A Systematic Approach to Identify Cooperatively Bound Homotrimers

ARTICLE *in* THE JOURNAL OF PHYSICAL CHEMISTRY A · DECEMBER 2012

Impact Factor: 2.69 · DOI: 10.1021/jp310067m · Source: PubMed

2 AUTHORS, INCLUDING:

Jack Yang
University of Southampton
16 PUBLICATIONS **225** CITATIONS

THE JOURNAL OF PHYSICAL CHEMISTRY A

Article

pubs.acs.org/JPCA

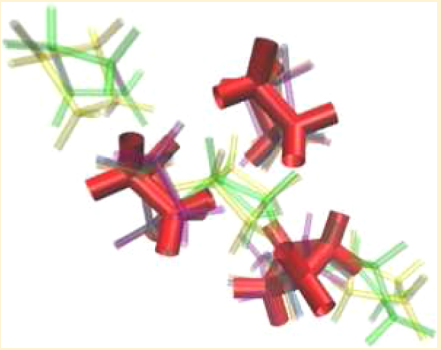
A Systematic Approach to Identify Cooperatively Bound Homotrimers

Jack Yang and Mark P. Waller*

Organisch-Chemisches Institut, Westfälische Wilhelms-Universität Münster, Corrensstraße 40, 48149 Münster, Germany

Supporting Information

ABSTRACT: A systematic multistage computational procedure is presented to investigate cooperativity within trimeric molecular complexes. It is an alternative to exhaustive or stochastic-based approaches. Trimeric clusters were extracted from known crystal structures and optimized in the gas phase, and subsequently filtered using energetic and RMSD structural cutoffs. Three-body interaction energies were computed for the subset of distinct low-energy trimer conformations. The procedure was validated using a set of 20 molecular crystals taken from the Cambridge Structural Database, and 25% of these strucutres gave rise to gas-phase homotrimers that showed a cooperative binding energy at the BP86-D3(BJ)/def2-SVP//TPSS-D3(BJ)/def2-TZVP level of theory.



INTRODUCTION

Noncovalent interactions (NCIs) are of crucial importance across many fields of science, such as supramolecular chemistry,[1] molecular recognition,[2] and nanostructure engineering.[3] The appearance of diverse structural conformations in molecular/cluster aggregates can be attributed to a delicate balance among various NCIs occurring between the aggregated components, which may influence their chemical and physical properties. Therefore, understanding the nature and energetic orderings of NCIs is essential for predicting the structures and properties of complex molecular systems.[4]

The many-body energy contribution could potentially determine the energetic ordering of the system when molecules aggregate together.[5] A particularly interesting case is the three-body interaction energy (TBE) in molecular trimers, which determines whether the binding of the dimer would facilitate/inhibit the binding of the third moiety. This gives rise to the concept of cooperativity, first proposed in the 1950s,[6] which has since gained increasing interest in chemical and biomolecular science.[7] For example, the cooperativity in ion-π systems is thought to stabilize the unexpected stacking geometry of tetrafluorophenyl substituents in the crystal structure of ferrocene derivatives.[8,9] Similarly, the growth kinetics of supramolecular polymers[10] could be driven by the cooperative effects in stacked aromatic cores.[11] Finally, cooperativity due to hydrogen bonding is important in understanding solvent−solute interactions,[12,13] which is required for developing accurate force fields in the area of macromolecular modeling.

From a theoretical perspective, understanding cooperative effects in supramolecular systems requires accurate calculations of the NCI energies. However, modeling NCIs to achieve so-called "chemical accuracy" has remained a long-term challenge for quantum chemistry,[14,15] particularly in the domain of density functional theory (DFT). Some notable efforts in improving the accuracy of DFT treatment upon NCI bounded complexes include Grimme's dispersion-corrected DFT[16,17] and Truhlar's Minnesota functionals.[18−20] The accuracy of these methods must be tested against high-level wave function results in benchmark studies. Examples of such benchmark sets include Grimme's GMTKN30,[21,22] Hobza's BEGDB,[23−26] and Schneebeli's gigantic NCI set.[27] Nevertheless, few theoretical benchmarking studies exist on investigating the density functional dependence of TBE in molecular systems.[28,29] Dimer conformations are often based on chemical intuition; however, trimer conformations are much more difficult to predict and pose a significant challenge. Most of the previous studies were focusing on trimers with high structural symmetry.[8,11,13,30,31]

However, constructing trimers without such symmetry constraints is a daunting task as the complexity of constituting monomers progressively increases. A method that is capable of indentifying cooperative trimer conformations at a moderate computational expense is clearly desirable. Many global optimization strategies have been developed over the years in searching global minimum-energy structures.[4,32−35] One of the key aspects in such a task is to generate a set of starting points,[33] either at random or from a high-temperature trajectory, for subsequent gradient-based local optimization. From a computational perspective, stochastic-based starting points are problematic due to the large number of possibilities if no constraints are applied. On the other hand, structure

Received: October 11, 2012
Revised: November 19, 2012
Published: December 11, 2012

generation from simulated-annealing-based approaches[36] would require special experitse in molecular dynamics simulations. Mahadevi et al. optimized clusters extracted from crystals to study cooperativity in acetamide clusters.[37] The authors found lower stability of the crystal-extracted clusters compared with that of circular or standard clusters. They concluded that the crystal starting point might lead to different preferences of aggregation patterns.

Extending the idea of considering the crystal structure database as a valuable tool for mining molecules,[31,38,39] we have developed a procedure to study the cooperativity of small-molecule homotrimers using crystal structures as a starting point. Our approach offers two distinct advantages; first, our procedure is much more efficient because the number of starting structures is significantly lowered compared to stochastic approaches. Second, because it is based on only three adjustable parameters, our approach serves as a versatile user-friendly tool that can be readily utilized by quantum chemists to study molecular cooperativity.

COMPUTATIONAL DETAILS

Procedure. On the basis of a starting monomer O extracted from a molecular crystal, a set of distinct trimers $\{\mathcal{T}\}$ can be generated by successively applying the symmetry operations and translation vectors of the crystal point group to the fractional coordinates of O. In this case, each trimer conformation $\mathcal{T}_i$ would consist of three distinct homodimer pairs $\{\mathcal{D}_j\}$ (j = 1, 2, 3). Denoting R_{min} as the minimum atom pairwise distance for a homodimer in $\{\mathcal{D}_j\}$, a valid trimer is selected based on the criteria such that at least two R_{min} must be less than the cutoff distance R_{cutoff}. All possible trimer conformations from a given molecular crystal structure were extracted within a certain cutoff radius. In order to achieve a balance with monomer size, trimer diversity, and computational power, R_{cutoff} was chosen to give a set of ∼50−100 starting conformations; therefore, the choice of R_{cutoff} was system-dependent. This procedure can be considered as a brute-force search in nature except that the search space granularity was predefined by lattice constants and symmetry operations.

The selected trimers were then freely optimized in the gas phase, with a subset of low-energy conformers selected based on a combined energy/root-mean-squared deviation (RMSD) filtering. RMSD values were computed using the quarternions-based algorithm from Coutsias et al.[76] For all trimers $\{\mathcal{T}'\}$ that were successfully optimized, the filtering process was performed as following:

1. Rank all trimers in $\{\mathcal{T}'\}$ based on the total energy of the final optimized structures.
2. Given that ΔE_k defines the difference between the total energy of conformation $k \in \{\mathcal{T}'\}$ and the lowest energy for the whole $\{\mathcal{T}'\}$ set, select the conformation with $\Delta E_k \leq E_{cutoff}$, where E_{cutoff} is a chosen energy cutoff value. This then provides us with a series of low-energy conformations $\{\mathcal{T}'_e\}$.
3. The RMSD filtering process is then performed for all trimer pairs (i,j) in $\{\mathcal{T}'_e\}$, and if $RMSD_{ij} < 0.5$ Å, then trimer j will be discarded from $\{\mathcal{T}'_e\}$. This would lead us to the final set of distinct low-energy trimer conformations.

Similar to R_{cutoff}, the choice of E_{cutoff} was also system-dependent. In general, E_{cutoff} was chosen to be 0.5 kcal/mol. However, in the case where the number of candidates below such a threshold was too low, then E_{cutoff} was raised. One possible way to identify such a threshold is from visual inspection of a cluster plot of ΔE_k against $RMSD_{k,0}$, where 0 denotes the conformation with the lowest total energy. This will be discussed in the following section.

Once a set of low-energy homotrimers had been identified, single-point energy calculations on the trimer were then performed to investigate the possible coorperative effects in optimized homotrimers. The cooperativity in an ABC trimer is indicated by the TBE term ΔE_{ABC}, which can be calculated as

$$\Delta E_{ABC} = E_{ABC} - (E_{AB} + E_{AC} + E_{BC}) + (E_A + E_B + E_C) \quad (1)$$

where E is the total energy of the system. The trimer can then be classified as either cooperative ($\Delta E_{ABC} < 0$), anticooperative ($\Delta E_{ABC} > 0$), or additive ($\Delta E_{ABC} \approx 0$).

Crystal Structures. Twenty molecular crystals were chosen from the Cambridge Structural Database (see Table1). The set includes examples to cover three important types of NCI, dispersion between aliphatic chains, aromatic π-stacking, and hydrogen bonding. Each crystal contained a complete monomer in its asymmetric unit.

Table 1. List of Small Molecular Crystals Selected for Trimer Generation[a]

molecule name	database code	number of conformations
1-carboxymethylthymine	ABEPIH	4
2-methoxy-3-nitrophenol	ADIPEJ	25
pyrazine *N*,*N*′-dioxide	AHEMAB	66
anthracene	ANTCEN07	10
8-azaguanine	AZGUAN	13
benzene	BENZEN	58
2,3-dihydroxybenzoic acid	CACDAM01	8
cyclohexane	CYCHEX	8
butane	DUCKOB	63
ethylene oxide	DUFBOV10	29
3-ethyladenine	ETADEN	9
ethane	ETHANE01	155
propane	JAYDUI	28
methylamine	METAMI	42
methanol	METHOL02	13
pentane	PENTAN01	19
pyridine *N*-oxide	PYRDNO11	17
cyclopropane	QQQCIS01	20
uracil	URACIL	12
cyclobutane	ZZZWEO02	6

[a]The number of distinct low-energy trimer conformations selected for each molecule is also listed.

Quantum Chemical Methods. Structure optimization and single-point energy calculations were performed with TURBOMOLE V6.3.1.[40] Resolution-of-identity[41,42] was used to speed up the calculations. The initial conformation searching was performed to find distinct low-energy conformations using the BP86[43−45] functional with Grimme's dispersion correction[46] [with Becke−Johnson (BJ) damping)[47]]. A small basis set of def2-SVP[48] was chosen to expedite the screening process. For comparison, calculations with BP86-D3(BJ)/def2-TZVP[48] were also performed on propane (see Figure S1 in the Supporting

dx.doi.org/10.1021/jp310067m | J. Phys. Chem. A 2013, 117, 174−182

Information), cyclobutane, and 8-azaguanine trimers, which filtered out a similar set of conformations compared to the BP86-D3(BJ)/def2-SVP level, and hence, the latter was applied for all subsequent optimizations. TBEs for all low-energy conformations were calculated with TPSS[49]-D3(BJ)/def2-TZVP. In principle, basis set superposition error may lead to an artificial increase in the cooperative effect. In regard to this, Antony et al.[28] showed that counterpoise (CP) correction has a minimum effect on the TBE calculated for nucleic acid trimers, while results from Mahadevi et al.[37] showed that it was important for calculating the interaction energy. However, the interaction energies calculated by Mahadevi et al.[37] were defined differently from the TBE defined by Antony et al.,[28] which is used in the present study. Therefore, we had chosen to not include CP correction in the present study.

To further validate our method, each of the 5 lowest-energy trimer conformations from 10 selected BP86-optimized trimer set was further optimized with the RI-MP2[50]/cc-pVDZ[51] method, which had been reported to result in accurate geometries compared to RI-MP2/cc-pVTZ.[36] The reoptimized geometries were compared with the starting BP86 geometries, and the results are shown in Table 2. It can be seen that on average, the RMSD between the MP2- and BP86-optimized structures was very low (0.124 Å across all 10 sets on average), indicating that BP86-D3(BJ)/def2-SVP is a reasonable choice for fast optimization of trimer conformations.

Table 2. Averaged RMSD between the Five Lowest-Energy BP86-D3(BJ)/def2-SVP-Optimized and MP2/cc-pVDZ-Reoptimized Trimer Conformations for 10 Selected Sets

database code	RMSD (Å)
ETHANE01	0.109 ± 0.075
JAYDUI	0.119 ± 0.035
DUCKOB	0.092 ± 0.009
QQQCIS01	0.114 ± 0.032
ZZZWEO02	0.212 ± 0.074
BENZEN	0.048 ± 0.008
METHOL02	0.109 ± 0.001
AZGUAN	0.151 ± 0.054
METAMI	0.192 ± 0.006
PYRDNO11	0.093 ± 0.004

The TBE for all reoptimized trimers were then calculated with the RI-MP2/def2-TZVPP method as reference values. MP2 is a computationally efficient method, which recovers a major part of electron correlation that is important for accurately describing noncovalent interactions. Although MP2 tends to overestimate the dispersion interactions by 10−20%,[52,53] it should not drastically alter the TBE orderings to be used as a cooperative motif filter. The reproducibility of the MP2 TBE values was further tested with 11 different dispersion-corrected DFTs, including one LDA functional SVWN,[54−57] five GGA functionals BLYP,[58,59] BVWN,[57,58] BP86, PBE[60] and B97D,[17] one meta-GGA functional TPSS, three hybrid functionals B3LYP,[61,62] PBE0,[63,64] and TPSSh,[49] together with double-hybrid functional B2PLYP.[65] All DFT calculations were performed with the def2-TZVPP basis set. The averaged mean absolute deviations (MADs) for each functional are shown in Figure 1. In general, the Jacob's ladder ordering could be found.[66] The VWN-based functionals led to averaged MADs of more than 2.5 kcal/mol for the TBEs. This was primarily due to the failure in treating the hydrocarbon

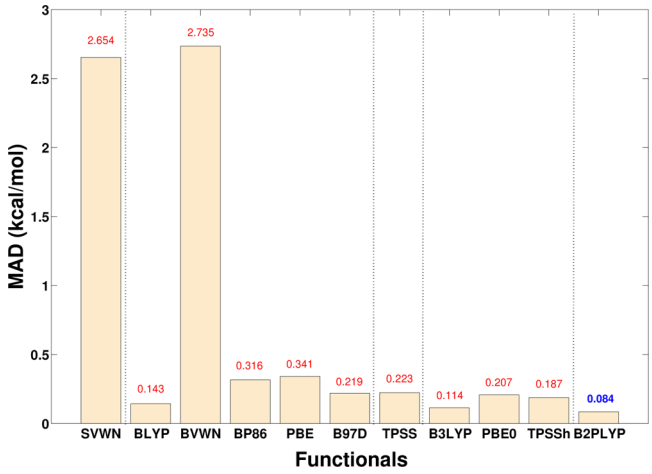


Figure 1. MADs of TBE for 10 reoptimized trimer sets tested with 11 different dispersion-corrected DFT methods, compared with MP2 results. All DFT calculations were performed with D3 corrections including BJ damping. The dotted lines separate functionals belonging to different rungs of Jacob's ladder.

trimers bounded by dispersion interactions (Figure 2, left), in which the absolute values of TBE had been significantly overestimated (up to ∼24 kcal/mol). However, their performances for heteroatomic systems (Figure 2, right) were reasonable. The double-hybrid B2PLYP-D3 performed best over all functionals tested with a MAD of 0.084 kcal/mol; however, it would be too expensive to be used for large-scale searching purposes. The most popular hybrid B3LYP-D3 functional gave the next lowest MAD of 0.114 kcal/mol. The overall MAD for TPSS-D3 was 0.223 kcal/mol.

The basis set dependence of RI-MP2 reference values was further tested with the METHOL02 set by extrapolating the TBE values from aug-cc-pVQZ[77] and aug-cc-pV5Z toward the complete basis set (CBS) limit, with the extrapolation scheme from Helgaker et al.[78] All five MP2-optimized conformations were tested. For methanol trimers, a 5% increase in TBE from −3.00 (MP2/def2-TZVPP) to −3.15 kcal/mol (MP2/CBS) was observed. This indicated that an incomplete basis set could lead to a slight underestimation of TBE.

It was shown previously that the semilocal TPSS functional in combination with dispersion correction was capable of providing good computational results with moderate costs.[67,68] Considering ±0.5 kcal/mol as a cutoff between additive and nonadditive TBE and taking into account a <1 kcal/mol error to be the generally accepted definition of chemical accuracy, the meta-GGA TPSS-D3 can be regarded as a good and affordable method for searching cooperative structural motifs.

Therefore, the proposed procedure in searching for cooperative structural motifs can be validated as a relatively cheap but reasonable computational method [BP86-D3(BJ)/def2-SVP//TPSS-D3(BJ)/def2-TZVP].

RESULTS AND DISCUSSIONS

Trimer Conformations. For all 2464 starting trimer conformations generated, an overall 75% success rate was achieved in optimizing the conformations with BP86-D3(BJ)/def2-SVP; these optimized conformations were subjected to subsequent filterings. While for some sets, such as ethane (351/381) and methanol (106/109), a success rate over 90% could be observed, the convergence for benzene (120/228) and

dx.doi.org/10.1021/jp310067m | J. Phys. Chem. A 2013, 117, 174−182

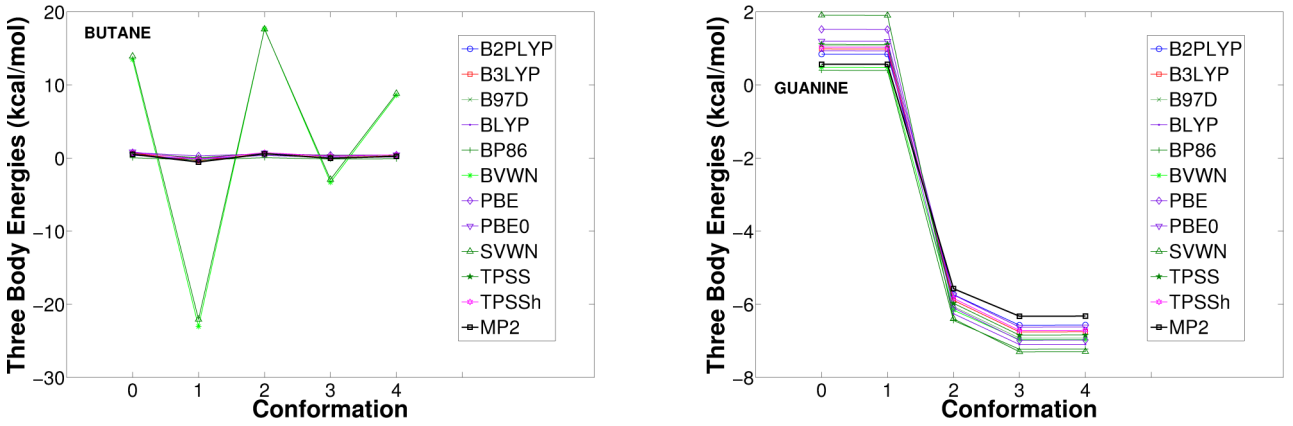


Figure 2. Functional dependence of TBE of the five lowest-energy conformations for (left) butane and (right) azaguanine trimers. Dispersion correction was used in all cases.

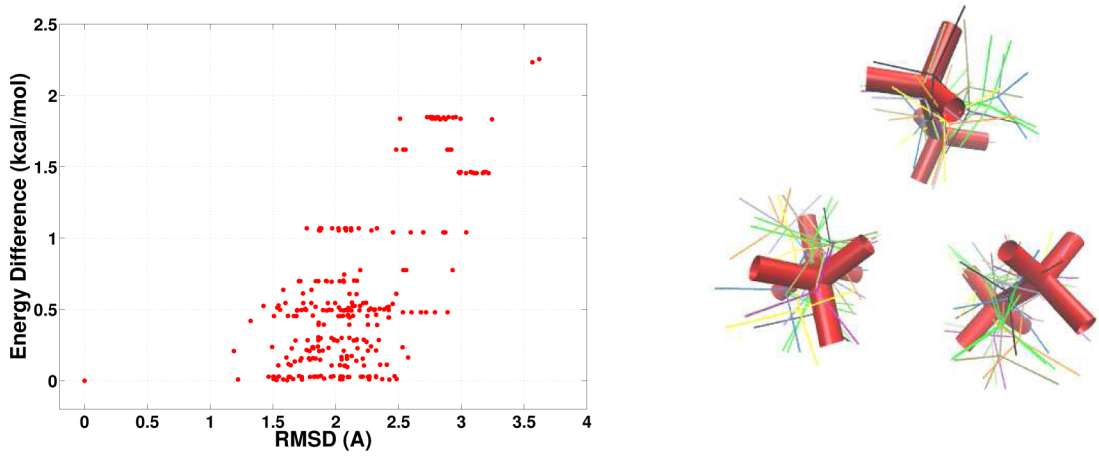


Figure 3. Results of conformation prescreening for ethane trimers. (left) Distributions of the final structures with energy and RMSD values compared to the conformations with the lowest energy among all optimized conformations. All optimizations were performed with BP86-D3(BJ)/def2-SVP. Low-energy conformations can be selected among conformations with energy differences less than 0.5 kcal/mol. Similar conformations can further be filtered out among the low-energy conformations with a pairwise RMSD less than 0.5 A. (right) Overlay of 11 low-energy unique trimer conformations for propane. The conformation with the lowest energy is represented by thick red tubes.

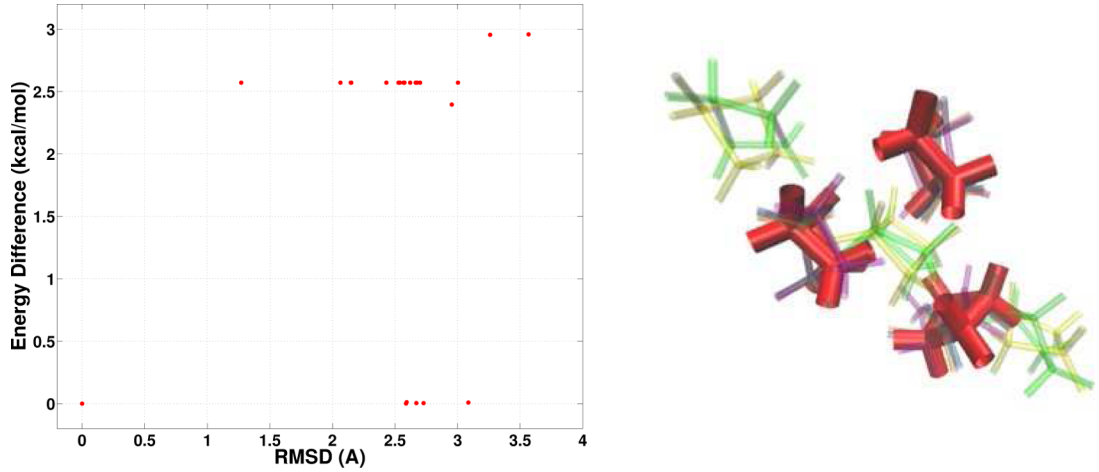


Figure 4. Results of conformation searching for cyclobutane trimers. (left) Distributions of the final structures with energy and RMSD values compared to the conformations with the lowest energy among all optimized conformations. All optimizations were performed with BP86-D3(BJ)/def2-SVP. Low-energy conformations can be selected among conformations with energy differences less than 0.2 kcal/mol. Similar conformations can further be filtered out among the low-energy conformations with a pairwise RMSD less than 0.5 A. (right) Overlay of unique trimer conformations for cyclobutane. The conformation with the lowest energy is represented by thick red tubes. Only six distinct low-energy structures can be identified, in this case.

dx.doi.org/10.1021/jp310067m | J. Phys. Chem. A 2013, 117, 174−182

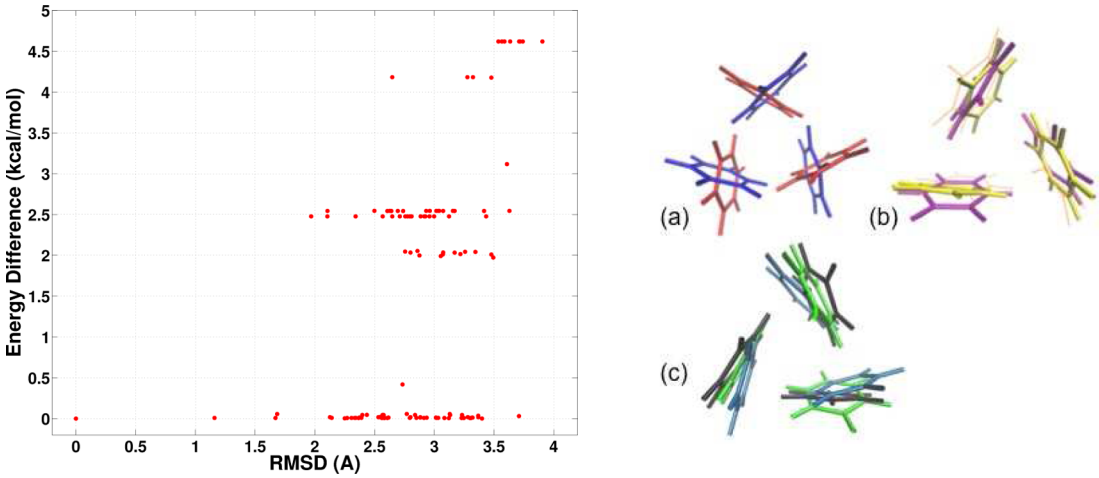


Figure 5. Results of conformation searching for benzene. (left) Distributions of the final structures with energy and RMSD values compared to the conformations with the lowest energy among all optimized conformations. All optimizations were performed with BP86-D3(BJ)/def2-SVP. Low-energy conformations can be selected among conformations with energy differences less than 0.5 kcal/mol. Similar conformations can further be filtered out among the low-energy conformations with a pairwise RMSD less than 0.5 Å. (a–c) Three groups of low-energy benzene trimer conformations, in which (a) represents two lowest-energy trimer conformations, which are mirror images of one another.

cyclobutane (23/76) was problematic, indicating that a higher level of theory, such as using larger basis sets including diffuse functions, might be necessary for treating these systems. However, the major focus of the present study is to demonstrate the methodology developed as a fast screening technique to efficiently construct a diverse set of trimer conformations by using crystal structures as starting points.

Combined energy/RMSD filtering is an effective method to identify unique conformations from a pool of candidates, and it has been applied to predict protein conformations in simulated docking/folding experiments.[69] An example of a cluster ΔE_k–$RMSD_{k,0}$ plot for identifying unique conformations among optimized ethane trimers is given in Figure 3 (left). A large number of low-energy conformations (ΔE_k < 0.5 kcal/mol) could be identified in the lower region of the cluster plot. These conformations spanned the RMSD range from 1.2 to 2.9 Å, indicating that most of them were quite structurally distinct from the lowest-energy trimer. Two trimer conformations were considered to be identical when pairwise RMSD < 0.5 Å. On the basis of the filtering criteria, 233 low-energy conformations were indentified, in which only 6 were discarded due to low pairwise RMSD values. An overlay of 11 low-energy trimer conformations for ethane are shown in Figure 3b.

The number of low-energy conformations identifiable from the cluster plots varied significantly among the 20 selected sets, which can be seen from the number of low-energy conformations selected from combined energy/RMSD filtering; see Table 1. The number of conformations seems to be related to the internal structural rigidity of each monomer, as well as the nature of intermolecular interactions within a trimer. In general, most lowest-energy conformations tended to adopt a cyclic rather than a linear arrangement (Figure 4), which maximized the pairwise noncovalent interactions for each dimer pair. This is also in agreement with previous reports that linear clusters of acetamide were less stable than the cyclic conformations.[37]

The trimer conformations for hydrocarbons exhibited strong dependency on the size and shape of the monomers. For saturated hydrocarbons, acyclic chains typically led to more trimer conformations, compared with the cyclic structures, due to the extra rotational flexibility of the sp^3 carbon in acyclic structures. This could be seen from Figure S2 (Supporting Information (SI)) for butane and Figure 4 for cyclobutane trimers. On the other hand, the number of good trimer conformations decreased as the molecular weight increased in both linear and cyclic saturated hydrocarbons. This was particularly true for the acyclic cases. As the molecule then became more rod-like in shape with increasing molecular weight, the carbon backbones tended to align parallel with the large surface areas facing each other, maximizing the dispersion interactions among the monomers, whereas for smaller molecules, such an interaction was more isotropic, giving rise to larger diversity in the final trimer conformations by allowing each monomer to freely rotate around its own center-of-mass. (see Figures 3 and S1–S3 in the SI).

The results for aromatic hydrocarbons, on the other hand, were quite different from the cases of saturated cyclic hydrocarbons. For benzene (Figure 5), 58 unique trimer conformations were identified below a cutoff energy value of 0.5 kcal/mol. The two conformations of the lowest energies were essentially mirror images of one another (Figure 5a). The number of low-energy conformations was also found to be greater than that in the case of saturated cyclic hydrocarbons (such as cyclohexane). This is probably because the planar aromatic rings could slide/tilt against each other more freely than an interlocked saturated cyclic hydrocarbon. It was also found that the lowest-energy conformations for benzene trimers tended not to adopt a π-stacked but a tilted-T-shape arrangement. Previous studies[70] on the π-interaction showed that for the benzene dimer, there was no strong energetic preference for T-shaped arrangements compared to stack dimers. However, as the number of rings increased in an aromatic system, the stacked arrangements became favored in order to maximize the dispersion interactions, whereas for benzene, the attractive dispersion interaction was compensated for by the repulsive Pauli exchange, making the dominant interactions be electrostatic.[71] This was indeed found for the case of anthracene trimers (Figure S6 in the SI) where the 10 lowest-energy conformations all adopted a certain kind of tilted-stack geometry. A recent survey of aromatic clusters in

dx.doi.org/10.1021/jp310067m | J. Phys. Chem. A 2013, 117, 174–182

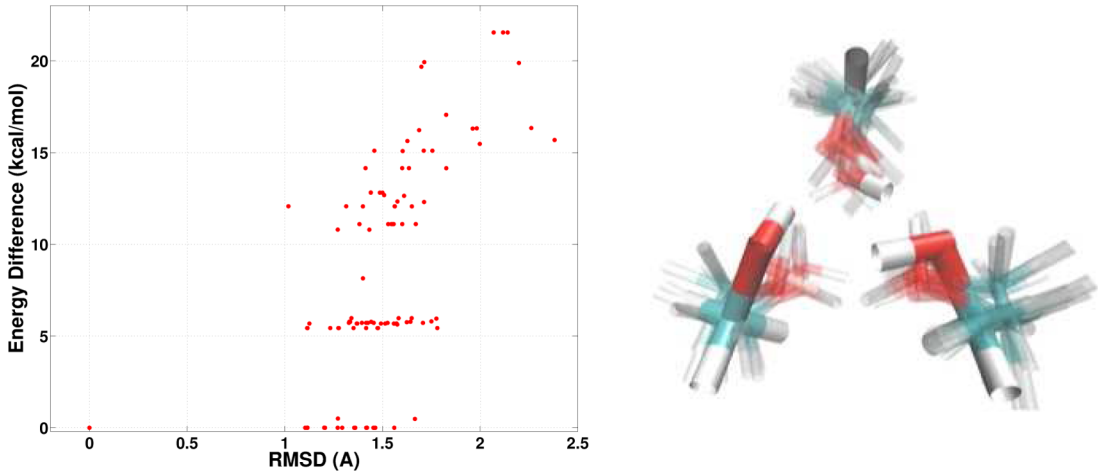


Figure 6. Results of conformation searching for methanol trimers. (left) Distributions of the final structures with energy and RMSD values compared to the conformations with the lowest energy among all optimized conformations. All optimizations were performed with BP86-D3(BJ)/def2-SVP. Low-energy conformations can be selected among conformations with energy differences less than 1 kcal/mol. Similar conformations can further be filtered out among the low-energy conformations with a pairwise RMSD less than 0.5 Å. (right) Overlay of 10 unique trimer conformations for methanol. The conformation with the lowest energy is represented by thick bonds.

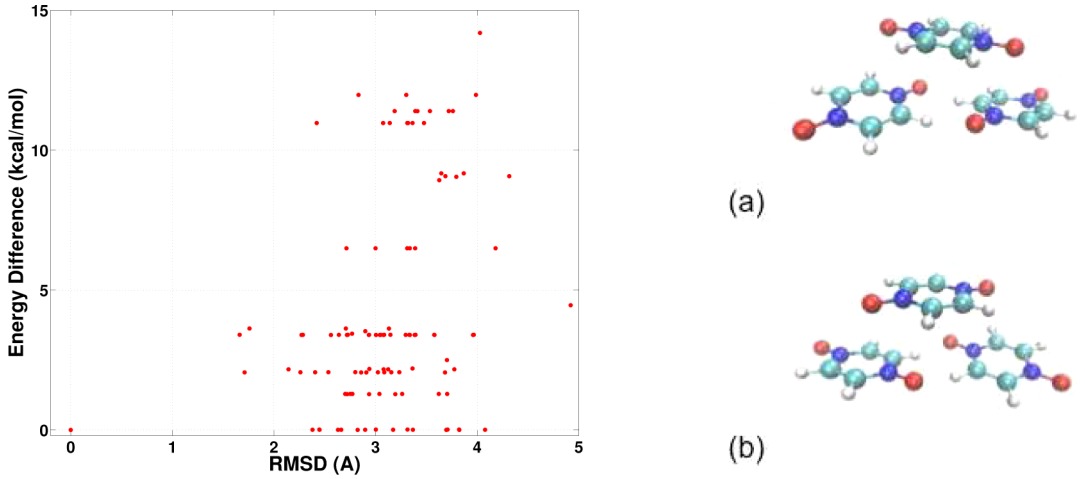


Figure 7. Results of conformation searching for PNO trimers. (left) Distributions of the final structures with energy and RMSD values compared to the conformations with the lowest energy among all optimized conformations. All optimizations were performed with BP86-D3(BJ)/def2-SVP. Low-energy conformations can be selected among conformations with energy differences less than 4 kcal/mol. Similar conformations can further be filtered out among the low-energy conformations with a pairwise RMSD less than 0.5 Å. (right) (a,b) Two lowest-energy conformations, which are essentially mirror images of one another.

protein structures deposited in the protein structure database also revealed a strong tendency for aromatic side chains to cluster in an edge-to-face-like fashion,[72] which further suggested that our procedure led to trimer conformations that are chemically sensible.

Intermolecular hydrogen bonding played an important role in determining the trimer conformations upon introducing electronegative atoms, such as O and N, into the monomer. This provided a more diverse test set to examine the dominant NCI in determining low-energy trimer conformations as well as the corresponding TBE orderings. For the simplest example, the lowest-energy cyclic conformation of the methanol trimer (Figure 6) had each hydrogen of the hydroxyl groups pointing directly to an oxygen in a neighboring molecule. This, as expected, would be a preferred conformation in a hydrogen-bonded complex. More importantly, the existence of such a preferred bonding pattern meant that each methanol monomer could not freely rotate itself about its own center-of-mass, as for the case of ethane trimers. This might have significantly lowered the number of low-energy homotrimer conformations for methanol.

It was also interesting to observe different results for a pair of heterocyclic structural isomers. Figures 7 and 8, for examples, demonstrate results for trimers of pyrazine *N*,*N*′-dioxide (PNO) and uracil. It was found that there were significantly more low-energy (<4 kcal/mol) conformers for PNO than for uracil. The lowest-energy conformer of the PNO trimers had a mirror image, which was the same in total energy but with a RMSD value greater than 2 Å. These trimers preferred a bridged structure, where the two *para*-oxygen atoms in the central PNO each interacted with a π-ring from a different monomer. This type of conformation led to more freedom for the pyrazine rings to tilt, which was probably the reason for the larger number of conformers found. In the case of the uracil trimer, however, there was an energy gap of ∼2 kcal/mol between the lowest-energy conformation and the next-lowest-

dx.doi.org/10.1021/jp310067m | J. Phys. Chem. A 2013, 117, 174−182

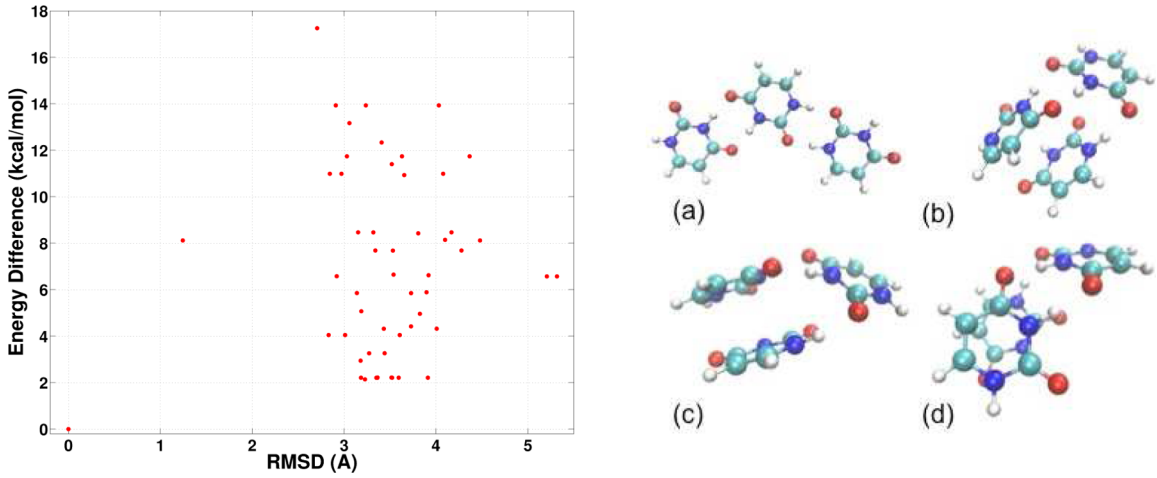


Figure 8. Results of conformation searching for uracil trimers. (left) Distributions of the final structures with energy and RMSD values compared to the conformations with the lowest energy among all optimized conformations. All optimizations were performed with BP86-D3(BJ)/def2-SVP. Low-energy conformations can be selected among conformations with energy differences less than 4 kcal/mol. Similar conformations can further be filtered out among the low-energy conformations with a pairwise RMSD less than 0.5 Å. (right) (a–d) Four lowest-energy conformations.

energy one, and these conformations were structurally very distinct from each other. The oxygen atoms in uracil are at the meta positions of the six-member ring. For the lowest-energy conformer, favorable hydrogen bondings occurred between the carbonyl groups and the −NH groups in two pairs of uracil dimers, which were very localized in space. The next-lowest-energy conformation of the uracil trimer involved a stacked geometry for a pair of heterocyclic rings, which presumably raised the total energy of the trimer due to Pauli repulsion between the stacked monomers.

This procedure was applied to the remaining molecules (listed in Table 1) to produce a relativley large set of homotrimers. As the monomeric structure became increasingly complex, it is presumably less likely to predict a diverse set of homotrimer conformations from chemical intuition.

Cooperativity in Optimized Homotrimers. The TBE for the filtered subset of distinct low-energy trimer conformations was calculated to identify structures that showed cooperative interaction energies. The averaged TBE for all low-energy trimers (as listed in Table 1) for each test set is presented in Figure 9. In the following discussion, $|\Delta E_{ABC}|$ = 0.5 kcal/mol was taken as a threshold for additivty. On the basis of such a criteria, it could be seen from the low averaged ΔE_{ABC} values in Figure 9 that 15 sets of trimers were mostly bound in an additive manner. The standard deviations were quite low (∼0.1 kcal/mol) for this subset of complexes. More specifically, the majority of hydrocarbons, including saturated acyclic, saturated cyclic, and aromatic trimers were found to be additive. The dominant interactions in these hydrocarbon trimers were dispersion interactions, which were weaker than hydrogen bonding; this could be the main reason for the low averaged TBE observed across the hydrocarbon sets. The TBEs for cyclic benzene trimers, for instance, were found to have an averaged TBE of −0.44 ± 0.15 kcal/mol, compared to the previously calculated TBE of −0.33 kcal/mol with MP2/cc-pVDZ+ for the cyclic benzene trimer.[73] Previous investigations on stacked thiophene trimers revealed weak additivity,[74] which can be attributed to the cancellation of the inductive component by dispersion interactions. Strong cooperativity might occur in ion−aromatic complexes due to electrostatic and ion-induced polarization.[8]

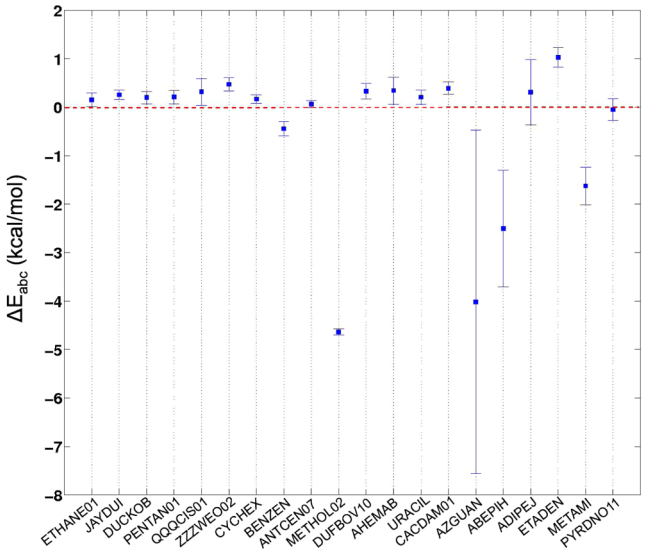


Figure 9. Mean three-body energies (ΔE_{ABC} in kcal/mol) calculated with TPSS-D3(BJ)/def2-TZVP for all 20 sets of trimers (optimzied with BP86-D3(BJ)/def2-SVP), each averaged over all low-energy conformations. The error bar indicates the standard deviation for each set.

The cooperative trends in trimers involving atoms other than hydrogen and carbons showed more variations. For example, in the simplest cases of methanol and ethylene oxide trimers, it was found that all 13 trimers for the former case were strongly cooperative (averaged TBE of −4.64 kcal/mol with a narrow standard deviation of 0.07 kcal/mol), but the later case was found to be additive for all 29 trimers (TBE of 0.33 ± 0.16 kcal/mol). In both cases, it can be seen that the hydrogen bond donors were directly pointing to the hydrogen bond acceptors in the neighboring monomer within a cyclic trimer conformation (Figures 6 and S7 in the SI). The electronegativity of the hydrogen bond acceptors has seemingly played a role in determining the TBE orderings for such a simple molecular trimer, whereby the hydrogen atom in an −OH group for methanol was more positive than the one in the −CH group for ethylene oxide. Mandal et al.[13] analyzed the cooperativity in

dx.doi.org/10.1021/jp310067m | J. Phys. Chem. A 2013, 117, 174−182

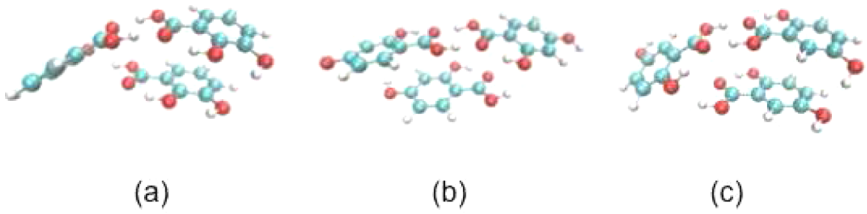


Figure 10. Lowest-energy trimer conformations for (a) 2,3-, (b) 2,4-, and (c) 2,5-dihydroxybenzoic acid.

methanol and water trimers. They also showed that the hydrogen bonding in the methanol trimer was stronger than that in the case of water trimers because methanol is a good hydrogen bond acceptor. Their estimated TBE for cyclic methanol trimer was −2.02 kcal/mol at the MP2/6-31++G(d,p) level. Our current findings can be further supported from the cooperative TBE found for methylamine trimers, with an average TBE of −1.62 ± 0.39 kcal/mol, a value that is between the two extreme cases of methanol and ethylene oxides. This matches nicely with the trend of electronegativities for C, N, and O.

For molecules with more complex structures, the cooperativity for homotrimers became difficult to predict. The guanine trimers, for instance, showed strong cooperative behavior with an averaged TBE of −4.02 kcal/mol. However, a large deviation of 3.54 kcal/mol was also seen among the 13 low-energy trimers. In particular, the two lowest total energy trimers were found to be anticooperative (TBE > 1 kcal/mol); these conformations have a π-stacked dimer and a third monomer perpendicular to the stacked dimer (Figure S13a and b, SI). The third lowest-energy conformations, on the other hand, showed a strong TBE of −7.22 kcal/mol, with the third monomer lying parallel to the stack dimer (Figure S13c, SI). This conformation had a total energy of about 5 kcal/mol higher than that of the lowest-energy conformation. The strongest cooperative guanine trimers were found to be in a cyclic arrangement (Figure S13d, SI) with a TBE of −8.75 kcal/mol. This highlights that the lowest-energy trimer conformations might not necessarily be a cooperatively bound system, and a rather comprehensive search strategy is essential.

The current findings suggested that parallel π-stacking might not be a good structural motif for promoting cooperative bindings in gas-phase homotrimers. For examples, the lowest-energy anthracene trimers, which adopted predominantly stacked geometries (Figure S7, SI) showed additive TBE (Figure S27, SI), similarly for the cases of pyrazine oxide and uracil trimers. On the other hand, homotrimers for 1-carboxymethylthymine showed strong cooperative binding. It can be seen from the low-energy conformations for the 1-carboxymethylthymine trimers that the carboxyl group was responsible for forming the hydrogen bonds among the monomers by tilting itself perpendicularly to the thymine ring, which prohibited the thymine ring from becoming aligned in parallel, thus minimizing the Pauli repulsion between stacked thymine rings (Figure S14, SI). On the other hand, for the 2-methoxy-3-nitrophenol trimers, the substituents were less bulky and also less efficient in promoting strong hydrogen bondings, compared with the carboxyl group; thus, a π-stacking geometry can be found in the homotrimer (Figure S15, SI), which significantly raised the TBE.

Finally, we investigated the cooperativity in gas-phase dihydroxybenzoic acid trimers. Our results for trimers of 2,3-dihydroxybenzoic acid revealed only additivity (averaged TBE of 0.392 ± 0.127 kcal/mol). In most cases, the dimer association from two −COOH groups through intermolecular hydrogen bondings was preferred, while the third molecule can only adopt a stacked geometry with one of the monomers in the hydrogen-bonded dimer (Figure 10a). For comparison, we performed the same procedure for the isomeric 2,4- and 2,5-dihydroxybenzoic acids based on the crystal structures provided by Adam et al.,[75] and we did not observe a significant difference compared with the 2,3-substituted case (Figure 10b and c). This showed that even though intermolecular dihydrogen bonds between the −COOH motifs in a pair of benzoic acid dimer are strongly preferred, there existed no other free −COOH group in the associated dimer to form effective hydrogen bonding with the third moiety. As a result, we observed a stacking arrangement from the third monomer, which led to TBE close to zero. We speculate that a dicarboxylic acid would be more effective in promoting cooperative interactions by offering two −COOH motifs to two other molecules.

CONCLUSIONS

In summary, we have developed a systematic approach for constructing homotrimers based on available molecular crystal structures. Although using crystal structures significantly lowers the number of starting structures, by setting a relatively large cutoff for the distance constraint, the probability for biasing toward a particular aggregation pattern is reduced. Similarly, because it is based on only three adjustable parameters, our approach serves as a versatile tool that can be readily utilized by quantum chemists. Overall, our approach proved to be a useful method for discovering cooperativity in homotrimers.

ASSOCIATED CONTENT

Supporting Information

Complete ref 3, combined energetic/RMSD filtering results for the remaining 18 sets of trimers, as well as the TBE for all of the low-energy trimers identified in Figure 1. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author

*E-mail: m.waller@uni-muenster.de.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

Generous financial support by the Deutsche Forschungsgemeinschaft (SFB858) is gratefully acknowledged.

REFERENCES

(1) Schneider, H.-J. *Angew. Chem., Int. Ed.* **2009**, *48*, 3924−3977.

(2) Meyer, E. A.; Castellano, R. K.; Diederich, F. *Angew. Chem., Int. Ed.* **2003**, *42*, 1210−1250.
(3) French, R. H.; Parsegian, V. A.; Podgornik, R.; Rajter, R. F.; Jagota, A.; Luo, J.; Asthagiri, D.; Chaudhury, M. K.; Chiang, Y.-m.; Granick, S.; Kalinin, S.; Kardar, M.; Kjellander, R.; Langreth, D. C.; Lewis, J.; et al. *Rev. Mod. Phys.* **2010**, *82*, 1887−1944.
(4) Catlow, C. R. A.; Bromley, S. T.; Hamad, S.; Mora-Fonz, M.; Sokol, A. A.; Woodley, S. M. *Phys. Chem. Chem. Phys.* **2010**, *12*, 786−811.
(5) Elrod, M. J.; Saykally, R. J. *Chem. Rev.* **1994**, *94*, 1975−1997.
(6) Frank, H.; Wen, W. *Discuss. Faraday Soc.* **1957**, *24*, 133−140.
(7) Hirschberg, J.; Brunsveld, L.; Ramzi, A.; Vekemans, J.; Sijbesma, R.; Meijer, E. *Nature* **2000**, *407*, 167−170.
(8) Frontera, A.; Quiñonero, D.; Costa, A.; Ballester, P.; Deyá, P. M. *New J. Chem.* **2007**, *31*, 556−560.
(9) Blanchard, M. D.; Hughes, R. P.; Concolino, T. E.; Rheingold, A. L. *Chem. Mater.* **2000**, *12*, 1604−1610.
(10) Markvoort, A.; ten Eikelder, H.; Hilbers, P.; de Greef, T.; Meijer, E. *Nat. Commun.* **2011**, *2*, 209.
(11) Kulkarni, C.; Reddy, S. K.; George, S. J.; Balasubramanian, S. *Chem. Phys. Lett.* **2011**, *515*, 226−230.
(12) Deshmukh, M. M.; Bartolotti, L. J.; Gadre, S. R. *J. Phys. Chem. A* **2008**, *112*, 312−321.
(13) Mandal, A.; Prakash, M.; Kumar, R. M.; Parthasarathi, R.; Subramanian, V. *J. Phys. Chem. A* **2010**, *114*, 2250−2258.
(14) Riley, K. E.; Pitoňák, M.; Jurečka, P.; Hobza, P. *Chem. Rev.* **2010**, *110*, 5023−5063.
(15) Cohen, A. J.; Mori-Sánchez, P.; Yang, W. *Chem. Rev.* **2012**, *112*, 289−320.
(16) Grimme, S. *J. Comput. Chem.* **2004**, *25*, 1463−1473.
(17) Grimme, S. *J. Comput. Chem.* **2006**, *27*, 1787−1799.
(18) Zhao, Y.; Truhlar, D. G. *J. Chem. Theory Comput.* **2007**, *3*, 289−300.
(19) Zhao, Y.; Truhlar, D. G. *Chem. Phys. Lett.* **2011**, *502*, 1−13.
(20) Zhao, Y.; Truhlar, D. G. *Acc. Chem. Res.* **2008**, *41*, 157−167.
(21) Goerigk, L.; Grimme, S. *J. Chem. Theory Comput.* **2010**, *6*, 107−126.
(22) Goerigk, L.; Grimme, S. *J. Chem. Theory Comput.* **2011**, *7*, 291−309.
(23) Jurečka, P.; Šponer, J.; Černý, J.; Hobza, P. *Phys. Chem. Chem. Phys.* **2006**, *8*, 1985−1993.
(24) Valdes, H.; Pluháčkova, K.; Pitoňák, M.; Řezáč, J.; Hobza, P. *Phys. Chem. Chem. Phys.* **2008**, *10*, 2747−2757.
(25) Gráfová, L.; Pitoňák, M.; Řezáč, J.; Hobza, P. *J. Chem. Theory Comput.* **2010**, *6*, 2365−2376.
(26) Řezáč, J.; Hobza, P. *J. Chem. Theory Comput.* **2012**, *8*, 141−151.
(27) Schneebeli, S. T.; Bochevarov, A. D.; Friesner, R. A. *J. Chem. Theory Comput.* **2011**, *7*, 658−668.
(28) Antony, J.; Brüske, B.; Grimme, S. *Phys. Chem. Chem. Phys.* **2009**, *11*, 8440−8447.
(29) de Lange, K.; Lane, J. *J. Chem. Phys.* **2011**, *135*, 064304.
(30) Alkorta, I.; Blanco, F.; Elguero, J.; Estarellas, C.; Frontera, A.; Quiñonero, D.; Deyá, P. M. *J. Chem. Theory Comput.* **2009**, *5*, 1186−1194.
(31) Prohens, R.; Portell, A.; Puigjaner, C.; Barbas, R.; Alcobé, X.; Font-Bardia, M.; Tomás, S. *CrystEngComm* **2012**, *14*, 5745−5748.
(32) Woodley, S.; Catlow, R. *Nat. Mater.* **2008**, *7*, 937−946.
(33) Hartke, B. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2011**, *1*, 879−887.
(34) Dieterich, J. M.; Hartke, B. *Mol. Phys.* **2010**, *108*, 279−291.
(35) Takeuchi, H. *J. Chem. Inf. Model.* **2007**, *47*, 104−109.
(36) Kabeláč, M.; Valdes, H.; Sherer, E. C.; Cramer, C. J.; Hobza, P. *Phys. Chem. Chem. Phys.* **2007**, *9*, 5000−5008.
(37) Mahadevi, A. S.; Neela, Y. I.; Sastry, G. N. *Phys. Chem. Chem. Phys.* **2011**, *13*, 15211−15220.
(38) Musumeci, D.; Hunter, C. A.; Prohens, R.; Scuderi, S.; McCabe, J. F. *Chem. Sci.* **2011**, *2*, 883−890.
(39) Li, H.; Lu, Y.; Liu, Y.; Zhu, X.; Liu, H.; Zhu, W. *Phys. Chem. Chem. Phys.* **2012**, *14*, 9948−9955.
(40) Ahlrichs, R.; Bär, M.; Häser, M.; Horn, H.; Kölmel, C. *Chem. Phys. Lett.* **1989**, *162*, 165−169.
(41) Eichkorn, K.; Weigend, F.; Treutler, O.; Ahlrichs, R. *Theo. Chem. Acc.* **1997**, *97*, 119−124.
(42) Sierka, M.; Hogekamp, A.; Ahlrichs, R. *J. Chem. Phys.* **2003**, *118*, 9136−9148.
(43) Becke, A. D. *Phys. Rev. A* **1988**, *38*, 3098−3100.
(44) Perdew, J. P. *Phys. Rev. B* **1986**, *33*, 8822−8824.
(45) Perdew, J. P. *Phys. Rev. B* **1986**, *34*, 7406−7406.
(46) Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. *J. Chem. Phys.* **2010**, *132*, 154104.
(47) Grimme, S.; Ehrlich, S.; Goerigk, L. *J. Comput. Chem.* **2011**, *32*, 1456−1465.
(48) Weigend, F.; Ahlrichs, R. *Phys. Chem. Chem. Phys.* **2005**, *7*, 3297−3305.
(49) Tao, J.; Perdew, J. P.; Staroverov, V. N.; Scuseria, G. E. *Phys. Rev. Lett.* **2003**, *91*, 146401.
(50) Møller, C.; Plesset, M. S. *Phys. Rev.* **1934**, *46*, 618−622.
(51) Thom H. Dunning, J. *J. Chem. Phys.* **1989**, *90*, 1007−1023.
(52) Cybulski, S. M.; Lytle, M. L. *J. Chem. Phys.* **2007**, *127*, 141102.
(53) Heßelmann, A.; Jansen, G.; Schütz, M. *J. Chem. Phys.* **2005**, *122*, 014103.
(54) Hohenberg, P.; Kohn, W. *Phys. Rev.* **1964**, *136*, B864−B871.
(55) Kohn, W.; Sham, L. J. *Phys. Rev.* **1965**, *140*, A1133−A1138.
(56) Slater, J. C. *The Self-Consistent Field for Molecular and Solids, Quantum Theory of Molecular and Solids*; McGraw-Hill: New York, 1974; Vol. *4*.
(57) Vosko, S. H.; Wilk, L.; Nusair, M. *Can. J. Phys.* **1980**, *58*, 1200−1211.
(58) Becke, A. D. *Phys. Rev. A* **1988**, *38*, 3098−3100.
(59) Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785−789.
(60) Perdew, J. P.; Burke, K.; Ernzerhof, M. *Phys. Rev. Lett.* **1996**, *77*, 3865−3868.
(61) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648−5652.
(62) Stephens, P. J.; Devlin, F. J.; Chabalowski, C. F.; Frisch, M. J. *J. Phys. Chem.* **1994**, *98*, 11623−11627.
(63) Ernzerhof, M.; Scuseria, G. E. *J. Chem. Phys.* **1999**, *110*, 5029−5036.
(64) Adamo, C.; Barone, V. *J. Chem. Phys.* **1999**, *110*, 6158−6170.
(65) Grimme, S. *J. Chem. Phys.* **2006**, *124*, 034108.
(66) Perdew, J. P.; Schmidt, K. *AIP Conf. Proc.* **2001**, *577*, 1−20.
(67) Grimme, S. *Chem.—Eur. J.* **2012**, *18*, 9955−9964.
(68) Wiegand, T.; Eckert, H.; Ekkert, O.; Fröhlich, R.; Kehr, G.; Erker, G.; Grimme, S. *J. Am. Chem. Soc.* **2012**, *134*, 4236−4249.
(69) Strunk, T.; Wolf, M.; Brieg, M.; Klenin, K.; Biewer, A.; Tristram, F.; Ernst, M.; Kleine, P. J.; Heilmann, N.; Kondov, I.; Wenzel, W. *J. Comput. Chem.* **2012**, DOI: 10.1002/jcc.23089.
(70) Grimme, S. *Angew. Chem., Int. Ed.* **2008**, *47*, 3430−3434.
(71) Sinnokrot, M. O.; Sherrill, C. D. *J. Phys. Chem. A* **2004**, *108*, 10200−10207.
(72) Lanzarotti, E.; Biekofsky, R. R.; Estrin, D. A.; Marti, M. A.; Turjanski, A. G. *J. Chem. Inf. Model.* **2011**, *51*, 1623−1633.
(73) Tauer, T. P.; Sherrill, C. D. *J. Phys. Chem. A* **2005**, *109*, 10475−10478.
(74) Rodríguez-Ropero, F.; Casanovas, J.; Alemá, C. *J. Comput. Chem.* **2008**, *29*, 69−78.
(75) Adam, M. S.; Gutmann, M. J.; Leech, C. K.; Middlemiss, D. S.; Parkin, A.; Thomas, L. H.; Wilson, C. C. *New J. Chem.* **2010**, *34*, 85−91.
(76) Coutsias, E. A.; Seok, C.; Dill, H. A. *J. Comput. Chem.* **2004**, *25*, 1849−1857.
(77) Dunning, T. H. *J. Chem. Phys.* **1989**, *90*, 1007−1023.
(78) Helgaker, T.; Klopper, W.; Koch, H.; Noga, J. *J. Chem. Phys.* **1997**, *106*, 9639−9646.