

Quantum/Classical Mechanical Comparison of Cation– π Interactions between Tetramethylammonium and Benzene

Clifford Felder,[†] Hua-Liang Jiang,^{†,§,||,⊥} Wei-Liang Zhu,^{§,||} Kai-Xian Chen,[§] Israel Silman,[‡] Simone A. Botti,^{†,‡} and Joel L. Sussman^{*,†}

Department of Structural Biology and Department of Neurobiology, Weizmann Institute of Science, 76100 Rehovot, Israel, Center for Drug Discovery and Design & Design and State Key Laboratory of Drug Research, Shanghai Institute of Materia Medica, Chinese Academy of Sciences, 294 Taiyuan Road, Shanghai 200031, People's Republic of China, and Chemical Process & Biotechnology Department, Singapore Polytechnic, 500 Dover Road, Singapore 139651

Received: August 11, 2000; In Final Form: December 4, 2000

To consider whether existing molecular force fields can adequately reproduce cation– π interactions without adding special interaction terms, theoretical calculations with geometry optimization were performed on three configurations of tetramethylammonium (TMA) interacting via one, two, or three *N*-methyl groups with a benzene ring, by use of density-functional theory (DFT) methods B3LYP/6-31G* and B3LYP/6-311G**, ab initio method MP2/6-31G*, and molecular mechanic methods EFF, Tinker's Amber and MM3. Only the first configuration was found to be stable from the DFT and MP2 results, and its geometry was found to be highly flexible. ESP CHELPG charges estimated from the DFT and MP2 calculations were used to modify the atomic charges of the force fields employed in the molecular mechanics calculations to improve agreement with the BSSE-corrected binding energies deduced from the DFT and MP2 results. After this modification, the molecular mechanics results were found to be in good agreement with those obtained by DFT and MP2, without a requirement to add any additional terms to the force fields. This was confirmed by comparing the energy profiles of the complex as benzene was moved away from TMA in 0.2 Å intervals. Hence it is possible to use existing force fields to represent cation– π interactions by a simple adjustment of certain partial atomic charge parameters. In this context, we discuss the high flexibility of the cation– π interactions in the framework of molecular recognition in biological systems.

Introduction

The interactions between cations and aromatic rings, generally referred to as cation– π interactions, and the forces involved, have been the focus of many investigations in the past decade due to their importance in molecular recognition.^{1–5} Extensive experimental and theoretical investigations have been performed with the objective of characterizing these interactions.^{3,5,6–13,14} Measurements of binding energies of cations to benzene and toluene have shown that, in the gaseous phase, cations bind preferentially to aromatic compounds rather than to water.¹⁵ Due to the sizable quadrupole moments of aromatic rings,¹⁶ it has been suggested that a pocket lined with the side chains of amino acids such as Trp, Phe, and Tyr can stabilize a positive charge as efficiently as full solvation by water, while still being compatible with the hydrophobic environment found in the interior of a protein molecule. Synthetic “hosts” with “walls” composed primarily of aromatic rings were found to bind acetylcholine (ACh) with affinities comparable to those of

natural ACh receptors.² Structural information concerning cation– π interactions between ligands and biological macromolecules has been provided by three-dimensional (3D) crystallographic structures, determined by Sussman and co-workers,^{4,6,17–19} for a series of complexes with cationic inhibitors of acetylcholinesterase (AChE), whose active-site gorge is lined by 14 aromatic residues. Analysis of the crystal structures of the edrophonium-AChE (PDB codes 2ACK and 1AX9), the phosphocholine-Fab McPC603 (PDB codes 2IMM and 2IMN), and the BW284C51-AChE (Felder et al., submitted) complexes revealed the same two most favorable orientations for an aromatic ring with respect to an R–N⁺(CH₃)₃ moiety observed by Verdonk et al.²⁰ through a statistical search of the small molecules in the Cambridge Structural Database (CSD).²¹ The rapid increase in the number of 3D structures deposited in the Protein Data Bank (PDB), coupled to data-mining techniques, has revealed the cation– π interaction to be widespread, not only in enzyme–ligand complexes, but also in the interiors of proteins.^{8,9} In a recent study,⁸ energy-based criteria were used to search a protein database composed of 593 proteins. It was found that >25% of all tryptophan residues experience an energetically significant cation– π interaction. Similarly, Minoux and Chipot⁹ analyzed a protein database containing 1718 representative structures for the association of phenylalanine, tyrosine, and tryptophan with arginine and lysine, and estimated close to 2500 cation–aromatic pairs.

Apart from computationally expensive ab initio quantum mechanical calculations, which are limited to small molecule

* Address correspondence to this author. Telephone: 972-8-934 2638.

FAX: 972-8-934 4159. E-mail: Joel.Sussman@weizmann.ac.il.

[†] Department of Structural Biology, Weizmann Institute of Science.

[‡] Department of Neurobiology, Weizmann Institute of Science.

[§] Chinese Academy of Sciences.

^{||} Singapore Polytechnic.

[⊥] Permanent address: Center for Drug Discovery and Design & Design and State Key Laboratory of Drug Research, Shanghai Institute of Materia Medica, Chinese Academy of Sciences, 294 Taiyuan Road, Shanghai 200031, People's Republic of China. Telephone: 86-21-64311833. FAX: 86-21-64370269. E-mail: jiang@iris3.simm.ac.cn or hljiang@mail.shnc.ac.cn.

complexes (up to 400 atoms, depending on the available computer resources), current macromolecular computational methods fail to account adequately for cation- π interactions when modeling the binding of ligands to their biological target. The importance of nonadditive effects in the accurate representation of cation- π interactions was shown in a study¹³ that used a version of the Amber force field,²² modified by employing a three-body exchange repulsion potential,^{23,24} to calculate alkali metal cation- $C(sp^2)_2$ interactions. A restrained electrostatic potential fit charge model (RESP)^{25,26} was used to describe the ammonium-benzene interaction by adding a new term to the electrostatic potential function for benzene to emphasize its polarizability, and thereby to mimic better the cation- π interaction. This improved force field successfully reproduced the binding energies of alkali metal cations and ammonium ions complexing with benzene. In another study, the CHARMM force field parameters²⁷ were modified²⁸ by the addition of empirical correction terms of the form $A/d^{n_{ij}}$ (where A is a constant, d_{ij} the ion to atom distance, and n an integer between 1 and 12) to describe K^+ -amino acid interactions and K^+ - π interactions. These empirical corrections are determined by a multiple linear regression analysis fit to density functional theory calculations (DFT).²⁹ Similarly, Minoux and Chipot⁹ used Lennard-Jones-like empirical correction terms (with $n = 4-12$) to improve the Amber force field potential in the determination of cation- π interactions. However, the validity of a nonadditive force field for modeling large molecular assemblies remains problematic because it is time-consuming, especially when it comes to the application of these force fields to docking programs employed in database screening. It is thus of significant interest to see whether it would be possible for existing additive force fields to accurately reproduce cation- π interactions, despite their lacking an explicit induction potential term. By suitable determination of their standard EFF van der Waals parameters, it was shown that one could obtain strong hydrogen bonding between amide NH moieties and carbonyl moieties, without addition of any hydrogen bond terms.³⁰ It was later demonstrated that a simple electrostatic model was adequate to describe iron(III) binding to three catecholate ligands,³¹ a model that was successfully extended to hydroxamate and other ligands, provided that they displayed octahedral or tetrahedral coordination geometry. We propose, by analogy, that cation- π interactions might also be reproduced accurately by current force fields by suitable modification of certain parameters, mainly the atomic charges, without the addition of extra terms. In other words, we suggest that existing molecular mechanics methods, which are empirical in nature, should be able implicitly to represent cation-aromatic interactions in their present formulation.

In this study we consider the capability of current molecular mechanics force fields to represent properly the cation- π interactions between the tetramethylammonium cation (TMA) and benzene. Toward this end we examined three idealized configurations of the TMA-benzene complex in which either one, two, or three *N*-methyl groups of TMA interact with the benzene ring (Figure 1). We applied DFT³² and Møller-Plesset second-order correlation (MP2) methods,³³ as implemented in the Gaussian98 software package (<http://www.gaussian.com>).³⁴ The basis sets employed are 6-31G* and 6-311G**, including full geometry optimization and calculation of the atomic charges. The results obtained are used to adjust the parameters for, and to compare the minimized energies and geometries from, three force fields, EFF,^{35,36} and versions of Amber²² and MM3,³⁷ as implemented by Tinker.^{38,39}

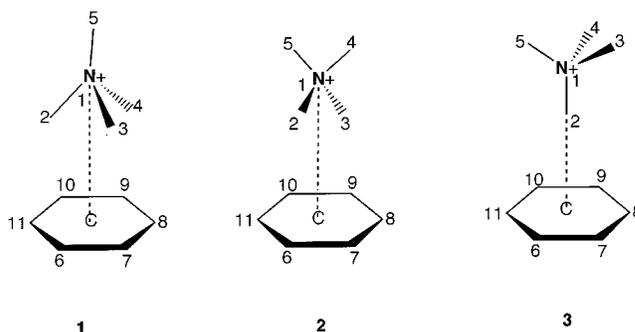


Figure 1. Three idealized configurations of tetramethylammonium-benzene complexes (C represents the center of the benzene ring).

Methods

Initial coordinate sets were prepared for three idealized configurations of the TMA-benzene complex (Figure 1): (1) three hydrogens from each of the three *N*-methyl groups interact with the benzene ring; (2) two hydrogens from each of the two *N*-methyl groups interact; (3) a single *N*-methyl group interacts. The DFT B3LYP parameter hybrid method was employed to optimize these structures fully at the 6-31G* basis set level, using the correlation functional.⁴⁰ The optimized structures were subsequently subjected to further optimization by the B3LYP/6-311G** and MP2/6-31G* methods. Binding energies were quantified by performing separate calculations on TMA and benzene alone. Frequency calculations were carried out at the same levels of B3LYP, based on the optimized geometries, so as to verify the reasonableness of the optimized structures, and to determine the zero-point and vibrational energies and, thus, the enthalpies and entropies. Due to computational limitations, we did not perform frequency calculations using the MP2 method. However, based on the experience of ourselves^{10,11} and others,⁴¹ the results of frequency calculations using B3LYP are as accurate as those obtained employing MP2. Accordingly, the results of the B3LYP frequency calculations were used to estimate the binding energies (ΔH) and free energy differences (ΔG) for the MP2 method. Due to the basis set superposition error (BSSE), ΔH values calculated by the DFT and MP2 methods are always overestimated.¹² To obtain reliable thermodynamic parameters for parametrizing our improved force fields, we also performed a BSSE correction, using⁴²

$$\text{BSSE} = [E_A - E_{A(\text{AB})}] + [E_B - E_{B(\text{AB})}] \quad (1)$$

where $E_{A(\text{AB})}$ (or $E_{B(\text{AB})}$) is the energy of fragment A (or B), based on the geometry extracted from the optimized structure, with its own basis set augmented by the basis set of B (or A). E_A (or E_B) is the energy of isolated fragment A (or B), with just its own basis set. ESP CHELPG (ElectroStatic Potential Charges from Electrostatic Potentials Generalized) charges⁴³ and natural population analysis (NPA) charges⁴⁴ were calculated based on the optimized structures of various configurations of the TMA-benzene complex, so as to parametrize the atomic charges in the force field. To reveal the binding behavior of TMA complexing with benzene, an additional series of BSSE-corrected MP2/6-31G* structure optimizations was performed, in which the TMA nitrogen to benzene centroid distance was fixed at given values that were increased in 0.2 Å increments.

Molecular mechanics calculations were performed for all three configurations of the TMA-benzene complex, using three separate force fields, EFF,^{35,36} Amber,²² and MM3.³⁷ The versions of Amber and MM3 employed were the ones implemented in the program Tinker (<http://dasher.wustl.edu/tinker>).^{38,39}

TABLE 1: Distances between the Nitrogen Atom of TMA and the Center of Benzene and C–C and N–C Bond Lengths Obtained from Quantum Chemical Calculations (Å)

methods		config 1	config 2	config 3	TMA	benzene
B3LYP/6-31G*	R ^a	4.38	4.53	4.96		
	C–C	1.40	1.40	1.40		1.40
	N–C	1.51	1.51	1.51	1.51	
B3LYP/6-311G**	R ^a	4.43	4.54	4.97		
	C–C	1.40	1.40	1.40		1.40
	N–C	1.51	1.51	1.51	1.51	
MP2/6-31G*	R ^a	4.24	4.37			
	C–C	1.40	1.40			1.40
	N–C	1.50	1.50		1.50	
MP2/6-31G* ^b	R ^a	4.23				
MP2/6-31G ^{ααc}	R ^a	4.18	4.33			
MP/6-311++G** ^d	R ^a	4.21				

^a N–centroid distances.

Based on the partial atomic charges calculated with DFT and MP2, modified parameter sets were prepared in which the partial atomic charges on the atoms were modified. As a control, the calculations were repeated using a second set of charges for EFF, as well as the original parameter sets for all three force fields. Finally, a series of energy calculations without minimization was made in which TMA was moved away from the benzene in 0.2 Å increments. Linear correlation analyses were performed comparing the DFT, MP2, EFF, Amber, and MM3 energy curves.

Results and Discussion

Quantum Chemistry Calculations. Geometry. Table 1 summarizes the distances between the N atom of TMA and the centroid of the benzene ring (N–centroid distances) and the bond lengths of the C–C bonds and N–C bonds. For configuration 1, the N–centroid distances derived from B3LYP/6-31G*, B3LYP/6-311G**, and MP2/6-31G* are 4.38, 4.43, and 4.24 Å, respectively. The MP2/6-31G* result is in agreement with that of one study¹³ at the same theoretical level, but about 0.06 Å longer than that of another.¹⁴ For configuration 2, the N–centroid distances calculated by the same three methods are 4.53, 4.54, and 4.37 Å, respectively, with the MP2 distance about 0.4 Å longer than that of Pullman and co-workers.¹⁴ This difference may be due to the different basis sets employed: whereas Pullman and co-workers used the modified 6-31G basis set 6-31G^{αα}, we used the standard Gaussian 6-31G* basis set. For the DFT results, the N–centroid distances are, in general, longer than those obtained by MP2. In particular, adding the polarization function to the basis set results in an increase in the interaction distance between TMA and benzene (Table 1). This is in contrast with the result for optimization of the NH₄⁺–benzene complex, where addition of the polarization function to the basis set results in a decrease in binding distance.^{11,45} The optimized geometries for all three configurations also show that complexation of TMA with benzene results in the hydrogen atoms of the benzene ring bending out of the plane of the aromatic ring away from TMA by about 2°. This may be ascribed to repulsion between these hydrogens and TMA, and was also observed in the NH₄⁺–benzene complex.⁴⁵ Compared to the free structures of TMA and benzene, the lengths of the C–C, N–C, and C–H bonds do not change significantly in the TMA–benzene complex (Table 1).

DFT methods employing both B3LYP/6-31G* and B3LYP/6-311G** recognized a stable local minimum energy structure for configuration 3, whereas MP2/6-31G* did not. Figure 2 shows the trajectory of the optimization process using MP2/6-31G*, starting from configuration 3 and converging into

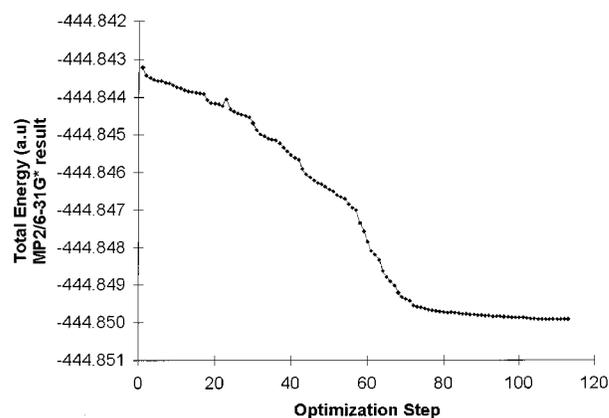


Figure 2. Trajectory of the MP2/6-31G* optimization process from configuration 3 to configuration 1.

configuration 1, with no energy barrier between them. This indicates that configuration 3 is not really a stable structure, as it can shift smoothly to configuration 1, depending on the method but not on the basis set. To investigate the dynamic properties of this shift, we isolated some arbitrary “intermediate” structures from the optimization trajectory (red squares in Figure 2), which are presented in Figure 3. The structures shown in Figure 3 indicate a simple pathway from configuration 3 to configuration 1: plane C2–N1–C5 of TMA (see Figure 1) turns up and around the N–centroid axis to the left, with the entire TMA moving initially to the right of the benzene ring and then back to the left. This observation is in good agreement with Cambridge database search results,²⁰ which indicated that R–N⁺(CH₃)₃ can bind to a phenyl group in any of these three configurations, or in any intermediate configuration between them. Both the results of our theoretical calculations employing DFT and MP2 and the database searches^{8,9,20} indicate that organic cation interactions with aromatic systems are remarkably flexible with respect to orientation. This is in marked contrast to such interactions as hydrogen bonding and salt bridges, whose interaction geometries and interatomic contact distances are quite rigid.

Stabilities of the Three Configurations. Table 2 lists the total and relative energies of the three configurations of the TMA–benzene complex. Among the three, configuration 1 is the most stable. Configuration 2 is less stable by 1.78 kcal/mol for B3LYP/6-31G*, by 1.30 kcal/mol for B3LYP/6-311G**, and by 2.21 kcal/mol for MP2/6-31G*. Configuration 3 is less stable than configuration 1 by 2.84 kcal/mol for B3LYP/6-31G* and by 1.89 kcal/mol for B3LYP/6-311G**. Frequency calculations on the optimized structures of the three configurations, using the two DFT methods (B3LYP/6-31G* and B3LYP/6-311G**),

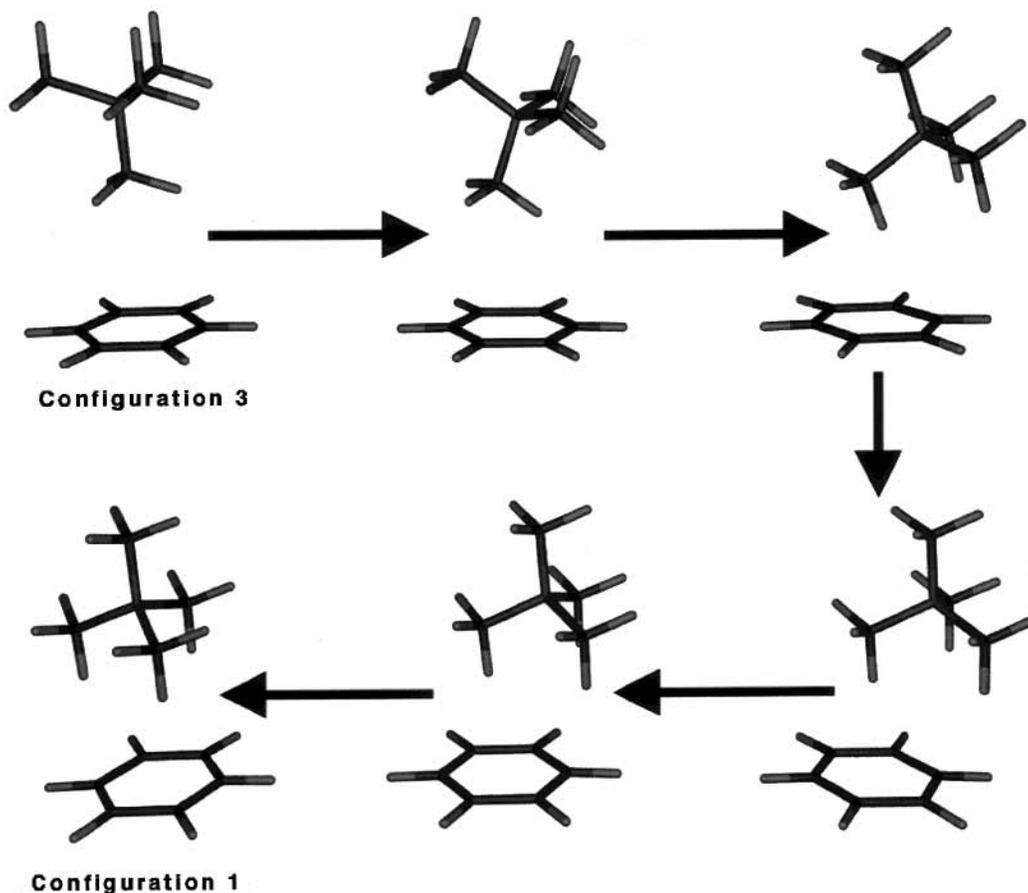


Figure 3. Pathway of configuration 1 changing to configuration 3. The structures correspond to the diamonds in Figure 2.

TABLE 2: TMA-Benzene: Total Energies (au) and Energy Differences Relative to the Most Stable Configuration 1 (kcal/mol) of the Calculated Complexes

methods	config 1	config 2	config 3
B3LYP/6-31G*	-446.425 918	-446.423 082	-446.421 386
relative energy	0.0	1.78	2.84
B3LYP/6-311G**	-446.541 456	-446.539 377	-446.538 464
relative energy	0.0	1.30	1.89
MP2/6-31G*	-444.849 934	-444.846 401	
relative energy	0.0	2.21	

showed that, for configuration 1, no negative frequency exists, indicating that the final optimized structure of this configuration is the minimum energy structure. However, for configurations 2 and 3, there are some negative frequencies. Reoptimizing these structures by reading off the force constants calculated from the previous calculation did not remove the negative frequencies of configurations 2 and 3. This further demonstrates that these two configurations are not true minimum energy structures. We did not therefore calculate their thermodynamic energy parameters. However, we can infer from the optimization trajectory of configuration 3 (Figure 2) that the potentials around configurations 2 and 3 are very flat, which means that there is no one precisely defined structure for these lowest energy minima.

Charge Transfer. ESP CHELPG charges were calculated on the optimized structures of the three configurations of the TMA-benzene complex using both the DFT and MP2 levels of theory. This was done in order to study the importance of charge transfer in the binding of cations to aromatic rings, as well as to help assign the atomic charges for the improved force fields. Each binding configuration was divided into two parts, an aromatic term and a TMA term, with the intention of

TABLE 3: Net Molecular Charges on TMA and Benzene Determined as the Sum of the CHELP Partial Atomic Charges of Their Calculated Complexes

	system	B3LYP/6-31G*	B3LYP/6-311G**	MP2/6-31G*
config 1	TMA	0.850 89	0.847 73	0.856 82
	benzene	0.149 11	0.152 27	0.143 18
config 2	TMA	0.883 72	0.870 94	0.884 68
	benzene	0.116 28	0.129 06	0.115 33
config 3	TMA	0.905 87	0.891 76	
	benzene	0.094 14	0.108 24	

investigating possible charge transfer during complex formation. Table 3 summarizes the total atomic charges of these two parts for each configuration.

For configurations 1, 2, and 3, charge transfer from TMA to the benzene ring is ~ 0.15 , ~ 0.12 , and ~ 0.10 au, respectively (Table 3). The larger the charge transfer, the more stable the configuration and the greater the binding energy. This is in agreement with the conclusion we derived from calculations for NH_4^+ -aromatic systems that, in addition to the electrostatic interaction, charge transfer also makes a major contribution to cation-aromatic interactions.^{11,45}

Binding Energy and the BSSE Effect. Table 4 presents the calculated thermodynamic parameters for configuration 1 at the B3LYP/6-31G*, B3LYP/6-31G**, and MP2/6-31G* levels, including the raw binding energy, E_{bind} , the change in thermal energy, E_{thermal} , the change in entropy, ΔS , the heat of formation, ΔH , and the change in free energy, ΔG . Details of the computations are presented in the footnotes to Table 4. The binding energy (E_{bind}) for the MP2/6-31G* level is -11.07 kcal/mol, very close to the result of Kim et al.¹² calculated by MP2/6-311+G** (-11.72 kcal/mol). The BSSE-corrected binding energy for the MP2/6-31G* level is -8.4 kcal/mol, which is

TABLE 4: Binding Energy of Configuration 1

methods	B3LYP/6-31G*	B3LYP/6-311G**	MP2/6-31G* ^j
E_{bind}^a	-7.583	-6.406	-11.073
$\Delta E_{\text{thermal}}^b$	2.025	2.039	
ΔS^c	-29.979	-27.574	
ΔH^d	-6.148	-4.947	-9.638
ΔG^e	2.786	3.260	-0.70
BSSE ^f	1.382	0.641	2.673
$E_{\text{corr}}^{\text{bind } g}$	-6.201	-5.765	-8.400
$\Delta H_{\text{corr}}^{\text{bind } h}$	-4.766	-4.306	-6.965
ΔH_{exp}^i	-9.40		

^a Binding energy (kcal/mol): $E_{\text{bind}} = E_{\text{total}}(\text{complex}) - E_{\text{benzene}} - E_{\text{TMA}}$. ^b Thermal energy change (kcal/mol): $\Delta E_{\text{therm}} = E_{\text{therm}}(\text{complex}) - E_{\text{inter}}(\text{benzene}) - E_{\text{therm}}(\text{TMA})$. ^c Entropy change (cal/(mol·K)): $\Delta S = S(\text{complex}) - S(\text{benzene}) - S(\text{TMA})$. ^d Heat of formation (kcal/mol): $\Delta H = \Delta E_{\text{bind}} + \Delta E_{\text{therm}} + \Delta(PV)$. ^e Free energy change (kcal/mol): $\Delta G = \Delta H - T\Delta S$. ^f Basis set superposition error correction (kcal/mol), calculated by eq 1. ^g BSSE-corrected binding energy (kcal/mol): $\langle \text{abv} \rangle E_{\text{bind}}^{\text{corr}} \langle \text{blw} \rangle = E_{\text{bind}} + \text{BSSE}$. ^h BSSE-corrected heat of formation (kcal/mol): $\Delta \langle \text{abv} \rangle H_{\text{bind}}^{\text{corr}} \langle \text{blw} \rangle = \Delta H_{\text{bind}} + \text{BSSE}$. ⁱ Experimental value of heat of formation. ^j The binding energy (E_{bind}) was calculated at the MP2/6-31G* level, and thermodynamic quantities were calculated using B3LYP/6-31G* vibrational frequencies.

also close to the value of -8.67 kcal/mol calculated previously, with the geometry optimized by MP2/6-311+G** and the frequency calculated by MP2/6-31G.¹² The MP2/6-311++G** method overestimated the N-centroid distance by about 0.02 Å, which increased the BSSE value. This is why our BSSE-corrected MP2/6-31G* binding energy is much closer to that from MP2/6-311++G**.¹²

The uncorrected binding enthalpy (ΔH) of our MP2 result is -9.64 kcal/mol, in good agreement with the experimental value of -9.4 kcal/mol.⁴⁶ However, the BSSE-corrected value for ΔH , -6.97 kcal/mol, is much lower than the experimental value. Thus, the DFT result of the binding energy may be underestimated, compared to both the MP2 results and the experimental data. However, Hoyau et al.⁴⁷ recently pointed out that binding enthalpies obtained either by high pressure mass spectrometry (HPMS) or by threshold collision-induced dissociation (CID) techniques are normally overestimated by up to 5 kcal/mol. They argue that enthalpies calculated with a medium-size basis set, against which the experimental values are compared generally, become much smaller when using an extended basis set and BSSE correction, conditions which they showed were able to reproduce the most accurate enthalpies for Na⁺ interacting with a number of simple organic compounds. This might also be true for the TMA-benzene complex. If so, both our DFT and MP2 results are in agreement with the experimental value.

Molecular Mechanics Calculations. *Parameter Determination.* The ESP CHELPG and natural population (NP) methods yielded very different calculated partial atomic charges (see Table 5) for the same geometry optimization method, though each charge determination method gave similar results for the different optimization methods. The CHELPG charges appear to be closer to values currently used by the force fields. In particular, the CHELPG charges for benzene are practically the same as those already present in the EFF and Amber force fields, meaning that only the TMA charges should be modified. The CHELPG charges for TMA are also more similar to those currently employed, while NP charges place a large negative charge on N. We chose, therefore, to work with the CHELPG charges (last row of Table 5), although as a check we also did an EFF minimization with a negative charge on N, as in the NP charges. Most current force fields do not include any explicit terms for cation- π interactions.

Energy Minimizations. The optimized coordinate sets were transferred to the EFF and Tinker force fields for minimization. Special modified parameter files were prepared for the EFF, Amber, and MM3 force fields, in which the partial atomic charges on the quaternary amine were changed to +0.28 for N, -0.30 for C, and +0.16 for H (in place of the standard EFF alkane: C, -0.33; H, +0.11), and those on benzene were left at their standard EFF values: C, -0.11; H, 0.11. For Amber, the benzene atomic charges were also changed to these values, so that all atoms would be parametrized in the same way. In the standard Amber force field, there is some variation in the values of the phenyl atomic charges. For MM3 these explicit atom charges replaced the bond dipole parameters normally used in this force field, since bond dipole-dipole interactions are a much poorer representation of electrostatic interactions than charge-charge interactions and give much worse agreement with experiment. To help justify our choice of CHELPG over NP charges, the EFF minimizations were repeated with a second set of atom charges, i.e., with -0.12 for N, -0.20 for C, and 0.16 for H in TMA, and, as a control, those of all three force fields were also done using their respective original parameters. These calculations were also performed with uniform sets of atom charges for Amber and MM3 (see footnotes to Table 6).

The results of these calculations are presented in Table 6, where they are compared with the quantum chemical results. It can be seen that (1) the calculated EFF and Amber binding energies, although lower than the experimental and MP2 values, are within the range of the DFT values (for MM3 they are smaller). These binding energies are affected little by the atom charges for the EFF force field, but much more so in the case of the Amber and MM3 force fields. In the case of MM3, this is probably due to use of bond dipoles rather than explicit charges. (2) The calculated energy differences between the first two configurations are somewhat smaller than those obtained by DFT and, particularly, by MP2. The Amber results agree best with the quantum chemistry calculations. The Amber binding energy for configuration 1 was close to that obtained with MP2, while the EFF energies were closer to the DFT values. Furthermore, the energy difference between the first two configurations, as determined with Amber, was almost as large as that obtained with DFT, while with the other force fields this value was much smaller. With all three force fields, using their unmodified or original charges led to all three configurations having practically the same energy, a worse agreement with experiment than when employing the modified charges. (3) The short contacts are slightly shorter by ca. 0.1–0.2 Å than those obtained from DFT, but in the range of the MP2 values. Furthermore, these distances are hardly affected by the atom charges used. (4) In all the force fields configuration 3 shifted into configuration 1 during minimization, indicating both that it is unstable and that there are no energy barriers or high-energy states between the two configurations, in agreement with the MP2 results and normal mode data for DFT described above. (5) The original MM3 parameters yielded a binding energy that was much too low and distorted geometries of the complexes. Apparently, use of bond dipoles rather than atomic charges prevented MM3 from being able to handle this kind of problem.

Minoux and Chipot⁹ suggested a cation- π correction term in their eq 2, $946511r^{-12} - 144.355r^{-4}$ (where r is the cation N-centroid distance), which, for $r = 4.38$ Å (from Table 1), is -2.2 kcal/mol. This value is smaller than the difference between the binding energies calculated by DFT and MP2, and also between those calculated by Amber using the original and modified atomic charges (see Table 6). Their correction thus

TABLE 5: Partial Atomic Charges Obtained from Quantum Calculations and Used for the Force Field Calculations^a

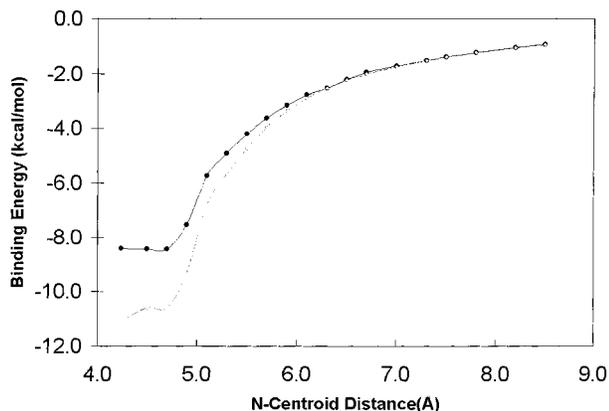
method	ESP CHELPG					natural population analysis				
	TMA			benzene		TMA			benzene	
	N	C	H	C	H	N	C	H	C	H
DFT/6-31G*	0.28	-0.30	0.16	-0.08	0.08	-0.30	-0.47	0.27	-0.08	0.08
DFT/6-311G**	0.26	-0.30	0.16	-0.08	0.08	-0.31	-0.35	0.23	-0.08	0.08
MP2/6-31G*	0.28	-0.31	0.16	-0.09	0.11	-0.37	-0.41	0.25	-0.24	0.24
used for FF	0.28	-0.30	0.16	-0.11	0.11			not used ^b		

^a The values presented are average values, calculated by different methods, for the TMA-benzene complex in configuration 1. ^b The CHELPG charges rather than the natural values were used to assign the atomic charge parameters, as described under Molecular Mechanics Calculations.

TABLE 6: Summary of Calculated Results for TMA-Benzene

method	binding energy (kcal/mol)	energy difference (kcal/mol)		short contact distances ^a (Å)	
		configs 1-2	configs 1-3	config 1	config 2
DFT/6-31G* ^b	-6.2	1.2	2.8	2.78 (3)	3.1 (4)
DFT/6-311G* ^b	-5.8	1.1	1.9	2.84 (3)	3.0-3.1 (4)
MP2/6-31G* ^b	-8.4	2.2	→config 1 ^{e,f}	2.64 (3)	2.9-3.1 (4)
EFF ^c	-6.1	0.2	→config 1 ^{e,f}	2.68 (3)	2.78 (4)
EFF no. 2 ^d	-5.9	0.04	→config 1 ^{e,f}	2.70 (3)	2.79 (4)
EFF orig ^e	-5.9	0.02	→config 1 ^{e,f}	2.69 (3)	2.78 (4)
Amber ^c	-7.8	0.7	→config 1 ^{e,f}	2.56 (3)	2.73 (4)
Amber orig ^g	-4.7	0.02	→config 1 ^{e,f}	2.59 (3)	2.75 (4)
Tinker/MM3 ^c	-4.5	0.1	→config 1 ^{e,f}	2.80 (3)	3.03 (4)
Tinker/MM3 orig ^g	-2.0	0.03		complexes come out distorted	
exptl enthalpy	-9.4	N/A	N/A		

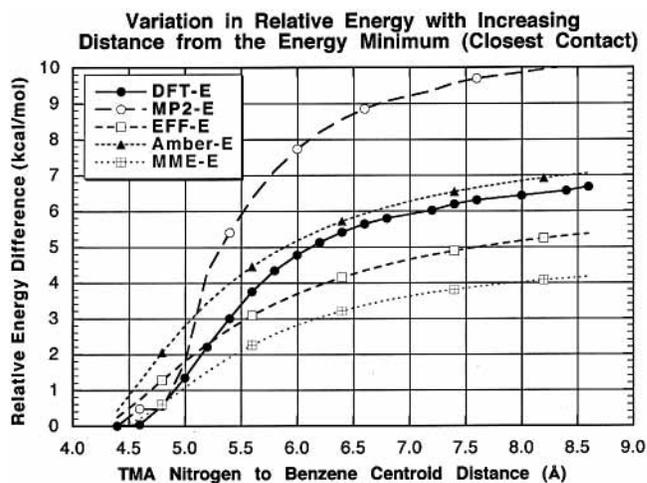
^a Between an *N*-Me H and a benzene C. The number in parentheses gives the number of contacts. ^b Includes BSSE quantum correction term. ^c Uses atom charges. For TMA: N, 0.28; C, -0.30; H, 0.16. For benzene: C, -0.11; H, 0.11. ^d Uses atom charges. For TMA: N, -0.12; C, -0.20; H, 0.16. For benzene: C, -0.11; H, 0.11. ^e Uses atom charges. For TMA: N, 0.20; C, -0.13; H, 0.16. For benzene: C, -0.11; H, 0.11. ^f During minimization, config 3 shifted to config 1. ^g Uses original force parameter set. Note that the net charge is not exactly 1.0.

**Figure 4.** MP2 binding energy curve varies with N-centroid distance (the upper curve is uncorrected, and the lower curve is BSSE corrected).

appears to offer hardly any improvement over the corrections to the atomic charges that we have implemented. We conclude that existing force fields, with suitably modified atomic charges, are able to handle the cation- π interaction satisfactorily. Furthermore, molecular mechanics parameters and calculations for cation- π interactions in the gas phase should be applicable in protein interiors, which are generally assumed to be fairly hydrophobic.^{48,49}

Calculated Energy Comparison as TMA Moves Away from Benzene. To study the dependence of binding energy on N-centroid distance, we optimized configuration 1 complexes with MP2/6-31G* for various fixed N-centroid distances, at 0.2 Å increments. We calculated the total and binding energies for each complex, as well as their BSSE corrections. The results are presented in Figure 4.

The binding energy of configuration 1 decreases with increasing N-centroid distance to about -1.2 kcal/mol at 7.8 Å. It is interesting that the BSSE also decreases when the

**Figure 5.** Variation in relative energy with increasing distance from the energy minimum (closest contact).

N-centroid distance becomes large, and becomes zero at 6.3 Å, suggesting that BSSE is a short-distance phenomenon. Figure 5 shows curves for binding energy vs N-centroid distance, from which we can see that the curves are flat in the range 4.24-4.70 Å. The BSSE correction enhances this phenomenon. Thus, even for the most stable configuration of the TMA-benzene complex, binding of TMA to benzene is quite flexible, inasmuch as the final N-centroid distance can assume any value in the 4.24-4.70 Å range. This finding explains the large variation found for N-centroid distances both in small molecules²⁰ and in proteins,^{8,9} since with such a flexible interaction both cations and aromatic moieties can adjust the equilibrium interaction distance in response to the environment they encounter in the process of binding.

Simple energy calculations, without either optimization or minimization, were also performed for the DFT (BSSE-

TABLE 7: Correlation Coefficients for Shifts in Energy as TMA Moves Away from Benzene

	MP2	DFT
DFT	0.995	
EFF	0.980	0.990
Amber	0.979	0.984
MM3	0.988	0.997

corrected) method and for the three molecular mechanics methods, with TMA being pulled away from the benzene ring by 0.2 Å increments along the TMA N-centroid axis. The resulting curves, shown in Figure 5, show an overall similar decrease in the binding energy with increasing separation of TMA from benzene for all five methods. However, at distances of ~0.2–0.6 Å from the point of closest contact, the shifts obtained by the DFT and MP2 methods are sharper than those obtained using molecular mechanics (Figure 5). Nevertheless, the magnitude of this effect is smaller than the disagreement between the binding energy values obtained by DFT and MP2. In addition, we made series of linear regression analyses between pairs of energy differences for the various methods. All the correlation coefficients were ≥ 0.98 , as shown in Table 7. These results indicate that the molecular mechanics results reproduce with a good degree of accuracy the binding behavior of TMA to benzene, despite the small discrepancies at close distances, because they agree strongly with the quantum mechanical calculations.

Biological Significance of Configurational Flexibility. A preliminary search of the Cambridge Database²¹ revealed a broad continuum in the N-centroid distances up to 5.2 Å and binding geometries in quaternary amino to aromatic interactions²⁰ (Felder et al., unpublished), spanning not only our three configurations but also many intermediate states. As these data included substituted quaternary amines and various aromatic groups such as tyrosine and tryptophan, our findings might be valid more generally. The fact that these searches included a wide spectrum of substituted quaternary amines and aromatics supports the general validity of our findings. Accordingly, we can advance the hypothesis that ligands with quaternary ammonium moieties, when binding to receptor sites lined with aromatic side chains, exploit both the specificity of the cation- π interaction and the ability to adopt several isoenergetic orientations with respect to the aromatic moieties. As an example, we can mention the process of binding of ACh to the active site of AChE, which is assumed to progress through a series of way stations composed of aromatic residues from the entrance along the length of its active-site gorge.^{4,50} A recent high-resolution study of the structure of the nicotinic ACh receptor (nAChR) revealed a series of narrow tunnels ~10–15 Å long connecting the ACh-binding cavities within the α subunits to the water-filled vestibule of the channel. It appears plausible that ACh is drawn into the vestibule of the channel⁵¹ along with other cations, but may be selectively guided into the two binding pockets by a combination of electrostatic effects and interactions of the hydrophobic quaternary ammonium group with aromatic residues lining the tunnel, in a fashion analogous to that proposed for AChE. We can thus envision a general mechanism by which ACh could adopt several isoenergetic orientations by interaction of one, two, or three methyls of its cationic moiety with the aromatic rings that line the active-site gorge and tunnels of AChE and the nAChR (Figure 3). This would allow ACh to reach its binding site smoothly and adopt its optimal and final orientation through the interaction of all of its three *N*-methyls with the aromatic moieties of residues such as W84 in the active site of AChE.⁴ We plan to extend our quantum and molecular

mechanical calculations to other amines and aromatic ring systems, to test the general validity of our calculations.

Conclusions

In this study, we have employed DFT and MP2 at the 6-31G* and 6-311G** basis set levels to investigate the interaction between TMA and benzene, and to deduce binding geometries and ESP CHELPG atomic charges in order to modify the atomic charge parameters of existing molecular mechanics force fields. This approach was used to test whether such modified force fields can accurately reproduce cation- π interactions in proteins, without the addition of computationally expensive nonlinear terms. We find that molecular mechanics methods can adequately and accurately describe this type of interaction, after a small adjustment to the atom charges of the cationic moiety, without the introduction of special correction terms. This means that one could take any force field off the shelf and, by suitable modification of the atomic charges based on calculated CHELPG charges, enable it to represent cation- π interactions without reprogramming.

The three TMA-benzene configurations that we have examined were all found to have about the same energy, and can readily interconvert since the energy barriers between them are very low. Moreover, over a wide range of intergroup distances near the point of closest contact, the energy hardly changes. The consequent unusual geometric flexibility of this kind of interaction, in comparison with other interatomic interactions such as hydrogen bonding and salt bridges, whose geometry is quite rigid, thus allows the quaternary ammonium-aromatic interaction to assume a broad repertoire of positions, interatomic distances, and orientations, all with practically the same energy. This allows cationic species to bind to an aromatic-rich region of a protein or surface in a variety of different positions and orientations relative to the aromatic moieties, which is in agreement with the CSD searches and AChE complexes mentioned in the Introduction. The binding of cationic species such as ACh to AChE, and to nAChR, are good examples of systems that can take advantage of the nonspecific nature of the cation- π interaction to allow the ligand to bind within a gorge or deep pocket lined with aromatic residues.

Acknowledgment. This work was supported by the U.S. Army Medical and Materiel Command under Contract No. DAMD17-97-2-7022, the EU Fourth Framework Program in Biotechnology, the Kimmelman Center for Biomolecular Structure and Assembly (Rehovot, Israel), and the Dana Foundation. I.S. is Bernstein-Mason Professor of Neurochemistry. The quantum chemistry calculations were performed on the SGI Origin 2000 at the Weizmann Institute of Science, Rehovot, Israel, and on the SGI Power Challenge R10000 at The Network Information Center, Chinese Academy of Sciences, Beijing, People's Republic of China. H.-L.J. and K.-X.C. gratefully acknowledge financial support from the National Natural Science Foundation of China (Grant 29725203), the "863" High-Tech Program of China (Grant 863-103-04-01), and the State Key Program of Basic Research of China (Grant 1998051115).

References and Notes

- (1) Shepodd, T. J.; Petti, M. A.; Dougherty, D. A. *J. Am. Chem. Soc.* **1988**, *110*, 1983–1985.
- (2) Dougherty, D. A.; Stauffer, D. A. *Science* **1990**, *250*, 1558–1560.
- (3) Dougherty, D. A. *Science* **1996**, *271*, 163–168.
- (4) Sussman, J. L.; Harel, M.; Frolow, F.; Oefner, C.; Goldman, A.; Toker, L.; Silman, I. *Science* **1991**, *253*, 872–879.
- (5) Ma, J. C.; Dougherty, D. A. *Chem. Rev.* **1997**, *97*, 1303–1324.

- (6) Ravelli, R. B. G.; Ravest, M. L.; Ren, Z.; Bourgeois, D.; Roth, M.; Kroon, J.; Silman, I.; Sussman, J. L. *Acta Crystallogr.* **1998**, *D54*, 1359–1366.
- (7) Roelens, S.; Torriti, R. *J. Am. Chem. Soc.* **1998**, *120*, 12443–12452.
- (8) Gallivan, J. P.; Dougherty, D. A. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 9459–9464.
- (9) Minoux, H.; Chipot, C. *J. Am. Chem. Soc.* **1999**, *121*, 10366–10372.
- (10) Tan, X. J.; Jiang, H. L.; Zhu, W. L.; Chen, K. X.; Ji, R. Y. *J. Chem. Soc., Perkin Trans. 2* **1999**, *1999*, 107–111.
- (11) Zhu, W. L.; Jiang, H. L.; Puah, C. M.; Tan, X. J.; Chen, K. X.; Cao, Y.; Ji, R. Y. *J. Chem. Soc., Perkin Trans. 2* **1999**, *1999*, 2615–2622.
- (12) Kim, K. S.; Lee, J. Y.; Lee, S. J.; Ha, T. K.; Kim, D. H. *J. Am. Chem. Soc.* **1994**, *116*, 7399–7400.
- (13) Caldwell, J. W.; Kollman, P. A. *J. Am. Chem. Soc.* **1995**, *117*, 4177–4187.
- (14) Pullman, A.; Berthier, G.; Savinelli, R. *J. Comput. Chem.* **1997**, *18*, 2012–2022.
- (15) Sunner, J.; Nishizawa, K.; Kebarle, P. *J. Phys. Chem.* **1981**, *85*, 1815–1820.
- (16) Luhmer, M.; Bartik, K.; Dejaegere, A.; Bovy, P.; Reisse, J. *Bull. Soc. Chim. Fr.* **1994**, *131*, 603–606.
- (17) Harel, M.; Schalk, I.; Ehret-Sabatier, L.; Bouet, F.; Goeldner, M.; Hirth, C.; Axelsen, P. H.; Silman, I.; Sussman, J. L. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 9031–9035.
- (18) Silman, I.; Harel, M.; Axelsen, P.; Ravest, M.; Sussman, J. L. *Biochem. Soc. Trans.* **1994**, *22*, 745–749.
- (19) Ravest, M. L.; Harel, M.; Pang, Y.-P.; Silman, I.; Kozikowski, A. P.; Sussman, J. L. *Nat. Struct. Biol.* **1997**, *4*, 57–63.
- (20) Verdonk, M. L.; Boks, G. J.; Kooijman, H.; Kanters, J. A.; Kroon, J. *J. Comput.-Aided Mol. Des.* **1993**, *7*, 173–182.
- (21) Allen, F. H.; Kennard, O. *Chem. Des. Autom. News* **1993**, *8*, 31–37.
- (22) Pearlman, D. A.; Case, D. A.; Caldwell, J. W.; Ross, W. S.; Cheatham, T. E., III; DeBolt, S.; Ferguson, D.; Seibel, G.; Kollman, P. *Comput. Phys. Commun.* **1995**, *91*, 1–41.
- (23) Caldwell, J.; Dang, L., X.; Kollman, P. A. *J. Am. Chem. Soc.* **1990**, *112*, 9144–9147.
- (24) Caldwell, J. W.; Kollman, P. A. *J. Phys. Chem.* **1995**, *99*, 6208–6219.
- (25) Bayly, C. I.; Cieplak, P.; Cornell, W. D.; Kollman, P. A. *J. Phys. Chem.* **1993**, *97*, 10269–10280.
- (26) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Kollman, P. A. *J. Am. Chem. Soc.* **1993**, *115*, 9620–9631.
- (27) MacKerrell, A. D., Jr. *J. Phys. Chem. B* **1998**, *102*, 3586–3616.
- (28) Donini, O.; Weaver, D. F. *J. Comput. Chem.* **1998**, *19*, 1515–1525.
- (29) Parr, R. G.; Yang, W. *Density-functional theory of atoms and molecules*; Oxford University Press: Oxford, UK, 1989.
- (30) Lifson, S.; Hagler, A. T.; Dauber, P. *J. Am. Chem. Soc.* **1979**, *101*, 5111–5121.
- (31) Shanzer, A.; Libman, J.; Lifson, S.; Felder, C. E. *J. Am. Chem. Soc.* **1986**, *108*, 7609–7619.
- (32) Head-Gordon, M.; Pople, J. A. *J. Chem. Phys.* **1988**, *89*, 5777.
- (33) Saebø, S.; Almlof, J. *Chem. Phys. Lett.* **1989**, *154*, 83–89.
- (34) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Zakrzewski, Y. G.; Montgomery, J. A.; 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, Y.; 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.; 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. G.; Chen, W.; Wong, M. W.; Andres, J. L.; Head-Gordon, M.; Replogle, E. S.; Pople, J. A. *Gaussian98*; Gaussian, Inc.: Pittsburgh, PA, 1998.
- (35) Lifson, S. Potential Energy Functions for Structural Molecular Biology. In *Structural Molecular Biology. Methods and Applications*; Davies, B. D., Saenger, W., Danyluck, S. S., Eds.; Plenum Publishing Co.: New York, 1982; pp 359–386.
- (36) Lifson, S. Potential Energy Functions for Molecular Biology. In *Supramolecular Structure and Function*; Pifat, G., Herak, J. N., Eds.; Plenum Press: New York, 1983; pp 1–44.
- (37) Allinger, N. L.; Yuh, Y. H.; Lii, J.-H. *J. Am. Chem. Soc.* **1989**, *111*, 8551–8566.
- (38) Dudek, M. J.; Ramnarayan, K.; Ponder, J. W. *J. Comput. Chem.* **1998**, *19*, 548–573.
- (39) Pappu, R. V.; Hart, R. K.; Ponder, J. W. *J. Phys. Chem. B* **1998**, *102*, 9725–9742.
- (40) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648–5652.
- (41) Amant, S. A. Density functional methods in biomolecular modeling. In *Computational Chemistry*; Lipkowitz, K. B., Boyd, D. B., Eds.; VCH: New York, 1996; Vol. 7; pp 721–260.
- (42) Boys, S. F.; Bernardi, F. *Mol. Phys.* **1970**, *19*, 553–566.
- (43) Breneman, C. M.; Wiberg, K. B. *J. Comput. Chem.* **1990**, *11*, 361–373.
- (44) Reed, A. E.; Curtiss, L. A.; Weinhold, F. *Chem. Rev.* **1988**, *88*, 899–926.
- (45) Zhu, W.-L.; Tan, X.-J.; Puah, C. M.; Gu, J.-D.; Jiang, H.-L.; Chen, K.-X.; Felder, C. E.; Silman, I.; L. Sussman, J. L. *J. Chem. Phys. A* **2000**, *104*, 9573–9580.
- (46) Meot-Ner (Mautner), M.; Deakyne, C. A. *J. Am. Chem. Soc.* **1985**, *107*, 469–474.
- (47) Hoyau, S.; Norrman, K.; McMahon, T. B.; Ohanessian, G. *J. Am. Chem. Soc.* **1999**, *121*, 8864–8875.
- (48) Kauzmann, W. *Adv. Protein Chem.* **1959**, *14*, 1–63.
- (49) Creighton, T. E. *Proteins Structures and Molecular Properties*, 2nd ed.; W. H. Freeman: New York, 1993.
- (50) Botti, S. A.; Felder, C. E.; Lifson, S.; Sussman, J. L.; Silman, I. *Biophys. J.* **1999**, *77*, 2430–2450.
- (51) Miyazawa, A.; Fujiiyoshi, Y.; Stowell, M.; Unwin, N. *J. Mol. Biol.* **1999**, *288*, 765–786. ^b From Caldwell and Kollman. ¹³ ^c From Pullman et al. ¹⁴ ^d From Kim et al. ¹² ^g Uses original force parameter set. Note that the net charge is not exactly 1.0.