Comparison of ab Initio and DFT Electronic Structure Methods for Peptides Containing an Aromatic Ring: Effect of Dispersion and BSSE

Ashley E. Shields and Tanja van Mourik*

School of Chemistry, University of St. Andrews, North Haugh, St. Andrews, Fife, KY16 9ST, Scotland, United Kingdom

Received: August 13, 2007; In Final Form: October 4, 2007

We establish that routine B3LYP and MP2 methods give qualitatively wrong conformations for flexible organic systems containing π systems and that recently developed methods can overcome the known inadequacies of these methods. This is illustrated for a molecule (a conformer of the Tyr-Gly dipeptide) for which B3LYP/6-31+G(d) and MP2/6-31+G(d) geometry optimizations yield strikingly different structures [*Mol. Phys.* **2006**, *104*, 559−570]: MP2 predicts a folded "closed-book" conformer with the glycine residue located above the tyrosine ring, whereas B3LYP predicts a more open conformation. By employing different levels of theory, including the local electron correlation methods LMP2 (local MP2) and LCCSD(T0) (local coupled cluster with single, double, and noniterative local triple excitations) and large basis sets (aug-cc-pVnZ, n = D, T, Q), it is shown that the folded MP2 minimum is an artifact caused by large intramolecular BSSE (basis set superposition error) effects in the MP2/6-31+G(d) calculations. The B3LYP functional gives the correct minimum, but the potential energy apparently rises too steeply when the glycine and tyrosine residues approach each other, presumably due to missing dispersion effects in the B3LYP calculations. The PWB6K and M05-2X functionals, designed to give good results for weak interactions, remedy this to some extent. The reduced BSSE in the LMP2 calculations leads to faster convergence with increasing basis set quality, and accurate results can be obtained with smaller basis sets as compared to canonical MP2. We propose LMP2 as a suitable method to study interactions with π-electron clouds.

1. Introduction

Density functional theory (DFT) has gained immense popularity over the last few decades due to its computational efficiency and similar accuracy compared to other correlated electronic structure methods. One of the most widely used density functionals is the B3LYP[1−3] functional, mainly due to its availability in all common quantum chemistry program packages and the extensive experience of the computational chemistry community in using this functional. The latter means there is vast experience concerning the accuracy that can be expected from calculations using the B3LYP functional. For example, it is generally known that B3LYP as well as many other density functionals are incapable of describing London dispersion interactions.[4−19] Thus, for interactions where dispersion is thought to be important (such as interactions with π-electron clouds) the MP2 (second-order Møller−Plesset perturbation theory) method is often used instead and found to give increased stabilization.[20−22,16] For nondispersive interactions, such as conventional hydrogen bonds (H bonds), B3LYP and MP2 are generally considered to be of comparable accuracy.[23,10]

To identify the most stable conformation of a flexible molecule often involves computation of the relative stability of many different conformers. It is generally thought that relative energies are more sensitive to the level of theory applied than molecular geometries. As such, a common strategy is to perform the geometry optimizations of the molecular conformations of interest with DFT and compute the energies with the more time-consuming MP2 method. This approach is followed even if the molecule contains an aromatic ring and may therefore contain π hydrogen bonds (which usually have a non-negligible dispersion contribution). A number of recent combined experimental/computational studies on gas-phase neurotransmitters and small peptides use this method of DFT geometry optimization followed by MP2 single-point calculation.[24−36] We recently followed the same procedure to compute the relative energies of the 20 conformers of the tyrosine-glycine (Tyr-Gly) dipeptide that were identified to be the most stable by our hierarchical selection scheme.[33] In this study we also optimized a selection of the Tyr-Gly conformers at the MP2 level employing the same moderately sized basis set (6-31+G(d)) as used in the B3LYP geometry optimizations and found surprisingly large geometric differences between the B3LYP- and MP2-optimized structures of some of the conformers. According to the MP2 single-point energies (using the B3LYP geometries), the six most stable Tyr-Gly conformers contain a folded "book" conformation, whereas B3LYP predicted extended conformations to be more stable. MP2 geometry optimization significantly altered the geometry of the six book conformers, increasing their degree of foldedness and stability relative to the extended conformers. The structures of the extended conformers were not notably influenced by the level of optimization. The fourth and sixth most stable conformers on the MP2 single-point energy surface, labeled "book4" and "book6", showed particularly large changes from B3LYP to MP2 geometry optimization (see Figure 1). The B3LYP-optimized structure of book6 is very similar to the corresponding book4 structure; the only difference lies in the orientation of the tyrosine hydroxyl group. Interestingly, MP2 optimization

* To whom correspondence should be addressed. E-mail: tanja.vanmourik@st-andrews.ac.uk.

10.1021/jp076496p CCC: $37.00
Published on Web 11/17/2007

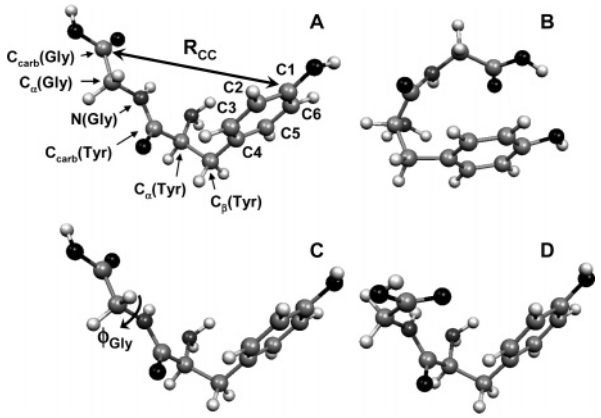

Figure 1. B3LYP/6-31+G(d)- and MP2/6-31+G(d)-optimized geometries of the Tyr-Gly conformers book4 and book6. (A) B3LYP structure of book6. (B) MP2 structure of book6. (C) B3LYP structure of book4. (D) MP2 structure of book4.

of the B3LYP structure of book6 yielded the folded structure shown in Figure 1B, whereas MP2 optimization starting from the B3LYP structure of book4 yielded a structure that mainly differs in the orientation of the C terminus, as characterized by the Ramachandran angle, ϕ_{gly} (Figure 1D).

Two possible explanations for the different book structures predicted by B3LYP and MP2 are missing dispersion interactions in the B3LYP calculations and large intramolecular basis set superposition error (BSSE) effects in the MP2 calculations. Both provide additional attractive forces in the MP2 calculations (though dispersion is a true physical effect, whereas BSSE is an artificial attraction), and therefore, both may be responsible for the more folded structures predicted by MP2. There are two possible remedies for solving the deficiencies in these two methods: including dispersion in the DFT calculations or eliminating/reducing BSSE in the MP2 calculations. We used the latter approach recently to investigate the different book4 structures predicted by B3LYP and MP2.[37] This was accomplished using density-fitting MP2 (df-MP2),[38] which allows use of larger basis sets (thereby reducing BSSE), and local MP2 (LMP2),[39−41,38] which produces much smaller BSSE values even with more limited basis sets.[42−50] These calculations revealed that large intramolecular BSSE values are responsible for hiding the B3LYP minimum on the MP2/6-31+G(d) potential-energy surface, whereas B3LYP misses the MP2 minimum due to lacking dispersion contributions. These results indicated that B3LYP/6-31+G(d) as well as MP2/6-31+G(d) may not give correct structures and energetics for molecules containing interactions with π-electron clouds. This has serious implications for applied computational chemistry, as these levels of theory are routinely used in computational (bio-)organic research.

In the present paper we focus on the book6 conformer, which showed the most dramatic change from B3LYP to MP2 geometry optimization (Figure 1A/B). We employ the same methodologies to reduce BSSE as in our previous paper.[37] In addition, we compute the potential-energy profile for MP2 ↔ B3LYP conformer transition at the CCSD(T) level (facilitated by density-fitting and local approximations). As coupled-cluster theory is considered to be the currently most accurate single-reference wavefunction-based method for calculating dispersion interactions, these calculations provide a reliable reference to compare the lower level results with.

An alternative approach to obtaining reliable transition energy profiles is the use of density functionals that can recover the dispersion contribution. Much effort has been directed in the past decade toward development of functionals that are suitable for noncovalent interactions, and many new functional forms have been proposed. Of these, the functionals developed by Truhlar and co-workers[51−54] have attracted considerable interest. The PBW6K[53] and M05-2X[54] functionals were found to be particularly promising, and we therefore used these two functionals in the current work. PBW6K was reported to outperform MP2 for nonbonded interactions in an assessment against several databases including H-bonded, charge-transfer, $\pi-\pi$ stacking, and weak-interaction databases.[53] The good performance of this functional was confirmed in subsequent studies on stacking and H-bonding interactions,[55,56] although MP2/aug-cc-pVDZ and MP2/aug-cc-pVTZ gave overall better results than PWB6K/6-311++G(2df,2p) for H-bonded and stacked structures of formic acid and formamide tetramers.[56] Several additional studies showed that both PWB6K and M05-2X gave excellent results for noncovalent interactions.[57−59]

The current study confirms the improved performance of the PWB6K and M05-2X functionals as compared to B3LYP. In addition, we show that the folded MP2 minimum is an artifact caused by large intramolecular BSSE effects in the MP2/6-31+G(d) calculations. As such, the B3LYP/6-31+G(d) and MP2/6-31+G(d) levels of theory are both not suitable for conformational studies on molecules containing an aromatic ring. We propose LMP2 as a suitable method to study such molecular systems.

2. Theoretical Calculations

The B3LYP/6-31+G(d) and MP2/6-31+G(d) structures of book6 were taken from ref 33. Potential-energy curves for transition between the B3LYP and MP2 minima were computed by optimizing the Tyr-Gly structure at fixed values of the distance between the C_{carb}(Gly) and C1 atoms (R_{CC}), ranging from 3.0 to 8.0 Å (see Figure 1 for the atom labeling). Initially we performed the partial optimizations at the B3LYP/6-31+G(d) as well as MP2/6-31+G(d) levels of theory. However, the MP2 geometry optimizations for R_{CC} = 5.0/5.5 Å converged to a structure in which the C terminus is reoriented with respect to the B3LYP minimum. This is in essence another conformer, closely related to the MP2-optimized structure of book4 (Figure 1D). The B3LYP-optimized structures, on the other hand, provided a smooth conversion from the MP2 to the B3LYP minimum, and therefore, only the B3LYP-optimized structures were used from this point onward.

The transition energy profiles were subsequently computed by performing single-point calculations at the df-MP2,[38] df-LMP2,[39−41,38] df-SCS-MP2,[60,38] df-SCS-LMP2,[39−41,60,38] and df-LCCSD(T0)[43,61−63] levels of theory using the aug-cc-pVnZ (n = D, T, Q) basis sets.[64,65] In the figures these basis sets are abbreviated as avdz, avtz, and avqz, respectively. The "df" (density fitting) approximation significantly reduces the cost of the two-electron-four-index integrals,[38] thereby allowing the use of much larger basis sets than would be feasible using canonical MP2 or CCSD(T). The "local" approximation also reduces computational cost and in addition decreases the size of the BSSE.[42−44,46] The spin-component-scaled (SCS) correction[60] attempts to correct for the overestimation of dispersion in the MP2 calculations. The default scaling factors (5/6 for antiparallel spins and 1/3 for parallel spins) were employed. The "(T0)" extension[61,62] to LCCSD (local coupled cluster theory with single and double excitations[43,63]) indicates the inclusion of perturbative noniterative local triple excitations. In the local calculations the two most diffuse functions of each angular momentum function were ignored in the localization to yield better-localized orbitals. A completion criterion of 0.99 was employed for the orbital

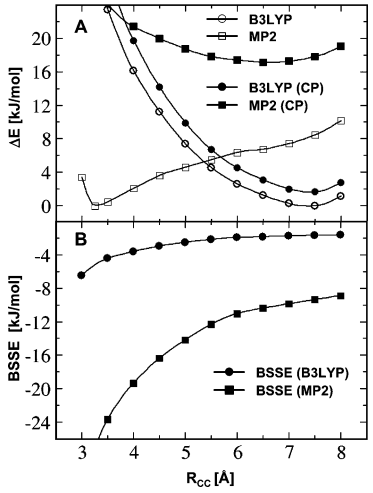

Figure 2. (A) B3LYP/6-31+G(d) and MP2/6-31+G(d) transition energy profiles: (circles) B3LYP results, (squares) MP2 results, (open symbols) uncorrected results, (closed circles) CP-corrected results (using the BSSE values computed for conformationally conforming complexes of *N*-formylglycine···phenol). (B) Variation of the BSSE values as a function of R_{CC}.

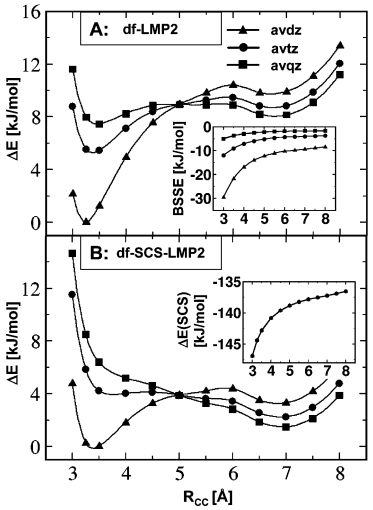

Figure 3. (A) df-MP2 transition energy profiles. The inset shows the variation of the BSSE as a function of R_{CC}. (B) df-SCS-MP2 transition energy profiles. The inset shows the SCS correction, $\Delta E(\mathrm{SCS}) = E(\text{df-MP2}) - E(\text{df-SCS-MP2})$, calculated with aug-cc-pVQZ.

domain selection. These options had been found to give smooth energy profiles for rotation around ϕ_{gly} in book4.[37] The transition energy profiles were also computed with the PWB6K[53] and M05-2X[54] hybrid meta-generalized gradient density functionals, which were developed for thermochemistry and nonbonded interactions.

As in our previous work,[66,37] the *intra*molecular BSSE in the partially optimized B3LYP/6-31+G(d) structures with fixed R_{CC} distances was estimated by *inter*molecular BSSE values in complexes of phenol and *N*-formylglycine with conformations and spatial arrangements identical to the partially optimized structures. For this the $C_\beta(\mathrm{Tyr})H_2$ and $NH_2(\mathrm{Tyr})H$ groups in the Tyr-Gly structures were replaced by hydrogen atoms. The positions of these two hydrogens were optimized for the isolated phenol and *N*-formylglycine fragments at the MP2/6-31+G(d) level of theory. The BSSE in the *N*-formylglycine···phenol complex was then computed using the counterpoise (CP) procedure.[67]

The B3LYP and canonical MP2 calculations were done with Gaussian 03.[68] The B3LYP calculations employed the "ultrafine" integration grid. The df-(SCS)-(L)MP2 and df-LCCSD(T0) calculations were done with Molpro.[69] The PWB6K and M05-2X calculations and optimization of the positions of the added hydrogens in the phenol and *N*-formylglycine fragments were performed with NWChem.[70]

3. Results

3.1. Reducing BSSE by Counterpoise Correction. Figure 2A shows the uncorrected and CP-corrected B3LYP/6-31+G(d) and MP2/6-31+G(d) transition energy profiles. The uncorrected B3LYP curve has a minimum at $R_{CC} = 7.4$ Å (corresponding to the open structure in Figure 1A). The uncorrected MP2 curve shows a sharp minimum at $R_{CC} = 3.3$ Å (corresponding to the folded structure displayed in Figure 1B). Counterpoise correction dramatically alters the MP2 curve: after subtracting the intermolecular BSSE values (calculated for conformationally conforming complexes of phenol and *N*-formylglycine) from the MP2 relative energies, the minimum at 3.4 Å has disappeared and in its place has emerged a shallow minimum at $R_{CC} = \sim 6.5$ Å. In contrast, counterpoise correction hardly affects the B3LYP curve. The main effect appears to be a parallel shift; the shape of the curve and position of the minimum remain nearly unaltered. Figure 2B shows the variation of the BSSE as a function of R_{CC}. The MP2 BSSE is very large and increases steeply at short R_{CC} distances. Presumably, the large BSSE at short distances is the main culprit of the artificial minimum at $R_{CC} = 3.3$ Å in the MP2 profile. In the area close to the B3LYP minimum the BSSE is almost constant, and it therefore does not affect the position of this minimum in the B3LYP profile.

3.2. Reducing BSSE Using Large Basis Sets. The df-MP2 transition energy profiles displayed in Figure 3A show two minima (at ~3.25−3.50 and ~6.5−7.0 Å). The short-distance minimum gradually decreases in depth when larger basis sets are employed but is still visible in the df-MP2/aug-cc-pVQZ profile. The position and depth of the shallow minimum at ~7.0 Å appears to be less dependent on basis set quality. The inset in Figure 3A shows that the BSSE is very large when calculated with the aug-cc-pVDZ basis set and only becomes small when using the aug-cc-pVQZ basis set. Even with this basis set the BSSE amounts to −5 kJ/mol at $R_{CC} = 3.0$ Å, indicating that the df-MP2/aug-cc-VQZ profile may still be distorted by intramolecular BSSE at short R_{CC} distances. Application of the SCS correction makes the short-distance minimum disappear completely when using the aug-cc-pVQZ basis set (Figure 3B). The SCS correction (inset, Figure 3B) shows a remarkably similar variation as the BSSE. Assuming that the empirical SCS correction properly remedies the overcorrection of dispersion in the MP2 calculations, the results suggest that the MP2 minimum at $R_{CC} = \sim 3-3.5$ Å is due to a combination of large BSSE values and overestimated dispersion contributions at short distances.

3.3. Reducing BSSE Using Local MP2. The df-LMP2/aug-cc-pVDZ profile displays two minima: a sharp minimum at $R_{CC} = \sim 4$ Å and a shallow minimum at $R_{CC} = \sim 7$ Å (Figure 4A). The short-distance minimum practically disappears when the larger aug-cc-pVTZ and aug-cc-pVQZ basis sets are used to create the transition energy profiles—with these basis sets the region between 3.5 and 6.0 Å resembles a plateau on the potential-energy surface. After applying the SCS correction the short-distance minimum completely disppears in the aug-cc-pVTZ and aug-cc-pVQZ profiles (Figure 4B).

The BSSE (inset, Figure 4A) is much smaller in the df-LMP2 calculations than in the corresponding nonlocal MP2 calcula-

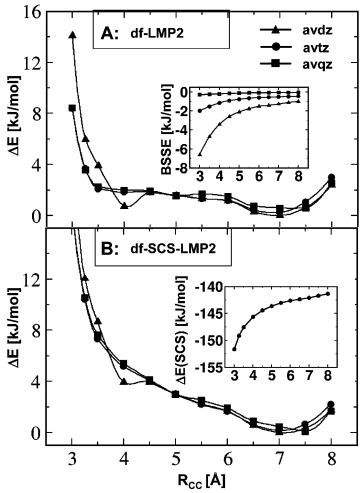

Figure 4. (A) df-LMP2 transition energy profiles. The inset shows the variation of the BSSE as a function of R_{CC}. (B) df-SCS-LMP2 transition energy profiles. The inset shows the SCS correction, $\Delta E(\mathrm{SCS}) = E(\mathrm{df\text{-}LMP2}) - E(\mathrm{df\text{-}SCS\text{-}LMP2})$, calculated with aug-cc-pVQZ.

tions, confirming the common observation[42−50] that the local approximation greatly reduces this error. The df-LMP2/aug-cc-pVTZ and df-LMP2/aug-cc-pVQZ curves are nearly identical, indicating that the results are effectively converged with respect to basis set size. In our study on the Tyr-Gly conformer book4[37] the df-LMP2/aug-cc-pVDZ and df-LMP2/aug-cc-pVTZ curves were almost identical (showing convergence with respect to basis set size), whereas the shape of the nonlocal MP2 profile gradually converged from aug-cc-pVDZ to aug-cc-pVQZ toward the df-LMP2/aug-cc-pVTZ result. From this we concluded that the main effect of basis set improvement in the nonlocal MP2 calculations is a reduction of the BSSE.[37] This is probably also the case in the present study, even though the df-MP2/aug-cc-pVQZ and df-LMP2/aug-cc-pVQZ profiles are not identical (the df-MP2/aug-cc-pVQZ profile shows a shallow minimum at ∼3.5 Å, which is nearly absent in the df-LMP2/aug-cc-pVQZ profile). However, the discrepancy between the two profiles could be due to lingering BSSE effects in the df-MP2 calculations. As shown in the previous section, the BSSE computed at the df-MP2/aug-cc-pVQZ level increases from −1.7 kJ/mol at R_{CC} = 7.5 Å to −5 kJ/mol at 3.0 Å. The df-MP2/aug-cc-VQZ profile may therefore be somewhat distorted by intramolecular BSSE at short R_{CC} distances. Likewise, Mata and Werner related distance-dependent differences between the MP2 and LMP2 energies (ΔE_{loc}) along the reaction coordinate of the $C_2H_5Cl + Cl^-$ reaction to BSSE variation in the MP2 calculations along the reaction coordinate (small average distances between the Cl^- atoms and the ethyl group lead to a large BSSE effect in the canonical MP2 calculations and large ΔE_{loc} values).[50] Comparison of the df-MP2 and df-LMP2 profiles indicates that reducing the BSSE in the calculations greatly decreases the basis set dependence of the Tyr-Gly relative energies. This corresponds with the earlier finding by Jensen that part of the basis set sensitivity of the relative energies of different conformations of a particular molecule, normally considered a basis set effect, should rather be interpreted as intramolecular basis set superposition error.[71]

3.4. Going Beyond MP2: df-LCCSD(T0) Calculations. The df-LCCSD(T0) profiles, computed with the aug-cc-pVDZ and aug-cc-pVTZ basis sets, are displayed in Figure 5. They are surprisingly similar to the corresponding df-LMP2 profiles: both show a broad minimum at 7.0−7.5 Å and a narrow, sharp, minimum at R_{CC} = 4.0 Å in the aug-cc-pVDZ curve; the short-distance minimum disappears on going to the aug-cc-pVTZ basis

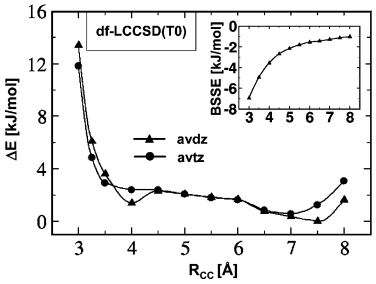

Figure 5. df-LCCSD(T0) transition energy profiles. The inset shows the variation of the BSSE (computed with aug-cc-pVDZ) as a function of R_{CC}.

set level. The df-LMP2 and df-LCCSD(T0) profiles computed with basis sets beyond aug-cc-pVDZ show a plateau in the region from 3.5 to 6.0 Å. In contrast, the SCS-(L)MP2 curves do not exhibit this plateau region. The close similarity of the df-LCCSD(T0) and df-LMP2 profiles seems to indicate that MP2 does *not* overestimate the interaction energy in this case (provided BSSE effects are effectively removed). SCS-MP2 only provides an improvement to MP2 in cases where MP2 tends to overestimate the energy relative to higher level correlation methods like CCSD(T).[72] MP2 is known to overbind stacked complexes such as the benzene dimer,[73] phenol dimer,[74] phenol−methanol complex,[74] and stacked conformations of DNA base pairs.[75] However, MP2 gives binding energies in close agreement with CCSD(T) for the water dimer,[76] whereas MP2 *under*estimates the interaction energy of the rare-gas dimers He_2 and HeNe.[77] For the latter systems the SCS correction leads to deteriorated potential-energy curves.[72]

When overlaying the df-LMP2/aug-cc-pVTZ and df-LCCSD-(T0)/aug-cc-pVTZ curves at R_{CC} = 8.0 Å (the distance with the smallest interaction between the two edges of the peptide), it can be seen that the df-LMP2 curve is increasingly more attractive at shorter distances. This could be a sign of overestimation of the dispersion interaction (which is expected to be larger at shorter distances). However, the differences between the two curves are small (ranging from 0.1 kJ/mol at 7.5 Å to 1.1 kJ/mol at 3.5 Å and 3.4 kJ/mol at 3.0 Å). A similarly small difference (1.7 kJ/mol) between the MP2 and CCSD(T) interaction energy has been observed for the π-bonded indole−H_2O minimum.[20] Also note that in most studies (including the current one) the structures were not re-optimized with CCSD(T), and it is possible that at the CCSD(T)-optimized geometry the CCSD(T) interaction energy is more attractive than the MP2 interaction energy. This was indeed found to be the case for the water dimer.[76] Thus, overestimation of dispersion by the MP2 method appears to be small in the present case, like in the π-bonded indole−water minimum. It may be even smaller than suggested by these results due to the error in the CCSD(T) curve introduced using MP2-optimized geometries. This does not necessarily indicate that the dispersion interaction itself is small in the current molecular system (considering that MP2 *under*-estimates the He_2 interaction, which is entirely caused by dispersion). It is also possible that inaccuracies caused by the local approximation in the df-LMP2 and df-LCCSD(T0) calculations are partially responsible for the (seemingly aberrant) close agreement between the df-LMP2 and df-LCCSD(T0) curves. However, such potential inaccuracies are expected to be similar in the df-LMP2 and df-LCCSD(T0) calculations and are therefore expected not to affect the differences between the df-LMP2 and df-LCCSD(T0) results. We would like to stress that other π-bonded systems (such as the NH−π bonded complexes formamide−benzene, *N*-methylformamide−benzene, and *N*-methylacetamide−benzene) show larger differences

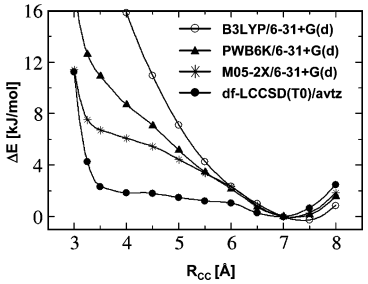

Figure 6. Transition energy profiles computed with PWB6K/6-31+G(d), B3LYP/6-31+G(d), df-SCS-LMP2/aug-cc-pVQZ, and df-LCCSD(T0)/aug-cc-pVTZ.

between the MP2 and CCSD(T) interaction energies,[78] and it is therefore recommended to investigate the potential overestimation of the dispersion interaction by MP2 on a case-by-case basis.

Also note the close similarity between the df-SCS-MP2/aug-cc-pVTZ and df-LCCSD(T0) profiles; in this case the SCS correction appears to yield improved results. This contrasts the above finding, where we concluded that the MP2 method does not notably overestimate the interaction energy of this particular Tyr-Gly conformer, and thus, the SCS correction should yield deteriorated results. However, the improved profile obtained by applying the SCS correction at the df-MP2/aug-cc-pVTZ level is presumably due to lingering BSSE effects. Figure 3 shows that the variation of the SCS correction along the R_{CC} coordinate is very similar to the BSSE variation. Apparently, the improved profile at this level of theory is due to partial cancellation of errors (BSSE vs SCS correction). When the BSSE is reduced (by applying the local approximation), the SCS correction leads to deteriorated potential-energy curves.

To check the adequacy of the noniterative local treatment of the triple excitations in the df-LCCSD(T0) calculations, we also computed the transition energy profile with df-LCCSD(T1), which includes one perturbative update of the triples amplitudes, employing the aug-cc-pVDZ basis set. The df-LCCSD(T0)/aug-cc-pVDZ and df-LCCSD(T1)/aug-cc-pVDZ profiles are nearly identical (Figure S1 in the Supporting Information), demonstrating the accuracy of the T0 triples.

The BSSE in the df-LCCSD(T0)/aug-cc-pVDZ calculations (see inset, Figure 5) is remarkably similar in magnitude and distance variation as the BSSE computed for the df-LMP2/aug-cc-pVDZ energy profile (inset, Figure 4). We therefore expect that the BSSE in the df-LCCSD(T0)/aug-cc-pVTZ calculations (not calculated) is similarly small as that in the df-LMP2/aug-cc-pVTZ calculations. Thus, we expect the df-LCCSD(T0)/aug-cc-pVTZ profile to be nearly BSSE free, and we consider this to be the most reliable profile obtained in the current work.

3.5. PWB6K and M05-2X Density Functionals. Figure 6 shows the transition energy profiles computed with PWB6K/6-31+G(d) and M05-2X/6-31+G(d). For comparison, the profiles computed with B3LYP/6-31+G(d) and df-LCCSD(T0)/aug-cc-pVTZ are shown as well. As mentioned above, the df-LCCSD(T0)/aug-cc-pVTZ curve is probably the most reliable B3LYP ↔ MP2 transition energy profile for the book6 Tyr-Gly conformer computed in this work. The B3LYP/6-31+G(d) curve rises very steeply at short distances. This is presumably due to underestimation of dispersion interactions in the B3LYP calculations, which is more pronounced at short distances. The PBW6K and M05-2X profiles increase less steeply at short distances, indicating that these functionals recover at least some of the dispersion interaction. The M05-2X profile shows a hint of the plateau region seen in the LCCSD(T0) curve. Overall, these two functionals, and particularly M05-2X, yield results closer to the reference LCCSD(T0) values than B3LYP. It should be noted that the DFT profiles have been calculated with a relatively small basis set (6-31+G(d)) and that we did not investigate the basis set dependence of the DFT results.

4. Summary

B3LYP/6-31+G(d) and MP2/6-31+G(d) geometry optimizations yield dramatically different conformations of the Tyr-Gly book6 conformer: MP2/6-31+G(d) predicts a folded "closed-book" conformer with the glycine residue located above the tyrosine ring (distance R_{CC} between the glycine C-terminal carbon atom and the tyrosine-ring C(OH) atom = 3.3 Å), whereas B3LYP predicts a more open conformation (R_{CC} = 7.4 Å).[33] Two possible explanations for the different structures predicted by these two methods are missing dispersion in the B3LYP calculations and large intramolecular BSSE effects in the MP2 calculations. We investigated these effects by computing energy profiles for MP2 ↔ B3LYP conformer transition at different levels of theory. The BSSE in the MP2 calculations was reduced in several ways: (i) by estimating the intramolecular BSSE in Tyr-Gly by intermolecular BSSE values in geometrically conforming complexes of phenol and *N*-formylglycine, (ii) by reducing BSSE using large basis sets, and (iii) by reducing BSSE through the use of LMP2 (local MP2). The transition energy profiles were also computed using the new PWB6K and M05-2X density functionals. The LCCSD(T0) method was used to supply reference values for the transition energy profile.

The results of our study show that the MP2 minimum is almost entirely an artifact of the large intramolecular BSSE in the MP2 calculations. This minimum (almost) disappears in the MP2 profile when larger basis sets are used or when the local approximation is employed, whereas the B3LYP minimum persists when using higher level methods. Thus, for this Tyr-Gly conformer, B3LYP appears to predict the correct minimum-energy structure. This contrasts the situation for the Tyr-Gly book4 conformer, where B3LYP failed to locate two conformer minima along the B3LYP ↔ MP2 transition energy profile.[37] The B3LYP transition energy profile of book6 rises too steeply at short R_{CC} distances, presumably due to missing dispersion in the B3LYP calculations. In addition, the plateau in the region from R_{CC} = 3.5 to 6.0 Å, evident in the df-LCCSD(T0)/aug-cc-pVTZ curve, is absent in the B3LYP profile. The M05-2X functional yields a profile in better agreement with df-LCCSD(T0): the profile increases less steeply at short distances and shows a hint of the plateau region.

The results confirm that neither B3LYP/6-31+G(d) nor MP2/6-31+G(d) is a suitable level of theory to describe interactions with π-electron clouds. One needs to use large basis sets (in this case, aug-cc-pVQZ or beyond!) in the nonlocal MP2 calculations to sufficiently reduce the BSSE. However, the reduced BSSE in the df-LMP2 calculations leads to faster convergence with increasing basis set quality, and accurate results can be obtained with smaller basis sets, though basis sets of the size of aug-cc-pVTZ are still needed to render the BSSE negligible. However, the density fitting and local approximations reduce computation time to such an extent that a df-LMP2/aug-cc-pVTZ calculation is not much more expensive than an MP2/6-31+G(d) calculation. For Tyr-Gly (using one dual-processor dual-core Opteron compute node), a single-point df-LMP2/aug-cc-pVTZ calculation was only two to three times more expensive than a canonical MP2 energy calculation using the much smaller 6-31+G(d) basis set. Thus, LMP2 seems to be a suitable method to study systems with π hydrogen bonds.

Acknowledgment. T.v.M. gratefully acknowledges the Royal Society for their support under the University Research Fellowship scheme and EaStCHEM for computational support via the EaStCHEM Research Computing Facility.

Supporting Information Available: Transition energy profiles computed at the CCSD(T0)/aug-cc-pVDZ and CCSD-(T1)/aug-cc-pVDZ levels of theory (Figure S1); Cartesian coordinates of the B3LYP/6-31+G(d)- and MP2/6-31+G(d)-optimized structures of the TyrGly conformer book6 (Tables S1A-B); Cartesian coordinates of the B3LYP/6-31+G(d)-optimized structures at fixed R_{CC} distances (Tables S2A-L); Total energies of the B3LYP/6-31+G(d)-optimized structures at fixed R_{CC} distances, computed at different levels of theory (Tables S3A-G). This material is available free of charge via the Internet at http://pubs.acs.org.

References and Notes

(1) Becke, A. D. *Phys. Rev. A* **1988**, *38*, 3098.
(2) Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785.
(3) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648.
(4) Kristyán, S.; Pulay, P. *Chem. Phys. Lett.* **1994**, *229*, 175.
(5) Hobza, P.; Šponer, J.; Reschel, T. *J. Comput. Chem.* **1995**, *16*, 1315.
(6) Pérez-Jordá, J. M.; Becke, A. D. *Chem. Phys. Lett.* **1995**, *233*, 134.
(7) Kohn, W.; Meir, Y.; Makarov, D. E. *Phys. Rev. Lett.* **1998**, *80*, 4153.
(8) Tsuzuki, S.; Uchimaru, T.; Tanabe, K. *Chem. Phys. Lett.* **1998**, *287*, 202.
(9) Millet, A.; Korona, T.; Moszynski, R.; Kochanski, E. *J. Chem. Phys.* **1999**, *111*, 7727.
(10) Tuma, C.; Boese, A. D.; Handy, N. C. *Phys. Chem. Chem. Phys.* **1999**, *1*, 3939.
(11) Rappé, A. K.; Bernstein, E. R. *J. Phys. Chem. A* **2000**, *104*, 6117.
(12) Kurita, N.; Sekino, H. *Chem. Phys. Lett.* **2001**, *348*, 139.
(13) Tsuzuki, S.; Lüthi, H. P. *J. Chem. Phys.* **2001**, *114*, 3949.
(14) van Mourik, T.; Gdanitz, R. J. *J. Chem. Phys.* **2002**, *116*, 9620.
(15) Johnson, E. R.; Wolkow, R. A.; DiLabio, G. A. *Chem. Phys. Lett.* **2004**, *394*, 334.
(16) van Mourik, T. *Chem. Phys.* **2004**, *304*, 317.
(17) Crespo-Otero, R.; Montero, L. A.; Stohrer, W.-D.; García de la Vega, J. M. *J. Chem. Phys.* **2005**, *123*, 134107.
(18) Cybulski, S. M.; Seversen, C. E. *J. Chem. Phys.* **2005**, *122*, 014177.
(19) Dabkowska, I.; Jurecka, P.; Hobza, P. *J. Chem. Phys.* **2005**, *122*, 204322.
(20) van Mourik, T.; Price, S. L.; Clary, D. C. *Chem. Phys. Lett.* **2000**, *331*, 253.
(21) Snoek, L. C.; van Mourik, T.; Çarçabal, P.; Simons, J. P. *Phys. Chem. Chem. Phys.* **2003**, *5*, 4519.
(22) van Mourik, T. *Phys. Chem. Chem. Phys.* **2004**, *6*, 2827.
(23) González, L.; Mó, O.; Yáñez, M. *J. Comput. Chem.* **1997**, *18*, 1124.
(24) Graham, R. J.; Kroemer, R. T.; Mons, M.; Robertson, E. G.; Snoek, L. C.; Simons, J. P. *J. Phys. Chem. A* **1999**, *103*, 9706.
(25) Snoek, L. C.; Kroemer, R. T.; Hockridge, M. R.; Simons, J. P. *Phys. Chem. Chem. Phys.* **2001**, *3*, 1819.
(26) MacLeod, N. A.; Simons, J. P. *Chem. Phys.* **2002**, *283*, 221.
(27) Snoek, L. C.; van Mourik, T.; Simons, J. P. *Mol. Phys.* **2003**, *101*, 1239.
(28) Gerhards, M.; Unterberg, C.; Gerlach, A.; Jansen, A. *Phys. Chem. Chem. Phys.* **2004**, *6*, 2682.
(29) MacLeod, N. A.; Simons, J. P. *Phys. Chem. Chem. Phys.* **2004**, *6*, 2878.
(30) Çarçabal, P.; Snoek, L. C.; van Mourik, T. *Mol. Phys.* **2005**, *103*, 1633.
(31) Chin, W.; Dognon, J.-P.; Piuzzi, F.; Tardivel, B.; Dimicoli, I.; Mons, M. *J. Am. Chem. Soc.* **2005**, *127*, 707.
(32) Simons, J. P. *Mol. Phys.* **2006**, *104*, 3317.
(33) Toroz, D.; van Mourik, T. *Mol. Phys.* **2006**, *104*, 559.
(34) Brenner, V.; Piuzzi, F.; Dimicoli, I.; Tardivel, B.; Mons, M. *Angew. Chem., Int. Ed.* **2007**, *46*, 2463.
(35) Toroz, D.; van Mourik, T. *Mol. Phys.* **2007**, *105*, 209.
(36) Vaden, T. D.; de Boer, T. S. J. A.; MacLeod, N. A.; Marzluff, E. M.; Simons, J. P.; Snoek, L. C. *Phys. Chem. Chem. Phys.* **2007**, *9*, 2549.
(37) Holroyd, L. F.; van Mourik, T. *Chem. Phys. Lett.* **2007**, *442*, 42.
(38) Werner, H.-J.; Manby, F. R.; Knowles, P. J. *J. Chem. Phys.* **2003**, *118*, 8149.
(39) Hetzer, G.; Pulay, P.; Werner, H.-J. *Chem. Phys. Lett.* **1998**, *290*, 143.
(40) Schütz, M.; Hetzer, G.; Werner, H.-J. *J. Chem. Phys.* **1999**, *111*, 5691.
(41) Hetzer, G.; Schütz, M.; Stoll, H.; Werner, H.-J. *J. Chem. Phys.* **2000**, *113*, 9443.
(42) Saebø, S.; Tong, W. *J. Chem. Phys.* **1993**, *98*, 2170.
(43) Hampel, C.; Werner, H.-J. *J. Chem. Phys.* **1996**, *104*, 6286.
(44) Pedulla, J. M.; Vila, F.; Jordan, K. D. *J. Chem. Phys.* **1996**, *105*, 11091.
(45) Hartke, B.; Schütz, M.; Werner, H.-J. *Chem. Phys.* **1998**, *239*, 561.
(46) Schütz, M.; Rauhut, G.; Werner, H.-J. *J. Phys. Chem. A* **1998**, *102*, 5997.
(47) Runeberg, N.; Schütz, M.; Werner, H.-J. *J. Chem. Phys.* **1999**, *110*, 7210−7215.
(48) Magnko, L.; Schweizer, M.; Rauhut, G.; Schütz, M.; Stoll, H.; Werner, H.-J. *Phys. Chem. Chem. Phys.* **2002**, *4*, 1006.
(49) Hrenar, T.; Rauhut, G.; Werner, H.-J. *J. Phys. Chem. A* **2006**, *110*, 2060.
(50) Mata, R. A.; Werner, H.-J. *J. Chem. Phys.* **2006**, *125*, 184110.
(51) Zhao, Y.; Schultz, N. E.; Truhlar, D. G. *J. Chem. Phys.* **2004**, *123*, 161103.
(52) Zhao, Y.; Truhlar, D. G. *J. Phys. Chem. A* **2004**, *108*, 6908.
(53) Zhao, Y.; Truhlar, D. G. *J. Phys. Chem. A* **2005**, *109*, 5656.
(54) Zhao, Y.; Schultz, N. E.; Truhlar, D. G. *J. Chem. Theor. Comput.* **2006**, *2*, 364.
(55) Zhao, Y.; Tischenko, O.; Truhlar, D. G. *J. Phys. Chem. B* **2005**, *109*, 19046.
(56) Zhao, Y.; Truhlar, D. G. *J. Phys. Chem. A* **2005**, *109*, 6624.
(57) Zhao, Y.; Truhlar, D. G. *J. Phys. Chem. A* **2006**, *110*, 5121.
(58) Zhao, Y.; Truhlar, D. G. *J. Chem. Theor. Comput.* **2006**, *2*, 1009.
(59) Zhao, Y.; Truhlar, D. G. *J. Chem. Theor. Comput.* **2007**, *3*, 289.
(60) Grimme, F. *J. Chem. Phys.* **2003**, *118*, 9095.
(61) Schütz, M. *J. Chem. Phys.* **2000**, *113*, 9986−10001.
(62) Schütz, M.; Werner, H.-J. *Chem. Phys. Lett.* **2000**, *318*, 370.
(63) Schütz, M.; Werner, H.-J. *J. Chem. Phys.* **2001**, *114*, 661.
(64) Dunning, T. H., Jr. *J. Chem. Phys.* **1989**, *90*, 1007.
(65) Kendall, R. A.; Dunning, T. H., Jr.; Harrison, R. J. *J. Chem. Phys.* **1992**, *96*, 6796.
(66) van Mourik, T.; Karamertzanis, P. G.; Price, S. L. *J. Phys. Chem. A* **2006**, *110*, 8.
(67) Boys, S. F.; Bernardi, F. *Mol. Phys.* **1970**, *19*, 553.
(68) Frisch, M. J. et al. *Gaussian 03*, Revision B.04 ed.; Gaussian Inc.: Pittsburgh, PA, 2003.
(69) 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, D. L.; Deegan, M. J. O.; Dobbyn, A. J.; Eckert, F.; Hampel, C.; Hetzer, G.; Lloyd, A. W.; McNicholas, S. J.; Meyer, W.; Mura, M. E.; Nicklass, A.; Palmieri, P.; Pitzer, R.; Schumann, U.; Stoll, H.; Stone, A. J.; Tarroni, R.; Thorsteinsson, T. *MOLPRO, a package of ab initio programs*, version 2006.1; http://www.molpro.net.
(70) *NWChem*, Version 4.5; High Performance Computational Chemistry Group, Pacific Northwest National Laboratory: Richland, WA, 2003.
(71) Jensen, F. *Chem. Phys. Lett.* **1996**, *261*, 633.
(72) Lochan, R. C.; Jung, Y.; Head-Gordon, M. *J. Phys. Chem. A* **2005**, *109*, 7598.
(73) Sinnokrot, M. O.; Sherrill, C. D. *J. Phys. Chem. A* **2004**, *108*, 10200.
(74) Kolář, M.; Hobza, P. *J. Phys. Chem. A* **2007**, *111*, 5851.
(75) Hobza, P.; Sponer, J. *J. Am. Chem. Soc.* **2002**, *124*, 11802.
(76) Klopper, W.; van Duijneveldt-van de Rijdt, J. G. C. M.; Van, Duijneveldt, F. B. *Phys. Chem. Chem. Phys.* **2000**, *2*, 2227.
(77) van Mourik, T.; Wilson, A. K.; Dunning, T. H., Jr. *Mol. Phys.* **1999**, *96*, 529.
(78) Bendová. L.; Jurečka, P.; Hobza, P.; Vondrášek, J. *J. Chem. Phys. B* **2007**, *111*, 9975.