*J. Phys. Chem. A* **2005,** *109*, 191−196

# Estimates of the ab Initio Limit for Sulfur−$\pi$ Interactions: The $H_2S$−Benzene Dimer

**Tony P. Tauer, M. Elizabeth Derrick,[†] 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: July 20, 2004; In Final Form: September 30, 2004*

The interaction between aromatic rings and sulfur atoms in the side chains of amino acids is a factor in the formation and stabilization of α-helices in proteins. We studied the $H_2S$−benzene dimer as the simplest possible prototype of sulfur−$\pi$ interactions. High-quality potential energy curves were obtained using coupled-cluster theory with single, double, and perturbative triple substitutions (CCSD(T)) and a large, augmented quadruple-$\zeta$ basis set (aug-cc-pVQZ). The equilibrium intermonomer distance for the hydrogens-down $C_{2v}$ configuration is 3.8 Å with an interaction energy of $-2.74$ kcal mol$^{-1}$. Extrapolating the binding energy to the complete basis set limit gives $-2.81$ kcal mol$^{-1}$. This binding energy is comparable to that of $H_2O$−benzene or of the benzene dimer, and the equilibrium distance is in close agreement with experiment. Other orientations of the dimer were also considered at less complete levels of theory. A considerable reduction in binding for the sulfur-down configuration, together with an energy decomposition analysis, indicates that the attraction in $H_2S$−benzene is best thought of as arising from a favorable electrostatic interaction between partially positive hydrogens in $H_2S$ with the negatively charged $\pi$-cloud of the benzene.

## 1. Introduction

Noncovalent interactions involving the aromatic side chains of certain amino acids are some of the most important factors in determining the dynamics of protein folding. The experimental and computational aspects of $\pi$−$\pi$, cation−$\pi$, alkyl−$\pi$, and amino−$\pi$ interactions have been a subject of much recent interest.[1] One type of interaction that has not received as much attention computationally is the sulfur−$\pi$ interaction, partly because it is not as common as the others in natural systems and partly because the presence of the sulfur atom increases the computational expense.

Morgan et al.[2] first proposed the hypothesis that strong and favorable sulfur−$\pi$ interactions exist after identifying chains of alternating “sulfur and $\pi$-bonded atoms” in the crystal structures of eight different proteins. This finding suggested that sulfur−$\pi$ stacking might play a significant role in stabilizing the folded conformations of these proteins. Database searches performed by Morgan et al.[3] and Reid et al.[4] on the Brookhaven Protein Data Bank,[5] and by Zauhar et al.[6] on the Cambridge Crystallographic Database,[7] all confirmed that sulfur−$\pi$ interactions occurred more frequently than expected from the random association of amino acids.

Viguera and Serrano[8] directly investigated the contribution of S−$\pi$ interactions to the stability of α-helices by calculating the helical content of a model protein from NMR and circular dichroism spectra. The AGADIR[9] algorithm, which calculates the helical content of peptides, was then parametrized in order to reproduce the experimental results; the optimized parameters gave interaction free energies of $-2.0$ kcal mol$^{-1}$ for phenylalanine−cysteine interactions and $-0.65$ kcal mol$^{-1}$ for phenylalanine−methionine.

Cheney et al.[10] performed a quantum mechanical study on the methanethiol−benzene system as a model of cysteine−aromatic interactions. They optimized various initial configurations using Hartree−Fock theory (HF) with the 3-21G* Pople basis set and subsequently performed single-point calculations using second-order Møller−Plesset perturbation theory (MP2) and the 6-31G* basis set. The optimum configuration was found at a distance of 4.4 Å between the sulfur and the center of the benzene ring and an angle of 56° between the line joining these two points and the plane of the benzene ring. The interaction energy for this geometry was computed as $-3.0$ kcal mol$^{-1}$. A more recent study by Duan et al.[11] utilized much larger basis sets, up to 6-311+G(2d,p). Using three different starting geometries, they first optimized the methanethiol−benzene dimer at the MP2/6-31G** level of theory and then performed single-point calculations using the larger basis sets in order to construct potential energy surfaces. Their results show that the equilibrium for the lowest-energy conformation (with the sulfur over the center of the ring) is at 3.73 Å separation, which gave an interaction energy of $-3.71$ kcal mol$^{-1}$. Using their results from a previous study of the methane−benzene dimer, they were able to isolate the contribution of only the SH−$\pi$ interaction, which they said “should be greater than 2.6 kcal mol$^{-1}$ ”. To our knowledge, these MP2/6-311+G(2d,p) calculations are the highest level of theory previously applied to a sulfur−$\pi$ complex. However, our previous experience with weak interactions in the benzene dimer suggests that this level of theory might be far from convergence.[12]

A molecular mechanics study of site-directed mutagenesis in staphylococcal nuclease by Yamaotsu et al.[13] reported a quite shocking result: they found that an M32L substitution (substituting leucine for the methionine at position 32) resulted in a structure that was 1.6 kcal mol$^{-1}$ *more* stable than the wild-type peptide, which is unusual because peptide mutations normally result in less stable protein structures. The mutant protein was subsequently synthesized by Spencer and Stites,[14] who reported a *decrease* in stability of 0.8 kcal mol$^{-1}$ compared to the wild type, a much more conventional result. These results

---

* Corresponding author. E-mail: sherrill@chemistry.gatech.edu.

[†] Visiting faculty from the Department of Chemistry at Valdosta State University, Valdosta, GA.

10.1021/jp046778e CCC: $30.25

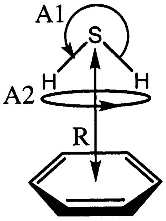

**Figure 1.** Geometry specification for $H_2S$−benzene dimer. The "swing" angle A1 is in the $C_{2v}$ plane of the complex, the "twist" angle A2 is centered on the $C_{2v}$ axis, and the intermonomer distance $R$ is measured from the center of the benzene ring to the sulfur atom. Experimental bond lengths and angles were used as described in the text.

inspired Pranata[15] to perform a theoretical study on the dimethyl sulfide−benzene system using both quantum mechanical (QM) and molecular mechanical (MM) methods. Although his MM results did not agree with Yamaotsu's results using the same force field, they were in good agreement with his MP2/6-31G* QM results, which showed that the M32L substitution was destabilizing.

Here, we present high-level quantum mechanical predictions for the simplest possible prototype of sulfur−$\pi$ interactions, the $H_2S$−benzene dimer. Not only is this system a prototype of sulfur−$\pi$ interactions in biological contexts, but $H_2S$ is a typical oil−gas component, and its interaction with benzene is important in modeling vapor−liquid equilibria relevant to oil and gas processing.[16]

At present, very few high-quality potential energy curves are available for prototype noncovalent interactions. However, such results are crucial for calibrating new methods aimed at modeling these interactions reliably and efficiently. Coupled-cluster theory through perturbative triple substitutions, CCSD(T),[17] is often referred to as the "gold standard of quantum chemistry" and is very reliable for such studies. In validating his new density functional theory plus dispersion model, Grimme[18] has observed that "very accurate CCSD(T) data are still missing" for complexes of benzene with small molecules. Here, we use CCSD(T) with very large basis sets, up through augmented correlation-consistent polarized valence quadruple-$\zeta$ (aug-cc-pVQZ).[19] Note that this augmented basis set, which includes a set of diffuse functions for every angular momentum present in the basis, is much larger than the cc-pVQZ basis set. The potential energy curves thus obtained should be of "subchemical" accuracy (within a few tenths of 1 kcal mol$^{-1}$). Our results for the equilibrium geometry of the complex will be compared to recent microwave experiments by Arunan et al.[20] In addition, the reliability of less complete levels of theory for sulfur−$\pi$ interactions will be evaluated in light of our benchmark results. These comparisons will be valuable in determining appropriate levels of theory for studies of larger models of sulfur−$\pi$ interactions.

## 2. Theoretical Methods

Energy computations using second-order Møller−Plesset perturbation theory (MP2), coupled-cluster theory with single and double substitutions (CCSD), and coupled-cluster including perturbative triple substitutions (CCSD(T)) were performed using various basis sets.[21] Rigid monomer geometries were used, according to the best values in the literature: $r_e$(C−C) = 1.3915 Å and $r_e$(C−H) = 1.0800 Å for benzene,[22] and $r_e$(S−H) = 1.3356 Å and $\theta_e$(H−S−H) = 92.12° for hydrogen sulfide.[23] The monomers were oriented with the sulfur atom directly over the center of the benzene ring, such that the $C_{2v}$ axis of $H_2S$ matches the $C_{6h}$ axis of benzene (Figure 1). Potential energy curves (PECs) for the "swing" angle, A1, and the "twist" angle, A2, were obtained at the CCSD(T)/aug-cc-pVDZ level of theory in order to determine the optimum values of these parameters for later computations. The intermonomer distance $R$ was held fixed at 3.9 Å for these computations.

PECs over the intermonomer distance $R$ were then obtained with the MP2, CCSD, and CCSD(T) methods in conjunction with the 6-31+G*, aug-cc-pVDZ, and aug-cc-pVTZ basis sets. MP2 curves were also obtained with the very large aug-cc-pVQZ basis set (932 functions). Taking advantage of the relative insensitivity to basis set of the difference between CCSD(T) and MP2 energies, we estimate the CCSD(T)/aug-cc-pVQZ energies as follows:

$$E_{\text{int}}^{\text{CCSD(T)/aug-cc-pVQZ}} = E_{\text{int}}^{\text{MP2/aug-cc-pVQZ}} + \delta_{\text{MP2}}^{\text{CCSD(T)}}$$

where

$$\delta_{\text{MP2}}^{\text{CCSD(T)}} = E_{\text{int}}^{\text{CCSD(T)/aug-cc-pVTZ}} - E_{\text{int}}^{\text{MP2/aug-cc-pVTZ}}$$

is calculated from the interaction energies computed with a smaller basis set, in this case, aug-cc-pVTZ.

With the availability of these high-quality results, we decided to assess the reliability of some smaller basis sets which have commonly been used for such calculations. Specifically, we obtained PECs for the 6-31++G** basis set (for comparison to aug-cc-pVDZ) and the 6-311+G(2d,p) basis set (used by Duan,[11] for comparison to aug-cc-pVTZ). We also obtained PECs for three modifications of the aug-cc-pVDZ basis: (1) aug(sp/sp)-cc-pVDZ, with the diffuse d-functions on carbon and sulfur removed; (2) aug(spd/s)-cc-pVDZ, with the diffuse p-function on hydrogen removed; and (3) aug(sp/s)-cc-pVDZ with both the d and p diffuse functions removed. The aug(sp/s)-cc-pVDZ basis set has the same number and types of contracted functions as 6-31++G**, with the only difference being in the number of primitive functions used, thus allowing us to directly compare the inherent quality of the Pople and Dunning basis sets for predictions of energies in van der Waals complexes.

The counterpoise (CP) correction method of Boys and Bernardi[28] was used to account for the basis set superposition error in all computations, since our previous results have shown that CP-corrected energies converge more quickly to the complete basis set limit for $\pi$−$\pi$ interactions.[12] Core orbitals were constrained to remain doubly occupied in all correlated calculations. Calculations were performed in *MOLPRO*[24] running on an IBM SP2 supercomputer.

Symmetry-adapted perturbation theory (SAPT)[25,26] was employed to decompose the energy into physically meaningful components, including electrostatic, induction, dispersion, and exchange energies. Our analysis of the SAPT results follows that of our previous study of substituted benzene dimers.[27] The SAPT calculations reported here used the correlation level technically designated as SAPT2, and they were carried out using the aug-cc-pVDZ basis set at the CCSD(T)/aug-cc-pVQZ geometry. SAPT calculations were performed using the *SAPT2002* program.[29]

## 3. Results and Discussion

The CCSD(T)/aug-cc-pVDZ curves showing the interaction energy as a function of the angles A1 and A2 are shown in Figures 2 and 3. The curve for A1 shows a shallow minimum around 30° from the starting geometry; this angle would have one of the hydrogens pointed almost directly down toward the center of the ring. However, the energy at this point is only

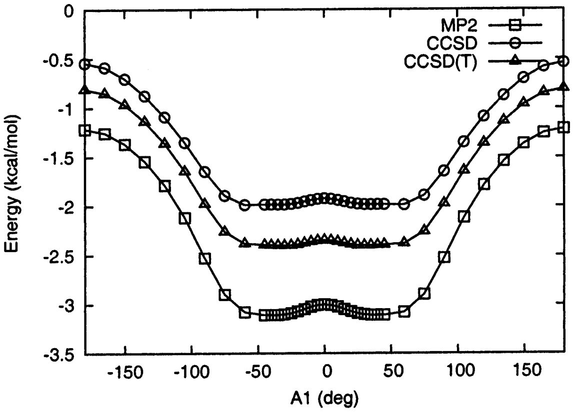

**Figure 2.** Potential energy curve over A1, aug-cc-pVDZ basis set.

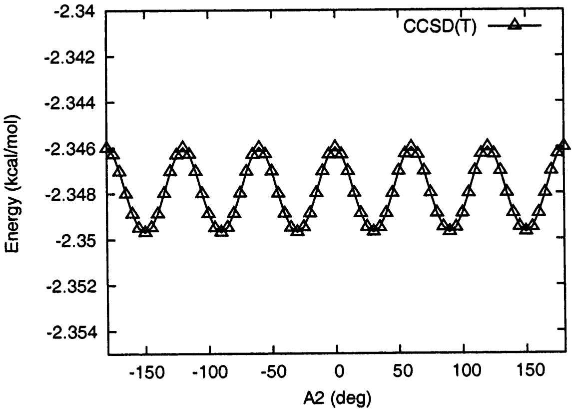

**Figure 3.** Potential energy curve over A2, aug-cc-pVDZ basis set.

0.06 kcal mol$^{-1}$ below the initial $E_{\text{int}}$ of −2.35 kcal mol$^{-1}$ at 0°. This difference is so small that the curve can be considered essentially flat near 0°. At 180°, the sulfur lone pairs are pointed down at the ring and the hydrogens are pointed away; the lone pair electrons interact much less favorably with the negatively charged $\pi$-cloud of the benzene, and the CCSD(T)/aug-cc-pVDZ interaction energy becomes only −0.81 kcal mol$^{-1}$. The corresponding curve for A2 is even flatter, showing very shallow minima (<0.01 kcal mol$^{-1}$) at angles that place the $H_2S$ hydrogens between the ring carbons. Because of this flatness in the potential energy surface for both parameters, and because setting A1 and A2 both equal to 0° gives the system $C_{2v}$ symmetry, we decided to use this geometry in order to reduce the cost of the computations.

The interaction energies as a function of intermonomer distance are shown in Figures 4 and 5. Figure 4 shows the effect of basis set size on the CCSD(T) results; the values obtained for $R_{\text{eq}}$ and $E_{\text{int}}$ are summarized in Table 1. The general trends in $R_{\text{eq}}$ and $E_{\text{int}}$ are readily observable: $R_{\text{eq}}$ decreases and the magnitude of $E_{\text{int}}$ increases ($E_{\text{int}}$ becomes more negative) as the size of the basis set increases. As the basis set becomes larger, the changes to $E_{\text{int}}$ become smaller: between 6-31+G* and aug-cc-pVDZ, $E_{\text{int}}$ increases by 0.8 kcal mol$^{-1}$, while it increases by only 0.3 kcal mol$^{-1}$ between aug-cc-pVDZ and aug-cc-pVTZ, and only 0.1 kcal mol$^{-1}$ between aug-cc-pVTZ and aug-cc-pVQZ. This is as expected, because the correlation-consistent basis sets were designed around the principle of systematically converging the correlation energy correction with increasing basis size.[30]

This convergence can be estimated by correcting for the two main types of basis set error. The first is basis set superposition error, or BSSE, which arises because each monomer in the complex can artificially lower its energy by “borrowing” basis functions from the other monomer, so that the attraction between the two monomers is overestimated; the recommended procedure for eliminating BSSE is the counterpoise correction,[31] which we have applied to all of our results. The second main basis set error is the basis set incompleteness error, or BSIE, which is a consequence of the incomplete description of the electronic Coulomb cusp. In an examination of hydrogen-bonded systems, Halkier and co-workers[32] developed a two-point extrapolation scheme to correct for the BSIE which has the following simple closed form:

$$E_{\text{corr,lim}} = \frac{X^3}{X^3 - (X-1)^3} E_{\text{corr},X} - \frac{(X-1)^3}{X^3 - (X-1)^3} E_{\text{corr},X-1}$$

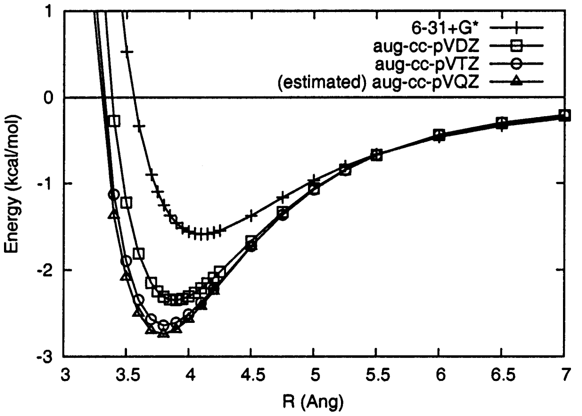

**Figure 4.** Effect of basis set choice for the CCSD(T) method.

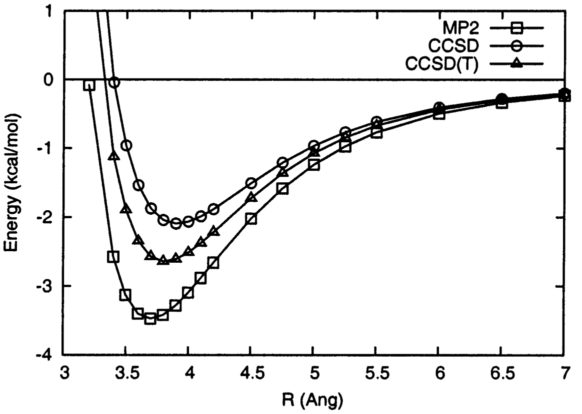

**Figure 5.** Effect of correlation method using the aug-cc-pVTZ basis set.

**TABLE 1: Intermonomer Distance (Å) and Interaction Energy (kcal mol$^{-1}$) at Equilibrium for Various Levels of Theory[a]**

| basis set | method | $R_{\text{eq}}$ | $E_{\text{int}}$ |
|---|---|---|---|
| 6-31+G* | MP2 | 4.00 | −1.92 |
| | CCSD | 4.15 | −1.42 |
| | CCSD(T) | 4.10 | −1.58 |
| aug-cc-pVDZ | MP2 | 3.80 | −3.06 |
| | CCSD | 3.95 | −1.94 |
| | CCSD(T) | 3.90 | −2.34 |
| aug-cc-pVTZ | MP2 | 3.70 | −3.47 |
| | CCSD | 3.90 | −2.09 |
| | CCSD(T) | 3.80 | −2.64 |
| aug-cc-pVQZ | MP2 | 3.70 | −3.60 |
| | CCSD(T) | (3.80)[b] | (−2.74)[b] |
| CBS | CCSD(T) | | −2.81[c] |

[a] All energies include counterpoise corrections. [b] CCSD(T)/aug-cc-pVQZ results are estimated as described in the text. [c] Complete basis set extrapolation at the CCSD(T)/aug-cc-pVQZ geometry.

where $E_{\mathrm{corr},X}$ is the correlation energy obtained with the correlation-consistent basis set with cardinal number $X$ (aug-cc-pVXZ). For the various hydrogen-bonded systems they studied, it was found that a "3−4" MP2 extrapolation (i.e., using the MP2/aug-cc-pVTZ and MP2/aug-cc-pVQZ correlation energies) always gave results within 0.05 kcal mol$^{-1}$ of the MP2−R12 basis set limit. Using the same 3−4 extrapolation here for the CCSD(T) correlation energies (with the CCSD(T)/aug-cc-pVQZ energies estimated as described previously), and taking the CP-corrected SCF/aug-cc-pVQZ energy as our reference, we obtained an extrapolated, complete-basis-set (CBS) CCSD(T) limit $E_{\mathrm{int}}$ of −2.81 kcal mol$^{-1}$. This is an improvement of only 0.07 kcal mol$^{-1}$ over our CCSD(T)/aug-cc-pVQZ results. On the basis of Halkier's results, and the good reliability of CCSD(T) for such problems, it seems certain that this result is within a few tenths of 1 kcal mol$^{-1}$ of the true value.

$R_{\mathrm{eq}}$ and $E_{\mathrm{int}}$ show consistent trends with regard to correlation method, as well. Figure 5 compares the MP2, CCSD, and CCSD(T) potential energy curves with the aug-cc-pVTZ basis set. MP2 binds more strongly than CCSD(T) ($R_{\mathrm{eq}}$ is shorter; $E_{\mathrm{int}}$ is more negative), which binds more strongly than CCSD. This finding is consistent with the results of Hopkins and Tschumper,[33] who found the same trend in their study of various $\pi$-bonded dimers. They also concluded that the effects of triple excitations, included here via the (T) term in CCSD(T), are required in order to determine $E_{\mathrm{int}}$ to chemical accuracy. From the figure, we see that the difference between CCSD(T) and MP2, $\delta_{\mathrm{MP2}}^{\mathrm{CCSD(T)}}$, is largest at short distances and dies off to zero at large distances. This coupled-cluster correction, which was added to the MP2/aug-cc-pVQZ results to estimate the CCSD(T)/aug-cc-pVQZ level of theory, was found to be quite insensitive to the basis set. If we compute this coupled-cluster correction in the smaller aug-cc-pVDZ basis set instead, the largest discrepancy from the aug-cc-pVTZ values is only 0.03−0.04 kcal mol$^{-1}$ at small $R$. This suggests that the errors in $\delta_{\mathrm{MP2}}^{\mathrm{CCSD(T)}}$ computed with the aug-cc-pVTZ basis set are smaller than this.

Our best theoretical results compare very well with the experimental geometry of Arunan et al.[20] Those authors reported geometrical parameters of A1 = 28.5° and $R_{\mathrm{eq}}$ = 3.818 Å; our CCSD(T) calculations showed minima at A1 = 30° (aug-cc-pVDZ basis) and $R_{\mathrm{eq}}$ = 3.80 Å (aug-cc-pVQZ basis set). The deviations from the results of Arunan et al. are well within the resolution of our curves, ±5° for A1 and ±0.1 Å for $R_{\mathrm{eq}}$. Unfortunately, we could not find any reports in the literature of experimental interaction energies for this dimer. We can, however, compare our results to the theoretical results of Duan et al.,[11] who determined that the SH−$\pi$ interaction in methyl sulfide should be ~2.6 kcal mol$^{-1}$ at the MP2/6-311+G(2d,p) level of theory. Their lower-level theoretical binding energy for the methylated model is roughly the same as our higher-level estimated CBS CCSD(T) limit of −2.81 kcal mol$^{-1}$. The binding energy of $H_2S$−benzene is also found to be very similar to that of $H_2O$−benzene, estimated by Tsuzuki et al.[34] as −3.17 kcal mol$^{-1}$ using computational techniques similar to those employed here.

An interesting result from the comparison of basis set effects is the large difference between the 6-31+G* and aug-cc-pVDZ binding energies of 0.76 kcal mol$^{-1}$. Both are double-$\zeta$ basis sets with polarization and diffuse functions, with the exception that 6-31+G* does not include diffuse and polarization functions for hydrogen. To investigate this discrepancy, we performed computations with the 6-31++G** basis set, which does include these functions. We also obtained results with the triple-$\zeta$ 6-311+G(2d,p) basis set, used by Duan et al.,[11] for comparison to the triple-$\zeta$ aug-cc-pVTZ basis. These results are summarized in Figure 6 and Table 2. It is readily apparent that the extra hydrogen functions provided by the 6-31++G** basis set are not particularly important, as they only increased the magnitude of $E_{\mathrm{int}}$ by 0.054 kcal mol$^{-1}$; there is still a large discrepancy (0.713 kcal mol$^{-1}$) between the Pople 6-31++G** and Dunning aug-cc-pVDZ double-$\zeta$ basis sets. The only other difference between the 6-31++G** and aug-cc-pVDZ basis sets is that 6-31++G** only includes diffuse functions for the valence function sets (1s1p/1s), while aug-cc-pVDZ also includes diffuse functions for the polarization sets (1s1p1d)/(1s1p). This led us to wonder whether these diffuse (1d/1p) functions could account for such a large difference, nearly a full kilocalorie per mole.

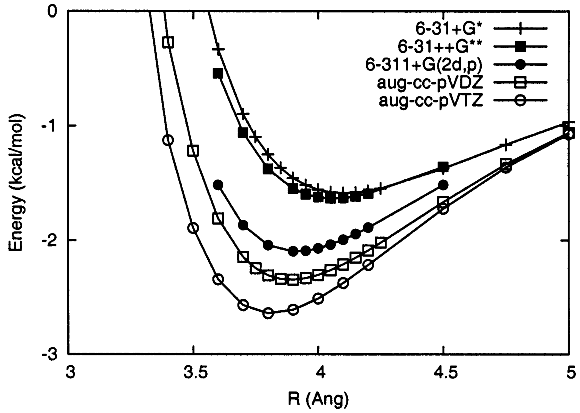

**Figure 6.** Comparison of Pople vs unmodified Dunning basis sets.

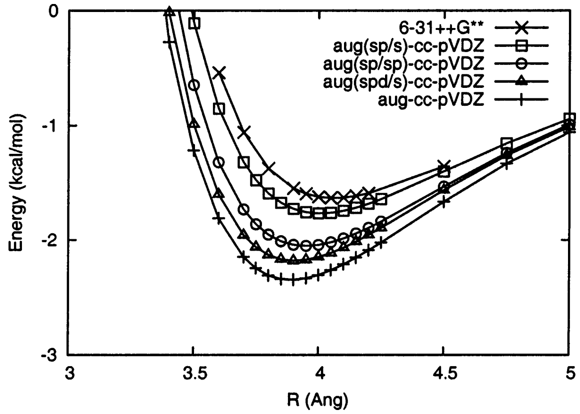

**Figure 7.** Comparison of Pople vs modified Dunning basis sets.

**TABLE 2: Intermonomer Distance (Å) and Interaction Energy (kcal mol$^{-1}$) at Equilibrium; Comparison between Pople and Dunning Basis Sets, CCSD(T) Method**

| basis set | $R_{\mathrm{eq}}$ | $E_{\mathrm{int}}$ |
|---|---|---|
| 6-31+G* | 4.10 | −1.58 |
| 6-31++G** | 4.10 | −1.63 |
| aug-cc-pVDZ | 3.90 | −2.34 |
| 6-311+G(2d,p) | 4.10 | −2.02 |
| aug-cc-pVTZ | 3.80 | −2.64 |

To investigate this possibility, we employed modified versions of the aug-cc-pVDZ basis set, as described in the Methods section. The results are displayed in Figure 7. Removing the diffuse d-functions on carbon and sulfur reduced the binding energy by ~0.30 kcal mol$^{-1}$, almost half of the total difference between the basis sets. Removing the diffuse p-functions on hydrogen had half as great of an effect, reducing $E_{\mathrm{int}}$ by ~0.16 kcal mol$^{-1}$. Removing both sets of functions at the same time reduced $E_{\mathrm{int}}$ by ~0.62 kcal mol$^{-1}$, leaving a difference of only ~0.1 kcal mol$^{-1}$ between 6-31++G** and aug(sp/s)-cc-

**TABLE 3: SAPT2/aug-cc-pVDZ Results for Contributions to the Interaction Energy (kcal mol$^{-1}$) at CCSD(T)/aug-cc-pVQZ Equilibrium Geometry**

| | A1 = 0° | 180° |
|---|---|---|
| $E_{\mathrm{elst}}$ | −2.37 | 0.01 |
| $E_{\mathrm{exch}}$ | 4.19 | 1.03 |
| $E_{\mathrm{ind}}$ | −0.81 | −0.17 |
| $E_{\mathrm{disp}}$ | −4.16 | −2.14 |
| $E_{\mathrm{int}}$(SAPT2) | −3.15 | −1.27 |
| $E_{\mathrm{int}}$(MP2) | −3.06 | −1.21 |

pVDZ. It therefore appears that both the Pople and Dunning basis sets are very similar in fundamental quality, even though they do not use the same number of primitive Gaussians for the contractions of valence orbitals. On the other hand, the extra diffuse functions present in the augmented Dunning basis set make a fairly large contribution to lowering the interaction energy.

A similar discrepancy also appears to exist between the two triple-$\zeta$ basis sets, with a difference in $E_{\mathrm{int}}$ at equilibrium of 0.62 kcal mol$^{-1}$. The difference in the number of basis functions in these two basis sets is greater than the difference in the number of functions in the double-$\zeta$ sets: as compared to aug-cc-pVDZ, aug-cc-pVTZ includes an additional set of (1d1f/1p1d) polarization functions, as well as an additional set of (1f/1d) diffuse functions. Besides the number of valence functions, 6-311+G(2d,p) only differs from 6-31++G** by an additional (1d) polarization function on heavy atoms and the lack of a diffuse (1s) function on hydrogen. The overall difference between 6-311+G(2d,p) and aug-cc-pVTZ is then composed of (1f/1p1d) polarization and (1d1f/1s1p1d) diffuse functions. Even though the difference in the number of functions is greater than that between the double-$\zeta$ basis sets, the magnitude of the difference in energies is slightly smaller; this is consistent with the systematic convergence of energies using the correlation-consistent basis sets. Overall, the higher-angular-momentum diffuse functions in the correlation-consistent basis sets, especially the diffuse d-functions, contribute significantly to the overall interaction energy and should remain important in other van der Waals complexes.

The SAPT-derived components of the binding energy are summarized in Table 3. Although we were only able to perform the SAPT analysis at the SAPT2/aug-cc-pVDZ level of theory, which gives total binding energies very similar to those from counterpoise-corrected MP2/aug-cc-pVDZ, this level of theory features a favorable cancellation of basis set and correlation errors and yields a binding energy similar to that of CCSD(T)/aug-cc-pVQZ. To simplify the analysis, for the present purposes, we have designated the exchange-dispersion and exchange-induction terms as dispersion and induction, respectively. Additionally, the term $\delta E^{\mathrm{HF}}_{\mathrm{int,resp}}$, which includes third- and higher-order HF induction and exchange-induction contributions, has been designated as induction. From the table, we see that electrostatic terms make a fairly strong attractive contribution, −2.37 kcal mol$^{-1}$, arising primarily from the interaction between the partial positive charge on the $H_2S$ hydrogens and the partial negative charge in the benzene $\pi$-cloud. The exchange energy is repulsive (4.19 kcal mol$^{-1}$) and has nearly twice the magnitude of the electrostatic energy. The induction energy is a product of the interaction between each monomer and the static electric field of the other; here, it contributes a modest attractive component (−0.80 kcal mol$^{-1}$) to the binding energy. The dispersion energy is by far the largest attractive component (−4.16 kcal mol$^{-1}$), with nearly twice the magnitude of the electrostatic energy. It is interesting that the magnitude of the dispersion energy is nearly equivalent to the exchange energy, which roughly holds for substituted benzene dimer systems also.[27]

We also performed an SAPT decomposition at the inverted, sulfur-down geometry, A1 = 180°. In this geometry, instead of the electron-deficient hydrogen atoms, the sulfur lone pairs are directed toward the benzene ring. As one might expect, this causes the electrostatic component of the interaction to decrease and even become slightly repulsive. The other three energy components also decrease in magnitude, because the electron density from the sulfur lone pairs does not extend as far from the sulfur as the electron density associated with the hydrogens in $H_2S$. This might be anticipated from simple valence-shell electron-pair repulsion (VSEPR) considerations, which would suggest that the very small H−S−H bond angle of 92 ° would imply a large angle between the sulfur lone pairs. We note that the exchange-repulsion is reduced in magnitude much more than the dispersion interaction, so that the sum of exchange-repulsion and dispersion is now somewhat attractive (−1.11 kcal mol$^{-1}$) rather than almost zero as in the hydrogens-down A1 = 0° configuration. However, the reduction in the electrostatic term outweighs this effect, so that, overall, the sulfur-down configuration is 1.88 kcal mol$^{-1}$ less favorable than the hydrogens-down configuration at the SAPT2/aug-cc-pVDZ level of theory (1.54 kcal mol$^{-1}$ less favorable for CCSD(T)/aug-cc-pVDZ). On the basis of these considerations, the sulfur−$\pi$ interaction, at least in this model system, is best thought of as being primarily an electrostatic attraction between the $H_2S$ hydrogens and the aromatic $\pi$-cloud.

## 4. Conclusions

In this study, we examined the $H_2S$−benzene dimer as the simplest model of sulfur−$\pi$ interactions. Calculations using several basis sets and different levels of electron correlation were performed to obtain potential energy curves for the intermonomer geometric parameters A1, A2, and $R$. Estimates of the CCSD(T)/aug-cc-pVQZ potential energy curves presented here for the $C_{2v}$ configuration represent a great leap forward in the reliability of theoretical data for this system, and they should be suitable as benchmarks for the calibration of new theoretical methods for noncovalent interactions. The results at our highest levels of theory, A1 = 30° for CCSD(T)/aug-cc-pVDZ, $R_{\mathrm{eq}}$ = 3.80 Å, and $E_{\mathrm{int}}$ = −2.74 kcal mol$^{-1}$ for CCSD(T)/aug-cc-pVQZ, are in good agreement with previous experimental and lower-level theoretical results. Complete basis set extrapolations yield a CCSD(T) interaction energy of −2.81 kcal mol$^{-1}$, which is very similar to our aug-cc-pVQZ result and suggests that errors due to basis set incompleteness are very small.

Analysis of the interaction using symmetry-adapted perturbation theory, together with the potential energy curve for rotation of the $H_2S$ unit relative to the benzene ring, suggests that the sulfur−$\pi$ interaction here is primarily an electrostatic attraction between the partial positive hydrogens in $H_2S$ and the negatively charged $\pi$ electrons of benzene.

Comparison of different theoretical treatments showed that MP2 overbinds and CCSD underbinds with respect to CCSD(T), in accord with studies on other van der Waals systems. The extra (1d/1p) diffuse functions present in the aug-cc-pVDZ basis set improve the overall quality of results obtained with that basis set over those obtained with the otherwise comparable 6-31++G** basis set by a significant amount. The extra functions in the aug-cc-pVTZ basis set produce a similar, but smaller, effect compared to the 6-311+G(2d,p) basis set. It is therefore recommended that the more complete aug-cc-pVXZ basis sets be employed when possible in future computational studies of this and similar van der Waals systems.

**Acknowledgment.** The authors thank Micah Abrams and Mutasem Sinnokrot for assistance with the computations. M.E.D. acknowledges support for this work as a 2003 Georgia Tech NSF-RSEC participant (grant no. CHE-9974841). C.D.S. acknowledges a Camille and Henry Dreyfus New Faculty Award and an NSF CAREER Award (grant no. CHE-0094088). The Center for Computational Molecular Science and Technology is funded through a Shared University Research (SUR) grant from IBM and by Georgia Tech.

**Supporting Information Available:** Potential energy curve data. This material is available free of charge via the Internet at http://pubs.acs.org.

## References and Notes

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

(2) Morgan, R. S.; Tatsch, C. E.; Gushard, R. H.; Mcadon, J. M.; Warme, P. K. *Int. J. Pept. Protein Res.* **1978**, *11*, 209.

(3) Morgan, R. S.; Mcadon, J. M. *Int. J. Pept. Protein Res.* **1980**, *15*, 177.

(4) Reid, K. S. C.; Lindley, P. F.; Thornton, J. M. *FEBS Lett.* **1985**, *190*, 209.

(5) Bernstein, F. C.; Koetzle, T. F.; Williams, G. J. B.; Meyer, E. F.; Brice, M. D.; Rodgers, J. R.; Kennard, O.; Shimanouchi, T.; Tasumi, M. *J. Mol. Biol.* **1977**, *112*, 535.

(6) Zauhar, R. J.; Colbert, C. L.; Morgan, R. S.; Welsh, W. J. *Biopolymers* **2000**, *53*, 233.

(7) Allen, F. H. *Acta Crystallogr. B* **2002**, *58*, 380.

(8) Viguera, A. R.; Serrano, L. *Biochemistry* **1995**, *34*, 8771.

(9) Munoz, V.; Serrano, L. *J. Mol. Biol.* **1995**, *245*, 275.

(10) Cheney, B. V.; Schulz, M. W.; Cheney, J. *Biochim. Biophys. Acta* **1989**, *996*, 116.

(11) Duan, G. L.; Smith, V. H., Jr.; Weaver, D. F. *Mol. Phys.* **2001**, *99*, 1689.

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

(13) Yamaotsu, N.; Moriguchi, I.; Kollman, P. A.; Hirono, S. *Biochim. Biophys. Acta* **1993**, *1163*, 81.

(14) Spencer, D. S.; Stites, W. E. *J. Mol. Biol.* **1996**, *257*, 497.

(15) Pranata, J. *Bioorg. Chem.* **1997**, *25*, 213.

(16) Prikhod′ko, I. V.; Vinogradova, I. V. *Russ. J. Appl. Chem. (Engl. Transl.)* **2002**, *75*, 1774.

(17) Raghavachari, K.; Trucks, G. W.; Pople, J. A.; Head-Gordon, M. *Chem. Phys. Lett.* **1989**, *157*, 479.

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

(19) Kendall, R. A.; Dunning, T. H.; Harrison, R. J. *J. Chem. Phys.* **1992**, *96*, 6796.

(20) Arunan, E.; Emilsson, T.; Gutowsky, H. S.; Fraser, G. T.; de Oliveira, G.; Dykstra, C. E. *J. Chem. Phys.* **2002**, *117*, 9766.

(21) Basis sets were obtained from the Extensible Computational Chemistry Environment Basis Set Database, Version 12/03/03, as developed and distributed by the Molecular Science Computing Facility, Environmental and Molecular Sciences Laboratory, which is part of the Pacific Northwest Laboratory, P.O. Box 999, Richland, WA 99352, U.S.A., and funded by the U.S. Department of Energy. The Pacific Northwest Laboratory is a multiprogram laboratory operated by Battelle Memorial Institute for the U.S. Department of Energy under contract DE-AC06-76RLO 1830. Contact David Feller or Karen Schuchardt for further information.

(22) Gauss, J.; Stanton, J. F. *J. Phys. Chem. A* **2000**, *104*, 2865.

(23) Edwards, T. H.; Moncur, N. K.; Snyder, L. E. *J. Chem. Phys.* **1967**, *46*, 2139.

(24) Amos, R. D.; Bernhardsson, A.; Berning, A.; Celani, P.; Cooper, D. L.; Deegan, M. J. O.; Dobbyn, A. J.; Eckert, F.; Hampel, C.; Hetzer, G.; Knowles, P. J.; Korona, T.; Lindh, R.; Lloyd, A. W.; McNicholas, S. J.; Manby, F. R.; Meyer, W.; Mura, M. E.; Nicklass, A.; Palmieri, P.; Pitzer, R.; Rauhut, G.; Schütz, M.; Schumann, U.; Stoll, H.; Stone, A. J.; Tarroni, R.; Thorsteinsson, T.; Werner, H.-J. *MOLPRO* (a package of ab initio programs designed by H.-J.Werner and P. J. Knowles), Version 2002.1; http://www.molpro.net.

(25) Jeziorski, B.; Moszynski, R.; Szalewicz, K. *Chem. Rev.* **1994**, *94*, 1887.

(26) Williams, H. L.; Szalewicz, K.; Jeziorski, B.; Moszynski, R.; Rybak, S. *J. Chem. Phys.* **1993**, *98*, 1279.

(27) Sinnokrot, M. O.; Sherrill, C. D. *J. Am. Chem. Soc.* **2004**, *126*, 7690.

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

(29) 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. *SAPT2002: An Ab Inito Program for Many-Body Symmetry-Adapted Perturbation Theory Calculations of Intermolecular Interaction Energies. Sequential and Parallel Versions*; 2003; http://www.physics.udel.edu/~szalewic/SAPT/SAPT.html.

(30) Dunning, T. H. *J. Chem. Phys.* **1989**, *90*, 1007.

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

(32) Halkier, A.; Klopper, W.; Helgaker, T.; Jørgensen, P.; Taylor, P. R. *J. Chem. Phys.* **1999**, *111*, 9157.

(33) Hopkins, B. W.; Tschumper, G. S. *J. Phys. Chem. A* **2004**, *108*, 2941.

(34) Tsuzuki, S.; Honda, K.; Uchimaru, T.; Mikami, M.; Tanabe, K. *J. Am. Chem. Soc.* **2000**, *122*, 11450.