JOC Article

Theoretical Study of Alkyl-π and Aryl-π Interactions. Reconciling Theory and Experiment

Jordi Ribas,[†] Elena Cubero,[†] F. Javier Luque,*[,‡] and Modesto Orozco*[,†]

Departament de Bioquímica, Facultat de Química, Universitat de Barcelona, Martí i Franquès 1, Barcelona 08028, Spain, Institut de Recerca Biomèdica, Parc Científic de Barcelona, Josep Samitier 1-5, Barcelona 08028, Spain, and Departament de Fisicoquímica, Facultat de Farmàcia, Universitat de Barcelona, Avgda Diagonal s/n, Barcelona 08028, Spain

modesto@mmb.pdb.ub.es; javier@far1.far.ub.es

Received February 22, 2002

Quantum mechanical and quantum mechanical/molecular mechanical calculations in conjunction with continuum solvation models have been used to analyze CH-π interactions in model systems of aryl- and alkyl-aromatic interactions, as well as in a model folding system designed to study those interactions. High level calculations reproduced accurately the interaction of CH-π interactions in both alkyl- and aryl-based model systems. Dispersion effects dominate the interaction, but the electrostatics term is also relevant for aryl CH-π interactions. Theoretical calculations were also used to examine the influence of CH-π interactions in determining the conformational flexibilitty of folding models. Finally, a critical comparison of the results obtained from high level calculations on model systems and the experimental data derived for folding models in apolar solvents was carried out, which allowed us to reconcile the apparent discrepancy found between both data.

Introduction

The weak interactions involving aromatic groups are key factors in determining many fundamental chemical and biochemical properties, such as the stabilization of the structure of proteins,[1−7] the conformational preferences of organic molecules,[5−15] and molecular recognition processes.[5,6,16−23] For instance, in a recent study Weiss and co-workers[4] found 3×10^5 CH-aromatic contacts in a database of 1000 proteins, which means that about three-quarters of all Trp and one-half of all Tyr and Phe residues participate in these interactions in the interior of proteins. This finding supports previous suggestions about the relevant contribution of CH-aromatic interactions to the folding of proteins[3] and reinforces their significant structural and functional role in biological macromolecules and supramolecular chemistry.

The π-electron distribution of aromatic rings can interact with charged species, yielding strong cation-π interactions, which are important in several biological and chemical processes.[24−29] This interaction is dominated by electrostatic and polarization effects,[16,24,25,30] while dispersion generally plays, at least for small cations, a secondary role.[30] Complexes of π-electron systems with standard hydrogen-bond (HB) donors have also been examined[31−35] and identified as a particular

[†] Departament de Bioquímica and Institut de Recerca Biomèdica.
[‡] Departament de Fisicoquímica.

(1) Burley, S. K.; Petsko, G. A. *Science* **1985**, *229*, 23.
(2) Blundell, T.; Singh, J.; Thornton, J.; Burley, S. K.; Petsko, G. A. *Science* **1986**, *234*, 1005.
(3) Burley, S. K.; Petskio, G. A. *J. Am. Chem. Soc.* **1986**, *108*, 7995.
(4) Brandl, M.; Weiss, M. S.; Jabs, A.; Sühnel, J.; Hilgenfeld, R. *J. Mol. Biol.* **2001**, *307*, 357.
(5) Kim, K. S.; Tarakeshwar, P.; Lee, J. Y. *Chem. Rev.* **2000**, *100*, 4145.
(6) Brutschy, B. *Chem. Rev.* **2000**, *100*, 3891.
(7) Tóth, G.; Murphy, R. F.; Lovas, S. *Protein Eng.* **2001**, *14*, 543.
(8) Suezawa, H.; Hashimoto, T.; Tsuchinaga, K.; Yoshida, T.; Yuzuri, T.; Sakakibara, K.; Hirota, M.; Nishio, M. *J. Chem. Soc., Perkin Trans. 2* **2000**, *2*, 1243.
(9) Paliwal, S.; Geib, S.; Wilcox, C. S. *J. Am. Chem. Soc.* **1994**, *116*, 4497−4498.
(10) Kim, E.; Paliwal, S.; Wilcox, C. S. *J. Am. Chem. Soc.* **1998**, *120*, 11192.
(11) Quan, R. W.; Li, Z.; Jacobsen, E. N. *J. Am. Chem. Soc.* **1996**, *118*, 8156.
(12) Adams, H.; Carver, F. J.; Hunter, C. A.; Morales, J. C.; Seward, E. M. *Angew. Chem., Int. Ed. Engl.* **1996**, *35*, 1542.
(13) Kobayashi, K.; Asakawa, Y.; Kikuchi, Y.; Toi, H.; Aoyama, Y. *J. Am. Chem. Soc.* **1993**, *115*, 2648.
(14) Newcomb, L. F.; Gellman, S. H. *J. Am. Chem. Soc.* **1994**, *116*, 4993.
(15) Gardner, R. R.; Christianson, L. A.; Gellman, S. H. *J. Am. Chem. Soc.* **1997**, *119*, 5041.
(16) Lindeman, S. V.; Kosynkin, D.; Kochi, J. K. *J. Am. Chem. Soc.* **1998**, *120*, 13268.
(17) Hunter, C. A. *Chem. Soc. Rev.* **1994**, *23*, 101.
(18) Hunter, C. A.; Sanders, J. K. M. *J. Am. Chem. Soc.* **1990**, *112*, 5525.
(19) Askew, B.; Ballester, P.; Buhr, C.; Jeong, K. S.; Jones, S.; Parris, K.; Williams, K., Rebek, J. *J. Am. Chem. Soc.* **1989**, *111*, 1082.
(20) Zimmerman, S. C.; VanZyl, C. M.; Hamilton, G. S. *J. Am. Chem. Soc.* **1989**, *111*, 1373.
(21) Shepodd, T. J.; Petti, M. A.; Dougherty, D. A. *J. Am. Chem. Soc.* **1989**, *110*, 1983.
(22) Rebek, J. *Angew. Chem., Int. Ed.* **1990**, *29*, 245.
(23) Diederich, F. *Angew. Chem., Int. Ed. Engl.* **1988**, *27*, 362.
(24) Ma, J. C.; Fougherty, D. A. *Chem. Rev.* **1997**, *97*, 1303.
(25) Dougherty, D. A. *Science* **1996**, *271*, 163.
(26) Sussman, J. L.; Harel, M.; Frolow, F.; Oefner, C.; Goldman, A.; Toker, L.; Silman, I. *Science* **1991**, *253*, 872.
(27) Kim, K. S.; Lee, J. Y.; Lee, S. J.; Ha, T. K.; Kim, D. H. *J. Am. Chem. Soc.* **1994**, *116*, 7399.
(28) Shepodd, T. J.; Petti, M. A.; Dougherty, D. A. *J. Am. Chem. Soc.* **1986**, *108*, 6085.
(29) Pullman, A.; Berthier, G.; Savinelli, R. *J. Am. Chem. Soc.* **1998**, *120*, 8553.
(30) Cubero, E.; Luque, F. J.; Orozco, M. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 5976.

10.1021/jo0201225 CCC: $22.00
Published on Web 09/10/2002

type of HB interaction. The stabilization energy in these complexes is weak (from −2 to −3 kcal/mol) and stems from both electrostatic (50−70% of the interaction energy depending on the donor) and dispersion forces.[36,37]

The CH group is described as a weak HB donor in complexes between alkyl or aryl groups with aromatic rings.[7,38,39] On the basis of theoretical calculations, some of these interactions have been identified as improper HBs.[40−42] The interaction is weak (interaction energies of 1−3 kcal/mol) for benzene complexes[3,43−52] and slightly stronger for larger aromatic rings.[44] Thus, for the CH_4-benzene dimer interaction energies of −1.2[51] and −1.4 kcal/mol[43] have been predicted from very high-level ab initio calculations. For the T-shaped benzene dimer interaction energies ranging from −2.1 to −3.2 kcal/mol have been determined from MP2 and CCSD(T) calculations performed with extended basis sets.[44−47] These values agree with empirical estimates determined from experimental properties of liquid benzene (−2.3 kcal/mol)[48] and from accurate ionization measurements for the benzene dimer (−2.4 ± 0.4 kcal/mol).[49]

Though the CH_4-benzene interaction is dominated by dispersion forces,[36,43,44,48,52] electrostatic effects still account for more than 20% of the interaction energy.[43,52] The importance of electrostatics is expected to be larger for the benzene dimer,[8,46,54,55] as well as for complexes involving unsaturated hydrocarbons as CH donors[43] or alkyl CH groups having electron-withdrawing substituents.[36,39] In summary, the bulk of theoretical data suggests that the CH_4-benzene interaction is weaker than the benzene-benzene interaction, and that, at least for the latter complex, electrostratic effects are not negligible. This picture, obtained mostly from high level quantum mechanical calculations, has been recently challenged by Wilcox and co-workers from[9,10,14] NMR data collected on folding models (see Figure 1). Their experimental results showed (i) no difference in the stability of alkyl-benzene and T-shaped benzene-benzene interactions and (ii) no change in benzene-benzene interaction when electron-donating or -withdrawing groups were added, suggesting that the electrostatic was irrelevant to the aryl-benzene interaction.

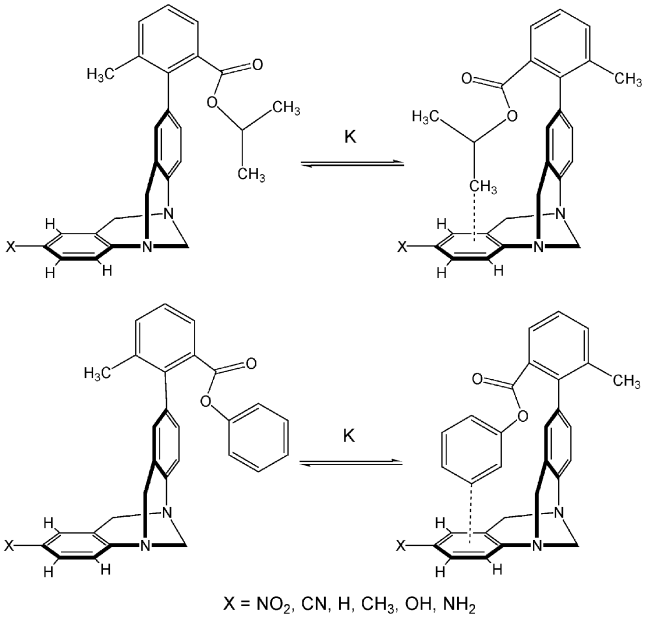

FIGURE 1. Schematic representation of the conformational equilibrium in Wilcox's folding models.[9,10] Top: derivatives of compound I. Bottom: derivatives of compound II. Left: *trans* conformation. Right: *cis* conformation.

In this paper we present a high-level ab initio study of the alkyl- and aryl-π dimerization, which compares theoretical results derived for small model systems in the gas phase with the experimental data determined for model folding systems in apolar solvents. The apparent discrepancy between both sets of data is reconciled.

Methods

The quantum chemical description of aryl- and alkyl-π interactions is difficult as a result of the magnitude of the dispersion energy, which makes it necessary to combine large basis sets with electron correlation effects,[36,43−46,47,56−59] precluding the use of density functional methods for determining interaction energies.[58,59] Fortunately, lower level correlated methods can give reliable interaction energies, since differences between MP2 and CCSD(T) results are generally small.[43,46−48,52] For instance, such a difference amounts to 0.3

(31) Oki, M.; Iwamura, H. *Bull. Chem. Soc. Jpn.* **1959**, *32*, 955.
(32) Perutz, M. F.; Fermi, G.; Abraham, D. J.; Poyart, C.; Bursaux, E. *J. Am. Chem. Soc.* **1986**, *108*, 1064.
(33) Burley, S. K.; Petsko, G. A. *FEBS Lett.* **1986**, *203*, 139.
(34) Suzuki, S.; Green, P. G.; Bumgarner, R. E.; Dasgupta, S.; Goddard, W. A., III; Blake, G. A. *Science* **1992**, *257*, 942.
(35) Rodham, D. A.; Suzuki, S.; Suenram, R. D.; Lovas, F. J.; Dasgupta, S.; Goddard, W. A., III; Blake, G. A. *Nature* **1993**, *362*, 735.
(36) Tsuzuki, S.; Honda, K.; Uchimaru, T.; Mikami, M.; Tanabe, K. *J. Am. Chem. Soc.* **2000**, *122*, 11450.
(37) Alhambra, C.; Luque, F. J.; Orozco, M. *J. Phys. Chem.* **1995**, *99*, 3084.
(38) Hirota, M.; Sakaibara, K.; Suezawa, H.; Yuzuri, T.; Ankai, E.; Nishio, M. *J. Phys. Org. Chem.* **2000**, *13*, 620.
(39) Lindeman, S. V.; Kosynkin, D.; Kochi, J. K. *J. Am. Chem. Soc.* **1998**, *120*, 13268.
(40) Hobza, P.; Selzle, H. L.; Schalg, E. W. *J. Phys. Chem.* **1996**, *100*, 18790.
(41) Cubero, E.; Orozco, M.; Hobza, P.; Luque, F. J. *J. Phys. Chem.* **1999**, *103*, 814.
(42) Hobza, P.; Havlas, Z. *Chem. Rev.* **2000**, *100*, 4253.
(43) Tsuzuki, S.; Honda, K.; Uchimaru, T.; Mikami, M.; Tanabe, K. *J. Am. Chem. Soc.* **2000**, *122*, 3746.
(44) Tsuzuki, S.; Uchimaru, T.; Matsumura, K.; Mikami, M.; Tanabe, K. *Chem. Phys. Lett.* **2000**, *319*, 547.
(45) Hobza, P.; Selzle, H. L.; Schlag, E. W. *J. Am. Chem. Soc.* **1994**, *116*, 3500.
(46) Hobza, P.; Selzle, H. L.; Schalg, E. W. *J. Phys. Chem.* **1996**, *100*, 18790.
(47) Tsuzuki, S.; Honda, K.; Uchimaru, T.; Mikami, M.; Tanabe, K. *J. Am. Chem. Soc.* **2002**, *124*, 104.
(48) Jorgensen, W. L.; Severance, D. L. *J. Am. Chem. Soc.* **1990**, *112*, 4762.
(49) Grower, J. R.; Walters, E. A.; Hui, E. T. *J. Phys. Chem.* **1987**, *91*, 3223.
(50) Krause, H.; Ernsterger, B.; Neusser, H. J. *Chem. Phys. Lett.* **1991**, *184*, 411.
(51) Ehama, R.; Yokoo, A.; Tsushima, M.; Yuzuri, T.; Suezawa, H.; Hitora, M. *Bull. Chem. Soc. Jpn.* **1993**, *66*, 814.
(52) Tarakeshwar, P.; Choi, H. S.; Kim, K. S. *J. Am. Chem. Soc.* **2001**, *123*, 3323.
(53) Karlström, G.; Linse, P.; Wallqvist, A.; Johnson, B. *J. Am. Chem. Soc.* **1983**, *105*, 3777.
(54) Janda, K. C.; Hemminger, J. C.; Winn, J. S.; Novick, S. E.; Harris, S. J.; Klemperer, W. *J. Chem. Phys.* **1975**, *63*, 1419.
(55) Steed, J. M.; Dixon, T. A.; Klemperer, W. *J. Chem. Phys.* **1979**, *70*, 4940.
(56) Tsuzuki, S.; Uchimaru, T.; Mikami, M.; Tanabe, K. *Chem. Phys. Lett.* **1996**, *252*, 206.
(57) Tsuzuki, S.; Uchimaru, T.; Tanabe, K. *THEOCHEM* **1994**, *307*, 107.
(58) Tsuzuki, S.; Uchimaru, T.; Tanabe, K. *Chem. Phys. Lett.* **1998**, *287*, 202.
(59) Tsuzuki, S.; Lüthi, H. P. *J. Chem. Phys.* **2001**, *114*, 3949.

kcal/mol for the methane-benzene dimer with the cc-pvdZ basis set[43] and 0.7−0.8 kcal/mol for the benzene dimer with the aug-cc-pvdZ and cc-pVTZ basis sets.[46,47]

Calculations for model systems were carried out at the MP2/6-311+G(2d,p) level. The CH group was oriented pointing toward the center of the ring, along the axis normal to the benzene ring. The rest of geometrical parameters were fully optimized. Single-point calculations at the MP2/aug-cc-pVDZ, MP2/6-311G(2d,2pd), and CCSD(T)/aug-cc-pVDZ levels were performed using those optimized geometries. Basis set superposition error (BSSE) was corrected in all the cases using the counterpoise method,[60] which was also considered in determining the optimum geometries.

The study of Wilcox's compounds is more difficult as a result of (i) its large size, which limits the size of the basis sets, and (ii) the problems derived from the correction of BSSE effects in intramolecular interactions.[61] Fortunately, test calculations showed that all of these problems can be solved simultaneously because of a fortuitous cancellation of errors. Thus, medium-size basis sets (at the MP2 level) typically underestimate the interaction energy as the result of an incorrect representation of dispersion effects,[36,43−47,56−59] and in fact, Kim and co-workers have suggested that BSSE does not have to be corrected completely[27,52] in these calculations. When we compared MP2/6-31G(d) and MP2/6-311+G(2d,p) profiles we found that in all cases BSSE-uncorrected MP2/6-31G(d) results reproduce almost perfectly BSSE-corrected MP2/6-311+G(2d,p) values, as well as single-point calculations at even higher levels of theory. This led us to completely ignore BSSE in calculations performed with Wilcox's compounds.

The impossibility to perform MP2/6-31G(d) optimizations on Wilcox's compounds introduces a priori some uncertainties in the calculations. To reduce errors, geometries in both *cis* and *trans* conformations were optimized using two different strategies: first, B3LYP/6-31G(d) optimizations[62] and, second, ONIOM optimizations[63] where the two interacting groups (two benzenes or one benzene and one isopropyl groups) were represented at the MP2/6-31G(d) level and the rest of the system at the HF/6-31G(d) level. Both geometries were in general similar (see below) and agree well with available crystal data. Once the geometries were optimized at either ONIOM or B3LYP levels molecular energies were computed using MP2/6-31G(d) calculations. Thermal and entropic effects on the conformational equilibrium were corrected using B3LYP frequencies and the standard harmonic procedure. Because of the mixed character of ONIOM frequencies, these were not used to obtain thermal or entropic corrections.

To complement the high-level ab initio calculations, the interactions in complexes of model compounds were analyzed by means of the Generalized Molecular Interaction Potential including polarization correction (GMIPp).[30,64−66] This simple QM/MM method is based on a perturbational treatment of the interaction between quantum and classical systems and allows a fast partition of the interaction energy into electrostatic, polarization, and dispersion components.[30,64−66] Besides its simplicity, the method provides a reliable representation of several π-based interactions.[30,66]

Finally, solvation effects were introduced by means of continuum solvation models. The AM1-optimized version of the Miertus−Scroco−Tomasi (MST) method[67,68] for chloroform[69,70] and the SM5.4A version of the continuum method developed by Cramer and Truhlar[71] were used to include solvent effects in the conformational equilibrium of Wilcox's compounds. Geometry optimization in chloroform was tried for both *cis* and *trans* conformers, but the semiempirical optimized geometries (both in gas phase and solution) were unreliable.[72] This forced us to use DFT gas-phase optimized geometries for all the continuum calculations. The small magnitude of the solvent effects revealed by experimental techniques[9,10] was confirmed by continuum calculations (see below), suggesting that no important geometrical changes can be expected due to solvation in chloroform.

Ab initio calculations were performed using Gaussian-98[73] and Jaguar[74] computer programs. GMIPp calculations were carried out using our MOPETE code;[75] AM1-MST calculations were carried out with a modified version of MOPAC.[76,77] Finally, SM5.4A calculations were performed using the Cramer and Truhlar's AMSOL computer program.[78]

Results and Discussion

Model Systems. Figure 2 shows the MP2/6-311+G(2d,p)//MP2/6-311+G(2d,p) interaction energy profiles for the CH-π complexes of methane (Figure 2, top) and benzene (Figure 2, bottom) interacting with the π distribution of a benzene molecule. Both uncorrected and BSSE-corrected profiles are shown. As expected, the BSSE correction reduces importantly the interaction energy and slightly increases the optimum interaction distance (0.1 and 0.2 Å for the methane-benzene and the T-shaped benzene-benzene dimers). The interaction profiles are smooth near the minimum, and displacements of up to ±0.3 Å from equilibrium positions change the interaction energy less than 0.3 (methane-benzene) and 0.4 (benzene-benzene) kcal/mol, suggesting a large flexibility in both types of interactions. This suggests that small errors in the optimized geometry of the dimer (like those expected in B3LYP and ONIOM optimizations) should not introduce important errors in the interaction energies.

(60) Boys, S. F.; Bernardi, F. *Mol. Phys.* **1989**, *214*, 15.
(61) LMP2 calculations, which are suggested to be BSSE-free, lead to spurious results in model calculations and were not used for the study of Wilcox's molecules.
(62) Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1998**, *37*, 785.
(63) Dapprich, S.; Komaromi, I.; Byun, K. S.; Morokuma, K.; Frisch, M. J. *THEOCHEM* **1999**, *461*, 1.
(64) Luque, F. J.; Orozco, M. *J. Comput.Chem.* **1998**, *19*, 866.
(65) Orozco, M.; Luque, F. J. In *Molecular Electrostatic Potentials: Concepts and Applications*; Murray, J. S., Sen, K., Eds.; Elsevier: Amsterdam; Vol. 3, pp 181−218.
(66) Hernández, B.; Luque, F. J.; Orozco, M. *J. Comput. Chem.* **1999**, *20*, 937.
(67) Miertus, S.; Scrocco, E.; Tomasi, J. *Chem. Phys.* **1981**, *51*, 117.
(68) Miertus, S.; Tomasi, J. *Chem. Phys.* **1982**, *65*, 239.
(69) Luque, F. J.; Zhang, Y.; Alemán, C.; Bachs, M.; Gao, J.; Orozco, M. *J. Phys. Chem.* **1996**, *100*, 4269.
(70) Curruchet, C.; Orozco, M.; Luque, F. J. *J. Comput. Chem.* **2001**, *22*, 1180.
(71) Chambers, C. C.; Giesen, D. J.; Cramer, C. J.; Truhlar, D. G. *J. Phys. Chem.* **1997**, *101*, 2061.
(72) Hobza, P.; Kabelac, M.; Sponer, J.; Mejzlik, P.; Vondrasek, J. *J. Comput. Chem.* **1997**, *18*, 19997.
(73) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Zakrzewski, V. G.; Montgomery, J. A., Jr.; Stratmann R. E.; Burant, J. C.; Dapprich, S.; Millam, J. M.; Daniels, A. D.; Kudin, K. N.; Strain, M. C.; Farkas, O.; Tomasi, J.; Barone, V.; Cossi, M.; Cammi, R.; Mennucci, B.; Pomelli, C.; Adamo, C.; Clifford, S.; Ochterski, J.; Petersson, G. A.; Ayala, P. Y.; Cui, Q.; Morokuma, K.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Cioslowski, J.; Ortiz, J. V.; Baboul, A. G.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Gonzalez, C.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Andres, J. L.; Gonzalez, C.; Head-Gordon, M.; Replogle, E. S.; Pople, J. A. *Gaussian 98*, rev. A.7; Gaussian, Inc.: Pittsburgh, PA, 1998.
(74) *Jaguar Computer Program*, version 3.5; Schrödinger, Inc.: Portland, OR, 1998.
(75) Luque, F. J.; Orozco, M. *Mopete Computer Program*; Universitat de Barcelona: Barcelona, 2001.
(76) Stewart, J. J. P. *MOPAC93*; Fujitsu Ltd.: 1993.
(77) Luque, F. J.; Orozco, M. *MOPAC version with optimized MST code*; 2001
(78) Hawkins, G. D.; Giesen, D. J.; Lynch, G. C.; Chamber, C. C.; Rossi, I.; Storer, J. W.; Li, J.; Rinaldi, D.; Liotard, D. A.; Cramer, C. J.; Truhlar, D. G. *AMSOL6.5.3*; University of Minnesota: Minneapolis, 1999.

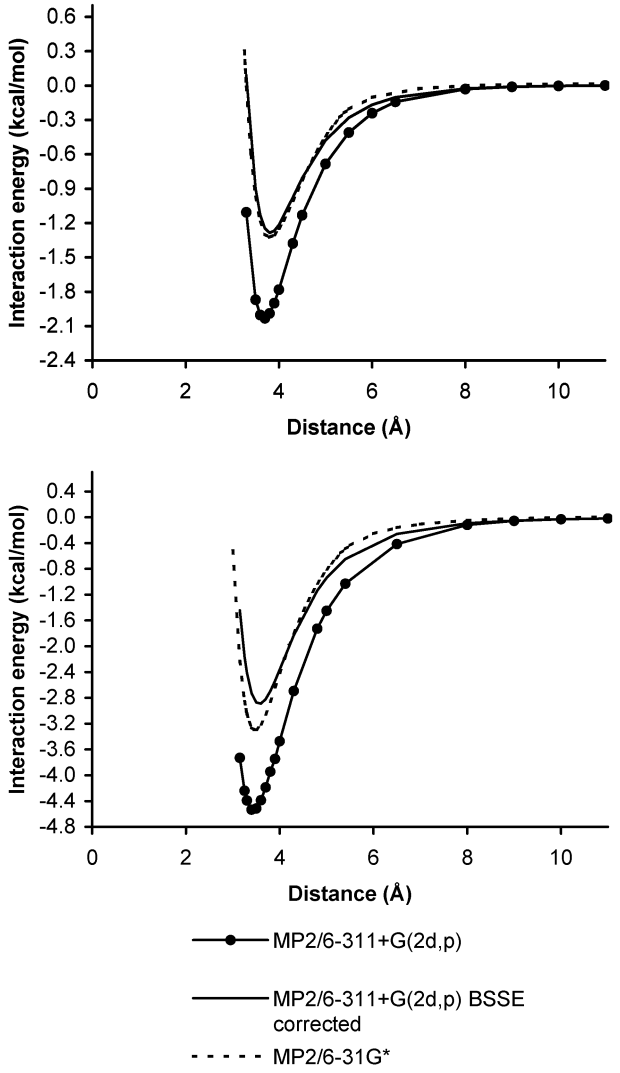

FIGURE 2. Interaction profiles for the methane-benzene and T-shaped benzene-benzene dimers computed at the MP2/6-311+G(2d,p)//MP2/6-311+G(2d,p) levels with and without BSSE correction. Values obtained using BSSE-uncorrected MP2/6-31G(d) calculations are also displayed.

Table S1 (see Supporting Information) displays the interaction energies at the optimum MP2/6-311+G(2d,p) geometries for the methane-benzene and (T-shape) benzene-benzene dimers computed at different levels of theory and compared with values reported in the literature. For the methane-benzene dimer, the MP2/6-311+G(2d,p) result (−1.3 kcal/mol) matches the values determined at higher levels of theory (−1.2 kcal/mol at the MP2/aug-cc-pVDZ level;[50] −1.4 kcal/mol at the MP2/complete basis level combined with CCSD(T)/ccpVDZ calculations).[43] The ability of the MP2/6-311+G(2d,p) method to reproduce higher level calculations is also very good for the benzene-benzene dimer. The value of −2.8 kcal/mol (MP2/6-311+G(2d,p)) agrees with previous values of −3.0 (MP2/aug-cc-pVDZ),[46] −3.0 (CCSD(T)/cc-pVTZ),[47] −3.2 (MP2/aug(df,p)-6-311G(d,p)),[44] and −2.5 (MP2/aug(df,p)-6-311G(d,p) combined with CCSD(T)/6-311G(d,p) corrections)[44] kcal/mol. There is also good agreement with our estimates at the MP2/6-31+G(2d,p)//MP2/aug-cc-pVDZ, MP2/6-31+G(2d,p)//MP2/6-311G(2d,-2pd), and MP2/6-31+G(2d,p)//CCSD(T)/aug-cc-pVDZ levels of theory (Table S1). It also agrees with the experimental value reported for the T-shaped interaction of benzene (2.4 ± 0.4 kcal/mol[49]). Finally, the optimum geometry found here at the MP2/6-311+G(2d,p) level also agrees with the experimental one (the optimum distance between the centers of mass of the benzene rings is 5.0 Å, as experimentally found).[79]

In summary, MP2/6-311+G(2d,p) calculations provide reliable interaction energies for the methane-benzene and benzene-benzene dimers. Unfortunately, this level of calculation is still too expensive for the study of large folding models. Therefore, we checked the behavior of several lower-level methods, and the best agreement was found with BSSE-uncorrected MP2/6-31G(d) values. The differences in the interaction energy profiles obtained between BSSE-free MP2/6-311G(2d,p) and BSSE-uncorrected MP2/6-31G(d) calculations are small (Figure 2). Using the optimized geometries the difference between the interaction energies determined at the two preceding levels of theory amounts to 0.0 and 0.3 kcal/mol for methane-benzene and benzene-benzene, respectively. The optimum distance for the benzene-benzene dimer from MP2/6-31G(d) calculations is 4.9 Å, which closely agrees with higher QM and experimental data.[79] In summary, BSSE-uncorrected MP2/6-31G(d) calculations give a very accurate description of the CH-π interaction.

Comparison of MP2 and HF interactions profiles (see Figure 3) clearly shows the importance of dispersion effects in CH-benzene interactions.[36,43,44,48,53] According to GMIPp calculations (Table 1), dispersion-repulsion interactions accounts for 60% (benzene) and 80% (methane) of the interaction energy at the optimum geometries. In turn, electrostatics account for 15% (methane) and 32% (benzene) of the interaction energy. It is worth noting the excellent quality of GMIPp calculations to reproduce both interaction geometries and energies (Tables S1 and 1). Furthermore, the GMIPp partition of the total interaction energy provides results very similar to those derived from other partitioning schemes,[43,52] confirming the goodness of the simple and computationally efficient GMIPp scheme.

Effect of Substituents. To analyze the influence of substituents attached to the (acceptor) benzene on the CH-π interaction, we computed the interaction energies at both MP2/6-31G(d) and GMIPp levels for the dimers CH_4 or C_6H_6 (donor) and tetrafluorinated and NO_2-, CN-, CH_3-, OH-, and NH_2-substituted benzenes (acceptor). The optimum geometries found at the MP2/6-31+G(2d,p) level for the isolated CH4-benzene and T-shaped benzene-benzene dimers were used (see above).

The interaction energy of methane with benzene is mostly unaffected by the presence of substituents on the aromatic ring (see Table S2; Supporting Information). Thus, the interaction energies vary from −1.1 to −1.4 kcal/mol, and the electrostatic component varies from +0.1 and −0.2 kcal/mol. These findings support the small relevance of electrostatic effects in alkyl-benzene interactions, even in cases where polar groups are introduced in the benzene core. It is worth noting the extreme quality of GMIPp calculations for the representation of this series of dimers, since in general GMIPp and MP2 calculations are identical (see Table S2).

(79) Arunan, E.; Gutowsky, H. S. *J. Chem. Phys.* **1993**, *98*, 4294.

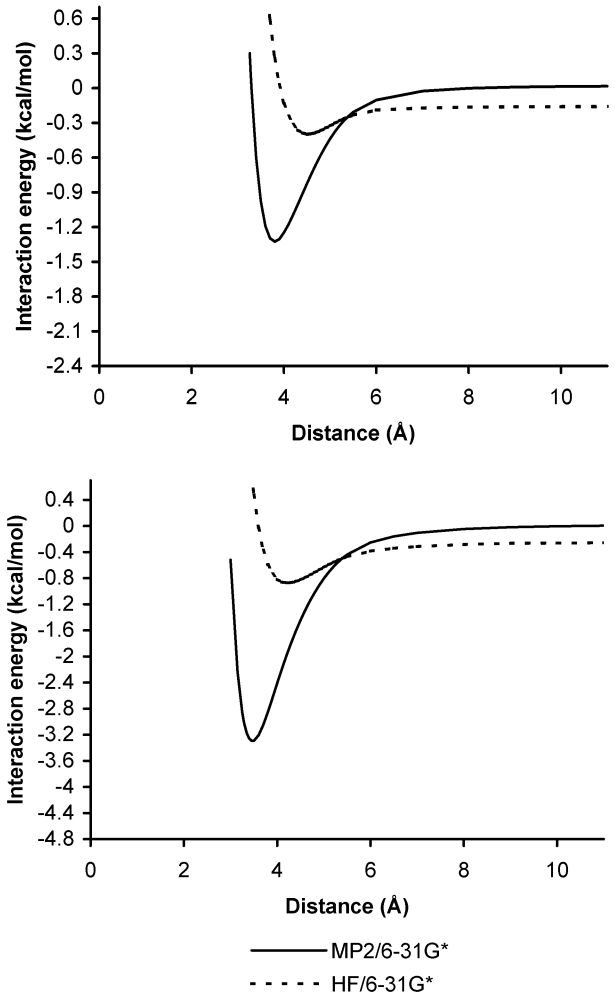

FIGURE 3. Interaction profiles for the methane-benzene and T-shaped benzene-benzene dimers computed at the HF and MP2 levels using the 6-31G(d) basis set.

TABLE 1. GMIPp Total Interaction Energy and Its Electrostatic, Polarization, and Dispersion-Repulsion Contributions; GMIPp Results Are Compared with MP2/6-31G(d) and MP2/6-311+G(2d,p) Estimates[a]

donor	E_{ele}	E_{pol}	$E_{disp-rep}$	E_{tot} (GMIPp)	E_{tot} (MP2)[b]	E_{tot} (MP2)[c]
CH_4	−0.2	−0.1	−1.0	−1.3	−1.3	−1.3
C_6H_6	−0.8	−0.2	−1.5	−2.5	−3.1	−2.8

[a] All values are in kcal/mol. [b] MP2/6-31G(d). [c] MP2/6-311+G(2d,p).

The situation when the CH donor is benzene is quite different (Table S3; Supporting Information). The presence of substituents changes the interaction energies up to 1.1 (MP2) or 1.2 (GMIPp) kcal/mol, which can be related to changes in the electrostatic term (see Table S3). For some compounds, like the amino derivative, the electrostatic term contributes more than half of the dispersion-repulsion term. In contrast, for the tetrafluorinated compound it destabilizes the dimer by about −17% of the magnitude of dispersion-repulsion term. In summary, polar substituents in the acceptor benzene modifies the strength of T-shaped benzene-benzene interactions by modulating the electrostatic interaction (see Figure 4), which is the key parameter to explain the

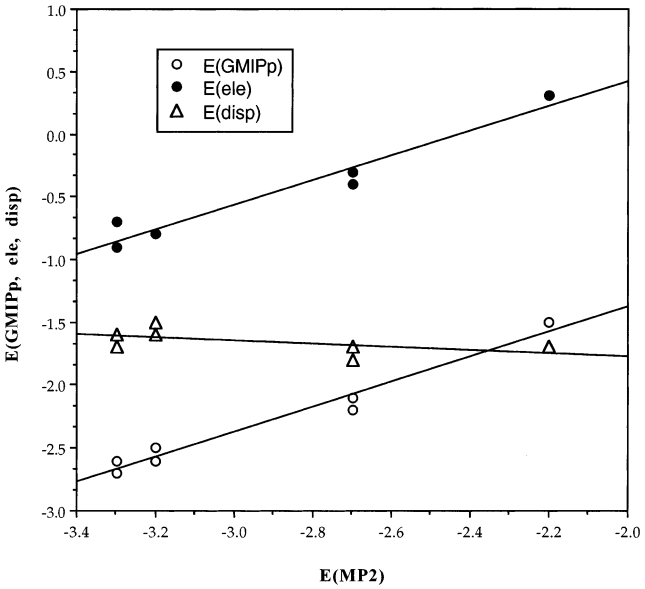

FIGURE 4. Representation of MP2 interaction ernergies for compound II series versus GMIPp total, electrostatic, and dispersion ernergies. All values are in kcal/mol.

change in binding energy along the series. Finally, it is worth noting again the ability of the GMIPp to reproduce the MP2 values.

The benzene-benzene interaction is stronger than the methane-benzene one for all the substituents owing to both dispersion and repulsion (see Tables S2 and S3). The dispersion-repulsion preference for the benzene-benzene series is always around 0.6 kcal/mol irrespective of the substituents attached to the acceptor benzene ring (see Tables S2 and S3). On the contrary, the electrostatic component is responsible for the change in the interaction energies along the series of compounds. In fact, more than 95% of the differences between the MP2 dimerization energies of the difference benzene-benzene and methane-benzene dimers (ΔMP2) is explained by the change in GMIPp electrostatic energy (ΔGMIPp) ($r = 0.98$ in Figure 5, see ref 80).

In summary, all the results reported above confirm the previously accepted picture (see Introduction) of CH-π interactions: (i) the T-shaped benzene-benzene dimer is more stable than the methane-benzene dimer, (ii) alkyl-π and aryl-π interactions are dominated by dispersion effects, (iii) electrostatic effects are not negligible for the benzene-benzene dimer, and (iv) substituents changing the π-electron population of the acceptor benzene strongly modify the strength of T-shaped benzene-benzene dimers.

Folding Models. Wilcox's structures[9,10] provide unique systems to study both aryl-π and alkyl-π interactions. The molecules can be in the *trans* conformation or in the *cis* conformation, this latter being stabilized by CH-π interactions between a benzene acceptor and either an isopropyl (compound I) or a benzene (compound II) ring (see Figure 1). Wilcox and co-workers determined folding free energies, i.e., free energy differences between *trans* and *cis* conformers in $CDCl_3$ with high accuracy (10% error)

(80) Relative energies are computed for each substituent (X in Figure 1) as $\Delta E = E^{int}_{cmpdII} - E^{int}_{cmpdI}$. Interaction energies are displayed in Tables S2 and S3.

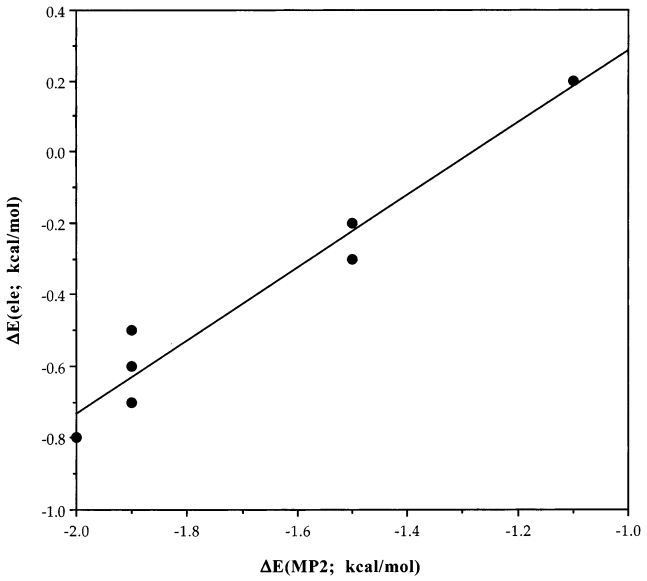

FIGURE 5. Representation of MP2 and electrostatic energy differences between benzene-benzene and methane-benzene interaction, when the acceptor benzene is substituted by chemical groups in Figure 1. All values are in kcal/mol.

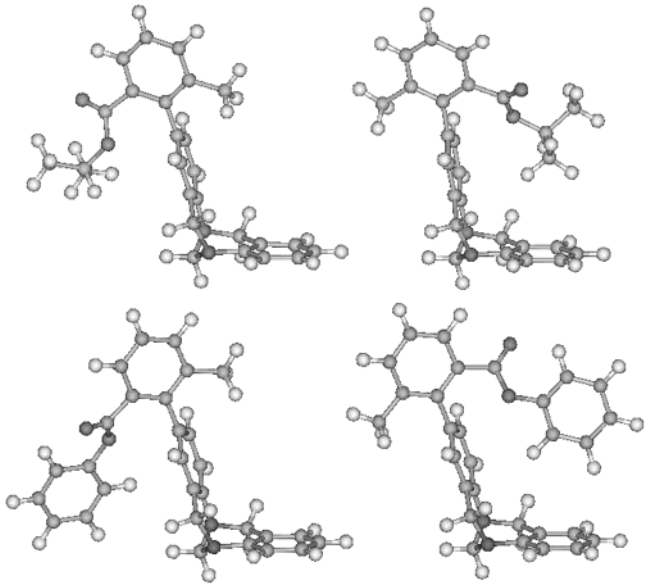

FIGURE 6. B3LYP optimized geometries for compound I (top) and compound II (bottom) in the *trans* (left) and *cis* (right) conformations. ONIOM optimized geometries (not shown) are very close to the B3YLP ones.

by NMR techniques. They claim, supported by overwhelming evidence,[9,10] that the effect of solvation is small, and the conclusions of their study can then be of general validity for gas phase or any apolar solvent.

Wilcox's group found that (i) the folding free energy of compound I (alkyl-benzene interaction) was slightly more negative than that of compound II (benzene-benzene interaction), and (ii) the folding energy of the two compounds was not largely affected by monosubstitutions in the acceptor benzene ring.[9,10] These results suggest that (i) alkyl-π and aryl-π interactions have similar interaction energies, and (ii) electrostatic effects are negligible for CH_4-benzene and (T-shape) benzene-benzene interactions, even in the presence of substituted acceptor benzenes. These conclusions apparently disagree with all the theoretical calculations in the gas phase summarized above.

To verify that current methods were able to reproduce model folding systems in solution and not only small isolated molecules in the gas phase, we built models of Wilcox's compounds in both *cis* and *trans* conformations and optimized them using B3LYP/6-31G(d) or ONION-(MP2:HF/6-31G(d)) calculations. The optimized structures of compounds I and II obtained by ONIOM and B3LYP calculations were almost identical (maximum RMSd 0.1 Å for the *cis* conformers and 0.3 Å for the *trans* conformer). The CH-π interactions are evident in both *cis* conformers (see Figure 6) and agree well with suggestions by Wilcox and co-workers. The optimized structures of the *cis* conformer of compound II are similar (see Figure 7, RMSd 0.5−0.6 Å) to the crystal structure of a nitro-methyl derivative.[9] Considering the smoothness of the potential energy surface of CH-π interactions these small geometrical uncertainties are expected to introduce very small energy changes.

Folding energies and free energies were determined from MP2/6-31G(d) calculations (see Methods) for several modified compounds (X = H, NO_2, CN, CH3, OH, and NH_2 in Figure 1) as well as the tetrafluorinated derivative. The effect of solvation was introduced by using continuum methods (see Methods). The results are summarized in Tables 2 and 3 for derivatives of compounds I and II, respectively.

For compound I folding energies lie in a range of 0.8 kcal/mol (0.3 kcal/mol if the amino derivative is excluded;[81] see Table 2). As expected, entropic effects disfavor folding, but to a similar extent for all of the derivatives. There are no major changes in the folding free energies due to the presence of substituents in the acceptor benzene, in agreement with previous calculations on model compounds (see above). The solvent effect in the folding free energy is small (especially at the AM1-MST level) as suggested by Wilcox and co-workers (see above). Folding free energies in chloroform are in reasonable agreement with Wilcox's experimental values (RMSd between 0.4 and 1.3 kcal/mol depending on the theoretical models used to compute the gas phase and solvation terms). The systematic overestimation of the *cis* stability detected by all the theoretical methods can be justified considering that entropic effects cannot be well introduced by the two states harmonic model when one of the conformers (*trans*) have many more accessible microstates than the other (*cis*). Finally, MP2/B3LYP and MP2/ONIOM results agree very well (RMSd 0.2 kcal/mol), suggesting that for all derivatives of compound II both ONIOM and B3LYP geometries are almost identical.

The folding of the derivatives of compound II is also dominated by the gas-phase energy term, since entropic and solvation contributions do not introduce large changes in the conformational equilibrium (see Table 3). Folding free energies for derivatives of compound II are not very

(81) There is a systematic (and unexplained) tendency of theoretical calculations to slightly overestimate the stability of the *cis* conformer of the amino derivatives of compounds I and II according to experimental results.

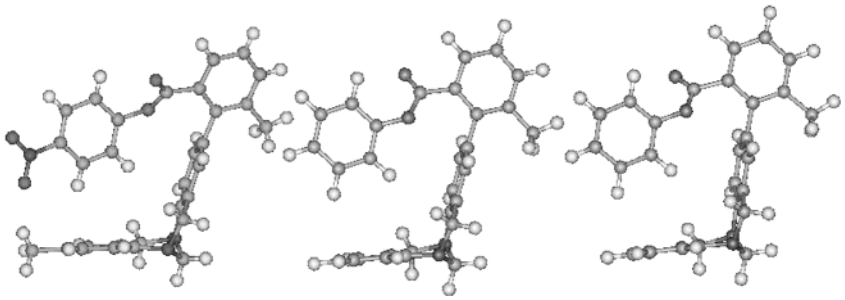

FIGURE 7. Representation of B3LYP (middle) and ONIOM (right) optimized geometries of compound II. The crystal structure[9] corresponding to the nitro(donor)-methy(acceptor) derivative of compound II is shown for comparison (left).

TABLE 2. Folding Energies and Free Energies in Gas Phase for Different Derivatives of Compound I Computed from MP2/6-31G(d) Calculations[a]

derivatives cmpd I	ΔE_{fold}	ΔG_{fold}^{gas}	$\Delta\Delta G_{solv}$	$\Delta G_{fold}^{CHCl_3}$	ΔG_{fold}^{exp}
tetrafluorinated	−2.0, **−2.0**	−1.6, **−1.6**	*0.0*, 0.1	*−1.6*, ***−1.6***, −0.9, **−0.9**	−0.56
X = NO_2	−2.1, **−2.2**	−1.8, **−1.9**	*0.1*, 0.9	*−1.7*, ***−1.8***, −0.9, **−1.0**	−0.51
X = CN	−2.1, **−2.3**	−1.9, **−2.1**	*0.1*, 0.6	*−1.8*, ***−2.0***, −1.3, **−1.5**	−0.64
X = H	−2.1, **−2.2**	−1.7, **−1.8**	*0.1*, 1.1	*−1.6*, ***−1.7***, −0.6, **−0.7**	
X = CH_3	−2.2, **−2.5**	−1.8, **−2.1**	*0.1*, 1.3	*−1.7*, ***−2.0***, −0.5, **−0.8**	−0.44
X = OH	−2.3, **−2.6**	−1.9, **−2.2**	*0.2*, 1.4	*−1.7*, ***−2.0***, −0.5, **−0.8**	−0.47
X = NH_2	−2.8, **−2.9**	−2.4, **−2.5**	*0.3*, 0.8	*−2.1*, ***−2.2***, −1.6, **−1.7**	−0.34

[a] Values in the gas phase are determined using both B3LYP (plain) and ONIOM (**bold**) optimized geometries. Differences in solvation free energies $\Delta\Delta G_{solv}$ in chloroform were computed using AM1-MST (*italics*) and AMSOL-SM5.4A (plain) methods. The resulting folding free energies in chloroform were computed by adding gas-phase free energies to differential solvation free energies. Four values are then obtained by combination of B3LYP/ONIOM and MST/SM5.4A values. Experimental results from refs 9 and 10 are presented for comparison. All values are in kcal/mol.

TABLE 3. Folding Energies and Free Energies in the Gas Phase for Different Derivatives of Compound II Computed from MP2/6-31G(d) Calculations[a]

derivatives cmpd II	ΔE_{fold}	ΔG_{fold}^{gas}	$\Delta\Delta G_{solv}$	$\Delta G_{fold}^{CHCl_3}$	ΔG_{fold}^{exp}
tetrafluorinated	−3.1,[b] **−4.7**[b]	−2.7, **−4.3**	*0.0*, 0.1	*−2.6*, ***−4.3***, −2.6, **−4.2**	−0.43
X = NO_2	−2.4,[b] **−4.9**[b]	−2.2, **−4.7**	*0.1*, 0.7	*−2.1*, ***−4.6***, −1.5, **−4.0**	−0.21
X = CN	−1.4, **−2.8**[b]	−1.3, **−2.7**	*−0.4*, −0.2	*−1.7*, ***−3.1***, −1.5, **−2.9**	−0.30
X = H	−1.7, **−1.9**	−1.3, **−1.5**	*0.0*, 0.8	*−1.3*, ***−1.5***, −0.5, **−0.7**	
X = CH_3	−2.0, **−2.5**	−1.4, **−1.9**	*0.1*, 0.7	*−1.3*, ***−1.9***, −0.7, **−1.2**	−0.27
X = OH	−2.1, **−2.8**	−1.8, **−2.5**	*0.1*, 0.9	*−1.6*, ***−2.3***, −0.9, **−1.6**	−0.23
X = NH_2	−2.6, **−3.1**	−2.5, **−3.0**	*0.0*, 0.7	*−2.5*, ***−3.0***, −1.8, **−2.3**	−0.18

[a] Values in the gas phase are determined using both B3LYP (plain) and ONIOM (**bold**) optimized geometries. Differences in solvation free energies $\Delta\Delta G_{solv}$ in chloroform were computed using AM1-MST (*italics*) and AMSOL-SM5.4A (plain) methods. The resulting folding free energies in chloroform were computed by adding gas-phase free energies to differential solvation free energies. Four values are then obtained by combination of B3LYP/ONIOM and MST/SM5.4A values. Experimental results from refs 9 and 10 are presented for comparison. All values are in kcal/mol. [b] Optimization yield to parallel stackings. Interaction energies obtained by fixing the cis conformation in the T-shaped orientation detected for the unmodified compound II (X = H) are X = 4F, −1.3 kcal/mol; X = NO_2, −1.5 kcal/mol; X = CN, −1.4 kcal/mol.

different from those found for compound I, in agreement with Wilcox's findings.[9,10] The inclusion of substituents into the acceptor benzene does not greatly affect the folding free energies, in agreement also with Wilcox's findings.[9,10] Interestingly (see Table 3), when electron-widthdrawing groups are introduced the folding free energy increases (in absolute value) despite the depletion in π-electrons. This seems to suggest that electrostatic interactions do not have any role in T-shaped benzene-benzene interactions. However, analysis of the optimized geometries show that, in fact, when the acceptor benzene is very depleted in π-electrons the optimum geometry corresponds to a parallel or quasiparallel π-stacking and not to a T-shaped interaction (see Figure 8). For instance, if the T-shaped orientation is forced[82] the MP2/6-31G(d) folding energies for the tetrafluoro and nitro derivatives are −1.3 and −1.5 kcal/mol respectively, values similar to those found for the corresponding derivatives of compound I and to experimental measures (see Table 3). Caution is then necessary before assuming that all the optimum *cis* geometries correspond to the T-shaped interactions. Otherwise the relative stability of the *cis* conformer can be underestimated.

B3LYP/MP2 and ONIOM/MP2 results agree reasonably well (see Table 3), except for the tetrafluorinated, nitro, and cyano derivatives. For instance, for the cyano derivative ONIOM predicts a quasiparallel stacking, while B3LYP suggests a T-shaped interaction. For the tetrafluorinated and nitro derivatives both methods predict π-stacking, but a more parallel arrangement is found in ONIOM optimizations (see Figure 8). The greater ability of the ONIOM procedure to introduce dispersion effects, which are not accounted for in B3LYP

(82) This is done by "chemical mutation" of the optimum conformation of compound II (with X = H) to the 4F, NO_2, or CN (ONIOM) derivative.

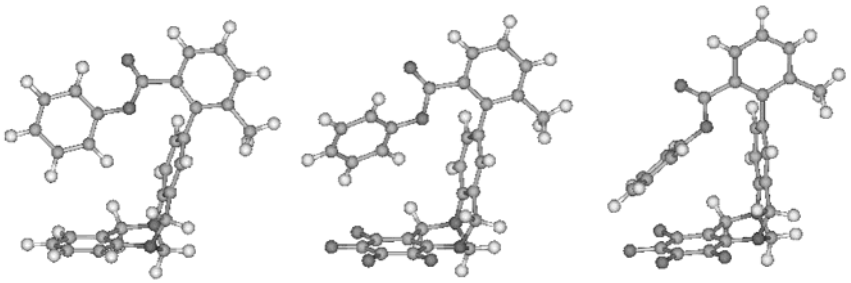

FIGURE 8. B3LYP optimized struture of compound II (left), B3LYP optimized geometry of the tetrafluorinated derivative (middle), and ONIOM optimized geometry of the tetrafluorinated derivative (right).

calculations, explains the difference between both types of optimizations for these compounds.

In summary, QM calculations are able to provide a reasonable description of Wilcox's experimental data, as they point out that (i) CH4-benzene and T-shaped benzene-benzene interactions are equally stable, (ii) the introduction of substituents (even the polar ones) in the acceptor benzene does not change the magnitude of the T-shaped benzene-benzene dimerization free energy. That is to say, current theoretical methods are able to provide a reasonably accurate representation of the folding process of Wilcox's compounds in solution. The surprising issue is, however, that these theoretical findings for Wilcox's molecules seem to disagree with our own previous analysis of model systems in the gas phase, which suggest that (i) T-shaped benzene-benzene interactions are stronger than CH4-benzene interactions, and (ii) the introduction of polar groups changes the magnitude of T-shaped interactions. Therefore, since the same methodology is used in both cases the apparent discrepancy might arise from the use of very different systems, thus precluding the extrapolation of results obtained for model systems in the gas phase to folding models in solution (and vice versa).

To analyze this effect in detail we recomputed the interaction energies at both MP2 and GMIPp levels for model systems consisting of the acceptor benzene and donor isopropyl (compound I) or benzene (compound II) groups placed at the same geometries found in Wilcox's compounds.[83] Results shown in Tables S4 and S5 (see Supporting Information) have to be compared with those obtained for the optimum gas phase conformations (Tables S2 and S3), as well as with gas-phase folding energies for the Wilcox's compounds (see Tables 2 and 3; B3LYP geometries)

The alkyl-π interaction is around 0.5 kcal/mol stronger when the CH_4 group is replaced by an isopropyl (see Table S4) as a result of the greater magnitude of dispersion forces. However, the changes in interaction energy due to the geometrical distortion of the interacting groups due to the covalent linkage of the folding model's backbone is negligible, as noted by comparing Tables S2 and S4. Interestingly, the folding energies in Table S2 are always more negative (from 0.5 to 1.4 kcal/mol) than values in Table S4. This suggests that the folding model's backbone is well designed for the alkyl-π interaction: it not only places the alkyl group at an optimum position for interaction but also provides additional interactions that stabilize the *cis* conformation with respect to the *trans*.

The effect of the folding model's backbone is less favorable for T-shaped benzene-benzene interaction. Thus, comparison of Tables S3 and S5 demonstrates that the optimum gas-phase T-shaped benzene-benzene interactions are 0.3−0.8 kcal/mol more stable than the corresponding interactions when the folding model geometries are used, mostly because of the electrostatic component (see Tables 3 and S5). These findings agree with the larger dependence of electrostatic interactions on orientation factors compared to dispersion interactions. Interestingly, the folding energies of derivatives of compound II (Table 3) are always less favorable than the corresponding dimerization energies (Table S5), which implies that the backbone's secondary interactions disfavor folding when the donor was benzene.

In summary, our results suggest that the backbone plays a nonnegligible effect in modulating the *cis* ↔ *trans* conformational equilibrium of Wilcox's compounds. Such an effect is favorable for the derivatives with isopropyl as donor but disfavorable for the derivatives with benzene owing to a 2-fold reason: (i) the benzene-benzene interaction has a stronger electrostatic component and is then more dependent on small backbone-related orientational changes, and (ii) the backbone's secondary interactions are favorable for the folding when isopropyl is the donor but destabilizing when the donor is benzene.

Conclusions

(i) High level ab initio calculations including correlation effects and simple QM/MM GMIPp computations are able to reproduce alkyl-π and aryl-π interactions not only in small model systems in the gas phase but also in large folding models in apolar solvents.

(ii) Dispersion effects dominate both alkyl-π and aryl-π interactions. However, electrostatic interactions are important at least for T-shaped benzene-benzene interactions. In all the cases alkyl-π are weaker than aryl-π interactions.

(iii) The introduction of polar substituents at the acceptor benzene introduces small changes in the methane-benzene dimerization energy but important alterations in the T-shaped benzene-benzene one. Such changes are clearly correlated with the differences in electrostatic interactions.

(iv) The effect of the backbone in the determination of the relative geometry of interacting groups and in the

(83) This is done by "removing all the backbone" of Wilcox's compounds, maintaining the relative orientation found in B3LYP optimizations. For the tetrafluoro and nitro derivatives also the perfect T-shaped orientation is considered.

formation of secondary interactions is not negligible and can change qualitatively the results. Extreme caution is then necessary before extrapolating conclusions derived from gas-phase studies to folding models and viceversa.

Acknowledgment. We thank Prof. J. Tomasi for a copy of his MST code, which was modified by us to perform AM1-MST calculations. We are also indebted to Prof. C. J. Cramer for a copy of his AMSOL code and helpful suggestions. This work was supported by the Centre de Supercomputació de Catalunya (CESCA; Molecular Recognition Project), and the Spanish Ministry of Science and Technology (PM99-0046 and PB98-1222).

Supporting Information Available: Optimized structures and energies of the eight more relevant Wilcox's compounds at the ONIOM and B3LYP levels (frequencies are also presented for B3LYP calculations). This material is available free of charge via the Internet at http://pubs.acs.org.

JO0201225