The Role of Polarization and Charge Transfer in the Solvation of Biomolecules

Arjan van der Vaart and Kenneth M. Merz, Jr.*

Contribution from the Department of Chemistry, 152 Davey Laboratory, The Pennsylvania State University, University Park, Pennsylvania 16802

Received April 19, 1999

Abstract: Herein we demonstrate that charge transfer from the protein to the first solvation layer is a significant contributor to the total solvation interaction energy between water and major cold shock protein A (CspA). Interestingly, we find that polarization and electrostatic interactions are predicted to be less important in protein-water interactions than charge transfer. Charge transfer is most prominent for charged residues, but also occurs between water molecules and the hydrophilic side chains and the carbonyl and amide groups of the main chain. The route of charge transfer is via hydrogen bonds between protein and solvent. These results are consistent with recent NMR and X-ray observations, which show that hydrogen bonding interactions have a significant covalent character as opposed to the traditional purely electrostatic view.

Introduction

Water plays a critical role in the stabilization of biomolecular systems[1] and the origin of this stabilization arises from the unique hydrogen bonding capability of water.[2] The precise nature of these hydrogen bonding interactions have been and continue to be of intense interest to both theoreticians as well as experimentalists.[2,3]

In the field of biomolecular simulations, it is generally assumed that hydrogen bonds are mostly electrostatic in nature.[4−8] On the basis of this assumption, stabilization of a biomolecular system by water is generally described by classical bonding models.[9−13] Intermolecular forces in these models consist of a combination of Coulomb and Lennard-Jones interactions. These models have been generally successful in obtaining structural and thermodynamical information for the system of interest.[9,10] Extension beyond this simple classical pair potential has been confined to the addition of polarization, which represents the intramolecular electronic reorganization that occurs when, for example, two molecules interact with one another.[14−23] Generally, this is assumed to contribute about 10% to the total intramolecular interaction energy.[24] However, there is some debate about whether polarization will advance classical potential functions significantly. An example is liquid water, in which polarization is estimated to contribute ∼17% to the interaction energy.[24] Berendsen and co-workers showed that the average effect of polarization can be included in the pair potential by the proper scaling of the atomic charges.[25] This procedure greatly improved solvent properties, like the density and diffusion constant. Another example of this is the determination of the solvation free energy of amines[26] which were problematic, and one of the proposed solutions to the problem was the inclusion of polarization. However, recently Jorgensen and co-workers demonstrated that a carefully parametrized classical amine model was able to correctly reproduce solvation free energies.[27] Jorgensen then argues that it may not be necessary to invoke polarization effects in many instances and that the current classical pair potentials are sufficient if parametrized carefully.[27]

The success of effective pair potentials clearly shows that in many instances many body effects, like polarization, can be incorporated in an average way by proper parametrization to

(1) Creighton, T. E. *Proteins, structures and molecular properties*; W. H. Freeman: New York, 1993.
(2) Jeffrey, G. A.; Saenger, W. *Hydrogen Bonding in Biological Structures*; Springer-Verlag: Heidelberg, 1991.
(3) Smith, D. A. *Modeling the Hydrogen Bond*; American Chemical Society: Washington, DC, 1994; Vol. 569.
(4) Buckingham, A. D.; Fowler, P. W.; Hutson, J. M. *Chem. Rev.* **1988**, *88*, 963−988.
(5) Singh, U. C.; Kollman, P. A. *J. Chem. Phys.* **1984**, *80*, 353−355.
(6) Kollman, P. *J. Am. Chem. Soc.* **1977**, *99*, 4875−4894.
(7) Rendell, A. P. L.; Bacskay, G. B.; Hush, N. S. *Chem. Phys. Lett.* **1985**, *117*, 400−408.
(8) Spackman, M. A. *J. Chem. Phys.* **1986**, *85*, 6587−6601.
(9) Pettitt, B. M.; Smith, J. C. *Comput. Phys. Commun.* **1995**, *91*, 1−344.
(10) Brooks, C. L., III.; Case, D. A. *Chem. Rev.* **1993**, *93*, 2487−2502.
(11) van Gunsteren, W. F.; Berendsen, H. J. C. *Angew. Chem., Int. Ed. Engl.* **1990**, *29*, 992−1023.
(12) Bowen, J. P.; Allinger, N. L. *Rev. Comput. Chem.* **1991**, *2*, 81−98.
(13) Allen, M. P.; Tildesley, D. J. *Computer Simulation of Liquids*; Clarendon Press: Oxford, 1987.
(14) Stuart, S. J.; Berne, B. J. *J. Phys. Chem.* **1996**, *100*, 11934−11943.
(15) Rick, S. W.; Stuart, S. J.; Berne, B. J. *J. Chem. Phys.* **1994**, *101*, 6141−6156.
(16) Caldwell, J. W.; Kollman, P. A. *J. Am. Chem. Soc.* **1995**, *117*, 4177−4178.
(17) Caldwell, J.; Dang, L. X.; Kollman, P. A. *J. Am. Chem. Soc.* **1990**, *112*, 9144−9147.
(18) Dang, L. X.; Rice, J. E.; Caldwell, J.; Kollman, P. A. *J. Am. Chem. Soc.* **1991**, *113*, 2481−2486.
(19) Applequist, J.; Carl, J. R.; Fung, K. K. *J. Am. Chem. Soc.* **1972**, *94*, 2952−2960.
(20) Applequist, J. *J. Phys. Chem.* **1993**, *97*, 6016−6023.
(21) Kuwajima, S.; Warshel, A. *J. Phys. Chem.* **1990**, *94*, 460−466.
(22) Wallqvist, A.; Ahlstrom, P.; Karlstrom, G. *J. Phys. Chem.* **1990**, *94*, 1649−1656.
(23) New, M. H.; Berne, B. J. *J. Am. Chem. Soc.* **1995**, *117*, 7172−7179.
(24) Gao, J.; Xia, X. *Science* **1992**, *258*, 631−635.
(25) Berendsen, H. J. C.; Grigera, J. R.; Straatsma, T. P. *J. Phys. Chem.* **1987**, *91*, 6289−6271.
(26) Tannor, D. J.; Marten, B.; Murphy, R.; Friesner, R. A.; Sitkoff, D.; Nicholls, A.; Ringnalda, M.; Goddard, I., W. A.; Honig, B. *J. Am. Chem. Soc.* **1994**, *116*, 11875−11882.
(27) Rizzo, R. C.; Jorgensen, W. L. *J. Am. Chem. Soc.* **1999**, *121*, 4827−4836.

10.1021/ja9912325 CCC: $18.00
Published on Web 09/15/1999

accurately reproduce thermodynamic data. However, it is not clear why this simple approach works so well. To understand the effectiveness of pair potentials, it is necessary to obtain insights into the different processes at the electronic level that govern interactions at the atomic and molecular level. Understanding these processes will be useful to develop increasingly better force fields for molecular simulations.

A number of quantum mechanical decomposition schemes have been developed to analyze intermolecular interactions at the electronic level.[28−32] The method of Kitaura and Morokuma[28] and Weinhold and co-workers[29−31] decomposes energies by deletion of specific Fock-matrix elements, while the method of van der Vaart and Merz relaxes the density matrix at appropriate steps in an SCF procedure (see below).[32] For example, the following decomposition is used in the Morokuma model:

$$\Delta E_{\text{total}} = \Delta E_{\text{el}} + \Delta E_{\text{pol}} + \Delta E_{\text{CT}} + \Delta E_{\text{ex}} \qquad (1)$$

where (from left to right) we have the total interaction energy, the electrostatic interaction energy, the polarization component, the charge-transfer component (intermolecular charge rearrangement) and the exchange-repulsion component. The Morokuma model early recognized the importance of nonelectrostatic components of the interaction energy.[28] It has been estimated (using a relative scale), for simple hydrogen bonded complexes, that the polarization component is ∼10%, charge transfer is ∼20%, exchange repulsion is ∼25%, and electrostatic interactions account for ∼45%.[33] From this analysis it is predicted that the so-called charge-transfer component is greater than the polarization component which has been of recent interest. Indeed, using Weinhold's natural bond order analysis charge transfer increases to upward of 60%,[29−31,34] a result also found with the decomposition method of van der Vaart and Merz.[32]

Given the early recognition of the magnitude of the CT term it is surprising more interest has not been given toward inclusion of this into classical models. The reasons for this are not clear, but the amount of charge actually transferred in small hydrogen bonded complexes is small (at most a few tenths of an electron), so that it was assumed that its overall net affect was quite small. It is also not clear how to include CT into a classical model, whereas there have been a number of approaches to incorporate polarization, for example.[14−23] It should be stressed that simple Coulombic terms are effective pair potentials in that they collapse the full interaction into the leading term in eq 1. Thus, in reality classical potentials have been attempting to capture all of these interactions, but in a highly simplified manner. Finally, it has not been realized how synergistic CT effects can be when large systems are considered. Nadig and co-workers[33] demonstrated (using the linear-scaling divide and conquer approach[35−37]) that at the protein−water interface there are a number of CT interactions which when considered in toto result in a significant amount of charge transferred. Charge-transfer effects are quantum mechanical in nature and result in the sharing of electrons between a pair of molecules that are hydrogen bonded with one another and, hence, introduces some covalency to this type of interaction. This result supports the recent experimental observation that the hydrogen bond is quantum mechanical in nature. Recent detection of large *J*-couplings (up to 10 Hz) transmitted through hydrogen-bond donor and acceptor atoms in NMR spectra of solvated biomolecular species first established the partial covalent character of hydrogen bonds.[38−42] The covalent nature of the hydrogen bond was also demonstrated in a direct X-ray study of ice crystals.[43] Given these recent observations it has become clear that further consideration of charge transfer is warranted.

Previous semiempirical calculations on a water−protein system showed that 1−2 electrons were transferred from the surface of the small protein CspA to water.[33] As mentioned above this observation demonstrated, for the first time, the importance of charge transfer at the biomolecule−water interface. However, this study only demonstrated the manifestation of charge transfer (i.e., electrons are transferred to and from the protein−water interface) and did not elucidate the energetic consequences of the charge-transfer interaction.

Here we investigate the energetic contribution of charge transfer to the total interaction energy between protein and water by means of the divide and conquer (D&C) interaction energy decomposition,[32] which is readily implemented into the D&C algorithm. This method decomposes the interaction energy into electrostatic, polarization, and charge-transfer energies. It also allows elimination of polarization or both polarization and charge transfer, from the intermolecular interactions. Analysis of the atomic charges obtained from these modified intermolecular interactions elucidates the effect of polarization and charge transfer on the charge distribution of the system.

Approach

Theory. For clarity, we will review the semiempirical divide and conquer interaction energy decomposition scheme.[32] In the D&C method,[35−37] the system of interest is divided into subsystems. Since density matrix elements between atoms of two subsystems are only different from zero when these subsystems overlap, charge flow between subsystems only occurs when the subsystems overlap. Suppose we have one solute molecule and one solvent molecule, which are each placed in a different subsystem. For this system, the overlap between the solute (p) and solvent (w) subsystems can be written as (*p*:*pw*, *w*:*pw*), which indicates that the density matrix elements between the subsystems are different from zero. The total number of electrons is constrained by the Fermi energy ϵ_F. Thus, the interaction energy between the solute and solvent is given by

$$E_{\text{int}} = E[\epsilon_F,(p{:}pw,w{:}pw)] - (E[\epsilon_F{}^{p},(p{:}p)] + E[\epsilon_F{}^{w},(w{:}w)]) \qquad (2)$$

The second term can be obtained with one calculation, by infinitely separating the solute and solvent system and employing two separate Fermi energies for the solute and solvent systems

(28) Kitaura, K.; Morokuma, K. *Int. J. Quantum Chem.* **1976**, *10*, 325−340.
(29) Reed, A. E.; Weinhold, F.; Curtiss, L. A.; Pochatko, D. J. *J. Chem. Phys.* **1986**, *84*, 5687−5705.
(30) Reed, A. E.; Curtiss, L. A.; Weinhold, F. *Chem. Rev.* **1988**, *88*, 899−926.
(31) Weinhold, F. *J. Mol. Struct.* **1997**, *399*, 181−197.
(32) van der Vaart, A.; Merz, K. M., Jr. *J. Phys. Chem. A* **1999**, *103*, 3321−3329.
(33) Nadig, G.; van Zant, L. C.; Dixon, S. L.; Merz, K. M., Jr. *J. Am. Chem. Soc.* **1998**, *120*, 5593−5594.
(34) King, B.; Weinhold, F. *J. Chem. Phys.* **1995**, *103*, 333−347.
(35) Yang, W.; Lee, T.-S. *J. Chem. Phys.* **1995**, *103*, 5674−5678.
(36) Dixon, S. L.; Merz, K. M., Jr. *J. Chem. Phys.* **1996**, *104*, 6643−6649.
(37) Dixon, S. L.; Merz, K. M., Jr. *J. Chem. Phys.* **1997**, *107*, 879−893.
(38) Blake, P. R.; Park, J. B.; Adams, M. W. W.; Summers, M. F. *J. Am. Chem. Soc.* **1992**, *114*, 4931−4933.
(39) Dingley, A. J.; Grzesiek, S. *J. Am. Chem. Soc.* **1998**, *120*, 8293−8297.
(40) Golubev, N. S.; Shenderovich, I. G.; Smirnov, S. N.; Denisov, G. S.; Limbach, H.-H. *Chem. Eur. J.* **1999**, *5*, 492−497.
(41) Cordier, F.; Grzesiek, S. *J. Am. Chem. Soc.* **1999**, *121*, 1601−1602.
(42) Cornilescu, G.; Hu, J.-S.; Bax, A. *J. Am. Chem. Soc.* **1999**, *121*, 2949−2950.
(43) Isaacs, E. D.; Shukla, A.; Platzman, P. M.; Hamann, D. R.; Barbiellini, B.; Tulk, C. A. *Phys. Rev. Lett.* **1999**, *82*, 600−603.

$$E_{int} = E[\epsilon_F,r,P(r),(p{:}pw,w{:}pw)] - E[(\epsilon_F^{p},\epsilon_F^{w}),\infty,P(\infty),(p{:}p,w{:}w)] \quad (3)$$

Here (p:p, w:w) indicates that the solute and solvent subsystems do not overlap, so that the density matrix elements between the two subsystems are zero. Energies are now shown as explicit functions of the solute−solvent separation (r) and the density matrix at this separation ($P(r)$). The electrostatic energy can be obtained by bringing the infinitely separated solute and solvent subsystems to the equilibrium distance, without relaxation of the density matrix

$$E_{es} = E[(\epsilon_F^{p},\epsilon_F^{w}),r,P(\infty),(p{:}p,w{:}w)] - E[(\epsilon_F^{p},\epsilon_F^{w}),\infty,P(\infty),(p{:}p,w{:}w)] \quad (4)$$

By allowing charge flow within the solute and solvent subsystems only, the polarization energy is obtained

$$E_{pol} = E[(\epsilon_F^{p},\epsilon_F^{w}),r,P(r),(p{:}p,w{:}w)] - E[(\epsilon_F^{p},\epsilon_F^{w}),r,P(\infty),(p{:}p,w{:}w)] \quad (5)$$

The charge-transfer energy is obtained by allowing electrons to flow between the solute and solvent

$$E_{CT} = E[\epsilon_F,r,P(r),(p{:}pw,w{:}pw)] - E[(\epsilon_F^{p},\epsilon_F^{w}),r,P(r),(p{:}p,w{:}w)] \quad (6)$$

We stress that this decomposition differs in several aspects from Morokuma's and Weinhold's method. First, in agreement with classical field theory, polarization is defined as intramolecular charge flow. This means that polarization contains the intramolecular reorganization of electrons through the mixing of occupied and virtual molecular orbitals. This term also contains the intramolecular exchange repulsion interaction, which is considered as a separate term in, for example, the Kitaura−Morokuma approach. While we could also formally remove this term in our analysis we feel that this is inappropriate for this discussion since we are considering the intramolecular rearrangement of electrons that is associated with polarization and, indeed, the exchange repulsion correction for like spins is a part of this rearrangement. Second, the charge-transfer contribution is obtained as a sum of all processes that effectively change the formal charge of each molecule. The charge-transfer contribution includes not only the intermolecular mixing between the occupied and virtual orbitals of the two systems but also intermolecular exchange repulsion interactions. Again, we could consider the intermolecular exchange repulsion component as a separate term, but we lump this term into the charge-transfer component of the total interaction energy. The electrostatic interaction only contains core−core repulsion, electron−electron repulsion, and core−electron interactions (no exchange repulsion). This scheme naturally decomposes the interaction energy into classical contributions ($E_{es} + E_{pol}$), that maintain the formal charge of each molecule and a quantum contribution (E_{CT}) that effectively changes the formal charge.

This method is easily extended to larger systems consisting of multiple subsystems. For example, in the analysis of the protein−water interface charge can freely flow between solvent molecules and protein residues during the polarization step of the analysis. However, consistent with our interest in the protein−water interface only, charge transfer between protein and water is prohibited during the polarization phase of the calculation.

We could decompose intrawater and intraprotein charge-transfer effects arising from water being in the field of protein, but our focus is on the interface in this manuscript. To calculate the intramolecular interaction between all solvent and solute molecules in a system, each solute and solvent molecule would need to be assigned a Fermi energy. Note that each solute and solvent molecule can still be divided into subsystems. Instead of the simple overlap vector in eqs 2−6, we now have (p_1:p_1, p_2:p_2, p_3:p_3, ..., w_1:w_1, w_2:w_2, w_3:w_3, ...) to indicate that subsystems of solute p_1 only overlaps with the subsystems of solute p_1, subsystems of solute p_2 only overlap with the subsystems of solute p_2, etc. Note that in this model all charge flow between molecules would be prohibited during polarization phase of the calculation.

Table 1. Number of Water Molecules for the Interaction Energy Decomposition

time[a]	no. waters
100	1124
200	1176
300	1193
400	1181
500	1113

[a] In ps.

Decomposition of the Protein−Water Interaction Energy. Snapshots were obtained from a MD simulation of CspA in TIP3P[44] water using the AMBER[45] force field. The formal charge of CspA was −1 *e*. Details of this simulation are described elsewhere.[33] All waters greater than 9 Å from the protein were removed, leaving between 1113 and 1193 waters total (Table 1). $P(\infty)$ was constructed from the density matrix of the protein in a vacuum and from the density matrix of the water system in which the protein was replaced by a continuum with dielectric constant 80. The continuum was used to mimic a water environment,[46] rather than a vacuum, for the water molecules close to the protein. In this way, polarization of the water molecules close to the protein is not artificially enhanced by changing the environment from a vacuum to solution, instead of from bulk water to solution. All calculations were performed with the semiempirical AM1[47] or PM3[48,49] Hamiltonian as implemented in our D&C program DivCon.[36] The continuum calculations were performed with a modified version of our coupled DelPhi[50]−DivCon program,[46] using CM2[51] charges. For all calculations, a cutoff of 8.0 Å was used for the off-diagonal elements of the Fock, 1-electron and density matrixes.

Charge Distribution Analysis. We introduce the polarization effect as the effect of polarization on the charge distribution of the system, and the charge-transfer effect as the effect of charge transfer on the charge distribution of the system. The polarization effect was calculated from the difference between charges obtained from calculation of $E[(\epsilon_F^{p},\epsilon_F^{w}),r,P(r),(p{:}p,w{:}w)]$, and those obtained from calculation of $E[(\epsilon_F^{p},\epsilon_F^{w}),r,P(\infty),(p{:}p,w{:}w)]$. The charge-transfer effect was obtained from the difference between charges from calculation of $E[\epsilon_F,r,P(r),(p{:}pw,w{:}pw)]$ and $E[(\epsilon_F^{p},\epsilon_F^{w}),r,P(r),(p{:}p,w{:}w)]$. We define the solvation effect as the difference in charge between the solvated and vacuum protein. The solvation effect is identical to the sum of the polarization and the charge-transfer effect on the system. A positive solvation effect indicates that a residue loses electrons upon solvation, i.e. a positive residue becomes more positive, a negative residue less negative. A negative solvation effect indicates the gain of electrons upon solvation, i.e. a positive residue becomes less positive, a negative residue more negative. Negative polarization and charge-transfer effects also indicate the gain of electrons, and positive polarization and charge-transfer effects indicate the loss of electrons. All reported charges are CM2[51] charges, rather than Mulliken charges, since CM2 charges provide a better representation of the dipole moment for tested compounds.[51]

Results and Discussion

Application of our interaction energy decomposition scheme to five configurations of solvated CspA showed that charge-

(44) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926−935.
(45) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M., Jr.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. *J. Am. Chem. Soc.* **1995**, *117*, 5179−5197.
(46) Gogonea, V.; Merz, K. M., Jr. *J. Phys. Chem. A* **1999**, *103*, 5171−5188.
(47) Dewar, M. J. S.; Zoebisch, E. G.; Healy, E. F.; Stewart, J. J. P. *J. Am. Chem. Soc.* **1985**, *107*, 3902−3909.
(48) Stewart, J. J. P. *J. Comput. Chem.* **1989**, *10*, 209−220.
(49) Stewart, J. J. P. *J. Comput. Chem.* **1989**, *10*, 221−264.
(50) Honig, B.; Nicholls, A. *Science (Washington, DC)* **1995**, *268*, 1144−1149.
(51) Li, J.; Cramer, C. J.; Truhlar, D. G. *J. Phys. Chem. A.* **1998**, *102*, 1820−1831.

Table 2. Interaction Energy Decomposition for Solvated CspA

time[a]	method	E_{int}[b]	E_{es}[b]	E_{pol}[b]	E_{CT}[b]	$\%E_{es}$[c]	$\%E_{pol}$[c]	$\%E_{CT}$[c]
100	PM3	−1501.495	678.918	−256.474	−1923.939	23.7	9.0	67.3
	AM1	−1394.793	−33.369	−235.601	−1125.823	2.4	16.9	80.7
200	PM3	−1489.493	698.240	−248.075	−1939.658	24.2	8.6	67.2
	AM1	−1354.087	8.731	−224.368	−1138.450	0.6	16.4	83.0
300	PM3	−1494.170	743.241	−234.367	−2003.044	24.9	7.9	67.2
	AM1	−1333.471	50.513	−214.238	−1169.746	3.5	14.9	81.6
400	PM3	−1550.705	869.492	−279.010	−2141.186	26.4	8.5	65.1
	AM1	−1414.340	81.045	−249.260	−1246.124	5.1	15.8	79.1
500	PM3	−1376.857	929.004	−235.015	−2070.846	28.7	7.3	64.0
	AM1	−1227.472	209.451	−217.185	−1219.738	12.7	13.2	74.1
average	PM3	−1482.544	783.779	−250.588	−2015.735	25.6	8.2	66.2
	AM1	−1344.833	63.274	−228.130	−1179.976	4.9	15.4	79.7

[a] In ps. [b] In kcal/mol. [c] $\%E_x = 100|E_x|/(|E_{es}| + |E_{pol}| + |E_{CT}|)$, x = es, pol, CT.

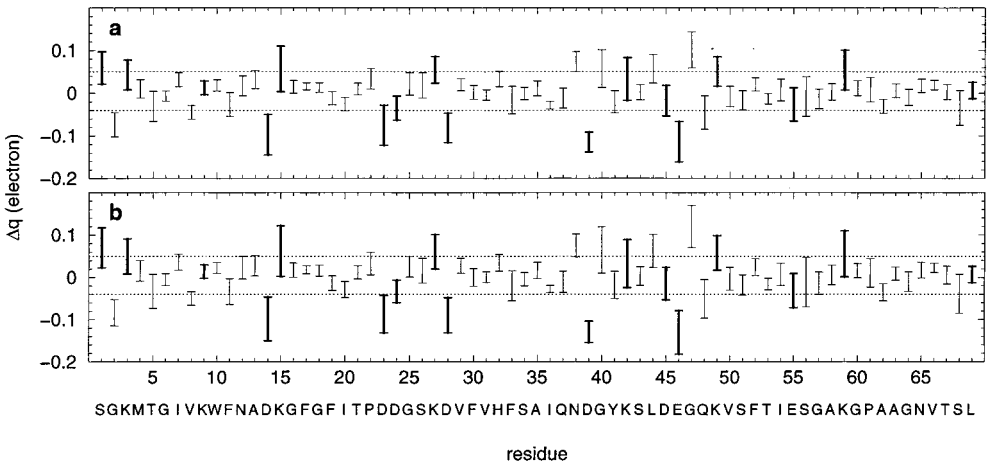

Figure 1. The effect of polarization on the charge distribution of CspA. Shown is the difference in the calculated CM2 charges when polarization is added to electrostatics. Charges are listed by residue. The bars indicate the observed range of charges from the five snapshots and the bold bars indicate the charged residues. (a) AM1. (b) PM3.

transfer constitutes between ∼67 (PM3) to ∼80% (AM1) of the interaction energy between water and CspA (percentages calculated on an absolute scale). Electrostatics contribute ∼25 (PM3) to ∼4% (AM1) of the interaction energy and polarization ∼8 (PM3) to 16% (AM1). Charge transfer and polarization are both stabilizing, while electrostatics are slightly destabilizing. These results roughly correspond to calculations performed on binary hydrogen bonded systems.[29,32] and a 64 water system.[32] However, the situation here is far more complex than in hydrogen bonded dimers. In the present case we have a surface that has some highly polar (i.e., charged), polar (i.e., uncharged polar groups), and nonpolar (i.e., hydrophobic groups). The complex nature of this interface is reinforced by the figures described in more detail below. Thus, it is tenuous at best to interpret our calculated numbers in terms of previous results obtained on small hydrogen bonded clusters.

Another interesting aspect of our calculations is the observation of a generally positive electrostatic component to the total interaction energy. Using the Kitaura−Morokuma scheme on small hydrogen bonded clusters the electrostatic component is negative,[28,33] but Weinhold has shown that the Kitaura−Morokuma electrostatic term contains some of the charge-transfer component, which may make this term more negative.[30] Moreover, Weinhold's scheme gives a positive electrostatic component[29,30] as we observe for many small hydrogen bonded clusters.[32] In our scheme the electrostatic component contains only the core−core and electron−electron repulsion and core−electron terms, and the value of this term reflects the balance between these terms. In PM3 the core−core repulsion term contains some "shoulders" in and around the minimum region which potentially make this term more repulsive than in AM1.[32,52] This is exactly what we see in Table 2—AM1 gives a more negative electrostatic term than does PM3. A further complication comes from the nature of the interface we are studying as outlined above. This may result in a larger number of unfavorable core−core, electron−electron, and core−electron interactions than would be present in a simple hydrogen bonded cluster. Another contributing factor is the use of AMBER-generated configurations which are not at the PM3 or AM1 minimum. In test minimizations we find that the electrostatic term tends to become more negative as we reach the AM1 or PM3 local minimum (however, the relationship among terms remains roughly constant). Finally, we note that there may be differences in the absolute percentages predicted by semiempirical models and ab initio or DFT methodologies. However, in terms of the amount of charge transferred Pearl and Zerner[53] demonstrated that semiempirical models and ab initio models transfer about the same amount of charge. In a more extensive study, van der Vaart and Merz have examined a series of hydrogen bonded clusters which show that, in general, the predicted amount of charge transfer is similar for semiempirical and ab initio based methods.[54]

Both polarization and charge transfer significantly alter the charge distribution of the system. The effect of polarization on the charge distribution of CspA is plotted in Figure 1. In this

(52) Csonka, G. I.; Angyan, J. G. *J. Mol. Struct.* **1997**, *393*, 31−38.
(53) Pearl, G. M.; Zerner, M. C. *J. Am. Chem. Soc.* **1999**, *121*, 399−404.
(54) van der Vaart, A.; Merz, K. M., Jr. *Int. J. Quantum Chem.*, **1999**, in press.

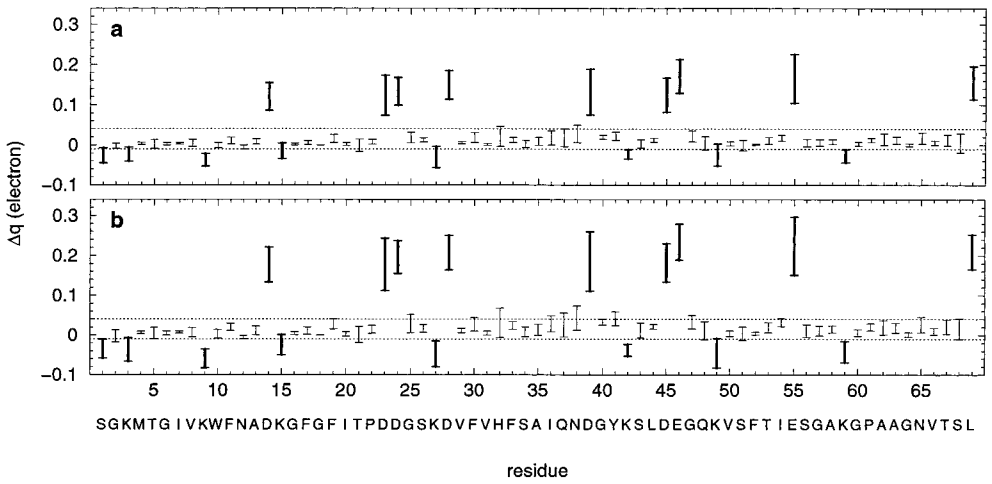

Figure 2. The effect of charge transfer on the charge distribution of CspA. Shown is the difference in the calculated CM2 charges when charge transfer is added to polarization + electrostatics. Charges are listed by residue. The bars indicate the observed range of charges from the five snapshots and the bold bars indicate the charged residues. (a) AM1. (b) PM3.

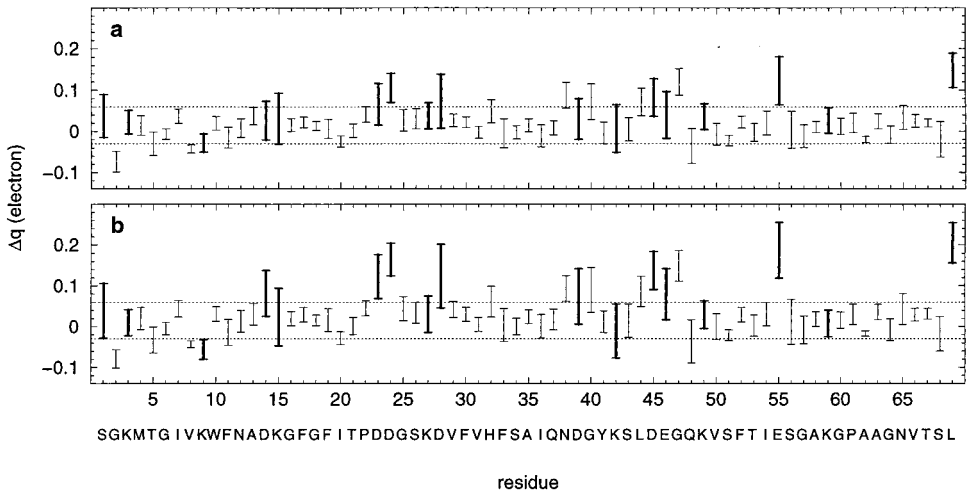

Figure 3. The effect of solvation on the charge distribution of CspA. Shown is the difference in the calculated CM2 charges between the solvated and vacuum protein. Charges are listed by residue. The bars indicate the observed range of charges from the five snapshots and the bold bars indicate the charged residues. (a) AM1. (b) PM3.

figure, each residue is represented by a vertical bar, bold bars indicate the charged residues. The bars represent the range of the polarization effect observed in the five snapshots. In Figure 2, the same is done for the charge-transfer effect.

Polarization has more impact on the charge distribution of CspA than charge transfer. The dotted lines in Figures 1−2 indicate the boundaries of the charge effects within which most residues appear. For polarization, the region is 0.09 electrons wide, for charge transfer, 0.05. Although polarization effects almost every residue significantly, the largest effects are observed for the charged residues (Lys, Asp, Glu, and terminal residues, indicated by the bold bars in Figure 1), and Gly, Leu, Asn, and Gln.

Figure 2 shows that the charged residues are responsible for most of the charge transfer from protein to water. The charge-transfer effect on charged residues were found to be well-separated from the noncharged residues. Charge transfer decreases the charge on charged residues, while polarization generally increases the charge on these residues. The sum of the charge-transfer effects on the charged residues accounts for 70−80% of the total charge-transfer effect.

The combined effect of polarization and charge transfer is shown in Figure 3, where the difference in charge between the solvated and vacuum protein (i.e., the solvation effect) is plotted. Figure 3 shows that most residues have a positive solvation charge effect, especially Asp and Glu. To further analyze the origin of this positive charge effect, we calculated the average polarization and charge-transfer effects for each residue type (Table 3).

Table 3 shows that the negatively charged Asp and Glu residues have the largest polarization and charge-transfer effects. The polarization effect is negative, and the charge-transfer effect is positive for these residues. This means that polarization makes Asp and Glu more negative, while charge transfer makes Asp and Glu less negative. Since the charge-transfer effect is about 2 (AM1) or 2.5 times (PM3) the magnitude of the polarization effect on Asp and Glu, the combined effect of polarization and charge transfer (i.e., the solvation effect) is positive: the charge of Asp and Glu become less negative upon solvation.

For the positively charged groups (Lys and the protonated terminal amine group on Ser 1), the polarization effect is positive, and the charge-transfer effect is negative. Polarization makes these groups more positive, while charge transfer makes these groups less positive. Since the polarization effect is about 1.5 (AM1) or about 1.2 (PM3) times larger in magnitude than the charge-transfer effect, the solvation effect on the positively charged groups is slightly positive. This means that the positively charged groups become slightly more positive upon solvation.

Table 3. Effect of Polarization and Charge Transfer on the Charge Distribution of CspA

		polarization effect[b]		charge-transfer effect[b]		solvation effect[b]	
type	res.[a]	AM1	PM3	AM1	PM3	AM1	PM3
negative	D	−0.069 ± 0.043	−0.076 ± 0.047	0.129 ± 0.032	0.186 ± 0.039	0.060 ± 0.048	0.111 ± 0.054
	E	−0.080 ± 0.054	−0.091 ± 0.064	0.155 ± 0.044	0.218 ± 0.049	0.075 ± 0.066	0.126 ± 0.076
	L 69	0.001 ± 0.016	0.004 ± 0.016	0.139 ± 0.033	0.195 ± 0.035	0.140 ± 0.032	0.199 ± 0.039
	all[c]	−0.064 ± 0.049	−0.070 ± 0.056	0.136 ± 0.036	0.194 ± 0.042	0.072 ± 0.056	0.124 ± 0.063
positive	K	0.040 ± 0.036	0.043 ± 0.040	−0.024 ± 0.015	−0.039 ± 0.021	0.017 ± 0.038	0.004 ± 0.045
	S 1	0.045 ± 0.031	0.053 ± 0.038	−0.033 ± 0.015	−0.045 ± 0.020	0.012 ± 0.044	0.009 ± 0.056
	all[c]	0.041 ± 0.035	0.045 ± 0.039	−0.025 ± 0.015	−0.040 ± 0.021	0.016 ± 0.038	0.005 ± 0.045
hydrophilic	H	0.034 ± 0.015	0.035 ± 0.018	0.011 ± 0.020	0.017 ± 0.029	0.044 ± 0.023	0.052 ± 0.030
	N	0.037 ± 0.032	0.040 ± 0.033	0.012 ± 0.017	0.018 ± 0.024	0.049 ± 0.039	0.058 ± 0.044
	Q	−0.028 ± 0.032	−0.032 ± 0.038	0.013 ± 0.016	0.021 ± 0.022	−0.015 ± 0.039	−0.012 ± 0.048
	S	−0.007 ± 0.028	−0.010 ± 0.032	0.005 ± 0.011	0.012 ± 0.014	−0.003 ± 0.030	0.002 ± 0.035
	T	−0.004 ± 0.020	−0.004 ± 0.024	0.006 ± 0.012	0.011 ± 0.016	0.002 ± 0.020	0.006 ± 0.025
	Y	−0.017 ± 0.026	−0.015 ± 0.031	0.019 ± 0.008	0.034 ± 0.014	0.002 ± 0.023	0.019 ± 0.023
	all[c]	0.001 ± 0.034	0.000 ± 0.038	0.008 ± 0.014	0.015 ± 0.019	0.009 ± 0.037	0.015 ± 0.042
hydrophobic	A	0.006 ± 0.025	0.006 ± 0.027	0.009 ± 0.008	0.013 ± 0.011	0.015 ± 0.024	0.019 ± 0.025
	F	−0.002 ± 0.023	−0.003 ± 0.027	0.010 ± 0.009	0.018 ± 0.013	0.008 ± 0.022	0.014 ± 0.025
	G	−0.003 ± 0.008	0.015 ± 0.054	0.005 ± 0.009	0.010 ± 0.014	0.018 ± 0.053	0.025 ± 0.063
	I	−0.003 ± 0.026	−0.003 ± 0.029	0.009 ± 0.010	0.015 ± 0.014	0.006 ± 0.029	0.012 ± 0.033
	L	0.052 ± 0.028	0.056 ± 0.031	0.013 ± 0.004	0.021 ± 0.005	0.065 ± 0.029	0.077 ± 0.032
	M	0.017 ± 0.017	0.022 ± 0.019	0.003 ± 0.002	0.006 ± 0.003	0.020 ± 0.019	0.028 ± 0.021
	P	0.018 ± 0.022	0.017 ± 0.025	0.010 ± 0.004	0.019 ± 0.007	0.028 ± 0.019	0.036 ± 0.019
	V	−0.003 ± 0.027	0.000 ± 0.031	0.003 ± 0.005	0.007 ± 0.007	0.000 ± 0.028	0.006 ± 0.033
	W	0.019 ± 0.011	0.025 ± 0.011	0.000 ± 0.006	0.005 ± 0.010	0.019 ± 0.014	0.030 ± 0.015
	all[c]	0.007 ± 0.034	0.008 ± 0.039	0.007 ± 0.008	0.012 ± 0.012	0.014 ± 0.036	0.020 ± 0.042

[a] Amino acids are indicated by their one-letter symbols. [b] Average CM2 charge (in electron) with the standard deviation. Averages are taken over all snapshots for the listed amino acids. A negative sign means electrons are gained, while a positive sign indicates that electrons are lost. [c] Average of all residues of the particular type over all snapshots.

Table 4. Average Charge-Transfer Effects per Residue Type

res. type[a]	AM1 (%)[b]	PM3 (%)[b]
negative	67.5	63.2
positive	11.1	11.5
hydrophilic	8.1	9.5
hydrophobic	13.3	15.8

[a] See Table 3 for definition of the residue type. [b] Percentage of the total charge-transfer effect. This percentage is defined as

$$100 \sum_{i \in \text{type A}} |CT_i| / \left(\sum_{A}^{\text{all types}} \sum_{i \in \text{type A}} |CT_i| \right)$$

where CT_i is the charge-transfer effect of residue i.

We see that both positively and negatively charged groups show positive solvation effects on the average, although the underlying mechanism for this effect is completely different for the two residue types. The situation is even more different for the terminal group Leu 69, since the average polarization effect is almost zero for this residue. Thus, the positive solvation charge effect of Leu 69 can be purely ascribed to charge transfer.

Table 4 shows that the negatively charged groups are responsible for most charge transfer from protein to water. This is caused by the high number of negatively charged groups (nine residues total) and the high charge-transfer effect per negatively charged group. Indeed, the charge-transfer effect on negatively charged groups is an order of magnitude larger than on other groups. Since the sign of the charge-transfer effect for the negatively charged groups is positive, the negatively charged groups become less negative, and negative charge flows from the protein to the water. Hydrophilic and hydrophobic groups slightly enhance the flow of electrons from protein to water, since their charge-transfer effect is also positive. The sign of the charge-transfer effect on positively charged groups is negative, meaning that these groups slightly decrease the flow of electrons from the protein to water. The total amount of charge transferred from protein to water is between 1.21 and 1.54 electrons for AM1 and between 1.81 and 2.26 electrons for PM3 (Table 5).

Table 5. Charge Transfer from Protein to Water

time[a]	AM1[b]	PM3[b]
100	1.494	2.214
200	1.213	1.814
300	1.415	2.134
400	1.420	2.179
500	1.536	2.265
average	1.416	2.121

[a] In ps. [b] CM2 charges in electrons. Water has accepted a net negative charge.

The charge-transfer effect on the hydrophilic and hydrophobic groups is positive, but much smaller than the charge-transfer effect on negatively charged groups. The polarization effect for the hydrophilic and hydrophobic groups is also smaller than for the charged groups, although His, Asn, Gln, and Leu still have significant polarization effects. Since the polarization effect for His, Asn, and Leu are positive, the solvation effects of these groups are rather large. For PM3 the solvation effect of His, Asn, and Leu are about half the size of the solvation effect of the negatively charged groups, for AM1 they are only slightly smaller than for the negatively charged groups. Since the polarization effect of Gln is negative and the charge-transfer effect is positive, the solvation effect of Gln is rather small.

Although AM1 and PM3 show very similar trends, the magnitude of charge effects is different for these Hamiltonians. Table 6 shows that the charge-transfer effect for negatively charged groups is about 0.06 electrons larger in PM3 than AM1. This difference is the main reason charge transfer from protein to water is larger in PM3 (Table 5). For the positively charged groups, the charge-transfer effect is about 0.015 electrons smaller in PM3. However, this difference is offset by the hydrophilic and hydrophobic groups for which the charge-transfer effect is about 0.006 electrons larger in PM3.

The polarization effect is very similar in AM1 and PM3: for most residues the difference is about 0.005 electrons, for Glu

Table 6. Dependence of Polarization and Charge-Transfer Effects on the Applied Hamiltonian

type	res.[a]	polarization effect[b]	charge-transfer effect[b]	solvation effect[b]
negative	D	-0.006 ± 0.008	0.057 ± 0.010	0.051 ± 0.013
	E	-0.011 ± 0.012	0.063 ± 0.007	0.052 ± 0.013
	L 69	0.003 ± 0.004	0.055 ± 0.007	0.058 ± 0.009
	all[c]	-0.006 ± 0.009	0.058 ± 0.009	0.052 ± 0.013
positive	K	0.003 ± 0.006	-0.016 ± 0.007	-0.013 ± 0.009
	S 1	0.008 ± 0.008	-0.012 ± 0.005	-0.003 ± 0.012
	all[c]	0.004 ± 0.006	-0.015 ± 0.007	-0.012 ± 0.010
hydrophilic	H	0.001 ± 0.003	0.006 ± 0.009	0.008 ± 0.009
	N	0.003 ± 0.005	0.006 ± 0.008	0.009 ± 0.006
	Q	-0.004 ± 0.006	0.008 ± 0.007	0.004 ± 0.010
	S	-0.003 ± 0.006	0.007 ± 0.005	0.004 ± 0.008
	T	0.000 ± 0.004	0.005 ± 0.005	0.004 ± 0.006
	Y	0.002 ± 0.005	0.015 ± 0.007	0.017 ± 0.002
	all[c]	-0.001 ± 0.006	0.007 ± 0.006	0.006 ± 0.008
hydrophobic	A	0.000 ± 0.005	0.004 ± 0.004	0.004 ± 0.007
	F	-0.001 ± 0.004	0.007 ± 0.005	0.006 ± 0.005
	G	0.003 ± 0.008	0.004 ± 0.005	0.007 ± 0.011
	I	0.000 ± 0.003	0.006 ± 0.005	0.006 ± 0.006
	L	0.004 ± 0.005	0.008 ± 0.001	0.012 ± 0.005
	M	0.005 ± 0.003	0.003 ± 0.002	0.008 ± 0.003
	P	-0.001 ± 0.003	0.008 ± 0.003	0.008 ± 0.003
	V	0.002 ± 0.005	0.003 ± 0.003	0.006 ± 0.006
	W	0.006 ± 0.002	0.005 ± 0.004	0.011 ± 0.002
	all[c]	0.001 ± 0.006	0.005 ± 0.005	0.007 ± 0.008

[a] Amino acids are indicated by their one-letter symbols. [b] Average difference (and standard deviation) between PM3 and AM1 CM2 charge effects. Averages are taken over all the snapshots of the listed amino acid, and the results are given in electrons. [c] Average difference between PM3 and AM1 charge effects of all residues of the particular type over all snapshots.

0.011 electrons. Differences in the solvation effect are therefore mainly caused by differences in the charge-transfer effect. Again, the negatively charged residues show the largest difference between AM1 and PM3, followed by the positively charged residues. The difference in solvation effects between AM1 and PM3 for the hydrophobic and hydrophilic residues are rather small.

It is not clear why AM1 and PM3 give different charge-transfer effects. However, this difference might be caused by the difference in hydrogen bonding geometry preference for these methods. AM1 favors bifurcated structures with nonlinear O···H and N···H hydrogen bonds, while PM3 favors linear hydrogen bonds.[55,56] Since AMBER generated structures favor linear hydrogen bonds too, results for AM1 and PM3 are expected to be different.

Water within 4 Å of the protein carries most of the charge that is transferred from the protein surface residues. The average charge residing on water molecules within 2 Å of the protein was ~ -0.009 for AM1 and ~ -0.013 electrons for PM3, between 2 and 4 Å ~ -0.002 for AM1 and ~ -0.003 electrons for PM3, and for layers further than 4 Å from the protein ~ 0.000 electrons. To assess the spatial extent of charge transfer and polarization on water, we calculated the charge changes experienced per solvation layer. In Figures 4 and 5 the averaged distribution of the polarization charge effect per water layer is shown, for AM1 and PM3 respectively; in Figures 6 and 7 this is done for the charge-transfer effect.

Charge-transfer and polarization-charge effects are mainly limited to the 0−2 Å layer, since this water layer had the largest percentage of either charge effect that differed from zero. Further away from the protein surface, both distributions become more sharply peaked around zero. Charge transfer, overall, has a larger effect on the dispersion of the charge distribution of the water layers, than polarization. Polarization effects range between −0.025 and 0.019 electrons for AM1 and between −0.030 and 0.023 electrons for PM3, while charge-transfer effects range from −0.100 to 0.059 electrons for AM1 and from −0.126 to 0.072 electrons for PM3. Furthermore, the intervals around zero are populated to a lesser extent for the charge-transfer effect, than for the polarization effect.

To assess which water molecules were most highly influenced by polarization, we identified those that had polarization charge effects in excess of 0.007 for AM1, 0.01 electrons for PM3, or less than −0.003 for AM1, −0.005 electrons for PM3. For all configurations, most water molecules with a polarization effect greater than 0.007 for AM1 or 0.01 electrons for PM3 were hydrogen bonded to the carboxylate groups from Asp, Glu, or Leu 69. The remainder were found to be hydrogen bonded to backbone carbonyl groups. Most water molecules with a polarization effect of less than −0.003 for AM1 or −0.005 electrons for PM3 were hydrogen bonded to the ammonium groups of Lys or Ser 1. All others were hydrogen bonded to Thr, Ser, Asn, or Gln side chains, or to backbone amide groups.

We performed a similar analysis to determine which water molecules were highly influenced by charge transfer. For all snapshots, most water molecules with charge-transfer effects exceeding 0.005 for AM1, 0.01 electrons for PM3 were hydrogen bonded to the ammonium groups of Lys or Ser 1. The remainder were hydrogen bonded to Thr, Ser, Asn, Gln side chains or amide backbone groups. Water molecules with charge-transfer effects less than −0.02 for AM1 or −0.03 electrons for PM3 were hydrogen bonded to the carboxylate groups of Glu, Asp, or Leu 69. The remainder were hydrogen bonded to the carbonyl backbone. Water molecules that were hydrogen bonded to positively charged groups lose electrons due to charge transfer and gain electrons due to polarization. The opposite is true for water molecules that were hydrogen bonded to negatively charged groups. This is in line with the polarization and charge-transfer effects observed in the protein.

Both AM1 and PM3 show very similar trends for the charge-transfer and polarization effects of the water molecules. However, for water that is directly hydrogen bonded to the protein, the magnitude of the polarization effect is up to 0.01 electron larger in PM3, and the magnitude of the charge-transfer effect is up to 0.02 electron larger in PM3, while the sign of both effects is the same in AM1 and PM3. For water that is not directly hydrogen bonded to the protein, both effects are almost equal in AM1 and PM3. These results are analogous with the difference in charge effects for the protein for AM1 and PM3. We again suspect that this difference can be explained by the difference in preferred hydrogen bonding geometry for AM1 and PM3.

Conclusions

Our analysis shows that charge transfer from CspA to water occurs mainly via the hydrogen bonds of charged groups with water, and to a lesser extent via hydrogen bonds between the protein backbone and Ser, Thr, Asn, and Gln side chains with water. Moreover, charge transfer is a localized phenomena, since it mainly effects waters that are directly hydrogen bonded to the protein. Since we predict that charge transfer constitutes a large portion of the interaction energy, the first solvation layer is critical for the stabilization of the solvated CspA system. It is tempting to draw an analogy between our results and those obtained from calculations on small hydrogen bonded com-

(55) Zheng, Y. J.; Merz, K. M. Jr. *J. Comput. Chem.* **1992**, *13*, 1151−1169.
(56) Dannenberg, J. J. *J. Mol. Struct.* **1997**, *401*, 279−286.

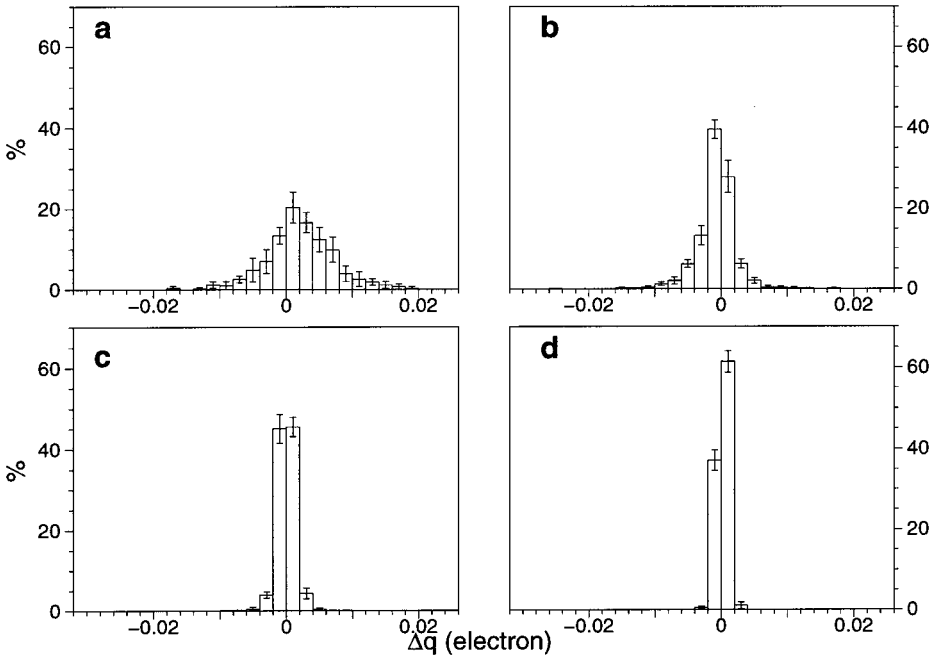

Figure 4. Effect of polarization on the charge distribution of water. AM1 results using CM2 charges. The length of the bars is twice the standard deviation. (a) 0−2 Å layer. (b) 2−4 Å layer. (c) 4−6 Å layer. (d) 6−9 Å layer.

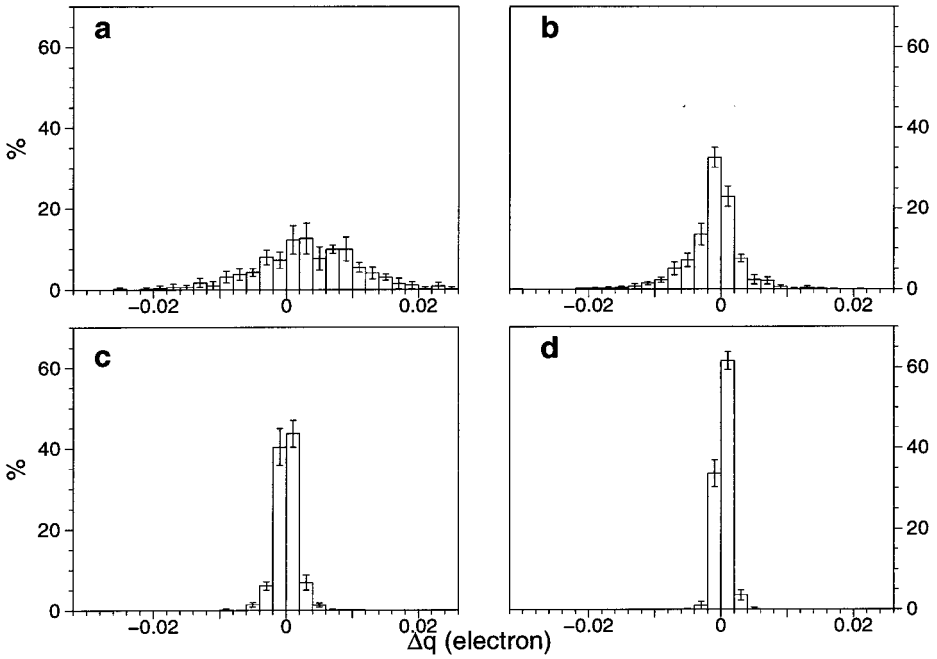

Figure 5. Effect of polarization on the charge distribution of water. PM3 results using CM2 charges. The length of the bars is twice the standard deviation. (a) 0−2 Å layer. (b) 2−4 Å layer. (c) 4−6 Å layer. (d) 6−9 Å layer.

plexes. However, given the complexity of the surface of a protein it is difficult to do this. Some regions of the protein surface will have hydrogen bonding features from amino acid side chains or from backbone groups, while other side chains have weak or hydrophobic interactions with the surrounding water molecules. Thus, the overall estimation of the role of charge transfer is likely larger than one might qualitatively estimate by drawing an analogy with small hydrogen bonded complexes.

Another important consideration is the use of semiempirical models. More sophisticated ab initio or DFT methods might give different results than those observed herein (though, early indications show the differences are relatively small[53,54]). Thus, we feel that the safest interpretation of our results at this time is the observation that charge transfer cannot be ignored when considering the solvation of biomolecules. Whether the percentage role of this energy component is as high as 60−80% or whether it might be as low as 20% (as predicted by the Kitaura−Morokuma analysis for small hydrogen bonded clusters) is the key issue under debate and not whether this interaction is present or not. However, to reconcile recent experimental observations[38−43] that hydrogen bonding has covalent character one needs to move away from a purely electrostatic view of a hydrogen and begin to incorporate charge-transfer effects.

We realize that charge-transfer effects may be overpronounced in CspA, since all residues are more or less solvent-exposed. However, given the ubiquitous hydrogen bonding capabilities of biomolecular systems, we anticipate that charge

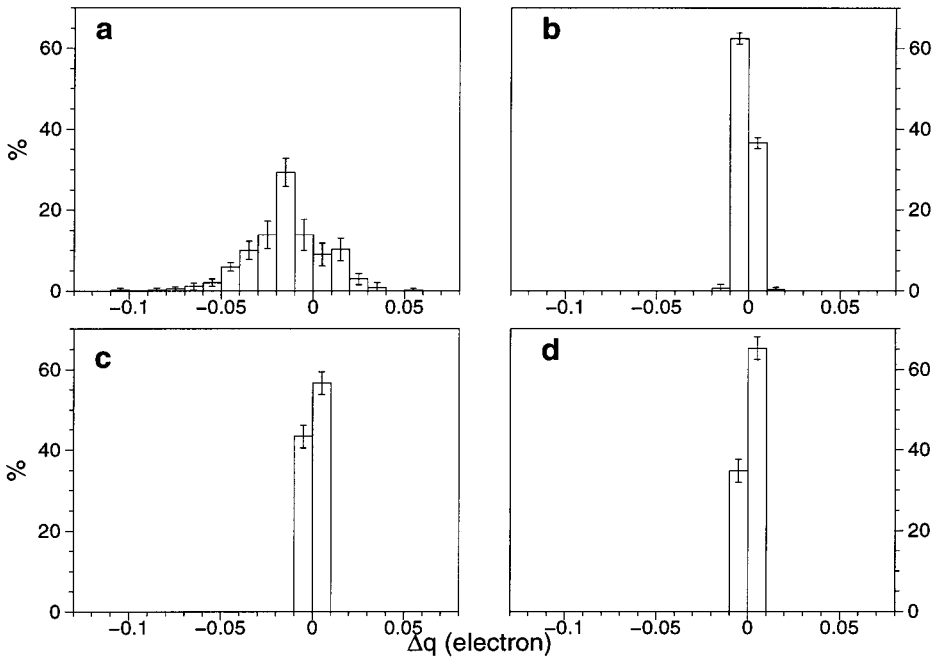

Figure 6. Effect of charge transfer on the charge distribution of water. AM1 results using CM2 charges. The length of the bars is twice the standard deviation. (a) 0−2 Å layer. (b) 2−4 Å layer. (c) 4−6 Å layer. (d) 6−9 Å layer.

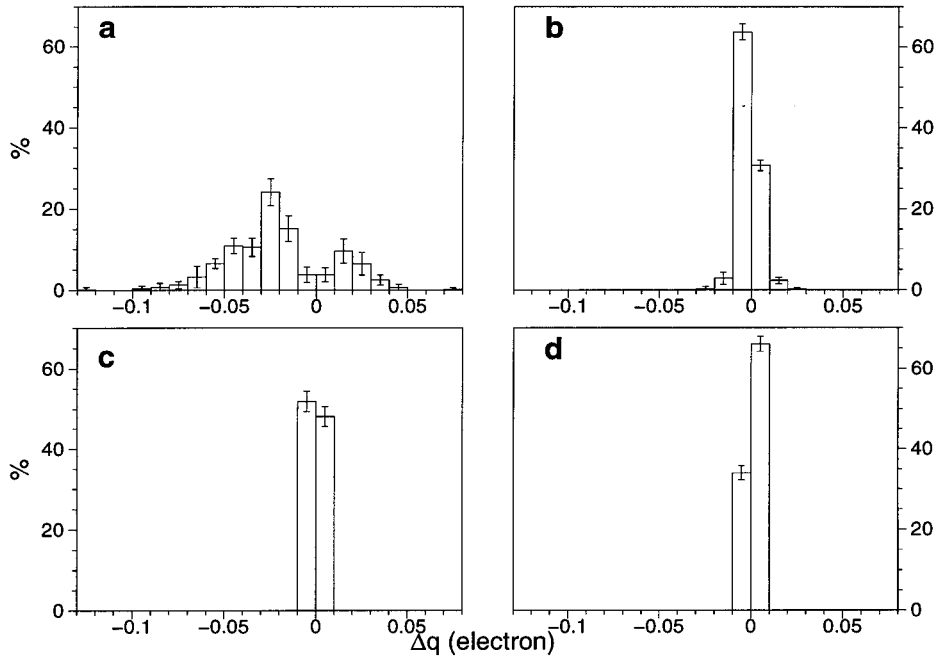

Figure 7. Effect of charge transfer on the charge distribution of water. PM3 results using CM2 charges. The length of the bars is twice the standard deviation. (a) 0−2 Å layer. (b) 2−4 Å layer. (c) 4−6 Å layer. (d) 6−9 Å layer.

transfer to the first solvation layer is important for other biomolecular systems as well. Moreover, these effects maybe important at critical interface regions other than the protein water interface. For example, these effects will clearly play a role at the DNA−water, DNA−protein, and protein−protein interfaces. These findings have important implications for techniques used to model biomolecular systems. Indeed, many of our basic theories of biomolecular solvation will need to be reformulated to account for these effects. In summary, our results indicate that quantum effects cannot be ignored in a theoretical description of solvated biomolecular systems.

Acknowledgment. We like to thank the National Center for Supercomputer Applications for generous allocation of supercomputer time. This work was supported by the Department of Energy (DE-FG02-96ER622670).

JA9912325