Aliphatic C−H/π Interactions: Methane−Benzene, Methane−Phenol, and Methane−Indole Complexes

Ashley L. Ringer, Michelle S. Figgs, Mutasem O. Sinnokrot,† and C. David Sherrill*

Center for Computational Molecular Science and Technology, School of Chemistry and Biochemistry, Georgia Institute of Technology, Atlanta, Georgia 30332-0400

Received: May 4, 2006; In Final Form: July 26, 2006

Noncovalent C−H/π interactions are prevalent in biochemistry and are important in molecular recognition. In this work, we present potential energy curves for methane−benzene, methane−phenol, and methane−indole complexes as prototypes for interactions between C−H bonds and the aromatic components of phenylalanine, tyrosine, and tryptophan. Second-order perturbation theory (MP2) is used in conjunction with the aug-cc-pVDZ and aug-cc-pVTZ basis sets to determine the counterpoise-corrected interaction energy for selected complex configurations. Using corrections for higher-order electron correlation determined with coupled-cluster theory through perturbative triples [CCSD(T)] in the aug-cc-pVDZ basis set, we estimate, through an additive approximation, results at the very accurate CCSD(T)/aug-cc-pVTZ level of theory. Symmetry-adapted perturbation theory (SAPT) is employed to determine the physically significant components of the total interaction energy for each complex.

1. Introduction

Noncovalent interactions are prevalent in biochemical molecules and play a role in numerous chemical processes. Of these, the classic hydrogen bond is considered one of the most important, but over the past few decades, evidence has accumulated in support of the significance of a much weaker "hydrogen bond" occurring between an aliphatic C−H group and an aromatic π system.[1,2] This type of noncovalent interaction has been shown to contribute to crystal packing, stereoselectivity, and protein stability and conformation.[3−6] The C−H/π bond also plays a vital role in molecular recognition for numerous ligand-binding proteins.[7,8] Muraki reported that the interaction is common in carbohydrate binding proteins where it affects both binding affinity and conformation.[9] The interaction has already been used in drug design,[10] where it is responsible for an increase in the affinity and selectivity of a thrombin inhibitor[11] and for a significant increase in the inhibitory activity of a tyrosine phosphatase inhibitor.[12] The importance of furthering the understanding of the C−H/π interaction and quantifying its energetics has been recognized.[13]

Analysis of known protein structures has shown the C−H/π interaction frequently occurs between the aliphatic and aromatic groups in protein side chains.[5] In this work, we study the simplest representation of these systems, using methane as a model of aliphatic side chains and benzene, phenol, and indole as the aromatic components of phenylalanine, tyrosine, and tryptophan, respectively. Full potential energy curves are of special interest given that the constrained environments of proteins give rise to individual interactions that may not be in the configurations that would be optimal if the interaction were considered in isolation. In addition to providing insight for drug design and supramolecular chemistry, these high-accuracy computations should be helpful for the calibration of molecular force fields[14] and the development of density functional theories that attempt to accurately model dispersion interactions.[15−28]

The highest-level computations performed previously for the prototype methane−benzene complex were reported by Tsuzuki and co-workers.[29,30] Potential energy curves were computed for six configurations of the complex, and the lowest energy orientation found was one in which the methane is centered on top of the benzene ring and one C−H bond points directly toward the center of the ring. The interaction energy for this configuration was computed using MP2 extrapolated to the complete basis set limit, with additional CCSD(T) correction terms. In recent work Tsuzuki and co-workers[30] determined potential energy curves for the complex using both correlation consistent (cc-pVXZ) and augmented correlation consistent (aug-cc-pVXZ) basis sets. The interaction energies were extrapolated to the complete basis set limit, using both the Helgaker[31] and Feller[32] basis set extrapolation techniques. To our knowledge, similar high-level studies have not been performed for the methane−phenol or methane−indole complexes.

In the present study of methane−benzene, methane−phenol, and methane−indole complexes, results are obtained using MP2 in conjunction with Dunning's augmented correlation-consistent basis sets, aug-cc-pVXZ (X = D, T). In addition, for the methane−benzene complex, we carefully explored basis set effects by using the very large aug-cc-pVQZ basis as well as extrapolation techniques to approximate the complete basis set (CBS) limit. This work expands upon the recent work of Tsuzuki and co-workers[30] for this complex by presenting high-quality aug-cc-pVTZ/aug-cc-pVQZ extrapolations to the CBS limit for the entire potential energy curve. Corrections to the MP2 energies were obtained using the robust CCSD(T) method with the smaller basis sets. Our best estimates should provide binding energies accurate to within a few tenths of a kcal mol^{-1}.

2. Computational Details

Monomer geometries were optimized using second-order perturbation theory (MP2) and the cc-pVDZ basis set, and these

* Corresponding author. E-mail: sherrill@chemistry.gatech.edu.
† Current address: Faculty of Science, Department of Chemistry, University of Jordan, Amman 11942, Jordan.

10.1021/jp062740l CCC: $33.50
Published on Web 08/30/2006

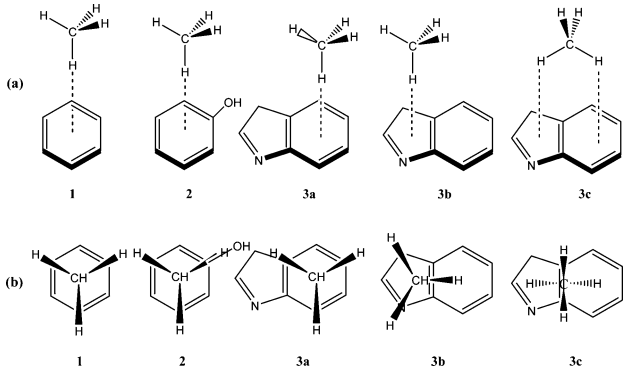

Figure 1. Configurations of methane−benzene, methane−phenol, and methane−indole complexes.

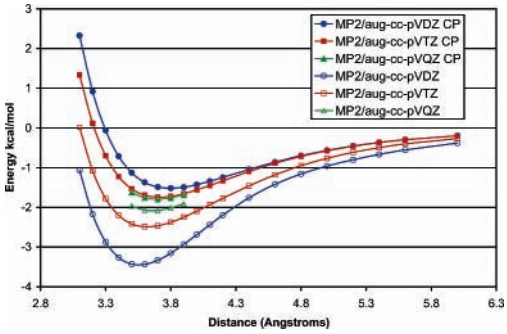

Figure 2. Effect of counterpoise (CP) correction on MP2 potential energy curves for the methane−benzene complex.

frozen monomer geometries were utilized in all computations of the complexes. To verify that the monomer geometry is not significantly changed in the complex, the methane−benzene complex was fully optimized using MP2 and the cc-pVDZ and aug-cc-pVDZ basis sets. No significant geometry changes were found with either basis set; for example, the length of the C−H bond pointing to benzene varied by no more than 0.002 Å and the hydrogens of benzene were bent by only 0.3°. The MP2/cc-pVDZ computational level was also used for single-point energy calculations to select low-energy complex configurations. Although this basis is not sufficient to determine accurate total binding energies (because it lacks diffuse functions), it is adequate to determine which are the low energy configurations.

MP2/aug-cc-pVXZ (where X = D and T) computations were performed for five selected complex configurations, depicted in Figure 1. For these configurations, the interfragment separation distance was varied over at least a 3 Å range using a 0.1 Å stepsize to find the equilibrium distances. CCSD(T) potential curves were determined explicitly using only the aug-cc-pVDZ basis set; the CCSD(T)/aug-cc-pVTZ potential curve was estimated for each complex by calculating a correlation correction term as the difference between the MP2 and CCSD(T) energies determined in the aug-cc-pVDZ basis. This change, denoted ΔCCSD(T), is then added to the MP2/aug-cc-pVTZ results, giving an estimated CCSD(T)/aug-cc-pVTZ interaction energy. This methodology is appropriate because the ΔCCSD(T) correction term is quite insensitive to basis set effects.[33] To further verify the validity of this ΔCCSD(T) addition method, the CCSD(T)/aug-cc-pVTZ interaction energy was explicitly determined for the benzene−methane complex at an equilibrium interfragment separation of 3.8 Å and was in excellent agreement (within 0.01 kcal mol^{-1}) with the estimated value.

Our experience with the benzene dimer[34,35] demonstrates that the interaction energies of noncovalent complexes frequently converge more rapidly when the Boys−Bernardi counterpoise correction[36] is employed. To determine if the counterpoise correction should be employed for C−H/π complexes, both counterpoise-corrected and noncorrected MP2 interaction energies were determined for the methane−benzene complex using the aug-cc-pVDZ, aug-cc-pVTZ, and aug-cc-pVQZ basis sets, as shown in Figure 2. The figure demonstrates that convergence with respect to basis set is greatly accelerated by the counterpoise correction; hence, we apply the counterpoise correction to all results reported here. Optimizations of monomer geometries were performed using Q-Chem 2.1,[37] and energy computations for the complexes were performed using MOLPRO.[38]

Symmetry-adapted perturbation theory (SAPT)[39,40] was applied using the program package SAPT2002[41] to divide the Hartree−Fock (HF) energy and the correlation energy into physically significant components, including electrostatic, induction, dispersion, and exchange energies, plus cross-terms for exchange-induction and exchange-dispersion. We have employed the SAPT2 approach, in which the correlated portion of the interaction energy is nearly equivalent to the supermolecular MP2 correlation energy. To simplify the discussion of the SAPT results, exchange-induction and exchange-dispersion will be counted as induction and dispersion, respectively. The $\delta E^{HF}_{int,resp}$ term, which includes the third-order and higher induction and exchange−induction contributions, is also counted as induction. Because SAPT analysis can be quite time consuming, a less expensive basis set was used to lower the computational cost. This basis set, denoted cc-pVDZ+, is the cc-pVDZ basis for hydrogen and an aug-cc-pVDZ basis minus diffuse d functions for all other atoms; this basis was used previously in our SAPT analysis of the benzene dimer.[35]

3. Results and Discussion

Methane−Benzene Complex. Tsuzuki and co-workers[29] found that for the methane−benzene complex, the preferred configuration has the methane directly above the center of the benzene with one hydrogen pointed at the center of the ring, and three directed away (complex **1** of Figure 1). On the basis of this result, we performed a series of additional computations to determine the effect of rotation of the methane about the axis containing the C atom of methane and the geometric center of benzene. The hydrogens of methane were rotated, in 10° increments, with the distance between methane carbon and the geometric center of benzene fixed at 3.8 Å. The results show less than a 0.001 kcal mol^{-1} variation in the energy. Therefore, the C_{3v} symmetric complex (as depicted as **1** of Figure 1) was selected for higher level analysis because of the greater computational efficiency afforded by its symmetry.

The potential energy curves determined using the MP2/aug-cc-pVDZ, MP2/aug-cc-pVTZ, and MP2/aug-cc-pVQZ, and CCSD(T)/aug-cc-pVDZ levels of theory are depicted in Figure 3. The figure demonstrates that the MP2 results are well converged with respect to the basis set for the aug-cc-pVTZ and aug-cc-pVQZ basis sets. Energies for these two basis sets are then used to extrapolate to the MP2 complete basis set (CBS) limit using the method of Helgaker.[31] This extrapolation procedure was also utilized by Tsuzuki and co-workers[30] with two pairs of basis sets (cc-pVTZ/cc-pVQZ and aug-cc-pVDZ/aug-cc-pVTZ), along with an aug-cc-pVTZ/aug-cc-pVQZ extrapolation for a single optimized geometry. In this work, a complete curve was determined using an aug-cc-pVTZ/aug-cc-pVQZ Helgaker extrapolation and is shown in Figure 3. The ΔCCSD(T) correction shown in Figure 3 is determined by subtracting the CCSD(T)/aug-cc-pVDZ and MP2/aug-cc-pVDZ curves. This correction can then be added to the MP2 results to

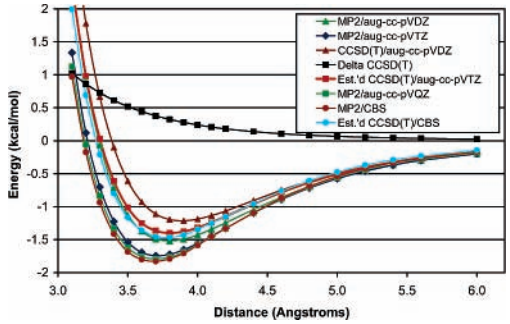

Figure 3. Potential energy curves of the benzene−methane complex.

TABLE 1: Interaction Energies (kcal mol^{-1}) for the Methane−Benzene Complex[a]

method	ΔE_{int}
MP2	
aug-cc-pVDZ	−1.519
aug-cc-pVTZ	−1.723
aug-cc-pVQZ	−1.763
CBS limit	−1.790
CCSD(T)	
aug-cc-pVDZ	−1.195
aug-cc-pVTZ	−1.387
ΔCCSD(T)	
aug-cc-pVDZ	0.324
aug-cc-pVTZ	0.336
Estimated CCSD(T)	
aug-cc-pVTZ	−1.387
aug-cc-pVQZ	−1.400
CBS limit	−1.454

[a] At an interfragment (methane C to the center of the benzene ring) separation of 3.8 Å, the equilibrium distance at the estimated CCSD(T)/CBS level of theory from Figure 3.

provide accurate estimations of the CCSD(T) interaction energy at the same basis set.[35] The ΔCCSD(T) correction decreases with increasing interfragment separation and goes to zero at large interfragment distances.

Results for the methane−benzene complex near equilibrium are presented in Table 1. All the results in this table are for a fixed interfragment separation of 3.8 Å, the equilibrium separation determined using the estimated CCSD(T) values extrapolated to the CBS limit. The MP2 results using the aug-cc-pVTZ (−1.723 kcal mol^{-1}) and aug-cc-pVQZ (−1.763 kcal mol^{-1}) basis sets show that the basis set is nearly converged, and extrapolating to the CBS limit (−1.790 kcal mol^{-1}) only changes the total interaction energy by 0.03 kcal mol^{-1}. These MP2 results are in reasonable agreement with those of Tsuzuki and co-workers,[30] who determined the total interaction energy of the methane−benzene complex as −1.699 kcal mol^{-1} using MP2/aug-cc-pVTZ and −1.759 kcal mol^{-1} using MP2/aug-cc-pVQZ. The small differences in the results are most likely an effect of slightly different geometries for the complex; Tsuzuki and co-workers optimized the complex geometry using the MP2/cc-pVTZ computational level, and the geometry in this work is the equilibrium geometry from our estimated CCSD(T)/CBS potential energy curve. The interaction energy for the complex at an interfragment separation of 3.8 Å was explicitly determined using CCSD(T) for the aug-cc-pVDZ and aug-cc-pVTZ basis sets, and the ΔCCSD(T) correction is shown for both basis sets in Table 1. These results differ by about 0.01 kcal mol^{-1}, confirming that the ΔCCSD(T) correction is insensitive to basis set effects. Adding the aug-cc-pVTZ ΔCCSD(T) correction to the MP2/CBS results gives our best estimate of the total binding

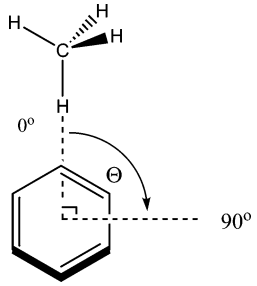

Figure 4. Angular space scanned for methane−benzene complex surface.

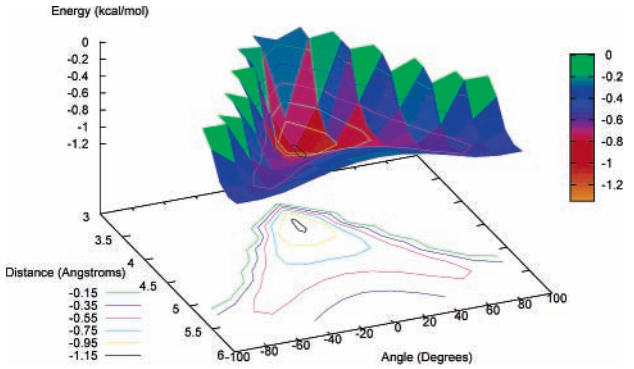

Figure 5. Methane−benzene potential energy surface; energy as a function of the distance between monomers measured from methane carbon to center of benzene and the angle between C−H bond of methane and normal to the benzene ring (see Figure 4).

energy of the complex, −1.454 kcal mol^{-1} at an equilibrium interfragment separation of 3.8 Å.

Thus far, we have only considered a particular slice of the methane−benzene potential surface. To more fully explore the surface, we took the initial complex configuration **1** and varied the angle between the C−H bond and the normal to the aromatic plane of benzene (see Figure 4). In our computations, the original configuration (C−H bond of methane perpendicular to the π system) is denoted 0°, and the configuration in which the C−H bond is in-plane with the aromatic ring is denoted 90°. This angular space was scanned in 15° increments with the interfragment separation held constant at 3.8 Å. At this short interfragment separation, the total interaction energy of the in-plane configuration (relative to benzene and methane at infinite separation) was repulsive by over 50 kcal mol^{-1}; further exploration of this configuration found the most attractive interaction energy for an in-plane configuration at 5.5 Å. The interfragment separation was then varied in 0.1 Å increments from 3.4 to 5.7 Å, for the same angular space. The potential surface is shown in Figure 5.

The surface confirms that, among configurations featuring one hydrogen pointed directly toward the benzene center, the minimum for the methane−benzene complex is the configuration in which the C−H is directly over the aromatic ring. This is reasonable, given that this configuration provides the best access for the partially positive hydrogen to interact with the negative π system. As one moves to longer interfragment separations, the preferred angle changes to one in which the methane is offset from the perpendicular. Even at the equilibrium interfragment separation for offset configurations (40−50°), these complexes are significantly less bound [maximum total CCSD(T)/aug-cc-pVDZ interaction energy is approximately −0.6 kcal mol^{-1}] than the minimum configuration where the C−H bond is perpendicular to the plane of the aromatic ring (−1.20 kcal mol^{-1} at the same level of theory), but they could

still play a stabilizing role in proteins or other complex systems in which the geometry is constrained to nonideal configurations.

Several studies have examined what C−H/π configurations are found in protein and peptide structures by analyzing databases of crystal structures.[6,5] Taking the methane−benzene complex as a model system to describe a general C−H/π interaction, we compared our computed interaction energies to the results of database studies of Brandl et al.[5] and Umezawa et al.[6] In the latter study, the authors examined a set of 130 peptide crystal structures from the Cambridge Structural Database (CSD) that contained a phenylalanine, tyrosine, or tryptophan residue. They counted intra- and intermolecular CH/π contacts separately and tabulated these results according to the distance between the hydrogen of the C−H contact and the nearest carbon atom in the aromatic ring. Considering the intra- and intermolecular contacts together, the greatest number of contacts was found for the 3.02−3.04 Å bin, which corresponds well to the same distance in our minimum methane−benzene complex structure of 3.04 Å. However, beyond this equilibrium distance, the number of contacts falls off very quickly, whereas our results would predict a gradual decrease in the number of contacts because complexes at slightly larger interfragment distances retain a significant interaction energy. We postulate that this discrepancy is due to the constraint of the searching parameters in the study, which would prevent counting of interactions with larger interfragment distances. In the study by Brandl et al.,[5] the authors examined a much larger set (1154) of protein structures from the Protein Data Bank (PDB) for close interactions between C−H-donors and π-acceptors. They defined a parameter d_{C-X} as the distance from the carbon of the C−H system to the center of the aromatic systems (Figure 2 of ref 5), the same parameter varied for our potential surfaces. They also constrain their selection criteria to select configurations above or below the π system, and not in-plane with the π system. This geometric search area corresponds to the well in our potential surface. The distribution of observed C−H/π contacts as a function of the d_{C-X} distance is shown in Figure 3 of ref 5. The maximum frequency was found for d_{C-X} distances of 3.7−3.8 Å depending on the resolution of the data set considered. This is in excellent agreement with the equilibrium distance of 3.8 Å our quantum mechanical results would predict. The frequency of contacts is low (near 0%) for distances shorter than 3.0 Å, distances at which we find positive interaction energies. Between 3.0 Å and the maximum value at 3.7−3.8 Å there is a steady increase in the frequency of contacts, as the predicted interaction energy becomes more attractive. At distances greater than 3.8 Å the frequency of contacts again begins to decrease, corresponding to less bound complexes on our potential energy surface. The qualitative agreement of this distribution with our potential energy surface is very encouraging and suggests that, despite a number of serious complicating factors (solvent effects, steric constraints, secondary interactions, etc.), there may nevertheless be a good correlation between the observed properties of noncovalent interactions in complex systems and the predicted properties of these interactions in small model systems.

Methane−Phenol Complex. The electrostatic potential above the ring in phenol is similar to that of benzene;[42] therefore it seems reasonable to expect that the C−H/π interaction in the methane−phenol complex might have geometric preferences similar to those of the methane−benzene complex. An analogous configuration (complex **2** of Figure 1) was examined, along with two additional configurations, both of which had two hydrogens directed toward the aromatic system. Both of these additional configurations positioned methane over the phenol ring and placed two hydrogens coplanar to the C−O bond of phenol. One configuration centered the methane carbon over the center of the ring, whereas the other configuration was shifted such that the methane carbon was over the substituted carbon of phenol. All three configurations were similar in energy (differences of about 0.1 kcal mol^{-1} at the MP2/cc-pVDZ level of theory), but the one hydrogen down configuration (**2** of Figure 1) was the only configuration chosen for higher-level analysis because it was the lowest in energy and was the most similar to the equilibrium benzene−methane configuration. A configuration similar to that, with methane directly above the center of the ring and with two hydrogens directed down toward benzene, was examined for the benzene−methane complex by Tsuzuki,[29] who also found this configuration slightly higher in energy than the one-hydrogen down configuration, except at short interfragment distances. The effect of rotating the methane over the phenol was examined in the same manner as for the methane−benzene complex, and at a separation distance of 3.8 Å the energy of the complex varied at most 0.007 kcal mol^{-1}. It is interesting to note that although rotational effects were not significant for the structure in which one hydrogen was directed toward the aromatic ring, for the two configurations in which two hydrogens were directed toward the ring, rotational effects were somewhat more pronounced, on the order of 0.2 kcal mol^{-1} at distances of 3.8 Å.

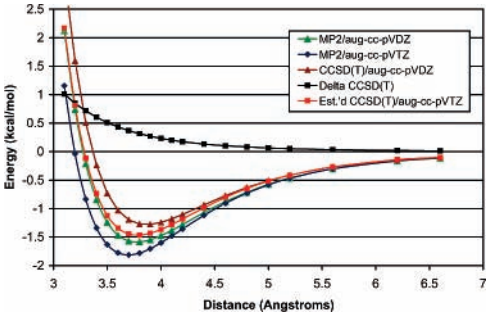

Figure 6. Potential energy curves of the phenol−methane complex.

For the selected one hydrogen down configuration (**2** in Figure 1), potential energy curves and the ΔCCSD(T) curve are illustrated in Figure 6. Our best estimate of the interaction energy is −1.47 kcal mol^{-1} at the estimated CCSD(T)/aug-cc-pVTZ level of theory with an equilibrium interfragment separation of 3.8 Å. These results are very similar to the interaction energy of −1.40 kcal mol^{-1} and interfragment separation of 3.8 Å found for the methane−benzene complex at the same level of theory, indicating that the hydroxyl substituent has only a minor effect. We note that a single hydroxyl group also had a minor effect in sandwich and T-shaped benzene complexes.[42,35]

Methane−Indole Complex. For the methane−indole complex, the two aromatic rings of indole necessitated more exploration of geometric binding preferences for the complex. Nine initial configurations were evaluated: methane centered over the six-membered ring, methane centered over the five-membered ring, and methane centered over the bond shared between the five- and six-membered rings, each with one, two, or three hydrogens directed toward the aromatic centers. Of these configurations, the lowest energy configuration centered the methane over the shared bond of indole with one hydrogen pointing toward the center of each ring (**3c**, Figure 1). This configuration, along with the one hydrogen down configurations centered over the five- (**3b**) and six-membered (**3c**) rings (those

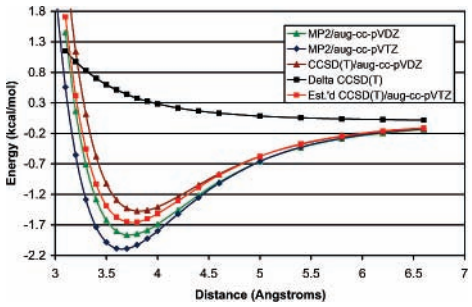

Figure 7. Potential energy curves of the indole−methane complex; configuration **(3a)**: methane centered over the six-membered aromatic system.

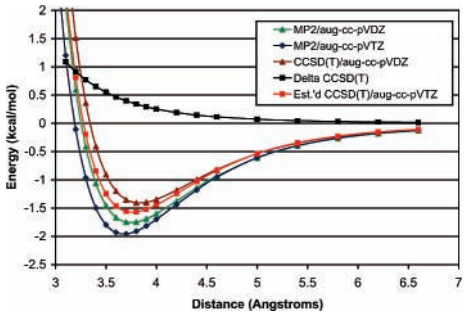

Figure 8. Potential energy curves of the indole−methane complex; configuration **(3b)**: methane centered over the five-membered aromatic system.

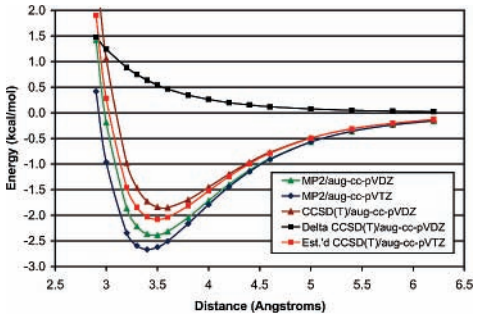

Figure 9. Potential energy curves of the indole−methane complex; configuration **(3c)**: methane is centered over the shared aromatic bond.

most analogous to the minimum configurations for methane−benzene and methane−phenol), was chosen for additional analysis.

The effect of rotating the methane hydrogens around the axis containing the methane carbon and the geometric center of the ring (for configurations **3a** and **3b**) or the axis containing the methane carbon and the center of the shared bond (for configuration **3c**) was considered for these three configurations by the procedure described in previous sections. Configuration **3c** was subject to the most significant rotational effects; rotation of the hydrogens of methane around the axis connecting the methane carbon and the center of the shared bond caused a maximum destabilization of 0.4 kcal mol^{-1}, when the hydrogens facing indole were coplanar with the shared bond. Rotational effects were not significant for either of the one hydrogen down methane−indole configurations (**3a** and **3b**).

The potential energy curves as a function of interfragment distance for these three indole−methane complex configurations (**3a, 3b, 3c**, Figure 1) are shown in Figures 7−9. Our best estimate for the most attractive interaction energy of the indole−methane complex is the estimated CCSD(T)/aug-cc-pVTZ interaction energy for configuration **3c**, −2.08 kcal mol^{-1}, with a separation (methane carbon to shared bond) of 3.5 Å. To examine the extent to which this interaction can be considered a sum of two individual C−H/π interactions, the methane−indole complex was divided into a new methane−benzene configuration and a methane−pyrrole complex. The orientation between the methane and the aromatic compound was fixed at the minimum for the methane−indole complex. At the MP2/aug-cc-pVDZ computational level, the total interaction energy for the methane−benzene complex (at the indole minimum geometry) was −1.08 kcal mol^{-1} and the methane−pyrrole complex was −0.95 kcal mol^{-1}, giving a total of −2.03 kcal mol^{-1}. At the same computational level and geometry, the methane−indole complex has a total interaction energy of −2.38 kcal mol^{-1}, only slightly larger than the sum of the two separate interactions.

TABLE 2: Equilibrium Interfragment Distances and Total Interaction Energies (kcal mol^{-1}) for All Complex Configurations[a]

	MP2/DZ		MP2/TZ		CCSD(T)/DZ		estd CCSD(T)/TZ	
	R^b	ΔE_{int}	R^b	ΔE_{int}	R^b	ΔE_{int}	R^b	ΔE_{int}
methane−benzene **(1)**	3.8	−1.52	3.7	−1.74	3.9	−1.21	3.8	−1.40
methane−phenol **(2)**	3.8	−1.58	3.7	−1.81	3.9	−1.20	3.8	−1.47
methane−indole **(3a)**	3.7	−1.87	3.7	−2.09	3.8	−1.47	3.8	−1.66
methane−indole **(3b)**	3.7	−1.75	3.7	−1.96	3.8	−1.41	3.8	−1.57
methane−indole **(3c)**	3.5	−2.38	3.4	−2.67	3.6	−1.85	3.5	−2.08

[a] Calculations preformed using the aug-cc-pVXZ basis set. [b] Equilibrium interfragment separation (using rigid monomers).

Comparison of Complexes. Table 2 shows the equilibrium interfragment separation for all five complex configurations determined at several computational levels. In all cases, the (counterpoise-corrected) MP2 interaction energies become more attractive as the basis set is improved from double-ζ to triple-ζ. When the MP2/aug-cc-pVDZ energy and the CCSD(T)/aug-cc-pVDZ results are compared, the more complete description of electron correlation predicts the complexes to be less bound (by about 0.3−0.5 kcal mol^{-1}) and have longer interfragment separations (by 0.1 Å). At the estimated CCSD(T)/aug-cc-pVTZ level of theory, the methane−benzene complex is the least bound of all the complexes, with a binding energy of −1.40 kcal mol^{-1}, but the interaction energies for all the configurations which feature one hydrogen down (**1, 2, 3a,** and **3b**) are within 0.20 kcal mol^{-1} of the methane−benzene complex (**1**) at this level of theory. Additionally, all four of these configurations have the same equilibrium interfragment separation of 3.8 Å. For these four complexes, the order of increasing stabilization is **1** < **2** < **3b** < **3a**. At every level of theory considered, the most stabilized complex is the indole−methane complex with one hydrogen directed toward each of the aromatic centers, configuration **3c**.

To provide further insight for the ordering of the configurations, SAPT analysis was performed to divide the total interaction energy into physically significant components. The results of this analysis are presented in Table 3. The similarity of the total interaction energies of the methane−benzene and methane−phenol complexes is reflected in most of the components of the interaction energy. The calculated electrostatic and induction components are almost identical for both complexes, with only slight variances in the exchange and dispersion components. Not surprisingly, the indole−methane complex configuration in which one hydrogen is directed toward the six-membered aromatic system (**3a**) also has very similar electrostatic, exchange, and induction contributions. The 0.34 kcal mol^{-1} difference in its total interaction energy (compared to that for methane−benzene) is primarily due to differing dispersion contributions.

TABLE 3: Physical Components (kcal mol^{-1}) of Total Interaction Energy Determined Using SAPT for All Complex Configurations[a]

	R	elst	exch	ind	disp	SAPT2
methane−benzene **(1)**	3.8	−0.898	2.164	−0.255	−2.025	−1.014
methane−phenol **(2)**	3.8	−0.898	2.144	−0.254	−2.064	−1.072
methane−indole **(3a)**	3.8	−0.893	2.116	−0.291	−2.286	−1.353
methane−indole **(3b)**	3.8	−1.165	2.881	−0.344	−2.614	−1.242
methane−indole **(3c)**	3.5	−1.349	3.221	−0.334	−3.229	−1.692

[a] All data were computed at the cc-pVDZ+ basis using the optimized MP2/aug-cc-pVDZ monomer geometries and the optimum interfragment separation as determined by the estimated CCSD(T)/aug-cc-pVTZ computation.

However, for the methane−indole complex in which one hydrogen is directed toward the five-membered aromatic system (**3b**), the electrostatic contributions are approximately 0.27 kcal mol^{-1} more stabilizing relative to the other one hydrogen down configurations (**1**, **2**, and **3a**). This is accompanied by a small stabilization (0.09 kcal mol^{-1}) in the induction contribution compared to the case for methane−benzene (**1**). Dispersion is more stabilizing by 0.59 kcal mol^{-1}, but this effect is countered by an additional 0.72 kcal mol^{-1} destabilization in the exchange term.

The most stable of all the complexes considered, the **3c** indole−methane complex, has stabilizing electrostatic (1.35 kcal mol^{-1}) and dispersion (3.23 kcal mol^{-1}) terms that are larger than for any of the other complexes. This configuration has a much shorter equilibrium interfragment separation ($R = 3.5$ Å), and shorter separation distances usually lead to more attractive dispersion terms, countered by a larger exchange-repulsion term (in this case 3.22 kcal mol^{-1}, almost completely canceling the dispersion term). The contribution from induction (stabilization of 0.33 kcal mol^{-1}) is similar to that of the other complex configurations considered.

Mulliken population analysis was performed to compare the charge distributions in the methane−benzene complex and in the separated monomers.[43] The SCF wave function determined using the cc-pVDZ basis set was analyzed (using the population analysis program in MOLPRO[38]) for the methane−benzene complex at an interfragment separation of 3.8 Å as well as for the separated molecules at their optimized geometries described above. The most significant difference was found for the charge distribution of methane. For the isolated methane molecule, the hydrogens all had equivalent charges of 0.039 au. However, in the methane−benzene complex, the methane hydrogen directed toward the center of the ring took on a greater positive charge (0.078 au) whereas the other methane hydrogens only had a partial charge of 0.030 au each. These results indicate that the electron distribution in methane polarizes somewhat to reinforce the favorable electrostatic interactions in the complex; this is reflected in the favorable −0.26 kcal mol^{-1} induction term from the SAPT analysis. The population analysis also indicates some transfer of negative electronic charge from methane to benzene, but only a very small amount (0.006 au).

Thus far, the complexes considered have modeled aliphatic C−H/π interactions and have not explored the possibility of aromatic C−H/π contacts, even though these contacts are also prevalent in protein structures.[5] The T-shaped benzene dimer provides a model for such an interaction, in that a hydrogen from the axial benzene interacts with the π cloud of the equatorial benzene. Previous work[33] has determined potential energy surfaces for the T-shaped benzene dimer, using methods similar to those used in this work for the methane−benzene complex. For the T-shaped benzene dimer, the equilibrium C−H/π distance (from the C of the upper benzene to the center of the ring of the lower benzene) is 3.5 Å, and the total interaction determined by adding the MP2/aug-cc-pVTZ energy and a ΔCCSD(T) correction is −2.53 kcal mol^{-1}.[33] In comparison with the methane−benzene complex at the same computational level, the methane−benzene complex has just over half the binding energy, indicating that a T-shaped benzene dimer may not be as simple as a C−H/π interaction. The results of SAPT analysis of these two systems are shown in Figure 10. Because SAPT analysis is quite dependent on interfragment separation, to enable a more direct comparison, both monomers were fixed at the T-shaped benzene dimer C−H/π distance of 3.5 Å. The electrostatic, exchange−repulsion, and induction terms for both systems are similar, within 0.5 kcal mol^{-1}. The electrostatic contribution differs by only 0.3 kcal mol^{-1} whereas the dispersion contributions differ by over 1 kcal mol^{-1}. This suggests that the increased interaction energy of the T-shaped benzene dimer is not primarily caused by the increased acidity of the benzene hydrogen over the methane hydrogen, but rather that an increased dispersion interaction (involving the electrons of the upper π system) and a decreased exchange-repulsion interaction are important in stabilizing the benzene dimer over the methane−benzene complex.

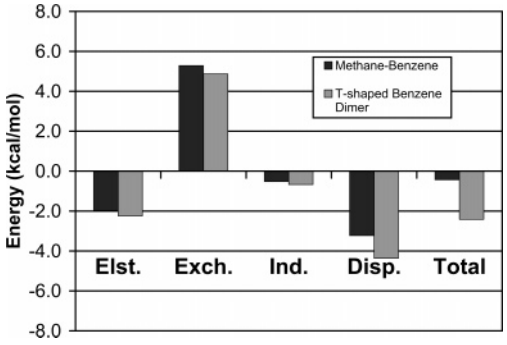

Figure 10. Electrostatic (−1.97, −2.24), exchange−repulsion (5.29, 4.87), induction (−0.53,−0.67), dispersion (−3.22, −4.37), and total interaction energies (−0.43, −2.41) for methane−benzene complex and T-shaped benzene dimer in kcal mol^{-1}. Both systems have a CH/π distance of 3.5 Å.

4. Conclusions

In this work, we have generated high-quality potential energy curves for methane−benzene, methane−phenol, and methane−indole complexes as the simplest prototype noncovalent C−H/π interactions between protein side chains. Curves were generated using MP2 and CCSD(T) in conjunction with the aug-cc-pVDZ basis set. By determining the difference between these two curves, we can capture the effect of higher electron correlation in a correction denoted by ΔCCSD(T). This correction is then applied to the MP2/aug-cc-pVTZ curve, which gives an accurate estimate of the interaction energy at the robust CCSD(T)/aug-cc-pVTZ level of theory.

For the methane−benzene complex, a two-dimensional potential surface was generated at the CCSD(T)/aug-cc-pVDZ computational level that varied both interfragment separation and the angle between the C−H bond of methane and the normal to the plane of benzene. This surface shows that the minimum is found for the configuration in which methane is located directly above the benzene ring. At our best computational level, estimated CCSD(T)/CBS, the interfragment separation (distance from the methane C to the center of the benzene ring) for the minimum configuration is 3.8 Å and the total interaction energy is −1.454 kcal mol^{-1}. As the interfragment separation increases, the preferred angle between the methane carbon and the aromatic

ring changes from directly perpendicular to offset. Comparing these results with those from the database study[5] of Brandl et al., we find a good correlation between the predicted interaction energies of our potential surface and the frequency of C−H/π contacts in crystal structures in the PDB.

The methane−benzene complex is the least bound of the complex configurations considered, but it still lies within 0.20 kcal mol^{-1} of methane−phenol and methane−indole complexes that have similar configurations in which only one hydrogen is directed toward the aromatic system. An indole−methane complex, which features two hydrogens directed toward the aromatic centers, is approximately 0.6 kcal mol^{-1} more stable than the methane−benzene complex. SAPT analysis shows that in complexes where electrostatics are similar (i.e., **1**, **2**, and **3a**), differences in the total interaction energy are caused by differences in the dispersion and induction contributions. SAPT analysis of the methane−benzene complex and the T-shaped benzene dimer indicates that the additional electron density provided by the π system of the upper benzene is important in stabilizing aromatic C−H/π interactions over aliphatic C−H/π interactions.

The high-quality potential energy curves presented here will aid in the analysis of C−H/π interactions in which other steric and geometric constraints prevent equilibrium structures from being attained. This information can also be used to calibrate force fields and to test new density functional theories and other techniques designed to model larger scale systems in which noncovalent interactions are critical.

Acknowledgment. A.L.R. acknowledges an NSF Graduate Research Fellowship. M.S.F. acknowledges an NSF Research Experience for Undergraduates (REU) Fellowship (Grant No. 0139123). C.D.S. gratefully acknowledges an NSF CAREER Award (Grant No. CHE-0094088), and an AC grant from the Petroleum Research Fund of the American Chemical Society (Grant No. 44262-AC6). The Center for Computational Molecular Science and Technology is funded through an NSF CRIF award (CHE 04-43564) and by Georgia Tech.

Supporting Information Available: Cartesian coordinates and potential energy curves of the complexes. This material is available free of charge via the Internet at http://pubs.acs.org.

References and Notes

(1) Nishio, M.; Hirota, M.; Umezawa, Y. *The CH/π Interaction;* Wiley-VCH: New York, 1998.
(2) See also Nishio's website (http://www.tim.hi-ho.ne.jp/dionisio/) for additional references.
(3) Toth, G.; Murphy, R. F.; Lovas, S. *Protein Eng.* **2001**, *14*, 543.
(4) Nishio, M. *Cryst. Eng. Commun.* **2004**, *6*, 130.
(5) Brandl, M.; Weiss, M. S.; Jabs, A.; Suhnel, J.; Hilgenfeld, R. *J. Mol. Biol.* **2001**, *307*, 357.
(6) Umezawa, Y.; Tsuboyama, S.; Takahashi, H.; Nishio, J. U. M. *Bioorg. Med. Chem.* **1999**, *7*, 2021.
(7) Nishio, M.; Umezawa, Y.; Hirota, M.; Takeuchi, Y. *Tetrahedron* **1995**, *51*, 8665.
(8) Shimohigashi, Y.; Nose, T.; Yamauchi, Y.; Maeda, I. *Biopolymers* **1999**, *51*, 9.
(9) Muraki, M. *Protein Peptide Lett.* **2002**, *9*, 195.
(10) Watanabe, T.; Suzuki, T.; Umezawa, Y.; Takeuchi, T.; Otsuka, M.; Umezawa, K. *Tetrahedron* **2000**, *56*, 741.
(11) Obst, U.; Banner, D. W.; Weber, L.; Diederich, F. *Chem. Biol.* **1997**, *4*, 287.
(12) Umezawa, K.; Kawakami, M.; Watanabe, T. *Pharm. Therapy* **2003**, *99*, 15.
(13) Meyer, E. A.; Castellano, R. K.; Diederich, F. *Angew. Chem., Int. Ed. Engl.* **2003**, *42*, 1210.
(14) Macias, A. T.; A. D. MacKerell, J. *J. Comput. Chem.* **2005**, *26*, 1452.
(15) Grimme, S. *J. Comput. Chem.* **2004**, *25*, 1463.
(16) von Lilienfeld, O. A.; Tavernelli, I.; Röthlisberger, U.; Sebastiani, D. *Phys. Rev. Lett.* **2004**, *93*, 153004.
(17) von Lilienfeld, A. O.; Tavernelli, I.; Rothlisberger, U.; Sebastiani, D. *Phys. Rev. B* **2005**, *71*, 195119.
(18) Dion, M.; Rydberg, H.; Schröder, E.; Langreth, D. C.; Lundqvist, B. I. *Phys. Rev. Lett.* **2004**, *92*, 246401.
(19) Podeszwa, R.; Szalewicz, K. *Chem. Phys. Lett.* **2005**, *412*, 488.
(20) Hesselmann, A.; Jansen, G.; Schütz, M. *J. Chem. Phys.* **2005**, *122*, 014103.
(21) Becke, A. D.; Johnson, E. R. *J. Chem. Phys.* **2005**, *123*, 154101.
(22) Zimmerli, U.; Parrinello, M.; Koumoutsakos, P. *J. Chem. Phys.* **2004**, *120*, 2693.
(23) Gonzalez, C.; Lim, E. C. *J. Phys. Chem. A* **2003**, *107*, 10105.
(24) Kumar, A.; Elstner, M.; Suhai, S. *Int. J. Quantum Chem.* **2003**, *95*, 44.
(25) Wu, Q.; Yang, W. *J. Chem. Phys.* **2002**, *116*, 515.
(26) Wu, X.; Vargas, M. C.; Nayak, S.; Lotrich, V.; Scoles, G. *J. Chem. Phys.* **2001**, *115*, 8748.
(27) Elstner, M.; Hobza, P.; Frauenheim, T.; Suhai, S.; Kaxiras, E. *J. Chem. Phys.* **2001**, *114*, 5149.
(28) Zhao, Y.; Truhlar, D. G. *J. Phys. Chem. A* **2005**, *109*, 4209.
(29) Tsuzuki, S.; Honda, K.; Uchimaru, T.; Mikami, M.; Tanabe, K. *J. Am. Chem. Soc.* **2000**, *122*, 3746.
(30) Shibasaki, K.; Fujii, A.; Mikami, N.; Tsuzuki, S. *J. Phys. Chem. A* **2006**, *110*, 4397.
(31) Halkier, A.; Klopper, W.; Helgaker, T.; Jørgensen, P.; Taylor, P. R. *J. Chem. Phys.* **1999**, *111*, 9157.
(32) Feller, D. *J. Chem. Phys.* **1992**, *96*, 6104.
(33) Sinnokrot, M. O.; Sherrill, C. D. *J. Phys. Chem. A* **2004**, *108*, 10200.
(34) Sinnokrot, M. O.; Valeev, E. F.; Sherrill, C. D. *J. Am. Chem. Soc.* **2002**, *124*, 10887.
(35) Sinnokrot, M. O.; Sherrill, C. D. *J. Am. Chem. Soc.* **2004**, *126*, 7690.
(36) Boys, S. F.; Bernardi, F. *Mol. Phys.* **1970**, *19*, 553.
(37) Kong, J.; White, C. A.; Krylov, A. I.; Sherrill, C. D.; Adamson, R. D.; Furlani, T. R.; Lee, M. S.; Lee, A. M.; Gwaltney, S. R.; Adams, T. R.; Daschel, H.; Zhang, W.; Korambath, P. P.; Ochsenfeld, C.; Gilbert, A. T. B.; Kedziora, G. S.; Maurice, D. R.; Nair, N.; Shao, Y.; Besley, N. A.; Maslen, P. E.; Dombroski, J. P.; Baker, J.; Byrd, E. F. C.; Voorhis, T. V.; Oumi, M.; Hirata, S.; Hsu, C.-P.; Ishikawa, N.; Florian, J.; Warshel, A.; Johnson, B. G.; Gill, P. M. W.; Head-Gordon, M.; Pople, J. A. *J. Comput. Chem.* **2000**, *21*, 1532.
(38) MOLPRO, a package of ab initio programs designed by H. J. Werner and P. J. Knowles, version 2002.6. see http://www.molpro.net.
(39) Jeziorski, B.; Moszynski, R.; Szalewicz, K. *Chem. Rev.* **1994**, *94*, 1887.
(40) Williams, H. L.; Szalewicz, K.; Jeziorski, B.; Moszynski, R.; Rybak, S. *J. Chem. Phys.* **1993**, *98*, 1279.
(41) Bukowski, R.; Cencek, W.; Jankowski, P.; Jeziorski, B.; Jeziorska, M.; Kucharski, S. A.; Misquitta, A. J.; Moszynski, R.; Patkowski, K.; Rybak, S.; Szalewicz, K.; Williams, H. L.; Wormer, P. E. S. SAPT2002: An Ab Initio Program for Many-Body Symmetry-Adapted Perturbation Theory Calculations of Intermolecular Interaction Energies. Sequential and Parallel Versions. See: http://www.physics.Udel.edu/~szalewic/SAPT/SAPT.html.
(42) Sinnokrot, M. O.; Sherrill, C. D. *J. Phys. Chem. A* **2003**, *107*, 8377.
(43) Although Mulliken analysis can be problematic (e.g., Mulliken charges can be very sensitive to the level of theory), we believe it should suffice for a general discussion of trends.