The Effect of Aqueous Solvation upon α-Helix Formation for Polyalanines

Pedro Salvador,[†] Amaparo Asensio,[‡] and J. J. Dannenberg*[,§]

Department of Chemistry, City University of New York, Hunter College and the Graduate School, 695 Park Avenue, New York, New York 10021, Institute of Computational Chemistry and Department of Chemistry, University of Girona, 17071 Girona (Catalonia), Spain, and Facultad de Farmacia, Departamento Quimica Organica, Universidad de Valencia, Avda Vicente Andres Estelles, 46100 Burjassot (Valencia), Spain

Received: March 8, 2007; In Final Form: April 13, 2007

The incremental free energies of aqueous solution for acetyl(ala)$_N$NH$_2$ in its extended unfolded and α-helical conformations are compared using the SM5.2 solvation method of Cramer and Truhlar. A combination of density functional theory (DFT) at the B3LYP/D95(d,p) and AM1 has been employed using the ONIOM method. The incremental solvation energies of α-helical structures are very similar for both ONIOM and AM1 optimized structures as these structures do not significantly change upon solution. However, the conformations of the unfolded peptides change from extended β-strand to polyproline II conformations upon aqueous solution. The incremental solvation free energy per residue of the polyproline II structure is about 2 kcal/mol/residue greater than that for the α-helix, representing an upper limit for the difference between the solvation energies. However, most of this difference disappears when the energy required to distort the optimized gas-phase extended β-strand structure to the optimized polyproline II solution structure is included in the analysis, leaving an estimated difference in incremental solvation free energy of 0.3−0.5 kcal/mol favoring the unfolded structure. The solution structure sacrifices the stability derived from the intramolecular C_5 H-bonds for more favorable interactions with the aqueous solvent.

The equilibrium between α-helical and unfolded polyalanines has been the subject of several recent experimental[1−11] and theoretical (i.e., nonempirical) reports.[12−14] Whereas the experimental observations generally pertain to polyalanines in (usually aqueous) solution (with some exceptions[15]), the theoretical calculations modeled the gas-phase system (again with some exceptions[16]). One needs to know the relative solvation energies for the unfolded and α-helical states to properly compare the gas-phase calculations to the data for the system in aqueous solution

Although the peptide/protein folding problem is often treated as if it were a two-state problem, clearly this is ambiguous at best. As it pertains to the α-helix, the helical structure might be reasonably considered one conformation or state despite the fact that it is probably dynamically folding and unfolding, especially near the termini. However, the unfolded "state" can consist of a myriad of different structures. Furthermore, while the extent of α-helicity is generally measured directly from its CD or NMR spectra, the unfolded state is rarely measured directly. We have suggested[17] that the energetic effects of solvation on the folded/unfolded equilibrium should depend mostly upon the solvation of the unfolded state as the solvent can influence which among the myriad of conformations might be most favored. Strictly speaking, the helix will be only one of these conformations. Although the solvent can affect its energy relative to the other (unfolded) conformations, the solvent cannot have much effect upon the α-helical structure as it would cease to be helical.

Recently, focus has shifted to solvation of the unfolded state.[18,19] Shi and co-workers recently reported polyproline II to be the predominant conformational structure in aqueous solutions of polyalanines that are too short to form helices.[20,21] Molecular dynamics calculations based upon the Monte Carlo procedure have reproduced this observation.[22,23] While Scheraga and co-workers have pointed out that many other conformations are surely present in aqueous solution,[24−28] there seems to be little dispute that polyproline II is one of those present.[25] As such, it must be representative of the energy of the short solvated polyalanines. Even if it did not represent the global minimum in solution, its energy could not be much higher if it is a significant component of the equilibrium mix.

In this paper, we assess the solvation energies of the α-helical and unfolded capped polyalanines acetyl(Ala)$_N$NH$_2$ ($N = 2-18$) that we have previously geometrically optimized[12] using an all quantum mechanical ONIOM[29−31] procedure, and whose relative enthalpies we have previously reported.[13]

Methods

The details of the ONIOM[30,32,33] geometric optimizations of α-helices and β-strands of the polyalanines were presented in a previous report.[12] We took the relevant molecular geometries from this study as starting points for the current calculations. We could not use the program available that couples the Cramer/Truhlar solvation scheme[34] with the GAUSSIAN 03[35] suite of programs used for the original optimizations because of conflict between the sizes of the peptides and the limitations of the program in its current form. Instead, we used the AMPAC 8.15 program,[36] which contains the SM 5.2 version of the AMSOL[34] solvent model. Using AMPAC, all calculations used the AM1[37] semiempirical Hamiltonian together with the SM 5.2 solvation method. Using the fixed ONIOM optimized helical geometries,

* To whom correspondence should be addressed. E-mail: jdannenberg@gc.cuny.edu.
[†] University of Girona.
[‡] Universidad de Valencia.
[§] Hunter College of CUNY.

10.1021/jp071899a CCC: $37.00
Published on Web 06/07/2007 PAGE EST: 4.1

TABLE 1: Calculated Solvation and Distortion Energies (kcal/mol) for Unfolded Acetyl(ala)$_N$NH$_2$[a]

	gas-phase structure		solution structure		distortion energy		solvation energy	
N	total	incremental	total	incremental	total	incremental	total	incremental
2	−25.89		−29.73					
3	−32.35	−6.46	−37.20	−7.47	4.35		−32.85	
4	−38.81	−6.46	−44.74	−7.55	5.92	1.57	−38.82	−5.98
5	−45.27	−6.46	−52.34	−7.60	7.56	1.63	−44.79	−5.96
6	−51.65	−6.38	−60.03	−7.69	9.26	1.71	−50.77	−5.98

[a] See Figure 1 for definitions.

TABLE 2: Incremental Change in Solvation Free Energy for α-Helices (kcal/mol)[a]

N	incremental $\Delta\Delta G_{solv}$	average of last three $\Delta\Delta G_{solv}$
10	*−5.51*	
12	*−5.78*	
13	−4.83	−5.60
14	−6.00	
15	−5.47	−5.43
16	−4.60	−5.36
17	−6.11	−5.39
18	−5.04	−5.25
average	−5.46	

[a] As no stable α-helices could be obtained for $N = 9$ and 11, the values for $N = 10$ and 12 represent average values over two increments. These entries are in italics and are counted twice in the overall average. The last column gives the average of the last three incremental values (where available).

we performed single-point AM1/SM5.2 calculations. These were then repeated allowing the helices to relax to their AM1-optimized geometries to test the sensitivity of the solvation model to these small differences in the helical geometries. We deemed the average differences on the calculated solvation energies between the helices in their ONIOM and AM1 unsolvated optimized geometries to be sufficiently small (0.19 kcal/mol/residue) to validate this procedure. We calculated the AM1/SM5.2 solvation energy in water of the β-strand (unfolded) structures for $N = 2-6$, both in their previously optimized ONIOM geometries and in their AM1-relaxed geometries. The distortion energies for the gas-phase structures were taken from the energy differences between the optimized gas-phase ONIOM calculations and similar structures that were constrained to retain the backbone dihedral angles, ϕ and ψ, of the AM1-optimized solvated geometries for the helices.

Results and Discussion

The relevant data are collected in Tables 1 and 2. The tables contain the incremental solvation free energies for the different species. We define the incremental free energies as the difference in solvation energies between acetyl(ala)$_N$NH$_2$ and acetyl(ala)$_{N-1}$NH$_2$. The importance of the incremental free energies lies in their independence from end effects (including the effect of the other amino acid residues used to enhance the solubilities of the peptides in the experimental studies). One might expect them to become constant for N's that exceed some threshold value. We shall discuss unfolded and α-helical states separately and then compare the two.

Figure 1 depicts a thermodynamic cycle that can be applied to either the helix or unfolded state. The energetic quantities are the vertical solvation energies, where the solvent interacts with a molecule in a fixed geometry (either that of the gas phase or the solvated molecule), and the distortion energies required to change the geometries from those of the gas phase to those in solution. These can either be destabilizing (the distortion from the extended β-stand of the gas phase to the polyproline II-like structure of the solution in the gas phase) or stabilizing (the same geometric change in solution becomes a relaxation energy).

Unfolded States. As previously reported, the minimum energy geometry calculated by the ONIOM (density functional theory (DFT)/AM1) method for the gas-phase completely unfolded state without traditional internal H-bonds corresponds to an extended β-strand (Figure 2).[12] This structure owes part of its stability to cyclic C_5 H-bonding interactions which interact somewhat cooperatively with each other.[38,39] When we calculate the hydration energies that correspond to these (ONIOM gas phase) structures using the single-point AM1/SM5.2 procedure, we find that the incremental solvation free energies become roughly constant at −6.4 kcal/mol/residue. The optimized solvated structures calculated using the AM1 Hamiltonian in water as simulated by the SM5.2 method assume a polyproline II conformation in accord with both experimental[20,21] and theoretical[22,23] reports (see Figure 3). This conformation replaces the cyclic C_5 H-bonding stabilizations of the extended β-strand with apparent H-bonding with solvent water as all the H-bond acceptors (C═O's) and donors (N−H's) are directed away from the peptide backbone into the solvent.

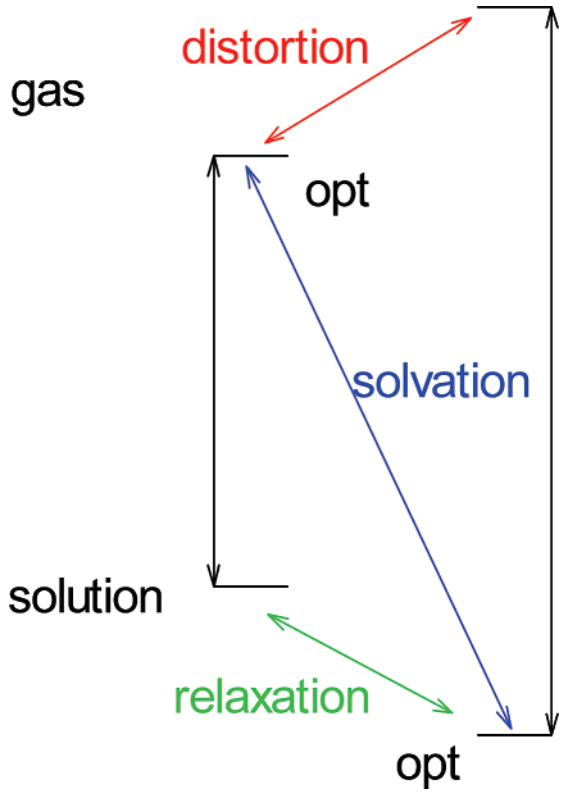

Figure 1. Thermodynamic cycle depicting the solvation energy of a gas-phase species and the four component steps discussed in this work. The cycle can be used to illustrate the behavior of any thermodynamic state function (e.g., enthalpy, free energy). The optimized geometries for both gas and solution phases are indicated with "opt". The vertical black arrows indicate the vertical solvation contributions to the state function for the (fixed) gas- and solution-phase structures.

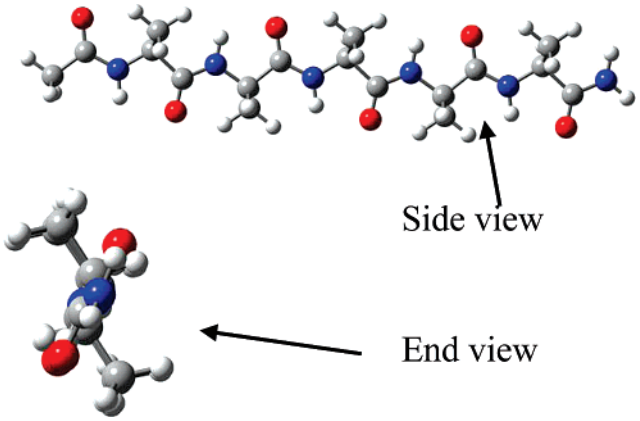

Figure 2. Gas-phase extended β-strand structure of an unfolded capped polyalanine.

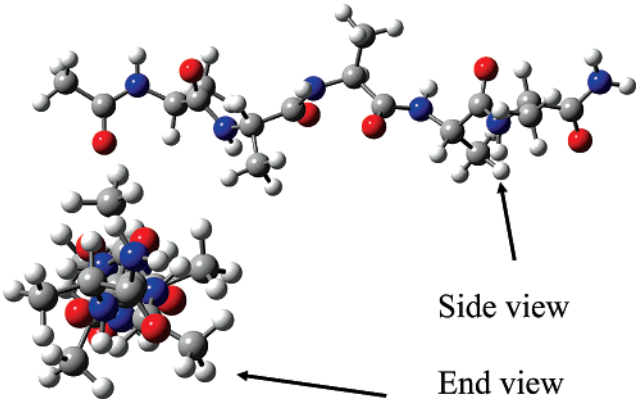

Figure 3. Polyproline II conformation of an unfolded capped polyalanine as optimized in aqueous solvent using AM1 and the SM5.2 solvation method.

Unlike the helices (discussed below), the gas-phase AM1-optimized open structures differ considerably from those calculated by DFT. Thus, the single-point AM1 distortion energy could not be reliably calculated by comparing the single-point AM1 energy of the open structures at their ONIOM geometries with the AM1-optimized solvated structures. We, therefore, calculated the vertical AM1 solvation energy of these structures at the optimized AM1 solvated geometries of Figure 3. Optimizations of these structures using AM1/SM5.2 increase the magnitude of the incremental solvation stabilization, which no longer is constant but very slightly increases in magnitude by 0.05 to 0.07 kcal/mol/residue. However, the incremental distortion energy in the gas phase increases by almost the same amount. Consequently, the incremental solvation energy (which is the sum of the vertical solvation and distortion energies) remains constant at −5.98 kcal/mol/ala. The constancy of both of these incremental solvation free-energy values allows us to use linear extrapolation to calculate the corresponding solvation parameters for longer β-strands in either of the two conformations.

The probable explanation for the systematic decrease in the incremental solvation free energy and matching increase in the incremental distortion energy as N increases lies in the small cooperative interactions previously reported for chains of C_5 H-bonds.[38−40] As the polyalanine becomes longer, the energy required to break the C_5 interactions in going from the extended chain (gas phase) structure to the polyproline II (aqueous structure) should increase.

The distortion energies posed an unanticipated problem, as the AM1 gas-phase optimized geometry differs from the quasi-planar structure obtained using the ONIOM calculations. To estimate the gas-phase distortion energies, we used the ONIOM procedure to reoptimize the relaxed AM1 solvated structures holding the dihedrals fixed (only the bond lengths and valence angles were varied).

The calculated incremental hydration values (−7.5 to −7.7 kcal/mol/residue) for the solution structure (see Table 1) are remarkably close to those originally reported (−7.9 kcal/mol/residue)[41] from an electrostatic model, while they are slightly smaller in magnitude than a more recent similar assessment (−8.51 kcal/mol/residue) by the same group.[42] These incremental hydration values differ substantially from those measured[43,44] and calculated for simple amides,[41,45] as previously noted.[41,42]

α-Helices. The solvation free energies of the α-helices fixed in either the AM1 or DFT gas-phase optimized geometries differ only by approximately 0.19 kcal/mol/residue. For example, ΔG_{solv} for $N = 18$ is −116.2 and −112.7 kcal/mol for the α-helix optimized using either AM1 or ONIOM, respectively, in the gas phase.

The results of single-point solvation calculations using the gas-phase ONIOM optimized structures appear in Table 2. We could only obtain optimized α-helices for polyalanines with $N = 8$, 10, and 12 or greater.[12] Consequently, we could only obtain incremental solvation free energies for $N = 13$ and higher. However, we can obtain solvation free-energy differences over

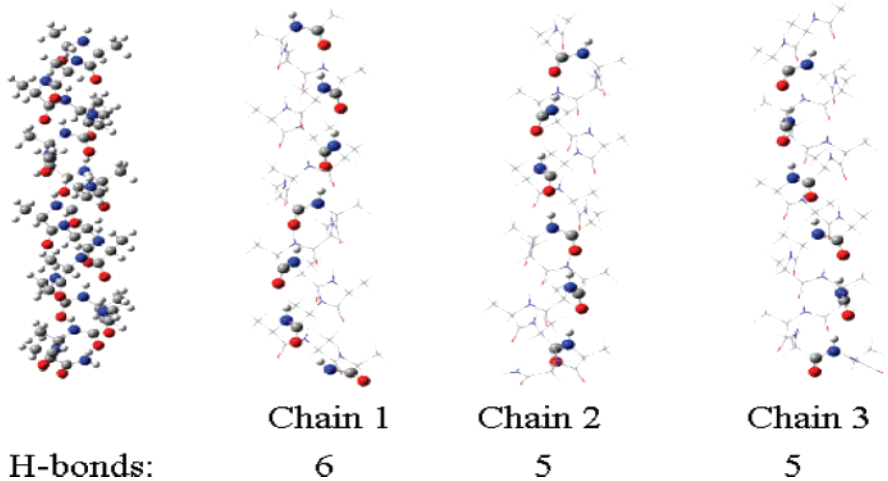

Figure 4. The structure of α-helical acetyl(ala)$_{18}$NH$_2$ (left) with the three H-bonding chains and their H-bonds individually emphasized as ball and sticks with the remaining atoms indicated as wire frame.

two increments for $N = 10$ and 12. We use the average over the two increments for these values of N in Table 2. The incremental solvation energies vary somewhat as N increased from 13 to 18. This apparent periodic variation probably reflects the periodic growth in the H-bonding chains of α-helices, which have three distinct chains of amide H-bonds (see Figure 4). Those helical structures with $N = 14$ and 17 have three H-bonding chains of the same length (4 and 5, respectively). Those with $N = 13$ and 16 have one H-bonding chain that contains one fewer H-bond than the others, while those with $N = 15$ and 18 have one H-bonding chain that contains one more H-bond than the others.

Relative Solvation Energies. The incremental solvation free energies for the α-helical and gas-phase optimized extended structures are remarkably similar to each other. If one uses the average value from Table 2, they differ by 0.52 cal/mol/residue. The values of Table 2 represent the single-point solvation of the gas-phase DFT optimized structures. As noted above, the solvation energy of the AM1 optimized structures is more stabilized by about 0.19 kcal/mol. Including this value would reduce the difference in incremental solvation free energies between the isomeric structures to 0.27 kcal/mol/residue. This finding is consistent with the previously reported[13] agreement of the calculated enthalpies/residue with those reported in aqueous solution for polyalanines[2,6,11] and with the observation that these values increase with N.[11] At least part of this agreement must be attributed to the coincidental and fortuitous choice of the optimized extended β-strands as the reference for comparison with the helices. The approximately 2 kcal/mol difference per residue in solvation energies between the polyproline II conformations and the α-helices could be taken as an upper limit to the overall difference in these energies as polyproline II represents the global minimum energy conformation for the hydrated unfolded polyalanines. Consequently, whatever other conformations that might be prevalent in the gas phase (including the fully extended β-strand that we have used as a reference) would be less stable in aqueous solution. Thus, the energy required to distort that conformation to the polyproline II structure favored in aqueous solution would reduce that upper limit.

Conclusion

The differential incremental solvation free energies for extended β-strands and α-helices of acetyl(ala)$_N$NH$_2$ in aqueous solution do not make significant contributions to the energetics of the equilibria between the helical and unfolded states. While the (arbitrary) choice of the extended β-strand conformations as the reference for the unfolded state in the gas phase might be fortuitous, the upper limit for the differential solvation per residue of about 2 kcal/mol will be reduced by whatever distortion energy is required to convert whatever gas-phase structure considered to that corresponding to the most stable solution conformation while in the gas phase. The solution structure sacrifices the stability derived from the intramolecular C_5 H-bonds for more favorable interactions with the aqueous solvent.

Acknowledgment. This work was supported in part by grants from PSC/CUNY. J.J.D. thanks DURSI of the Generalitat de Catalunya for a visiting professorship at the Universitat de Girona. J.J.D. thanks Prof. Christopher Cramer for helpful discussions.

References and Notes

(1) Rohl, C. A.; Fiori, W.; Baldwin, R. L. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 3682−3687.
(2) Marqusee, S.; Robbins, V. H.; Baldwin, R. L. *Proc. Natl. Acad. Sci. U.S.A.* **1989**, *86*, 5286−5290
(3) Lopez, M. M.; Chin, D.-H.; Baldwin, R. L.; Makhatadze, G. I. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 1298−1302.
(4) Kemp, D. S.; Oslick, S. L.; Allen, T. J. *J. Am. Chem. Soc.* **1996**, *118*, 4249−55.
(5) Kemp, D. S.; Allen, T. J.; Oslick, S. L.; Boyd, J. G. *J. Am. Chem. Soc.* **1996**, *118*, 4240−8.
(6) Kennedy, R. J.; Tsang, K.-Y.; Kemp, D. S. *J. Am. Chem. Soc.* **2002**, *124*, 934−944.
(7) Miller, J. S.; Kennedy, R. J.; Kemp, D. S. *J. Am. Chem. Soc.* **2002**, *124*, 945−962.
(8) Wallimann, P.; Kennedy, R. J.; Miller, J. S.; Shalongo, W.; Kemp, D. S. *J. Am. Chem. Soc.***2003**, *125*, 1203−1220.
(9) Heitmann, B.; Job, G. E.; Kennedy, R. J.; Walker, S. M.; Kemp, D. S. *J. Am. Chem. Soc.* **2005**, *127*, 1690−1704.
(10) Kennedy, R. J. W.; Sharon, M.; Kemp, D. S. *J. Am. Chem. Soc.* **2005**, *127*, 16961−16968
(11) Job, G. E.; Kennedy, R. J.; Heitmann, B.; Miller, J. S.; Walker, S. M.; Kemp, D. S. *J. Am. Chem. Soc.* **2006**, *128*, 8227−8233.
(12) Wieczorek, R.; Dannenberg, J. J. *J. Am. Chem. Soc.* **2004**, *126*, 14198−14205.
(13) Wieczorek, R.; Dannenberg, J. J. *J. Am. Chem. Soc.* **2005**, *127*, 14534−14535.
(14) Morozov, A. V.; Tsemekhman, K.; Baker, D. *J. Phys. Chem. B* **2006**, *110*, 4503−4505.
(15) Chin, W.; Piuzzi, F.; Dognon, J.-P.; Dimicoli, I.; Tardivel, B.; Mons, M. *J. Am. Chem. Soc.* **2005**, *127*, 11900−11901.
(16) Wu, X.; Wang, S. *J. Phys. Chem. B* **2001**, *105*, 2227−2235.
(17) Dannenberg, J. J. *Adv. Protein Chem.* **2006**, *72*, 227−273.
(18) Baldwin, R. L. Weak interactions in protein folding: hydrophobic free energy, van der Waals interactions, peptide hydrogen bonds, and peptide solvation. In *Protein Folding Handbook*; Buchner, J., Kiefhaber, T., Eds.; Wiley-VCH Verlag GmbH & Co.: Weinheim, Germany, 2005; Vol. 1, pp 127−162.
(19) Chin, D.-H.; Woody, R. W.; Rohl, C. A.; Baldwin, R. L. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 15416−15421.
(20) Shi, Z.; Woody, R. W.; Kallenbach, N. R. *Adv. Protein Chem.* **2002**, *62*, 163−240
(21) Shi, Z.; Olson, C. A.; Rose, G. D.; Baldwin, R. L.; Kallenbach, N. R. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 9190−9195.
(22) Mezei, M. F.; Patrick, J.; Srinivasan, R.; Rose, G. D. *Proteins: Struct., Funct., Bioinf.* **2004**, *55*, 502−507.
(23) Kentsis, A.; Mezei, M.; Gindin, T.; Osman, R. *Proteins: Struct., Funct., Bioinf.* **2004**, *55*, 493−501.
(24) Makowska, J.; Rodziewicz-Motowidlo, S.; Baginska, K.; Makowski, M.; Vila, J. A.; Liwo, A.; Chmurzynski, L.; Scheraga, H. A. *Biophys. J.* **2007**, biophysj.106.097550.
(25) Makowska, J.; Rodziewicz-Motowidlo, S.; Baginska, K.; Vila, J. A.; Liwo, A.; Chmurzynski, L.; Scheraga, H. A. *PNAS* **2006**, *103*, 1744−1749.
(26) Scheraga, H. A.; Vila, J. A.; Ripoll, D. R. *Biophys. Chem.* **2002**, *101−102*, 255−265.
(27) Ingwall, R. T.; Scheraga, H. A.; Lotan, N.; Berger, A.; Katchalski, E. *Biopolymers* **1968**, *6*, 331−368.
(28) Platzer, K. E. B.; Ananthanarayanan, V. S.; Andreatta, R. H.; Scheraga, H. A. *Macromolecules* **1972**, *5*, 177−187.
(29) Vreven, T.; Morokuma, K. *J. Comput. Chem.* **2000**, *21*, 1419−1432.
(30) Dapprich, S.; Komiromi, I.; Byun, K. S.; Morokuma, K.; Frisch, M. J. *THEOCHEM* **1999**, 0166−1280.
(31) Morokuma, K. *Philos. Trans. R. Soc. London, Ser. A: Math., Phys. Eng. Sci.* **2002**, *360*, 1149−1164.
(32) Svensson, M.; Humbel, S.; Froese, R. D. J.; Matsubara, T.; Sieber, S.; Morokuma, K. *J. Phys. Chem.* **1996**, *100*, 19357−19363.
(33) Morokuma, K. *Bull. Korean Chem. Soc.* **2003**, *24*, 797−801.
(34) Cramer, C. J.; Truhlar, D. G. *Chem. Rev.* **1999**, *99*, 2161−2200.
(35) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, Jr., J. A.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A.

G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; Pople, J. A.; *Gaussian 03*, revision C.01; Gaussian, Inc.: Wallingford, CT, 2004.

(36) Dewar, M. J. S. *Mod. Tech. Comput. Chem. MOTECC* **1991**.

(37) Dewar, M. J. S.; Zoebisch, E. G.; Healy, E. F.; Stewart, J. J. P. *J. Am. Chem. Soc.* **1985**, *107*, 3902−3909.

(38) Viswanathan, R.; Asensio, A.; Dannenberg, J. J. *J. Phys. Chem. A* **2004**, *108*, 9205−9212.

(39) Horvath, V.; Varga, Z.; Kovacs, A. *J. Phys. Chem. A* **2004**, *108*, 6869−6873.

(40) Wieczorek, R.; Dannenberg, J. J. *J. Am. Chem. Soc.* **2005**, *127*, 17216−17223.

(41) Avbelj, F.; Luo, P.; Baldwin, R. L. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 10786−10791.

(42) Avbelj, F.; Baldwin, R. L. *Proteins: Struct., Funct., Bioinf.* **2006**, *63*, 283−289.

(43) Della Gatta, G.; Barone, G.; Elia, V. *J. Solution Chem.* **1986**, *15*, 157−167.

(44) Makhatadze, G. I.; Privalov, P. L. *J. Mol. Biol.* **1993**, *232*, 639−659.

(45) Dannenberg, J. J. *J. Phys. Chem. A* **2006**, *110*, 5798−5802.