Salting in Peptides: Conformationally Dependent Solubilities and Phase Behavior of a Tripeptide Zwitterion in Electrolyte Solution

John S. Perkyns, Yiyang Wang, and B. Montgomery Pettitt*

Contribution from the Chemistry Department, University of Houston, Houston, Texas 77204−5641

Received July 19, 1995. Revised Manuscript Received December 4, 1995[⊗]

Abstract: The dielectrically consistent interaction site model theory is applied to a model system consisting of a zwitterion tripeptide of sequence Gly-Ala-Gly at infinite dilution in a solvent mixture of water and sodium chloride with only the central ϕ,ψ dihedrals for conformational degrees of freedom. The peptide is found to be salted into solution by the cosolvent, with its solubility depending strongly on conformation of the central ϕ,ψ pair. The distribution of cosolvent relative to the bulk solvent mixture is examined and relatively little specific association is found. Instead, the ionic concentration around the peptide is increased, especially near the phase boundary, and the increased concentration extends up to eight solvent diameters into the bulk. The consequences of such an ionic distribution on thermodynamic measures of association are discussed. Similarities between this model system and the unusual solubility behavior of β-lactoglobulin found in experiments are shown. The molecular distributions calculated are found to be consistent with a separation of the solvent mixture into two non-miscible phases, one of which contains the solute and has a higher cosolvent concentration than the other. Since turbid solutions have been observed in the β-lactoglobulin system, it is suggested that a separation into two liquid phases could be common to both systems.

(1) Introduction

The stability and solubility of biological molecules has been the subject of intense interest in both experimental and theoretical science for some time.[1−11] Enough experimental work has now been done, particularly on proteins and peptides in solvent mixtures, that the overall trends in the results have given rise to several general principles, as well as clarifying the thermodynamic quantities that exhibit them. Most stability and solubility phenomena have been codified by Timasheff and collaborators[5−8] in terms of the complex balance between cosolvent exclusion from a region around a protein and specific interactions between the cosolvent and specific sites on the surface of a protein. Many cosolvents, which in the literature include ionic species such as NaCl, have been heuristically classified according to their relative tendency to become excluded from the surface region, or to associate with the protein. These are the general principles we wish to examine for the present model system. We wish to discover whether the molecular level information available from statistical mechanical theory is consistent, in terms of these principles, with the energetic information calculated from the same theory. Specifically we would like to see if the exclusion or association of Na^+ or Cl^- ions in the region around a conformationally restricted model tripeptide zwitterion of sequence Gly-Ala-Gly is correlated with the trends in the solvation free energy of the zwitterion that would be predicted from a knowledge of the trends seen in experimental results. Direct comparison to experiments on Gly-Ala-Gly would be inappropriate due to the conformational restrictions imposed on the Gly residues.

Theoretical work in this field has lagged behind experimental work, largely because established methods were incapable of examining, in a calculationally tractable manner, the important cases of biological molecules in molecular solvent mixtures. A recent extension in the area of statistical theories of site−site models of the liquid state has made it possible to examine many interesting systems using ion/water mixtures as solvent. The statistical mechanical theory we use here to investigate all these considerations is the Dielectrically Consistent Reference Interaction Site Model theory (DRISM).[12,13] The DRISM theory is an approximate theory which has given qualitatively correct structural and thermodynamic quantities for a variety of systems[12−16] and has given insight into the nontrivial trends in the solubilities of alkali halides[17] and nonpolar molecules in salt solution[14,16,18] and into the common ion effect.[18] Recently the theory was used to demonstrate that peptide conformations can be restricted by overall solution stability.[15] This phenomenon will be elucidated further here in the general context of the salting-in and salting-out of peptides and proteins.

Also of particular interest is the comparison of our results to those from the experiments[5] on β-lactoglobulin (β-LG). At first sight one would imagine a relatively small molecule such as Gly-Ala-Gly would have little in common with a much larger peptide or globular protein. The solubility behavior of β-LG

[⊗] Abstract published in *Advance ACS Abstracts,* January 15, 1996.

(1) Schellman, J. A. *Biopolymers* **1975**, *14*, 999.
(2) Schellman, J. A. *Biopolymers* **1978**, *17*, 1305.
(3) Schellman, J. A. *Biopolymers* **1987**, *26*, 549.
(4) Schellman, J. A. *Biophys. Chem.* **1990**, *37*, 121.
(5) Arakawa, T.; Timasheff, S. N. *Biochemistry* **1987**, *26*, 5147.
(6) Arakawa, T.; Bhat, R.; Timasheff, S. N. *Biochemistry* **1990**, *29*, 1914.
(7) Arakawa, T.; Bhat, R.; Timasheff, S. N. *Biochemistry* **1990**, *29*, 1924.
(8) Timasheff, S. N. *Curr. Opin. Struct. Biol.* **1992**, *2*, 35.
(9) Marlow, G. E.; Perkyns, J. S.; Pettitt, B. M. *Chem. Rev.* **1993**, *93*, 2503.
(10) Marlow, G. E.; Pettitt, B. M. *Adv. Comp. Biol.* **1994**, *1*, 231.
(11) Schellman, J. A. *Biopolymers* **1994**, *34*, 1015.
(12) Perkyns, J. S.; Pettitt, B. M. *Chem. Phys. Lett.* **1992**, *190*, 626.
(13) Perkyns, J. S.; Pettitt, B. M. *J. Chem. Phys.* **1992**, *97*, 7656.
(14) Perkyns, J. S.; Pettitt, B. M. *Biophys. Chem.* **1994**, *51*, 129.
(15) Perkyns, J. S.; Pettitt, B. M. *J. Phys. Chem.* **1995**, *99*, 1.
(16) Perkyns, J. S.; Pettitt, B. M. *J. Phys. Chem.*, in press.
(17) Perkyns, J. S.; Pettitt, B. M. *J. Phys. Chem.* **1994**, *98*, 5147.
(18) Valdeavella, C. V.; Perkyns, J. S.; Pettitt, B. M. *J. Chem. Phys.* **1994**, *101*, 5093.

0002-7863/96/1518-1164$12.00/0 © 1996 American Chemical Society

is, however, distinctive.[19,20] It shows a strong tendency to be salted in by NaCl and its solubility is also very dependent on solution pH. Previous theoretical work[15] has shown that a model of our chosen zwitterion tripeptide also has a strong tendency to be salted in by NaCl, and its solubility is strongly dependent on molecular conformation. Some time ago it was proposed for the β-LG case that the solubility is a consequence of the size of the molecular dipole[19] which in turn is determined, in part, by solution pH. We wish to see here if conformationally induced changes in the molecular dipole of Gly-Ala-Gly give rise to solubility behavior corresponding to pH-induced changes in the molecular dipole of β-LG.

In summary, this paper has two goals. The first is to demonstrate that a statistical mechanical theory of the liquid state predicts the salting-in of a model tripeptide in NaCl solution, and to examine the similarities between this model system and the experimental results for β-lactoglobulin. The second is to determine whether the molecular quantities directly obtainable from calculations agree with the trends deduced from experiment for protein stability and solubility in general.

The remainder of this paper is divided as follows. Section 2 contains the background necessary to describe the DRISM theory as well as those quantities we have chosen to use in our comparison to experimental results. Section 3 contains selected results from calculations and a discussion of their interpretation. Section 4 briefly summarizes the conclusions which have been drawn.

(2) Theory and Thermodynamic Quantities: The DRISM Theory

It is our main purpose here to interpret the results obtained by a statistical mechanical theory of the liquid state to see whether the molecular level structural quantities obtained directly agree with the molecular interpretation of the thermodynamic results of experiment. For this purpose we will use the DRISM integral equation theory. The DRISM theory has been presented in detail elsewhere[12,13] and need only be summarized here.

Interaction site model (ISM) theories are used to calculate radial distribution functions, $g_{ij}(r)$, which give the probability density of finding site or atom i on one molecule at a distance r from site or atom j on another molecule (possibly of the same species). The distribution functions are sought because, in principle, they can be used to calculate all equilibrium thermodynamic properties of the system while simultaneously yielding a structural perspective. Models for intermolecular interactions are chosen with atom based sites and are defined by the pair potential between sites. Here we choose the common Coulomb plus Lennard-Jones form

$$u_{ij}(r) = \frac{q_i q_j}{r} + 4\epsilon_{ij}\left[\left(\frac{\sigma_{ij}}{r}\right)^{12} - \left(\frac{\sigma_{ij}}{r}\right)^{6}\right] \tag{1}$$

where σ_{ij} and ϵ_{ij} are the usual Lennard-Jones diameter and well depth parameters. For intramolecular interactions we use a common molecular mechanics form and parameters.[21,22]

Distribution functions are calculated by simultaneously solving a set of two equations which relate the total correlation functions, $h_{ij}(r) = g_{ij}(r) - 1$, to the direct correlation functions, $c_{ij}(r)$. The direct correlation functions can be thought of as defined by one of a related set of site−site Ornstein−Zernike Fourier space integral equations. The appropriate equation from the DRISM theory can be written as[12,13]

$$(\boldsymbol{\rho}\tilde{\boldsymbol{h}}\boldsymbol{\rho} - \tilde{\boldsymbol{\chi}}) = (\tilde{\boldsymbol{\omega}} + \tilde{\boldsymbol{\chi}})\tilde{\boldsymbol{c}}(\tilde{\boldsymbol{\omega}} + \tilde{\boldsymbol{\chi}}) + (\tilde{\boldsymbol{\omega}} + \tilde{\boldsymbol{\chi}})\tilde{\boldsymbol{c}}(\boldsymbol{\rho}\tilde{\boldsymbol{h}}\boldsymbol{\rho} - \tilde{\boldsymbol{\chi}}) \tag{2a}$$

where $\tilde{\boldsymbol{\chi}}$ is a matrix consisting of elements

$$\tilde{\chi}_{ij}(k) = -j_0(-kd_{ix})j_0(-kd_{iy})j_1(-kd_{iz})j_0(kd_{jx})j_0(kd_{jy})j_1(kd_{jz})[\rho_v \tilde{h}_c(k)] \tag{2b}$$

and d_{ix} is the distance between site i and its molecular center of charge in the x-direction (with the dipolar direction arbitrarily chosen as the z-direction). Also, $j_0(kd)$ and $j_1(kd)$ are zeroth and first order spherical Bessel functions, respectively. The form of eq 2b guarantees that the functions $\tilde{h}_{ij}(k)$ will have the correct asymptotic form in the limit of small k. The function $\tilde{h}_c(k)$, common to all elements of $\tilde{\chi}$, is given by

$$\tilde{h}_c(k) = \left[\frac{\epsilon - 1}{y\rho_v} - 3\right]\exp(-ak^2), \qquad a > 0 \tag{2c}$$

and we use a value of $a = 0.15$ throughout. The matrix $\tilde{\boldsymbol{h}}$ is constructed so that its (i,j)th element is the function $\tilde{h}_{ij}(k)$, the Fourier transform of the corresponding total correlation function, which can be calculated via

$$\tilde{h}_{ij}(k) = 4\pi \int_0^\infty r^2 h_{ij}(r)\frac{\sin(kr)}{kr}\,dr \tag{3}$$

The matrix $\tilde{c}$ is defined in a similar way. The matrix $\tilde{\boldsymbol{\omega}}$, representing intramolecular correlations, has elements $\tilde{\omega}_{ij}(k) = j_0(kd_{ij})$ for rigid molecules, where d_{ij} is the (positive, fixed) distance between site i and site j within the same molecule, and when $d_{ij} = 0$, $\tilde{\omega}_{ij}(k) = 1$. For mixtures, when sites i and j are on different species, $\tilde{\omega}_{ij}(k) = 0$. The matrix $\boldsymbol{\rho}$ is a diagonal matrix with the (i,i)th element being the molecular number density of the species in which site i is located. Throughout, v and u denote the solvent and solute, respectively.

The second equation, known as a closure, relates $c_{ij}(r)$ and $h_{ij}(r)$ in a mathematically independent way. For this work we shall use the hypernetted-chain (HNC) equation[23] exclusively,

$$c_{ij}(r) = \exp[-\beta u_{ij}(r) + t_{ij}(r)] - 1 - t_{ij}(r) \tag{4}$$

where $t_{ij}(r) = h_{ij}(r) - c_{ij}(r)$, and $\beta = 1/k_B T$ with k_B the Boltzmann constant.

Equations 2a and 4, when solved together, will be referred to below as the dielectrically consistent RISM, or DRISM theory.

For all the following discussions the free energy of adding a molecule to the solution is a central quantity. The solvation free energy for placing a single solute molecule in an infinite solvent bath, in the DRISM/HNC formulation, can be written as[14,24]

$$\beta\Delta\mu^{(0)}_{u,\mathrm{sol}} = \rho_v \sum_{i=1}^{n_u}\sum_{j=1}^{n_v} \int \left(\frac{1}{2}[h^{(0)}_{ij}(r)]^2 - c^{(0)}_{ij}(r) - \frac{1}{2}h^{(0)}_{ij}(r)c^{(0)}_{ij}(r)\right) d\mathbf{r} \tag{5}$$

where $\Delta\mu$ is the chemical potential (free energy per particle) and the superscript (0) indicates infinite dilution.

We choose, as a measure of the relative level of interaction between the peptide and the solvent−cosolvent mixture, the quantity $\rho_i G_{iu}$, which can be written in terms of distribution functions as

$$\rho_i G_{iu} = 4\pi\rho_i \int_0^\infty r_{iu}^2 h_{iu}(r_{iu})\,dr_{iu} \tag{6}$$

and is the product between the number density of either the solvent or cosolvent and the appropriate "Kirkwood G".[23] This quantity can be interpreted simply as the difference between the number of molecules of species i that are found in the region of solution around the solute and the number that would have been there if the entire region were replaced by the equilibrium bulk solvent mixture. There is no specific boundary around the region of solution where the solute resides. The integrand used to calculate the quantity G_{iu} approaches zero as the

(19) Cohn, E. J.; Ferry, J. D. Interactions of Proteins with Ions and Dipolar Ions. In *Proteins, Amino Acids and Peptides*; Cohn, E. J., Edsall, J. T., Eds.; Reinhold: New York, 1943; Chapter 24, p 586.

(20) Palmer, A. H. *J. Biol. Chem*. **1934**, *104*, 359.

(21) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. *J. Comput. Chem*. **1983**, *4*, 187.

(22) MacKerell, A. D., Jr.; Karplus, M. All-hydrogen empirical potential for Molecular Modeling and Dynamics Studies of Proteins: CHARMM22 force field; submitted.

(23) Hansen, J. P.; McDonald, I. R. *Theory of Simple Liquids*, 2nd ed.; Academic Press Inc.: London, 1986.

(24) Singer, S. J.; Chandler, D. *Mol. Phys*. **1985**, *55*, 621.

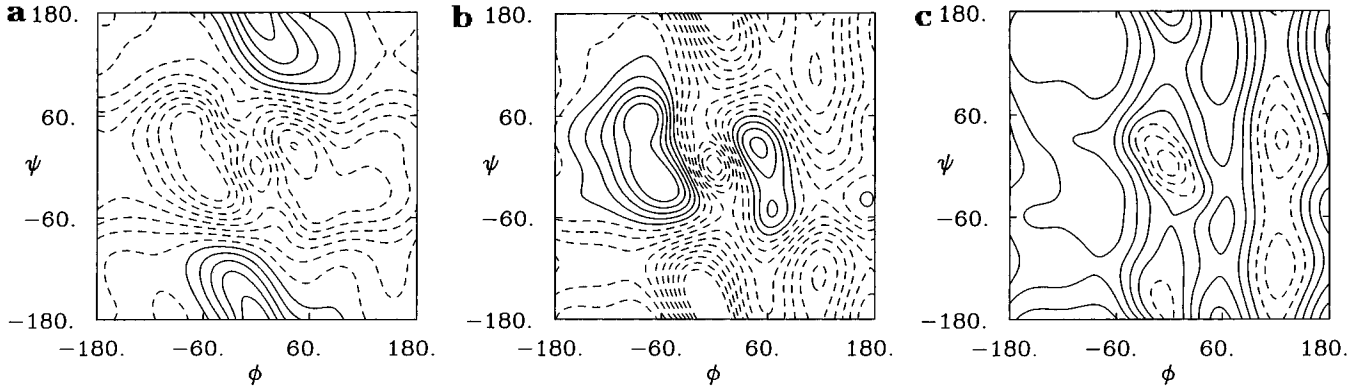

Figure 1. Calculated Ramachandran type plots for fixed structures of the Gly-Ala-Gly zwitterion at infinite dilution in pure water: (a) relative chemical potential of solvation, (b) relative gas-phase conformational energy, and (c) relative total free energy of solvation. All energy quantities are in kcal/mol and all contours are 1.5 kcal/mol apart. The five lowest contours are solid lines, and the rest are dashed lines. See the text for the definitions of the angles ϕ and ψ and the methods of calculation.

average solvent structure in the perturbed region around the solute approaches the average bulk solvent structure, typically between 3 and 20 solvent diameters away from the solute surface.

Three things should be noted before we examine the calculated values. Firstly, the quantity $\rho_i G_{iu}$ is usually negative, because it includes both the number of molecules (or ions) which have been excluded from the space taken up by the solute and those molecules displaced in the region around the solute. It is, therefore, only the differences between values for different solvent mixtures but the same solute structure which are easily interpreted. Secondly, the number of cosolvent cations found in the region around the solute is equal to the number of cosolvent anions in this region, thus maintaining global charge neutrality for the solvent mixture itself. This reflects the situation which would occur if a single solute molecule were added to an infinite bath of salt/water solution. For this study, the solute Gly-Ala-Gly is neutral so we have overall charge neutrality. Global charge neutrality among finite concentration components must be maintained or the DRISM/HNC theory will not permit a solution since such a condition implies that at least one of the Fourier transform integrals (eq 3) will not converge. The strict enforcement of charge neutrality conditions that accurately reflect real systems is a property common to most statistical mechanical theories of the liquid state which employ the HNC equation. Finally, the peptide structures with identical fixed values of the central conformational dihedral angles ϕ and ψ, when calculated using different solute parameters representing the different salt solutions (see above), were virtually identical to each other and so displaced negligibly different volumes of solvent. Therefore, differences in $\rho_i G_{iu}$ are almost completely due to changes in the solvent structure for the model in the same (ϕ,ψ) conformation with different salt concentrations.

The thermodynamic and structural quantities presented below were calculated from statistical mechanical theory in the following manner. Each model structure of the zwitterion tripeptide with sequence Gly-Ala-Gly was created using the united atom parametrized force field developed for the computer program CHARMM.[21,22] All the backbone dihedral angles were fixed in the trans position, except the angles $CNC_\alpha C$ and $NC_\alpha CN$ in the central alanine residue, which we denote ϕ and ψ, respectively. The angles ϕ and ψ were each fixed to multiple values throughout their possible ranges of −180° to 180°. Each structure was then minimized in energy with the solvent influence included (for this peptide structural minimization procedure only) as a continuum with a dielectric constant equivalent to that found experimentally[25] for each salt and water mixture. The water and ion distributions around these minimized structures were then calculated using the statistical mechanical DRISM theory.[12,13] The (molecular) solvent model used was similar to that of TIP3P[26] water and the ion models for Na^+ and Cl^- were from the program QUANTA.[27] The ion and water interaction parameters are given in Table 1. The dielectric constant inherent in the molecular solvent mixture distributions thus obtained (see ref 13) was consistent with that used in the energy minimization of each particular tripeptide structure. All calculations were done using the experimental densities of water and NaCl at 1 atm of pressure and a temperature of 298 K.[28] A series of structures was calculated at salt concentrations of 0.0, 4.0, and 5.0 M, in all cases with Gly-Ala-Gly at infinite dilution. The dielectric constants used for the solvent mixtures were 78.54,[28] 39.65, and 33.55 (interpolated from ref 25), respectively. The solvent mixture contributions to the thermodynamics of the system were calculated from the distribution functions as described in Section 2. To be consistent with the literature on proteins in solvent mixtures,[1−4,6−8,11] we will often refer below to water as the solvent, NaCl as the "cosolvent", and the Gly-Ala-Gly zwitterion as the solute.

Table 1. Intermolecular Interaction Parameters for the NaCl[27] and H_2O[26] Models Used

i	q_i (electrons)	ϵ_{ii} (kcal/mol)	σ_{ii} (Å)
Na^+	1	0.026	2.940
Cl^-	−1	0.260	3.671
O	−0.834	0.152	3.151
H	0.417	0.020	0.400

(3) Results and Discussion

a. The Salting-In of Gly-Ala-Gly by NaCl. The chemical potential of solvation μ_{sol}, relative to the minimum value, as a function of the angles ϕ and ψ is given in Figure 1a for Gly-Ala-Gly in pure water. The most striking feature of this figure is the similarity of the contours in the region of highest chemical potential to the corresponding region of smallest dipole moment, which is shown in Figure 2. An inverse relation between strength of local electric field and solvation chemical potential has also been observed for simple ions in water.[14] The solvent responds to the solute electric field by restructuring in such a way as to oppose the local field, and in doing so decreases the internal energy of the system. The new solvent structure is also more ordered, and such a decrease in entropy opposes the decrease in internal energy, but it is not large enough to dominate the internal energy and the net effect on the free energy is a decrease. Figure 1a shows only the solvent contribution to the relative solvation free energy, however. This basic trend in internal energy vs entropy was seen some time ago in pairwise additive free energies of solvation.[29]

It should be reemphasized that the free energy surfaces displayed here using eq 5 are not explicitly pairwise additive.

(25) Behret, H.; Schmithals, F.; Barthel, J. *Z. Phys. Chem. Neue. Folge.* **1975**, *96*, 73.

(26) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926.

(27) The parameters for Na^+ and Cl^- were taken from page 153 of the Parameter Handbook of the computer program QUANTA, Version 3.3. QUANTA has been developed by Molecular Simulations, Inc.

(28) Weast, R. C., Ed. *CRC Handbook of Chemistry and Physics*, 66th ed.; CRC Press, Inc.: Boca Raton, FL, 1986.

(29) Pettitt, B. M.; Karplus, M. *J. Phys. Chem.* **1988**, *92*, 3994.

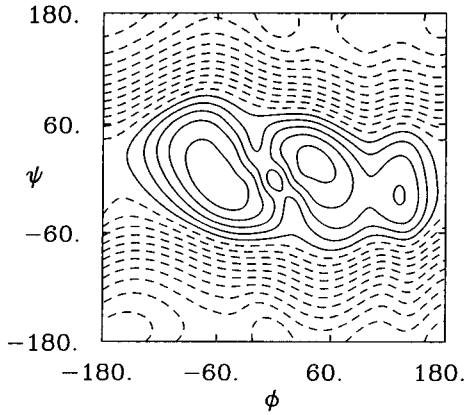

Figure 2. The dipole moment (in D) of the Gly-Ala-Gly zwitterion at infinite dilution in 5 M NaCl, relative to the lowest value represented. The contours are 1 D apart with the lowest five solid lines and the rest dashed lines.

The contribution to the total free energy surface from the intramolecular energy of the solute only is given in Figure 1b. Regions (for instance near $\phi = \psi = 0$) where the minimized structure cannot avoid the repulsive overlap of electron clouds exhibit the highest (least favorable) energies. Regions of lowest energy occur when the solute is in conformations which most successfully compromise the opposing tendencies to have oppositely charged regions of the molecule in close proximity and also avoid the short−range repulsive overlap for a given valence force field structure. The largest negative contribution to the former energetic tendency will be to have the molecule fold in such a way as to bring the charged ends of the molecule and hydrogen bonding groups together, thus minimizing the energy penalty incurred when the dipole is increased. This dipole minimization effect dominates the intramolecular energy when there is no significant spacial overlap of atoms on the tripeptide. Thus we have two contributions to the net relative free energy of solvation, one which is roughly proportional and the other which is roughly inversely proportional to the size of the dipole. The dipole dependent contributions in Figures 1a and 1b are approximately equal and opposite in size, so the effect of the dipole on each of these energy contributions is largely cancelled when they are added, leaving the more subtle structural effects exemplified by the relative total free energy of solvation in Figure 1c. The global minimum in the total energy occurs near the point ($\phi = -90°$; $\psi = 150°$) where the molecule is almost completely extended into the β-sheet form. There is also a relative minimum consistent with the right-handed α-helix form located nearby ($\phi = -80°$; $\psi = -70°$). The energy barrier between these two structures is only about 2 kcal/mol, so facile transitions between them can be expected in pure water at normal temperatures. Another pair of local minima can be near the points ($\phi = 60°$; $\psi = -60°$), which is near the C_7 axial form, and ($\phi = 60°$; $\psi = 120°$) which is near the left-handed α-helix form. These two minima have an energy barrier between them of a similar height. The transition between these two sets of minima is much less likely, since the path with the lowest barrier is roughly 4 kcal/mol in size.

Figures 3a and 3b give results corresponding to Figures 1b and 1c, except they show results for a high (5 M) cosolvent concentration. It is somewhat surprising to see that the overall shape of the contours representing the solvent chemical potential, μ_{sol}, including contributions from water and Na^+ and Cl^- ions (Figure 3a), is similar to that for the pure water case (Figure 1a). The solution with high cosolvent concentration has a larger energy range than the pure water case with a surface that is noticeably steeper due to the higher cost of charge separation

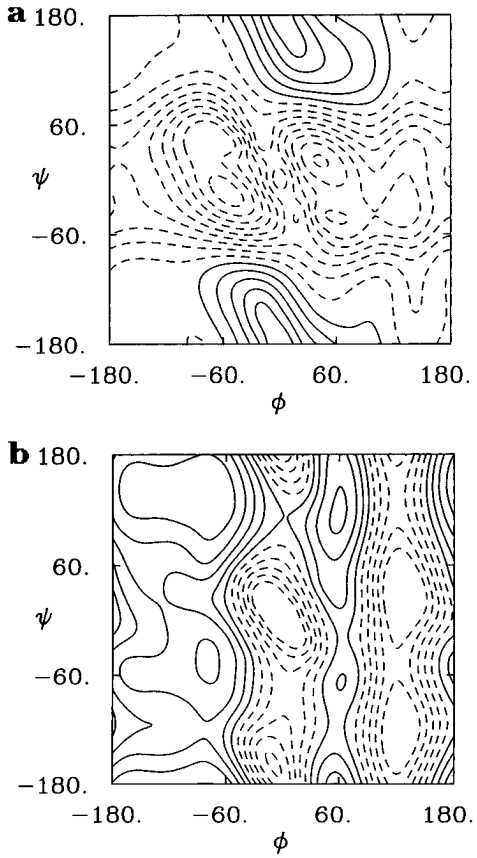

Figure 3. Calculated Ramachandran type plots for rigid structures of the Gly-Ala-Gly zwitterion at infinite dilution in a 5 M NaCl solution: (a) relative chemical potential of solvation, and (b) relative total free energy of solvation. The contours are as in Figure 1.

(enlarging the dipole) in a medium with a lower dielectric constant. The relative intramolecular energies for peptide structures in the case of 5 M salt solution (not shown) are very similar to those of Figure 1b except once again the range of energies is larger due to the decreased dielectric constant of the solvent medium. The minimized structures of the zwitterion are quite insensitive to changes in dielectric constant especially when the backbone dihedral angles are constrained to fixed values. Since the two contributions to the total relative solvation have differences from the pure water case which are largely equal and opposite, the total relative free energy surface for the peptide in 5 M NaCl solvent (Figure 3b) is also similar in overall shape to that of the pure water solvent case (Figure 1c). The range of energies is larger for the high salt concentration case by about 10%. One consequence of the wider range in energies is that local minima have higher energetic boundaries, and thus represent structures which would be longer lived in solution.

When the salt concentration lies between the two examples given above, the calculated results can differ quite considerably. As an example, the equivalent calculations for a 4 M NaCl solvent mixture have been performed and are shown in Figures 4a and 4b. The conformational contribution (not shown) is again very similar to both the 0 and 5 M salt solution cases for the same reasons as were given above. The solvent contribution is quite different, however. There are large regions of the energy surface, where ϕ and ψ angles are close to the trans configuration, where no solution to the DRISM theory was possible. This type of phenomenon has been demonstrated before in many similar contexts,[15,30−32] and the lack of solution usually corre-

(30) Stell, G.; Wu, K. C.; Larsen, B. *Phys. Rev. Lett.* **1976**, *37*, 1369.

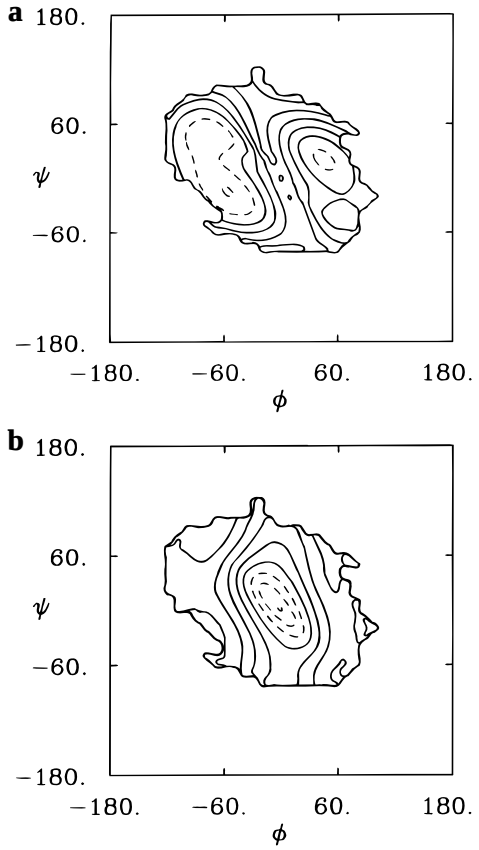

Figure 4. Calculated Ramachandran type plots for rigid structures of the Gly-Ala-Gly zwitterion at infinite dilution in a 4 M NaCl solution: (a) relative chemical potential of solvation, and (b) relative total free energy of solvation. The contours are as in Figure 1. The thick solid line outlines the region within which solutions to the DRISM theory were possible.

sponds to some form of thermodynamic phase instability. Those regions of the energy surface which are available, however, are also remarkably similar to the 5 M solution result. For solvent mixtures with NaCl concentrations below 4 M the region of accessible solutions decreases, and when the concentration has been decreased to 3 M no solutions to the theory for this model system are possible. This is a clear indication of the salting-in of a peptide with increasing salt concentration. It is surprising since NaCl is usually classified among those cosolvents which salt peptides out of solution.

The apparent thermodynamic stability boundary seems to have little effect on the relative stability of those conformations which can exist in a single liquid phase with the solvent mixture. The free energy of solvation, therefore, is unlikely to be the thermodynamic criterion which gives rise to such a phase stabilaity phenomenon. The search for more sensitive thermodynamic criteria which govern such phenomena has been of much interest for some time.[17,18,31,33,34] Below we argue that a possibility for the type of phase instability which apparently occurs, and is consistent with the solvent and cosolvent distributions, is the separation of the solution into two liquid phases, one in the region of the peptide with higher salt concentration, and the other of lower salt concentration in the remaining bulk solvent. Experimentally, exceptions to the classification of NaCl as a salting-out agent, such as the case of β-lactoglobulin, have been found.[5,19] We will examine the implications of our resutls in this context below.

b. Examination of Specific Association and Stability. The salt NaCl has been classified as a cosolvent which is both a salting-out agent and a stabilizing agent.[35,36] Furthermore, the explanation of these properties in general is given in terms of the relative hydration of, or cosolvent association with, the biological molecule in question. NaCl is expected to be preferentially excluded from the surface of a biological solute and to stabilize those structures which minimize the volume of the preferentially excluded zone. For experimental results these molecular level interpretations have to be inferred from thermodynamic data. Statistical mechanical theory can demonstrate both molecular level phenomena and their associated thermodynamic effects. We will now examine thermodynamic and average structural quantities to determine if this model of NaCl has other unexpected behavior, in addition to the phase phenomenon described above, for this model peptide system. In particular we wish to see if the principles established in the literature[8] for predicting solvent/cosolvent structure from thermodynamic quantities do indeed predict the structure which is actually found.

Figure 5a shows the free energy associated with dissolving one Gly-Ala-Gly zwitterion in solvents of pure water, 4 M NaCl, and 5 M NaCl solutions. The cross-section of peptide structures used is the available range of ψ for a fixed value of $\phi = -60°$. The method of creating peptide structures is described in the previous section. Where mathematical solutions are possible, an increase in the NaCl concentration increases the solvation chemical potential of all the structures of the peptide by an approximately uniform amount. Since essentially all structures are energetically less favorable, NaCl must destabilize the solvation of the peptide. Despite this destabilization the overall stability of the system is greater, because the chemical potential of solvation of the additional salt added is favorable and relatively large.

The sense in which the term stability is used to classify cosolvents, however, regards the change in free energy which occurs when a biological molecule *folds*, or when its volume to surface area ratio increases. The calculated results do not display this type of stability either, because the minimum energy point on each curve corresponds to an extended structure, although some folding is not allowed in our conformationally restricted model. The minimum energy point on the cross-section shown is also close to the global minimum for the entire space of (ϕ,ψ) angles (for pure water and for 5 M NaCl solution). Inspection of Figures 1c, 3b, and 4c shows that there is *no* cross-section where more compact structures are more energetically stable than extended ones.

Figure 5c shows $\rho_v G_{vu}$, the (negative) number of water molecules displaced by the peptide, for the same three solvent mixtures. Almost all the structures in the 5 M salt solution displace more water molecules than the corresponding structures in pure water. In each case the average number of molecules displaced is relative to the average number that would have taken up the space (in the relevant bulk solvent mixture) not only that the peptide occupies but also that the peptide reorders in the surrounding solvent. The counting of the average number of water molecules goes to the distance where the average solvent structure becomes indistinguishable from the bulk structure. The significant range of the integrand used to calculate this quantity is typically about 25 Å. The overall

(31) Ursenbach, C.; Patey, G. N. *J. Chem. Phys.* **1994**, *100*, 3827.
(32) Caccamo, C.; Malescio, G. *J. Chem. Phys.* **1989**, *90*, 1091.
(33) Prigogine, I.; Defay, R. *Chemical Thermodynamics*; translated by D. H. Everett; Longmans Green and Co.: London, 1954.
(34) Tisza, L. *Generalized Thermodynamics*; M.I.T. Press: Cambridge, 1966.
(35) von Hippel, P. H.; Schleich, T. In *Structure and Stability of Biological Macromolecules*; Timasheff, S. N., Fasman, G. D., Eds.; Marcel Dekker: New York, 1969.
(36) Robinson, D. R.; Jencks, W. P. *J. Am. Chem. Soc.* **1965**, *87*, 2470.

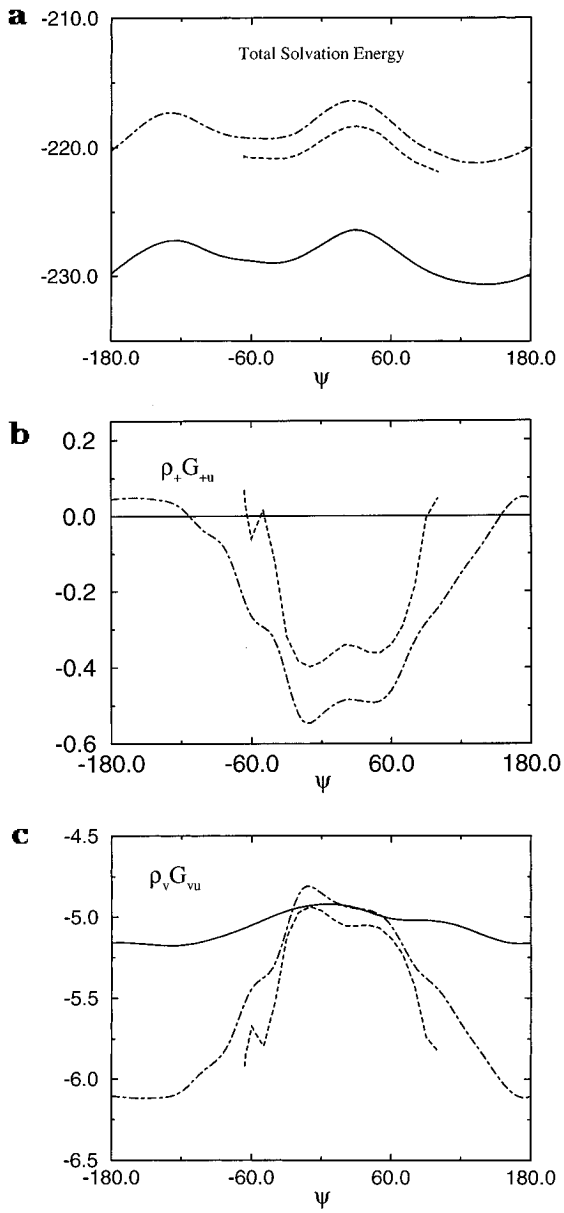

Figure 5. (a) The cross-section of the total solvation energy of Gly-Ala-Gly when ϕ is fixed at an angle of $-60°$. The solid line represents immersion in pure water, the dashed line represents a 4 M NaCl solvent solution, and the dot-dashed line represents a 5 M NaCl solvent solution. (b) The quantity of ρ_+G_{+u}, the (negative) number of cations *or anions* displaced by the peptide, for the same conformations and solvent mixtures as in (a). (c) The quantity ρ_vG_{vu} is the (negative) number of solvent molecules displaced by the peptide, for the same conformations and solvent mixtures as in (a).

decreases below the values for pure water are not primarily due to the decrease in the water density with increased salt concentration, because all solvent solutions have very similar ρ_vG_{vu} values for the most compact (least extended) structures. The more extended structures, those with larger dipoles, show a significant relative decrease in the number of water molecules nearby. The decrease in hydration is accompanied by a corresponding increase in the number of both cations and anions (Figure 5b). While Figure 5b gives the quantity ρ_+G_{+u}, it should be noted that global charge neutrality, rigorously enforced, requires that $\rho_-G_{-u} = \rho_+G_{+u}$. We see, therefore, a general increase in the ionic atmosphere around the peptide when its dipole increases in size. The increase in the ionic atmosphere is accompanied by a rearrangement of the solvent and cosolvent which on average, however, shows that energetic conformational stability is not proportional to the size of the dipole (compare Figures 5a, 1c, 3b, and 4b with Figure 2). It should be noted that these figures include the energy penalty for the separation of the charges at the ends of the peptide when its structure is extended. Figures 5a−c show how the screening and rearrangement in solvents with increased cosolvent concentration stabilize many energetically unfavorable peptide structures with large dipoles. The reorganization of pure water alone in response to an increased dipole does not yield as much free energy with which to offset the energetic cost of increasing the dipole, but correspondingly less energy is needed because the dielectric constant of pure water is so much higher. The net effect is that the free energy difference between two structures in either 5 M NaCl solution or pure water is surprisingly similar, evidenced by the almost constant separation between the solid and dot-dashed lines in Figure 5a.

It is therefore unlikely that the salting-in phase phenomenon discussed above is characterized by precipitation of the peptide. Not only is the total free energy of solvation less negative for higher concentrations of salt, but also the relative free energy between peptide structures is quite insensitive to the salt concentration, whereas the calculated phase behavior is not. The preferential hydration properties of this peptide (Figures 5b,c) are clearly opposite to the preferential hydration more often seen for proteins in salt solution.[6−8] It is also inconsistent with the classification of NaCl as a stabilizing and salting-out agent[35,36] which is expected to be preferentially excluded from the surface of solutes. It is, however, consistent with the findings above, where Gly-Ala-Gly is shown to be both salted into solution and destabilized by the addition of salt *for a given particular set of rigid structures*. Since the peptide minimum energy structures are essentially unchanged between 5 M NaCl solution and pure water solvents (Figures 1c and 3b) it is unlikely that the addition of high concentration salt induces structural change in the peptide. The ensemble average structure of the peptide is most likely therefore to be neither stabilized nor destabilized, and hence there is no tendency to fold or unfold due to the change of solvent from pure water to 5 M NaCl solution as might be expected for this small peptide.

c. Comparison to the Case of *β*-Lactoglobulin. For those peptide structures which are thermodynamically stable in a single liquid phase in 4 M salt solution, when approached from higher salt concentration, the number of water molecules displaced decreases, despite the increase in water density (Figure 5c). The change in the number of displaced water molecules, ρ_vG_{vu}, does not follow the trend one would expect, falling outside the range outlined by the pure water and 5 M NaCl solution results everywhere, instead of between them, as a consequence of the phase stability of the solution. Figure 5b demonstrates the corresponding increase in ion numbers in the region around the solute, which also unexpectedly lies at a higher value everywhere than for the higher concentration (5 M) result. The average number of ions found close to the peptide has increased so much near the end points of the available range of ψ for in the 4 M case that they have offset the number of ions displaced by the peptide's excluded volume and are positive and increasing rapidly. This demonstrates the intrusion of many ions into the surface region and is consistent with the specific association of the ions and solute. It should be noted that the increase in cosolvent association, or decrease in specific hydration, is much more than the law of mass action alone would predict for the change from 4 to 5 M NaCl solution.

It is clear that an increased concentration of cosolvent is formed in the region surrounding our example zwitterion as the concentration of the cosolvent is decreased, while simultaneously

the solubility of the zwitterion decreases. These same phenomena were discovered experimentally for β-lactoglobulin in NaCl/water solution while near neutral pH. Neutral pH in the β-LG system corresponds to the GLy-Ala-Gly system with the molecule extended, since this is when both biological molecules have their dipoles maximized. The dependence of the unusual solubility behavior of β-LG on the size of its dipole was first suggested by Cohn and Ferry,[19] and appears to apply to the Gly-Ala-Gly zwitterion as well.

We now turn to the molecular level interpretations of these experimental results. We wish to determine whether or not the increase in ion concentration around our peptide is due to associations which can be truly thought of as specific, or localized, with respect to the peptide. We also wish to determine whether the simple substitution model of Schellman[4] can be reasonably applied to this particular case.

The approach to the solubility boundary of Gly-Ala-Gly can be examined from several points of view, so we attempt to use quantities which are as intuitive as possible. We choose to look at the change in cosolvent ion distribution as the structure of the peptide is extended rather than as the NaCl concentration is decreased. The picture is therefore simplified by the lack of mass action considerations and is similar in nature to the approach to neutral pH for the β-LG case, where the dipole also enlarges. We will consider part of the same cross-section of the domain of (ϕ,ψ) angles as given in Figures 4a and 4b for the 4 M NaCl case. For a fixed value of $\phi = -60°$, the minimum value of the angle ψ for which a solution could be found was $-66.6°$, when ϕ was lowered by 0.1° increments. We denote the probability, relative to the bulk, of finding an ion i at a distance r away from atom j on the peptide as $P_{ji}(r,\phi,\psi) = 4\pi r^2 g_{ji,\phi,\psi}(r)\,dr$ for a given fixed (ϕ,ψ) pair. Figure 6a gives the difference quantities

$$\Delta_1(r) = \frac{\rho_-[P_{N-}(r,-60°,-66°) - P_{N-}(r,-60°,-60°)]}{dr} \quad (7a)$$

and

$$\Delta_2(r) = \frac{\rho_-[P_{N-}[r,-60°,-66.6°) - P_{N-}(r,-60°,-60°)]}{dr} \quad (7b)$$

where N denotes the nitrogen atom in the NH_3^+ group at the positive end of the zwitterion, and a subscript minus denotes a Cl^- ion. The integral of these functions can be interpreted as the change in the average number of chloride ions around the nitrogen atom as the peptide is extended in a 4 M NaCl solution.

Figure 6a shows the only relatively large contact peak in these number difference quantities for cations or anions around any site on the peptide. The peak represents an increase in the average number of Cl^- ions in contact with the three hydrogen atoms which, in this model, have equal partial charges, and are bonded with equal equilibrium distances from the N-terminus nitrogen. The equivalent functions for Cl^- ions around each of the three hydrogen atoms (not shown) have their contact peak radially split into two smaller peaks representing anions in contact with themselves and, at a slightly larger separation, with the other hydrogen atoms. It is particularly informative to examine these same functions with the negatively charged end of the peptide as the chosen origin. The change in anion distribution around the partially negatively charged oxygen atoms at the C-terminus (Figure 6b) shows a significant peak at a separation of about 10 Å. The reason for this peak is the same as for the contact peak in Figure 6a, namely the increased association of Cl^- ions near the N-terminus. This identification

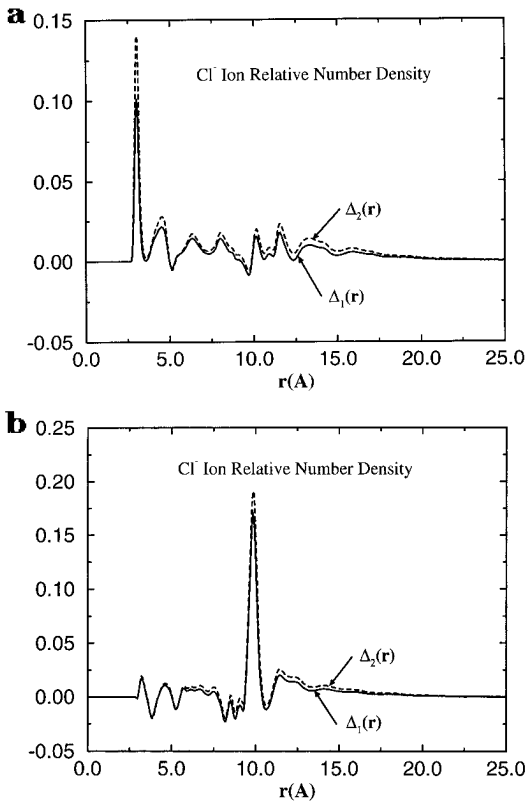

Figure 6. (a) The difference in the probability of finding a Cl^- ion at a distance r away from the nitrogen atom at the N-terminus of Gly-Ala-Gly in a 4 M NaCl solution, relative to the case when the intramolecular dihedral angle ψ is $-60°$. The solid line represents the case when the reference structure probability ($\psi = -60°$) is subtracted from the probability for the structure with $\psi = -66°$. The dashed line represents the case when the reference structure probability is subtracted from the probability for the structure with $\psi = -66.6°$. (b) The difference in the probability of finding a Cl^- ion at a distance r away from an oxygen atom at the C-terminus of Gly-Ala-Gly in a 4 M NaCl solution. The lines represent probabilities relative to the same structures as in (a).

can be made because the sum of the distance from the N-terminus nitrogen atom to the Cl^- contact peak and the distance from this nitrogen atom to the oxygen atom in question at the C-terminus is very close to 10 Å for these particular peptide structures. Since there are no other significant peaks in Figure 6a or 6b, we can conclude that there are no other significant increases in anion number at any other localized association positions as the solubility boundary is approached.

Figures 7a and 7b show the functions corresponding to Figures 6a and 6b for cation distribution around the same N-terminus nitrogen and C-terminus oxygen atoms. The relatively broad peak near 5 Å separation in Figure 7a is consistent with an increase in Na^+ ions further out in solution induced by the increase in contact Cl^- ions; there is no contact peak corresponding to any other atom center on the peptide molecule. In the case shown in Figure 7b there is no significant single peak, but rather a wide region between contact and a separation distance of about 8 Å where there is a significant difference in the average number of Na^+. This is probably due to increased average numbers of Na^+ ions in the general region of the C-terminus, without any one specific local association position. Once again there is a broad peak at a separation of 11-12 Å, consistent with an increase in average Na^+ numbers just beyond the opposite end of the zwitterion. Neither Figure 7a nor 7b shows any evidence of an important association location for Na^+ ions somewhere else on the peptide molecule.

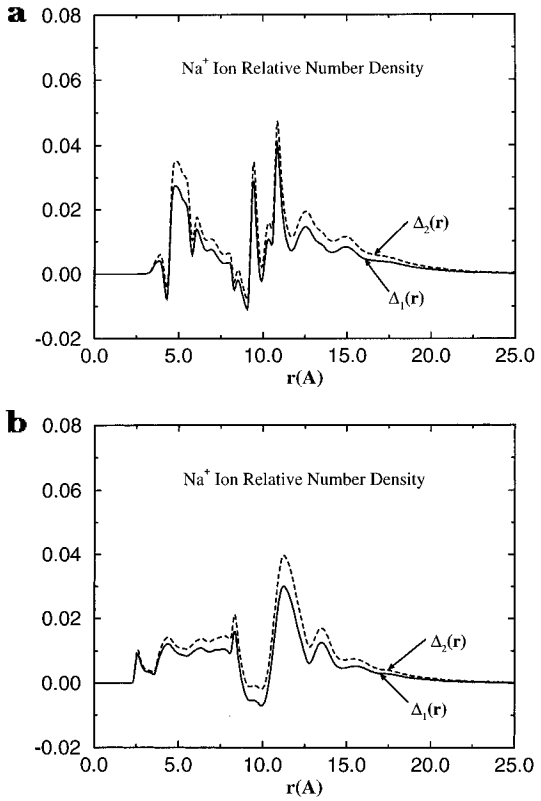

Figure 7. (a) The difference in the probability of finding a Na^+ ion at a distance *r* away from the nitrogen atom at the N-terminus of Gly-Ala-Gly in a 4 M NaCl solution. The lines represent probabilities relative to the same structures as in Figure 6a. The difference in the probability of finding a Na^+ ion at a distance *r* away from an oxygen atom at the C-terminus of Gly-Ala-Gly in a 4 M NaCl solution. The lines represent probabilities relative to the same structures as in Figure 6a.

There are three important additional features demonstrated by all the figures discussed in this section. The first is that, as the solution stability limit is approached, the change in the average numbers of both anions and cations around the peptide becomes increasingly more sensitive to smaller and smaller changes in the structure of the peptide. The difference between the dashed and solid lines is due to a change in the angle ψ of only 0.6°, whereas the change represented by a solid line is due to an change in ψ an order of magnitude larger. This sensitivity to small changes in the model is consistent with the behavior seen in the approach to a phase boundary in other systems.[17,18] The second is that no regular long-range structure characteristic of the formation of a salt crystal lattice is forming. Such structure has been seen in similar calculations when the phase behavior concerned is one of salt precipitation.[31] We note that such a structure, if present, would indicate that the peptide was nucleating a salt precipitation reaction, not that a peptide precipitation was occurring. Our model calculations are all performed with the peptide rigorously held at infinite dilution and so no such peptide precipitation could be seen. Nevertheless, *some* phase phenomenon depending strongly on the peptide structure is occurring. The third is that there is a significant increase in the average number of both cations and anions beyond a separation of 10 Å, which is approximately the largest separation distance at which a contact structure is possible. Perhaps 50% or more of the total of the locally associated ions (the integrals of these functions) are due to increases in ion concentration which are in the region of the peptide, but may be as far as 25 Å (≈8 solvent diameters) away from contact. This is not the typical kind of behavior which one would interpret as specific association. Yet, the thermodynamic quantities which are often used to measure the degree of association[8,11] cannot distinguish between an ion which moves into a contact position next to the peptide and one which simply moves closer to the peptide. Furthermore, those contact peaks which are present represent a relatively small contribution to the integral of the functions, and thus those increases in specific association which are present will not be, in this case, the most significant contributors to the related thermodynamic quantities. An increase in ionic strength in the region of the peptide, when such an increase is not primarily due to the increased association of ions with the surface, is consistent with the separation of the solution into two liquid phases. In this picture, the phase which contains the peptide would have a higher NaCl concentration than the other. The separation of an ionic solution into two non-miscible liquid phases has been seen before both experimentally[37−39] and theoretically.[15,30−32]

Schellman's simple exchange model[4] has been shown to be of great utility in the examination of results for the binding of solvent and cosolvent to molecules of biological interest.[8] While the simple intuitive picture which is presented in the context of the mathematical model is one of single solvent molecules being completely replaced by single cosolvent species, there is nothing explicit in the formulation which necessitates such exchange happening directly at the solute surface. For instance, if the model were one where cosolvent species could replace solvent molecules in the second solvation shell, say, the mathematical terms and derivation would be essentially unchanged. Thermodynamic quantities that indicate such exchange, when they are derived from experiments and are consistent with the exchange model, could represent, therefore, many such types of solvent/cosolvent exchange. Furthermore, there is no way to determine from thermodynamic quantities alone if the molecular association thus discovered represents interactions at the surface or some distance away.

In summary, there is an overall tendency of cosolvent to migrate into the vicinity of the peptide which concurs with a decreased solubility of the peptide in a single liquid phase when the dipole is extended. This is in agreement with the experiments on β-LG when its dipole is increased by the approach to neutral pH. However, the action of the cosolvent is not just that of specific association at the surface of the peptide. In addition there is a general increase in the average cosolvent concentration in a large region around the peptide, with significant contributions extending to at least 8 solvent diameters from the surface. Furthermore, the restructuring of the solvent and cosolvent around the peltide appears to be dominated by the general ion migration into the region rather than the specific association which is present. Such a widespread restructuring is likely to contribute significantly to any thermodynamic measure of ion association. Since there appears to be more than one possible molecular level explanation consistent with the experimental data, the distinction between these distinct phenomena cannot be made from the interpretation of purely thermodynamic quantities. In particular, Schellman's model[4] of ion exchange will not distinguish between ions which migrate to a specific location on a biological molecule's surface and those which simply replace water molecules in the surrounding region, perhaps at a significant distance of separation.

When one considers the changes in ionic distribution around the peptide studied here, and the behavior consistent with a separation into two non-miscible liquid phases, it is natural to

(37) Schettler, P. D., Jr.; Patterson, A., Jr. *J. Phys. Chem.* **1964**, *68*, 2870.
(38) Doumaux, P. W.; Patterson, A., Jr. *J. Phys. Chem.* **1967**, *71*, 3535.
(39) Doumaux, P. W.; Patterson, A., Jr. *J. Phys. Chem.* **1967**, *71*, 3540.

|wonder if this liquid−liquid separation is also occurring in the case of β-LG near neutral pH, especially when there appear to be many similarities between these two systems. Turbid solutions have been observed for β-LG with NaCl as cosolvent[5] which lend support to this conjecture.

(4) Summary and Conclusions

The DRISM/HNC statistical mechanical theory were applied to a model system consisting of a conformationally restricted zwitterion tripeptide of sequence Gly-Ala-Gly at infinite dilution in a solvent/cosolvent mixture of 5 M NaCl solution and in pure water. Solvent−solute distribution functions and solvation thermodynamic quantities were examined for peptide conformations covering the entire available range of the central (ϕ,ψ) dihedral angles. Relative free energies of each conformation in the mixed solvents were shown decomposed into the intrmolecular energy of the peptide itself and the energy of restructuring of the neighboring solvent. Each of these contributions to the total relative free energy were found to be closely related to the size of the dipole, but approximately cancelled each other when added leaving a free energy surface with four local minima. Each of these minima represented open, relatively extended structures. Between the minima were energy barriers which grouped them in two pairs with transition within each pair likely, but between pairs unlikely, in a room temperature solvent solution. The energy minima for the pure water and the 5 M NaCl cases corresponded to very similar structures demonstrating that the minimum energy conformations were insensitive to the presence of cosolvent. However, for concentrations between 3 and 5 M a solution to the DRISM theory for many peptide conformations, particularly the more extended ones with larger molecular dipoles, was unavailable. Below 3 M no solutions for the theory for this model system seemed possible. The loss of solution for this class of liquid state theories has been seen before in many contexts,[15,30−32] and is usually associated with a phase boundary in the composition range of this study. This phase behavior was interpreted as a salting-in by NaCl, which is unusual since NaCl belongs to the class of cosolvents which often salt many peptide molecules out of solution.

The degree of ionic association with the peptide was investigated by examining the quantity $\rho_v G_{vu}$, the average number of molecules of solvent or cosolvent displaced by the peptide relative to their bulk average distribution. For most conformations in the 5 M NaCl solution there was a significant increase in ionic concentration near the peptide, which, in conjunction with extended structures being most stable, is consistent with the trends seen for a variety of biological in NaCl solutions experimentally.[35,36] For solutions with NaCl concentrations below 5 M, however, the number of ions in the vicinty of the peptide increased dramatically, especially near the phase boundary. For these lower salt concentrations the consistency with the usual cosolvent classification of NaCl was lost.

When the distribution of ions around the peptide was examined, the only specific association found was that of Cl^- ions near the (positively charged) N-terminus. This association, however, accounted for only a small fraction of the redistribution of ions with respect to bulk structure around the peptide. Significantly increased average ion numbers were found all around the peptide as the conformation was changed in such a way as to increase the molecular dipole toward the size where the phase boundary was reached. This increased ionic atmosphere around the peptide was found to extend up to eight solvent diameters into the surrounding solvent mixture. The calculation of such a large restructured region around the peptide has an important bearing on the interpretation of thermodynamic association quantities derived from experimental observations. In particular in the context of Schellman's solvent/cosolvent exchange model, which has proven extremely useful in the interpretation of experiments, it is impossible to tell whether the migration of cosolvent species toward a biological solute is primarily a surface interaction or an exchange with a solvent molecule somewhere in the surrounding region, perhaps in the second or a more distant solvation shell. Only the former fits the intuitive picture of specific association arguments. Since no regular structure appeared in the ionic distributions near the peptide, which would have indicated the nucleation of a salt crystal, the phase change was interpreted as a separation into two non-miscible liquid phases, one with a higher salt concentration than the other. Such liquid phase separations for ionic solutions have also been seen both in the theoretical[15,30−32] and the experimental[37−39] literature for systems with dielectric behavior similar to our model system.

When the experimental results for the behavior of β-lactoglobulin in NaCl solution were examined there were several solubility and cosolvent behavior properties common to the peptide zwitterion case. In particular, β-LG is salted in by NaCl[5] and appears to be related to the molecular dipole.[19] The solubility behavior of β-LG in a solvent mixture where the pH is manipulated, and consequently the molecular dipole is changed, paralleled the behavior or the Gly-Ala-Gly model as its dipole is changed. The degree of cosolvent intrusion into the region around each molecule near the solubility limit was also similar in the two systems. Since β-LG forms turbid solutions, which are often characteristic of mixtures of two nonmiscible liquid phases, when the molecule is close to its maximum dipole, it was suggested that both our model system and the experimental β-LG system could have this property in common as well.

Acknowledgment. The authors thank the Robert A. Welch Foundation and the N.I.H. for partial support of this work. The N.S.F. and the Keck Center for Computational Biology are thanked for local computational support.

JA952392T