Hydrogen-Bonding Capabilities Based on Polarizability Studies of Model Peptide Systems

Shahul H. Nilar*,‡ and Tadeusz S. Pluta*,§

Contribution from the Biotechnology Research Institute, National Research Council of Canada, 6100 Ave Royalmount, Montreal, Canada H4P 2R2, and Department of Chemistry, Silesian University, 40-006 Katowice, Poland

Received February 14, 1995[⊗]

Abstract: The peptide bond is a fundamental unit in understanding the interactions between proteins and the surrounding medium. In this paper, models for the main chain and polar and nopolar side chains of amino acid residues were carefully chosen and complexes with each other and water clusters were calculated using a Monte Carlo/Simulated Annealing technique. The dipole polarizabilities, α, of these clusters were evaluated within the Coupled Perturbed Hartree−Fock (CPHF) method using special basis sets designed for the evaluation of electric response properties. The change in the interaction modified average polarizability per electron for each of the interacting subsystems, $\Delta\Delta\alpha$, is defined in the text, evaluated for each of the interacting systems, and used as a measure of the strength of the hydrogen bond and the identification of the hydrogen bond donor in each complex. The relative importance of main-chain, side-chain, and solvent effects in the models used to describe protein folding is discussed.

I. Introduction

The folding of peptides and proteins is closely associated with their biological activity, so much so, that the same system in a globular state can have properties which are very different from those exhibited in a different conformation. In certain cases, a change in the conformation of a protein or peptide can completely destroy the biological activity of the system.[1]

Molecular dynamic studies of biological systems[2,3] have shed some light on the problem of protein folding providing a better understanding of the mechanisms involved in the processes that contribute to the conformational changes in such large systems. These studies require a parametrized potential energy surface (PES), and consequently the results obtained are dependent on the quality of the PES in evaluating thermodynamic quantities, such as the free energy change, associated with the folding process. Recent simulations which incorporate solvent effects in the parametrization of the PES[4] have been shown to be promising, but the systems treated are small, essentially dipeptides.

The shape of a folded protein is largely dependent on the nature of the solvent and the primary structure or the amino acid sequence. If the residues are hydrophobic in nature and the solvent is polar, e.g., water, the protein tends to fold into a globular state; polar/charged side chains in the residues tend to favor a conformation that allows more exposure to a polar solvent. The interaction between the solvent molecules and the amino acid residues is one of the important factors that needs to be considered in evaluating the stability of a particular conformation, the others being back-bone−back-bone, side-chain−back-bone, and side-chain−side-chain interactions. The dominant factor among these effects will be the main contribution that decides the preferred conformation of the protein in the solvent.

Experimental measurements of the relative free energy changes that occur during hydrogen bond exchanges between atoms on amide systems and various solvents have been measured recently.[5] The amides were used as models for the back-bone structure of the protein and particular side chains of the constituent amino acid residues. The results showed that amides form stronger bonds with water molecules than with other amides, indicating that back-bone−water interactions are more important than back-bone−back-bone interactions. Furthermore, formamide, which resembles to some extent the side chain of asparagine and glutamine, has been shown to be as good a hydrogen bond donor as water, suggesting that side-chain−back-bone interactions can be as or more important in the folding process as back-bone−back-bone hydrogen bond interactions.

Guo and Karplus[6] evaluated the cooperative effect (ΔE_{coop}) using two antiparallel molecules of *N*-methylacetamide as a model for the peptide bond. ΔE_{coop} was defined as the difference between the energy required to break the hydrogen bond in their model for the peptide bond in the presence of solvent molecules such as water, ethanol, and formamide and in their absence. Based on their ab initio quantum chemical calculations, formamide as a donor molecule had a larger cooperative energy than water, even though the former is known to be a weaker hydrogen bond donor.

As the conformational flexibility of peptides is very dependent on the strength of the various hydrogen-bonding contributions, it is desirable to define a scheme that ranks the strength of the hydrogen bond between the solvent cluster, side chains, and peptide back bone. In this paper, we have approached the problem of protein back-bone, solvent cluster, and side-chain

‡ Biotechnology Research Institute. 1. Present Address: Department of Chemistry, The University of Alberta, Edmonton, AB T6G 2G2, Canada.

§ Silesian University.

⊗ Abstract published in *Advance ACS Abstracts*, November 15, 1995.

(1) Creighton, T. E. *Proteins: Structure & Molecular Properties*; W. H. Freeman: New York, 1993.

(2) Brooks, C. L., III; Karplus, M.; Pettitt, B. M. *Proteins: A Theoretical Perspective of Dynamics, Structure and Thermodynamics*; J. Wiley: New York, 1988.

(3) Sali, A.; Shakhnovich, E.; Karplus, M. *Nature* **1994**, *369*, 248−251.

(4) Smith, P. E.; Pettitt, B. M.; Karplus, M. *J. Phys. Chem.* **1993**, *97*, 6907−6913.

(5) Eberhardt, E. S.; Raines, R. T. *J. Am. Chem. Soc.* **1994**, *116*, 2149−2150.

(6) Guo, H.; Karplus, M. *J. Phys. Chem.* **1994**, *98*, 7104−7105.

0002-7863/95/1517-12603$09.00/0

interactions using a quantum mechanical approach. The distortion of the electron density, and hence the change in the dipole polarizabilities of the interacting systems during the formation of the hydrogen bond, was used to define the interaction modified average polarizability per electron $\Delta\alpha$ defined in the next section. Using this quantity, it has been possible to determine the relative hydrogen-bonding capabilities of molecules that have been used as models for the peptide bond and certain side chains and water clusters.

The very large number of atoms found in the molecular systems involved in biochemical processes necessitates the modeling of the protein by a simpler system and the solvent environment has to be modified accordingly. We were primarily interested in the interactions of the models for peptide bond moiety and the side chain of representative amino acids, with explicit solvent molecules and the resulting changes in the electron density of the interacting molecules.

N-Acetylalanine *N'*-methylamide (*N*-Ace-Ala-*N'*-Me) was used as a model for the backbone of the protein and the geometry was maintained at the fully extended configuration. This molecule is the smallest system, amenable to ab initio calculations, in which the structural effects on the central motif are similar to those found in a polypeptide. Formamide was used to represent the side chains of asparagine and glutamine, and benzene was chosen as a model for the hydrophobic side chains as found in phenylalanine, tryptophan, and tyrosine. In using such simplified models, the contribution to stabilization due to the conformational flexibility of these side chains and structural information specific to the aromatic part is lost. Nevertheless, these simplified models do give insight into the relative ease of hydrogen bonding as will be presented in Section 3.

Explicit water molecules were used to simulate the immediate solvent environment of the model systems described above using a Simulated Annealing/Monte Carlo method[7-9] based on the Fraga potential.[10] The geometries of the interacting systems were unchanged in the simulation, and this method of constructing solvent clusters has been shown to be useful in studying various molecular properties.[9,11,12]

II. Method of Calculation

When two systems A and B interact, their electron clouds or electronic charge distributions distort, and the extent of the distortion can be used as a measure of the strength of the interaction. In the case of hydrogen bonding, the donor will acquire a partial negative charge during interaction (due to the labile nature of the hydrogen atom/s being donated), and hence will have a larger polarizability than in the absence of such interactions. We intend using this change in the polarizability of the interacting systems to not only rank the interacting molecules by their hydrogen bonding capabilities but also decide which system is the donor in the interaction process. It should be noted that the total energy differences used in the literature to measure the strength of the hydrogen bond cannot decide on the donor in the process.

Initially, we calculate the polarizability per electron (ξ) of each isolated molecule defined as the average polarizability divided by the number of electrons in the molecule. The dipole polarizability tensors for the hydrated formamide clusters, the dimer of (*N*-Ace-Ala-*N'*-Me), its cluster with formamide and water, the benzene dimer, benzene–H_2O and benzene–NH_3 generated by the Simulated Annealing/Monte Carlo technique are evaluated next. The interaction modified average polarizability per electron, $\Delta\alpha^B$, of system B in the complex [A ⟛ B] defined as

$$\Delta\alpha^B = \sum_{i=x,y,z} \{(\alpha_{ii} - \alpha_{ii}^A)^2\}^{1/2}/\sqrt{2}N_{el,B}$$

is obtained from the dipole polarizability tensor for each system B in each cluster. α_{ii} and α_{ii}^A in the above equation are the *i*th components of the polarizability of the complex [A ⟛ B] and A, respectively, where system A is in the same orientation with respect to the laboratory frame as in the complex [A ⟛ B]. $N_{el,B}$ is the number of electrons in subsystem B.

The system A for the clusters which contain more than two molecules, e.g. formamide–$(H_2O)_n$ (n = 2, 3), is defined as that part, say C, of the cluster which does not contain the particular molecule B for which $\Delta\alpha^B$ is being calculated. For example, in evaluating $\Delta\alpha^B$ for the third water molecule in the trihydrated formamide system, system A (in the above equation) is obtained by deleting the third water molecule from formamide–$(H_2O)_3$. In this definition, the geometry of the rest of the cluster is maintained in the configuration obtained by the minimization process.

The interaction modified average polarizability per electron defined above gives a measure of the distortion in the electron cloud of system B due to interaction with system A. Our definition is not derived from first principles, but it is an intuitive manner in which to quantify the distortion of the electron density of the subsystems during hydrogen bonding and follows the definition of the experimentally accessible property, the polarization anisotropy.

The difference between $\Delta\alpha$ and ξ for subsystem B gives the *ease* of distortion of the electronic density of B during interaction with A. This difference, denoted $\Delta\Delta\alpha$, is the quantity we will use to compare various systems for their hydrogen-bonding capabilities. If $\Delta\Delta\alpha$ for a subsystem in a complex is larger than the corresponding value for the other subsystems in the same cluster, then the former molecule has a greater ease of distortion of its electron cloud, and hence a higher charge density (compared to its isolated or non-interacting state), and is therefore a better hydrogen bond donor. In this paper, we have focussed attention on hydrogen bonding between models for the main chain (back bone), side chain and explicit water molecules as described in the Introduction.

We have used the Coupled Perturbed Hartree–Fock (CPHF) method[13,14] to calculate the polarizabilities of the isolated systems and their clusters. The essence of the method is to use an optimized Hartree–Fock wave function obtained in the absence of the perturbation as the reference state in the perturbational expansion. First-order corrections to the unperturbed molecular orbitals are retained and the wave function of the system is calculated for the perturbed hamiltonian. CPHF, as an accurate method of including external perturbations at the Hartree–Fock level, does not introduce electron correlation energy[15] which is responsible for 10–15% of the total values for typical closed shell systems.[15,16] As our goal was not to obtain absolute values for the molecules/clusters but rather a set of consistent values for a series of systems for purposes of comparison, we could neglect the importance of electron correlation. The strength of the external electric field used as the perturbation was 0.001 au.

The choice of the one-electron basis set proved to be quite a challenge as the polarization functions are very important in obtaining a reasonable wave function for a system in the presence of an external interaction.[17-19] Although these functions do not follow directly from the Self Consistent Field Hartree–Fock (SCF-HF) calculations for

(7) Metropolis, N.; Rosenbluth, A. W.; Rosenbluth, M. N.; Teller, A. H.; Teller, E. J. *J. Chem. Phys.* **1953**, *21*, 1087–1092.
(8) Kirkpatrick, S. *J. Stat. Phys.* **1984**, *34*, 975–986.
(9) Nilar, S. H. *J. Comp. Chem.* **1991**, *12*, 1008–1013.
(10) Fraga, S. *J. Comp. Chem.* **1982**, *5*, 261–275.
(11) Fraga, S.; Thornton, S. *Theor. Chim. Acta* **1993**, *85*, 61–67.
(12) Nilar, S. H. *J. Chem. Soc., Perkin Trans. 2* **1992**, 2197–2199.
(13) Stevens, R. M.; Pitzer, R. M.; Lipscomb, W. N. *J. Chem. Phys.* **1983**, *78*, 550–559.
(14) Dykstra, C. E. *Ab Initio Calculation of the Structures and Properties of Molecules*; Elsevier: Amsterdam, 1988.
(15) Bartlett, R. J.; Purvis, G. D., III *Phys. Rev. A* **1979**, *20*, 1313–1322.
(16) Werner, H. J.; Meyer, W. *Mol. Phys.* **1976**, *31*, 855–872.
(17) van Duijneveldt, F. *IBM J. Res. Dev.* **1971**, RJ-945.
(18) Huzinaga, S. *Comput. Phys. Rep.* **1985**, *2*, 279–286.
(19) Dunning, T. H., Jr.; Hay, P. J. *Methods of Electronic Structure Theory*; Schaefer, F., III, Ed.; Pleneum Press: New York, 1977.
(20) Archibong, E. F.; Thakkar, A. J. *Phys. Rev. A* **1991**, *44*, 5478–5484.

Table 1. Energy, Cartesian Components of the Static Dipole Polarizability, α, Average Polarizability $\langle\alpha_{av}\rangle$, and Polarizability per Electron ξ (All Values in au)[a]

system	energy	α_{xx}	α_{yy}	α_{zz}	$\langle\alpha_{av}\rangle$	ξ
H_2O	−75.999 695	7.820	9.008	7.250	8.026	0.8026
	−76.052 914	8.491	9.172	7.822	8.495	0.8495
MBPT(2) (ref 31): Spackman basis		9.205	10.012	9.170	9.462	0.946
MBPT(2) (ref 31): Sadlej basis		9.760	10.064	9.572	9.799	0.9799
expt (ref 39)		9.91	10.32	9.55	9.93	0.993
NH3	−56.172 702	12.467	12.467	12.471	12.468	1.2468
	−56.212 139	12.742	12.742	13.268	12.918	1.2918
expt (ref 16)					14.82	1.482
benzene	−230.636 704	80.114	43.932	80.114	68.053	1.620
	−230.745 518	70.686	41.541	70.686	60.971	1.452
expt (ref 40)		79.16	44.13	79.16	67.48	1.607
formamide	−168.868 759	28.624	17.473	25.616	23.904	0.996
	−168.983 959	29.696	18.613	26.404	24.904	1.038
MBPT(2) (ref 31): Spackman basis		33.635	19.882	28.838	27.450	1.144
MBPT(2) (ref 31): Sadlej basis		34.376	20.649	29.643	28.223	1.176
expt (ref 41)					28.626	1.193
N-Ace-Ala-*N*′-Me	−492.676 906	101.305	95.176	69.051	88.511	1.135
	−492.984 456	103.189	96.427	70.399	90.005	1.154

[a] The first and second rows for each molecule are the results obtained using the Spackman and Sadlej basis sets, respectively.

atoms, they are very important in describing the distortion of the electron cloud when the system is subjected to an external field.[20-22]

Two one-electron basis sets were chosen for this work, the first due to Spackman[23] and the second set to Sadlej.[24] The Spackman basis augments the 6-31G basis set of Pople et al.[25] with s,d functions on the heavy atoms and s,p functions on the hydrogen atom, the values of the exponents of these augmented functions being optimized so as to maximize the average polarizability $\langle\alpha_{av}\rangle$ of first- and second-row AH_n hydrides. The Sadlej basis set uses the basis set polarization method[26,27] to augment the (5s) and (9s,5p) basis of van Duijneveldt[17] for H and C−F, respectively. The basis has been shown to give excellent values for the dipole, quadrupole moments, and dipole polarizability of small molecules.[28,29] The polarized basis sets are saturated with respect to their ability to respond to the perturbation caused by the external electric field, and the basis set superposition effects are small for interaction modified electric properties.[30] Details of the method and the effect of the various contractions are clearly presented in reference (26) and will not be repeated here.

Our preliminary studies[31] on formamide−$(H_2O)_n$ (n = 1, 2, 3), formamide−NH_3, and formamide clusters used the Finite Field (FF) method[14,32,33] of evaluating the dipole polarizabilities and first hyperpolarizabilities of molecules at the Self Consistent Field (SCF) and Second Order Many Body Perturbation Theory (MBPT) levels. The results indicated that the Interaction Modified Polarizability can be used as a measure of hydrogen-bonding capability.

In this paper we have evaluated the components of the dipole polarizabilities α and the first hyperpolarizabilities β of benzene, water, benzene−H_2O, benzene−NH_3, formamide−$(H_2O)_n$ (n = 1, 2, 3), *N*-Ace-Ala-*N*′-Me, *N*-Ace-Ala-*N*′-Me−H_2O complex, *N*-Ace-Ala-*N*′-Me−formamide, and the dimer of *N*-Ace-Ala-Nme. The choice of these systems was dictated by the models used for the amino acid side chains and the back bone of the peptide as discussed in the Introduction. Both one-electron basis sets were used and the response properties evaluated using the Coupled Perturbed Hartree−Fock (CPHF) method[13] as implemented in the TURBOMOLE[34,35] suite of programs. In the case of (*N*-Ace-Ala-*N*′-Me)$_2$, we could use only the Spackman basis due to the computational limitations in the TURBOMOLE programs. All calculations presented in this paper were obtained on an SGI 4D-280 machine.

Table 2. Interaction Energy ΔE (kcal/mol) Obtained Using the Simulated Annealing/Monte Carlo Method

complex	ΔE calcd	ΔE expt
benzene−H_2O	−1.78 (local minimum)	−1.63 to −2.78 (refs 46,51)
	−2.68 (global minimum)	
benzene−NH_3	−4.50	∼−5 (ref 47)
formamide−H_2O	−8.59	
formamide−$(H_2O)_2$	−21.32	
formamide−$(H_2O)_3$	−34.59	
formamide−$(H_2O)_4$	−47.84	
formamide−$(H_2O)_5$	−61.81	
(formamide)$_2$	−7.95	
N-Ace-Ala-*N*′-Me−H_2O	−11.08	
N-Ace-Ala-*N*′-Me−formamide	−9.76	
(*N*-Ace-Ala-*N*′-Me)$_2$	−17.00	

Experimental geometries of benzene, water, and NH_3 were obtained from the literature[36-38] and standard geometries were used for formamide and *N*-Ace-Ala-*N*′-Me. The reasoning for the latter choice was that in the extension of the method of calculation for the properties of larger peptide systems, e.g., $(Gly)_4$ in an α-helical conformation for which experimental geometries are not available, the same standard values could then be used without introducing a source of arbitrariness into the results due to a different geometry.

III. Results and Discussion

The results of the calculations are presented in Tables 1−4. The values obtained for the polarizabilities of the isolated

(21) Castro, M. A.; Canuto, S. *Phys. Lett. A* **1993**, *176*, 105−108.
(22) Darling, C. L.; Schlegel, H. B. *J. Phys. Chem.* **1994**, *98*, 5855−5861.
(23) Spackman, M. A. *J. Phys. Chem.* **1989**, *93*, 7594−7603.
(24) Sadlej, A. J. *Collect. Czech. Chem. Commun.* **1988**, *53*, 1995−2016.
(25) Hehre, W. J.; Radom, L.; Schleyer, P. v. R.; Pople, J. A. *Ab initio Molecular Orbital Theory*; John Wiley: New York, 1986; and references therein.
(26) Roos, B. O.; Sadlej, A. *J. Chem. Phys.* **1985**, *81*, 43−51.
(27) Wolinski, K.; Roos, B. O.; Sadlej, A. J. *Theor. Chim. Acta* **1985**, *68*, 431−442.
(28) Wolinski, K.; Sadlej, A. J.; Karlstrom, G. *Mol. Phys.* **1991**, *72*, 425−432.
(29) Wolinski, K.; Sadlej, A. *J. Mol. Phys.* **1992**, *75*, 221−231.
(30) Sadlej, J.; Roos, B. *Theor. Chim. Acta* **1989**, *76*, 173−189.
(31) Nilar, S. H.; Pluta, T. S. Symposium on Computational Chemistry, Proceedings of the Fifth Asian Chemical Congress, Kuala Lumpur, Malaysia 1993.
(32) Cohen, H. D.; Roothan, C. C. J. *J. Chem. Phys.* **1965**, *43*, S34−S38.
(33) Kurtz, H. A.; Stewart, J. J. P.; Dieter, K. M. *J. Comput. Chem.* **1990**, *11*, 82−87.
(34) Ahlrichs, R.; Bar, M.; Haser, M.; Horn, H.; Kolmel, C. *Chem. Phys. Lett.* **1989**, *162*, 165−169.
(35) TURBOMOLE 2.3; Biosym Technologies: San Diego, CA, 1993.
(36) Cabana, A.; Bachand, J.; Giguere, J. *Can. J. Phys.* **1974**, *52*, 1949−1955.
(37) Cook, R. L.; DeLucia, F. C.; Helminger, P. *J. Mol. Spectrosc.* **1974**, *53*, 62−76.
(38) Benedict, W. S.; Plyler, E. K. *Can. J. Phys.* **1957**, *35*, 1235−1241.

Table 3. Energy, Cartesian Components of the Static Dipole Polarizability, α, and Average Polarizability $\langle\alpha_{av}\rangle$ for the Molecular Complexes (All Values in au)[a]

complex	energy	α_{xx}	α_{yy}	α_{zz}	α_{av}
formamide−H_2O	−244.880 260	35.332	25.231	34.550	31.705
	−245.045 011	36.879	26.399	35.709	32.996
formamide−$(H_2O)_2$	−320.875 681	45.564	32.539	40.771	39.625
	−321.089 692	47.465	33.822	42.307	41.198
formamide−$(H_2O)_3$	−396.877 179	53.038	40.615	48.189	47.281
	−397.140 647	55.177	42.424	49.748	49.116
$(formamide)_2$	−337.739 328	49.788	41.186	49.768	46.914
	−337.966 749	51.337	42.748	51.576	48.553
benzene−H_2O	−306.640 269	85.416	54.729	86.004	75.383
	−306.801 147	86.386	56.454	86.860	76.567
benzene−NH_3	−286.790 927	86.577	59.363	86.329	77.423
	−286.940 089	87.426	61.062	87.398	78.628
$(benzene)_2$	−461.274 491	151.664	118.624	135.686	135.325
	−461.493 009	152.842	120.643	137.777	137.087
N-Ace-Ala-*N'*-Me−H_2O	−568.685 344	108.512	106.599	75.521	96.877
	−569.041 378	110.574	108.155	77.068	98.599
N-Ace-Ala-*N'*-Me−formamide	−661.535 506	122.688	122.759	88.514	111.320
	−661.954 803	124.821	124.585	90.741	113.382
$(N$-Ace-Ala-N'-Me$)_2$	−985.275 517	178.957	194.916	166.806	180.226

[a] The first and second rows for each complex are the results obtained using the Spackman and Sadlej basis sets, respectively.

Table 4. Values of the Interaction Modified Average Polarization, $\Delta\Delta\alpha$, for Each Subsystem in Each Complex

		$\Delta\Delta\alpha$	
complex	subsystem	Spackman	Sadlej
formamide−H_2O	formamide	0.240	0.240
	H_2O	0.158	0.148
formamide−$(H_2O)_2$	formamide	0.251	0.248
	$H_2O(1)$	0.192	0.180
	$H_2O(2)$	0.126	0.118
formamide−$(H_2O)_3$	formamide	0.217	0.216
	$H_2O(1)$	0.135	0.118
	$H_2O(2)$	0.148	0.131
	$H_2O(3)$	0.159	0.154
benzene−H_2O	benzene	0.389	0.576
	H_2O	0.143	1.062
benzene−NH_3	benzene	0.310	0.497
	NH_3	0.011	0.877
$(benzene)_2$	benzene(1)	0.443	0.632
	benzene(2)	0.357	0.776
$(formamide)_2$	formamide(1)	0.217	0.171
	formamide(2)	0.177	0.210
N-Ace-Ala-*N'*-Me−H_2O	*N*-Ace-Ala-*N'*-Me-H_2O	0.118	0.281
		0.256	0.238
N-Ace-Ala-*N'*-Me− formamide	*N*-Ace-Ala-*N'*-Me- formamide	0.254 0.179	0.249 0.168
$(N$-Ace-Ala-N'-Me$)_2$	*N*-Ace-Ala-*N'*-Me(1)	0.313	
	N-Ace-Ala-*N'*-Me(2)	0.266	

molecules are listed in Table 1. The first row under each molecule contains the results obtained using the Spackman basis while the Sadlej results for the same molecule are given in the corresponding second row. Our values for H_2O, NH_3, and formamide compare favorably with recent calculations.[42,43]

The Spackman basis gives results in excellent agreement with experiment for benzene. However, the results obtained using the Sadlej basis are a disappointment. The near-linear dependence of this very large basis set makes it a difficult calculation to converge, especially in dealing with multiple bonded systems as observed by other workers as well.[44,45]

Second Order Many Body Perturbation Theory (MBPT(2)) results for water and formamide are listed in Table 1, as an illustration of the applicability of these basis sets in evaluating polarizabilities. As the size of the systems in this study is large by quantum chemical standards, we have confined our attention to calculating the response properties at the SCF level only.

The results of the Monte Carlo simulations in Table 2 were obtained after repeating the calculations using various values for the step size—rotational and translational movement, the number of moves per "temperature", and the rate of cooling in the annealing step. From Table 2, three molecules of water can be considered to be closely associated with the formamide molecule as beyond this number the interaction energy per water molecule changes by less than 0.5 kcal/mol. The potential energy surface is not expected to be more accurate than this threshold value. We have thus focused our attention on the formamide−$(H_2O)_n$ (n = 1, 2, 3) complexes as being an adequate representation of the solvation of hydrophilic side chains of asparagine and glutamine.

Structural studies of benzene−H_2O and benzene−NH_3 complexes have recently shed some light on the hydrogen-bonding capabilities of benzene.[46,47] The importance of such interactions is profound in maintaining biologically active structures of proteins.[48,49] The agreement between the experimental and theoretical interaction energies for the model complexes is very satisfying indeed. Furthermore, the energetics and the orientation of the water and ammonia molecules over the aromatic ring are in excellent agreement with experiment.[50,51]

The interaction energies for formamide and *N*-Ace-Ala-*N'*-Me with one molecule of water obtained from the simulation technique show that amides form stronger hydrogen bonds with water than with other amides in agreement with experiment.[5] The formamide dimer and *N*-Ace-Ala-*N'*-Me−formamide have

(39) Murphy, W. F. *J. Chem. Phys.* **1977**, *67*, 5877−5882.
(40) Alms, G. R.; Burnham, A. K.; Flygare, W. H. *J. Chem. Phys.* **1975**, *63*, 3321−3326.
(41) Aroney, M. J.; LeFevre, R. J. W.; Singh, A. N. *J. Chem. Soc.* **1965**, 3179−3184.
(42) Voisin, C.; Cartier, A.; Rivail, J.-L. *J. Phys. Chem.* **1992**, *96*, 7966−7971.
(43) de Meras, A. M. S.; Aa Jensen, H. J.; Jorgensen, P.; Olesen, J. *Chem. Phys. Lett.* **1991**, *186*, 379−385.
(44) Lazzeretti, P.; Malagoli, M.; Zanasi, R. *Chem. Phys. Lett.* **1990**, *167*, 101−104.
(45) Dougherty, J.; Spackman, M. *Mol. Phys.* **1994**, *82*, 193−209.
(46) Sukui, S.; Green, P. G.; Bumgarner, R. E.; Dasgupta, S.; Goddard, W. A., III; Blake, G. A. *Science* **1992**, *257*, 942−945.
(47) Rodham, D. A.; Suzuki, S.; Suernam, H. D.; Lovas, F. J.; Dasgupta, S.; Goddard, W. A., III; Blake, G. A. *Nature* **1993**, *362*, 735−737.
(48) Mitchell, J. B. O.; Nandi, C. L.; Ali, S.; McDonald, I. K.; Thornton, J. M.; Price, S. L.; Singh, J. *Nature* **1993**, *366*, 413−413.
(49) Flocco, M. M.; Mowbray, S. L. *J. Mol. Biol.* **1994**, *235*, 709−717.
(50) Gotch, A.; Zwier, T. S. *J. Chem. Phys.* **1992**, *96*, 3388−3401.
(51) Cheng, B.-M.; Grover, J. R.; Walters, E. A. *Chem. Phys. Lett.* **1995**, *232*, 364−369.

significantly lower values for the energy of interaction than their aqueous complexes. The exception, however, is in the case of the dimer of *N*-Ace-Ala-*N'*-Me. Although the experimental results have been interpreted to indicate that main-chain−main-chain interactions are not as strong as main-chain−water interactions,[5] our calculations do not support this conclusion in agreement with other ab initio results on model systems.[6]

Results obtained for the polarizabilities of complexes are given in Table 3. Although it is possible to calculate the energies of interaction for these systems from the results listed, we chose not to do so as the effects due to Basis Set Superposition Error were not calculated.

In order to quantify the distortion of the electron density of the interacting molecules in the complexes studied, we evaluated $\Delta\Delta\alpha$ defined as

$$\Delta\Delta\alpha = \Delta\alpha - \xi$$

for all the complexes based on the results from each basis set. These results are listed in Table 4.

The better hydrogen bond donor will accumulate a greater quantity of electron density during bond donation and hence will have a higher value for $\Delta\Delta\alpha$. Considering the formamide−$(H_2O)_n$ complexes, formamide is the better hydrogen bond donor in each case (i.e., $n = 1, 2, 3$), and based on our model for the side chains of asparagine and glutamine, these amino acid residues are predicted to be more prevalent at the C-termini of α-helices in agreement with experiment.[52] In the case of the di- and trihydrated complexes of formamide, the different hydrogen bond accepting properties of the water molecules are clearly reflected in their values of $\Delta\Delta\alpha$, with one molecule of water being the most effective acceptor.

In the case of the nonpolar side chain (as modeled by benzene), the results obtained for the benzene−H_2O and benzene−NH_3 complexes using the Spackman basis predict the water and NH_3 molecules behave as a proton acceptor which is not in agreement with experimental observations.[46,47] Although the Sadlej basis set predicts the correct behavior, we do not intend laying emphasis on these results hereafter in this paper due to the inherent problems associated with this basis applied to benzene as mentioned in the preceding paragraphs of this section. Polar and nonpolar side-chain interactions were modeled as the formamide and benzene dimers, respectively, and according to the Spackman basis, the second benzene molecule behaves as the acceptor of the hydrogen bond, as in the case of formamide. The Sadlej basis predicts the opposite behavior.

Based on the Spackman results of the clusters with at least one benzene molecule, the presence of more polarization functions that sample the intermolecular region of space is necessary for the evaluation of response properties of the clusters. Although the Spackman basis is adequate in evaluating the dipole polarizability of single molecules, more diffuse functions are necessary, especially for nonpolar molecular clusters.

The Spackman basis set gives a higher degree of distortion for the water molecule in the complex *N*-Ace-Ala-*N'*-Me−H_2O, thereby indicating that the water molecule behaves as a hydrogen bond donor in the interaction, in agreement with experiments.[5] *N*-Ace-Ala-*N'*-Me−formamide is the model used in this work to describe the main-chain−polar side-chain interactions, and based on our results, the main chain is predicted to behave as a H bond donor as in the case of (*N*-Ace-Ala-*N'*-Me)$_2$. The strength of this interaction is predicted to be of the same order of magnitude as the polar side-chain−water term and weaker than the main-chain−water interaction which is much less than the main-chain−main-chain interaction (Table 2). It is interesting to note that the main chain is always the hydrogen bond donor, except in the case of the interaction with water.

In this paper we have evaluated the strength of interactions in hydrated model systems and complexes which have been used as models that represent important molecular structures that are implicated in the processes of protein folding. The distortion of the electron clouds of the interacting systems, based on ab initio calculations of their polarizabilities using two basis sets, have been used to predict which of the systems would act as hydrogen bond donors. The results obtained give a method by which one can decide on the entity which is behaving as a hydrogen bond donor. The comparative study of the basis sets indicates that the Spackman basis is adequate and economical for the study of most intermolecular complexes, except those with a nonpolar subsystem, for which more diffuse functions are required. The Sadlej basis, though it works well for small singly bonded systems, is difficult to converge due to linear dependence problems and is very time consuming, in comparison.

JA950523K

(52) Chakrabartty, A.; Doig, A. J.; Baldwin, R. L. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 11332−11336.