JCTC Journal of Chemical Theory and Computation

On the Structure and Geometry of Biomolecular Binding Motifs (Hydrogen-Bonding, Stacking, X−H···π): WFT and DFT Calculations

Kevin E. Riley,*,† Michal Pitoňák,‡,§ Jiří Černý,|| and Pavel Hobza*,‡,⊥

Department of Chemistry, University of Puerto Rico, P.O. Box 23346, Rio Piedras, Puerto Rico 00931, Institute of Organic Chemistry and Biochemistry, Academy of Sciences of the Czech Republic and Center of Biomolecules and Complex Molecular Systems, Flemingovo nam. 2, 166 10 Prague 6, Czech Republic, Department of Physical and Theoretical Chemistry, Faculty of Natural Sciences, Comenius University, Mlynska Dolina CH-1, 842 15 Bratislava, Slovak Republic, Institute of Biotechnology, Academy of Sciences of the Czech Republic, 142 00 Prague 4, Czech Republic, and Department of Physical Chemistry, Palacký University, Olomouc, 771 46 Olomouc, Czech Republic

Received July 20, 2009

Abstract: The strengths of noncovalent interactions are generally very sensitive to a number of geometric parameters. Among the most important of these parameters is the separation between the interacting moieties (in the case of an intermolecular interaction, this would be the intermolecular separation). Most works seeking to characterize the properties of intermolecular interactions are mainly concerned with binding energies obtained at the potential energy minimum (as determined at some particular level of theory). In this work, in order to extend our understanding of these types of noncovalent interactions, we investigate the distance dependence of several types of intermolecular interactions, these are hydrogen bonds, stacking interactions, dispersion interactions, and X−H···π interactions. There are several methods that have traditionally been used to treat noncovalent interactions as well as many new methods that have emerged within the past three or four years. Here we obtain reference data using estimated CCSD(T) values at the complete basis set limit (using the CBS(T) method); potential energy curves are also produced using several other methods thought to be accurate for intermolecular interactions, these are MP2/cc-pVTZ, MP2/aug-cc-pVDZ, MP2/6-31G*(0.25), SCS(MI)-MP2/cc-pVTZ, estimated MP2.5/CBS, DFT-SAPT/aug-cc-pVTZ, DFT/M06-2X/6-311+G(2df,2p), and DFT-D/TPSS/6-311++G(3df,3pd). The basis set superposition error is systematically considered throughout the study. It is found that the MP2.5 and DFT-SAPT methods, which are both quite computationally intensive, produce potential energy curves that are in very good agreement to those of the reference method. Among the MP2 techniques, which can be said to be of medium computational expense, the best results are obtained with MP2/cc-pVTZ and SCS(MI)-MP2/cc-pVTZ. DFT-D/TPSS/6-311++G(3df,3pd) is the DFT-based method that can be said to give the most well-balanced description of intermolecular interactions.

Introduction

The structure, stability, and dynamic properties of biomolecular systems, such as proteins, DNA/RNA, and protein−ligand complexes, are influenced by several physical factors, the most important of which are solvation effects[1,2] and noncovalent interactions.[3–6] The mode of action of solvation effects in stabilizing biomacromolecules is generally seen

* Corresponding author. E-mail: kev.e.riley@gmail.com (K.E.R.); pavel.hobza@uochb.cas.cz (P.H.).

† Department of Chemistry, University of Puerto Rico.

‡ Institute of Organic Chemistry and Biochemistry, Academy of Sciences of the Czech Republic and Center of Biomolecules and Complex Molecular Systems.

§ Department of Physical and Theoretical Chemistry, Comenius University.

|| Institute of Biotechnology, Academy of Sciences of the Czech Republic.

⊥ Department of Physical Chemistry, Palacký University.

10.1021/ct900376r
Published on Web 12/09/2009

as being nonspecific in character, playing roles, for example, in the aggregation of hydrophobic amino acids in the core of globular proteins.[4,6] The roles that noncovalent interactions play in the structures and stabilities of biomacromolecules can be quite different than those played by solvation effects because of the presence of certain specific binding motifs that commonly occur in proteins and DNA (as well as other biomolecular structures), that lead to very stable interactions. The formation of these strong interactions can have a large impact on structure and, in the case of protein receptors that interact with particular ligands, can determine whether or not the receptor is activated.[7–9] Among the most common of these specific types of interactions are hydrogen-bonds (H-bonds) and stacking and X−H···π interactions (X is usually O, N, S, or C). It should be noted that dispersion, or van der Waals, interactions, which are generally fairly weak, represent a class of noncovalent interaction that is geometrically nonspecific, that is to say that they do not depend heavily on the relative orientation of the monomers, such as in the case of, for example, H-bonds. Although these types of interactions are weak, they are very important in biomolecular structure because of their pervasiveness throughout the structures of proteins, DNA and other biostructures. We will note here that, when we refer to dispersion interactions, we are describing the types of weak interactions, such as those between aliphatic molecules, whose attractive nature is largely attributable to London dispersion forces. In general, all types of noncovalent interactions contain some degree of a dispersion-type component. Likewise, even interactions between aliphatic molecules contain some contribution from electrostatic forces.

Noncovalent interactions are characterized by a very subtle energetic scale (with respect to geometric parameters), a property that is necessary for the fine-tuning and the diversity of biochemical processes.[10] As noted above, there are four classes of noncovalent interactions that play the largest roles in biomolecular structure, these are H-bonding and dispersion, stacking, and X−H···π interactions. We will note here that σ-hole bonding, which has been the subject of many recent investigations, also plays important roles in biology but, because it is fairly specialized and is not as ubiquitous as the other noncovalent bonding classes, this type of interaction will not be discussed here.[11–14] Among the interaction types, H-bonding is the best characterized and is known to be chiefly attributable to electrostatic forces (dipole−dipole interactions).[10,15,16] Dispersion interactions, as indicated by the name, are stabilized principally by London dispersion (part of van der Waals) forces.[10,15,16] Both dispersion and electrostatic forces contribute to the stabilization of stacked and X−H···π structures, with the largest energetic contribution for both these types of interactions coming from dispersion. It should be noted that, because of the enhanced electrostatic landscape of heterocyclic aromatic groups, interactions involving these moieties tend to be more attractive and to have larger electrostatic contributions than those involving phenyl rings. This is especially important when considering the extremely attractive stacking interactions between the nucleobases contained in DNA and RNA.[10,17]

The characterization of noncovalent interactions in biomolecules has been the subject of many experimental and theoretical investigations in (at least) the past two decades.[8–10,18–30] On the computational side, it has been possible for many years to properly characterize H-bonding interactions because these dipole−dipole dependent interactions can be described relatively well using one-particle methods, such as Hartree−Fock (HF) and density functional theory (DFT). Dispersion, stacking, and X−H···π interactions are largely dependent on dispersion forces, which can only be accurately described by (computationally expensive) high-level theoretical methods, such as the coupled cluster theory (CC) method using single, double, and perturbative triple excitations (i.e., CCSD(T)) along with large basis sets (at least aug-cc-pVTZ).[17,31] Because of the prohibitive cost of these types of calculations for all but the smallest complexes, there has been relatively little work done seeking to accurately characterize interactions that are heavily based on London dispersion forces. Over the past 15 years or so there have been many studies describing dispersion, stacking, and X−H···π interactions using the second-order Møller−Plesset perturbation theory method (MP2), a method that can be said to be of intermediate computational cost, with various basis sets.[10,16,32] It has been shown (for several different types of intermolecular interactions) that the results obtained with the MP2 method can be semiquantitative, with accuracies that are highly dependent on the basis sets that are employed.[32,33] Recently it has become possible to compute binding energies for molecular complexes with increasing accuracy by using techniques that take advantage of the fact that the CCSD(T) and MP2 binding energies exhibit very similar basis set behavior.[31,34] That is to say that the difference in binding energy computed using, for example, the aug-cc-pVDZ and aug-cc-pVTZ basis sets is roughly the same for both the MP2 and CCSD(T) methods. This basis set behavior allows one to compute the MP2 binding energy using the largest possible basis set (or extrapolate to the complete basis set limit (CBS)) and then add a CC correction term (ΔCCSD(T)), corresponding to the difference between the CCSD(T) and MP2 binding energies for a given (generally small or medium) basis set. At present, this scheme represents the most accurate technique for the determination of interaction energies for systems that cannot be treated using the CCSD(T) method, along with large basis sets. The use of this type of scheme along with MP2 binding energies that have been extrapolated to the complete basis set has been termed the CBS(T) method. The accuracy of the method was recently confirmed by performing the direct extrapolation of the CCSD(T) energies determined with the aug-cc-pVDZ and aug-cc-pVTZ basis sets.[17,35]

Most investigations concerned with the accurate characterization of noncovalent interactions in biomacromolecules have focused on obtaining accurate binding energies either by using the potential energy minimum (as determined at some lower level of theory) or the experimentally derived complex structures (such as those obtained from X-ray crystal structures). In this study, we investigate the types of noncovalent interactions that are relevant in biomolecular structure, focusing on the potential energy curves of these

interactions along the most important geometrical axis (i.e., directly along the dissociation pathway), meaning that the structure of a biomolecular cluster was optimized with respect to one geometry coordinate. There are several reasons that it is important to characterize not only minimum energy structures but also potential energy curves for these interactions. First, as noted above, noncovalent interactions are very sensitive to geometric parameters, and their strengths can often vary significantly with only a small geometric perturbation. This sensitivity can have a tremendous influence on the structures and the stabilities of proteins and nucleic acid compounds (DNA/RNA) and may be a large factor in determining whether or not a ligand (such as a hormone or pharmaceutical compound) successfully binds to a protein receptor. Formulating a deeper understanding of the behavior of noncovalent interactions as a function of geometric parameters can give us insights into the dynamics of biomolecular systems, giving us information that could be very valuable in the interpretation of vibrational (infrared) spectra of peptides, proteins, and nucleic acid compounds. Second, studying the potential energy curves for a variety of noncovalent interaction types can aid in determining the accuracy, in terms of converging to the geometric energy minimum, that can be expected of lower level (less computationally expensive) methods. This last point is very important because the structures obtained at these lower levels are often used for high-level binding energy analyses (as noted above) and because lower level theory is often used to obtain theoretical infrared spectra, which can potentially be useful in assigning peaks in experimentally obtained spectra. Finally, and this point is particularly significant for complex molecular systems, interactions at long ranges play a key role in complexes of extended systems, where the number of contacts at these distances grows extremely quickly.

It will be noted here that there are many degrees of freedom that must be considered in studying geometrical relationships in noncovalently bound systems. The goal of this work is to study potential energy curves along the dissociation pathways of several complexes, this is the coordinate that is generally considered to be the most important in terms of complex formation and dissociation. Further studies are underway in our laboratories to investigate the full geometrical dependence of noncovalent interactions on structures that have been fully gradient optimized at very high levels of theory (including estimated CCSD(T)/CBS).[36]

As noted above, the MP2 method has long been the method of choice for the computation of intermolecular interactions, producing binding energies that are generally semiquantitatively accurate at a reasonable computational cost. It has been shown that for the S26 test set of complexes, which contains H-bonded, dispersion-bound, and mixed (contributions from both electrostatics and dispersion) interactions, the MP2 method yields the best results when it is paired with the medium-sized cc-pVTZ and aug-cc-pVDZ basis sets (the S26 test set is related to the S22 test set described below).[32] The use of larger basis sets usually results in overestimation of binding energies, with electronic energies for complexes being too high relative to those of the monomers. Generally the MP2 method treats H-bonding interactions fairly well but often greatly underestimates the binding energies of cyclic H-bonds, such as those found in nucleic acid base pairs. In terms of dispersion and stacking interactions, the MP2 method generally tends to (sometimes strongly) overestimate binding energies for these types of complexes, this is especially true when larger basis sets are used. It should be stressed that much of the success of the MP2 method can be attributed to error compensation effects stemming from the relative energies of a complex and from its constituent monomers. As a result of this, MP2 binding energies for intermolecular interactions do not generally converge to the correct value (as determined with CCSD(T)) with increasing basis set size. For example, the aug-cc-pVDZ basis set has been observed to obtain a more balanced description of binding energies for the S26 set than that of the aug-cc-pVTZ basis.

The use of smaller basis sets, such as those of the Pople-type 6-31G* family, along with the MP2 method allow for the treatment of larger systems and have been used with some frequency in past years when using larger bases was not possible. In some cases, these types of bases have been shown to yield very good binding energies. One example of a small basis set that has been extensively used for the treatment of noncovalent interactions is 6-31G*(0.25), which is a modified 6-31G* basis set for which the polarization functions have been modified to be more diffuse (change in exponential parameter from 0.80 to 0.25).[37] This basis has been shown to give reasonable results for binding energies of molecular complexes and has performed especially well for stacking interactions.[32] The surprisingly good agreement of MP2/6-31G*(0.25) and CCSD(T)/CBS binding energies for stacked systems has recently been shown.[38]

The past several years have seen the development of many new computational techniques that promise to provide well balanced and accurate descriptions of a wide variety of different types of noncovalent interactions at much lower computational costs than the CCSD(T), or even the CBS(T), method. A great number of these methods have been parametrized and/or tested using S22,[33] S26,[32] and JSCH2005[33] benchmark data sets; all complexes presented there are systematically given in their (estimated) global energy minima. A similar situation also exists for other noncovalent databases. It is, thus, highly desirable to test the performance of these methods not only for the stabilization energy but also for the geometry.

It is well know that one-particle methods, such as HF and DFT, generally fail to describe interactions that are strongly dependent on dispersion forces,[39] however, recently several DFT techniques seeking to take dispersion interaction contributions into account have been developed; here we will discuss two of these methods, DFT-D[40,41] and M06-2X.[22,42,43]

The DFT-D method deals with dispersion by using an empirical term describing the London dispersion energy. The DFT-D empirical dispersion term has been parametrized against the S22 binding energy test set, which includes H-bonded, dispersion-bound, and mixed (electrostatic and dispersion) complexes. The M06-2X functional is based on the reparameterization of the DFT

functional in order to take dispersion effects into account; the parametrization was made on various data sets including a set of small noncovalent complexes. The M06-2X functional is a member of the M06 family of functionals, which, along with several other functionals (described at http://comp.chem.umn.edu/info/DFT.htm), represent an extensive effort by Truhlar and co-workers to develop density functionals with improved reliability for the computation of many molecular properties.[22,23,42–47] The performance of the M06-2X functional (as well as other functionals from the M06 family) was tested using the S22 data set.[43] In a recent assessment, Sherrill and co-workers note that the M05−2X and M06-2X descriptions of variously configured nucleic acids from the JSCH-2005 test set are not as well-balanced as that of the DFT-D/PBE/aug-cc-pVDZ method by Grimme.[48–50]

The DFT-symmetry adapted perturbation theory method (DFT-SAPT)[51–56] is the only method considered in this work treating molecular interaction differently than by the supermolecular approach. This technique has been shown to compute accurate binding energies for a variety of interaction types and has the great advantage of determining the total intermolecular interactions as a sum of physically meaningful components, such as electrostatic, exchange, induction, and dispersion terms. The method provides very good estimates of stabilization energies close to the CCSD(T) benchmark data. A very important advantage of the procedure is the fact that it is almost a genuine ab initio procedure, i.e., it does not contain any empirical parameter, except for those in underlying DFT functional, e.g., in the DFT-SAPT procedure.

The overestimation of the stabilization energy in dispersion-dominated complexes by MP2 was shown to be due to the fact that the supermolecular MP2 interaction energy includes the dispersion energy determined only at the uncoupled HF level. Dispersion energies are generally overestimated by 10−20% in comparison with accurate values.[57] In the past few years, several methods have been developed with the aim of improving the performance of MP2, in terms of their abilities, to accurately describe intermolecular interactions in a well balanced way (across all interaction types).[57,58]

The basis for the spin-component scaled MP2 method (SCS-MP2) is the parametrization of the parallel and antiparallel spin components of the MP2 correlation energy.[59] The parameters for the family of SCS-MP2 methods have been deduced from either theory or fitted against many test sets describing several atomic and molecular properties. In this work, we will only be concerned with the molecular interactions (SCS(MI)-MP2) variant of the method,[60] though there are several other variants that may produce good potential energy curves for intermolecular interactions (for example, SCSN-MP2).[61,62] This method, like DFT-D, has been parametrized against the S22 molecular interactions test set. The SCS(MI)-MP2 method has been shown to reduce the systematic overestimation of binding energies for dispersion-bound complexes seen with the MP2 technique and, thus, should be suitable for the description of a wide variety of molecular interaction motifs. The SCS(MI)-MP2 method provides very good stabilization energies for stacked as well as H-bonded complexes, in contrast to the original SCS-MP2 method, which fails for the latter complexes.[17] All methods of the SCS-MP2 family contain empirical parameter(s). Sherrill and co-workers have recently carried out studies in which various SCS-MP2 methods (as well as DFT-based methods) are compared in terms of their ability to accurately produce potential energy curves for molecular complexes containing benzene as (at least) one of the monomers and the methane dimer .[61,63] One of the main conclusions of these studies is that SCS-MP2 methods, and particularly SCS(MI)-MP2, give reasonable potential energy curves for the systems considered, although binding energies for the methane dimer are strongly underestimated.

Recently an interesting property of the interaction energy calculated at the supermolecular MP3 level was recognized.[64] Tests carried out on the S22 as well as the JCSH2005 test sets revealed that MP3 underestimates stacking interactions roughly to the same extent as the MP2 overestimates them.[64] At the same time MP3 typically slightly increases the accuracy of the interaction energies of the H-bonded complexes. This was the basis for formulating the MP2.5 (or in general SMP3, Scaled MP3) method, i.e. the MP2 corrected by scaled $E^{(3)}$ (third-order correlation contribution). In the case of MP2.5, the scaling factor is 0.5, while in SMP3, the optimal scaling factor typically ranges from 0.45 to 0.65, depending on the type of molecular complex and the basis set applied. MP2.5 in general reproduces the CCSD(T) values very well (outperforms SCS(MI)-MP2 and all DFT methods mentioned above), but the scaling factor 0.5 is known not to be optimal for all kinds of molecular complexes and cannot be determined a priori, which could lead to errors of about ±10% of $E^{(3)}$. Fortunately (as shown further), SMP3, with a particular choice of the scaling factor, reproduces the CCSD(T) potential energy curves with almost a constant error along a wide range of geometry displacements. However, one main drawback of the method is in its N^6 scaling with system size, which means an order of magnitude slowdown compared to MP2 but a dramatic speedup compared to CCSD(T). The other advantage of the method is that it contains only one empirical parameter, the scaling factor.

There have been a number of studies carried out within the past several years in which high-quality potential energy curves for intermolecular interaction are produced.[17,31,65–81] In a recent work, Pitoňák et al. described both the (cyclic) H-bonding and stacking potential energy curves for the uracil dimer, the smallest nucleic acid complex, at various levels of theory, including the estimated CCSD(T)/aug-cc-pVTZ level.[17] One of the main findings made in this study is that the DFT-D, M06−2X, and SCS(MI)-MP2 methods produce potential energy curves for these interactions that are at least semiquantitatively accurate. The SCS(MI)-MP2 technique yielded particularly accurate results for both H-bonded and stacked systems, while the results obtained with the DFT-D and M06-2X methods were substantially better for the H-bonded complex than for the stacked one. It should be noted that Sherrill and co-workers have produced a number of high-quality potential energy curves for several interesting

intermolecular interactions,[31,63,65–70] among these are various configurations of the (substituted and unsubstituted) benzene dimer[31,67–69] and the H_2S−benzene[70] and methane−benzene complexes.[66] Extremely high-quality geometries and energies for the benzene dimer in various configurations have also been computed by Pulay and Janowski.[71] The geometries and interaction energies of stacked and H-bonded uracil dimers and stacked adenine−thymine pairs were studied by means of high-level quantum chemical calculations including CCSD(T) by Dabkowska et al.[72] It was found that geometry optimization with extended basis sets at the MP2 level underestimates the intermolecular distances compared to the reference CCSD(T) results, whereas the MP2/counterpoise-corrected gradient optimization agrees well with the reference geometries; therefore, this level (MP2/cc-pVTZ) was recommended for geometry optimizations. In a recent study Sponer and co-workers produced potential energy curves near the potential energy minima for several configurations of the uracil dimer using several electronic structure methods (including CBS(T)) and using an empirical potential-based method.[82] For these complexes, it was observed that the DFT-D, DFT-SAPT, and SCS(MI)-MP2 methods all generated curves that were in very good agreement with reference data. Tekin and Jansen produced high-quality, CCSD(T) and DFT-SAPT (both with aug-cc-pVTZ), potential energy curves for various configurations of the acetylene−benzene complex.[83] Tsuzuki and co-workers have produced high-quality CCSD(T) binding energies for a number of alkane dimers, including the propane dimer considered in this work, and have also generated MP2 potential energy curves for a number of conformations of the propane dimer.[76,84] Very recently Fusti Molnar et al. produced high-level estimated CCSD(T) potential energy curves for 20 of the 22 structures found within the S22 molecular interactions test set.[85]

One of the main goals of this article is to compute accurate potential energy curves for the most important classes of noncovalent interaction motifs relevant to biomolecular structure, in order to elucidate the properties of these types of interactions. To this end we have selected seven model systems representing the four major interaction categories to be studied here, these are: cytosine−benzene (stacked), adenine−benzene (stacked), and water−benzene (X−H$\cdots\pi$) and propane (dispersion), methanol (H-bond), methylamine (H-bond), and formamide (H-bond, cyclic) dimers. Potential energy curves for each of these complexes have been computed at the estimated CCSD(T)/CBS level of theory, the highest level currently possible for the largest of these systems. Another principal aim of this work is to compare the performance of several lower-level methods in reproducing the potential energy curves of these complexes. The methods considered here include the MP2, which has long been used for the computation of binding energies of intermolecular interactions, and the relatively new SCS(MI)-MP2, DFT-SAPT, DFT-D, and DFT/M06-2X techniques. More specifically, the method/basis combinations that will be treated in this work are: MP2/cc-pVTZ, MP2/aug-cc-pVDZ, MP2/6-31G*(0.25), SCS(MI)-MP2/cc-pVTZ, DFT-SAPT/aug-cc-pVTZ, DFT-D/TPSS/6-311++G(3df,3pd), and DFT/M06-2X/6-311+G(2df,2p). It should be noted that some of these methods, for example SCS(MI)-MP2, may yield better results when they are used along with larger basis sets. Our main purpose here is to evaluate the performance of several methods that could be used (and have been used) to treat relatively large systems relevant to biochemistry, as such we have chosen to use medium-sized basis sets for all of these methods.

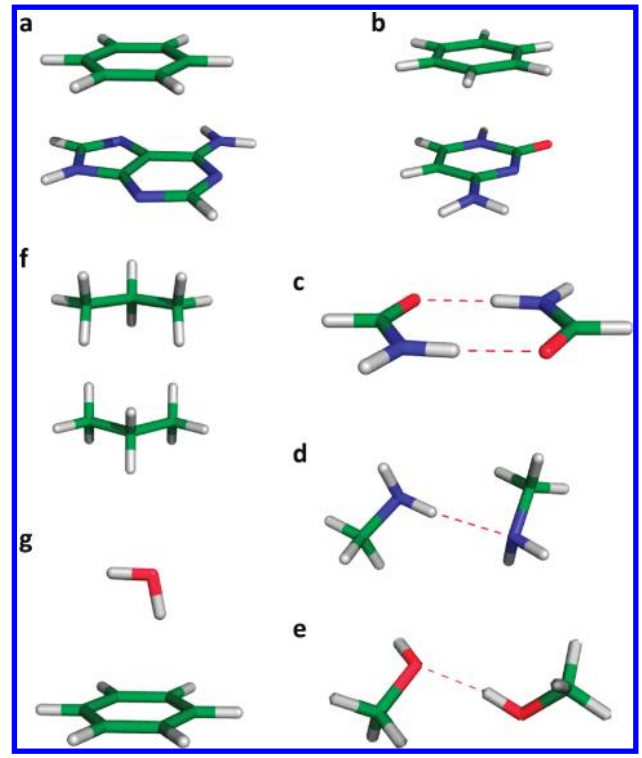

Figure 1. Molecular complexes considered in this work: (a) adenine−benzene, (b) cytosine−benzene, (c) formamide dimer, (d) methylamine dimer, (e) methanol dimer, (f) propane dimer, and (g) benzene−water.

Computational Methods

Structures of Studied Complexes. In order to investigate the noncovalent interactions of varying character, ranging from strongly electrostatic to strongly dispersive, we have included examples of four different interaction types into our study, these are:

i. Stacking interaction: adenine−benzene and cytosine−benzene.
ii. H-bonding interaction: methanol, methylamine, formamide dimers.
iii. Dispersion interaction: propane dimer.
iv. X−H$\cdots\pi$ interaction: benzene−water.

Structures of all complexes investigated are visualized in Figure 1. Initial geometries for the stacking systems were prepared by positioning the benzene ring in an ideal stacking position (i.e., perfectly flat) with its center directly above the center of either cytosine or adenine. The center positions of benzene and cytosine were determined as the average position of all atoms within the ring; in the case of adenine, the center of each ring was determined, and the overall molecular center was taken to be the position in the middle of these two points. The geometries of these monomers were

determined at the B3LYP/6-31+G* level of theory. In order to generate the points for the potential energy curves of these systems, the monomers were simply separated in such a way that they remained parallel to one another.

The initial geometries of the methanol and methylamine complexes were determined at the estimated CCSD(T)/CBS level of theory, while the initial geometry of the formamide complex was taken directly from the S22 data set and computed at the MP2/cc-pVTZ level, using the counterpoise correction (CP) to account for the basis set superposition error (BSSE). Potential energy points for these systems were produced by modifying the H···O or H···N distances, such that the O−H···O or N−H···N angles remained constant. It should be noted that in the case of the formamide dimer, which contains a cyclic double H-bond, the H···N distances were modulated such that both H-bonds were consistently of the same length.

The initial structure for the propane dimer was obtained at the (CP-corrected) MP2/cc-pVTZ level of theory. Here, potential energy points were generated by modifying the distances between the monomers, such that the molecular planes defined by the three carbon atoms in each of the monomers were always parallel to one another and the molecules' centers of mass formed a line perpendicular to the two molecular planes.

In the case of the benzene−water complex, the initial geometry, as determined at the (CP-corrected) MP2/cc-pVTZ level of theory, was taken from the S22 data set. Here, points along the potential energy curve were produced by modulating the distance between the water and benzene monomers in a direction perpendicular to the plane defined by the benzene ring.

Electronic Structure Methods. High-level reference data for each of these curves were obtained using the CBS(T) method to estimate CCSD(T)/CBS results. These values are obtained by first computing the binding energies at the MP2/CBS level and then by adding a ΔCCSD(T) correction term:[10,34]

$$\Delta E^{\mathrm{CCSD(T)}} = \Delta E^{\mathrm{MP2}}_{\mathrm{CBS}} + (\Delta E^{\mathrm{CCSD(T)}} - \Delta E^{\mathrm{MP2}})_{\text{small bassis set}} \quad (1)$$

Here, $E^{\mathrm{MP2}}_{\mathrm{CBS}} = E^{\mathrm{HF}}_{\mathrm{CBS}} + E^{\mathrm{corr,\ MP\,2}}_{\mathrm{CBS}}$ quantities were determined by extrapolating MP2 electronic energies to the complete basis set limit, using the extrapolation scheme of Helgaker and co-workers.[86] All MP2/CBS extrapolations were performed using electronic energies produced at the MP2/aug-cc-pVDZ and MP2/aug-cc-pVTZ levels of theory. As noted above, all ΔCCSD(T) correction terms were computed using the relatively large aug-cc-pVDZ basis set.

The success of this method in providing accurate binding energies is based on the fact that the CCSD(T) and MP2 methods exhibit basis set dependence that is very similar. Although the binding energies obtained using these two methods (with a particular basis set) can be quite different, the difference between these binding energies $(\Delta E^{\mathrm{CCSD(T)}} - \Delta E^{\mathrm{MP2}})_{\text{small bassis set}}$ remains nearly constant regardless of the basis set used. It should be noted that binding energy results generally improve when larger basis sets are used to compute the ΔCCSD(T) term.

In a recent study, Pitoňák et al. produced very accurate interaction energies for the stacked and H-bonding uracil dimers by directly extrapolating CCSD(T) results to the complete basis set limit (using the aug-cc-pVDZ and aug-cc-pVTZ basis sets).[17] Here it was shown that the CBS(T) method, as implemented in this study (i.e., based on eq 1), produces accurate binding energies for these complexes, with errors no larger than 0.1 kcal/mol for the stacked structure and 0.2 kcal/mol for the H-bonding structure (both corresponding to errors of roughly 1.0%).

It should also be noted that the MP2 extrapolation to the complete basis set limit based on the aug-cc-pVDZ and aug-cc-pVTZ basis sets may cause some small binding energy errors, as these are the smallest basis sets for which this type of extrapolation can be carried out. Based on the H-bonding and stacking structures of the uracil dimer (as described above), it is seen that this extrapolation scheme underestimates the MP2/CBS binding energy of both structures compared to the aug-cc-pVTZ to aug-cc-pVQZ extrapolation by about 1%.[17] Jurečka et al. investigated the binding energies of four H-bonding and nine stacked nucleic acid pairs using MP2 extrapolation based both on aug-cc-pvDZ and aug-cc-pVTZ and on aug-cc-pVTZ and aug-cc-pVQZ.[33] For the H-bonding pairs, the smaller basis extrapolation scheme underestimated binding energies, with an average error of 1.8% (compared to the larger basis MP2 extrapolation), the largest error for these interactions was 2.0%. For the stacked structures, the average error of the smaller basis extrapolation was 1.4% (underbinding), with a maximum error of 2.6%. It should also be noted that, for one of the stacking pairs, the smaller basis extrapolation overestimated the binding energy by 1.2%. We expect that the errors present in our studies would not exceed those observed for the nucleic acid complexes described above. Fusti Molnar et al. have also noted the relatively small errors associated with MP2 extrapolation based on aug-cc-pVDZ and aug-cc-pVTZ.[85]

Here all CCSD(T)/aug-cc-pVDZ and MP2/aug-cc-pVDZ calculations were performed using the Molpro electronic structure package (Molpro 2006),[87] while MP2/aug-cc-pVTZ values were computed using the resolution-of-the-identity approach (RI-MP2)[88] with the Turbomole package.[89] The RI-MP2 approximation has been shown to introduce negligible errors.[90] MP2 and CCSD(T) computations were carried out using the frozen-core approximation in which the lowest lying molecular orbitals are constrained to remain doubly occupied in all configurations.

Here, we study the performance of MP2 along with the aug-cc-pVDZ, cc-pVTZ, and 6-31G*(0.25) basis sets for the description of the potential energy curves associated with the separation of molecular complexes. All MP2 curves were produced using the Molpro electronic structure package and incorporate the counterpoise correction in order to account for the BSSE.[91]

The DFT-D technique incorporates an empirical London dispersion energy term, which describes dispersion using the well-known C_6/R^6 formula.[41] The dispersion energy term also contains a damping function to account for overlap effects (most importantly at small values of R).

DFT-D parameters were fitted against the S22 data set of intermolecular interactions. In this work, we employ the DFT-D/TPSS/6-311++G(3df,3pd) technique using empirical coefficients optimized for noncounterpoise-corrected binding energy computations. DFT energies were computed using the Gaussian package (G03),[92] while the dispersion terms were obtained using an in-house Fortran program. For DFT calculations, fine grids and tight convergence were utilized for all calculations (i.e., INT(GRID=ULTRAFINE), SCF(CONVER=TIGHT)).

The M06-2X functional of Truhlar and co-workers (along with several other functional, M05, M05-2X, M06, M06-L, etc.) was developed to give improved results for several molecular properties.[22,42,43,45,47] One of the large goals achieved with many of these functionals was a much-improved description of dispersion forces. Here, we have employed the M06-2X functional along with the 6-311+G(2df,2p), using the Gaussian electronic structure package.[92] The counterpoise technique was employed to account for BSSE.

Both DFT-based methods considered here have been parametrized to be used along with specific basis sets and, in the case of DFT-D, functionals. Here we have used the basis sets (and functionals) recommended by the developers of these two methods in order to give the best results at reasonable computational costs.

The SCS(MI)-MP2 method obtains improved results for molecular interactions by scaling the MP2 parallel and antiparallel contributions to the correlation energy.[60] The main result of the spin parametrization in the SCS(MI)-MP2 method is the reduction of the overstabilization of dispersion interactions seen with MP2. The SCS(MI)-MP2 parameters were optimized against the S22 data set of molecular interactions. Here, SCS(MI)-MP2 calculations were performed along with the cc-pVTZ basis set, using Molpro. The counterpoise correction for BSSE was included.

DFT-SAPT uses monomer properties and electronic densities from DFT in order to compute interaction energies using the symmetry adapted perturbation theory (SAPT).[51–56] This is the only variant of the SAPT methods that can be practically used for systems containing more than a few atoms and is, thus, the most useful for computations on biomolecular systems. DFT-SAPT has been shown in several studies to obtain accurate binding energies for a wide variety of intermolecular interaction types. This method determines the total interaction energy as a sum of physically meaningful components, such as those arising from electrostatics, dispersion, induction, and exchange. The DFT-SAPT interaction energy is given as the sum of these components:

$$E_{\mathrm{int}} = E^1_{\mathrm{pol}} + E^1_{\mathrm{ex}} + E^2_{\mathrm{ind}} + E^2_{\mathrm{ex-ind}} + E^2_{\mathrm{disp}} + E^2_{\mathrm{ex-disp}} + \delta\mathrm{HF} \quad (2)$$

Some of these terms can be combined in order to define values that correspond to commonly understood physical quantities. The terms are commonly combined as such:

$$\begin{aligned} E(\mathrm{elec}) &= E^1_{\mathrm{pol}} \\ E(\mathrm{ind}) &= E^2_{\mathrm{ind}} + E^2_{\mathrm{ex-ind}} \\ E(\mathrm{disp}) &= E^2_{\mathrm{disp}} + E^2_{\mathrm{ex-disp}} \end{aligned}$$

and

$$E(\mathrm{exch}) = E^1_{\mathrm{ex}}$$

These four quantities refer to the electrostatic (elec), induction (ind), dispersion (disp), and exchange−repulsion (exch) contributions, respectively, to the total interaction energy. The δHF term is an estimate of higher-order Hartree−Fock contributions and is determined as the difference between the HF interaction energy and the sum of all the first- and second-order contributions (obtained with the HF wave functions), with the exceptions of the dispersion and exchange−dispersion terms. Since the HF interaction energy is determined using a supermolecular description (including counterpoise corrections), the DFT-SAPT interaction energy constructed, as in eq 2, is not BSSE free. It is, however, true that the BSSE for the HF interaction energy is much smaller than that of the correlation interaction energy.

All DFT-SAPT computations have been carried out using the LPBE0AC potential along with the aug-cc-pVTZ basis set. This basis set can generally be viewed as the smallest basis that gives meaningful results with SAPT methods; the use of smaller basis sets will result in significant underestimation of the dispersion term and, thus, the binding energy. In a study of the binding in several configurations of the acetylene−benzene complex by Tekin and Jansen, it was found that DFT-SAPT/aug-cc-pVTZ produces binding energies that are up to ~5% lower than those of DFT-SAPT/CBS.[83]

The density fitting procedure was used to significantly reduce the computational cost of these calculations. It is necessary to compute a shift term involving the ionization potentials and the highest occupied molecular orbital (HOMO) energies for interacting monomers. These terms were obtained using the PBE0 functional along with the aug-cc-pVDZ basis set. DFT-SAPT calculations were performed using the Molpro package of programs. Here, it should be noted that DFT-SAPT/aug-cc-pVTZ computations on the adenine−benzene complex were not possible to obtain because of technical (convergence) difficulties.

MP3 (and thus MP2.5) calculations were performed using the L-CCD (linearized coupled clusters singles and doubles) module based on the Cholesky decomposed two-electron integrals implemented in the MOLCAS 7 program package,[93] where a 1.10^{-7} threshold for integral decomposition was used. Overall estimated MP2.5/CBS results were obtained analogously to eq 1:

$$\Delta E^{\mathrm{MP2.5}}_{\mathrm{CBS}} = \Delta E^{\mathrm{MP2}}_{\mathrm{CBS}} + \frac{1}{2}(\Delta E^{\mathrm{MP3}} - \Delta E^{\mathrm{MP2}})_{\mathrm{small\ basis\ set}} \quad (3)$$

assuming that the $E^{(3)}$ term ($\Delta E^{\mathrm{MP3}} - \Delta E^{\mathrm{MP2}}$) converges, analogously to higher-order correction terms from CCSD(T), faster with the basis set than with the MP2. Just to illustrate the speedup of the MP3 calculation compared to the CCSD(T), a MP3 step of the single-point calculation of the cytosine−benzene complex using the aug-cc-pVDZ basis set

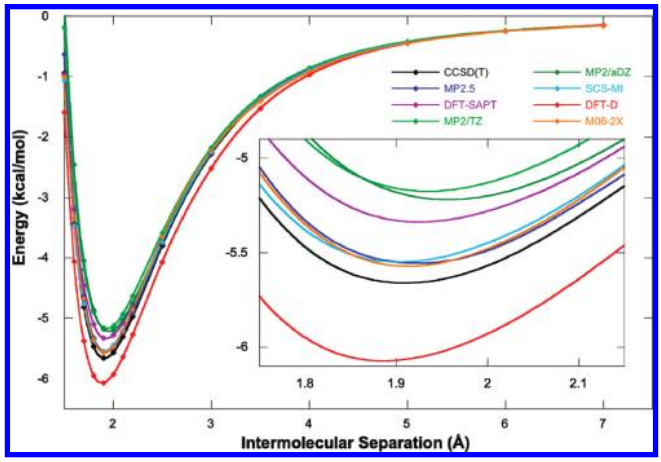

Figure 2. Potential energy curves for the methanol dimer using several electronic structure methods (see text for exact description of methods used). CCSD(T) (black), MP2.5 (dark blue), DFT-SAPT (purple), MP2/TZ (light green), MP2/aDZ (dark green), SCS-MI (light blue), DFT-D (red), M06-2X (orange).

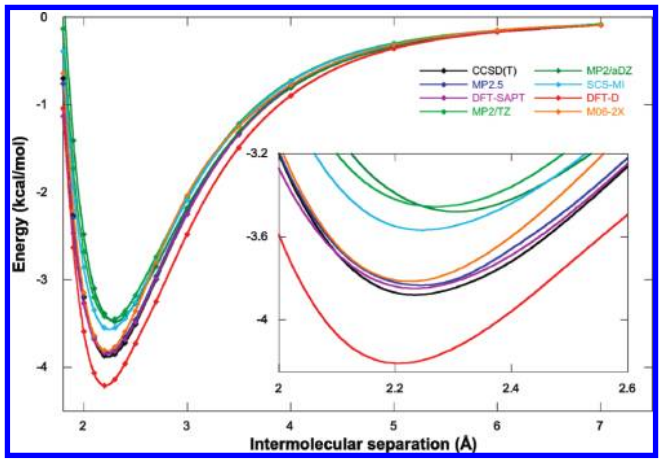

Figure 3. Potential energy curves for the methylamine dimer using several electronic structure methods (see text for exact description of methods used).

took 45 min on four cores of a 2.4Ghz Intel Core2 Quad processor machine, while the CCSD(T) step took 39 h on four 2.66 GHz four-core Intel Xeon E5430 processors. It should be noted, that the MP2 method is still roughly by an order of magnitude faster than MP3.

Results

Performance of MP2-, DFT-, and SAPT-Based Methods. *H-bonding Interactions.* Inspection of Figures 2−4 reveals that all of the computational methods studied here give similar interaction energy curves for the methanol (Figure 2), methylamine (Figure 3), and formamide (Figure 4) dimers, which are the three H-bonding complexes considered here. As would be expected for H-bonds whose properties are generally well described even by low-level methods, all of the computational techniques treated in this work give results that are, at least, semiquantitatively correct. It should also be noted that all of the methods considered here exhibit proper long-range behavior.

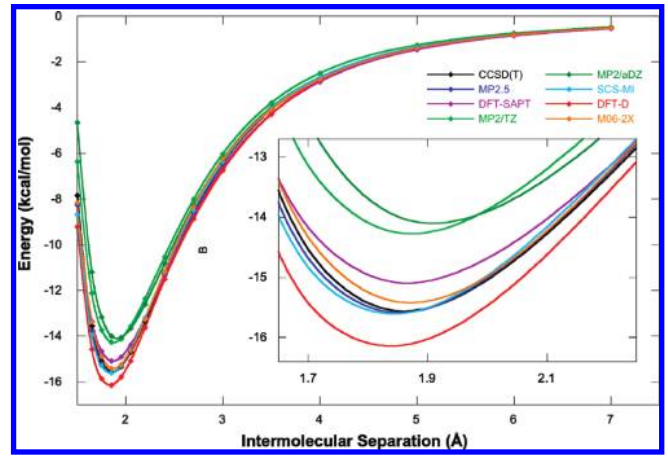

Figure 4. Potential energy curves for the formamide dimer using several electronic structure methods (see text for exact description of methods used).

The estimated MP2.5/CBS method produces extremely high-quality potential energy curves for these H-bonding systems, that are generally slightly underbound. The potential energy minima for all three complexes are at the same intermolecular separation as those of CCSD(T), with binding energies that are in error by no more than 2%. In general, MP2/aug-cc-pVDZ and MP2/cc-pVTZ tend to underbind these complexes by about 8−12%, with potential energy minima at slightly too large an intermolecular separation (by no more than about 0.01 Å, see Table 1). Still considering MP2, the minimum energy separation is slightly larger for the aug-cc-pVDZ basis set than for cc-pVTZ for all H-bonding systems. This finding agrees well with our previous conclusion mentioned above (Dabkowska et al.),[72] showing that counterpoise-corrected MP2/cc-pVTZ optimization yields reliable geometries well comparable with the CBS(T) ones. The SCS(MI)-MP2 method generally produces accurate results for the H-bonding complexes, matching the CBS(T) curves for the formamide and methanol dimers extremely closely. SCS(MI)-MP2 results for the methylamine dimer are not quite as accurate, with binding energies that are approximately 8% too high near the potential energy minimum. It should be noted that the curve produced with this method, although too shallow, is still in good agreement with CCSD(T) in terms of the location of the potential energy minimum. The fact that all calculations based on the MP2 procedure yield reliable distances of the minima is promising, since it allows one to optimize the structure of H-bonded complexes at this (rather cheap) level and then to perform a single-point calculation with some higher-level method providing accurate energies. DFT-SAPT produces very accurate potential energy curves for these H-bonding systems, with minima located at the same locations as those of CCSD(T) (to within 0.1 Å) and with binding energies that are slightly too high (underbound) for all complexes. The largest error in the binding energy at the potential energy minimum occurs for the methanol dimer, which is underbound by about 6%.

For each of the H-bonding complexes, DFT-D yields binding energies that are too large by about 4−9%. It should, however, be noted that this method predicts the proper

Table 1. Equilibrium Separation Distances and Binding Energies (in Parentheses) for Complexes/Methods Considered in This Work[a]

method	methanol	methylamine	formamide	adenine−benzene	cytosine−benzene	propane	benzene−water
CCSD(T)	1.91	2.23	1.86	3.40	3.56	3.81	2.41
	(5.66)	(3.88)	(15.56)	(5.28)	(3.97)	(2.04)	(3.26)
MP2.5	1.92	2.24	1.85	3.38	3.54	3.81	2.41
	(5.64)	(3.83)	(15.59)	(5.58)	(4.16)	(1.92)	(3.30)
DFT-SAPT	1.93	2.23	1.87	n/a	3.54	3.85	2.42
	(5.34)	(3.85)	(15.09)		(3.89)	(1.91)	(3.23)
MP2/TZ	1.94	2.27	1.87	3.30	3.51	3.87	2.47
	(5.17)	(3.46)	(14.27)	(6.84)	(4.33)	(1.64)	(2.93)
MP2/aDZ	1.96	2.30	1.91	3.30	3.49	3.89	2.50
	(5.22)	(3.48)	(15.00)	(7.69)	(5.19)	(1.72)	(2.98)
MP2/6-31	2.02	2.37	1.97	3.37	3.57	4.08	2.54
	(4.75)	(3.01)	(13.16)	(5.74)	(3.64)	(1.02)	(2.64)
SCS/TZ	1.90	2.25	1.84	3.36	3.58	3.90	2.47
	(5.66)	(3.57)	(15.62)	(5.87)	(3.79)	(1.49)	(3.06)
DFT-D	1.89	2.21	1.84	3.37	3.52	3.78	2.45
	(5.72)	(3.90)	(15.95)	(5.40)	(4.15)	(2.87)	(3.61)
M06-2X	1.91	2.23	1.87	3.31	3.46	3.65	2.36
	(5.57)	(3.81)	(15.43)	(5.10)	(3.31)	(2.09)	(3.58)

[a] CCSD(T) = estimated CCSD(T)/CBS, TZ = cc-pVTZ, aDZ = aug-cc-pVDZ, 6-31 = 6-31G*(0.25). MP2.5 refers to estimated MP2.5/CBS, DFT-SAPT refers to DFT-SAPT/aug-cc-pVTZ, DFT-D refers to DFT-D/TPSS/6-311++G(3df,3pd), and M06-2X refers to DFT/M06-2X/6-311+G(2df,2p). (See text for detailed descriptions of the complexes and the methods used; distances in Å and binding energies in kcal/mol.)

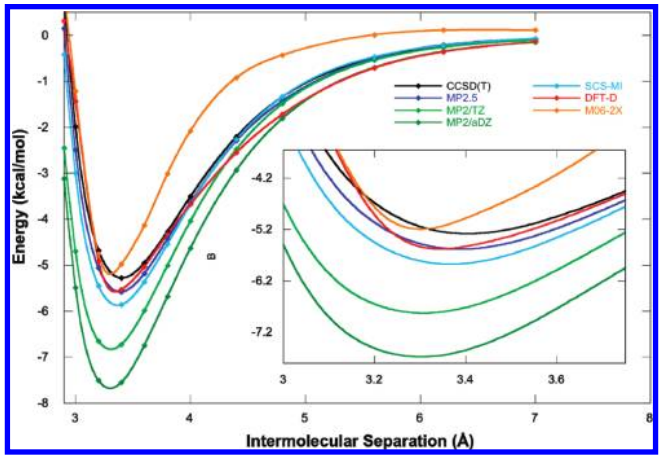

Figure 5. Potential energy curves for adenine−benzene using several electronic structure methods (see text for exact description of methods used).

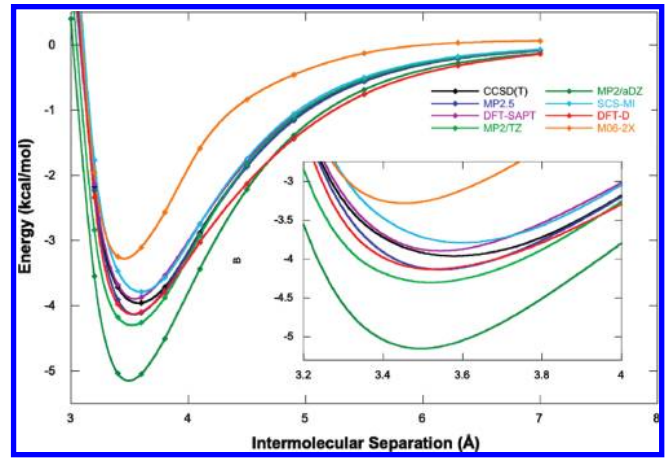

Figure 6. Potential energy curves for cytosine−benzene using several electronic structure methods (see text for exact description of methods used).

locations of the potential energy minima for all three complexes (to within 0.02 Å, Table 1). The M06-2X functional gives excellent results for each of the H-bonding complexes, producing binding energies that are only very slightly too small. The potential energy curves given by this functional track the estimated CCSD(T) curves almost perfectly for both the formamide and methanol dimers and produce very good results for the methylamine dimer. In the case of the methylamine dimer, the M06-2X interaction energies, at points in the region of ascent from the potential energy minimum (moving radially outward), are too high (underbound) by about 0.15−0.20 kcal/mol. The fact that both DFT methods (and specifically the M06-2X) yield proper long-range behavior is not surprising, since the long-range contribution originates here in the dipole−dipole interaction, which is well described by the DFT functionals considered.

Stacking Interactions. Figures 5 and 6 give the potential energy curves for the interactions of adenine and cytosine with benzene, respectively. In these figures it can be seen that the estimated MP2.5/CBS method yields very good potential energy curves for these complexes, with binding energies at the potential energy minimum that are overbound by approximately 4−6%. The fact that these binding energies are too low indicates that the magnitudes of the scaled $E^{(3)}$ corrections to the MP2/CBS results are smaller than those of the CCSD(T) corrections. One of the most prominent aspects of the data presented in these figures is the dramatic overbinding exhibited by the MP2 method when paired with both the cc-pVTZ and aug-cc-pVDZ basis sets. Overbinding is particularly strong for MP2/aug-cc-pVDZ, with minimum energy binding energies that are off approximately by 42% for the adenine−benzene complex and by 27% for the cytosine−benzene complex. This result is not surprising, as it is well documented that the MP2 method, along with medium basis sets, generally tends to overbind for stacking interactions. It should also be noted that the MP2 method, with both basis sets, predicts the optimum intermolecular

separations for these complexes to be slightly too small compared to those of estimated CCSD(T)/CBS results, but the difference is fairly small. As with H-bonded complexes, the counterpoise-corrected MP2/cc-pVTZ geometry is closer to the reference data than that of the MP2/aug-cc-pVDZ. In the case of cytosine−benzene, the MP2/cc-pVTZ method overbinds the minimum geometry by only about 5−10%. One of the goals in developing the SCS(MI)-MP2 method was to correct for the large overbinding effects seen with the MP2 method for stacked systems. In the figures, it can be seen that SCS(MI)-MP2 produces potential energy curves that are generally better than those given by MP2, with minimum energy binding energies that are within 10% of our reference values. Interestingly, this method overbinds by about 10% for the adenine−benzene complex, while underbinding by about 5% for the cytosine−benzene complex. For both complexes, the potential energy minima are located at approximately the same points for both SCS(MI)-MP2 and CCSD(T). DFT-SAPT results are available only for cytosine−benzene. For this complex, DFT-SAPT yields a very accurate potential energy curve, with a binding energy that is too high at the minimum by about 2%. One interesting aspect of the DFT-SAPT data depicted here is that the curve is slightly too shallow moving out (increasing the intermolecular separation) from the potential energy minimum. This shallowness is observed from the minimum (∼3.6 Å) out to about 4.2−4.4 Å.

The DFT-D technique obtains very good results for both of these stacking complexes, with minimum binding energies that are overbound by no more than about 5% for both the adenine−benzene and the cytosine−benzene systems. The locations of the potential energy minima are also in good agreement with estimated CCSD(T) results, although it should be noted that the optimum separation for the cytosine−benzene complex is slightly too short. Brief inspection of Figures 5 and 6 reveals that the features of the potential energy curves generated using the M06-2X functional are very different than those produced with estimated CCSD(T) interaction energies. For both complexes, this functional produces curves that are too steep near the minima, resulting in very narrow potential wells. In the case of the adenine−benzene complex, the optimum separation is predicted to be slightly too short, with a binding energy that is in good agreement with CCSD(T) results. However, for the cytosine−benzene complex, the minimum energy separation is too short by about 0.1 Å, with a binding energy that is approximately 20% too low. The incorrect long-range behavior of the M06-2X functional is due to the fact that the dispersion energy was covered by reparametrization of the exchange functional and not by the correlation one.

Dispersion Interactions. Potential energy curves of the propane dimer, whose chief mode of interaction is dispersion, for all methods considered here are given in Figure 7. It can directly be seen that MP2.5 and DFT-SAPT are the only methods producing good potential energy curves and that, among all of the MP2- and DFT-based methods, none can be said to be in excellent agreement with estimated CCSD(T) results. Both MP2.5 and DFT-SAPT give the correct location for the potential energy minimum, with binding energies that

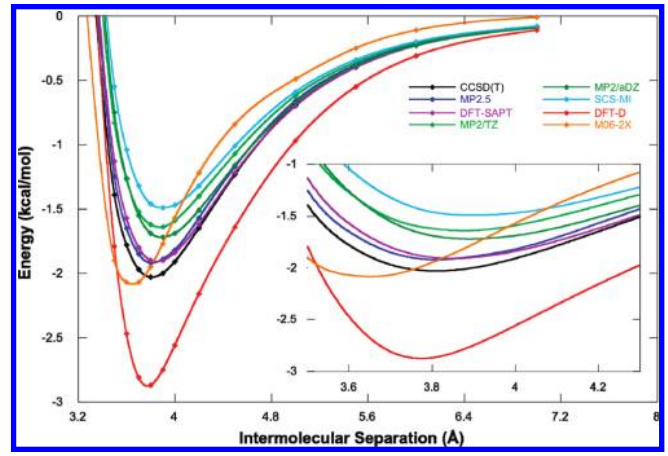

Figure 7. Potential energy curves for the propane dimer using several electronic structure methods (see text for exact description of methods used).

are underbound about approximately 6%. Among all other methods considered here, only DFT-D predicts the correct point for the potential energy minimum, at a separation of about 3.8 Å, but this method greatly overbinds the complex (by 0.84 kcal/mol or approximately 41%). This overbinding is most probably attributable to the fact that the S22 test set, from which DFT-D was parametrized, is heavily weighted toward sp^2-hybridized carbons (aromatic systems). The MP2/aug-cc-pVDZ, MP2/cc-pVTZ, and SCS(MI)-MP2/cc-pVTZ methods all underestimate the binding energy of the propane dimer and all predict the potential energy minimum to be located at a separation close to 3.9 Å. The performance of SCS(MI)-MP2 is particularly disappointing, with a binding energy that is 0.54 kcal/mol too low. In terms of the binding energy at the potential energy minimum, the M06-2X DFT gives the best result, with a binding energy that is only 0.04 kcal/mol higher than that of the CCSD(T) result. This minimum is located at 3.65 Å, which is too small a separation. It should be noted that at longer ranges the potential energy curve produced by the M06-2X functional deviates significantly from that of CCSD(T) (and those of the other methods), with energies that rise too sharply in the range between the minimum and about 4.5 Å. The result is a potential energy well that is too narrow near the minimum. Notice that this method has very similar behavior for both stacked complexes described in the previous paragraph.

$O-H\cdots\pi$ Interactions. One of the most noteworthy aspects of the curves shown for the interaction between benzene and water in Figure 8 is the fact that, as in the case of the dispersion interactions, MP2.5 and DFT-SAPT are the only computational techniques whose potential energy curves closely match the CCSD(T) results. The MP2.5 method overbinds near the potential energy minimum, whereas DFT-SAPT tends to underbind, however, neither of these methods is in error by more than about 1% of the minimum. All MP2-based methods tend to underbind this complex, while the DFT-based methods both overbind. MP2/cc-pVTZ, MP2/aug-cc-pVDZ, and SCS(MI)-MP2 all produce curves whose minimum energy separations are too large (by about 0.05−0.1 Å) and whose binding energies are too

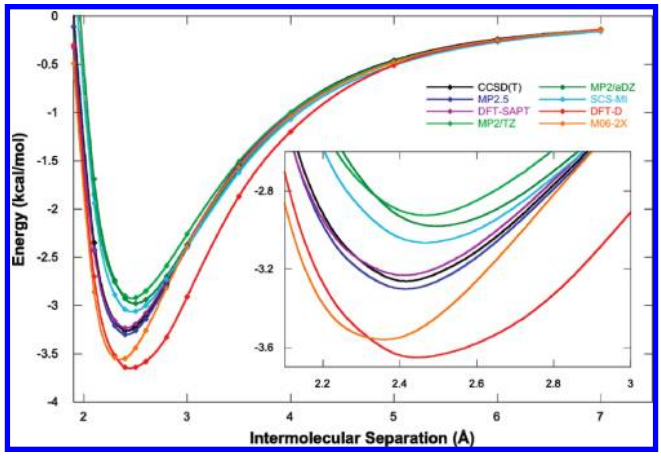

Figure 8. Potential energy curves for benzene−water using several electronic structure methods (see text for exact description of methods used).

low. The strongest underbinding tendencies are produced by the MP2/cc-pVTZ method, which is in error by approximately 10%. SCS(MI)-MP2 produces the best binding energy results for this mixed dispersion/electrostatic complex, with a binding energy that is within approximately 6% of the reference value. The binding energies produced at the potential energy minima for both DFT-D and M06-2X are overbound by approximately 10%. While DFT-D, like the MP2 methods, gives an optimum separation that is slightly too large, M06-2X gives a separation that is slightly too small (by 0.05 Å). It is also interesting to note that the curve produced with the DFT-D method appears to be somewhat too broad compared to the reference curve. Notice that, in this case, the long-range behavior is correct and well comparable with the reference CCSD(T) calculations. This might be explained by the fact that the electrostatic term represents the leading energy contribution.

Performance of the MP2/6-31G(0.25) Method.* In the Supporting Information, Figures S1−S7 give the MP2/6-31G*(0.25) potential energy curves, along with estimated CCSD(T)/CBS, MP2/cc-pVTZ, and MP2/aug-cc-pVDZ data, for all of the complexes considered in this work. One of the most striking aspects of these data is the fact that this method outperforms MP2/cc-pVTZ and MP2/aug-cc-pVDZ for both stacking interactions, being slightly overbound for the adenine−benzene complex and slightly underbound for the cytosine−benzene complex. This is in good agreement with previous results for stacked systems, where it was found that the 6-31G*(0.25) basis is among the best performers for MP2 binding energies of stacked systems.[32,72] It can also be seen in Table 1 that the potential energy minima obtained with MP2/6-31G*(0.25) are in very good agreement with the reference data. Unfortunately this method's exceptional performance for stacked systems does not translate to the other interaction motifs. It can be seen in the Supporting Information, Figures S3−S5, that MP2/6-31G*(0.25) is significantly underbound for all of the H-bonding complexes and also gives optimum intermolecular separations that are too wide (see also Table 1). Binding curves for the propane dimer and benzene−water complex are given in the Supporting Information, Figures S6−S7, respectively. MP2/6-31G*(0.25) is underbound and gives too large an optimum intermolecular separation for both of these complexes. The most problematic case is clearly the propane dimer for which the method gives a binding energy that is far too weak (by a factor of 2) and an intermolecular separation that is about 0.3 Å too wide.

DFT-SAPT Decomposition of Interactions. As noted above, the DFT-SAPT technique determines the binding energy of a complex as a sum of physically meaningful terms, namely the electrostatic, exchange, dispersion, and induction contributions. Figure 9 gives the curves for DFT-SAPT decomposition terms for each of the interactions considered in this work (with the exception of adenine−benzene). In Figure 9, it can be seen that the two main components of all of these interactions are electrostatics and dispersion. However, in terms of their interaction energy components, the interaction types are quite different, being dominated either by electrostatics or dispersion or, as in the case of the benzene−water complex, by having large contributions from both electrostatic and dispersive forces.

Figure 9a, b, and c gives the DFT-SAPT decompositions for the H-bonded systems considered in this work. Here it can be seen that, as would be expected, electrostatics play the dominant role in stabilizing these complexes. At their potential energy minima, electrostatic effects account for about 59% of the total attractive forces in the methanol and formamide dimers. It is somewhat surprising that the electrostatic contribution found in the formamide dimer, which is bound by a cyclic network of two H-bonds, is not greater than that of the methanol dimer. It should be noted, however, that the induction and δHF contributions are both larger for the formamide dimer (13% induction) than for the methanol dimer (11% induction). It should be pointed out that induction is generally the biggest contributor to the higher-order terms within the δHF term. The methylamine dimer interaction is the weakest H-bonding interaction considered in this work and is also the least electrostatic in nature (54% electrostatic contribution to the total attractive energy). Dispersion interactions are ubiquitous throughout intermolecular interaction types and play a role in the stabilization of H-bonded complexes. Dispersion accounts for 22, 30, and 17% of the attractive interactions in the methanol, methylamine, and formamide dimers, respectively (at their potential energy minima).

The DFT-SAPT interaction energy analysis for the cytosine−benzene dimer is given in Figure 9d. Here it is apparent that dispersion is the dominant contributor to this stacking interaction, with electrostatics playing a lesser role. At its potential energy minimum (3.6 Å), the attractive interactions present in this complex are about 74% dispersion, 19% electrostatic, and 4% induction (2% δHF). Interestingly, the electrostatic interaction increases relative to the dispersion interaction as the separation distance grows shorter. For example, at a separation of 3.4 Å, dispersion is responsible for only 67% and electrostatics about 25% of the attractive interaction. It should be noted that, in cases where two heterocyclic aromatic groups are stacked, the contribution from electrostatics will generally be higher than in this heterocyclic aromatic and aromatic complex. As an example, in DFT-SAPT computations recently carried out on the

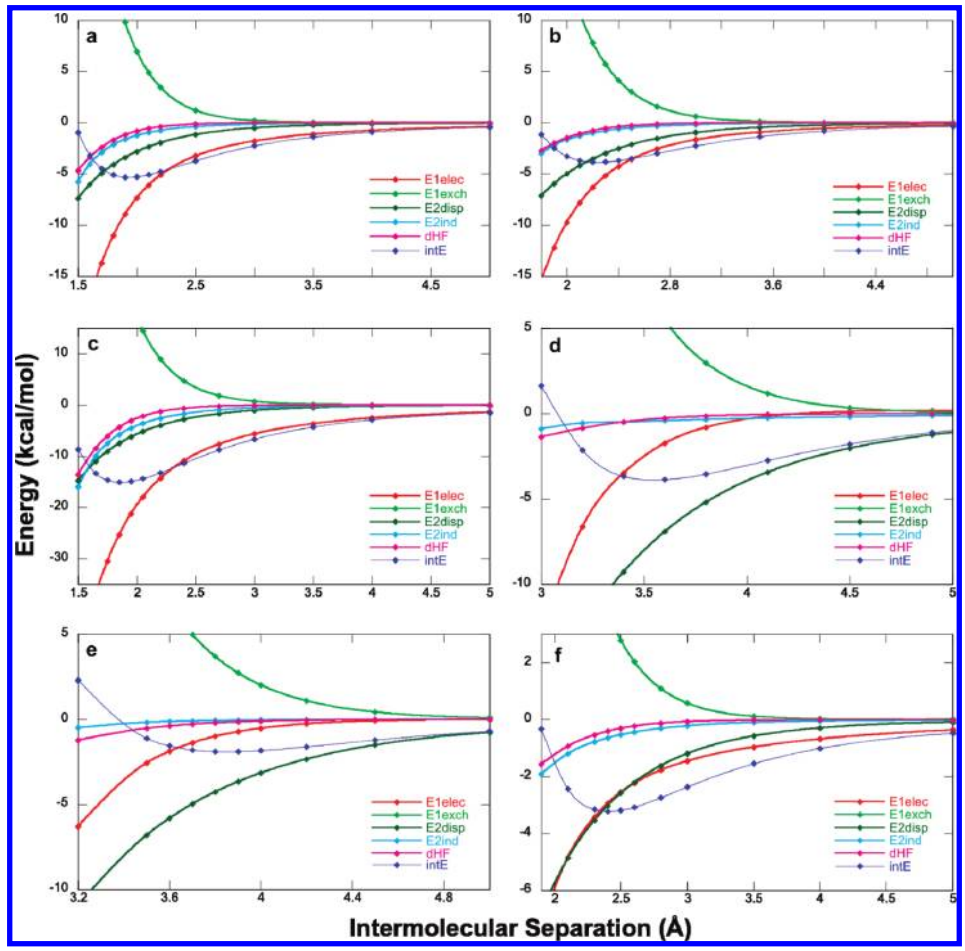

Figure 9. DFT-SAPT interaction energy decompositions for: (a) methanol dimer, (b) methylamine dimer, (c) formamide dimer, (d) cytosine−benzene, (e) propane dimer, and (f) benzene−water. E1elec (red), E1exch (light green), E2disp (dark green), E2ind (light blue), dHF (pink), intE (dark blue).

stacked uracil dimer by Pitoňák et al., it was found that, at the potential energy minimum, dispersion describes approximately 52% and electrostatics 41% of the total attractive interaction.[17]

As in the case of the cytosine−benzene complex, the interaction in the propane dimer is dominated by dispersion forces (Figure 9e). This is, of course, an expected result as there are no strong dipole moments found in propane. At the potential energy minimum (3.9 Å), dispersion accounts for about 80% of the attraction in this complex, with electrostatics describing about 16%. The electrostatic contribution grows slightly more quickly than the dispersion contribution with decreasing intermolecular separation but not nearly as quickly as in the case of the cytosine−benzene complex. At a separation of 3.5 Å, dispersion describes approximately 67% and electrostatics 25% of the total attractive interaction.

The O−H$\cdots\pi$ interaction found in the benzene−water complex depends strongly on both dispersion and electrostatic contributions, with induction also playing an appreciable role in stabilizing the complex (Figure 9f). Between the distances of 1.9 and 2.6 Å, the contributions from dispersion and electrostatics to the overall attraction in this complex are almost identical (approximately 40−44%). Beyond 2.6 Å, the electrostatic term begins to have a larger relative attractive contribution relative to dispersion. The increase in the relative electrostatic stabilization of the complex at larger distances can most likely be explained by the large spacial extent of benzene's π density and by the interaction between this π ring and water's positively charged hydrogen atom.

Conclusions

One of the major goals of this investigation is to determine the quality of several modern quantum chemical methods in describing potential energy curves for a variety of different types of intermolecular complexes. It has been found that, generally speaking, each of the methods applied in this work is capable of describing the H-bonding, stacking, dispersion, and O−H$\cdots\pi$ interactions considered here at least qualitatively. That is to say that each of the methods predicts bound states that correspond, upon gross inspection, to the particular interaction type being studied. In the cases of the two DFT-based methods, DFT-D and M06-2X, this is a major accomplishment considering the fact that just a few years ago it was widely acknowledged that existing DFT methods were incapable of describing noncovalent interactions, whose attractive forces are largely attributable to dispersion.

Here, we will summarize some of the general trends that have been observed in this study. As would be expected, the H-bonding interactions are generally well described by

each of the methods used here. MP2, combined with both the cc-pVTZ and aug-cc-pVDZ basis sets, has a strong tendency to underbind these types of interactions but provides reliable geometries. The former basis set yields better equilibrium geometries. The use of the spin-component scaled technique (SCS(MI)-MP2) with the cc-pVTZ basis improves the performance of MP2, but these binding energies are also generally slightly too low. The potential energy curves produced by MP2.5, M06-2X, and (to a lesser extent) DFT-SAPT match those obtained using the estimated CCSD(T)/CBS method (CBS(T)) very well. DFT-D tends to (sometimes strongly) overbind H-bonding interactions.

In terms of stacking interactions, the MP2 methods, as has been previously observed, have a strong tendency to overbind, this is especially true when MP2 is paired with the aug-cc-pVDZ basis set. The SCS(MI)-MP2 method greatly improves on the MP2 results, reducing the amount of overbinding significantly and, in the case of the cytosine−benzene complex, actually underbinding at the potential energy minimum. All calculations based on the MP2 procedure yield reliable distances of the minima, which allows for the structural optimization of various complexes at a rather cheap level (inclusion of the counterpoise correction is, however, necessary). The cc-pVTZ basis set yields better geometries than aug-cc-pVDZ ones. The MP2.5, DFT-SAPT, and DFT-D methods all produce potential energy curves that are in good agreement with those of CBS(T). The DFT/M06-2X potential energy curves for stacked systems are generally not in good agreement with the reference data, being strongly underbound for the cytosine−benzene complex, predicting incorrect minimum energy separations for both complexes, and are generally having curves with the wrong overall shape.

Only the relatively expensive MP2.5 and DFT-SAPT methods can be said to produce high-quality potential energy curves for the propane dimer. All MP2 methods, including SCS(MI)-MP2, tend to strongly underestimate the binding energy of this complex, while DFT-D very strongly overestimates it. DFT/M06-2X, on the other hand, obtains a reasonable value for the binding energy at the potential energy minimum but predicts the minimum to be at too small an intermolecular separation.

For the benzene−water complex MP2.5 and DFT-SAPT are, once again, the methods that produce the best results relative to those of CBS(T). As in the cases of H-bonding complexes and the propane dimer, all of the MP2 methods studied here underbind this O−H$\cdots\pi$ complex. Both DFT based methods, on the other hand, tend to overbind the complex, with DFT-D having a potential energy curve that appears to be much too broad.

Generally speaking, the only two methods that can be said to provide accurate potential energy curves for all of the complexes considered here, apart from the reference CBS(T) method, are MP2.5 and DFT-SAPT. Unfortunately, these methods are computationally very expensive and can only be used on complexes containing relatively few atoms (up to ∼60−80). Furthermore, DFT-SAPT has another two major disadvantages compared to other methods investigated in this work. First, the analytic gradients needed for optimization of geometries have not been formulated or implemented yet, and second, in DFT-SAPT potential energy surface calculations, only rigid monomers can be considered (the deformation energy cannot be included). The MP2 method, long the "workhorse" used for computations on molecular complexes, only produces very good potential energy curves for H-bonding complexes, otherwise its performance can be said to be semiquantitatively accurate. MP2 results are generally better when the method is used in conjunction with the cc-pVTZ basis set, and this result agrees well with our previous finding.[72] The SCS(MI)-MP2/cc-pVTZ method, which seeks to improve the results of MP2/cc-pVTZ, is largely successful in this task, with improved potential energy curves for all of the noncovalently bound complexes with the exception of the propane dimer. In two of our previous studies, we thoroughly investigated the PES's of the uracil and adenine dimers, and in both studies, the SCS (MI)-MP2 method provided very good results, well comparable to those of the benchmark CCSD(T)/CBS method.[82,94] In the present study, SCS(MI)-MP2 gives very good results for all of the complexes with exception of the propane dimer. We believe that the reason for this is the same as disscussed above for the DFT-D method, which is the fact that SCS(MI)-MP2 as well as DFT-D were parameterized against the S22 set, which lacks systems containing carbon atoms having sp3 hybridization. On the other hand, the aromatic systems, such as DNA bases and benzene, are well represented in the set, and the method provides good results, even for complexes not included within the S22 set (e.g., adenine dimer or adenine−benzene complex). This finding is important and should be kept in mind when preparing data sets of the second generation. Among the much less computationally expensive DFT-based methods, DFT-D can be said to yield the best performance, giving accurate potential energy curves for H-bonding and stacking interactions. This method, however, tends to strongly overbind for both the propane dimer and the benzene−water complex. The M06-2X functional produces good results for H-bonding and O−H$\cdots\pi$ interactions but produces curves for stacked and dispersion-bound complexes that generally have the wrong overall shape.

Acknowledgment. This work was a part of the research project no. Z40550506 of the Institute of Organic Chemistry and Biochemistry, Academy of Sciences of the Czech Republic, and it was supported by grant nos. LC512 and MSM6198959216 from the Ministry of Education, Youth and Sports of the Czech Republic. This work was also supported by the Institutional Research concept no. AV0Z50520701 of the Academy of Sciences of the Czech Republic. The support of Praemium Academiae, Academy of Sciences of the Czech Republic, awarded to P.H. in 2007 is also acknowledged. M.P. gratefully acknowledges the support of the Slovak Research and Development Agency (contract no. APVV-20-018405) and the Slovak Grant Agency VEGA (contract no. 1/0428/09). K.R. gratefully acknowledges the support of the National Science Foundation EPSCOR program (EPS-0701525). This work was also supported by Korea Science and Engineering Foundation (World Class University program: R32-2008-000-10180-0).

The authors are grateful for computer time given through the 'Chinook' supercomputer at the Environmental Molecular Sciences Laboratory of the Pacific Northwest National Laboratory.

Supporting Information Available: The MP2/6-31G*(0.25) potential energy curves, along with estimated CCSD(T)/CBS, MP2/cc-pVTZ, and MP2/aug-cc-pVDZ data, for all of the complexes considered in this work. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Baldwin, R. L. *J. Mol. Biol.* **2007**, *371*, 283.

(2) Loladze, V. V.; Ermolenko, D. N.; Makhatadze, G. I. *J. Mol. Biol.* **2002**, *320*, 343.

(3) Berka, K.; Laskowski, R.; Riley, K. E.; Hobza, P.; Vondrasek, J. *J. Chem. Theory Comput.* **2009**, *5*, 982.

(4) Vondrasek, J.; Bendova, L.; Klusak, V.; Hobza, P. *J. Am. Chem. Soc.* **2005**, *127*, 2615.

(5) Vondrasek, J.; Kubar, T.; Jenney, F. E.; Adams, M. W. W.; Kozisek, M.; Cerny, J.; Sklenar, V.; Hobza, P. *Chem.−Eur. J.* **2007**, *13*, 9022.

(6) Riley, K. E.; Merz, K. M. *J. Phys. Chem. B* **2006**, *110*, 15650.

(7) Wilchek, M.; Bayer, E. A.; Livnah, O. *Immunol. Lett.* **2006**, *103*, 27.

(8) Meyer, E. A.; Castellano, R. K.; Diederich, F. *Angew. Chem., Int. Ed.* **2003**, *42*, 1210.

(9) Williams, D. H.; Stephens, E.; O'Brien, D. P.; Zhou, M. *Angew. Chem., Int. Ed.* **2004**, *43*, 6596.

(10) Sponer, J.; Riley, K. E.; Hobza, P. *Phys. Chem. Chem. Phys.* **2008**, *10*, 2595.

(11) Murray, J. S.; Lane, P.; Clark, T.; Politzer, P. *J. Mol. Model.* **2007**, *13*, 1033.

(12) Politzer, P.; Lane, P.; Concha, M. C.; Ma, Y. G.; Murray, J. S. *J. Mol. Model.* **2007**, *13*, 305.

(13) Riley, K. E.; Hobza, P. *J. Chem. Theory Comput.* **2008**, *4*, 232.

(14) Riley, K. E.; Murray, J. S.; Politzer, P.; Concha, M. C.; Hobza, P. *J. Chem. Theory Comput.* **2009**, *5*, 155.

(15) Sponer, J.; Leszczynski, J.; Hobza, P. *Biopolymers* **2001**, *61*, 3.

(16) Sponer, J.; Leszczynski, J.; Hobza, P. *Theochem* **2001**, *573*, 43.

(17) Pitonak, M.; Riley, K. E.; Neogrady, P.; Hobza, P. *ChemPhysChem* **2008**, *9*, 1636.

(18) Hobza, P.; Zahradnik, R. *Chem. Rev.* **1988**, *88*, 871.

(19) Hofstadler, S. A.; Griffey, R. H. *Chem. Rev.* **2001**, *101*, 377.

(20) Kollman, P. A. *Acc. Chem. Res.* **1977**, *10*, 365.

(21) Muller-Dethlefs, K.; Hobza, P. *Chem. Rev.* **2000**, *100*, 143.

(22) Zhao, Y.; Truhlar, D. G. *J. Chem. Theory Comput.* **2007**, *3*, 289.

(23) Zhao, Y.; Truhlar, D. G. *J. Chem. Theory Comput.* **2008**, *4*, 1849.

(24) Lee, E. C.; Kim, D.; Jurecka, P.; Tarakeshwar, P.; Hobza, P.; Kim, K. S. *J. Phys. Chem. A* **2007**, *111*, 3446.

(25) Dykstra, C. E.; Lisy, J. M. *Theochem* **2000**, *500*, 375.

(26) Hartmann, M.; Wetmore, S. D.; Radom, L. *J. Phys. Chem. A* **2001**, *105*, 4470.

(27) McConnell, T. L.; Wheaton, C. A.; Hunter, K. C.; Wetmore, S. D. *J. Phys. Chem. A* **2005**, *109*, 6351.

(28) Tsuzuki, S.; Uchimaru, T. *Curr. Org. Chem.* **2006**, *10*, 745.

(29) Rozas, I. *Phys. Chem. Chem. Phys.* **2007**, *9*, 2782.

(30) Rappe, A. K.; Bernstein, E. R. *J. Phys. Chem. A* **2000**, *104*, 6117.

(31) Sinnokrot, M. O.; Sherrill, C. D. *J. Phys. Chem. A* **2004**, *108*, 10200.

(32) Riley, K. E.; Hobza, P. *J. Phys. Chem. A* **2007**, *111*, 8257.

(33) Jurecka, P.; Sponer, J.; Cerny, J.; Hobza, P. *Phys. Chem. Chem. Phys.* **2006**, *8*, 1985.

(34) Jurecka, P.; Hobza, P. *Chem. Phys. Lett.* **2002**, *365*, 89.

(35) Pitonak, M.; Janowski, T.; Neogrady, P.; Pulay, P.; Hobza, P. Journal of Chemical Theory and Computation 2009, In Press: DOI 10.1021/ct900126q.

(36) Cerny, J.; Pitonak, M.; Riley, K. E.; Hobza, P. *J. Chem. Theory Comput.* **2009**, To be submitted.

(37) Sponer, J.; Leszczynski, J.; Hobza, P. *J. Phys. Chem.* **1996**, *100*, 5590.

(38) Rutledge, L. R.; Durst, H. F.; Wetmore, S. D. *J. Chem. Theory Comput.* **2009**, *5*, 1400.

(39) Johnson, E. R.; Wolkow, R. A.; DiLabio, G. A. *Chem. Phys. Lett.* **2004**, *394*, 334.

(40) Cerny, J.; Jurecka, P.; Hobza, P.; Valdes, H. *J. Phys. Chem. A* **2007**, *111*, 1146.

(41) Jurecka, P.; Cerny, J.; Hobza, P.; Salahub, D. R. *J. Comput. Chem.* **2007**, *28*, 555.

(42) Zhao, Y.; Truhlar, D. G. *J. Chem. Phys.* **2006**, *125*, 194101.

(43) Zhao, Y.; Truhlar, D. G. *Theor. Chem. Acc.* **2008**, *120*, 215.

(44) Zhao, Y.; Pu, J. Z.; Lynch, B. J.; Truhlar, D. G. *Phys. Chem. Chem. Phys.* **2004**, *6*, 673.

(45) Zhao, Y.; Schultz, N. E.; Truhlar, D. G. *J. Chem. Theory Comput.* **2006**, *2*, 364.

(46) Zhao, Y.; Truhlar, D. G. *J. Chem. Theory Comput.* **2005**, *1*, 415.

(47) Zhao, Y.; Truhlar, D. G. *Acc. Chem. Res.* **2008**, *41*, 157.

(48) Grimme, S. *J. Comput. Chem.* **2004**, *25*, 1463.

(49) Grimme, S. *J. Comput. Chem.* **2006**, *27*, 1787.

(50) Hohenstein, E. G.; Chill, S. T.; Sherrill, C. D. *J. Chem. Theory Comput.* **2008**, *4*, 1996.

(51) Hesselmann, A.; Jansen, G. *Phys. Chem. Chem. Phys.* **2003**, *5*, 5010.

(52) Hesselmann, A.; Jansen, G. *Chem. Phys. Lett.* **2003**, *367*, 778.

(53) Hesselmann, A.; Jansen, G. *Chem. Phys. Lett.* **2002**, *362*, 319.

(54) Hesselmann, A.; Jansen, G. *Chem. Phys. Lett.* **2002**, *357*, 464.

(55) Jansen, G.; Hesselmann, A. *J. Phys. Chem. A* **2001**, *105*, 11156.

(56) Hesselmann, A.; Jansen, G.; Schutz, M. *J. Chem. Phys.* **2005**, *122*.

(57) Cybulski, S. M.; Lytle, M. L. *J. Chem. Phys.* **2007**, *127*, 141102.

(58) Hesselmann, A. *J. Chem. Phys.* **2008**, *128*, 9.

(59) Grimme, S. *J. Chem. Phys.* **2003**, *118*, 9095.

(60) Distasio, R. A.; Head-Gordon, M. *Mol. Phys.* **2007**, *105*, 1073.

(61) Takatani, T.; Sherrill, C. D. *Phys. Chem. Chem. Phys.* **2007**, *9*, 6106.

(62) Hill, J. G.; Platts, J. A. *J. Chem. Theory Comput.* **2007**, *3*, 80.

(63) Sherrill, C. D.; Takatani, T.; Hohenstein, E. G. *J. Phys. Chem. A* **2009**, *113*, 10146.

(64) Pitonak, M.; Neogrady, P.; Cerny, J.; Grimme, S.; Hobza, P. *ChemPhysChem* **2009**, *10*, 282.

(65) Arnstein, S. A.; Sherrill, C. D. *Phys. Chem. Chem. Phys.* **2008**, *10*, 2646.

(66) Ringer, A. L.; Figgs, M. S.; Sinnokrot, M. O.; Sherrill, C. D. *J. Phys. Chem. A* **2006**, *110*, 10822.

(67) Ringer, A. L.; Sinnokrot, M. O.; Lively, R. P.; Sherrill, C. D. *Chem.*−*Eur. J.* **2006**, *12*, 3821.

(68) Sinnokrot, M. O.; Sherrill, C. D. *J. Phys. Chem. A* **2006**, *110*, 10656.

(69) Sinnokrot, M. O.; Valeev, E. F.; Sherrill, C. D. *J. Am. Chem. Soc.* **2002**, *124*, 10887.

(70) Tauer, T. P.; Derrick, M. E.; Sherrill, C. D. *J. Phys. Chem. A* **2005**, *109*, 191.

(71) Janowski, T.; Pulay, P. *Chem. Phys. Lett.* **2007**, *447*, 27.

(72) Dabkowska, I.; Gonzalez, H. V.; Jurecka, P.; Hobza, P. *J. Phys. Chem. A* **2005**, *109*, 1131.

(73) Peverati, R.; Baldridge, K. K. *J. Chem. Theory Comput.* **2008**, *4*, 2030.

(74) Tapavicza, E.; Lin, I. C.; von Lilienfeld, O. A.; Tavernelli, I.; Coutinho-Neto, M. D.; Rothlisberger, U. *J. Chem. Theory Comput.* **2007**, *3*, 1673.

(75) Williams, R. W.; Malhotra, D. *Chem. Phys.* **2006**, *327*, 54.

(76) Tsuzuki, S.; Uchimaru, T.; Mikami, M.; Tanabe, K. *J. Phys. Chem. A* **2002**, *106*, 3867.

(77) Jalkanen, J. P.; Mahlanen, R.; Pakkanen, T. A.; Rowley, R. L. *J. Chem. Phys.* **2002**, *116*, 1303.

(78) Jalkanen, J. P.; Pakkanen, T. A.; Rowley, R. L. *J. Chem. Phys.* **2004**, *120*, 1705.

(79) Jalkanen, J. P.; Pulkkinen, S.; Pakkanen, T. A.; Rowley, R. L. *J. Phys. Chem. A* **2005**, *109*, 2866.

(80) Rowley, R. L.; Tracy, C. M.; Pakkanen, T. A. *J. Chem. Phys.* **2006**, *125*, 154302.

(81) Rowley, R. L.; Tracy, C. M.; Pakkanen, T. A. *J. Chem. Phys.* **2007**, *127*, 025101.

(82) Morgado, C. A.; Jurecka, P.; Svozil, D.; Hobza, P.; Sponer, J. *J. Chem. Theory Comput.* **2009**, *5*, 1524.

(83) Tekin, A.; Jansen, G. *Phys. Chem. Chem. Phys.* **2007**, *9*, 1680.

(84) Tsuzuki, S.; Honda, K.; Uchimaru, T.; Mikami, M. *J. Chem. Phys.* **2006**, *124*, 7.

(85) Fusti Molnar, L.; He, X.; Wang, B.; Merz, K. M. *J. Chem. Phys.* **2009**, *131*, AN 065102.

(86) Halkier, A.; Helgaker, T.; Jorgensen, P.; Klopper, W.; Koch, H.; Olsen, J.; Wilson, A. K. *Chem. Phys. Lett.* **1998**, *286*, 243.

(87) Werner, H.-J.; Knowles, P. J.; Lindh, R.; Manby, F. R.; Schütz, M.; Celani, P.; Korona, T.; Rauhut, G.; Amos, R. D.; Bernhardsson, A.; Berning, A.; Cooper, M. J. O.; Deegan, A. J.; Dobbyn, A. J.; Eckert, F.; Hampel, C.; Hetzer, G.; LLoyd, A. W.; McNicholas, S. J.; Meyer, W.; Mura, M. R.; Nicklass, A.; Palmieri, P.; Pitzer, R.; Schumann, U.; Stoll, H.; Stone, A. J.; Taroni, T.; Thorsteinsson, T. *Molpro 7*, version 2006, University College Cardiff Consultants Limited: Cardiff, U.K.

(88) Feyereisen, M.; Fitzgerald, G.; Komornicki, A. *Chem. Phys. Lett.* **1993**, *208*, 359.

(89) Ahlrichs, R.; Bar, M.; Haser, M.; Horn, H.; Kolmel, C. *Chem. Phys. Lett.* **1989**, *162*, 165.

(90) Jurecka, P.; Nachtigall, P.; Hobza, P. *Phys. Chem. Chem. Phys.* **2001**, *3*, 4578.

(91) Boys, S. F.; Bernardi, F. *Mol. Phys.* **1970**, *19*.

(92) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Chesseman, J. R.; Zakrzewski, V. G.; Montgomery Jr., J. A.; Stratmann, R. E.; Burant, J. C.; Dapprich, S.; Millam, J. M.; Daniels, A. D.; Kudin, K. N.; Strain, M. C.; Farkas, O.; Tomasi, J.; Barone, V.; Cossi, M.; Cammi, R.; Mennucci, B.; Pomelli, C.; Adamo, C.; Clifford, S.; Ochterski, J.; Petersson, G. A.; Ayala, P. Y.; Cui, Q.; Morokuma, K.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Forseman, J. B.; Cioslowski, J.; Ortiz, J. V.; Baboul, A. G.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Keith, T.; AlLoham, M. A.; Peng, C. Y.; Nanayakkara, A.; Gonzalez, C.; Challacombe, M.; Gill, P. M. W.; Johnson, B. G.; Chen, W.; Wong, M. W.; Andres, J. L.; Head-Gordon, M.; Replogle, E. S.; Pople, J. A. *Gaussian 03*; Gaussian Inc.: Wallingford, CT, 2003.

(93) Aquilante, F.; De Vico, L.; Ferré, N.; Ghigo, G.; Malmqvist, P.; Neogrády, P.; Pedersen, T. B.; Pitonák, M.; Reiher, M.; Roos, B. O.; Serrano-Andrés, L.; Urban, M.; Veryazov, V.; Lindh, R. *J. Comput. Chem.* **2009**, In Press.

(94) Morgado, C. A.; Jurecka, P.; Svozil, D.; Hobza, P.; Sponer, J. *Phys. Chem. Chem. Phys.* **2009**, submitted.

CT900376R