JCTC Journal of Chemical Theory and Computation

Basis Set Superposition Error along the Free-Energy Surface of the Water Dimer

Jens Thar,[†] Rainer Hovorka,[‡] and Barbara Kirchner*[,†]

Lehrstuhl für Theoretische Chemie, Wilhelm-Ostwald Institut für Physikalische und Theoretische Chemie, Universität Leipzig, Linnestrasse 2, D-04103 Leipzig, Germany, and Organische Chemie, Kekulé Institut für Organische Chemie und Biochemie, Universität Bonn, Gerhard-Domagk-Strasse 1, D-53121 Bonn, Germany

Received August 9, 2006

Abstract: In this article we review the behavior of static plane wave basis set calculations in comparison to Gaussian basis set calculations. This was done in the framework of density functional theory for description of hydrogen bonds with the water dimer as an example. Furthermore we carried out molecular dynamics simulations enforcing the self-dissociation reaction of the water dimer to study the influence of the basis set onto the reaction. Not surprisingly, we find strongly varying results of the calculated forces for a chosen cutoff along the reaction coordinates. The basis set superposition errors of the dimer interaction energy are analyzed along the free-energy surface, i.e., along the trajectories. Based on the analysis along the trajectories a qualitative and quantitative estimate depending on the particular point of the free-energy surface can be provided. Namely, at the intermolecular O···H distance close to the equilibrium geometry the errors are smaller than at shorter O···H distances. However, the distribution at the equilibrium distance is more unsymmetrical than the distribution at short distances. It is wider, and the standard deviation is larger than at shorter distances where the basis set superposition error is larger.

1. Introduction

The plane wave basis set (PWBS) combined with pseudopotentials and density functional theory (DFT) is the standard method used in many first-principles simulations (FPMD).[1−6] A reason for the heavy use of this combination lies in the nice property of low computational costs together with the advantage of easy technical applicability.[7] The importance of low computational costs is understandable when keeping in mind that along a trajectory in each step a quantum chemical calculation has to be carried out. For a total simulation time of 10 ps with a time step of 0.1 fs about 100 000 of such calculations are needed.

Basis sets used in standard quantum chemical calculations are usually built up by atom-centered functions.[8,9]

When comparing interaction energies (E_I) at different geometries the atom-centered basis set (e.g., the Gaussian Basis Set (GBS)) introduces the well studied basis set superposition error (BSSE).[10] The quality of the basis set is not the same at all geometries, owing to the fact that the electron density around one nucleus may be described by functions centered at another nucleus, see Figure 1. The counterpoise correction procedure introduced by Boys and Bernardi is the standard method to correct for the BSSE.[11] The BSSE represents a strong disadvantage of the Gaussian basis set and a strong argument in favor of the plane wave basis set especially for simulations of large systems and many electronic structure calculations. Since plane waves are not atom centered functions, such an effect does not appear in calculations employing them. On the other hand, plane wave basis sets have an extended dimension and cannot describe the compact charge densities as accurately as the Gaussian basis set.

* Corresponding author e-mail: bkirchner@uni-leipzig.de.
[†] Universität Leipzig.
[‡] Universität Bonn.

10.1021/ct600259x CCC: $37.00
Published on Web 06/12/2007

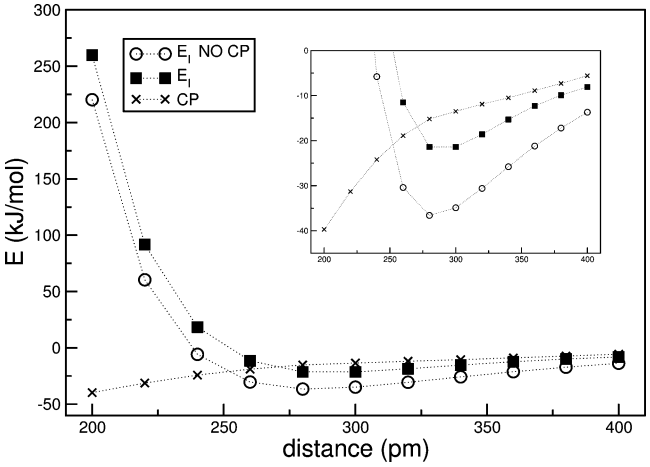

Figure 1. Interaction energy E_I in kJ/mol not counterpoise corrected (NO CP) (empty circles), counterpoise corrected (filled squares), and counterpoise correction (CP) (crosses) of the water dimer at different $r_{O^\star H^\star}$ intermolecular distances in pm but otherwise in the global minimum conformation. All calculations were performed with the BP86/SV(P) combination of density functional and basis set.

Attempts to combine the advantages of both types of basis sets—GBS as well as PWBS—were successfully carried out for example in the Hutter group.[12−14] One needs to add here that all plane-wave and mixed basis set calculations can be used in combination with pseudopotentials, for a detailed description see refs 15−17. An approach that allows O(N)-scaling which employs pure Gaussians was introduced by Schlegel and co-workers.[18] A different route is taken by the Tuckerman group where a complete basis set limit for simulations is achieved with a discrete variable representation basis set.[19−21] While ref 19 is dedicated to the development and implementation of this real-space approach for the electronic structure calculations in FPMD, refs 20 and 21 deal with the simulation of liquid water. It was found by Lee and Tuckerman that less overstructuring in the radial pair distribution functions occurs with this kind of basis set ansatz. In ref 19 the authors also provide an informative overview over alternatives to the PWBS approach in condensed phase. BSSE-free methods which, however, are not described in the context of condensed phase are discussed in the literature, see ref 22 for an overview and some selected examples in refs 23−27. There are more technical issues than the question of the BSSE to investigate when electronic structure is calculated on the fly. A recent article by Kuo is highlighted here as an example.[28] In a very important and reliable comparison of the Car−Parrinello method[1] versus the Born−Oppenheimer molecular dynamics simulation technique the authors could show that, despite some beliefs, the Car−Parrinello method and the Born−Oppenheimer simulation technique produce equivalent results.[28]

Due to the problems associated with FPMD it is worth mentioning that other approaches to describe the liquid-phase such as water[29−32] exists. Of course there is the wealth of simulations applying empirical water potentials[33,34] and polarizable force fields.[35−38] Next to these important methods, Weinhold and Ludwig developed and refined an alternative to treat the liquid state by considering the so-called quantum cluster equilibrium (QCE) theory.[39,40] So far many problems using the QCE theory[39−42] were studied, for example, it was possible to determine the triple point of water and describe an icelike phase of water.[41] Investigating isotopically substituted water[42] and studying the cooperative versus dispersion effects[43] in liquid water was also the scope of the QCE studies. The importance of cooperative effects to be correctly described was also demonstrated by quantum calculations.[44] Employing pair potentials from ab initio calculations can be considered to be another area of liquid-phase simulations.[45] For water pioneering work was carried out by the Clementi group.[46,47] An ab initio constructed force field in order to provide "chemical accuracy" was constructed by Liu and co-workers.[48] Furthermore the very important and BSSE-free symmetry-adapted perturbation method[49] was also used to provide ab initio pair potentials for simulations of water.[50]

It can be recognized from Figure 1 that the size of the BSSE strongly depends on the geometry of the water dimer. It would therefore be difficult to work with Gaussian basis sets in simulations. So far no detailed study was published concerning the basis set superposition error in simulations. The reliability of (Gaussian-type) basis sets is usually only tested for static quantum chemical calculations and in the majority of cases for *intra*molecular properties like bond lengths; see, e.g., refs 51−53. However, the very recent article of Haynes and co-workers is interesting in our context, because the authors showed that the BSSE is eliminated by optimizing the local orbitals in situ using a basis set that is similar to a PWBS.[23]

The objective of this article is to try to assess the reliability of simulations as obtained from different PWBS calculations with varying basis set size and to assess the BSSE along the trajectories. We do that by simulating the barrier for the autoprotolysis within a water dimer employing the technique of thermodynamic integration. This means that we force one proton to move from one water molecule to the other water molecule within a distance constraint to form the OH^- + H_3O^+ ion-pair:

$$H_2O^\star\cdots H^\star-O-H \rightarrow H_3O^+ + OH^-$$

The reaction coordinate is herewith determined by the stepwise increased hydrogen bond distance $r_{O^\star H^\star}$. The technique of thermodynamic integration is a standard technique so we recall here only the essential equation. For a detailed discussion see refs 34 and 54. The difference in free energy ΔA is obtained by integrating the negative averaged force $-\langle f \rangle_{r_{O^\star H^\star}}$

$$-\langle f \rangle_{r_{O^\star H^\star}} = \frac{\partial A}{\partial r_{O^\star H_\star}} \qquad (1)$$

The self-dissociation of water is not a reaction that takes place under usual condition—standard pressure, temperature, and absent solvent—in the gas phase. Sobolewski and Domcke observed in their study about the hydrated hydronium that the ground state of the dissociation reaction correlates adiabatically with the formation of the OH^- + H_3O^+ ion-pair, while the excited state of the water dimer correlates with the biradical $OH-H_3O$ complex.[55] Their

studies are based on CASPT2 calculations using a modified ANO-L basis set upon B3LYP/6-311++G** structures. They found that the ion-pair configuration does not exist as a minimum on the ground-state potential energy surface. Nevertheless, the observed shallow plateau around 180 pm develops into a local minimum for a larger water cluster. Thus the $OH^- + H_3O^+$ ion-pair is stabilized by solvation. An overview over the present knowledge about the autoprotolysis can be found in refs 56 and 66.

Despite the fact that this reaction does not take place in the gas phase, it serves us to obtain many chemically different situations. During its course, each of this situation is associated with a unique BSSE when a Gaussian type orbital is applied, thus this reaction allows for the study of the BSSE during the course of the simulations. We are able to model the influence of the basis set quality onto the simulations by increasing or decreasing the number of plane waves in a particular calculation. In subsequent quantum chemical calculations applying the "cluster ansatz"[57,58] we then analyze the BSSE with two particular GBSs along the trajectories. It can be expected that by changing the chemical situation in the system we create situations which are sometimes more and sometimes less affected by the quality of the basis set.

2. Technical Details

The general setup for the simulations was chosen to be the same as for the static PWBS calculations, see the Supporting Information. Molecular dynamics simulations were performed in the NVT ensemble at 300 K using a Nosé−Hoover chain thermostating scheme.[59−61] All simulations were performed with the CPMD code.[62] It was shown previously that BLYP provides the best results for liquid water; therefore, our dynamical calculations are mainly done for this functional. Other functionals are only tested at one cutoff.[63] The quality of the plane wave basis set was determined by the energy cutoff E_{cut} which we selected to be 20, 50, 70, and 90 Ryd. We simulated each run with a time step of 5 au (=0.12094 fs) and with a fictitious mass of 600 au. The total number of time steps per simulation was 50 000, i.e., 6 ps. To obtain the free energy difference for each point such a trajectory was carried out, i.e., we ran in total 45 trajectories. Subsequently to this we calculated along one set (BLYP/70 Ryd with ranging from $r_{O^\star H^\star}$=100−180 pm) of the obtained trajectories the BSSE, i.e., we carried out 45000 single point quantum chemical calculation for each of the TZVPP and the SVP basis sets, see the Supporting Information.

3. Results

3.1. Static Considerations Revisited. We recall here well-known results for static calculations, because they provide the basis for the dynamical calculations. For the interested reader we provide a full discussion in the Supporting Information.

We start by summarizing the total energy results. For the BLYP functional the basis set limit as shown previously by Lee and Tuckerman[20] of 300 Ryd deviates from the value at 150 Ryd only by approximately 5 kJ/mol. We see that independent of the functional and of the water dimer structures, the calculations are converged within the first digit before the decimal point for E_{cut} = 50 Ryd, with respect to the reference (150 Ryd) calculation. The difference between 50 Ryd and 70 Ryd is thus in the range of 800 kJ/mol, except for the BP86 functional which only shows a difference of 250 kJ/mol. Using the standard cutoff of 70 Ryd, the energies converge within the first digit after the decimal point. This means that the error lowers to about 180 kJ/mol. Choosing the cutoff of 90 Ryd improves the convergence behavior to 0.01 hartree, i.e., an error of about 25 kJ/mol. Therefore we also recommend using a cutoff of 90 Ryd in calculations with systems containing water molecules if computer time is available as it is in general recommended but currently seldomly used. For a qualitative discussion a cutoff of 70 Ryd might be sufficient to capture all important chemical effects. The behavior for all functionals is similar, except that the BP86 values start at lower energies and converge faster.

The difference in interaction energies between the two structures converges to approximately 3 kJ/mol independently of the particular functional. We also see the usual functional dependencies, for example both PBE functionals give stronger binding energies than all other functionals and BP86 yields a higher binding energy than BLYP. Obviously more than one structure should be investigated, because for all values of E_{cut} the global minimum structure yields a reasonable interaction energy. Inspecting the results for a local minimum structure we realize that a cutoff energy of 20 Ryd is leading to absurd results, while cutoff 20 Ryd yields good results for global minimum, see the Supporting Information. From cutoff 70 Ryd and even 50 Ryd on we obtain comparable interaction energies for all chosen cutoffs of a particular functional. This is the reason why we expect cutoff 70 Ryd to capture important "chemical effects" despite the fact that for all chosen cutoffs and some functionals the results are not within "chemical accuracy" which is simply due to density functional theory and has nothing to do with basis set convergence. Similar trends were found by the Hutter group in their study of hybrid functionals applied to water simulations.[64] The authors calculated interaction energies of 18.16 kJ/mol for BLYP, 19.37 kJ/mol for B3LYP, 19.08 kJ/mol for PBE, and 19.96 kJ/mol for PBE0. The difference of the latter two values to our values are about 3 kJ/mol and might be attributed to the choice of the pseudopotentials. From our considerations (see the Supporting Information) we advise to compare the convergence behavior of the SVP basis set to a cutoff E_{cut} of 50 Ryd. The TZVP basis set convergence may be compared to E_{cut} = 70 Ryd and the TZVPP to E_{cut} = 90 Ryd. The interaction energies for the global minimum structure calculated with MP2/TZVPP and CCSD(T) in the basis set limit are −19.2 kJ/mol and −20.7 kJ/mol, respectively.[43] A recent and high-level correlated (R_{12} method) thus trustworthy value for the water dimer was provided by Klopper et al. with −21.00 kJ/mol.[65] PBE and PBE0 energies compare best with the CCSD(T) basis set limit values, see the Supporting Information. BLYP provides values that least correspond to the CCSD(T) data. If we compare the stability of the water dimer given by the different functionals, all the PWBS calculations

Table 1. Constraint Force $\langle f \rangle$ in au at Hydrogen Bridge Distance $r_{O^*H^*}$ in pm for Different Functionals and Energy Cutoffs E_{cut} in Ryd

	BLYP				BP86	B3LYP
$r_{O^*H^*}$	20	50	70	90	50	50
100	0.1922	0.1354	0.1394	0.1384	0.1353	0.1349
110	0.1427	0.1017	0.1022	0.1011	0.1001	0.1021
120	0.1014	0.0784	0.0752	0.0766	0.0772	0.0805
130	0.0648	0.0573	0.0555	0.0551	0.0563	0.0587
140	0.0370	0.0421	0.0397	0.0378	0.0404	0.0422
150	0.0193	0.0295	0.0274	0.0255	0.0271	0.0293
160	0.0104	0.0184	0.0167	0.0172	0.0183	0.0191
170	0.0075	0.0115	0.0110	0.0095	0.0127	0.0114
180	0.0072	0.0063	0.0066	0.0050	0.0076	0.0062

as well as GBS results show the same order. The intermolecular bond of the water dimer is strongest using the PBE functional, followed by PBE0, B3LYP, BP86, and BLYP. It is remarkable that the differences in binding energy between $E_{cut} = 150$ Ryd PWBS calculations and the largest GBS calculations are less than 0.5 kJ/mol for all functionals. Again we want to recall that for all functional the difference between 70 Ryd and 150 Ryd is less (<1 kJ/mol) and so is the difference between the plane wave basis set results at 150 Ryd and the TZVPPP Gaussian basis set. Importantly, the deviations provided by different functionals (5 kJ/mol) and thus the difference to the benchmark value of Klopper and co-workers[65] is much larger.

Turning now to the geometry, we find in general that the PWBS geometries do not show trends like distances becoming shorter with a larger basis set. We note that the BP86 geometries independent of the basis set are relatively constant. This is the reason that usually BP86 is preferred over other GGA functionals if reliable structures are sought for. Comparing now TZVPPP and cc-pV5Z with PWBS/150 we find that for all functionals the geometry values agree within 3 pm. The angle varies for the GBS between 2 and 9 degrees. There is one very large angle of 38 degrees for the PBE/SVP combination. The plane wave basis set angles are a bit smaller ranging from 2 to 4 degrees. Again there is no convergence behavior. The difference between PWBSs and GBSs is mostly within 2°.

3.2. Thermodynamic Integration. We now investigate the influence of the basis set quality onto the outcome of the simulations. Table 1 and Figure 3 show the mean absolute constraint values $\langle f \rangle$ for a given distance and at a given E_{cut} obtained from a 6 ps long trajectory.

Please note that the first value for the cutoff of 20 Ryd was obtained at a distance of 103 pm instead of 100 pm. Simulations at 100 pm led to the enforced proton transfer and the subsequent back-transfer of another proton. We recall that at $E_{cut} = 20$ Ryd the dimer geometries show larger distances than at larger E_{cut}. This means that at 20 Ryd the chemically stable ion H_3O^+ can only be formed at larger distances. Obviously, from this it can be deduced that the chemistry of a system is altered by the basis set. At short enforced distances we see the largest deviations between the calculations with different cutoffs for the PWBS. These are situations where bond cleavage of O*−H* occurs, and the proton is transferred to the second water molecule in order to form the Eigen ion.[66]

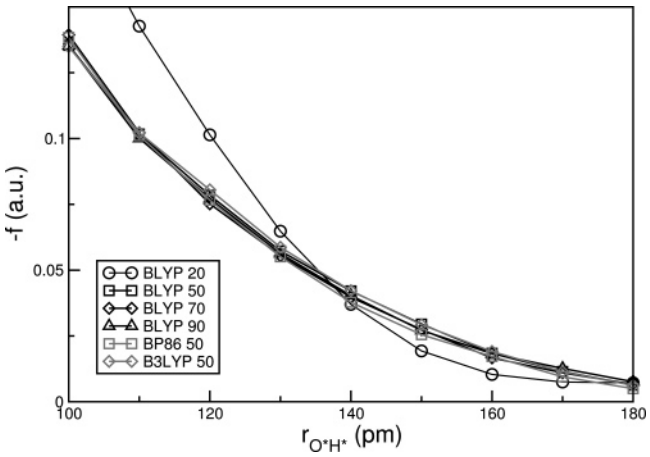

Figure 2. The negative constraint force $-\langle f \rangle$ in au for different energy cutoffs with functional BLYP (black) and different functionals with energy cutoff of 50 Ryd (grey) obtained from PWBS trajectories.

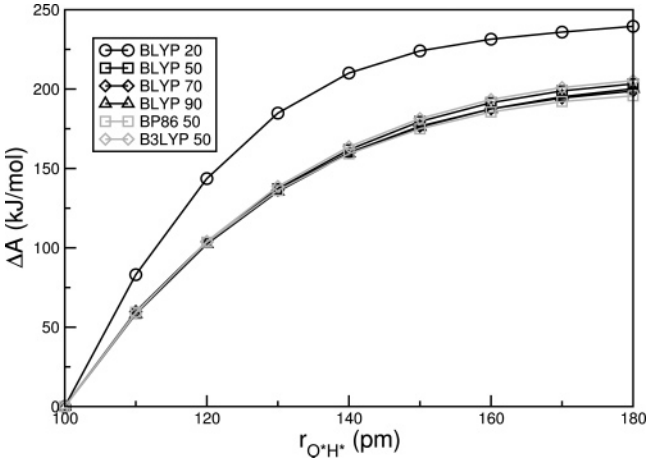

Figure 3. The free energy difference ΔA in kJ/mol of different functionals and different energy cutoffs in Ryd from PWBS molecular dynamics simulations.

Table 2. Free Energy Difference ΔA in kJ/mol for Different Functionals and Different Energy Cutoffs E_{cut} in Ryd

	BLYP				BP86	B3LYP
E_{cut}	20	50	70	90	50	50
ΔA	239.5	203.3	198.8	200.2	195.8	205.4

Surprisingly, BP86 values at a cutoff of 50 Ryd resemble more the BLYP data at 70 Ryd than the BLYP values at 50 Ryd. The integrated free energy difference ΔA is given in Table 2 and in Figure 3. All chosen cutoff and functional results give free energies around 200 kJ/mol, except for the value obtained with a cutoff of 20 which is 40 kJ/mol above the other, see the first entry in Table 2. The data compare well with the single-minimum-path difference of 220 kJ/mol−230 kJ/mol obtained by Sobolewski and Domcke within these distances. GGA functionals are said to underestimate reaction barriers as compared to exact-exchange functionals.[67] Comparing the GGA functional at $E_{cut} = 50$

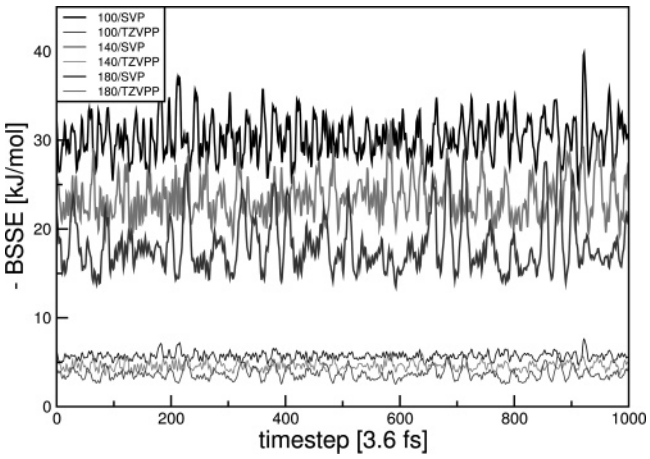

Figure 4. Development of the BSSE with the SVP and the TZVPP basis set along PWBS trajectories of different $r_{O^\star H^\star}$ distance (100 pm, 140 pm, and 180 pm) constraints.

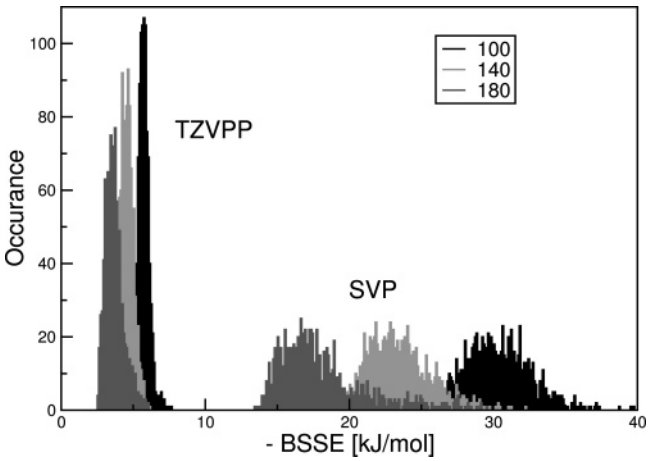

Figure 5. Distribution of the BSSE obtained with the SVP and the TZVPP basis set along PWBS trajectories of different $r_{O^\star H^\star}$ distance (100 pm, 140 pm, and 180 pm) constraints.

Ryd with B3LYP, we find that the B3LYP functional gives the largest value being approximately 10 kJ/mol above the BP86 data and only 2 kJ/mol above the BLYP free energy difference ΔA.

3.3. BSSE along Trajectories. For every trajectory of the 70 Ryd constraint dynamics 1000 snapshots were taken and the BSSE of each snapshot was calculated. The results are presented in Figures 4 and 5.

The BSSE-development along some trajectories of constraint distance ($r_{O^\star H^\star}$=100, 140, and 180 pm) is depicted in Figure 4 for the two basis sets SVP and TZVPP. Obviously the BSSE varies during the course of the simulation. For the SVP basis set along the $r_{O^\star H^\star} = 180$ pm trajectory there are several regions where the BSSE changes rapidly from about 14 kJ/mol to over 29 kJ/mol. It is also interesting that its behavior does not appear like fluctuations around an average of about 21 kJ/mol but resembles more a series of fluctuations around an average considerably smaller than 20 kJ/mol with large outliners to higher values. The same pattern can be observed for the TZVPP basis set but with a more confined range of maximum and minimum values due to the overall smaller average BSSE. Because the BSSE calculations for both basis sets are based on the same trajectory for each distance, the local maxima and minima of the BSSE are found at the same time step. For the BSSE along the $r_{O^\star H^\star} = 100$ pm trajectory large fluctuations are also present; however, these values vary in average around the mean value for the $r_{O^\star H^\star} = 100$ pm trajectory. The pattern for $r_{O^\star H^\star} = 140$ pm trajectory lies between that of the $r_{O^\star H^\star} = 180$ pm and the $r_{O^\star H^\star} = 100$ pm trajectories.

Table 3. Statistical Analysis of the BSSE of the TZVPP and SVP Basis Set for Different $r_{O^\star H^\star}$ (in pm) Distances[a]

	TZVPP				SVP			
$r_{O^\star H^\star}$	⟨⟩	σ	min	max	⟨⟩	σ	min	max
100	5.7	0.41	4.7	7.6	30.3	2.25	24.6	39.8
110	5.5	0.29	4.6	6.7	28.9	1.66	24.5	35.4
120	5.2	0.38	4.3	6.6	27.3	2.11	22.7	34.1
130	4.8	0.43	3.8	6.2	24.9	2.11	20.5	32.5
140	4.6	0.44	3.5	6.1	23.5	2.20	18.9	32.2
150	4.3	0.51	3.1	6.1	21.9	2.66	16.8	30.5
160	4.1	0.46	3.0	6.2	20.4	2.47	16.1	32.0
170	3.8	0.56	2.7	5.7	19.1	2.73	14.5	29.3
180	3.7	0.64	2.5	6.1	18.1	3.02	13.5	29.6

[a] ⟨⟩: mean value of BSSE; σ: standard deviation; min: minimum value of BSSE; max: maximum value of BSSE. All values are in kJ/mol.

In order to investigate the BSSE behavior further, we provide histograms with 0.1 kJ/mol bins of the BSSE data, see Figure 5. Here it is shown more clearly what has been observed before in Figure 4. Obviously the BSSE occurs at larger values for the SVP basis set than for the TZVPP basis set. This is also true for the smaller distances, i.e., for shorter constraint distances the BSSE is larger than for longer constraint values as can be expected also from Figure 1. The fluctuations of the BSSE for the TZVPP basis set is smaller compared to these at the SVP basis set. All distributions are not symmetric with respect to the average value but instead fade out toward higher BSSE errors. This observation is most present for $r_{O^\star H^\star} = 180$ pm and least obvious for $r_{O^\star H^\star} = 100$ pm. It is also apparent that the distribution of different constraint trajectories is closer for the TZVPP series of BSSE than for the SVP basis set values.

In Table 3 we list the maximum, average, and minimum BSSE and its standard deviation σ obtained from the trajectories at all different values of $r_{O^\star H^\star}$. We observe a decrease of the BSSE with increasing $r_{O^\star H^\star}$ from 30 kJ/mol to 18 kJ/mol for the SVP basis set and from 5.7 kJ/mol to 3.7 kJ/mol for the TZVPP basis set. The standard deviation σ being in the 10% range of the mean BSSE value shows the opposite trend, i.e., it rises while the distance enlarges. This nicely corresponds with the behavior reflected in the histogram, see Figure 5, that the distribution fades out for larger distances of $r_{O^\star H^\star}$. The observations from Figure 4 regarding the deviations of local maxima and minima from the assumed average value can also be quantified: For the SVP basis set the difference between the average BSSE value and the minimal BSSE value decreases from 5.7 kJ/mol to 4.6 kJ/mol while changing from $r_{O^\star H^\star} = 100$ pm to $r_{O^\star H^\star} = 180$ pm, whereas the difference between the maximal BSSE value and the mean BSSE value increases from 9.5 kJ/mol to 11.5 kJ/mol. The same trend although with smaller values is observed for the TZVPP basis set. Comparing both basis

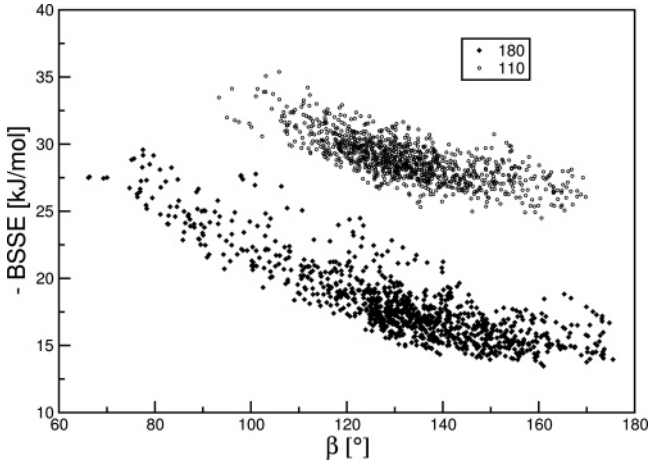

Figure 6. BSSE of the SVP basis set plotted against the angle β depicted in Figure 1 of the Supporting Information obtained from PWBS trajectories at $r_{O^\star H^\star} = 180$ pm and 110 pm.

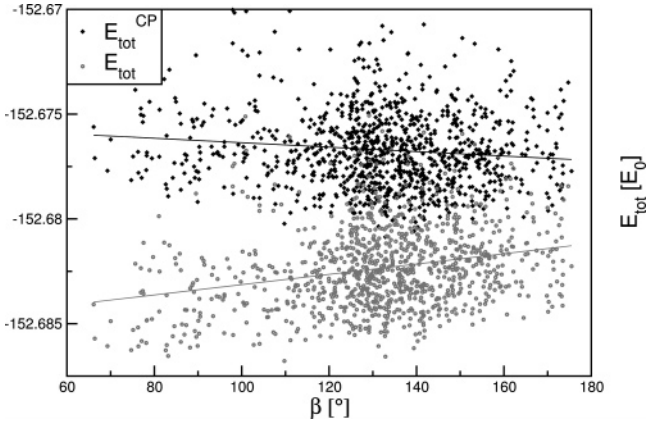

Figure 7. Total energies along the constraint trajectory $r_{O^\star H^\star} = 180$ pm plotted against the angle β depicted: E_{tot}, uncorrected total SVP energy; E_{tot}^{CP}, counterpoise corrected total SVP energy.

sets at a given distance, the average BSSE as well as the standard deviation is considerably smaller for the TZVPP basis set for obvious reasons.

We now select the trajectories with the lowest and highest standard deviation σ, i.e., the trajectories at constraint $r_{O^\star H^\star} = 110$ pm and $r_{O^\star H^\star} = 180$ pm to gain further insight into the unsymmetric distribution of the BSSE. Figure 6 depicts the BSSE plotted against the angle β as defined in Figure 1 of the Supporting Information. β is the angle between the hydrogen bond vector and the bisector of the accepting water molecule. We observe two regions of the BSSE for the $r_{O^\star H^\star} = 180$ pm trajectory, see the filled black diamonds in Figure 6: left and right of $\beta = 145°$. Right of 145° the BSSE behaves almost constant, whereas left of 145° the BSSE increases with smaller values of β. The BSSE of the $r_{O^\star H^\star} = 110$ pm trajectory shows a similar behavior. However, it completely lacks geometries with $\beta < 95°$, which are responsible for most of the high BSSE values at the $r_{O^\star H^\star} = 180$ pm trajectory. This behavior provides a qualitative answer for the question why σ rises with increasing constraint distance: In water dimers with larger values of the distance $r_{O^\star H^\star}$ the hydrogen atoms of the water molecule that accepts the hydrogen bond can still approach the water molecule that donates the hydrogen bond as closely as in dimers of shorter $r_{O^\star H^\star}$. The average BSSE becomes smaller at $r_{O^\star H^\star} = 180$ pm, because the centers of the basis sets of the two molecules are further away from each other than for example in geometries of the $r_{O^\star H^\star} = 110$ pm trajectory. This means that conformations which are sterically unfavorable with regard to β can rather be populated at the $r_{O^\star H^\star} = 180$ pm trajectory, i.e., conformations where the hydrogen atoms of the accepting water are rather close to the donating water. These conformations then show a comparatively high BSSE, because the centers of the basis functions at the hydrogen atoms are closer to the first molecule than in any other conformation of a given $r_{O^\star H^\star}$ distance.

The remaining question is why the distribution of the BSSE only fades out into the direction of large BSSE values. This can be understood with the aid of Figure 6. Minimal values of the BSSE are found for geometries with $\beta > 155°$. For these angles structures are populated independent of the size of distance $r_{O^\star H^\star}$. On the other hand angles of $\beta < 90°$ are only populated for the $r_{O^\star H^\star} = 180$ pm but not for the shorter $r_{O^\star H^\star} = 110$ pm distance geometries This shows why also at favorable $r_{O^\star H^\star}$ distances close to the equilibrium large BSSE values can occur and do occur.

In Figure 7 we plotted the total energies as well as the total energies with subtracted counterpoise corrections against the angle β for the constraint simulation with $r_{O^\star H^\star} = 180$ pm. The corrected values (black dimonds) are higher in the graph than the uncorrected values (gray circles). The straight lines in Figure 7 show linear regressions in order to clarify the tendencies. It is apparent from Figure 7 that configurations with small distances between atoms of different molecules (represented here by small angles β) would be overpopulated since their energy is more favorable than if correct energies would be calculated, i.e., the BSSE is not a simple shift of the potential energy surface.

4. Discussion and Conclusion

We recalled static calculation results of the energies and geometries for the water dimer and investigated the dynamical behavior of the water self-dissociation dependent on the basis set size and BSSE. We can confirm that the plane wave basis set is able to describe the structure as reasonable as a Gaussian basis set provided that a reliable value for the cutoff is chosen. This is also true for the interaction energies. Pulay, who investigated three small peptide molecules, found similar results with larger deviations between the plane wave basis set and Gaussian basis sets only in the dihedral angle.[68] In an assessment of the bulk properties of lithium tetraborate, Islam and co-workers found good results for geometries using the PWBS as compared to GBS.[69] The differences in energetics appeared to be a little more pronounced. More interestingly, Geissler et al. observed in the transition state region of a proton-transfer reaction in $(H_2O)_3H^+$ good agreement between plane wave DFT results and MP2 data using a GBS.[70] The authors compared these combination of methods and basis sets for geometries as well as frequencies for two transition states and the global minimum structure.

Judging from the behavior of the total energies we recommend cutoff 90 Ryd in general for calculations involving hydrogen bonds. The smaller cutoffs of 50 Ryd and 70 Ryd, however, lead to reasonable structures and interaction energies which is important for a discussion of chemical effects. Comparing PWBSs with GBSs the cutoff 50 Ryd for plane waves provides results similar to the SVP basis set, 70 Ryd similar to the TZVP basis set, and 90 Ryd similar to the TZVPP basis set.

Controlling the accuracy of liquid water simulations is still a delicate issue. There are several parameters changeable to improve the simulations which are the electronic structure method, the quality of the basis set, and the usual molecular dynamics approximations, such as the pairwise additivity.[43] Improving just one of these can worsen the previously found results because changing only the method or the basis set can undo the error compensation of fitted approximations for the given system and will then lead away from a realistic behavior of the system. Of course changing all of the variables to amount to more precise results always goes along with additional costs in computer time. This is the reason why compromises with respect to simulation time, system size, and accuracy have to be made. Thereby other errors can be introduced.

Using different cutoffs to obtain the free energy difference by thermodynamic integration, i.e., along several trajectories, shows that a change in the basis set can indeed lead to "different chemical behavior". It was for example found that the proton transfer at cutoff 20 Ryd occurs already at larger distances than if larger cutoff values are chosen. The analysis of the BSSE along several trajectories revealed that the basis set superposition errors might introduce new problems when a proper trajectory should be calculated, because the errors are not of the same size for different chemical (hydrogen-bonding) situations. For the investigated water dimer the distribution of the BSSE is more unsymmetrical, and most importantly the fluctuations are bigger for larger intermolecular distances than for shorter intermolecular distances. Conformations in which atoms of the two different water molecules approach each other closer than they approach each other in the equilibrium geometry or in structures from the attractive region of the potential energy surface show a higher BSSE. In simulations these structures will be overpopulated when the BSSE plays a role, i.e., geometries with smaller average distances or less probable angles will be sampled more often. This would then lead to overstructured liquids which corresponds indirectly to the observations of Lee and Tuckerman.[20,21] Based on similar forces of a GAPW and a PWBS simulation for liquid water it was assumed that the BSSE is mostly a shift of the potential energy surface.[17] For our autoprotolysis reaction of the water dimer the opposite was observed. The question whether the basis set superposition error cancels out in a fully solvated reactions remains open. The amount of the BSSE cannot be estimated a priori, and the obtained trajectory might deviate more from the "true" trajectory when using atom-centered basis functions instead of plane waves. This may result in a stronger violation of the ergodic principle.[34,71]

Acknowledgment. The authors gratefully acknowledge the financial support of the DFG priority program SPP 1191 "Ionic Liquids" and the ERA Chemistry program that allows fruitful collaboration under the project "A Modular Approach to Multi-responsive Surfactant/Peptide (SP) and Surfactant/Peptide/Nanoparticle (SPN) Hybrid Materials". We furthermore like to acknowledge the financial support under the collaborative research center SFB 624 "Templates" at the University of Bonn. Computational time is gratefully acknowledged from the NIC supercomputers in Jülich. J.T. thanks funding by a "Chemiefonds-Stipendium" of the Fonds der Chemischen Industrie.

Supporting Information Available: Plane waves, computational details, static calculations, Tables 1−8, and Figures 1 and 2. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Car, R.; Parrinello, M. *Phys. Rev. Lett.* **1985**, *55*, 2471.

(2) Marx, D.; Hutter, J. Ab Initio Molecular Dynamics: Theory and Implementation. In *Modern Methods and Algorithms of Quantum Chemistry*; Grotendorst, J., Ed.; John von Neumann Institute for Computing, Forschungszentrum Jülich: Jülich, 2000. See http://www.theochem.rub.de/go/cprev.html (accessed May 1, 2007).

(3) Thar, J.; Reckien, W.; Kirchner, B. *Top. Curr. Chem.* **2007**, *268*, 133.

(4) Carloni, P.; Röthlisberger, U.; Parrinello, M. *Acc. Chem. Res.* **2002**, *35*, 455.

(5) Colombo, M. C.; Guidoni, L.; Laio, A.; Magistrato, A.; Maurer, P.; Piana, S.; Röhrig, U.; Spiegel, K.; Sulpizi, M.; Vondele, J. V.; Zumstein, M.; Röthlisberger, U. *Chimia* **2002**, *56*, 13.

(6) Röhrig, U. F.; Guidoni, L.; Laio, A.; Frank, I.; Röthlisberger, U. *J. Am. Chem. Soc.* **2004**, *126*, 15328.

(7) Hutter, J.; Curioni, A. *Chem. Phys. Chem.* **2005**, *6*, 1788.

(8) Szabo, A.; Ostlund, N. S. *Modern Quantum Chemistry;* Dover: New York, U.S.A., 1996; p 480.

(9) Jensen, F. *Introduction to Computational Chemistry;* Wiley-VCH: Chichester, 2002; p 624.

(10) van Duijneveldt, F. B.; van Duijenveldt-van de Rijdt, J. G. C. M.; van Lenthe, J. H. *Chem. Rev.* **1994**, *94*, 1873.

(11) Boys, S. F.; Bernardi, F. *Mol. Phys.* **1970**, *19*, 553.

(12) Lippert, G.; Hutter, J.; Parrinello, M. *Mol. Phys.* **1997**, *92*, 477.

(13) Lippert, G.; Hutter, J.; Parrinello, M. *Theor. Chem. Acc.* **1999**, *103*, 124.

(14) Blöchl, P. E. *Phys. Rev. B* **1994**, *50*, 17953.

(15) Krack, M.; Parrinello, M. *Phys. Chem. Chem. Phys.* **2000**, *2*, 2105.

(16) VandeVondele, J.; Krack, M.; Mohamed, F.; Parrinello, M.; Chassaing, T.; Hutter, J. *Comput. Phys. Commun.* **2005**, *167*, 103.

(17) VandeVondele, J.; Krack, M.; Mohamed, F.; Parrinello, M.; Chassaing, T.; Hutter, J. *Comp. Phys. Comm.* **2005**, *167*, 103.

(18) Schlegel, H. B.; Millam, J. M.; Iyengar, S. S.; Voth, G. A.; A. D. D.; Scuseria, G. E.; Frisch, M. J. *J. Chem. Phys.* **2001**, *114*, 9758.

(19) Liu, Y.; Yarne, D. A.; Tuckerman, M. E. *Phys. Rev. B* **2003**, *68*, 125110.

(20) Lee, H. S.; Tuckerman, M. E. *J. Phys. Chem. A* **2006**, *110*, 5549.

(21) Lee, H. S.; Tuckerman, M. E. *J. Chem. Phys.* **2006**, *125*, 154507.

(22) Helgaker, T.; Jørgensen, P.; Olsen, J. *Molecular Electronic-Structure Theory;* Wiley-VCH: Chichester, 2002; p 938.

(23) Haynes, P. D.; Skylaris, C.-K.; Mostofi, A. A.; Payne, M. C. *Chem. Phys. Lett.* **2006**, *422*, 345.

(24) Bende, A.; Knapp-Mohammady, M.; Suhai, S. *Int. J. Quantum Chem.* **2003**, *92*, 152.

(25) Salvador, P.; Asturiol, D.; Mayer, I. *J. Comput. Chem.* **2006**, *27*, 1505.

(26) Famulari, A.; Gianinetti, E.; Raimondi, M.; Sironi, M. *Int. J. Quantum Chem.* **1998**, *69*, 151.

(27) Famulari, A.; Raimondi, M.; Sironi, M.; Gianinetti, E. *Chem. Phys.* **1998**, *232*, 275.

(28) Kuo, I.-F. W.; Mundy, C. J.; McGrath, M. J.; Siepmann, J. I. *J. Chem. Theory Comput.* **2006**, *2*, 1274.

(29) Ludwig, R. *Angew. Chem., Int. Ed. Engl.* **2006**, *45*, 3402.

(30) Ludwig, R. *Chem. Phys. Chem.* **2007**, *8*, 44.

(31) Ludwig, R.; Paschek, D. *Chem. Unserer Zeit* **2005**, *39*, 164.

(32) Ludwig, R.; Paschek, D. *Angew. Chem., Int. Ed. Engl.* **2001**, *40*, 1809.

(33) Rahman, A.; Stillinger, F. *J. Chem. Phys.* **1971**, *55*, 3336.

(34) Frenkel, D.; Smit, B. *Understanding Molecular Simulation − From Algorithms to Applications*; Academic Press: San Diego, 2002; p 638.

(35) Halgren, T. A.; Damm, W. *Curr. Opin. Struct. Biol.* **2001**, *11*, 236.

(36) Yu, H. B.; Hansson, T.; van Gunsteren, W. *J. Chem. Phys.* **2003**, *118*, 221.

(37) Yu, H. B.; van Gunsteren, W. *J. Chem. Phys.* **2004**, *121*, 9549.

(38) Yu, H. B.; van Gunsteren, W. *Comput. Phys. Commun.* **2005**, *172*, 69.

(39) Weinhold, F. *J. Chem. Phys.* **1998**, *109*, 367.

(40) Weinhold, F. *J. Chem. Phys.* **1998**, *109*, 373.

(41) Ludwig, R.; Weinhold, F. *J. Chem. Phys.* **1999**, *110*, 508.

(42) Ludwig, R.; Weinhold, F. *Z. Phys. Chem.* **2002**, *216*, 659.

(43) Kirchner, B. *J. Chem. Phys.* **2005**, *123*, 204116.

(44) Znamenskiy, V. S.; Green, M. E. *J. Chem. Theory Comput.* **2007**, *3*, 103.

(45) Huber, H.; Dyson, A.; Kirchner, B. *Chem. Soc. Rev.* **1999**, *28*, 121.

(46) Matsuoka, O.; Clementi, E.; Yoshimine, M. *J. Chem. Phys.* **1976**, *64*, 1351.

(47) Lie, G. C.; Clementi, E.; Yoshimine, M. *J. Chem. Phys.* **1976**, *64*, 2314.

(48) Liu, Y. P.; Kim, K.; Berne, B. J.; Friesner, R. A.; Rick, S. W. *J. Chem. Phys.* **1998**, *108*, 4739.

(49) Mas, E. M.; Bukowski, R.; Szalewicz, K.; Groenenboom, G. C.; Wormer, P. E. S.; van der Avoird, A. *J. Chem. Phys.* **2000**, *113*, 6687.

(50) Bukowski, R.; Szalewicz, K.; Groenenboom, G. C.; van der Avoird, A. *J. Chem. Phys.* **2006**, *125*, 044301.

(51) Helgaker, T.; Gauss, J.; Jørgensen, P.; Olsen, J. *J. Chem. Phys.* **1997**, *106*, 6430.

(52) Bak, K. L.; Gauss, J.; Jørgensen, P.; Olsen, J.; Helgaker, T.; Stanton, J. F. *J. Chem. Phys.* **2001**, *114*, 6548.

(53) Pawlowski, F.; Halkier, A.; Jørgensen, P.; Bak, K. L.; Helgaker, T.; Klopper, W. *J. Chem. Phys.* **2003**, *118*, 2539.

(54) Sprik, M.; Ciccotti, G. *J. Chem. Phys.* **1998**, *109*, 7737.

(55) Sobolewski, A. L.; Domcke, W. *Phys. Chem. Chem. Phys.* **2002**, *4*, 4.

(56) Ludwig, R. *Angew. Chem., Int. Ed. Engl.* **2003**, *42*, 258.

(57) Hermansson, K.; Knuts, S.; Lindgren, J. *J. Chem. Phys.* **1991**, *95*, 7486.

(58) Eggenberger, R.; Gerber, S.; Huber, H.; Searles, D.; Welker, M. *J. Chem. Phys.* **1992**, *97*, 5898.

(59) Nosé, S. *J. Chem. Phys.* **1984**, *81*, 511.

(60) Martyna, G. J.; Klein, M. L.; Tuckerman, M. E. *J. Chem. Phys.* **1992**, *97*, 2635.

(61) Hoover, W. G. *Phys. Rev. A* **1985**, *31*, 1695.

(62) CPMD V3.8; Copyright IBM Corp. 1990−2003, Copyright MPI für Festkörperforschung Stuttgart 1997−2001. See also www.cmpd.org (accessed May 1, 2007).

(63) Sprik, M.; Hutter, J.; Parrinello, M. *J. Chem. Phys.* **1996**, *105*, 1142.

(64) Todorova, T.; Seitsonen, A. P.; Hutter, J.; Kuo, I.-F. W.; Mundy, C. J. *J. Phys. Chem. B* **2006**, *110*, 3685.

(65) Klopper, W.; van Duijneveldt- va de Rijdt, J. G. C. M.; van Duijneveldt, F. B. *Phys. Chem. Chem. Phys.* **2000**, *2*, 2227.

(66) Kirchner, B. *ChemPhysChem* **2007**, *8*, 41.

(67) Koch, W.; Holthausen, M. C. *A Chemist's Guide to Density Functional Theory;* Wiley-VCH: Weinheim, 2000; p 528.

(68) Pulay, P.; Saebo, S.; Malagoli, M.; Baker, J. *J. Comput. Chem.* **2005**, *26*, 599.

(69) Islam, M. M.; Maslyuk, V. V.; Bredow, T.; Minot, C. *J. Phys. Chem. B* **2005**, *109*, 13597.

(70) Geissler, P. L.; Van Voorhis, T.; Dellago, C. *Chem. Phys. Lett.* **2000**, *324*, 149.

(71) Allen, M. P.; Tildesley, D. J. *Computer Simulation of Liquids;* Clarendon Press: Oxford, 1987; reprinted 1990; p 408.

CT600259X