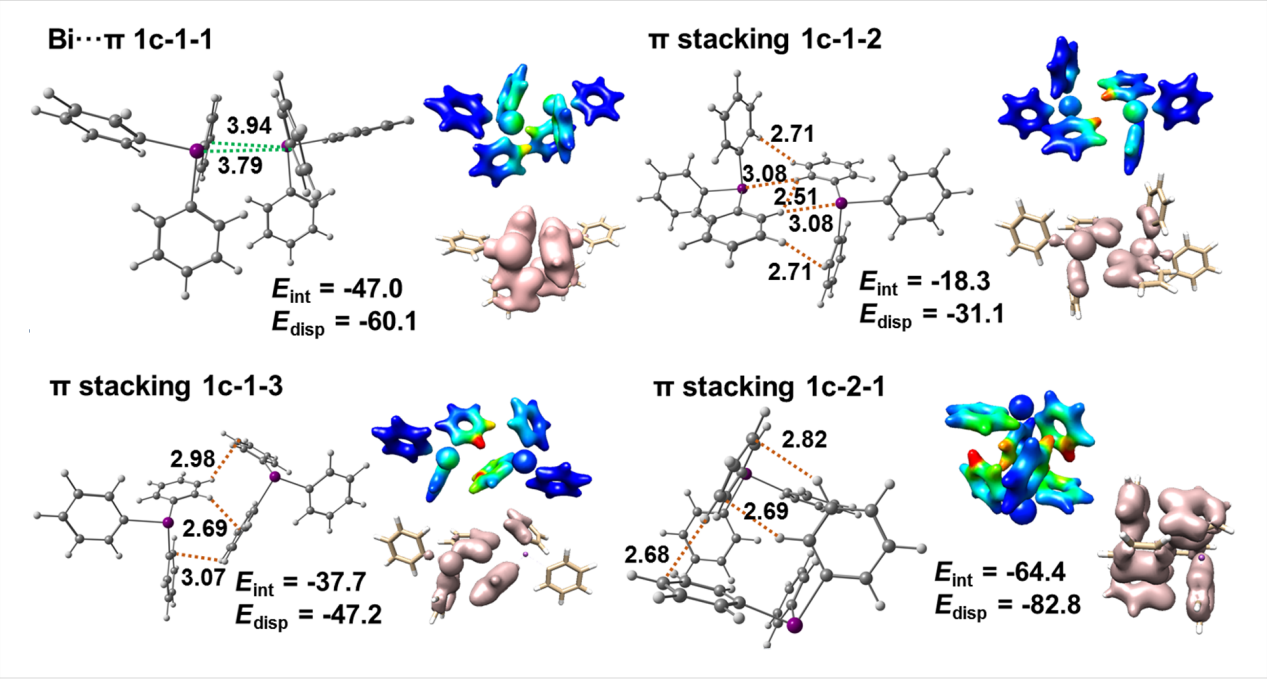


Figure 16: Detailed structures of selected dimers extracted from the crystal structure of the polymorph **1c** and dispersion energy plots. See Figure 12 for details.

in this case packing effects are not very strong compared to the other two polymorphs of $BiPh_3$.

In case of polymorph **1c** energies required to remove one molecule from the tetramer are significantly higher and are roughly the sum of the interactions that involve a specific monomer. Note that the sum of interaction energies of dimers in tetramer **1c-2** amounts to −158.3 kJ mol^{-1} which is much smaller than the interaction energy of the tetramer ($E_{tetramer}$ = −171.7 kJ mol^{-1}). Most probably there are two possible interactions between slightly remote monomers 1 (green) and 4 (grey), and 2 (red) and 3 (blue) each accounting for about −7 kJ mol^{-1}. In case of the other two tetramers of polymorph **1c** all interaction energies of dimers add up to roughly the interaction energy of the tetramer.

Figure 16 depicts the most important dimers found in the structure of polymorph **1c**, for which the interaction and dispersion energies are given. The dispersion energy plots show the spatial distribution of the dispersion interaction within each dimer. An especially high dispersion energy contribution is observed for dimer **1c-2-1** (−83 kJ mol^{-1}). By looking at the distribution of the dispersion energy (dispersion energy density plots) for this dimer it is noticed that almost the entire monomers contribute to the overall dispersion from π···π interactions.

In case of polymorph **1c** the results of the quantification of the interaction energies reveal that this structure is actually not dominated by Bi···π arene interaction, but rather consists of dimers connected by strong π···π interaction (Figure 16, **1c-2-1** with −64 kJ mol^{-1}) which are connected by two weaker Bi···π arene contacts (Figure 16, **1c-1-1**, −47 kJ mol^{-1}).

Figure 17 shows the comparison of the distortion and the interaction energies of Bi···π arene and π stacking dimers. As for $BiPh_3$, Bi···π arene and π···π interactions are of comparable magnitude and the existence of the different polymorphs can be explained by a balance of two competing interactions. While in case of polymorph **1a** the Bi···π arene interaction dominates as a structure building factor, for polymorph **1b** Bi···π arene and π···π interactions are in the same energy range. In case of polymorph **1c** the π-stacking interaction dominates. Figure 17 also demonstrates the differences between distortion energies of monomers (E_{prep}) as found for a specific polymorph and shows how packing effects decrease from polymorph **1a** to **1c**.

Conclusion

Herein, we have shown that the dispersion type Bi···π arene interactions provide an important contribution to the structure formation of arylbismuth compounds. In the absence of stronger donors such as -OR and -Cl, the dispersion type bismuth···π arene interaction is supplemented by other weak interactions such as π···π or C–H_{Ph}···π. Each Bi···π arene contact with bismuth as a strong dispersion energy donor (DED) provides a higher interaction energy than a single C–H_{Ph}···π contact, but most often several of the latter compete with the single Bi···π

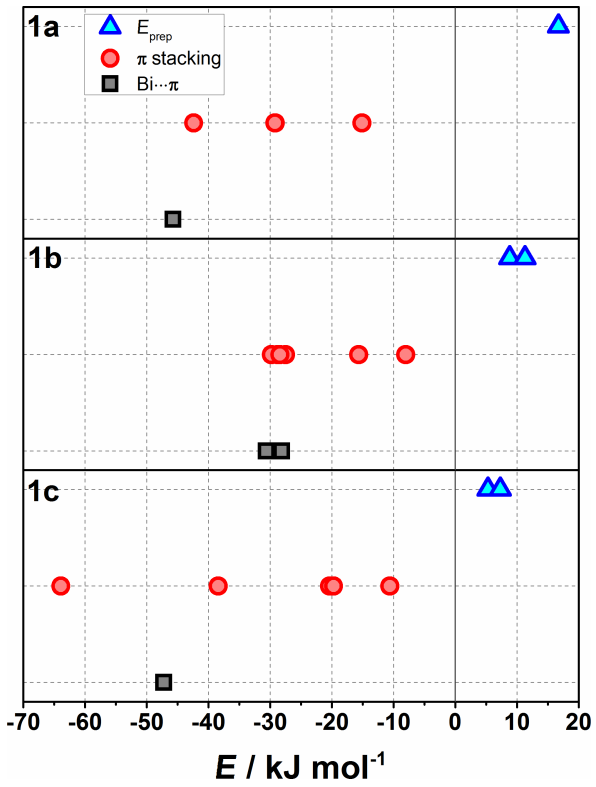


Figure 17: Distortion energies of monomers (E_{prep}, blue triangles) and interaction energies of Bi···π and π-stacking dimers (grey squares and red dots, accordingly) computed at the PBE-D3/def2-TZVP level of theory for polymorphs **1a**, **1b**, and **1c** of compound **1**. Note that every point denotes a distortion energy or dimer interaction energy obtained for the given polymorph.

arene contact. In case of multiple C–H_{Ph}···π contacts these can become dominating. As a result, triorganobismuth compounds show a diversity of polymorphs.

In compounds with bulky ligands the formation of Bi···π arene contacts is hindered and multiple C–H_{Ph}···π contacts dominate. In the presence of strong structure directing donor–acceptor bonds the role of bismuth as DED can usually be neglected. However, compounds of the type Ar_2BiX (Ar = C_6H_3-*t*-Bu_2-3,5, X = halide) show some special features. In these compounds the formation of a one dimensional ribbon as a result of Bi···X···Bi coordination is typical, which is supplemented by Bi···π arene interactions between two neighboring bismuth atoms in the chain. Thus, a sort of X, π-pincer system is obtained. In order to strengthen the Bi···π arene interaction and to induce directionality in structure formation it is important to introduce electron-withdrawing substituents. Otherwise, a subtle interplay between Bi···π arene and the dispersion type forces must be considered.

Analysis of the Bi···π arene interaction in the $BiPh_3$–benzene complex shows that it is of moderate strength (−17 kJ mol^{-1}). Comparing the $BiPh_3$–benzene complex with other BiR_3-benzene systems (with R = Me, OMe, Cl) exhibits that the nature of this complex is mainly dispersive with small addition of donor–acceptor character which brings it closer to the $BiMe_3$ rather than to $Bi(OMe)_3$ as a dispersion energy donor. The weak donor–acceptor character of $BiPh_3$ causes that the Bi···π arene interactions compete with π···π and C–H_{Ph}···π interactions. Inspection of the intermolecular interactions in polymorphs **1a**, **1b**, and **1c** of $BiPh_3$ (**1**) confirms that Bi···π arene interactions are very important building blocks of the bulk. These are rather strong with interaction energies in the range from −28 kJ mol^{-1} to −47 kJ mol^{-1} and are purely dispersive. These energies are much higher than the interaction energy obtained for the model $BiPh_3$···benzene system. An analysis of selected tetramer units reveals that also π-stacking interactions and contacts between layers of $BiPh_3$ molecules are crucial in the formation of the crystal structures. The interaction energies of the π-stacking dimers are as high as interaction energies of Bi···π arene complexes or even larger (−64 kJ mol^{-1}). The energy of such dimers depends strongly on the distance and the contact area between two monomers. Both types of dimers are exclusively dispersive as shown by LED analysis performed at the DLPNO-CCSD(T) level of theory. Analysis of tetrameric units also reveals that the interaction energy of tetramers is additive and can be described as a sum of interaction energies of particular dimers.

In the polymorphs of compound **1** the energetics of interactions is balanced between Bi···π arene and π···π interactions that are of comparable strength. In case of polymorph **1a** the Bi···π arene interaction dominates, in case of polymorph **1b** the Bi···π arene and the π···π interactions are of similar magnitude. For polymorph **1c**, π···π interactions dominate the intermolecular interactions.

Overall, the compounds and structures discussed in this work demonstrate that a broad range of intermolecular interaction motifs are accessible by tuning the donor–acceptor properties of bismuth as a dispersion energy donor. Using electronic structure theory, these interactions can be quantified and studied in detail.

Experimental

Crystallographic studies

Crystal data, data collection and refinement parameters for Ph_3Bi (polymorphs **1a**, **1b**, **1c**, **1d**), **1b**, **2a**, **2b** and **3, 4**, **5**·$2CH_2Cl_2$ are given in Table S1, Table S2 and Table S3 (in Supporting Information File 1), respectively. All data for the new structures were collected with an Oxford Gemini S diffractometer at 123 K (**1b**), 120 K (**2a**, **3**), 115 K (**4**, **5**·$2CH_2Cl_2$) and 100 K (**2b**) using Cu Kα radiation (λ = 1.54184 Å) for **2a** and

Mo Kα radiation (λ = 0.71073 Å) for **1b**, **2b**, **3**, **4**, **5**·2CH_2Cl_2. The structures were solved by direct methods using SHELXS-2013 [68,69] and refined by full-matrix least-square procedures on F^2 using SHELXL-2014 [68,70] and SHELXL-2016/6 [71]. All non-hydrogen atoms were refined anisotropically. All hydrogen atoms were geometrically placed and refined isotropically in riding modes using default parameters. The drawings were created with the Diamond program [72]. The identity of Ph_3Bi (polymorphs **1a**, **1b**), **2a**, **2b**, and **3** was confirmed by PXRD analyses. The simulated diffraction patterns of three polymorphs of Ph_3Bi (**1a**, **1b**, **1c**) are illustrated in Figure S6 (see Supporting Information File 1). The diffraction patterns of the measured diffractograms are in good agreement with those simulated from the single crystal X-ray crystallographic data (see Supporting Information File 1, Figures S7–S11). The crystal structure of **2b** shows one disordered aryl ring over the whole aryl ligand with an occupancy ratio of 0.689:0.311 (69:31%). CCDC 1828668 (**1b**), 1824685 (**2a**), 1824684 (**2b**), 1824683 (**3**), 1824221 (**4**), 1824222 (**5**).

Supporting Information

Synthesis of compounds **1**–**5**. Molecular structures of **2a**, **2b**, **4**, and **5** (Figures S1–S4). Temperature dependent PXRD of Ph_3Bi (**1a**, Figure S5), PXRD pattern of the three Ph_3Bi polymorphs (Figure S6), PXRD pattern of **1a**, **1b**, **2a**, **2b**, and **3** (Figures S7–S11). Crystallographic data and structure refinement details for (**1a**), [42] (**1b**), [45] (**1c**) [44] and (**1d**) [39], **1b**, **2a**, **2b** and **3**–**5**, respectively (Tables S1, S2, and S3). Computational details. Structures of π-stacking dimers found for polymorph **1a**, **1b** and **1c** of Ph_3Bi (Figures S12–S14). Interaction energies (with respect to $BiPh_3$ in crystal geometry) and total energies (with respect to fully relaxed $BiPh_3$) in kJ mol^{-1} of π-stacking dimers (Table S4). Cartesian coordinates.

Supporting Information File 1

Additional material.
[https://www.beilstein-journals.org/bjoc/content/supplementary/1860-5397-14-187-S1.pdf]

Acknowledgements

We gratefully acknowledge financial support from the DFG SPP1807 "*Control of London Dispersion Interactions in Molecular Chemistry*". We thank Prof. S. Spange for access to IR instruments. We are grateful to Anja Veit for measuring the IR spectra. We also thank Ute Stöß and Janine Freytag for performing the CHN analyses. L. W. thanks the Deutscher Akademischer Austauschdienst (DAAD - Group of Eight Australia – Germany Joint Research Co-operation) for financial support. P. K. gratefully acknowledges financial support from the Deutsche Akademie der Naturforscher Leopoldina–Nationale Akademie der Wissenschaften for a Leopoldina Postdoc Fellowship. M. Korb thanks the Fonds der Chemischen Industrie for a Ph.D. Chemiefonds fellowship.

ORCID® iDs

Ana-Maria Preda - https://orcid.org/0000-0003-0047-0152
Phil C. Andrews - https://orcid.org/0000-0002-3971-7311
Marcus Korb - https://orcid.org/0000-0001-5453-4137
Alexander A. Auer - https://orcid.org/0000-0001-6012-3027
Michael Mehring - https://orcid.org/0000-0001-6485-6156

References

1. Auer, A. A.; Mansfeld, D.; Nolde, C.; Schneider, W.; Schürmann, M.; Mehring, M. *Organometallics* **2009,** *28,* 5405–5411. doi:10.1021/om900536r
2. Schmidbaur, H.; Schier, A. *Organometallics* **2008,** *27,* 2361–2395. doi:10.1021/om701044e
3. Silvestru, C.; Breunig, H. J.; Althaus, H. *Chem. Rev.* **1999,** *99,* 3277–3328. doi:10.1021/cr980083q
4. Cangelosi, V. M.; Pitt, M. A.; Vickaryous, W. J.; Allen, C. A.; Zakharov, L. N.; Johnson, D. W. *Cryst. Growth Des.* **2010,** *10,* 3531–3536. doi:10.1021/cg100444n
5. Zukerman-Schpector, J.; Tiekink, E. R. T. Intermolecular C–H · · · π(Chelate) Interactions – Prevalence in the Crystal Structures of Metal 1,1-Dithiolates. *The Importance of Pi-Interactions in Crystal Engineering;* John Wiley & Sons, Ltd.: Chichester, U.K., 2012; pp 275–299. doi:10.1002/9781119945888.ch11
6. Schneider, W. B.; Bistoni, G.; Sparta, M.; Saitow, M.; Riplinger, C.; Auer, A. A.; Neese, F. *J. Chem. Theory Comput.* **2016,** *12,* 4778–4792. doi:10.1021/acs.jctc.6b00523
7. Bistoni, G.; Auer, A. A.; Neese, F. *Chem. – Eur. J.* **2017,** *23,* 865–873. doi:10.1002/chem.201604127
8. Grimme, S.; Hansen, A.; Brandenburg, J. G.; Bannwarth, C. *Chem. Rev.* **2016,** *116,* 5105–5154. doi:10.1021/acs.chemrev.5b00533
9. Liptrot, D. J.; Power, P. P. *Nat. Rev. Chem.* **2017,** *1,* No. 0004. doi:10.1038/s41570-016-0004
10. Caracelli, I.; Haiduc, I.; Zukerman-Schpector, J.; Tiekink, E. R. T. *Coord. Chem. Rev.* **2013,** *257,* 2863–2879. doi:10.1016/j.ccr.2013.05.022
11. Caracelli, I.; Zukerman-Schpector, J.; Haiduc, I.; Tiekink, E. R. T. *CrystEngComm* **2016,** *18,* 6960–6978. doi:10.1039/C6CE01460G
12. Tiekink, E. R. T. *Coord. Chem. Rev.* **2017,** *345,* 209–228. doi:10.1016/j.ccr.2017.01.009
13. Sadler, P. J.; Li, H.; Sun, H. *Coord. Chem. Rev.* **1999,** *185–186,* 689–709. doi:10.1016/S0010-8545(99)00018-1
14. Briand, G. G.; Burford, N. *Chem. Rev.* **1999,** *99,* 2601–2658. doi:10.1021/cr980425s
15. Yang, N.; Sun, H. *Coord. Chem. Rev.* **2007,** *251,* 2354–2366. doi:10.1016/j.ccr.2007.03.003
16. Monakhov, K. Y.; Gourlaouen, C.; Zessin, T.; Linti, G. *Inorg. Chem.* **2013,** *52,* 6782–6784. doi:10.1021/ic401001f
17. Fanfrlik, J.; Sedlak, R.; Pecina, A.; Rulíšek, L.; Dostál, L.; Moncoľ, J.; Růžička, A.; Hobza, P. *Dalton Trans.* **2016,** *45,* 462–465. doi:10.1039/C5DT04381F

18. Arora, J. S.; Gaikar, V. G. *RSC Adv.* **2016,** *6,* 39663–39674. doi:10.1039/C6RA01960A
19. Senevirathna, D. C.; Werrett, M. V.; Pai, N.; Blair, V. L.; Spiccia, L.; Andrews, P. C. *Chem. – Eur. J.* **2017,** *23,* 8171–8175. doi:10.1002/chem.201701952
20. Srinivas, K.; Suresh, P.; Babu, C. N.; Sathyanarayana, A.; Prabusankar, G. *RSC Adv.* **2015,** *5,* 15579–15590. doi:10.1039/C4RA17144F
21. Hering-Junghans, C.; Schulz, A.; Villinger, A. *Chem. Commun.* **2015,** *51,* 13834–13837. doi:10.1039/C5CC04516A
22. Hering-Junghans, C.; Schulz, A.; Thomas, M.; Villinger, A. *Dalton Trans.* **2016,** *45,* 6053–6059. doi:10.1039/C6DT00229C
23. Schwamm, R. J.; Coles, M. P.; Fitchett, C. M. *Dalton Trans.* **2017,** *46,* 4066–4074. doi:10.1039/C7DT00230K
24. Luqman, A.; Blair, V. L.; Brammananth, R.; Crellin, P. K.; Coppel, R. L.; Kedzierski, L.; Andrews, P. C. *Eur. J. Inorg. Chem.* **2015,** 725–733. doi:10.1002/ejic.201402958
25. Mansfeld, D.; Mehring, M.; Schürmann, M. *Z. Anorg. Allg. Chem.* **2004,** *630,* 1795–1797. doi:10.1002/zaac.200400214
26. Toma, A. M.; Pop, A.; Silvestru, A.; Rüffer, T.; Lang, H.; Mehring, M. *Dalton Trans.* **2017,** *46,* 3953–3962. doi:10.1039/C7DT00188F
27. Breunig, H. J.; Haddad, N.; Lork, E.; Mehring, M.; Mügge, C.; Nolde, C.; Raţ, C. I.; Schürmann, M. *Organometallics* **2009,** *28,* 1202–1211. doi:10.1021/om800934c
28. Preda, A. M.; Schneider, W. B.; Rainer, M.; Rüffer, T.; Schaarschmidt, D.; Lang, H.; Mehring, M. *Dalton Trans.* **2017,** *46,* 8269–8278. doi:10.1039/C7DT01437F
29. Preda, A. M.; Schneider, W. B.; Schaarschmidt, D.; Lang, H.; Mertens, L.; Auer, A. A.; Mehring, M. *Dalton Trans.* **2017,** *46,* 13492–13501. doi:10.1039/C7DT02567J
30. Bauzá, A.; Quiñonero, D.; Deyà, P. M.; Frontera, A. *Phys. Chem. Chem. Phys.* **2012,** *14,* 14061–14066. doi:10.1039/c2cp42672b
31. Bauzá, A.; Quiñonero, D.; Deyà, P. M.; Frontera, A. *CrystEngComm* **2013,** *15,* 3137–3144. doi:10.1039/C2CE26741A
32. Neese, F.; Hansen, A.; Liakos, D. G. *J. Chem. Phys.* **2009,** *131,* 064103. doi:10.1063/1.3173827
33. Neese, F.; Hansen, A.; Wennmohs, F.; Grimme, S. *Acc. Chem. Res.* **2009,** *42,* 641–648. doi:10.1021/ar800241t
34. Hansen, A.; Liakos, D. G.; Neese, F. *J. Chem. Phys.* **2011,** *135,* 214102. doi:10.1063/1.3663855
35. Liakos, D. G.; Hansen, A.; Neese, F. *J. Chem. Theory Comput.* **2011,** *7,* 76–87. doi:10.1021/ct100445s
36. Riplinger, C.; Neese, F. *J. Chem. Phys.* **2013,** *138,* 034106. doi:10.1063/1.4773581
37. Riplinger, C.; Sandhoefer, B.; Hansen, A.; Neese, F. *J. Chem. Phys.* **2013,** *139,* 134101. doi:10.1063/1.4821834
38. Riplinger, C.; Pinski, P.; Becker, U.; Valeev, E. F.; Neese, F. *J. Chem. Phys.* **2016,** *144,* 024109. doi:10.1063/1.4939030
39. Wenzel, J. *Z. Kristallogr., Kristallgeom., Kristallphys., Kristallchem.* **1942,** *104,* 305.
40. Hawley, D. M.; Ferguson, G.; Harris, G. S. *Chem. Commun.* **1966,** 111–112. doi:10.1039/c19660000111
41. Hawley, D. M.; Ferguson, G. *J. Chem. Soc. A* **1968,** 2059–2063. doi:10.1039/j19680002059
42. Jones, P. G.; Blaschette, A.; Henschel, D.; Weitze, A. *Z. Kristallogr.* **1995,** *210,* 377–378. doi:10.1524/zkri.1995.210.5.377
43. Bučinský, L.; Jayatilaka, D.; Grabowsky, S. *J. Phys. Chem. A* **2016,** *120,* 6650–6669. doi:10.1021/acs.jpca.6b05769
44. Stammler, H.-G.; Neumann, B., *Private Communication to the CSD (BITRPH11)* 2014, DOI: 10.5517/cc13k8qn.
45. Andrews, P. C.; MacLellan, J. G., *personal communication.*
46. Michaelis, A.; Polis, A. *Ber. Dtsch. Chem. Ges.* **1887,** *20,* 54–57. doi:10.1002/cber.18870200110
47. Michaelis, A. *Justus Liebigs Ann. Chem.* **1889,** *251,* 323–335. doi:10.1002/jlac.18892510305
48. Pfeiffer, P. *Ber. Dtsch. Chem. Ges.* **1904,** *37,* 4620–4623. doi:10.1002/cber.19040370468
49. Florinskij, F. S.; Koton, M. M. *Preparation of tris(p-vinylphenyl)antimony or –bismuth.* Soviet Patent SU 165716, 1964.
50. Supniewski, J. V.; Adams, R. *J. Am. Chem. Soc.* **1926,** *48,* 507–517. doi:10.1021/ja01413a031
51. Gillmeister, A. *Ber. Dtsch. Chem. Ges.* **1897,** *30,* 2843–2850. doi:10.1002/cber.18970300380
52. Hassan, A.; Breeze, S. R.; Courtenay, S.; Deslippe, C.; Wang, S. *Organometallics* **1996,** *15,* 5613–5621. doi:10.1021/om960641w
53. Petiot, P.; Dansereau, J.; Gagnon, A. *RSC Adv.* **2014,** *4,* 22255–22259. doi:10.1039/C4RA02467B
54. Grimme, S. *Angew. Chem., Int. Ed.* **2008,** *47,* 3430–3434. doi:10.1002/anie.200705157
55. Martinez, C. R.; Iverson, B. L. *Chem. Sci.* **2012,** *3,* 2191–2201. doi:10.1039/c2sc20045g
56. Ogawa, T.; Ikegami, T.; Hikasa, T.; Ono, N.; Suzuki, H. *J. Chem. Soc., Perkin Trans. 1* **1994,** 3479–3483. doi:10.1039/p19940003479
57. Stavila, V.; Thurston, J. H.; Prieto-Centurión, D.; Whitmire, K. H. *Organometallics* **2007,** *26,* 6864–6866. doi:10.1021/om7009667
58. Nishio, M. *Phys. Chem. Chem. Phys.* **2011,** *13,* 13873–13900. doi:10.1039/c1cp20404a
59. Nishio, M. *CrystEngComm* **2004,** *6,* 130–158. doi:10.1039/b313104a
60. Alvarez, S. *Dalton Trans.* **2013,** *42,* 8617–8636. doi:10.1039/c3dt50599e
61. Bondi, A. *J. Phys. Chem.* **1964,** *68,* 441–451. doi:10.1021/j100785a001
62. Mantina, M.; Chamberlin, A. C.; Valero, R.; Cramer, C. J.; Truhlar, D. G. *J. Phys. Chem. A* **2009,** *113,* 5806–5812. doi:10.1021/jp8111556
63. Haack, R.; Schulz, S.; Jansen, G. *J. Comput. Chem.* **2018,** *39,* 1413–1423. doi:10.1002/jcc.25209
64. Hébert, M.; Petiot, P.; Benoit, E.; Dansereau, J.; Ahmad, T.; Le Roch, A.; Ottenwaelder, X.; Gagnon, A. *J. Org. Chem.* **2016,** *81,* 5401–5416. doi:10.1021/acs.joc.6b00767
65. Rao, M. L. N.; Dhanorkar, R. J. *RSC Adv.* **2016,** *6,* 1012–1017. doi:10.1039/C5RA23311A
66. Neese, F. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2012,** *2,* 73–78. doi:10.1002/wcms.81
67. Krasowska, M.; Schneider, W. B.; Mehring, M.; Auer, A. A. *Chem. – Eur. J.* **2018,** *39,* 1413–1423. doi:10.1002/chem.201801758
68. Sheldrick, G. M. *Acta Crystallogr., Sect. A: Found. Crystallogr.* **2008,** *64,* 112–122. doi:10.1107/S0108767307043930
69. *SHELXL-2013, Program for Crystal Structures Refinement;* University of Göttingen: Göttingen, Germany, 2013.
70. *SHELXL-2014, Program for the Solution of Crystal Structures;* University of Göttingen: Göttingen, Germany, 2014.
71. Sheldrick, G. M. *Acta Crystallogr., Sect. C: Struct. Chem.* **2015,** *71,* 3–8. doi:10.1107/S2053229614024218
72. *DIAMOND Visual Crystal Structure Information System,* 31d; Crystal Impact: Bonn, Germany, 2006.

BEILSTEIN JOURNAL OF ORGANIC CHEMISTRY

Dispersion-mediated steering of organic adsorbates on a precovered silicon surface

Lisa Pecher, Sebastian Schmidt and Ralf Tonner*

Full Research Paper

Open Access

Address:
Fachbereich Chemie and Material Sciences Center, Philipps-Universität Marburg, Hans-Meerwein-Straße 4, 35032 Marburg, Germany

Email:
Ralf Tonner* - tonner@chemie.uni-marburg.de

* Corresponding author

Keywords:
bonding analysis; cyclooctyne; density functional theory; dispersion; organic/inorganic interfaces

Beilstein J. Org. Chem. **2018,** *14,* 2715–2721.
doi:10.3762/bjoc.14.249

Received: 06 August 2018
Accepted: 16 October 2018
Published: 26 October 2018

This article is part of the Thematic Series "Dispersion interactions".

Guest Editor: P. Schreiner

Abstract

The chemistry of organic adsorbates on surfaces is often discussed in terms of Pauli repulsion as limiting factor regarding the packing of molecules. Here we show that the attractive part of the van der Waals potential can be similarly decisive. For the semiconductor surface Si(001), an already covalently bonded molecule of cyclooctyne steers a second incoming molecule via dispersion interactions onto the neighbouring adsorption site. This helps in understanding the nonstatistical pattern formation for this surface–adsorbate system and hints toward an inclusion of dispersion attraction as another determining factor for surface adsorption.

Introduction

The creation of organic/inorganic interfaces is one of the main endeavours in enhancing the application range of modern electronic devices for silicon-based technology [1,2]. One way to achieve this is covalent attachment of bifunctional organic molecules on bare silicon surfaces and subsequent reaction with a second molecule with both reactions being chemoselective (layer-by-layer, LbL, approach) [3-5]. To achieve an interface structure with predictable properties, it is important that the molecules used for the first layer show well-defined surface chemistry without side reactions and lead to densely packed and well-ordered structures.

Cyclooctyne (**1**), the smallest stable cyclic alkyne, on Si(001) is a system where this is the case and it has previously been thoroughly studied by experiment and theory [6-8]. Even though **1** is missing a second functional group necessary for the LbL approach, previous studies have shown that synthetic routes exist for derivatization and that the reactivity of the strained triple bond of **1** with the surface is not affected by the second functional group [4,5,9]. Studying the adsorption behaviour of the parent system **1** thus gives crucial insight that is expected to be transferable to the bifunctional derivatives.

The adsorption of a molecule on a surface can proceed either via a direct pathway or via an intermediate species that is crucial for selectivity and the description of adsorption dynamics (Figure 1). The dominant interaction between molecule and surface changes with the distance: For surface–adsorbate dis-

tances at which there is no significant orbital overlap but already rather close contact, dispersion attraction dominates since the numerous rather weak interactions add up to a significant stabilization, especially for larger adsorbates and/or polarizable substrates. At shorter distances, covalent bond formation leads to a steeper attractive potential well. The resulting covalently bonded state is usually called chemisorbed [10]. Bonding of intermediate states can be dominated by dispersion or covalent interactions. At very short distances, Pauli repulsion creates the repulsive potential wall.

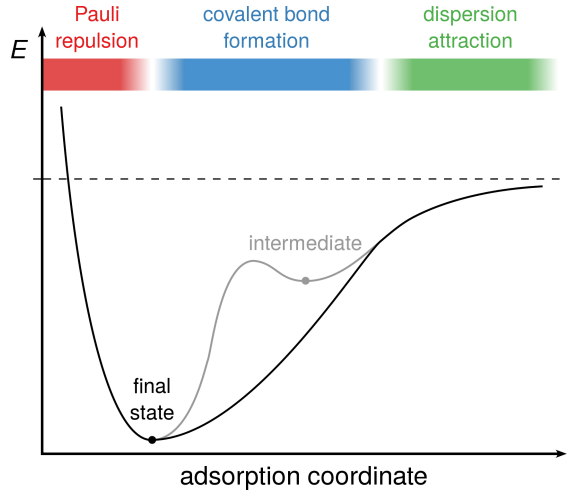


Figure 1: Adsorption energy profile of a direct adsorption (black, e.g., cyclooctyne/Si(001)) in comparison with adsorption via an intermediate (grey, e.g., methanol/Ge(001)). On top, the dominant type of interaction in the different regimes of the profile is given.

Experimental studies in combination with Monte Carlo simulations have shown that growth of **1** on Si(001) results in non-statistical formation of chains with an average distance of 1.5 to 2 dimers between adsorbates [6]. For the adsorption of methanol on Ge(001), where a similar behaviour is observed, it was shown that in an intermediate state, interactions with other adsorbed molecules lead to a reduction of the energy barrier for conversion into the final chemisorbed state [11]. This accelerates the adsorption next to occupied sites and leads to the formation of 1D chains, as derived from computations. However, since the adsorption of cyclooctyne on Si(001) is direct or pseudo-direct and does not proceed via an intermediate [5,6,8], this explanation is not applicable. Previously, it was proposed that the occupied sites might "steer" impinging molecules via an attractive adsorption potential close to an already adsorbed molecule [6].

Here, we will show that this steering potential is indeed found and is caused by attractive dispersion interactions. To this end, we investigated the adsorption of a molecule on a pre-covered surface using density functional theory (DFT) approaches with and without dispersion correction terms. In contrast to the above-mentioned intermediate-based selectivity, the steering-type interaction takes place before covalent bonds between molecule and surface are formed. The results show that dispersive interactions can be decisive in building novel organic structures on surfaces by tweaking the potential energy surface.

Computational Details

All calculations were performed with the Vienna Ab Initio Simulation Package (VASP) [12-15] version 5.3.5 using the PBE functional [16,17], the DFT-D3 dispersion correction [18,19] and the PAW formalism [20,21] with a basis set cutoff of E_{cutoff} = 400 eV. Electronic k space was sampled using a Γ(221) grid. Some calculations used the DFT-TS scheme for comparison [22]. Self-consistent field (SCF) and structural optimization convergence criteria were set to 10^{-6} eV and 10^{-2} eV·Å^{-1}, respectively. Structures were optimized using the Conjugate Gradient algorithm [23] and Gibbs energies were calculated at T = 300 K, p = 1 bar using an approach described elsewhere [24]. Harmonic vibrational frequencies used in the calculation of Gibbs energies were derived by numerical construction of the Hessian using Cartesian displacements of 0.01 Å from the equilibrium structure. The Si(001) surface was modelled as a six-layer slab in $c(4 \times 2)$ reconstruction with 4 × 4 atoms per layer. The frozen double layer approximation was applied (i.e., the bottom two layers were not relaxed in structural optimizations) and the bottom layer saturated with hydrogen atoms in tetrahedral arrangement at d(Si-H) = 1.480 Å, the experimental equilibrium distance in silane [25]. Cell constants a and b (in x and y direction) were set to 15.324 Å, derived from an optimized bulk parameter of 5.418 Å for this computational setup [7], while in z direction, a vacuum layer of at least 10 Å was ensured. The bonding energy E_{bond} was defined as the energy difference between the relaxed structures of the total system (E_{tot}) and the isolated molecule (E_{mol}) and surface (E_{surf}):

$$E_{\text{bond}} = E_{\text{tot}} - E_{\text{mol}} - E_{\text{surf}}.$$

Please note that in case of a precovered surface, E_{surf} also includes the already adsorbed molecule, and that surface science convention is the use of the adsorption energy E_{ads} with inverse sign convention ($E_{\text{ads}} = -E_{\text{bond}}$).

Adsorption energy profiles were calculated by placing the cyclooctyne molecule in an upright orientation (molecular C_2 axis aligned parallel to the z axis of the cell), with the triple bond aligned parallel to the y axis of the cell, the triple-bond centre located vertically above a lower surface atom (Si_{down}) at a height corresponding to a vertical distance between the triple-

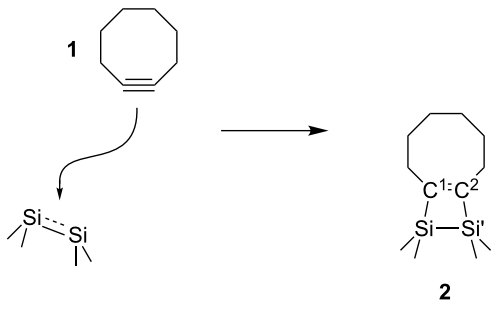
bond carbon atoms and the uppermost surface atoms (Si_{up}) of $\Delta z(Si_{up}–C_{triple}) = 4$ Å. The system was then optimized using the Conjugate Gradient algorithm. In a previous study, we have shown that this approach yields an energy profile that is in qualitative agreement with the true minimum energy path for this system [8].

Potential energy surface scans were performed by displacing a cyclooctyne molecule in *x* and *y* direction while retaining the orientation (equivalent to the starting point of the adsorption energy profile) and a fixed distance $\Delta z(Si_{up}–C_{triple})$ above the surface. The displacement grid was chosen to consist of 20 × 20 equidistant points spanning the whole unit cell, corresponding to a distance of 0.766 Å between individual grid points. Since the system was not optimized at each grid point, this corresponds to a so-called frozen scan. The approach outlined here has delivered accurate results for organic/semiconductor systems in the past [7-9,24].

Results and Discussion

Bonding and the adsorption path

The reactivity of the Si(001) surface is dominated by Si surface dimers with an electronic structure that is well represented by an electrophilic and a nucleophilic Si atom. The adsorption of a first molecule of **1** on Si(001) is characterized by a direct adsorption path without intermediate structure leading to a strongly covalently bonded [2 + 2] cycloaddition product **2** as summarized in Scheme 1 with ring strain being decisive for the high reactivity of **1** [5,6,8]. Not reflected in the Lewis structure is the tilting of the molecule upon adsorption leading to a chair-like conformer bending over the dimer rows on the surface [7].



Scheme 1: The reaction of cyclooctyne (**1**) with a Si(001) surface dimer, yielding a [2 + 2] cycloadduct **2**.

Starting from this precovered surface (i.e., decorated with one adsorbate in the unit cell), we now investigate the adsorption of a second molecule of **1** on a neighbouring dimer leading to structure **3** (Figure 2). Although repulsive interactions might be expected for adsorption close to a rather large adsorbate, we find this mode to be the most stable adsorption mode for two molecules of **1** in the unit cell. Due to their conformational flexibility, both molecules **1** and **1′** bend away from each other (Figure 2), thus reducing steric repulsion as further discussed below. An alternative structure where both cyclooctyne molecules bend in the same direction is higher in energy (+4 kJ·mol^{-1}), although dispersion attraction is slightly more stabilizing compared to **3** (by 3 kJ·mol^{-1}). We will thus focus our discussion on the minimum-energy structure.

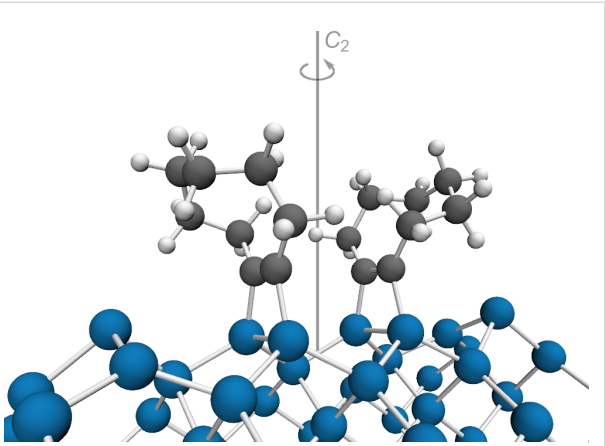


Figure 2: Optimized (PBE-D3/PAW) structure of two molecules of **1** on Si(001) on neighbouring surface dimers (**3**).

Comparison of key structural parameters (Table 1) shows that the C–C as well as the C–Si bond lengths are essentially unaffected by the presence of the second molecule. Interestingly, the energy minimization without symmetry constraints leads to a structure with a local C_2 rotational axis resulting in symmetry-equivalent molecules **1** and **1′**.

Table 1: Selected interatomic distances (in Å) of **2** and **3**.[a]

	$d(C^1–C^2)$	$d(C^1–Si)$	$d(C^2–Si')$
2[b]	1.368	1.916	1.900
3[c]	1.368	1.915	1.904

[a]See Scheme 1 for nomenclature; [b]values taken from [7]; [c]the two molecules are symmetry-equivalent.

Energies and Gibbs energies of adsorption were previously found to support the notion of strong covalent bonding for the [2 + 2] cycloaddition of **1** on the silicon surface [7]. The adsorption energy for a second molecule on the precovered surface is now found to be even slightly larger by 11 kJ·mol^{-1} (Table 2). This is surprising at first since the presence of a rather bulky adsorbate on the surface should lead to a blocking of neighbouring sites by Pauli repulsion. As we will see later, this is indeed the case for one of the neighbouring dimers. But due to

the tilting of the first molecule of **1**, the second adsorbate is not hindered by repulsive interactions. The electronic contribution to the adsorption energy (E_{bond}(PBE) in Table 2) is indeed unchanged. On the contrary, the dispersion contribution shows an increase for the second adsorbate (−55 kJ·mol^{-1} vs −47 kJ·mol^{-1} for adsorption on the clean surface), which is the main cause for the slightly larger bonding energy.

Table 2: Adsorption energies (in kJ·mol^{-1}) of **1** on a precovered Si(001) surface leading to **3** compared with the corresponding values for adsorption on a clean surface leading to **2**.

	clean surface[a]	precovered surface
E_{bond}(PBE)[b]	−261	−264
E_{bond}(D3)[b]	−47	−55
E_{bond}(PBE-D3)	−308	−319
G_{bond}(PBE-D3)	–238	−249

[a]Values taken from [7]; [b]electronic (PBE) and dispersive (DFT-D3) contributions adding up to E_{bond}, derived from the PBE-D3 structure.

Energy decomposition analysis for both structures (Table 3) confirms that indeed Pauli repulsion is virtually the same for adsorption on the clean (ΔE_{Pauli} = 1468 kJ·mol^{-1}) and precovered surface (ΔE_{Pauli} = 1467 kJ·mol^{-1}) while small changes in electrostatic (ΔE_{elstat}) and orbital (ΔE_{orb}) contributions compensate each other. This leaves the increase in dispersion interaction by 8 kJ·mol^{-1} for the precovered surface as the major, albeit small, contribution to the slightly larger interaction energy thus confirming the finding above. Thus, the changes in the pEDA energy terms are rather small but the most important observation is that Pauli repulsion does not significantly rise as is often found for the adsorption of molecules on precovered surfaces [24].

Table 3: Energy decomposition analysis (pEDA) results (PBE-D3/TZ2P) for the adsorption of **1** on a clean and precovered Si(001) surface. All values in kJ·mol^{-1}.

	clean surface[a]	precovered surface
ΔE_{int}	−658	−668
ΔE_{int}(disp)	−43 (7%)	−51 (8%)
ΔE_{int}(elec)	−615 (93%)	−616 (92%)
ΔE_{Pauli}	1468	1467
ΔE_{elstat}	−936 (45%)	−949 (46%)
ΔE_{orb}	–1148 (55%)	−1134 (54%)
ΔE_{prep}(mol.)	313	312
ΔE_{prep}(surf.)	26	30
E_{bond}[b]	−319 (−308)	−325 (−319)

[a]Values taken from [7]; [b]PAW values (in parentheses) given for comparison.

The bonding in the covalent [2 + 2] cycloaddition product (i.e., the final state of adsorption) is thus very similar for clean and precovered surfaces. But the reaction path leading to this state might still be qualitatively changed by the presence of a molecule **1** on the surface. The comparison of optimized adsorption paths for clean and precovered surface in Figure 3 shows that this is not the case. In agreement with experimental observation and our previous findings, a direct pathway is observed for the adsorption of **1** on the silicon surface without an intermediate

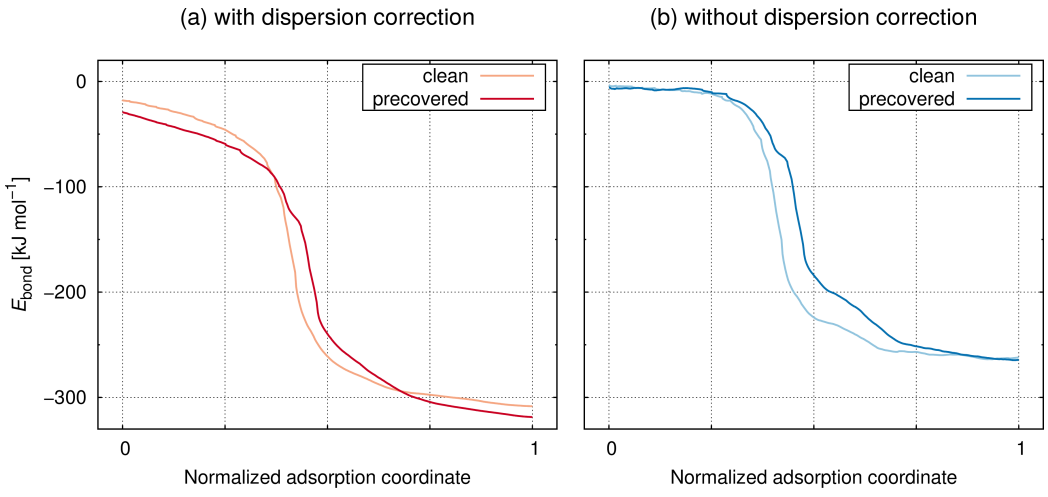


Figure 3: Energy profile of the adsorption pathway depicted in Scheme 1 on the clean and precovered Si(001) surface computed with (a) PBE-D3 and (b) PBE.

that would show up as stationary point in the energy profile [5-7]. The only difference is found in the adsorption paths when dispersion corrections are included in the computation (Figure 3a).The curve is rather constantly shifted by 5–10 kJ·mol^{-1} towards more negative bonding energies E_{bond} in case of the precovered surface. This is not found in the computation that omits dispersion forces (Figure 3b). Thus, dispersion interactions not only stabilize product **3** but act along the whole adsorption path of **1** onto Si(001). This leads us to a comprehensive investigation of the potential energy surface of adsorption.

The potential energy surface

Since both the product and the pathway are influenced by dispersion interactions, the question arises if these forces can tweak the potential energy surface (PES) in a way to steer the second adsorbate onto a certain position on the surface. We investigated this by conducting rigid PES scans on the clean and precovered surface by systematically placing **1** on a grid of possible positions at a fixed distance to the surface (Figure 4). We considered different adsorption heights and orientations of **1** (see "Computational Details" and Supporting Information File 1).

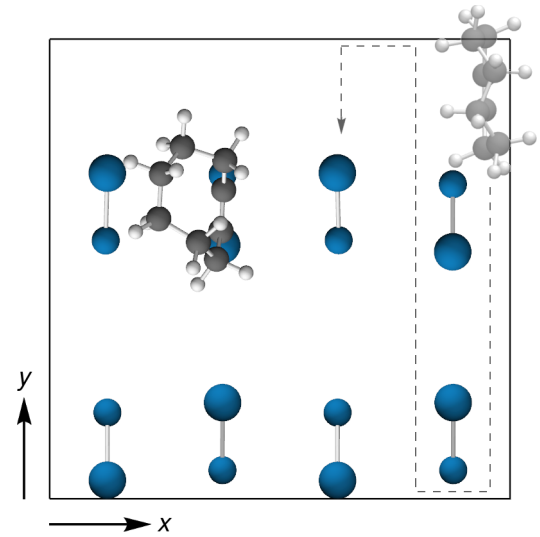


Figure 4: Si(001) Surface precovered with one adsorbate **1** and unit cell used in the PES scans (orientation of second adsorbate and scan path indicated) shown in Figure 5.

The most interesting data set is found for the case where the triple bond of **1** points toward the surface at an adsorption height of $\Delta z(Si_{up}–C_{triple})$ = 5 Å (Figure 5). For the clean surface, the PES is essentially featureless (Figure 5a) and only minor effects (stabilization by less than 5 kJ·mol^{-1}) are found by including dispersion in the computation. If one molecule of **1** is already present at the surface, parts of the surface are inaccessible due to strong Pauli repulsion and overlapping molecules. This is indicated by the white areas around the first adsorbate in Figure 5b. Now, significant differences can be found between the PES scan with (top panels) and without (bottom panels) considering dispersion effects. For the computations without dispersion correction, the PES is again rather featureless and the interaction between adsorbate and surface is very weak. This

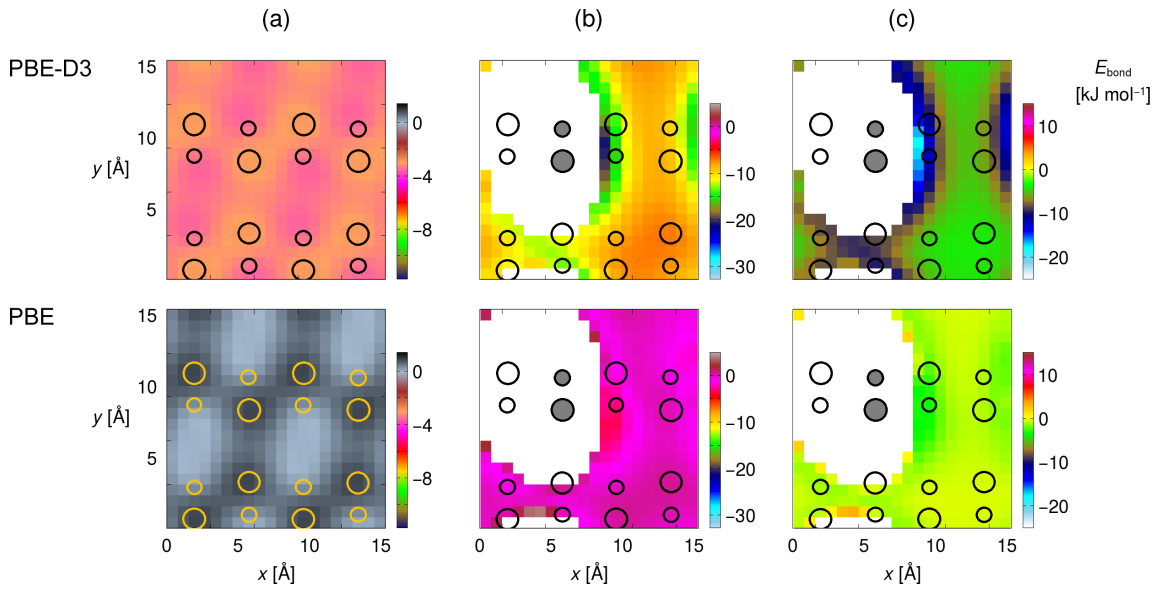


Figure 5: Frozen PES scans of **1** along path indicated in Figure 4 (C_2 axis parallel to *z*, C≡C parallel to *y*, $\Delta z(Si_{up}–C_{triple})$ = 5 Å) on (a) the clean and (b) the precovered surface, and (c) the difference between (a) and (b). Circles denote surface atoms, grey shading denotes the occupied site. White region: no values given due to highly repulsive interactions or overlapping molecules.

can be seen in the difference plot between the PES of the clean and precovered surface in Figure 5c. Only weak preference for the surface dimer adjacent to the already adsorbed molecule is found (less than 5 kJ·mol^{-1}), which can be attributed to weak electrostatic attraction between the two molecules.

The picture changes completely when dispersion attraction is considered. The PBE-D3 computations show a pronounced feature in the PES scan on the precovered surface with a strong energetic preference for adsorption on the surface dimer next to the first adsorbate. The stabilization can be seen in the difference to the PES of the clean surface (Figure 5c) and amounts to ca. 20 kJ·mol^{-1} out of a total molecule–surface attraction of ≤25 kJ·mol^{-1}. Notably, the tilting of **1** in structure **2** (Figure 4) thus leads to a blocking of one adjacent dimer in *x* direction, but an adsorption preference on the other adjacent dimer leading to structure **3** (Figure 2).

This preference is most pronounced for an adsorption height of 5 Å shown here but is also found for vertical distances of 3 and 7 Å to the surface (Figure S1, Supporting Information File 1). It is also not an artifact of the dispersion correction method chosen (DFT-D3) since a scan with a second method (DFT-TS) leads to the same picture with only slight numeric differences (Figure S2, Supporting Information File 1).

The double-adsorption structure **3** will now lead to a blocking of two dimers and thus result on average in a distance of two dimers between adsorbates. As was shown before, **1** can also adsorb in the twist-boat conformation leading to an arrangement of three molecules on three consecutive dimers [7]. The resulting coverage is thus in agreement with the coverages derived from analysis of the experimental structure [6].

Conclusion

We have shown that dispersion effects are not only important for the thermodynamic stability of molecule–adsorbate complexes but they also crucially influence the adsorption path. While Pauli repulsion is often discussed as important effect for determining surface adsorption, the attractive part of the van der Waals potential can be of similar importance. For the system cyclooctyne on Si(001), attractive dispersion interactions lead to a preferred adsorption of a second molecule in the neighbourhood of a first adsorbate – an arrangement that is often excluded due to Pauli repulsion arguments. Experimental observation of nonstatistical chain formation can thus be explained. Especially for larger adsorbates, these attractive interactions are expected to play an important role in determining the surface arrangement of molecules and might thus be even used for designing patterned surfaces. To this end, ab initio modelling that accounts for dispersion interactions plays an important role.

Supporting Information

The supporting information shows PES scans comparing DFT-D3 and DFT-TS, scans at different adsorption heights as well as Cartesian coordinates and total energies for the equilibrium structures presented.

Supporting Information File 1

Additional calculational data.
[https://www.beilstein-journals.org/bjoc/content/supplementary/1860-5397-14-249-S1.pdf]

Acknowledgements

We thank the German Research Foundation (DFG) for funding via SFB 1083. Computational resources provided by HRZ Marburg, CSC-LOEWE Frankfurt and HLRS Stuttgart are kindly acknowledged.

ORCID® iDs

Lisa Pecher - https://orcid.org/0000-0001-7130-0287
Ralf Tonner - https://orcid.org/0000-0002-6759-8559

References

1. Wolkow, R. A. *Annu. Rev. Phys. Chem.* **1999,** *50,* 413–441. doi:10.1146/annurev.physchem.50.1.413
2. Teplyakov, A. V.; Bent, S. F. *J. Vac. Sci. Technol., A* **2013,** *31,* 050810. doi:10.1116/1.4810784
3. Decher, G. *Science* **1997,** *277,* 1232–1237. doi:10.1126/science.277.5330.1232
4. Münster, N.; Nikodemiak, P.; Koert, U. *Org. Lett.* **2016,** *18,* 4296–4299. doi:10.1021/acs.orglett.6b02048
5. Reutzel, M.; Münster, N.; Lipponer, M. A.; Länger, C.; Höfer, U.; Koert, U.; Dürr, M. *J. Phys. Chem. C* **2016,** *120,* 26284–26289. doi:10.1021/acs.jpcc.6b07501
6. Mette, G.; Dürr, M.; Bartholomäus, R.; Koert, U.; Höfer, U. *Chem. Phys. Lett.* **2013,** *556,* 70–76. doi:10.1016/j.cplett.2012.11.029
7. Pecher, J.; Schober, C.; Tonner, R. *Chem. – Eur. J.* **2017,** *23,* 5459–5466. doi:10.1002/chem.201605952
8. Pecher, L.; Schmidt, S.; Tonner, R. *J. Phys. Chem. C* **2017,** *121,* 26840–26850. doi:10.1021/acs.jpcc.7b09148
9. Pecher, L.; Tonner, R. *Theor. Chem. Acc.* **2018,** *137,* 48. doi:10.1007/s00214-018-2212-5
10. Pecher, J.; Tonner, R. *ChemPhysChem* **2017,** *18,* 34–38. doi:10.1002/cphc.201601129
11. Shong, B.; Bent, S. F. *J. Phys. Chem. C* **2013,** *117,* 949–955. doi:10.1021/jp3078503
12. Kresse, G.; Hafner, J. *Phys. Rev. B* **1993,** *47,* 558–561. doi:10.1103/PhysRevB.47.558
13. Kresse, G.; Hafner, J. *Phys. Rev. B* **1994,** *49,* 14251–14269. doi:10.1103/PhysRevB.49.14251
14. Kresse, G.; Furthmüller, J. *Phys. Rev. B* **1996,** *54,* 11169–11186. doi:10.1103/PhysRevB.54.11169
15. Kresse, G.; Furthmüller, J. *Comput. Mater. Sci.* **1996,** *6,* 15–50. doi:10.1016/0927-0256(96)00008-0

16. Perdew, J. P.; Burke, K.; Ernzerhof, M. *Phys. Rev. Lett.* **1996,** *77,* 3865–3868. doi:10.1103/PhysRevLett.77.3865
17. Perdew, J. P.; Burke, K.; Enzerhof, M. *Phys. Rev. Lett.* **1997,** *78,* 1396. doi:10.1103/PhysRevLett.78.1396
18. Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, S. *J. Chem. Phys.* **2010,** *132,* 154104. doi:10.1063/1.3382344
19. Grimme, S.; Ehrlich, S.; Goerigk, L. *J. Comput. Chem.* **2011,** *32,* 1456–1465. doi:10.1002/jcc.21759
20. Blöchl, P. *Phys. Rev. B* **1994,** *50,* 17953–17979. doi:10.1103/PhysRevB.50.17953
21. Kresse, G.; Joubert, D. *Phys. Rev. B* **1999,** *59,* 1758–1775. doi:10.1103/PhysRevB.59.1758
22. Tkatchenko, A.; Scheffler, M. *Phys. Rev. Lett.* **2009,** *102,* 073005. doi:10.1103/PhysRevLett.102.073005
23. Hestenes, M. R.; Stiefel, E. *J. Res. Natl. Bur. Stand. (U. S.)* **1952,** *49,* 409–436. doi:10.6028/jres.049.044
24. Pecher, J.; Mette, G.; Dürr, M.; Tonner, R. *ChemPhysChem* **2017,** *18,* 357–365. doi:10.1002/cphc.201601134
25. Boyd, D. R. J. *J. Chem. Phys.* **1955,** *23,* 922–926. doi:10.1063/1.1742148

License and Terms

BEILSTEIN JOURNAL OF ORGANIC CHEMISTRY

Adhesion, forces and the stability of interfaces

Robin Guttmann[1], Johannes Hoja[1,2], Christoph Lechner[1], Reinhard J. Maurer[3] and Alexander F. Sax[*1]

Full Research Paper

Open Access

Address:
[1]Department of Chemistry, University of Graz, Heinrichstrasse 28, 8010 Graz, Austria, [2]Present address: Physics and Materials Science Research Unit, University of Luxembourg, 1511 Luxembourg, Luxembourg and [3]Department of Chemistry and Centre for Scientific Computing, University of Warwick, Gibbet Hill Road, Coventry, CV4 7AL, United Kingdom

Email:
Alexander F. Sax[*] - alexander.sax@uni-graz.at

* Corresponding author

Keywords:
adhesion energy; adhesive force; dispersion interaction; weak molecular interaction

Beilstein J. Org. Chem. **2019,** *15,* 106–129.
doi:10.3762/bjoc.15.12

Received: 17 August 2018
Accepted: 12 December 2018
Published: 11 January 2019

This article is part of the Thematic Series "Dispersion interactions".

Guest Editor: P. Schreiner

Abstract

Weak molecular interactions (WMI) are responsible for processes such as physisorption; they are essential for the structure and stability of interfaces, and for bulk properties of liquids and molecular crystals. The dispersion interaction is one of the four basic interactions types – electrostatics, induction, dispersion and exchange repulsion – of which all WMIs are composed. The fact that each class of basic interactions covers a wide range explains the large variety of WMIs. To some of them, special names are assigned, such as hydrogen bonding or hydrophobic interactions. In chemistry, these WMIs are frequently used as if they were basic interaction types. For a long time, dispersion was largely ignored in chemistry, attractive intermolecular interactions were nearly exclusively attributed to electrostatic interactions. We discuss the importance of dispersion interactions for the stabilization in systems that are traditionally explained in terms of the "special interactions" mentioned above. System stabilization can be explained by using interaction energies, or by attractive forces between the interacting subsystems; in the case of stabilizing WMIs, one frequently speaks of adhesion energies and adhesive forces. We show that the description of system stability using maximum adhesive forces and the description using adhesion energies are not equivalent. The systems discussed are polyaromatic molecules adsorbed to graphene and carbon nanotubes; dimers of alcohols and amines; cellulose crystals; and alcohols adsorbed onto cellulose surfaces.

Introduction

Any change of the state of motion of a particle, described in an inertial frame, is caused by a force acting on the particle. The change of motion, i.e., the acceleration, causes a change of the position of the particle in space. If in a system of particles all particles exert forces on each other, these forces are called internal forces. In the simplest system, particle A at position $\mathbf{r}_A$

exerts a force $\mathbf{F}_{A\rightarrow B}$ on particle B and particle B at position $\mathbf{r}_B$ exerts a force $\mathbf{F}_{B\rightarrow A}$ on particle A. Both forces obey Newton's third law, $\mathbf{F}_{A\rightarrow B} = -\mathbf{F}_{B\rightarrow A}$, expressed by the Latin phrase "actio est reactio". This process of mutually exerting internal forces is called an interaction. If the internal forces are conservative, a potential function $V_{AB}(\mathbf{r}_A,\mathbf{r}_B)$ can be defined, and both internal forces can be calculated as gradients with respect to the position of the particles. With the advent of the science of energetics, as promoted by William Rankine, forces were nearly completely replaced by potentials in the description of interacting systems, the term "force" remained as a synonym for interaction. When we speak of a force in this paper, however, we always mean the physical vector quantity or its magnitude.

The obvious advantage of using the scalar quantity, energy, instead of the vector quantity, force, is that it is simpler to describe and categorize system stabilization by using properly defined stabilization energies calculated as differences in the values of the energy functions. In chemistry, stabilizing interactions are roughly classified as strong or weak according to the magnitude of stabilization energies [1]. Strong interactions are 1) Coulomb interactions in ionic solids ranging between 600 kJ/mol (CsI) and 3900 kJ/mol (MgO), 2) covalent interactions in molecules ranging between about 150 kJ/mol (I_2) and 950 kJ/mol (N_2), and 3) metallic interactions ranging between 65 kJ/mol (Hg) and 850 kJ/mol (W). Ionic and metallic interactions are the interactions in extended systems, mostly solids, whereas covalent interactions are between molecular subsystems (fragments, radicals) at localized positions, mostly atom positions. Interactions between atoms or small molecules with closed-shell electron configurations having stabilization energies of up to 50 kJ/mol are typical weak interactions. They are smaller by a factor of roughly ten than Coulomb interactions or covalent bonding.

It is a characteristic of attractive, weak molecular interactions (WMIs) that the molecules involved retain their integrity. This may mean one of three things: 1) that the geometries of the interacting molecules differ very little from the equilibrium geometries of the isolated species, e.g., an interacting molecule changes only its conformation; 2) that the neutral molecules do not undergo an electron-transfer interaction to form cation–anion pairs; or 3) that there is no significant change in the electronic structure of the interacting molecules, such as that caused by electronic excitation or covalent bonding. The absence of covalent bonding (case 3) is also the reason for using the term "non-covalent interaction", another frequently used term is "weak intermolecular interaction". Both terms have disadvantages. Weak intermolecular interaction does not cover those cases in which intramolecular interactions cause stabilization, e.g., when a large *n*-alkane changes from the linear to the hairpin structure; non-covalent interaction, on the other hand, does not exclude creation and stabilization of cation–anion pairs or zwitterions and their stabilization by Coulomb interaction. Weak molecular interaction is certainly the best term for describing any attractive interaction in which the interacting subsystems retain their integrity.

WMI does not have a single physical cause. Instead, several basic interactions are responsible for the interactions between molecules, which can be seen as extended charge distributions consisting of nuclei and electrons. When interactions between saturated molecules in their electronic ground states are considered, there are four basic interactions: 1) electrostatics, which are the interactions between static multipoles without any charge shift in the interacting molecules; 2) induction or polarization interactions, which are those between static multipoles in one and multipoles in the other molecule that are induced by charge shifts; 3) dispersion interactions, which are those between non-static multipoles in one molecule and induced multipoles in the other molecule; and 4) exchange repulsions or Pauli repulsions, which describe the tendency of electrons to avoid coming spatially close due to their Fermion character, not due to their charge [2]. Electrostatics and induction can be explained with classical physics, whereas dispersion and exchange repulsions are pure quantum effects. Induction and dispersion are also called polarization interactions, because both involve polarizations in at least one interacting molecule. An intricate aspect of WMI is that the four basic interactions may contribute with different weights; moreover, in each group, different "flavors" can be found due to the different distance dependencies of the various multipole–multipole interactions. The WMI for a certain pair of interacting molecules is like a cocktail composed of four basic ingredients, the characteristics of the cocktail are due not only to the different bar measures of the basic ingredients, but also due to their different flavors.

Hydrogen bonding is a typical WMI. As such, it is composed of the abovementioned basic interactions, each having its own strength and range. Nevertheless, it is common practice in chemistry to speak about hydrogen bonding as if it was indeed a genuine basic interaction rather than a composed interaction. Instead of stressing the different compositions of the basic interactions, chemists speak of strong, moderate or weak hydrogen bonding [3]; sometimes even further divisions are made [4]. Hydrogen bonding was introduced nearly 100 years ago to explain the stabilization of complexes of, e.g., water, alcohol or amine molecules. The stabilization was first explained solely by electrostatic attraction, but this simplistic view was already corrected in 1952 by Coulson [5], who stressed the need to also consider induction and dispersion as attractive interactions.

Nevertheless, even today it is more often claimed than actually demonstrated that hydrogen-bonded complexes are predominantly stabilized by electrostatics [6]. If any other interaction but electrostatics is considered, it is "charge-transfer", which suggests that the dimer stabilization is caused by an electron transfer, although this would mean the creation of a cation–anion pair and, thus, a loss of molecular integrity. What is meant, however, is a polarization of the electron density due to a charge shift, which is covered by the basic induction interactions [2]. Although dispersion interaction is a ubiquitous attractive interaction, it is frequently considered to be less important than electrostatics when explaining hydrogen bonding. However, we have shown that this is not the case in our studies on the stabilization of alcohol and amine dimers [7,8].

Another type of WMI is the hydrophobic interaction, which was introduced by Kauzmann [9] to explain protein folding in analogy with the transfer of a non-polar solute from water into a non-polar solvent. This process was attributed to the poor solubility of the solute in water. Wolfenden and Lewis [10], on the other hand, assumed "that a strong favorable interaction among alkane molecules in liquid alkanes gives a strong favorable transfer energy for passage of an alkane from vapor into liquid alkane", explaining the poor solubility of hydrocarbons in water and the good solubility of alkane molecules in liquid alkane [11]. Nonetheless, this interaction is nothing more than a dispersion-dominated WMI.

On the other hand, electrostatic interactions are often ignored, unless the interacting molecules have obvious dipolar structures. For example, the fact that there is electrostatic interaction between the quadrupoles of benzene molecules is mostly ignored or not even known. Instead, attraction is attributed to π–π interactions or CH–π interactions of unclear physical origin. That deformation of molecules induces static multipoles is also not well known; the bending of non-polar planar molecules that have a quadrupole as their lowest static multipole (e.g., polyaromatic hydrocarbons) induces a dipole moment; likewise, when a spherical charge distribution is deformed to an ellipsoid, a quadrupole is induced. Discussion of WMI, as found in the chemical literature, often suffers from a profound confusion of tongues due to the preference of a folkloristic [12] instead of a physically sound language.

With respect to extended systems, one has to consider an important modification of the theory of WMI. The standard calculation of the contributions to WMI is based on the multipole expansion of the charge distributions involved with respect to a single expansion center. This is justified for small molecules, but this expansion slowly converges or fails for large molecular systems. In molecular orbital theory, the slow convergence of single-center expansions of molecular orbitals, which is mathematically equivalent to the multipole expansion, was cured by the use of atom-centered basis functions in the linear combination of atomic orbitals (LCAO) approximation. This approximation allows the expansions to be stopped at much smaller angular momentum quantum numbers than in a single-center expansion. In the context of WMI, replacements of single-center expansions by multicenter expansions are termed distributed multipole analysis, distributed polarizabilities, and distributed dispersion interaction [2]. The possibility of calculating electrostatic, induction and dispersion interactions by dividing molecules into subsystems, mostly atoms or atom groups, which are characterized by their own short multipole expansion, together with the short range of attractive induction and dispersion interactions in particular, explains our findings of an approximate additivity of the stabilization energy and the adhesive forces [13-15]. Adhesion is the term for the attractive interaction between unlike subsystems, e.g., a graphene sheet and adsorbed molecules, whereas the attractive interaction between like subsystems, e.g., graphene sheets in graphite, is called cohesion. Nonetheless, the basic interactions are the same for adhesion and cohesion.

We attributed the additivity to the "near-sightedness" of WMIs, and defined the contact zone of two interacting molecules as the set of all atom pairs making non-negligible contributions to the adhesion energy and adhesive forces. We showed that the contact zone is a useful means for discussing the origin of stabilization of parallel alkane chains, as well as the stabilization of aromatic molecules adsorbed to graphene or carbon nanotubes. Furthermore, we found that the stabilization energy of an adsorbent and several small adsorbate molecules increases when the latter are in close contact with each other. This cooperative effect agrees well with the approximate isotropy of dispersion interactions.

In this paper, we discuss the implications of WMIs on structure and stability of different systems we studied in the past. We discuss the physical origin of WMIs, that is, their composition of different basic interaction types laying the focus on the role of dispersion interactions. We show that dispersion interactions are essential for the correct description of the structure and stability of systems composed of subsystems, such as dimers or clusters of small molecules, or interfaces between large adsorbents and adsorbates of different sizes. We discuss the different roles of adhesion energies and adhesive forces and friction forces for the description of the stability of condensed matter systems, and we show that use of the vector quantity force is essential for the understanding of mechanical stability of solids, and for many properties such as boiling point or viscosity of liquids.

Basics of Weak Molecular Interaction

Description of interaction through forces and potentials

Interactions in a system consisting of two or more subsystems cause changes of the spatial positions of the subsystems relative to each other. Attractive interactions reduce the distance between the centers-of-mass of two subsystems, whereas repulsive ones increase the distance. There may also be changes in the relative orientation of the subsystems due to rotations without any change in the distance between the centers of mass. The internal forces that each subsystem exerts on the others change their atomic positions; thus changes of the atomic positions are an indicator of interactions in the system.

For the moment, we assume that the two subsystems are structureless and completely described by the center-of-mass coordinates $\mathbf{R}_A$ and $\mathbf{R}_B$, the structure of the total system is presented by $\mathbf{R} = (\mathbf{R}_A, \mathbf{R}_B)$. Forces $\mathbf{F}_{A\to B}$, exerted by subsystem A on subsystem B, and $\mathbf{F}_{B\to A}$, exerted by subsystem B on subsystem A, depend in general on both subsystems, $\mathbf{F}_{A\to B} = \mathbf{F}_{A\to B}(\mathbf{R}_A, \mathbf{R}_B)$. Each is the negative of the other, the relation $\mathbf{F}_{A\to B}(\mathbf{R}_A, \mathbf{R}_B) = -\mathbf{F}_{B\to A}(\mathbf{R}_A, \mathbf{R}_B)$, expressing Newton's third law, can also be written as $\mathbf{F}_{A\to B} + \mathbf{F}_{B\to A} = 0$. Since this relation defines balanced forces, all internal forces are balanced forces.

An alternative way of describing interaction in a system uses a potential energy function (PEF) $V^{\text{int}}(\mathbf{R}_A, \mathbf{R}_B)$, called the interaction potential, for structureless subsystems, the potential depends only on the distance $r = |\mathbf{R}_B - \mathbf{R}_A|$ between the particles, $V^{\text{int}}(\mathbf{R}) = V^{\text{int}}(r)$. The internal forces are the negative gradients of the PEF with respect to the center-of-mass coordinates, $\mathbf{F}_{A\to B}(\mathbf{R}_A, \mathbf{R}_B) = -\nabla_A V^{\text{int}}(\mathbf{R}_A, \mathbf{R}_B)$ and $\mathbf{F}_{B\to A}(\mathbf{R}_A, \mathbf{R}_B) = -\nabla_B V^{\text{int}}(\mathbf{R}_A, \mathbf{R}_B)$. All elementary electrostatic potentials are strictly monotonic functions in r, such as reciprocal powers or exponentially decreasing functions, and they obey the asymptotic boundary condition $\lim_{r\to\infty} V(r) = 0$. For all finite values of r, they have either only positive or only negative values. Monotonically decreasing PEFs represent repulsive interactions, monotonically increasing PEFs represent attractive interactions.

PEFs $V^{\text{int}}(r)$ describing realistic molecular interactions, also called effective potentials, are always a sum of elementary attractive and repulsive components, $V^{\text{int}} = V^{\text{rep}} + V^{\text{att}}$, not all of them need be true potentials. In general, effective potentials have a local minimum at r_{equ} and are, accordingly, not monotonic, however, they always have a repulsive branch left of the local minimum and an attractive branch right of it. Furthermore, they obey the asymptotic boundary condition. Examples are the Lennard-Jones potential or the Morse potential, see Figure 1. Because of the asymptotic boundary conditions, the constant interaction energy for large r is chosen as zero. Any system geometry $\mathbf{R}^{\text{diss}}$ with $V^{\text{int}}(\mathbf{R}^{\text{diss}}) = 0$ represents the dissociated system, and the energy difference $\Delta V = V^{\text{int}}(\mathbf{R}^{\text{diss}}) - V^{\text{int}}(\mathbf{R}^{\text{equ}}) = -V^{\text{int}}(\mathbf{R}^{\text{equ}})$ is the adhesion energy.

If the potential depends only on the distance r between the particles, $V^{\text{int}} = V^{\text{int}}(r)$, the internal forces are central forces and automatically obey Newton's third law. The first derivative or the slope function of $V^{\text{int}}(r)$ is the negative force function, $-F(r) = [V^{\text{int}}(r)]'$. In this paper, we will always show PEFs together with their first derivatives instead of the force functions:

$$\mathbf{F}_{A\to B}(r) = -\nabla_A V^{\text{int}}(r) = -\left[V^{\text{int}}(r)\right]' \frac{\mathbf{r}}{r} = F(r)\frac{\mathbf{r}}{r}$$
$$= \nabla_B V^{\text{int}}(r) = -\mathbf{F}_{B\to A}(r).$$

Because the interaction potential is the sum of attractive and repulsive components, the same is true for the internal forces, which are the sum of attractive and repulsive components,

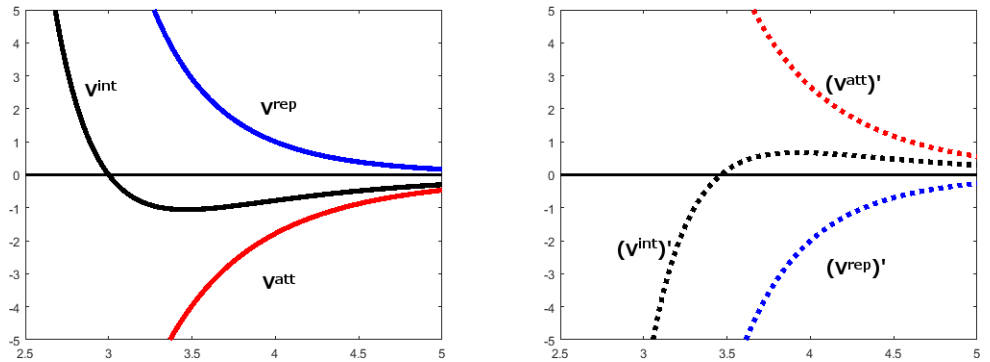


Figure 1: Left: The graphs of an interaction potential V^{int} composed of an attractive component V^{att} and a repulsive component V^{rep}. Right: The corresponding slope functions.

$F(r) = -[V^{\text{int}}(r)]' = -[V^{\text{rep}}(r)]' V^{\text{rep}} - [V^{\text{att}}(r)]'$. The first derivatives of the components are also monotonic and they obey the asymptotic boundary conditions.

For the description of interaction between subsystems, the forces corresponding to the attractive and the repulsive branch of the interaction potential are more important than the force components. For all distances $r < r_{\text{min}}$, that is for the repulsive branch of $V^{\text{int}}(r)$, the force function has positive values, $F(r) > 0$ and the internal forces are repulsive. For the distances of the attractive branch, $r > r_{\text{min}}$, the force function has negative values, $F(r) < 0$, and the internal forces are attractive. The attractive branch of the interaction potential $V^{\text{int}}(r)$ has an inflection point at r_{infl}, where the slope function has a maximum. The maximum internal force is equal to the negative slope at the inflection point, $F_{\text{max}} = F(r_{\text{infl}})$. In a complex of interacting molecular subsystems, attractive internal forces are called adhesive forces. At the local minimum of $V^{\text{int}}(r)$ the force function $F(r)$ has a zero because the non-zero repulsive and attractive components of the internal force are equal in magnitude and, therefore, cancel out each other. For large distances r, that is for the dissociation of the system, the internal forces become zero because both force components become zero.

To separate subsystems from each other, an external force, called a pull-off force, must act on a subsystem and pull it off the other one. The point at which a pull-off force acts on the subsystem is called the pull-off point. External forces do not necessarily occur in pairs; thus, they are not genuinely balanced. Whenever a pull-off force acts on a system in its equilibrium the latter responds by inducing a pair of adhesive forces, see Figure 2.

Hence, both an external and an internal force act on the pull-off point, but in opposite directions. Stretching stops as soon as the adhesive forces are equal in magnitude to the pull-off force. Then, the pull-off force and the adhesive force are balanced and the system is in a new, stretched equilibrium structure. However, if the external force is larger in magnitude than the maximum adhesive force, the system dissociates and there is no stabilizing adhesive force. The maximum adhesive force therefore provides another measure of the system stability, which may differ considerably from that using the stabilization energy, ΔV. After all, F_{max} depends not only on ΔV, but also on the curvature of the potential curve at the minimum, see Figure 3. Therefore, interaction potential curves with the same ΔV can have different $\mathbf{F}_{\text{max}}$, or as Israelachvili says: ".. a bond may have a high bond energy, but a low force needed to break it. Thus, simply talking about the 'strength' of a bond may not mean anything" [6].

Theoretical methods for the description of weak molecular interaction

Interactions in molecular system cause spatial displacements in the subsystems due to changes of the geometries and changes of

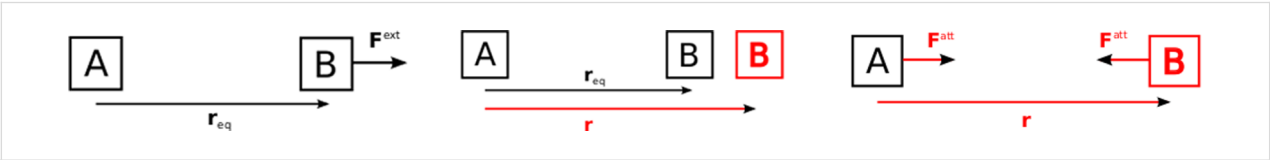


Figure 2: From left to right: An external pulling force acting on the system in its equilibrium structure increases the distance between the subsystems and induces an attractive internal force.

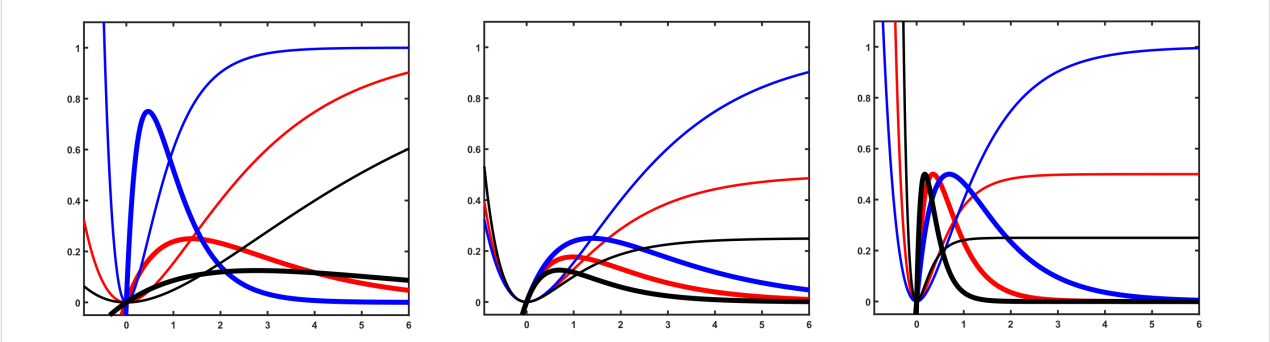


Figure 3: Potential functions (thin lines) and the first derivatives (thick lines). Left: For constant ΔV the maximum force decreases with decreasing curvature at the minimum. High curvature (blue), middle curvature (red), low curvature (black). Middle: For constant curvature the maximum force decreases with decreasing ΔV. Large ΔV (blue), middle ΔV (red), small ΔV (black). Right: Combination of large ΔV and small curvature may yield the same maximum force as small ΔV and large curvature.

the electron distributions. These intramolecular effects will, in turn, influence the intermolecular interactions between the distorted subsystems. Any interaction in a molecular system is the sum of intermolecular and intramolecular interactions. Molecules that strongly resist geometric distortions are called rigid or stiff, the resistance of electron distributions against distortion is called its hardness. By freezing the geometries of the interacting subsystems the interaction energy is calculated as if the interacting subsystems were ideally rigid. The intermolecular contributions to the interaction energy can be calculated in different ways.

In the supermolecule approach, the interacting system is treated as a large molecule and the stabilization energy is simply the difference between the energy of supermolecule E^{AB} and the sum of the energies of the isolated molecules E^{A} and E^{B}:

$$\Delta E = E^{AB} - \left(E^{A} + E^{B}\right). \quad (1)$$

The mutual deformation of the electron distributions of the interacting molecules in the supermolecule is caused by attractive electrostatic interaction between electrons and nuclei, by mutual repulsion of electrons due to the charge, called Coulomb correlation, and by the mutual influence of the electrons due to the spin, called Fermi correlation. The advantage of high-level electron-structure methods is that they cover all these contributions and that they allow one to calculate weak molecular interactions for large distances between the interacting molecules, as well as strong molecular interactions when the molecules come very close. The disadvantage is that they are costly and, to explain the physical origin of the interaction energy, one has to split up the energy difference into physically meaningful contributions, which cannot be done in a unique way.

An alternative way of calculating the interaction energy is to make a multipole expansion of the interaction potential $V_{AB}(\mathbf{r}_A, \mathbf{r}_B)$ for the supermolecule and to calculate the energy contributions using perturbation theory:

$$V_{AB}\left(\mathbf{r}_A, \mathbf{r}_B\right) = \sum_{I=1}^{n_A} \sum_{J=1}^{n_B} \frac{q_I q_J}{r_{IJ}}, \quad (2)$$

where r_{IJ} is the distance between particles I and J. $V_{AB}(\mathbf{r}_A, \mathbf{r}_B)$ represents the interaction between the charge distributions of molecules A and B due to both the nuclei and the electrons. The nuclei are assumed to be point charges in space whereas the electrons form a continuous charge distribution. According to classical electrostatics, this gives rise to two basic contributions, called electrostatics and induction, also called polarization. However, there are purely quantum theoretical contributions to the interaction energy, both have to do with electron correlation. These two basic interactions are called dispersion interaction and exchange repulsion. With this approach, the interaction energy can be calculated at different orders, for all contributions a physical interpretation can be given. An obvious disadvantage of this approach is, however, that the multipole expansion can be done in different ways, and that the multipole terms have singularities when the distance between the expansion centers goes to zero.

Basic interactions

Range of interactions

For interactions that only depend on the distance r between the interacting particles, and that can be represented by discontinuous model potentials, one can define the range of the interaction as the length of the interval of r values for which the interaction energy is negative. This definition is convenient for hard-sphere model potentials with a rectangular potential well, but it is less useful for continuous interaction potentials only going to zero for infinite distances. When the definition of range is based on forces, the mentioned hard sphere potentials are less useful because the derivative of such a potential is non-zero only at the discontinuities; that is, at the borders of the intervals where the interaction energy is negative. Anywhere else, the forces are zero. For continuous potentials, one can define the range as the length of the interval for which the potential or the force is significantly larger than zero. Using the extension of the Yukawa potential to general screened potentials, $V(r) \propto e^{-(r/r_0)} / r^n$, with a power $n \geq 0$, allows interactions to be classified as being of infinite range when $r_0 \to \infty$, otherwise they have the finite range r_0. According to this definition, all potentials depending on powers of the inverse distance, e.g., all electrostatic, induction and dispersion interactions, are of infinite range, whereas exchange repulsion is of finite range. For all functions of infinite range, the power n can be used to distinguish between shorter and longer ranges: the smaller n, the longer the range. Another caveat by Israelachvili is the following: "It is […] wrong to associate long-range *effects* with long-range *forces*. In fact, the opposite is usually the case – for what is more important is the strength of the interaction, and […] short-range forces tend to be stronger than long-range forces" [6].

Exchange repulsion

Exchange repulsion, or Pauli repulsion, is a consequence of the Pauli exclusion principle, which states that Fermions avoid coming spatially close to each other. Thus, exchange repulsion has an enormous impact on the spatial distribution of electrons in molecular systems. The effect of keeping electrons at a distance "plays the role of a fictitious, although highly effective, mutual repulsion being exerted within the system, irrespective

of any other actual forces of interactions […] that might be present" [16]. Exchange repulsion is a nonlocal effect of purely quantum origin, it is ubiquitous and it is fundamental. As Lennard-Jones wrote 1954: "This effect is most powerful, much more powerful than that of electrostatic forces. It does more to determine the shapes and properties of molecules than any other single factor. It is the exclusion principle which plays the dominant role in chemistry" [17]. Exchange repulsion can be described by a repulsive potential–energy function with exponentially decaying dependence on the interatomic distance [18,19]. Its representation by a potential–energy function is similar to the use of local, repulsive "pseudo"-potentials. The assumption that exchange repulsion between any two molecules can be represented by a single exponential is not justified, there need to be more.

Electrostatics

Electrostatics is the classical interaction between static electric multipoles, which are obtained by a multipole expansion of the charge distribution of a molecule about a convenient expansion point, usually the center of mass. Static multipoles 2^l of rank $l = 0, 1, 2,\ldots$ are monopoles ($l = 0$), dipoles ($l = 1$), quadrupoles ($l = 2$), and so on. The interaction potential for the interaction between an l-pole and an L-pole has a distance dependence of $1/r^{l+L+1}$. The higher the the multipoles, the shorter the range of interaction. The Coulomb interaction, i.e., the interaction between electric monopoles, that is, charges, has the longest range. The interaction between static multipoles may be attractive or repulsive. The sign of the Coulomb interaction depends only on the signs of the charges. If at least one higher multipole is involved, the interaction also depends on the relative orientation of the multipoles, meaning that it can be attractive, repulsive or that there is no interaction at all. The interaction potential between an l-pole and an L-pole can be quite generally written as

$$V_{lL}^{\text{int}}(r) = M_l M_L \cdot 1/r^{l+L+1} f(\theta_l, \phi_l, \theta_L \phi_L), \qquad (3)$$

where M_l and M_L are the magnitudes of the l-pole and the L-pole, and $f(\theta_1,\phi_1,\theta_2\phi_2)$ is the geometric factor describing the relative orientation of the two multipoles with respect to the line connecting the centers of mass using local spherical polar coordinates. The product of the magnitudes of the multipoles is often used as a measure of the strength of interaction, which is modulated by the angular dependence of the geometric factor. The product $M_l M_L \cdot 1/r^{l+L+1}$ must have the physical dimension energy, the geometric factor is a bare number. The strength of Coulomb interaction is proportional to the product of the charges.

Although any spatial charge distribution can be expanded into a series of multipoles, the number of terms can be large when the symmetry of the charge distribution is low. One can avoid working with high-rank multipoles when the single-center expansion is replaced by a multicenter expansion, also called a distributed multipole expansion, in which several meaningful expansion centers are chosen, for example the positions of the nuclei in a molecule or the centers of mass of atom groups. Each expansion then contains only few multipoles. Regardless of whether single-center or multicenter expansions are used, the electrostatic interaction energy becomes singular only when the distances between the expansion points become zero. For extended charge distributions, the multipole expansion of the electrostatic interaction energy is in error as soon as the charge distributions overlap. Classical electrostatics shows that the interaction energy for extended charge distributions is much smaller in magnitude than that of point-multipoles. Correctly calculated electrostatic interaction energies do not have singularities. The difference between the interaction energy for extended charge distributions and the multipole expansion is called the penetration error. It can be corrected either by using damping functions or by applying a penetration error correction [20]. However, one should not overestimate the physical significance of this correction, the overlap of hard electron distributions is purely fictional, after all, both Fermi and Coulomb interactions are not considered.

Induction

The second class of classical interactions covers those between the static electric multipoles in molecule A and the induced multipoles in molecule B. The latter are the result of charge shifts (polarization) in the polarizable electron distribution of molecule B. The strength of the interaction is measured by the respective static polarizabilities, which describe the ability of polarizable systems to create induced multipoles under the influence of nonuniform electric fields. The interaction between a static l-pole and an induced L-pole has a $1/r^{2(l+L+1)}$ distance dependence, and again it depends on the relative orientation of the multipoles. Induction interactions are therefore always of much shorter range than the interactions between the corresponding static multipoles. The polarizability that describes the interaction between a static and an induced dipole is called the dipole–dipole polarizability. Likewise, for the interaction between a static dipole and an induced quadrupole the dipole–quadrupole polarizability is responsible, and so on. The interaction with the longest range is again the dipole–dipole interaction. However, at short distances, e.g., the equilibrium distance, the short-range interactions become important. Unfortunately, the corresponding polarizabilities are seldom tabulated. After all, they are tensor quantities, and one cannot infer from tabulated dipole–dipole polarizabilities whether or not the

higher short-range interactions are important or not. As for electrostatic interactions, the induction energy at short distances between the multipoles is in error for point-multipoles, but can be corrected when damping functions are used [2].

Dispersion

The dispersion interaction is a ubiquitous interaction of purely quantum origin. It is a kind of dynamic electron correlation, and between ground-state molecules it is always attractive. Its description is far from simple [21]. A catchy albeit disputable explanation starts with short-time deformations of the electron density of one molecule caused by the non-deterministic motion of electrons. These fluctuations are represented by spontaneously created multipoles that will interact with induced multipoles in the electron distribution of the other molecule. Any nonsymmetric deformation leads at least to a dipole. The induced multipole of lowest rank in the other molecule is again a dipole. The $1/r^6$ distance dependence of the corresponding dipole–dipole dispersion interaction again has the longest range and is the leading contribution at large distances. At shorter distances, higher-order dispersion interactions of much shorter range are again important. For two interacting atoms, the interaction energy is isotropic because it depends only on the interatomic distance. For molecules, an effective isotropic dispersion interaction follows from averaging over all relative orientations of the multipoles. The distance dependence of the dispersion interactions is the same as that of the corresponding induction interactions. The strength of the interaction between atoms is proportional to the product of the dynamic polarizabilities [2]. It is much larger between noble-gas atoms from the higher periods, than between atoms with hard electron densities [22,23]. For molecules, one has to consider the anisotropy of molecular polarizabilities, which is strongly pronounced for molecules with delocalized pi-electron densities, the polarizability component along the molecular axis, that is the polarizability of the p-electrons, is always considerably larger than the components orthogonal to it [24,25]. Dispersion energies calculated with this method at short distances between the multipoles are in error. Again, damping functions help to avoid these errors.

Note, that the $1/r^6$ distance dependence does not hold for atoms or small molecules interacting with extended metals or perfect graphene, for such systems one finds a $1/r^3$ distance dependence [22-24].

Combination of the basic interactions

For non-charged systems with spherical electron distribution (atoms), there are no electrostatic or induction interactions. There are only dispersion interactions, starting with the long-range dipole–dipole interaction. This interaction exists between any two molecular systems. Every non-charged and non-spherical molecule has static multipoles of different ranks, in polar molecules, the series starts with dipoles, whereas in non-polar molecules it starts with quadrupoles or higher multipoles. Accordingly, there will always be electrostatic interactions of different ranges between molecules with static multipoles. For example, the T-shaped equilibrium structure of the benzene dimer is favored by the geometric factor of the quadrupole–quadrupole interaction [2]. Every molecular system has a polarizable electron distribution, in which multipoles can be induced. Therefore, if at least one subsystem has static multipoles of any rank, there will be induction interactions.

If we combine the products of the magnitudes of the multipoles and the geometric factors to prefactors P_n, the interaction potential for two uncharged molecules can be written as a series

$$\begin{aligned} V^{\text{int}}(r) &= V^{\text{exc}}(r) + V^{\text{elst}}(r) + V^{\text{ind}}(r) + V^{\text{disp}}(r) \\ &= \sum_i P_i^{\text{exc}} e^{-a_i r} + \sum_{n=3}^{\infty} P_n^{\text{elst}} \frac{1}{r^n} \\ &\quad + \sum_{n=6}^{\infty} P_n^{\text{ind}} \frac{1}{r^n} + \sum_{n=6}^{\infty} P_n^{\text{disp}} \frac{1}{r^n}. \end{aligned} \tag{4}$$

The number of terms in the series that contribute significantly depends on the magnitude of the corresponding prefactors and also defines the “flavor” of the interaction.

Many-body effects

Many-body systems [25] are composed of particles of different kind. Each particle interacts with all others, that is, all particles are highly correlated, otherwise one would have many one-body systems. The nature of the particles depends on how the systems is modeled. In an electron gas the particles will be electrons. If an atom is regarded as a many-electron system, they will be electrons. If a molecule is regarded as being composed of atoms, the particles will be atoms, but if the molecule is modeled as a many-electron system, the particles will be electrons again. In a liquid or a molecular crystal, the particles may be molecules, they may be the atoms or they may be electrons. Frequently, many-body systems behave as if the particles interact only weakly or do not interact at all. But these particles are not the real, strongly interacting particles but fictitious particles, called quasi-particles. Calculating the energy of the many-body system by summing up the interactions between all real particles is impossible. Weak interactions between quasi-particles can, however, be calculated using conventional techniques, e.g., perturbation methods. A simple introduction to the idea of quasi-particles goes as follows: All interacting particles are in motion, so any particle may interact with two or more other par-

ticles at the same time, and a certain interaction may occur repeatedly in a certain time interval. A strategy for defining quasi-particles is to identify and keep a few of the most important interaction types between the real particles and to neglect all others. Furthermore, it is assumed that it is easy to calculate the sum of all repeated occurrences of these interactions. By adding these partial sums of interactions with the other particles to the real particle it is transformed into a quasi-particle. Some properties of quasi-particles may be different from those of real particles, others are unchanged. A good introduction to quasi-particles in many-body systems can be found in the book by Mattuck [26]. Depending on what kind of real particle is transformed into a quasi-particle, different properties are of interest and different theoretical quantities are used to describe them. For example, electrons as described by Hartree orbitals, Hartree–Fock orbitals or Kohn–Sham orbitals are quasi-particles. They exhibit a different extent of interaction with other electrons, they have different (orbital) energies but the charge is not changed. (Quasi)-atoms in a molecule have, for example, volumes and polarizabilities that differ from those of free atoms in the gas phase. Properties of such quasi-atoms are often calculated by using propagators (Green's functions) or response functions that were obtained by using a special summation of important interactions, for example by using the random phase approximation. The weak interactions not absorbed into quasi-particles are calculated as interactions between quasi-particles. The magnitude of these interactions depends essentially on the way the quasi-particles are created. Frequently, it is assumed that the weak interactions are dominated by pair contributions, and that interaction between three or more quasi-particles can be reduced to sums of pair interactions (additivity of interaction). Whether or not this assumption is justified depends on the many-body system, and on the extent to which the interaction between the real particles is included in the quasi-particles.

Many dispersion-correction strategies assume pair-wise additivity of the long-range electron correlation energy. The properties of the quasi-atoms may be obtained by fitting them to interaction energies calculated with other high-level methods. This strategy is used, for example, in Grimme's D2 method [27], for the calculation of the dispersion energy in the dlDF+D method by Szalewicz [28,29], and for the dispersion correction to DFTB [30-33]. In the Tkatchenko and Scheffler (TS) method [34], the C_6 coefficients for atoms in a molecule are set proportional to those of the corresponding free atoms. The proportional constant is a function of the ratio between the volume of the free atom and the Hirshfeld volume of the atom in the molecule. According to Dobson [35], one can distinguish between three different types of non-additivity of dispersion interactions. Type-A non-additivity originates from the fact that the dispersion coefficients of free atoms are different from those of atoms in molecules. This type of non-additivity is captured for example by the TS model and Grimme's D3 method [36] by employing environment-dependent dispersion coefficients. Type-B non-additivity occurs, when the interaction between two particles is screened by a third particle, giving a three-center angularly dependent interaction contribution. The most simple three-body term is a triple–dipole contribution, the so called Axilrod–Teller–Muto term, which, because of the angular dependence, can give attractive and repulsive contributions. This three-body correction is included in Grimme's D3 method. When N perturbing particles are considered, one gets N-center contributions. In diagrammatic many-body theory, interactions of that kind are represented by ring diagrams [26], summation of ring diagrams to infinite order gives the correlation energy in the random phase approximation. Type-C effects, according to Dobson's classification, can be found in nanostructures of low dimensionality with degenerate electronic ground states where any perturbation causes delocalized density fluctuations or density waves, also called collective excitations [25,26]. Often they are found in one- or two-dimensional structures such as graphene or metallic nanotubes with easily polarizable electron densities, and they are less frequently found in three-dimensional metals [35,37,38]. Delocalized density fluctuations allow for the induction of large dipoles or higher multipoles that enhance weak molecular interaction: It is characteristic of interactions between such extended density waves that the range of the interactions is much longer than that of dispersion interactions between localized structures [35]. Dispersion interactions are a type of electron correlation, but dispersion interaction is not a synonym for electron correlation. Therefore it is clear that there must be other types of electron correlation beyond dispersion interaction. It is also clear that there are many different types of collective motions in extended systems [25]. It should not be surprising that interactions between different density fluctuations may have different ranges.

Different strategies can be used for improvement of the description in many-body systems. One is to go beyond the triple-dipole term in the calculation of three-body energies, formulas for the dipole–dipole–quadrupole or dipole–quadrupole–quadrupole terms are given, e.g., in the book by Salam [21]. Due to the distance dependence of these terms they are only significant at short range, and they are strongly anisotropic [2]. Another strategy is to keep the description of the interacting atoms in molecular systems as simple as possible but to include the interaction between many of these atoms. This route is followed in the many-body-dispersion (MBD) method by Tkatchenko and co-workers [34,39,40]. The atoms are considered to be isotropic, oscillating charge distributions represented by 3D harmonic oscillators, the polarizabilities are obtained with the TS method. Interaction between the atoms considered as

vibrating dipoles yields screened atomic polarizabilities that are finally used to calculate long-range correlation energies from diagonalizing the Hamiltonian of the coupled oscillators with the screened polarizabilities as input. Although by this procedure many-body contributions are captured that go beyond the three-body ATM term and improve, for example, cohesive energies considerably [37,41], some open questions concerning the calculation of correlation energies using the MBD method remain. For example, it is not yet clear how well fluctuating dipoles represent fluctuations of anisotropic charge in general, or whether molecular polarizabilities entering the expressions for dispersion interaction in the single-center expansion can be replaced by fragment polarizabilities, analogous to the multicenter expansion of charge distributions [38].

With respect to the calculation of adhesive forces, no detailed MBD studies are available, especially it is not clear, how strong the many-body effects change the shape and slope of the adhesion energy curves around the inflection point.

Range of electrostatic interactions

Electrostatic potentials of 2^l-poles depend on the distance according to $1/r^{l+1}$, and the electric fields depend according to $1/r^{l+2}$. High-rank multipoles can be approximately represented by multipoles of lower rank at different spatial positions, i.e., a dipole can be represented by two charges (monopoles), a quadrupole by two dipoles or four monopoles, and so on. But when this is done, one must not forget the correct distance dependence of the high-rank multipole–multipole interaction. Since the interaction between an l-pole and an L-pole is proportional to $1/r^{l+L+1}$, the interaction between a dipole and a charge is proportional to $1/r^2$. If this is ignored, one could believe that there is a Coulomb interaction between monopoles, which has, however, a $1/r$ distance dependence. That the field of spatially close charges has a different distance dependence than isolated charges far apart shows the electrostatic potential of an ionic crystal, which is composed of a large number of monopoles. The interaction between a test charge and, e.g., a rock-salt crystal, operates at very short distance, and not at distances as large as one might assume, considering the long range of Coulomb interactions. However, close to each charge in the crystal, there is a charge of opposite sign forming a dipole with a field that is proportional to $1/r^2$. Close to each dipole is another dipole and the resulting quadrupole field is proportional to $1/r^3$. Two quadrupoles close to each other form an octopole with a $1/r^4$ distance dependence, and so on. This means that the potential of an ionic lattice decays faster with r than any power of $1/r$, which means an exponential decay. The finite range of such a potential is smaller than the spacing between the ions in the crystal [6]. Elementary classical electrostatics shows, thus, that superpositions of low-rank multipoles with large range located at different positions in space are equivalent to high-rank multipoles with a much shorter range. But this is frequently ignored in chemistry, where, for example, interactions between two molecular quadrupoles ($1/r^5$ distance dependence) are reduced to interactions between bond dipoles having a $1/r^3$ distance dependence.

Contact zone

Interactions between atoms or finite molecules are dominated by pair contributions, even when many-body contributions are shown to be important, as, for example, in the case of the nonadditive induction interaction [2]. We will now consider the pair contributions to the long-range dipole–dipole dispersion interactions between molecules A and B with n_A and n_B atoms, respectively, which are used to define the contact zone (CZ) of atoms or interacting molecules:

$$
\begin{aligned}
V_{\mathrm{AB}}^{\mathrm{disp}}(r) &= \sum_{I=1}^{n_\mathrm{A}}\sum_{J=1}^{n_\mathrm{B}} V\left(r_{IJ}\right) = \sum_{I=1}^{n_\mathrm{A}}\left(\sum_{J=1}^{n_\mathrm{B}} V\left(r_{IJ}\right)\right) \\
&\approx \sum_{I=1}^{n_\mathrm{A}}\left(\sum_{J=1}^{\mathrm{CZ}(I)} V\left(r_{IJ}\right)\right) \\
&= \sum_{I=1}^{n_\mathrm{A}} V_I^{\mathrm{CZ}}.
\end{aligned}
\tag{5}
$$

For atom B_J in molecule B nearest to atom A_I in molecule A with pair distance $r = r_{IJ}$, the pair contribution to the dipole–dipole dispersion interaction is proportional to $1/r^6$. For all atoms B_J with a pair distance larger by a factor of f with $f > 1$, the pair contribution is reduced by $1/f^6$. Due to the sixth power, the magnitude of the pair contribution decreases strongly with increasing distance: When r increases by 10%, the pair contribution is reduced by 44%; when r increases by 50%, the pair contribution is reduced by 91%. The contributions to the attractive short-range dispersion drop even faster, as do the contributions to the exchange repulsion. For the $1/r^{12}$ term in the Lennard-Jones potential, if r increases by 20%, the interaction is reduced by 89%. Therefore, each atom in molecule A will "see" only few atoms from molecule B; the others can safely be neglected. Dispersion interactions, induction interactions and exchange repulsion are "near-sighted", as are electrostatic interactions between high multipoles. Whenever $f \geq 1.5$, that is, when the pair distance r is more than 50% larger than the equilibrium distance r_{equ}, the contributions will be close to zero. From here on, we will always speak of a distance $r = 1.5 r_{\mathrm{equ}}$ as the threshold value. All atoms B_J that give nonnegligible contributions to the interactions with atom A_I make the contact zone CZ(I) of atom A_I in molecule B. The sum over J in Equation 5 can thus be limited to the atoms in the CZ

causing only a small and acceptable loss of accuracy. The sum over atoms I shows that the CZs are approximately additive as a consequence of the "near-sightedness" of WMI.

The concept of near-sightedness of electrons was introduced by Kohn [42] in the description of many-atom systems, and "[i]t can be viewed as underlying such important ideas as Pauling's 'chemical bond', 'transferability'…" [43], about which Prodan and Kohn say: "Understanding the physics and chemistry of large molecules and solids would have been practically impossible if not for the principle of transferability" [43]. In the language of density functional theory, the concept of near-sightedness of electrons "…describes the fact that, for fixed chemical potential, local electronic properties, such as the density $n(r)$, depend significantly on the effective external potential only at nearby points. Changes of that potential, no matter how large, beyond a distance R have limited effects on local electronic properties, which rapidly tend to zero as a function of R" [43]. In their 2005 paper, Prodan and Kohn list what near-sightedness of electronic matter is not. For example, it is not screening of charges, as it applies also to neutral fermions, it "does not apply to systems of few electrons" and "it is not limited to macroscopically homogeneous systems" [43]. We explain the approximate additivity of dispersion interactions between molecular systems by a similar near-sightedness of WMI, caused by the short range of the basic interactions. The concept of near-sightedness of WMI is not the same as the near-sightedness of electrons, the distance R mentioned by Prodan and Kohn is different from our threshold value described above. Near-sightedness of electronic matter is of finite range, it explains why linear scaling in electronic structure methods works. The near-sightedness of the attractive basic interactions in WMI, on the other hand, is of infinite range, but it allows to understand the transferability of group contributions of, for example, pairs of CH_2 in two parallel aligned alkane chains. For the interaction of atoms or small molecules with extended metal surfaces, the concept of near-sightedness of dispersion does not apply, because the polarization of the metal due to the small interaction partner is not local, there are collective polarizations in the metal, rather than local ones [22,23].

For an atom, the shape of its CZ in a planar molecule is a disk that is the base of a cone with a lateral surface composed entirely of lines of length fr, which is the "vision cone" of the atom, see Figure 4. The disk contains all atoms B_J that make a contribution larger than $1/f^6$. If atom A_I interacts with atoms B_J of a curved molecule B, say a fullerene or a carbon nanotube, the CZ is smaller than when the molecule is planar. See the right-hand side of Figure 4. Of course, one can do the same with the roles of molecules A and B reversed. Therefore, the CZ of two interacting molecules can be defined as the set of all atom pairs contributing significantly to the interaction energy. This is in accord with the success of distributed multipole expansions of all basic interactions.

The maximum interaction energy is proportional to the size of the CZ at the equilibrium geometry of the complex. Any decrease in the size of the CZ brought about by increasing the distance between the interacting molecules reduces the interaction energy and reduces the adhesive forces in the complex. The change in the interaction energy, and therefore the magnitude of the adhesive forces, is proportional to the changing part of the CZ where the pair distances r_{IJ} increase and the adhesion energy decreases; this is the reduced contact zone [44], see Figure 5.

Other representations of basic interactions

A problem in speaking about WMI is that, in chemistry, often a stinted and frequently unphysical language is used. Although electrostatics and induction have very different ranges, induction is often, incorrectly, included under electrostatics, rather than being separately discussed. If induction is considered, it is described in terms of the dominant theoretical means mastered by chemists, namely orbitals. Polarization of the electron distribution of atoms manifests itself, for example, in an increase of the weight of the polarization functions in the occupied atomic

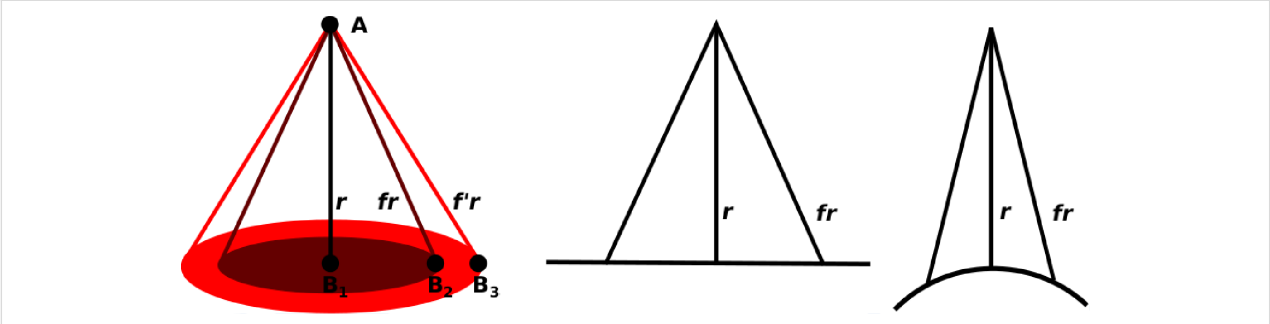


Figure 4: Left: The disk covering the atoms of molecule B seen by an atom in molecule A expands with increasing lateral height fr of the "vision cone". Right: The aperture of the vision cone becomes smaller when molecule B is not planar but bent.

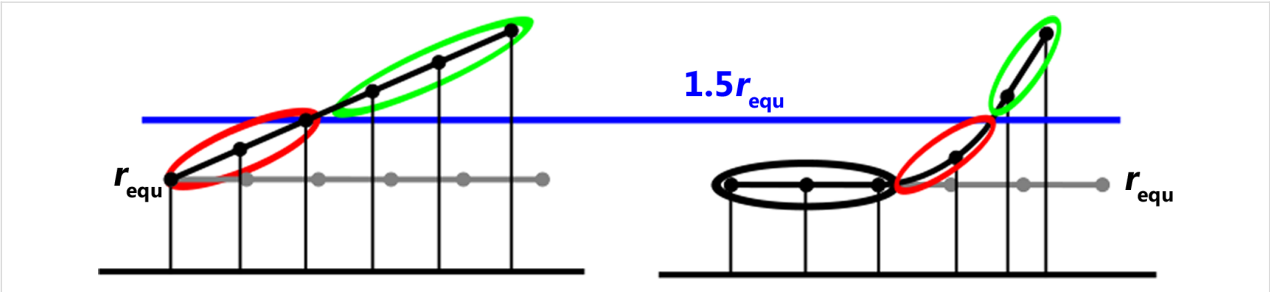


Figure 5: Demonstration of the contact zone and the reduced contact zone of an adsorbate/adsorbent complex with equilibrium distance r_{equ} during separation in modes S2 and S3. Left: In separation mode S2, all atom pairs, except the one with the pivot atom, change their distance r_{IJ}. Atom pairs with distances larger than the threshold value $1.5r_{equ}$ (adsorbate atoms encircled in green) do not contribute to the stabilization energy or to the adhesive forces. Non-zero contributions come from atom pairs in the reduced contact zone (adsorbate atoms encircled in red). Right: In separation mode S3, some part of the contact zone remains unchanged (adsorbate atoms encircled in black). These atom pairs contribute most to the stabilization energy, but not at all to the adhesive forces. The atom pairs with distances larger than the threshold value (adsorbate atoms encircled in green) do not contribute to the stabilization. Atom pairs in the reduced contact zone (adsorbate atoms encircled in red) contribute less to the stabilization energy than atom pairs from the (black) contact zone but they are the origin of the change in the adhesive forces.

orbitals (AOs). This could be shown, for instance, by adding p-type or d-type basis functions to occupied AOs having purely s-character in the unpolarized atom. This is nothing else than the hybridization of AOs. The molecular orbitals (MOs) of a complex of non-interacting molecules are, in general, linear combinations of the occupied fragment MOs, that is, the MOs of the isolated molecules. If such a complex MO is dominated by a fragment MO of one molecule, the complex MO is localized on that molecule. Induction or polarization will change the weights of the fragment MOs in the complex MOs. Localized complex MOs may then become delocalized, which is frequently called by chemists "charge transfer", and it is claimed that the charge-transfer interaction is an important, stabilizing interaction. Charge transfer, however, refers to an ion pair stabilized by a strong Coulomb interaction with a much larger stabilization energy than that of a weakly interacting system. Describing a charge shift in the electron density of a molecular system as a charge transfer incorrectly twists the semantics of the word transfer.

Quantum theory says that states of subsystems may interfere whenever the subsystem wave functions overlap significantly. Because the wave functions of atoms or molecules decay exponentially, this only happens at short distances between the subsystems. Ruedenberg et al. [45-51] showed that covalent bonding is a one-electron effect, and that the so-called accumulation of charge between the atoms connected by a covalent bond is a charge shift caused by constructive interference of exponentially decaying AOs or hybrid AOs. Thus, covalent bonding operates only at much shorter distances than those between weakly interacting molecules. At distances as large as those between weakly interacting molecules, the overlap of the molecular wave functions and the ensuing stabilization are very small, given that there is indeed constructive (and not destructive) interference of the many-electron state functions of the interacting molecules. Nevertheless, it is frequently claimed, but not proven, that strong covalent bonding is important for hydrogen bonding.

Rather curious are so-called orbital–orbital interactions such as π–π interactions, because orbitals are one-electron state functions, which do not interact but may be used to describe interacting states. However, it is never quite clear what kind of "interactions" they are describing. Are they describing static attractive multipole–multipole interactions between orbital contributions to the molecular electron densities, as Anthony Stone suggests [2]; or are they describing constructive or destructive interference of orbitals similarly as for the explanation of reactions using the Woodward–Hoffmann rules? Are they describing attractive dispersion interactions between the π-densities, or the exchange repulsion of π-densities?

Methods to describe WMI

WMI stabilization energies for interacting molecules A and B are calculated either with the supermolecule method or with perturbation methods. In the supermolecule approach, the interacting complex is treated as a supermolecule and the stabilization energy is simply the difference between the energy of the supermolecule E^{AB} and the sum of the energies of the isolated molecules E^{A} and E^{B}:

$$\Delta E = E^{AB} - \left(E^{A} + E^{B}\right). \quad (6)$$

The energies can be calculated with any high-level electron structure method. The Hamiltonian of the supermolecule is

$$\hat{H}^{AB}\left(\mathbf{r}_A, \mathbf{r}_B\right) = \hat{H}^{A}\left(\mathbf{r}_A\right) + \hat{H}^{B}\left(\mathbf{r}_B\right) + V_{AB}\left(\mathbf{r}_A, \mathbf{r}_B\right), \quad (7)$$

where $\hat{H}^{\mathrm{A}}(\mathbf{r}_{\mathrm{A}})$ describes the isolated molecule A with particles I having position vectors $\mathbf{r}_{\mathrm{I}}$ and charges q_I. Analogously, $\hat{H}^{\mathrm{B}}(\mathbf{r}_{\mathrm{B}})$ describes molecule B, and $V_{\mathrm{AB}}(\mathbf{r}_{\mathrm{A}}, \mathbf{r}_{\mathrm{B}})$ describes the Coulomb interaction between all particles of A with all particles of B:

$$V_{\mathrm{AB}}\left(\mathbf{r}_{\mathrm{A}}, \mathbf{r}_{\mathrm{B}}\right) = \sum_{I=1}^{n_{\mathrm{A}}} \sum_{J=1}^{n_{\mathrm{B}}} \frac{q_I q_J}{r_{IJ}}, \tag{8}$$

where r_{IJ} is the distance between particle I and J. Whereas the geometry of the supermolecule is nearly always optimized, the isolated molecules may either be in their corresponding equilibrium geometries or in deformed geometries, depending on whether the interaction energy includes the deformation energies of the interacting molecules or not. A well-known problem with the supermolecule approach is the basis set superposition error (BSSE). Because of the finite one-particle basis, counterpoise corrections (CPC) are necessary to get reliable interaction energies.

In the perturbation approach, the unperturbed Hamiltonian for the complex is $\hat{H}^{(0)} = \hat{H}^{\mathrm{A}} + \hat{H}^{\mathrm{B}}$. Here, the geometry of the interacting molecules determines the geometry of the complex. It is assumed that the ground- and excited-state functions Ψ_i^{A} and Ψ_j^{B} of the interacting molecules are known, the wave functions of the complex are then simply the products $\Psi_{ij}^{\mathrm{AB}} = \Psi_i^{\mathrm{A}}\Psi_j^{\mathrm{B}}$, they are eigenfunctions of $\hat{H}^{(0)}$. The energy of the interacting complex is the sum of the energy contributions of different order:

$$E^{\mathrm{AB}} = E^{(0)} + E^{(1)} + E^{(2)} + \cdots, \tag{9}$$

with

$$\begin{aligned} E^{(0)} &= E_0^{\mathrm{A}} + E_0^{\mathrm{B}} = \left\langle \Psi_0^{A} \middle| \hat{H}^{\mathrm{A}} \middle| \Psi_0^{\mathrm{A}} \right\rangle + \left\langle \Psi_0^{\mathrm{B}} \middle| \hat{H}^{\mathrm{B}} \middle| \Psi_0^{\mathrm{B}} \right\rangle, \\ E^{(1)} &= \left\langle \Psi_0^{\mathrm{A}}\Psi_0^{\mathrm{B}} \middle| V^{\mathrm{AB}} \middle| \Psi_0^{\mathrm{A}}\Psi_0^{\mathrm{B}} \right\rangle, \\ E^{(2)} &= -{\sum_{ij}}' \frac{\left| \left\langle \Psi_0^{\mathrm{A}}\Psi_0^{\mathrm{B}} \middle| V^{\mathrm{AB}} \middle| \Psi_i^{\mathrm{A}}\Psi_j^{\mathrm{B}} \right\rangle \right|^2}{E_{ij}^{(0)} - E^{(0)}}, \end{aligned} \tag{10}$$

where $E_{ij}^{(0)} = E_i^{\mathrm{A}} + E_j^{\mathrm{B}} = \langle \Psi_i^{\mathrm{A}} \mid \hat{H}^{\mathrm{A}} \mid \Psi_i^{\mathrm{A}} \rangle + \langle \Psi_j^{\mathrm{B}} \mid \hat{H}^{\mathrm{B}} \mid \Psi_j^{\mathrm{B}} \rangle$ and the prime on the summation sign indicates that i and j are not zero at the same time.

Since all perturbation contributions are calculated by using the wave functions of the isolated molecules, there is no BSSE and no CPC is needed. The first-order correction $E^{(1)}$ is simply the electrostatic interaction energy, whereas the second-order contributions are the sum of the induction and dispersion energies. This perturbation series is correct for interacting molecules far apart, because for them, the simple product $\Psi_i^{\mathrm{A}}\Psi_j^{\mathrm{B}}$ is an eigenfunction of $\hat{H}^{(0)}$. For shorter distances, the exchange of electrons between the two molecules must be considered, and the correct wave function for the interacting complex is $\mathcal{A}\Psi^{A}\Psi^{B}$ where $\mathcal{A}$ enforces the exchange of all electrons of A with those of B. But this wave function is no longer an eigenfunction of $\hat{H}^{(0)}$. There are many perturbation approaches with correctly antisymmetrized wave functions. One of them is symmetry-adapted perturbation theory (SAPT) [52]. We use the variant in which the intramolecular energies $E_0^{\mathrm{A}} + E_0^{\mathrm{B}}$ are calculated with density functional theory (DFT) and only the intermolecular contributions are calculated with perturbation theory. This approach is called SAPT(DFT) [53,54].

The electronic-structure methods used together with the supermolecule approach must be able to cover the largest part of electron correlation. Among wave-function methods, the coupled cluster method at the CCSD(T) level is currently the best method available. Conventional DFT methods cover short- and medium-range electron correlation, but not long-range correlation, which includes dispersion interactions. To correct this deficiency, empirical dispersion corrections were developed [27,34,36,39,40,55-57], which, when added to the DFT energy, yield energies of comparable quality to CCSD(T). This class of methods is called DFT+D: They are discussed in reviews such as those by Grimme and Tkatchenko [38,58]. Empirical dispersion corrections are often the sum of pair contributions calculated with strongly parameterized functions that depend only on the positions of the atom pairs, and are independent of the basis functions used with the electronic structure methods. Only the latter require CPC. For large systems, conventional DFT is often too costly and therefore semiempirical DFT methods such as DFTB (density functional tight binding) [30,31] are used, together with empirical dispersion corrections [32,33]. With these methods, intramolecular dispersion interactions in large molecules can be embraced.

If one is only interested in intermolecular dispersion corrections, one could directly calculate the pair contributions, instead of first calculating the dispersion contributions for each interacting molecule and the supermolecule, and then calculating the difference. This is the basis of the dlDF+D approach [28,29], with a dispersionless density functional that reproduces the CCSD(T) correlation energy of an interacting system without any dispersion contributions. The dispersion contributions to the interaction energy are calculated pairwise with a function that was fitted to SAPT(DFT) dispersion energies. The dlDF contributions are calculated using the supermolecule approach. These

energies require CPC. All methods mentioned have been used in our studies on weakly interacting systems.

Results

All investigations on adhesion of aromatic molecules onto carbon nanotubes and graphene presented here have been published [13-15,44]. All calculations were done with the DFTB+D method as implemented in the DFTB+ code [59]. In [13,14,44] periodic boundary conditions were used. In [15] all systems were treated as large clusters.

The description of system stabilization due to adhesion can be done by using either adhesion energies or adhesive forces. We use both concepts to describe stabilization of the systems investigated.

Adhesion energies

The starting point for these investigations was the claim, that (8,0)-carbon nanotubes (CNT) can be dissolved in aniline [60,61]. If this was true, the CNT should be more strongly bound to the aniline molecules in the first solvation shell than to other CNTs in a bundle, to prohibit solvated CNTs from aggregating and eventually precipitating. However, this was found not to be the case [13]. The stabilization energy for a (8,0)-CNT tightly covered with an aniline monolayer was only 40% of that of a bundle of CNTs in which one CNT is hexagonally surrounded by six other CNTs. We used stabilization energies normalized to the unit length. We have pointed out [13] that two parallel CNTs in their equilibrium geometry have one CZ, while three parallel CNTs with their molecular axis lying in a plane have two CZs. Accordingly, in a complex of a CNT surrounded by six CNTs, that is, covered by a monolayer of CNTs, there are six CZs between the central CNT and the monolayer, but there are also six further CZs within the monolayer – altogether 12 CZs. We found that the stabilization is indeed twelve times larger than that of a CNT dimer with one CZ. The aniline monolayer was found to consist of six strips of aniline molecules, similar to the monolayer of six CNTs. However, the stabilization energies show the differences between the two systems: for two CNTs in contact, the stabilization energy (7.33 kJ/mol·Å) is about 20% larger than for an aniline strip in contact with a CNT (5.02 kJ/mol·Å). For a monolayer of aniline molecules, the stabilization energy per aniline strip is 6.08 kJ/mol·Å. The 20% increase is caused by the interaction between the six aniline strips touching each other at the edges. Likewise, for a CNT covered by six CNTs the interaction per CNT in the monolayer is 15.07 kJ/mol·Å. The 106% increase is caused by the interactions between the CNTs in the monolayer. Both increases reflect collective effects due to interactions between molecules forming the monolayer. They also show that the edge-to-edge interaction between aniline strips is much smaller than the face-to-face interaction between aromatic molecules. Therefore, it is not surprising that a complex of an aniline strip inserted between two CNTs is less stable than two CNTs in contact with each other and the aniline strip in contact with one CNT. Accordingly, it would be highly unfavorable for an aniline molecule to separate two CNT molecules and insert itself between them, as it would need to happen if aniline were indeed a solvent for solid CNT. Although aniline has a permanent dipole moment (1.56 D) slightly smaller than that of water (1.87 D), there is no significant difference in the stabilization energies of parallel and antiparallel orientations of two linearly arranged aniline molecules. In the complex of a CNT and an aniline strip there will be stabilizing contributions from the permanent aniline dipole and the induced CNT dipole. At the CZ of two parallel CNTs, there will be a stabilizing interaction between the permanent dipole moments originating in the curvature of the CNT molecules. Nonetheless, dispersion interactions are the major stabilizing contribution for both systems, and they are also the origin of the difference in the stabilization. In a strip of aniline molecules, there is a large distance between the phenyl rings caused by the CH bonds and the NH_2 groups, and in this gap there are far fewer atoms contributing to dispersion interactions than in the underlying CNT molecule. Furthermore, many of these atoms are hydrogen atoms, which have a considerably smaller dipole polarizability than carbon atoms [62]. This explains the 20% difference between the stabilization energies and the fact that solid CNT cannot be dissolved by simple aromatic solvents. Note that solid CNT produced in electric arcs is amorphous, it consists of randomly arranged nanotubes or bundles of nanotubes. In our studies, we did not consider such irregularly arranged nanotubes, instead we studied only clusters of crystalline CNTs.

In a second paper [14], we studied the dependence of the stabilization energy on the number of atoms for a series of six aromatic and polyaromatic molecules benzene, naphthalene, anthracene, phenanthrene, pyrene and tetracene with a (8,0)-CNT molecule, see Figure 6 for the (8.0)-CNT/tetracene complex. For the series of acenes with the growth direction parallel to the CNT molecular axis, we found an excellent correlation with the number of carbon atoms. The energies for phenanthrene and especially for pyrene were, however, not well repro-

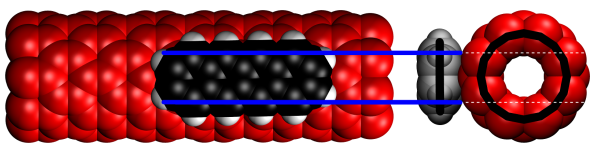


Figure 6: The contact zone of an (8.0)-CNT/tetracene complex. The bold black lines in the traverse section represent the positions of the nuclei. Reprinted with permission from [44], copyright 2017 Elsevier.

duced by the regression function because the shape of these molecules, and therefore the area of the CZ, is different from that of the four acenes. We also showed that when a planar molecule comes into contact with a CNT, it will bend towards the CNT, and this increases the number of atom pairs in close contact or, in other words, the size of the CZ.

Adhesive forces

Although CNTs cannot be dissolved in common organic solvents, it is possible to achieve dissolution by adding small amounts of a third substance, i.e., a solubilizer. Very different substance classes are claimed to be efficient solubilizers and the main question is: How can small amounts of these substances achieve separation of CNT molecules from the bulk solid? A comparison of the total energies of the systems with and without solvated CNT molecules does not explain the process of separating CNT molecules from the bulk. This can only be done with the help of forces. Every CNT molecule in the bulk is a subsystem in a large interacting system that is stabilized by adhesive forces. A CNT can be separated from the bulk only if the pull-off force is larger than the maximum adhesive force. A satisfactory explanation of the process of dissolution must include not only the origin of such pull-off forces but also show which point an external force can act on. For answering both questions, papers from the group of Nakashima [63,64] provide valuable insights. The solubilizers used by these researchers to dissolve bulk CNTs consisted of an aromatic moiety with at least three condensed aromatic rings connected by a very short aliphatic chain to a so-called solvophilic group, which could have a very different polarity. Embedding the solvophilic group into the solvent bulk is essential for the solubilizer to facilitate dissolution. Nonpolar solvophilic groups enabled CNT molecules to be dissolved in nonpolar solvents, while strongly polar or charged solvophilic groups allowed the CNTs to be dissolvated even in polar solvents. The aromatic moiety, on the other hand, is attached to a CNT molecule. The collisions of solvent molecules and the solvophilic moiety result in the generation of stochastic impulses that may add up to a net pull-off force that acts via the solvophilic group and the short connecting chain on the aromatic moiety. If the maximum adhesive force in the solubilizer/CNT(molecule) system is larger than the pull-off force, the solubilizer will not be separated, but the pull-off force will act on the CNT molecule and try to pull it off the bulk. This will happen if the maximum adhesive force in the CNT(molecule)/CNT(bulk) system is smaller than the pull-off force. Then the solubilizer is efficient. One can assume that more than one solubilizer molecule will stick to a CNT molecule and that external forces acting via several solubilizer molecules will separate a CNT molecule from the bulk. After separation of the CNT from the bulk, the solubilizer molecules will remain attached to the dissolved CNT molecule and thus avoid immediate aggregation.

To find out how the efficiency of a Nakashima-type solubilizer depends on the number of condensed aromatic rings in the aromatic moiety, we calculated the adhesive force functions for the separation of benzene, anthracene, tetracene and pyrene adsorbed to (8,0)-CNT and graphene [14,15,44]. The basic features are best understood by considering the separation of an adsorbate from graphene, where, in the equilibrium geometry, the CZ is the intersection of the area of the adsorbate and the graphene sheet, and is, therefore, proportional to the area of the adsorbate. An adsorbate can be rigid or flexible, and the pull-off point can be at the edge or in the middle of an adsorbate. Thus, four different separation modes can be formulated. Figure 7 shows the four separation modes for the separation of tetracene from graphene, Figure 8 shows the slope functions for the four separation modes.

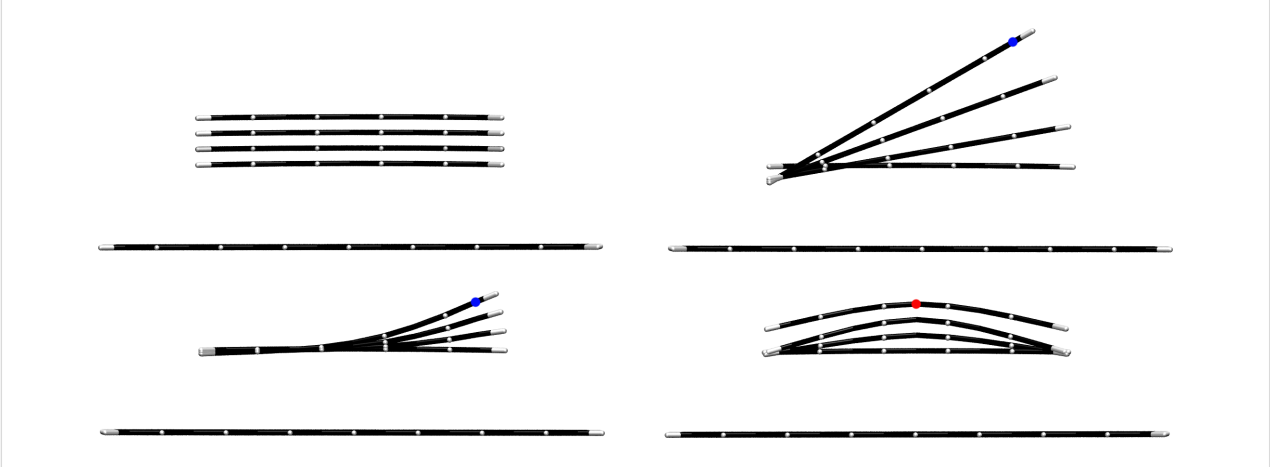


Figure 7: The separation of tetracene from graphene. Top row: Mode S1 (left), mode S2 (right). Bottom row: mode S3 (left), mode S4 (right). The blue and red dots indicate the pull-off points.

1. In separation mode S1, a rigid adsorbate is separated so that the distances of all atom pairs in the CZ increase by the same factor *f*. Think of the separation of two glass plates in contact without tilting. In such a separation, the interaction across the whole CZ changes equally, and, therefore, the reduced CZ is identical to the CZ, which is zero as soon as the separation is larger than the threshold value. Since all atom pairs contribute equally to the change in the interaction energy, the adhesive forces are directly proportional to the size of the CZ.

2. Separation mode S2 can be regarded as the separation of two glass plates by tilting. The pull-off force acts at one end of the rigid adsorbate and causes a rotation about the pivot at the other end. The distances of all atom pairs in the CZ increase at different rates. For each fixed tilt angle, the factor *f* is proportional to the tilt angle and the distance of the moving atom from the pivot. The distances of all atom pairs increase linearly along the length of the adsorbate, while the atom pairs furthest from the pivot reach the threshold value first, after which these atom pairs no longer contribute to the adhesion energy. The reduced CZ is maximal. Only the distances between the pivot atom pairs remain constant. In this separation mode, the position of the adsorbate changes from initially parallel to orthogonal with respect to the adsorbent. Only then is the adsorbate separated from the adsorbent. The adhesion energy changes less strongly than in mode S1 and, accordingly, the adhesive forces are smaller. Note that, in molecular systems, the pivot is slightly shifted.

3. Separation mode S3 is similar to separation mode S2, in that the external force acts at the edge of the adsorbate. However, in S3 the adsorbate is flexible, and bends during separation (peeling). Therefore, in all atom pairs far from the pull-off point, the distances remain largely unaffected. These atom pairs form the CZ and contribute most to the adhesion energy, but not at all to the adhesive forces. The pair distances of all other atom pairs are stretched; if the stretched distance is shorter than the threshold value, the adhesive forces resist the pulling, and in this region the adsorbate is bent. The atom pairs of the convex adsorbate form the reduced CZ. For distances larger than the threshold, no adhesive forces resist the pulling, and the adsorbate relaxes. See Figure 5. During relaxation, the bending energy is released. By continuously pulling at the pull-off point, the non-interacting part of the adsorbate increases steadily. The CZ is steadily reduced but remains as large as possible, and the small bent area of the adsorbate, i.e., the reduced CZ, remains approximately constant in size and moves towards the pivot. The change in the stabilization energy of the reduced CZ is not only due to the separation of the atom pairs but also due to the bending of the adsorbate. The energy needed for doing this, the bending energy, is stored in the adsorbate. When the pair distance is larger than the threshold value, the bending energy is released during relaxation of the adsorbate. The change in the stabilization energy therefore depends strongly on the stiffness of the adsorbate. A flexible adsorbate can be easily bent. The bending energy stored in the adsorbate is small and therefore only little bending energy will be released. The reduced CZ is small. For a stiff adsorbate, the bending energy and the reduced CZ are large. For infinite stiffness of the adsorbate separation mode S3 becomes separation mode S2.

4. In separation mode S4, the external force acts at the middle of the flexible adsorbate. Only the atoms close to the non-terminal pull-off point are displaced. The reduced CZ is symmetrical to the pull-off point, and the CZ is farther away. If only the distances of the atom pairs close to the pull-off point increase, then sufficiently large adsorbates are bell-shaped, which means that the center of the adsorbate is concave, further out, it is convex. This causes strong bending of the adsorbate and a substantial reduction in the stabilization energy. Although only small parts of the CZ are reduced, the increase in the bending energy makes this separation mode less favorable than S3 but still more favorable than S1. In separation modes S1, S2 and S3, dragging and thus friction can be avoided. In mode S4, the left and right wings will always slide over the adsorbent unless stretching of the adsorbate is less costly than dragging the parts into contact with the adsorbent. For infinite stiffness of the adsorbate, separation mode S4 becomes separation mode S1.

The slope functions for the separation of tetracene from graphene are shown in Figure 8. Small and isotropic adsorbates such as benzene or pyrene are stiffer than long, anisotropic

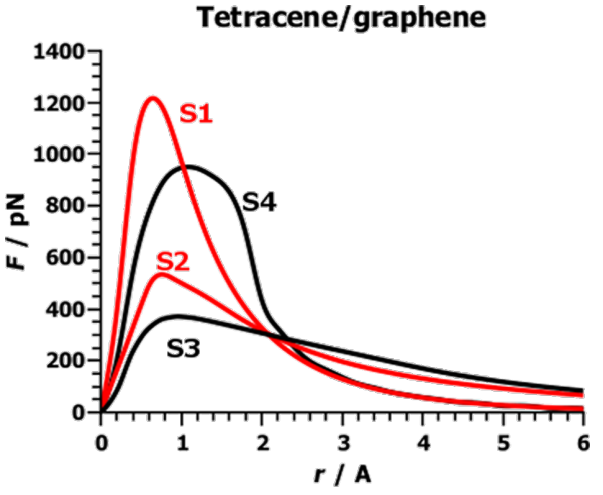


Figure 8: The slope functions for the separation of tetracene from graphene for the four separation modes. Red: Rigid adsorbate. Black: Flexible adsorbate.

acenes and bending costs more energy. Only in complexes with large adsorbates will a large part of the complex be nearly parallel to the adsorbent, causing bending (mode S3) instead of tilting (mode S2). For mode S1, we found an increase in the maximum adhesive force with the number of aromatic rings. The force increases from 451 pN (benzene) through 962 pN (anthracene) and 1059 pN (pyrene) to 1219 pN (tetracene), For mode S3, the force increases from 214 pN through 353 pN (anthracene) and 362 pN (pyrene) to 371 pN (tetracene). This demonstrates, firstly, that in mode S1, the reduced CZ increases with the size of the adsorbate. Therefore, the maximum adhesive force is proportional to the size of the adsorbate, and, secondly, that in mode S3, the maximum adhesive force increases strongly from benzene to anthracene, but that the difference between the large adsorbates anthracene, pyrene and tetracene is much smaller than between benzene and anthracene. These observations are in accord with the finding that the aromatic moiety of a Nakashima-type solubilizer should have at least three condensed aromatic rings to be efficient. This connection of size and shape of the adsorbate and its elastic properties is true for all classes of molecules when an adsorbate comes into contact with nonplanar adsorbents, because bending can increase the CZ, improving the stabilization of the complex. If bending produces more stabilization (due to the larger CZ) than it costs, the adsorbate will change its form to maximize both adhesion energy and adhesive forces.

Hydrophobic interaction, hydrogen bonding and properties of liquids

Condensed-matter properties are strongly influenced by cooperative effects caused by more than two interacting particles (many-body effects). In statistical physics, these effects are represented by the cluster expansions of the partition function or the thermodynamic potentials [65]. The cluster expansion of the interaction potential of a condensed matter system composed of molecules,

$$V_{\text{int}} = \sum_{\text{A>B}} V_{\text{AB}} + \sum_{\text{A>B>C}} V_{\text{ABC}} + \sum_{\text{A>B>C>D}} V_{\text{ABCD}} + \cdots, \quad (11)$$

says that the properties of a liquid cannot be described solely by two-body contributions, that is, contributions of two solvent molecules. Clusters of molecules with low spatial symmetry have, in general, several stable structures that vary in their stabilities and molecular properties, e.g., electric multipole moments and polarizabilities, and thus contribute differently to the stabilization energy. Electrostatic interactions are strictly additive for all distances at which the molecular electron distributions do not overlap, all other basis interactions are non-additive and contributions of larger clusters are essential. Whereas long-range dispersion interactions are approximately additive, induction interactions are strictly non-additive. This means that it is not possible to add up all electric fields due to the static moments of the surrounding molecules and then calculate the induction energy for a given molecule. However, in the case of less polar or less polarizable molecules, approximate additivity seems to be reasonable [2]. Liquid alkanes are such systems. All straight-chain alkanes (*n*-alkanes) can be derived from the parent substance methane by substituting one hydrogen atom for *n*-alkyl chains of increasing length. Under standard conditions, macroscopic amounts of alkanes occur in all three phases. *n*-Alkanes with up to four carbon atoms are gases, *n*-alkanes from five to 17 carbon atoms are liquids, and all longer *n*-alkanes are solids. The melting point and particularly the boiling point (BP) reflect the degree of interaction between the molecules in the condensed phase, the degree of interactions itself depends on the size of the CZs. The boiling point is especially significant because it is related to the process of separating molecules from the bulk, i.e., to working against adhesion forces. The melting point is related to the change of the short-range order in the two condensed phases. The dynamic (shear) viscosity is another property that strongly depends on intermolecular interactions. It is related to the resistance of the molecules to moving relative to each other, which is nothing but friction caused by either attractive interactions between the molecules or mechanical locking caused by surface roughness.

All experimental data mentioned in this section were taken from the online databases GESTIS [66], EngineeringToolBox [67] and ChemicalBook [68]. The BP of *n*-alkanes is a monotonically increasing function of the chain length. The changes in BP for the first four *n*-alkanes are rather large: 73 °C between methane and ethane, 47 °C between ethane and propane and 41 °C between propane and butane. The low BP of methane reflects very weak molecular interactions. The largest contribution to electrostatics are octopole–octopole interactions. In the liquid phase, this interaction varies as $1/r^7$ with the intermolecular distance. The lowest contribution to induction is the interaction of a static octopole and an induced dipole with a $1/r^8$ distance dependence, so the dipole–dipole dispersion interaction with the $1/r^6$ distance dependence yields the largest attractive contribution. Due to the free rotation in the gas phase, the electrostatic interaction is of much shorter range, the thermal averaged interaction is proportional to $1/r^{14}$. In ethane, the lowest multipole is a quadrupole with a very low quadrupole moment. Accordingly, contributions to electrostatics and induction are also very small, and dispersion again yields the largest contribution. But in this case, the size of the CZ becomes important. The more atoms with large polarizability there are in the CZ of an atom, the larger is the dispersion contribution, and the polarizability of the carbon atom is much larger than that of the hydro-

gen atom. In a large alkane molecule, many more carbon atoms are close to each other because of the short covalent bonds between them, in contrast to the large distances between the carbon atoms in liquid methane. In the latter case, any atom seeing a carbon atom will see also some hydrogen atoms but no other carbon atom close by, whereas an atom seeing a carbon atom in a large alkane molecule will probably see a second or a third carbon atom. The CZ of an atom seeing methane molecules is much smaller than that of an atom seeing large alkanes. Therefore, a larger number of carbon atoms in the CZ will increase the stabilization much more than the same number of hydrogen atoms. The differences in the boiling points of the first few *n*-alkanes also show a strong influence of the shape of the molecules on the CZs. Methane is much less anisotropic than ethane, which itself is less anisotropic than propane. However, the degree of anisotropy becomes less important, the larger the alkane chain is. Then, the size of the CZ becomes decisive, and the successive increases in BP become roughly constant. The importance of the anisotropy of the interacting molecules and, consequently, of the shape of the CZ, can be seen in the differences in the boiling points of isomers of a certain alkane. Straight-chain isomers can lie parallel to each other, achieving a larger CZ than branched, globular molecules. The boiling points of *n*-pentane, isopentane and neopentane are 40 °C, 28 °C, and 10 °C. On the other hand, all disk-shaped, cyclic alkanes have higher boiling points than the straight-chain molecules. The boiling point of cyclopentane is 49 °C. Likewise, the boiling points of *n*-hexane and cyclohexane are 69 °C and 81 °C, respectively.

The dependence of friction on the size of the CZ also explains why the viscosity of straight-chain alkanes increases with chain length. Surface roughness and mechanical locking are a second cause of viscosity, and they explain why the viscosity of large branched alkanes is larger than that of the corresponding straight-chain alkanes. The much larger CZ of disk-shaped cyclic alkanes such as cyclopentane, cyclohexane or cycloheptane explains the larger dynamic viscosity of cycloalkanes compared to that of straight-chain alkanes. The importance of both causes is nicely demonstrated by the viscosities of cyclohexane (1.20 mPa·s), benzene (0.65 mPa·s) and *n*-hexane (0.33 mPa·s). The cyclohexane molecule is disk-shaped and, because of the axial CH groups, has a higher roughness than the benzene molecule. Linear *n*-hexane, finally, has the smallest CZ of the three molecules. The shape and size of the CZ also explains the low viscosity of spherical molecules, such as neopentane, or quasi-spherical molecules, such as isopentane and isohexane, compared to *n*-pentane, *n*-hexane or *n*-heptane [69-71], which is surprising when one assumes that branched alkanes always have a higher viscosity. MD results for the pentane isomers are 0.2667, 0.2445 and 0.1500 mPa·s for *n*-pentane, isopentane and neopentane, respectively [70]. Experimentally, the viscosity of branched isobutane (0.166 mPa·s) was found to be slightly larger than that of *n*-butane (0.162 mPa·s) [71]. However, one can speculate that this small difference in the viscosities of the two isomers is not caused by branching but by the disk-shape of isobutane.

Since the largest contribution to the interaction between alkane molecules is the dispersion interaction, it is not only responsible for liquefaction and solidification of alkanes but also for the stabilization of hairpin structures of large *n*-alkanes with about 18 to 20 carbon atoms. In this conformation, the CZ, and hence the attractive interaction, is maximized for the carbon atoms in the arms. Only the atoms in the loop of the hairpin are further away from other atoms and, moreover, the carbon skeleton in the loop is strongly bent, which destabilizes the hairpin structure. Only if the arms are long enough and the interaction between them outweighs the destabilization in the loop does the hairpin become the most stable structure. One can assume that the stabilization of these conformations is only important in the gas phase, because in a liquid, every alkane molecule will be in contact with several other alkane molecules, which makes the interaction between non-bent alkane molecules more probable. This dispersion-dominated interaction between alkane molecules is the physical origin of the so-called hydrophobic interaction.

Polar liquids are systems in which many-body contributions cannot be neglected because of the non-additivity of polarization effects especially for induction interaction. We discuss straight-chain primary alcohols and straight-chain primary amines and their corresponding parent substances, water and ammonia. Figure 9 shows 1) that primary alcohols exhibit a higher BP than primary amines and alkanes having the same number of heavy atoms (C, N or O); 2) an apparent convergence of the BPs of amines, alcohols and alkanes with increasing size of the molecules; 3) large differences between the BPs for the respective small members of the homologous series; and 4) an exceptional BP of water, the parent substance of the alcohols. Except for water, the BPs of all three series increase monotonically with increasing size of the molecules. Ammonia and the first two amines are gases, amines with three to twelve heavy atoms are liquids, while all higher amines are solids at room temperature. Water and all alcohols with up to twelve heavy atoms are liquids, all higher alcohols are solids at room temperature. The difference in the BPs of the parent substances methane and ammonia is about 130 °C; between ammonia and water, the difference is a further 130 °C. Due to the different shapes and volumes of the three parent molecules, the densities of the liquids are rather different: H_2O (1.00 g/cm^3), NH_3 (0.73 g/cm^3) and CH_4 (0.42 g/cm^3). This

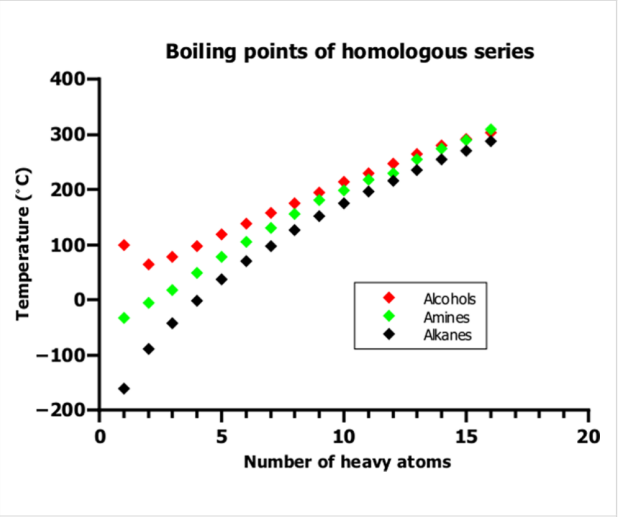


Figure 9: Boiling points of straight-chain primary alcohols, straight-chain primary amines and straight-chain alkanes. Heavy atoms are C, N or O.

implies for the number of molecules in a certain volume a ratio of 1:0.77:0.47. Accordingly, the average intermolecular distances in ammonia is 9% greater than in water, but 28% smaller than in methane. In other words, the attraction between ammonia molecules is much smaller than that between water molecules, but greater than the attraction between methane molecules. The water molecule has a larger dipole moment (1.85 D) than ammonia (1.47 D) and the anisotropies of the quadrupole moments are very different. For water the quadrupole components are (−2.12, 2.32, −0.20) DÅ, whereas for ammonia they are (1.27, 1.27, −2.54) DÅ. The shape of the water molecule, the magnitude of the multipole moments and the anisotropy of the quadrupole moment together with the shorter distance between the molecules enable much stronger intermolecular attractions in liquid water than in liquid ammonia. This is also in line with the fact that the magnitude of the dispersion contribution, E_D, in the equilibrium structures of the dimers of water, ammonia and methane varies as follows: E_D(water) > E_D(ammonia) > E_D(methane) [8,72], although the magnitude of the dipole–dipole polarizabilities shows the inverse trend: $\alpha(H_2O) = 1.501$ Å^3 < $\alpha(NH_3) = 2.103$ Å^3 < $\alpha(CH_4) = 2.448$ Å^3. This means that at equilibrium short-range dipole–quadrupole and quadrupole–quadrupole dispersion contributions are more important than the long-range dipole–dipole dispersion contributions. Substitution of one hydrogen atom by a methyl group in each parent molecule increases the size of the molecule and thus also the long-range dispersion interactions. This causes the larger BP of ethane, but it does not outweigh the loss of attractive interactions in methanol relative to the interaction in water, and thus causes the lower BP of methanol. In the amine series, we see that the increase in the molecular size is more important than a possible reduction of electrostatic, induction and short-range dispersion interactions. Further increasing alkyl chains leads to increasing BPs in all three homologous series.

In ice and, to a lesser extent, also in liquid water, each water molecule is surrounded by four other water molecules such that their dipole moments do not cancel each other out. Due to the resulting dipole-induced dipole interaction, many-body contributions, especially three-body contributions, dominated by the induction energy, are essential [2]. In methanol, the alkyl groups disturb the short-range order found in water. The mean distance between the OH groups is increased and, accordingly, the attractive electrostatic and induction interactions are reduced. Although the volume of the methanol molecule is larger than the volume of the water molecule, and the dispersion interactions are increased, the total interaction is decreased and the BP is lower. With increasing length of the alkyl groups, the BP increases again, *n*-propanol having a BP nearly equal to that of water (97 °C). In chemistry, system stabilization is preferably attributed to local molecular substructures. The most prominent example is the successful explanation of covalent bonding using groups of two atoms. Adopting this localized approach, the stability of water or alcohol dimers can be attributed to a group of three atoms forming a hydrogen bridge, A–H···B, where A and B are atoms with higher electronegativity than the bridging hydrogen atom, and B has an electron lone pair. The claims that charge transfer and covalent bonding are relevant for hydrogen bonding have their origins in this interpretation of bonding with the three-atom-four-electron group using the Lewis structures A–H|B, A|$^-$ H–B$^+$, and A|$^-$ H$^+$|B. However, these Lewis structures are simply necessary to describe the polarization of the atom group $A^{\delta-}$–$H^{\delta+}$···$B^{\delta-}$. The latter formulation also facilitates the interpretation that hydrogen bonding is predominantly electrostatic in character. In the MO description, the three Lewis structures are contained in a four-electron-three-MO CASSCF wave function, which is the lowest-level wave function including polarization effects in this atom group. However, all these simple wave functions ignore 1) all atoms attached to atoms A and B, and, in consequence, 2) the electric multipoles of the whole molecules containing A and B, 3) the polarization of the electron density of these molecules, and 4) any kind of dynamic electron correlation that covers dispersion interaction. The quantum chemical models developed to explain strong chemical bonding in localized regions of a molecule cannot cover the non-localized and non-additive bonding contributions that are typical for WMI. Using them to explain WMI leads to serious errors.

In two papers [7,8], we investigated the contributions of the four basic interactions to the stabilization of alcohol and amine dimers, demonstrating that the whole molecules contribute to the stabilization, not just the atoms of the central moiety

A–H···B. We found that, for the hairpin structures of amine dimers R^1-NH–H···NH_2–R^2 with up to four carbon atoms, the stabilization energy increases by a constant increment for each added CH_2 group. For the stretched hairpin structures going from the ammonia dimer to the methylamine dimer gave a substantial stabilization, but any further growth of the alkyl chains did not improve stabilization. In the alcohol series, R^1-O–H···OH-R^2, we investigated only the stretched hairpin structures and found improvements of the stabilization up to the ethanol dimer, but no change for the higher alcohols. These findings agree with those for the CZ of alkanes. In the hairpin structure each carbon atom in one arm is close to the corresponding carbon atom in the opposite arm and, thus, the CZ is maximal, and increases with each inserted methylene group. In the stretched hairpin structure, only the α and maybe the β carbon atoms, and the attached hydrogen atoms, of one monomer are in contact with the nearest atoms in the opposite monomer. Hence, the CZ is minimal. This is true for the growth of straight-chain alkyl substituents. However, in *tert*-butylamine dimer or the *tert*-butanol dimer, each primary methyl group is equally close to the oxygen or nitrogen atom of the opposite monomer, but the distances between the methyl groups of the two monomers are larger than the distances between opposite carbon atoms in the hairpin structure. Accordingly, we find the following order of dimer stabilities: *n*-butyl dimer(hairpin) > *tert*-butyl dimer > *n*-butyl dimer (stretched hairpin). The higher stability of the hairpin structures due to intramolecular interaction is, however, not relevant for the properties of the liquids, which depend on intermolecular interactions, and we do not even know whether two-body clusters in the cluster expansion are indeed dimers. But we do know for sure that the BP of *tert*-butanol (83 °C) is 35 °C lower than the BP of *n*-butanol and that the BP of *tert*-butylamine (45 °C) is 33 °C lower than the BP of *n*-butylamine. This is comparable to the 30 °C difference between the BPs of *n*-pentane and neopentane.

Comparison of the viscosities of alkanes with those of alcohols and amines is difficult because far less experimental data are available for the latter two groups of substances, and the data found in the literature vary considerably. Nonetheless, the available data do allow the following conclusions to be drawn: First, polar groups increase the viscosity due to electrostatic and induction interactions, as the values for *n*-pentane, *n*-butylamine and *n*-butanol show, which are 0.240 mPa·s (25 °C), 0.470 mPa·s (20 °C), and 2.95 mPa·s (20 °C), respectively. The assumption that the increase in viscosity is caused by hydrogen bonding is unjustified. Replacing a CH_2 group in cyclopentane by a sulfur atom doubles the viscosity from 0.413 mPa·s (25 °C) to 0.973 mPa·s (25 °C) [73]. Second, the combination of polar groups and branching enhances the increase in viscosity. The viscosity of isobutanol (3.95 mPa·s at 20 °C) is considerably larger than the viscosity of *n*-butanol. These data show that any attempt to attribute properties of condensed-phase systems to a single cause, e.g., hydrogen bonding, fails.

Bonding in cellulose

Cellulose is a material showing polymorphism, crystals of the Iα and the Iβ allomorphs are composed of layers of parallel aligned cellulose chains, which are chains of D-glucose rings connected by 1→4 glycosidic bonds. In each glucose ring, there are five axial CH groups, and two OH groups and one hydroxymethyl group in equatorial position. The cellulose chains are stabilized by intrachain hydrogen bridges, while interchain hydrogen bridges connect the cellulose chains. All equatorial OH groups are involved in intra- and interchain hydrogen bridges. In addition to these hydrogen bridges, the hydroxymethyl group can also contribute to intersheet hydrogen bridges. This flexibility is due to the different possible conformations the hydroxymethyl group can adopt. It is common belief among cellulose scientists that this hydrogen-bonding network is responsible for the stability of cellulose [74,75] and also for the insolubility of cellulose fibers, but this view has recently been criticized [76]. After all, the crystal structures of cellulose Iα, cellulose Iβ and cellulose II vary considerably and so do their hydrogen bonding networks. Only in cellulose II can one speak of a three-dimensional network. In cellulose I, hydrogen bridges are nearly exclusively found within the sheets, with only very few hydrogen bridges connecting the sheets. These so called hydrogen-bonding networks are, however, neither unique nor static [75,77]. MD simulations show that hydrogen bridges are dynamically created and broken [77]. While cellulose Iα is made of one type of layers, cellulose Iβ is composed of two types of sheets, each of which seems to favor a different hydrogen-bonding network [75]. The few intersheet hydrogen bridges cannot explain the stability of cellulose, so other attractive interactions must be responsible for the attraction of the sheets, which in Iβ crystals have Miller indices (100). The (100) surfaces are described as hydrophobic because they are dominated by the axial CH groups (57% accessibility). Both the hydroxyl and acetal oxygen atoms are lying deeper in the sheet and are, consequently, less accessible (43% accessibility) [78]. The interaction between the stacked glucose rings is dominated by dispersion interactions, similar to the interaction between cyclic alkanes, but there are, of course, also electrostatic and induction contributions from all atoms involved, not just from the interaction between CH groups and oxygen atoms. Nonetheless, the importance of dispersion interactions can easily be seen when important properties of all four principal cleavage planes of cellulose [79] are investigated, such as the surface energy, the attachment energy or the surface roughness. In Iβ crystals, these planes have Miller indices (100), (010), (110) and (1−10). The accessibility of oxygen atoms on the surface increases the

electrostatic and induction interactions. Consequently, the surface energies of the four surfaces vary by a factor of almost two, according to molecular dynamics studies [78]. The surface energy of the most hydrophobic (100) surface is about 190 mJ/m^2 whereas for the most hydrophilic (010) surface it is about 350 mJ/m^2 [78]. Likewise, the energies for the attachment of a new cellulose layer, which is a measure for the cohesive energy, were found to be about 125 kcal/mol for the (100) and about 270 kcal/mol for the (010) surface [78]. These are the energies for the unrelaxed surfaces. Thus, electrostatics and induction only enhance the interaction between different crystallographic planes by a factor of two, which demonstrates the importance of dispersion interactions for the stability of cellulose crystals. This motivated us to make an analogy between bonding in hydrogen-bridged systems and reinforced concrete, which is made of concrete and rebars. The role of concrete can be seen as being played by the largely isotropic dispersion interactions, while the anisotropic electrostatic interactions play the role of rebars. Neglecting dispersion is like forgetting the concrete, while neglecting electrostatics and induction would be forgetting the rebars. French recently criticized inconsistent claims such as "cellulose fibers are insoluble because they are held together by hydrogen bonds", which he calls a truism. He raised the question, "if the three hydrogen bonds per glucose unit in cellulose Iβ […] explain the insolubility, then why is β-glucose, with five conventional hydrogen bonds per glucose unit […] so soluble?" [76]. He concluded: "other factors such as unconventional C–H···O hydrogen bonding and van der Waals interactions must also be important, and the truism does not bring them into consideration". In our opinion, the above statement is not a truism, which by definition is frequently true. It is simply wrong because it considers only the rebars and forgets the concrete.

The stability of cellulose crystals is, however, not only due to adhesive forces between layers but also due to dry friction, which describes the processes that hinder relative lateral motions of two solid surfaces moving against each other. The cause for dry friction can be strong adhesion, entanglement of the surfaces due to roughness, or strong interaction between localized parts of the surfaces such as heteroatoms or atom groups. Although the (100) surface in cellulose Iβ crystals is the smoothest of the four principal cleavage planes, the planes do not slide against each other as do the graphene sheets in graphite, because there are adhesive interactions between the sheets as well as friction due to the surface roughness caused by the axial CH groups.

Adsorption to cellulose

Similar to bonding in the cellulose bulk, hydrogen bonding is regarded as the dominant type of interaction responsible for the adsorption of small molecules with polar groups onto cellulose surfaces. We investigated the adsorption of glucose, cellobiose and cellotetraose onto the hydrophilic (100) surface of Iα cellulose and the hydrophobic (100) surface of cellulose Iβ by using the BP86-D2 density functional and the GLYCAM06 force field [80]. For the adsorption of D-glucose onto the hydrophilic Iα surface, the most stable structure was the one in which the glucose ring was perpendicular to the cellulose surface. At least two hydrogen bridges were found for the structure, depending on the method used. Also on the hydrophobic (100) surface of Iβ, a structure with the glucose ring perpendicular to the cellulose surface was most stable, but with increasing size of the adsorbate the situation changes considerably. In the most stable structures of cellobiose adsorbed to both surfaces, the glucose rings are parallel to the surfaces. Structures with perpendicular glucose rings are markedly less stable, and for the adsorption of cellotetraose, this trend is intensified. Although for the small adsorbates, bonding to the hydrophilic surface is markedly stronger than to the hydrophobic surface, this difference vanishes for large adsorbates. All these findings are consistent with an increasing contribution of dispersion interactions with increasing size of the adsorbate, that is, with increasing size of the CZ. This demonstrates that, as we have frequently emphasized [7,8,80], WMIs cannot be described by a single basic interaction. The contribution of electrostatics to the bonding of complexes with hydrogen bridges is large, but it is not sufficient to explain their stability.

Discussion

In the Discussion we consider only WMIs between finite molecules. A WMI between the two arms in the hairpin structure of a large alkane is an example of an intramolecular interaction. The interactions between the alkyl groups in alcohol or amine dimers in hairpin structures, or between two parallel alkane molecules, are examples for intermolecular interactions. The physical origin is the same in both cases. The strength of the interaction depends only on the size of the CZ. The near-sightedness of the WMI makes it a local interaction, and the approximate additivity justifies the assumption that the WMI is dominated by pair contributions. If the WMI is dominated by dispersion interactions, it depends only on the distance between the interacting atoms. If the interaction between multipoles of low rank (e.g., dipoles) contributes significantly, the WMI will be anisotropic and will depend strongly on the relative orientation of the multipoles. Adhesion is a process where WMIs stabilize a system consisting of different subsystems. It can be described by using energies or forces, but the descriptions are not equivalent. Forces are vital for describing the perturbations of such systems by external forces, as well as for describing the response of the systems to this perturbation. We showed how a change of the pull-off point can influence the magnitude of the

internal force holding the system together. We also showed that, in general, the magnitude of adhesive forces does not depend on the whole CZ. Rather, it depends only on that part of the CZ where the attractive interaction changes but is not yet zero. Thus, the elastic properties of extended molecular systems are directly related to internal forces and WMIs. Although dispersion interactions play a dominant role in WMIs, there are fields of chemistry in which dispersion interaction are systematically ignored as soon as polar atom groups occur, such as those involved in hydrogen bridges, at which point all stabilizing interactions are attributed to the hydrogen bridges. In contrast, we found that, in systems such as cellulose crystals, the electrostatic and induction contributions of hydrogen bridges amplify the stabilizing dispersion interactions, a finding that is confirmed by conjectures and observations of others [76]. The stability of cellulose crystals is, however, not only due to attractions between cellulose chains within and between the layers, but also due to friction. In solids like cellulose, static friction can hinder the lateral movement of the layers against each other. In liquids, kinetic friction is responsible for viscosity, static and kinetic friction are types of dry friction. Friction changes the state of motion of the subsystems involved, it always slows down the speed, the (negative) acceleration is due to a force called a friction force. If the relative lateral motion is due to a pulling or pushing force, in engineering this is called the "load", the system responds with an opposing friction force. This friction force is directly proportional to the applied load. Friction forces are not conservative and cannot be derived from a potential. Detailed information about friction can be found, e.g., in the Handbook of Tribology [81]. The cause for dry friction can be strong attraction between the surfaces. This strong attraction can be due to adhesive forces, entanglement of the surfaces due to roughness, or strong interaction between localized parts of the surfaces, such as heteroatoms or atom groups. The relation between the magnitude of the friction force, F_f, and the magnitude of the adhesive force F_a is $F_f \leq \mu F_a$. The constant μ is the coefficient of friction and is an empirical quantity of the interacting materials. In most cases it is smaller than 1 [82]. The viscosity of a liquid is caused by friction between the molecules of the liquid. The higher viscosity of branched alkanes compared to linear alkanes can be attributed to entanglement of the molecules, while the higher viscosity of alcohols or amines relative to alkanes is mainly caused by the interactions of the polar atom groups with the other molecules. In any case, all these interactions are WMIs, and in many cases they are dominated by dispersion interactions. Irrespective of how the WMIs in given systems are composed, WMIs are responsible for many properties of condensed matter systems, and forces are important for their description. For the description of processes such as friction or the stability of interfaces, forces are indispensable.

ORCID® iDs

Robin Guttmann - https://orcid.org/0000-0002-7305-9324
Christoph Lechner - https://orcid.org/0000-0001-6160-2409
Reinhard J. Maurer - https://orcid.org/0000-0002-3004-785X
Alexander F. Sax - https://orcid.org/0000-0002-6652-5710

References

1. Chemical Bonding. http://www.smallscalechemistry.colostate.edu/PowerfulPictures/ChemicalBonding.pdf (accessed Aug 2, 2018).
2. Stone, A. *The Theory of Intermolecular Forces,* 2nd ed.; Oxford University Press: Oxford, UK, 2013.
3. Jeffrey, G. A. *An Introduction to Hydrogen Bonding;* Oxford University Press: New York, NY, USA, 1997.
4. Gilli, G.; Gilli, P. *The Nature of the Hydrogen Bond;* Oxford University Press: New York, NY, USA, 2009.
5. Coulson, C. A. *Valence;* Clarendon Press: Oxford, UK, 1952.
6. Israelachvili, J. N. *Intermolecular and Surface Forces,* 3rd ed.; Academic Press: Amsterdam, Netherlands, 2011.
7. Hoja, J.; Sax, A. F.; Szalewicz, K. *Chem. – Eur. J.* **2014,** *20,* 2292–2300. doi:10.1002/chem.201303528
8. Guttmann, R.; Sax, A. F. *ChemistryOpen* **2017,** *6,* 571–584. doi:10.1002/open.201700052
9. Kauzmann, W. *Adv. Protein Chem.* **1959,** *14,* 1–63. doi:10.1016/s0065-3233(08)60608-7
10. Wolfenden, R.; Lewis, C. A., Jr. *J. Theor. Biol.* **1976,** *59,* 231–235. doi:10.1016/s0022-5193(76)80035-5
11. Baldwin, R. L. *FEBS Lett.* **2013,** *587,* 1062–1066. doi:10.1016/j.febslet.2013.01.006
12. In 1946, Eugene Blank published "The folklore of chemistry" in *School Science and Mathematics* **1946**, (*46*), 493-499. Referring to the definition of "folklore" in Webster's Collegiate Dictionary, 5th ed. (1940) as *Traditional customs, beliefs, tales, or sayings, preserved unreflectively among a people*, he stated "On the basis of this definition there exists a vast body of unrecognized chemical folklore as this paper will endeavor to record." This statement has not lost its relevance.
13. Maurer, R. J.; Sax, A. F. *Phys. Chem. Chem. Phys.* **2010,** *12,* 9893–9899. doi:10.1039/c001066a
14. Glanzer, S.; Sax, A. F. *Mol. Phys.* **2013,** *111,* 2427–2438. doi:10.1080/00268976.2013.831499
15. Lechner, C.; Sax, A. F. *J. Phys. Chem. C* **2014,** *118,* 20970–20981. doi:10.1021/jp505894p
16. Levy-Leblond, J.-M.; Balibar, F. *Quantics, Rudiments of Quantum Physics;* North-Holland: Amsterdam, Netherlands, 1900.
17. Lennard-Jones, J. *Adv. Sci.* **1954,** *11,* 136–148.
18. Ahlrichs, R. *Theor. Chim. Acta* **1976,** *41,* 7–15. doi:10.1007/bf00558020
19. Böhm, H.-J.; Ahlrichs, R. *J. Chem. Phys.* **1982,** *77,* 2028–2034. doi:10.1063/1.444057
20. Stone, A. J. *J. Phys. Chem. A* **2011,** *115,* 7017–7027. doi:10.1021/jp112251z
21. Salam, A. *Non-Relativistic QED Theory of the van der Waals Dispersion Interaction;* Springer, 2016.
22. Zaremba, E.; Kohn, W. *Phys. Rev. B* **1976,** *13,* 2270–2285. doi:10.1103/physrevb.13.2270
23. Zaremba, E.; Kohn, W. *Phys. Rev. B* **1977,** *15,* 1769–1781. doi:10.1103/physrevb.15.1769

24. Persson, B. N. J.; Apell, P. *Phys. Rev. B* **1983,** *27,* 6058–6065. doi:10.1103/physrevb.27.6058
25. Thouless, D. J. *The quantum mechanics of many-body systems,* 2nd ed.; Academic Press: New York, NY, USA, 1972.
26. Mattuck, R. D. *A Guide to Feynman Diagrams in the Many-Body Problem,* 2nd ed.; McGraw-Hill: New York, NY, USA, 1976.
27. Grimme, S. *J. Comput. Chem.* **2006,** *27,* 1787–1799. doi:10.1002/jcc.20495
28. Pernal, K.; Podeszwa, R.; Patkowski, K.; Szalewicz, K. *Phys. Rev. Lett.* **2009,** *103,* 263201. doi:10.1103/physrevlett.103.263201
29. Podeszwa, R.; Pernal, K.; Patkowski, K.; Szalewicz, K. *J. Phys. Chem. Lett.* **2010,** *1,* 550–555. doi:10.1021/jz9002444
30. Seifert, G.; Porezag, D.; Frauenheim, T. *Int. J. Quantum Chem.* **1996,** *58,* 185–192. doi:10.1002/(sici)1097-461x(1996)58:2<185::aid-qua7>3.0.co;2-u
31. Elstner, M.; Porezag, D.; Jungnickel, G.; Elsner, J.; Haugk, M.; Frauenheim, T.; Suhai, S.; Seifert, G. *Phys. Rev. B* **1998,** *58,* 7260–7268. doi:10.1103/physrevb.58.7260
32. Elstner, M.; Hobza, P.; Frauenheim, T.; Suhai, S.; Kaxiras, E. *J. Chem. Phys.* **2001,** *114,* 5149–5155. doi:10.1063/1.1329889
33. Stöhr, M.; Michelitsch, G. S.; Tully, J. C.; Reuter, K.; Maurer, R. J. *J. Chem. Phys.* **2016,** *144,* 151101. doi:10.1063/1.4947214
34. Tkatchenko, A.; Scheffler, M. *Phys. Rev. Lett.* **2009,** *102,* 073005. doi:10.1103/physrevlett.102.073005
35. Dobson, J. F. *Int. J. Quantum Chem.* **2014,** *114,* 1157–1161. doi:10.1002/qua.24635
36. Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. *J. Chem. Phys.* **2010,** *132,* 154104. doi:10.1063/1.3382344
37. Reilly, A. M.; Tkatchenko, A. *Chem. Sci.* **2015,** *6,* 3289–3301. doi:10.1039/c5sc00410a
38. Hermann, J.; DiStasio, R. A., Jr.; Tkatchenko, A. *Chem. Rev.* **2017,** *117,* 4714–4758. doi:10.1021/acs.chemrev.6b00446
39. Tkatchenko, A.; DiStasio, R. A.; Car, R.; Scheffler, M. *Phys. Rev. Lett.* **2012,** *108,* 236402. doi:10.1103/physrevlett.108.236402
40. Ambrosetti, A.; Reilly, A. M.; DiStasio, R. A., Jr.; Tkatchenko, A. *J. Chem. Phys.* **2014,** *140,* 18A508. doi:10.1063/1.4865104
41. Maurer, R. J.; Ruiz, V. G.; Tkatchenko, A. *J. Chem. Phys.* **2015,** *143,* 102808. doi:10.1063/1.4922688
42. Kohn, W. *Phys. Rev. Lett.* **1996,** *76,* 3168–3171. doi:10.1103/physrevlett.76.3168
43. Prodan, E.; Kohn, W. *Proc. Natl. Acad. Sci. U. S. A.* **2005,** *102,* 11635–11638. doi:10.1073/pnas.0505436102
44. Lechner, C.; Sax, A. F. *Appl. Surf. Sci.* **2017,** *420,* 606–617. doi:10.1016/j.apsusc.2017.05.170
45. Ruedenberg, K. *Rev. Mod. Phys.* **1962,** *34,* 326–376. doi:10.1103/revmodphys.34.326
46. Ruedenberg, K.; Schmidt, M. W. *J. Comput. Chem.* **2007,** *28,* 391–410. doi:10.1002/jcc.20553
47. Ruedenberg, K.; Schmidt, M. W. *J. Phys. Chem. A* **2009,** *113,* 1954–1968. doi:10.1021/jp807973x
48. Bitter, T.; Ruedenberg, K.; Schwarz, W. H. E. *J. Comput. Chem.* **2007,** *28,* 411–422. doi:10.1002/jcc.20531
49. Bitter, T.; Wang, S. G.; Ruedenberg, K.; Schwarz, W. H. E. *Theor. Chem. Acc.* **2010,** *127,* 237–257. doi:10.1007/s00214-010-0758-y
50. Schmidt, M. W.; Ivanic, J.; Ruedenberg, K. *J. Chem. Phys.* **2014,** *140,* 204104. doi:10.1063/1.4875735
51. Schmidt, M. W.; Ivanic, J.; Ruedenberg, K. The Physical Origin of Covalent Bonding. *The Chemical Bond;* Wiley-VCH Verlag GmbH: Weinheim, Germany, 2014; pp 1–68. doi:10.1002/9783527664696.ch1
52. Jeziorski, B.; Moszynski, R.; Szalewicz, K. *Chem. Rev.* **1994,** *94,* 1887–1930. doi:10.1021/cr00031a008
53. Misquitta, A. J.; Podeszwa, R.; Jeziorski, B.; Szalewicz, K. *J. Chem. Phys.* **2005,** *123,* 214103. doi:10.1063/1.2135288
54. Heßelmann, A.; Jansen, G.; Schütz, M. *J. Chem. Phys.* **2005,** *122,* 014103. doi:10.1063/1.1824898
55. Caldeweyher, E.; Bannwarth, C.; Grimme, S. *J. Chem. Phys.* **2017,** *147,* 034112. doi:10.1063/1.4993215
56. Becke, A. D.; Johnson, E. R. *J. Chem. Phys.* **2007,** *127,* 154108. doi:10.1063/1.2795701
57. Becke, A. D.; Johnson, E. R. *J. Chem. Phys.* **2007,** *127,* 124108. doi:10.1063/1.2768530
58. Grimme, S.; Hansen, A.; Brandenburg, J. G.; Bannwarth, C. *Chem. Rev.* **2016,** *116,* 5105–5154. doi:10.1021/acs.chemrev.5b00533
59. Aradi, B.; Hourahine, B.; Frauenheim, T. *J. Phys. Chem. A* **2007,** *111,* 5678–5684. doi:10.1021/jp070186p
60. Sun, Y.; Wilson, S. R.; Schuster, D. I. *J. Am. Chem. Soc.* **2001,** *123,* 5348–5349. doi:10.1021/ja0041730
61. Perepichka, D. F.; Wudl, F.; Wilson, S. R.; Sun, Y.; Schuster, D. I. *J. Mater. Chem.* **2004,** *14,* 2749–2752. doi:10.1039/b403509g
62. Schwerdtfeger, P. *Table of experimental and calculated static dipole polarizabilities for the electronic ground states of the neutral elements (in atomic units)*. http://ctcp.massey.ac.nz/Tablepol2015.pdf (accessed March 1, 2018).
63. Nakashima, N. *Sci. Technol. Adv. Mater.* **2006,** *7,* 609–616. doi:10.1016/j.stam.2006.08.004
64. Tomonari, Y.; Murakami, H.; Nakashima, N. *Chem. – Eur. J.* **2006,** *12,* 4027–4034. doi:10.1002/chem.200501176
65. Huang, K. *Statistical Mechanics,* 2nd ed.; Wiley: New York, 1987.
66. GESTIS-Stoffdatenbank. https://www.dguv.de/ifa/gestis/gestis-stoffdatenbank/index.jsp (accessed Aug 2, 2018).
67. Engineering ToolBox. https://www.engineeringtoolbox.com/ (accessed Aug 2, 2018).
68. ChemicalBook—Chemical Search Engine. https://www.chemicalbook.com/ProductIndex_EN.aspx (accessed Aug 2, 2018).
69. Lee, S. H.; Lee, H.; Pak, H. *Bull. Korean Chem. Soc.* **1997,** *18,* 501–509.
70. Lee, S. H.; Lee, S. K. *Bull. Korean Chem. Soc.* **1999,** *20,* 897–904.
71. Yoo, C.-D.; Kim, S.-C.; Lee, S.-H. *Bull. Korean Chem. Soc.* **2008,** *29,* 1059–1062. doi:10.5012/bkcs.2008.29.5.1059
72. Guttmann, R.; Sax, A. F., in preparation.
73. Schobert, H. *Chemistry of Fossil Fuels and Biofuels;* Cambridge University Press: Cambridge, UK, 2013.
74. Nishiyama, Y.; Langan, P.; Chanzy, H. *J. Am. Chem. Soc.* **2002,** *124,* 9074–9082. doi:10.1021/ja0257319
75. Nishiyama, Y.; Johnson, G. P.; French, A. D.; Forsyth, V. T.; Langan, P. *Biomacromolecules* **2008,** *9,* 3133–3140. doi:10.1021/bm800726v
76. French, A. D. *Cellulose* **2017,** *24,* 4605–4609. doi:10.1007/s10570-017-1450-3
77. Maurer, R. J.; Sax, A. F.; Ribitsch, V. *Cellulose* **2013,** *20,* 25–42. doi:10.1007/s10570-012-9835-9
78. Mazeau, K. *Carbohydr. Polym.* **2011,** *84,* 524–532. doi:10.1016/j.carbpol.2010.12.016
79. Da Silva Perez, D.; Ruggiero, R.; Morais, L. C.; Machado, A. E. H.; Mazeau, K. *Langmuir* **2004,** *20,* 3151–3158. doi:10.1021/la0357817
80. Hoja, J.; Maurer, R. J.; Sax, A. F. *J. Phys. Chem. B* **2014,** *118,* 9017–9027. doi:10.1021/jp5025685

81. Wang, Q. J.; Chung, Y.-W., Eds. *Enzyclopedia of Tribology;* Springer: Boston, MA, USA, 2013.
82. Coefficients of Friction. https://www.roymech.co.uk/Useful_Tables/Tribology/co_of_frict.htm (accessed May 4, 2018).

License and Terms