# Continuum Solvation Model for Studying Protein Hydration Thermodynamics at High Temperatures

**Adrian H. Elcock* and J. Andrew McCammon**

*Department of Chemistry and Biochemistry, Department of Pharmacology, University of California at San Diego, La Jolla, California 92093-0365*

*Received: June 10, 1997; In Final Form: August 12, 1997*[⊗]

A macroscopic solvation model that combines a solvent-accessible surface area term to describe hydration of nonpolar groups with a continuum electrostatics term to describe the hydration of polar groups has previously been shown to provide an excellent description of amino acid hydration free energies at 25 °C (Sitkoff et al. *J. Phys. Chem.* **1994**, *98*, pp 1978−1988). This paper describes the extension of this method and its accompanying parameter set (known as PARSE) to handle temperatures in the range from 5 to 100 °C. For the neutral amino acids, hydration free energies were taken from the literature; for the charged amino acids Asp, Glu, and Lys, hydration free energies were obtained by combining results for the neutral analogues with information on the p$K_a$ value at the required temperature. An important result of this analysis is that the hydration free energies of the charged residues are much more strongly affected by increasing temperature than their neutral analogues. In extending the PARSE method to reproduce hydration free energies over a range of temperatures, a number of alternative models were investigated: best results were obtained when separate surface area dependent terms were used to represent the hydration of aliphatic and aromatic regions and when the continuum electrostatics term was made strongly temperature dependent. The temperature dependence of the electrostatic component stems partly from changes in the dielectric constant of water but appears to be rather better described when the atomic radii are also made temperature dependent, increasing in size as the temperature rises. This requirement for temperature-dependent radii gains important support from previous studies using the Born model to describe the entropies of hydration of simple ions. The extension of the PARSE method described here permits its use in investigating the effects of hydration on protein stability over a wide range of biologically relevant temperatures.

## Introduction

Hydration effects are one of the key determinants of protein stability.[1−3] Although the hydrophobic hydration associated with nonpolar groups is commonly viewed as the most important contributor to stability, the changes in the hydration of polar and charged groups that accompany protein folding are also thought to play significant roles. Hydrophobic contributions are typically expressed empirically as being proportional to the solvent-accessible surface area (SASA) buried on folding of the protein,[4] an approach that is inspired by the linear relationship[5] between the hydration free energies of linear alkanes and their SASAs. Similar surface area-based approaches have also been applied to investigate the hydration contribution of polar groups to protein stability.[6−8] These may be expected to work reasonably well when a polar group becomes *completely* buried on folding of a protein, since transfer from solution to the protein environment is in some respects similar to transfer to the gas phase, from which process the SASA proportionality constants have been derived. However, it is less likely to work well for the intermediate cases when a polar group becomes only partially buried or remains close to the protein surface, for the simple reason that a halving of the surface area of a polar group on folding need not result in a halving of its hydration energy.[9]

Alternative methods for describing the hydration of polar molecules exist, the most prominent of which are based on continuum electrostatics treatments.[10,11] Continuum methods make use of the fact that hydration of polar and charged groups is an essentially electrostatic phenomenon so that the gas-to-water transfer process can be adequately described simply as a transfer from a medium of relative dielectric 1 to one of 78. The energetics of this transfer process are given by the difference between the electrostatic energies calculated in the two phases, determined for each by solving the Poisson equation (or the Poisson−Boltzmann equation if the aqueous phase contains dissolved ions). Of course, treating water as a dielectric continuum represents a significant approximation, and hydration effects which are thought to reflect aspects of water structure, such as the hydrophobic effect, are clearly beyond such methods. Nevertheless, studies have indicated that for many polar molecules the approach provides hydration free energies in excellent agreement both with experiment and with the results of much more sophisticated (and computationally demanding) free energy simulation methods.[12−15]

Recent work has described the use of a combination of a continuum electrostatics approach for polar hydration and a SASA method for nonpolar hydration to reproduce the hydration free energies of a wide range of small molecules.[11] Of particular relevance for the present work was the method's success, when suitably parametrized, in reproducing the hydration free energies of the amino acid side chains with an average error of only 0.1 kcal/mol. This method, and its accompanying parameter set, known as PARSE, opens the way for investigating a number of aspects relating to protein hydration, assuming, that is, that parameters suitable for isolated amino acids are also appropriate for proteins. The method has been applied with success to investigations of electrostatic effects at the termini of α-helixes[16] and to studies of the factors contributing to α-helix[17] and β-sheet stability.[18]

The purpose of this paper is to describe the extension of the PARSE parameter set to temperatures other than 25 °C,

[⊗] Abstract published in *Advance ACS Abstracts,* October 1, 1997.

S1089-5647(97)01903-2 CCC: $14.00 © 1997 American Chemical Society

specifically, the range 5−100 °C. To properly parametrize the method for other temperatures, we use as our basis the experimental thermodynamic data derived by Privalov and Makhatadze (referred to in the text hereafter as P&M) to describe the hydration free energies of the *neutral* amino acid side chains in the same temperature range.[6,7] These authors used the experimental hydration enthalpies ($\Delta H_{hyd}$) and free energies ($\Delta G_{hyd}$) of a wide range of small molecules at 25 °C to derive a set of parameters that describes the hydration contributions of each of the various functional groups present in the amino acid side chains. Importantly, in the purely empirical approach of P&M, the total hydration free energy of a given residue is assumed to be equal to the sum of the contributions from each of its component functional groups. For example, the hydration of the glutamine side chain is described as a straightforward additive sum of contributions from $-CH_{2}-$, $-C{=}O$, and $-NH_2$ groups. Using these group contributions to obtain values of $\Delta G_{hyd}$ and $\Delta H_{hyd}$ for each of the neutral amino acid side chains at 25 °C, the hydration entropies $\Delta S_{hyd}$ were calculated using the relation $\Delta G = \Delta H - T\Delta S$.

In calculating the temperature dependence of each of the terms $\Delta G_{hyd}$, $\Delta H_{hyd}$, and $\Delta S_{hyd}$, P&M made use of hydration heat capacities $\Delta C_{p,hyd}$ that they had previously either measured directly or derived.[19] Standard thermodynamic relationships were then used to obtain $\Delta H_{hyd}$ and $\Delta S_{hyd}$ at each of the temperatures 5, 50, 75, and 100 °C (and additionally 125 °C), from which $\Delta G_{hyd}$ follows straightforwardly:

$$\Delta H_{hyd}(T) = \Delta H_{hyd}(25\ ^\circ\mathrm{C}) + \int_{25\ ^\circ\mathrm{C}}^{T} \Delta C_{p,hyd}(T)\, \mathrm{d}T$$

$$\Delta S_{hyd}(T) = \Delta S_{hyd}(25\ ^\circ\mathrm{C}) + \int_{25\ ^\circ\mathrm{C}}^{T} \frac{\Delta C_{p,hyd}(T)}{T}\, \mathrm{d}T$$

The results of P&M's analysis are confined to the neutral amino acids. However, we show here that by combining their results with experimental information on p$K_a$ values, it is possible to derive reasonable estimates of the temperature-dependent hydration free energies of the *charged* amino acid side chains. Given this set of $\Delta G_{hyd}$ values for temperatures from 5 to 100 °C for each of the amino acids, we then set about modifying the existing PARSE parameter set to reproduce these results, while keeping alterations to a minimum. This paper describes this parametrization process in detail and gives parameters suitable for describing amino acid hydration free energies in the range 5−100 °C.

## Methods

**(a) Evaluation of P&M's Results.** Figure 1 shows the correspondence between P&M's derived values for $\Delta G_{hyd}$ and those obtained experimentally by Wolfenden et al.[20] at 25 °C. On the whole, the agreement is good, a result that lends support to the overall approach adopted by P&M. However, in certain cases agreement is quite poor: especially noticeable are the residues Arg and Trp, for which the errors amount to 5.5 and 4.1 kcal/mol, respectively. Because of these discrepancies, and for additional reasons outlined below, we omitted Arg and Trp (and also Met, see below) from the set of molecules used to derive parameters.

In the particular case of Arg, the overestimation of $\Delta G_{hyd}$ obtained with P&M's parameters is perhaps not very surprising. Molecules such as Arg, which have a number of functional groups in close proximity, are expected to cause problems for simple additive hydration models of the type adopted by P&M, particularly when, as in the present case, a similar degree of close grouping is not present in any of the compounds used to derive the parameters. In this connection, it is worth pointing out that in the related group contribution method followed by Cabani et al.,[21] it was found that correction factors as high as 17 kJ/mol (∼4 kcal/mol) were required for closely spaced functional groups. A similar concern applies also to the $\Delta H_{hyd}$ value for Arg,[6] which it appears might also be substantially overestimated, although owing to the lack of experimental data it is not possible to decide on the magnitude of any error that may be present. It is important to have confidence in the values of *both* $\Delta G_{hyd}$ and $\Delta H_{hyd}$ for a particular residue, since they are used together to determine $\Delta S_{hyd}$ through use of the relation $T\Delta S_{hyd} = \Delta H_{hyd} - \Delta G_{hyd}$. Any errors in $\Delta S_{hyd}$ will naturally lead to errors in the temperature dependence of $\Delta G_{hyd}$. The degree of uncertainty in the value of $\Delta H_{hyd}$ and the clear error in the value of $\Delta G_{hyd}$ are by themselves sufficient reason to exclude Arg from the parametrization, but an additional reason is that there are no data for $\Delta C_{p,hyd}$ of the neutral residue (data are reported only for the charged analogue; ref 19).

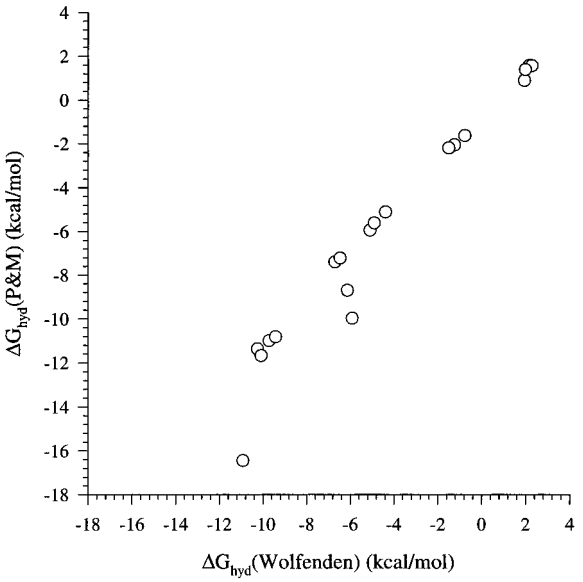


**Figure 1.** Correspondence between hydration free energies of the neutral amino acid residues calculated by Privalov and Makhatadze (ref 7) and obtained experimentally by Wolfenden et al. (ref 20).

For Trp, the difference between the estimates of P&M and Wolfenden et al. may result in part from the fact that a single $\Delta G_{hyd}$ contribution is assumed by P&M for all $-NH-$ groups, irrespective of whether they are in aliphatic or aromatic environments. The work of Cabani et al.,[21] while not explicitly reporting different values for $-NH-$ groups, does indicate that the contribution of groups such as $-N{=}$ and $-OH$ are substantially more favorable in aliphatic (−24.79 and −25.95 kJ/mol) than aromatic (−17.02 and −19.35 kJ/mol, respectively) environments. P&M's use of a single value for the contribution of $-NH-$ groups, derived using aliphatic model compounds, might therefore overestimate the contribution of the $-NH-$ group in Trp. Again, this argument carries over to the values of $\Delta H_{hyd}$ and is further compounded by the fact that, owing to an accumulation of errors in their method, P&M themselves had to set (quite arbitrarily) the $\Delta S_{hyd}$ value for Trp to be identical with that of His (see the footnote to Table 4 in ref 7). The temperature dependencies that they derived for Trp then are not without considerable uncertainty.

Our reasons for omitting Met from the set of residues used for parametrization are not due to concerns over the $\Delta G_{hyd}$ value obtained by P&M: in fact, their value (−2.20 kcal/mol) is in reasonable agreement with the experimental value of −1.49 kcal/

mol obtained by Wolfenden et al.[20] Instead, the omission was based on the fact that during the course of our parametrization work, the errors were consistently greatest for Met (along with Arg and Trp). This seems to result from the fact that the temperature dependence of $\Delta G_{hyd}$ for Met is unusually large: according to P&M's results, of all the amino acid residues only Arg shows a stronger temperature dependence than Met. It is difficult to establish precise reasons for why this might be the case: for example, the $\Delta H_{hyd}$ and $\Delta S_{hyd}$ values obtained by P&M at 25 °C do not appear unreasonable when compared with other residues. However, the $\Delta H_{hyd}$ and $\Delta S_{hyd}$ contributions from the *polar* part of Met do show a far greater temperature dependence than any other residue (see especially Table 4 of ref 7): this appears to be a consequence of its compensating for the opposing temperature dependence of the nonpolar part. Aside from noting this, we have no compelling reason for omitting Met, except for the fact that as shown under the Results, the experimental results cannot be adequately reproduced by models which otherwise accurately describe all of the amino acid residues.

**(b) Parametrization of Neutral Residues.** With the exception of Arg, Trp, and Met then, we took the temperature dependencies of the hydration free energies of the neutral side chains directly from the work of P&M.[7] To modify the PARSE parameter set to reproduce these results, we tried a number of different approaches discussed more fully below. While differing in details, each approach consisted of some combination or other of a SASA-dependent term and an electrostatic contribution, similar to the expression shown below, which was used in the original PARSE parametrization:[11]

$$\Delta G_{hyd} = \gamma \times \mathrm{SASA} + \Delta G_{elec} + 0.92\ \mathrm{kcal/mol}$$

Here $\gamma$ represents a proportionality constant (for which PARSE adopts the value 5.4 cal/mol·Å$^2$), while $\Delta G_{elec}$ denotes the change in electrostatic free energy for transfer of the molecule from the gas phase to the aqueous phase. When one is considering the temperature dependence of the hydration free energy, it is worth remembering that the above expression assumes that there is no contribution from conformational entropy; i.e., we assume that there is effectively no difference between the solute's conformational freedom in the gas phase and the aqueous phase. This is likely to be a good approximation for the solutes investigated here, but would be less appropriate, for example, where a hydrogen bond is formed in the gas phase which is lost on transfer to the aqueous phase: in such cases conformational entropy could make a significant contribution to the temperature dependence of the hydration free energy.

Since we are interested here only in modifying the original PARSE parameter set, rather than in deriving a completely new set, we concentrated on reproducing the *differences* between the hydration free energies obtained at one particular temperature, *T*, and those obtained at 25 °C. In other words, we use the work of P&M *only for the temperature dependencies* of the hydration free energies and retain PARSE in an unaltered form for the temperature 25 °C.

The simplest model that one can adopt is to assume that the change in $\Delta G_{hyd}$ that results from changing the temperature ($\Delta\Delta G_{hyd}$) is simply proportional to the total SASA of the side chain, i.e., that the change can be expressed solely in terms of the PARSE nonpolar contribution:

$$\Delta\Delta G_{hyd}(25\ ^\circ\mathrm{C}\rightarrow\mathrm{T}) = \Delta\gamma^*\mathrm{SASA} \qquad \text{(model 1)}$$

Notice that the proportionality constant is here termed $\Delta\gamma$ to indicate that it is the *change* in $\gamma$ from 25° to *T*. With the above model, we used a least-squares regression procedure implemented in SigmaPlot[22] to find a single optimal value of $\Delta\gamma$ that would simultaneously describe the temperature dependencies of the hydration free energies of *all* the neutral side chains. As described under Results, and as would be expected from the detailed analysis of P&M, such a simple model performs very poorly, so the next approach adopted made use of an expression for $\Delta\Delta G_{hyd}$ that contains an additional contribution due to a temperature dependence of $\Delta G_{elec}$:

$$\Delta\Delta G_{hyd}(25\ ^\circ\mathrm{C}\rightarrow\mathrm{T}) = \Delta\gamma^*\mathrm{SASA} + \Delta\Delta G_{elec}(25\ ^\circ\mathrm{C}\rightarrow\mathrm{T}) \qquad \text{(model 2)}$$

The calculation of the $\Delta\Delta G_{elec}$ term is discussed in more detail below. As above, a regression procedure was used to find the single value of $\Delta\gamma$ most capable of reproducing the $\Delta\Delta G_{hyd}$ values of all the amino acids simultaneously.

An alternative to including a $\Delta\Delta G_{elec}$ term, while maintaining flexibility in the fitting, is to follow the approach adopted by P&M of using separate $\gamma$ values for each of the aliphatic, aromatic, and polar regions. Of course, PARSE uses only a single $\gamma$ value applied to the entire SASA:[11] we can use P&M's approach and still maintain consistency with PARSE by simply requiring that at 25 °C $\gamma_{aliphatic} = \gamma_{aromatic} = \gamma_{polar}$.

$$\Delta\Delta G_{hyd}(25\ ^\circ\mathrm{C}\rightarrow\mathrm{T}) = \Delta\gamma_{aliphatic}^*\mathrm{SASA}_{aliphatic} + \Delta\gamma_{aromatic}^*\mathrm{SASA}_{aromatic} + \Delta\gamma_{polar}^*\mathrm{SASA}_{polar} \qquad \text{(model 3)}$$

With the above model, each $\Delta\gamma$ was allowed to vary freely in the fitting procedure. This approach is essentially the same as that used by P&M, with the exception that it attempts to find a single value of $\Delta\gamma_{polar}$ appropriate for all residues. Finally, we considered a combination of three separate $\gamma$ values coupled with a $\Delta\Delta G_{elec}$ component:

$$\Delta\Delta G_{hyd}(25\ ^\circ\mathrm{C}\rightarrow\mathrm{T}) = \Delta\gamma_{aliphatic}^*\mathrm{SASA}_{aliphatic} + \Delta\gamma_{aromatic}^*\mathrm{SASA}_{aromatic} + \Delta\gamma_{polar}^*\mathrm{SASA}_{polar} + \Delta\Delta G_{elec}(25\ ^\circ\mathrm{C}\rightarrow\mathrm{T}) \qquad \text{(model 4)}$$

SASA values for each of the amino acid side chains were calculated from models constructed in the program QUANTA,[23] using atomic radii from the PARSE parameter set and a solvent probe radius of 1.4 Å. For the purposes of subdividing the SASA into separate contributions from polar, aliphatic, and aromatic groups, the only atoms considered to be polar were the following: heteroatoms and the H's attached to heteroatoms, plus the carbonyl carbon of the acid and amide groups, since this atom carries a substantial partial charge of +0.55 e. For the particular case of histidine, only the =CH− groups of the imidazole ring were treated as aromatic. The calculated SASA values are reported in Table 1.

The electrostatic component $\Delta\Delta G_{elec}$ was obtained using the program UHBD,[24] by carrying out separate calculations in the gas and aqueous phases and subtracting the calculated electrostatic energies. For each phase, calculations were performed in two stages using the technique of 'focusing'. First, the Poisson equation was solved on a $40^3$ grid of spacing 1.2 Å with boundary potentials set by applying Coulomb's law to each atom. Second, the results of this calculation were used to determine boundary potentials for a much finer $80^3$ grid of spacing 0.3 Å. In all calculations the internal relative dielectric of the amino acid side chain was set to 2.0 (as in the original PARSE parametrization). For the gas-phase calculations, the external (solvent) dielectric was set to 1.0, while for the aqueous phase it was set to the value appropriate for water at the given

**TABLE 1: SASA Values for the Amino Acid Residues**

| residue | $SASA_{aliphatic}$ | $SASA_{aromatic}$ | $SASA_{polar}$ | $SASA_{total}$ |
|---|---|---|---|---|
| Ala | 146 | 0 | 0 | 146 |
| Arg | 175 | 0 | 109 | 284 |
| Asn | 90 | 0 | 107 | 197 |
| Asp | 95 | 0 | 98 | 193 |
| Cys | 97 | 0 | 83 | 180 |
| Gln | 127 | 0 | 101 | 228 |
| Glu | 130 | 0 | 94 | 224 |
| His | 83 | 119 | 32 | 234 |
| Ile | 237 | 0 | 0 | 237 |
| Leu | 235 | 0 | 0 | 235 |
| Lys | 200 | 0 | 55 | 255 |
| Met | 206 | 0 | 37 | 243 |
| Phe | 85 | 176 | 0 | 261 |
| Ser | 112 | 0 | 46 | 158 |
| Thr | 148 | 0 | 42 | 190 |
| Trp | 81 | 192 | 28 | 301 |
| Tyr | 85 | 139 | 50 | 274 |
| Val | 210 | 0 | 0 | 210 |
| Nma[a] | 183 | 0 | 53 | 236 |

[a] Nma, *N*-methylacetamide.

temperature. Values for the latter, which decrease from 80.20 at 0 °C to 55.12 at 100 °C, were taken from the *CRC Handbook*.[25]

It was assumed initially that simply accounting for the temperature dependence of the dielectric constant of water would be sufficient to provide a good description of the temperature dependence of $\Delta G_{elec}$. Our calculations (discussed under Results) suggested, however, that this approach results in a considerable underestimation of $\Delta G_{elec}$'s temperature dependence. We found that better results were obtained if not only the solvent dielectric but also the atomic radii were made temperature dependent. The latter was implemented by uniformly applying a radius scaling factor (RSF) to every atom of the molecule. However, since we might intuitively expect that the need for a scaling of atomic radii would reflect temperature-dependent changes in the nature of the solute−solvent interaction, *we did not apply this scaling procedure to the gas-phase calculation*: that is, the atomic radii are considered to be temperature dependent *only* in the aqueous phase. It is also important to note that the RSF was applied *only* in the electrostatics calculations: for calculating the SASA-dependent contribution at different temperatures, we retain the atomic radii appropriate to 25 °C, thus making $\gamma$ temperature dependent but leaving the SASA temperature independent. The use of separate radii for the surface area and electrostatics calculations might at first sight seem inconsistent; in fact, it should not be thought so, since the two represent completely separate and independent terms.

**(c) Temperature Dependence of Hydration Free Energies for Charged Residues.** P&M's analysis of hydration free energy temperature dependencies was restricted to the neutral side chains. Similar experimental information is not available for the charged residues, at least in a direct form, so estimates have to be obtained from a combination of sources. As we outline below, the approach we have adopted is not without assumptions, and these are expected to place something of a limit on the reliability of the numbers obtained.

In the original parametrization of PARSE, hydration free energies for the charged residues were extracted by using experimental information on the p$K_a$ of the residue in both the gas phase and aqueous phase, together with a thermodynamic cycle. Here we use a similar approach to extract values of $\Delta G_{hyd}$ as a function of temperature [denoted $\Delta G_{hyd}(T)$], adopting as our starting point the thermodynamic cycle illustrated in Figure 2 for acetic acid, which we use as a model for the acidic residues Asp and Glu. The upper and lower horizontal processes describe the ionization (dissociation) process in the gas and aqueous phases, respectively. Experimental information on the free energy change for the former process comes from gas-phase acidity[26] and basicity data;[27] at 25 °C this amounts to 182.7 kcal/mol (see Table 2). The free energy change for the latter process can be obtained using the experimental aqueous phase p$K_a$:

$$\Delta G_{dissoc} = 2.303RT(pK_a) \quad (1)$$

The vertical processes describe transfer from the gas phase to the aqueous phase and therefore constitute the hydration free energies in which we are interested. $\Delta G_{hyd}$ for acetic acid comes from the literature,[21] while that for the acetate ion was taken as −80.65 kcal/mol (see ref 11 for the derivation of this number). $\Delta G_{hyd}$ for water (−6.33 kcal/mol) comes from the work of Marcus and Ben-Naim.[28] Closure of the thermodynamic cycle gives the value of $\Delta G_{hyd}$ for $H_3O^+$ as −108.60 kcal/mol; this is in reasonable agreement with the experimental value of −102 kcal/mol quoted by Pearson.[29] The ∼7 kcal/mol discrepancy is not particularly important providing that a consistent value of $\Delta G_{hyd}$ for $H_3O^+$ is used throughout what follows. We note, however, that in the original PARSE parametrization scheme (which used $H^+$ instead of $H_3O^+$ in the thermodynamic cycle), a range of experimental values of $\Delta G_{hyd}$ for $H^+$ from −254 to −261 kcal/mol was reported,[11] so some degree of uncertainty in the value for $H_3O^+$ is also to be expected.

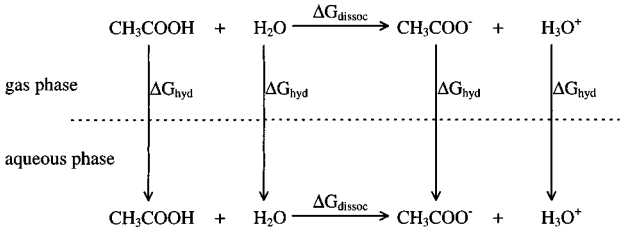


**Figure 2.** Thermodynamic cycle for ionization of acetic acid.

**TABLE 2: Free Energy, Enthalpy, and Entropy Changes for the Gas-Phase Processes Shown[a]**

| process | $\Delta G$ (kcal/mol) | $\Delta H$ (kcal/mol) | $\Delta S$ (EU)[b] |
|---|---|---|---|
| $H_2O + H^+ \rightarrow H_3O^+$ | −159.0 | −166.5 | −25.0 |
| $H_2O \rightarrow H^+ + OH^-$ | 384.1 | 390.7 | 22.0 |
| $CH_3COOH \rightarrow CH_3COO^- + H^+$ | 341.7 | 348.8 | 23.7 |
| $CH_3CH_2CH_2CH_2NH_3^+ \rightarrow CH_3CH_2CH_2CH_2NH_2 + H^+$ | 210.6 | 218.4 | 26.0 |

[a] All data from Bartmess (Ref 26) or Lias et al. (ref 27). [b] cal/mol.K.

To use the thermodynamic cycle of Figure 2 to calculate the temperature dependence, $\Delta G_{hyd}(T)$, for acetate ion, we need to know, among other things, $\Delta G_{hyd}(T)$ for each of the other species present in the cycle, i.e., acetic acid, $H_2O$, and $H_3O^+$. The former presents no problem as it is essentially identical with that of the side chain of *neutral* Asp, for which data were derived by P&M. $\Delta G_{hyd}(T)$ for $H_2O$ is taken from the experimentally derived data given by Marcus and Ben-Naim.[28] $\Delta G_{hyd}(T)$ for $H_3O^+$, however, represents a more difficult problem, and to estimate it we have to make use of a second thermodynamic cycle describing the autoionization process of water (Figure 3).

As with the thermodynamic cycle for acetic acid, we have separate gas-phase and aqueous-phase ionization reactions connected by hydration processes. Again, $\Delta G_{dissoc}$ in both the gas (225.1 kcal/mol; see Table 2) and aqueous phases (19.1 kcal/mol) is obtained directly from experimental data. $\Delta G_{hyd}$

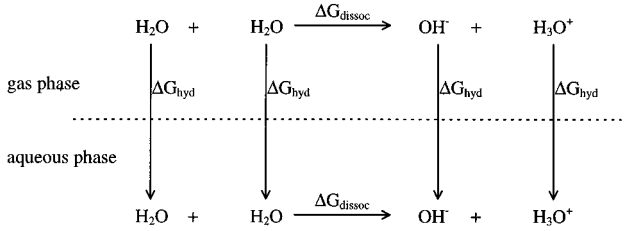


**Figure 3.** Thermodynamic cycle for autoionization of water.

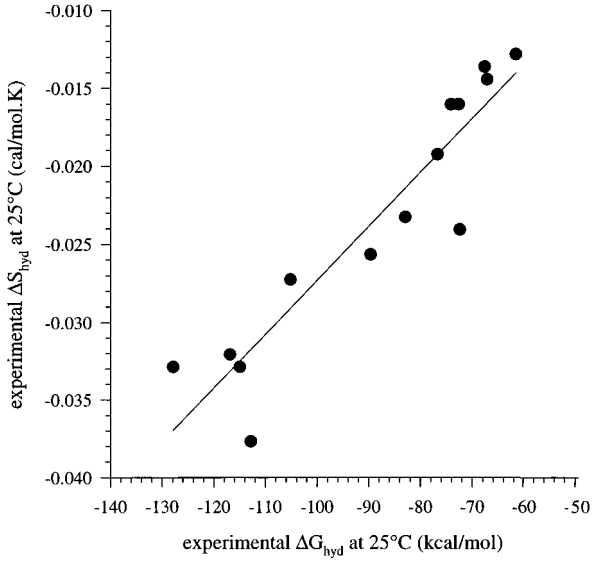


**Figure 4.** Correspondence between free energy of hydration and entropy of hydration for the following ions: $Li^+$, $Na^+$, $K^+$, $Rb^+$, $Cs^+$, $Cu^+$, $Ag^+$, $Tl^+$, $F^-$, $Cl^-$, $Br^-$, $I^-$, $OH^-$, $SH^-$. All data were taken from Marcus (ref 30).

values at 25 °C for both $H_2O$ and $H_3O^+$ were discussed above, so closure of the thermodynamic cycle gives a value of $\Delta G_{hyd}$ for $OH^-$ at 25 °C of −110.08 kcal/mol; again, this is in quite good agreement with the experimental value (−104 kcal/mol) quoted by Pearson.[29]

To obtain the temperature dependence of the gas-phase autoionization process, we make use of the experimental enthalpy and entropy values reported in Table 2. These numbers can be used together with the relation $\Delta G = \Delta H - T\Delta S$ to obtain the temperature dependence $\Delta G_{dissoc}(T)$, since the heat capacity change for the reaction is negligible (i.e. $\Delta H$ and $\Delta S$ are effectively independent of temperature). For the aqueous-phase process, the temperature dependence of the free energy can be obtained directly from the experimental temperature dependence of $K_w$, the autoionization constant of water,[25] using the relation $\Delta G = -RT \ln K_w$. Since, as noted above, we have values of $\Delta G_{hyd}(T)$ for $H_2O$, we can now close the thermodynamic cycle to give us the temperature dependence of [$\Delta G_{hyd}$ ($H_3O^+$) + $\Delta G_{hyd}$ ($OH^-$)]. At this point we have to make a further assumption to separate out the temperature dependence of $\Delta G_{hyd}$ ($H_3O^+$) from that of $\Delta G_{hyd}$ ($OH^-$). We assume that $H_3O^+$ and $OH^-$ make contributions to the temperature dependence in direct proportion to their relative contributions to their combined hydration free energy. That is, we assume that $H_3O^+$ contributes [108.6/(108.6 + 110.08)] × 100%, i.e., 49.7%. We base this assumption on the approximately linear relationship that is observed between $\Delta G_{hyd}$ and $\Delta S_{hyd}$ for simple monovalent ions[30] at 25 °C (see Figure 4) and the expectation that $\Delta S_{hyd}$-(25 °C) should provide a good indication of the temperature dependence of $\Delta G_{hyd}$. Beyond this, it is difficult to justify the assumption, except to say that it does not seem unreasonable

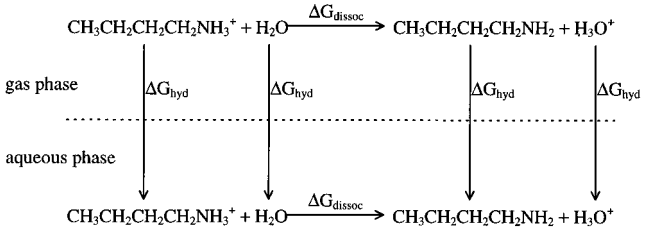


**Figure 5.** Thermodynamic cycle for ionization of *n*-butylamine.

that $H_3O^+$ and $OH^-$ should contribute in approximately equal measure to their combined $\Delta G_{hyd}$ temperature dependence.

Having now obtained a reasonable estimate of $\Delta G_{hyd}(T)$ for $H_3O^+$, we can return to the thermodynamic cycle for acetic acid shown in Figure 2 to extract $\Delta G_{hyd}$(T) for acetate ion. Again, we obtain the temperature dependence of the gas-phase dissociation process using the experimental free energy, enthalpy, and entropy of reaction (Table 2) with the relation $\Delta G = \Delta H - T\Delta S$. Again, we obtain the temperature dependence of the aqueous-phase dissociation process using the experimental data on the p$K_a$ shift of acetic acid with temperature.[31] Knowing all other temperature dependencies in the cycle then, we can calculate $\Delta G_{hyd}(T)$ for acetate ion straightforwardly.

The above discussion has shown how we can obtain estimates of $\Delta G_{hyd}(T)$ for acetate, which we can use as a model for the acidic residues Asp and Glu. In attempting to extend the approach to deal with the basic residues, we were unable to find suitable data to describe Arg; however, we outline below the procedure used to obtain $\Delta G_{hyd}(T)$ for protonated *n*-butylamine, which we adopt as a model for Lys. The thermodynamic cycle used in this case is illustrated in Figure 5. Unfortunately, we were unable to find experimental data on the aqueous-phase p$K_a$ of *n*-butylamine over an extended temperature range; however, we were able to find suitable data for diglycolamine[32] ($HOCH_2CH_2OCH_2CH_2NH_2$), which, despite the slight complication due to the presence of an ether linkage, should be a good substitute. This is supported by the fact that the value of $\Delta S_{dissoc}$ at 25 °C for diglycolamine (−2.9 cal/mol·K; ref 32) is similar to that for *n*-butylamine (−2.0 cal/mol·K; ref 33) All other quantities and temperature dependencies entering into the thermodynamic cycle of Figure 5 are those of the lysine side chain (e.g. the gas-phase data used are that of *n*-butylamine): only the *temperature dependence* of $\Delta G_{dissoc}$ in the aqueous phase is that of diglycolamine.

**(d) Parametrization of Charged Residues.** The parametrization of charged residues was carried out in exactly the same way as described earlier for neutral residues (section b), with two exceptions: first, only model 4 was used, since this worked best for neutral residues; second, the RSF applied to the atoms most closely associated with the molecule's charge was different from that applied to the rest of the molecule. In the case of the carboxylic acids, this meant the atoms of the $-CO_2^-$ group; in the case of lysine, this meant the $-NH_3^+$ group.

## Results and Discussion

P&M have discussed at length the temperature dependencies of the neutral amino acid hydration free energies,[6,7] so it is not our aim to repeat their detailed analysis. Instead, we restrict our attention initially to the process of parametrizing the PARSE continuum solvation model to reproduce their results. Before doing so, however, it is worth briefly stating the qualitative behavior noted by P&M in their study: in the range 5−100 °C, *hydration free energies become less favorable as the temperature increases*, regardless of whether the molecule is

**TABLE 3: Calculated and Experimental (Ref 7) Changes in Hydration Free Energies Resulting from Increasing Temperature from 25 to 100 °C[a]**

| residue | 1 | 2 | 3 | 4 | 5 | 6 | 7 | exptl |
|---|---|---|---|---|---|---|---|---|
| Ala | 0.60 | 0.56 | 0.44 | 0.62 | 0.47 | 0.52 | 0.50 | 0.44 |
| Arg | | | | | | 1.79 | | 1.49 |
| Asn | 0.81 | 0.86 | 1.00 | 0.88 | 1.10 | 1.06 | 1.09 | 1.18 |
| Asp | 0.79 | 0.82 | 0.92 | 0.86 | 1.00 | 0.96 | 0.99 | 0.92 |
| Cys | 0.74 | 0.72 | 0.67 | 0.80 | 0.49 | 0.53 | 0.58 | 0.39 |
| Gln | 0.93 | 0.98 | 1.09 | 1.01 | 1.24 | 1.20 | 1.22 | 1.31 |
| Glu | 0.92 | 0.94 | 1.04 | 0.99 | 1.14 | 1.12 | 1.13 | 1.03 |
| His | 0.96 | 1.00 | 1.13 | 0.81 | 1.20 | 1.02 | 1.09 | 1.23 |
| Ile | 0.97 | 0.91 | 0.71 | 1.00 | 0.77 | 0.84 | 0.81 | 0.79 |
| Leu | 0.96 | 0.90 | 0.70 | 0.99 | 0.76 | 0.84 | 0.80 | 0.83 |
| Lys | 1.04 | 1.04 | 1.05 | 1.10 | 1.25 | 1.25 | 1.22 | 1.29 |
| Met | | | | | | 1.01 | | 1.39 |
| Phe | 1.07 | 1.03 | 0.90 | 0.81 | 0.52 | 0.38 | 0.52 | 0.55 |
| Ser | 0.65 | 0.68 | 0.77 | 0.69 | 0.99 | 0.96 | 0.94 | 0.99 |
| Thr | 0.78 | 0.80 | 0.87 | 0.82 | 1.12 | 1.09 | 1.06 | 1.05 |
| Trp | | | | | | 0.72 | | 0.36 |
| Tyr | 1.12 | 1.13 | 1.17 | 0.95 | 0.98 | 0.82 | 0.93 | 0.92 |
| Val | 0.86 | 0.80 | 0.63 | 0.88 | 0.68 | 0.75 | 0.72 | 0.68 |
| Nma | 0.97 | 1.02 | 1.11 | 1.10 | 0.69 | 0.67 | 0.84 | 0.74 |
| av error | 0.25 | 0.23 | 0.17 | 0.23 | 0.05 | 0.13 | 0.07 | |
| max error | 0.52 | 0.48 | 0.37 | 0.42 | 0.10 | 0.38 | 0.19 | |

[a] All values are in kcal/mol. The various models used are discussed in the text.

polar, aliphatic, or aromatic in nature. If we omit Arg, Met, and Trp from consideration, the change in $\Delta G_{hyd}$ that accompanies the temperature increase from 25 to 100 °C averages around 0.90 kcal/mol for the neutral amino acid side chains, with a maximum change of 1.31 kcal/mol for glutamine.

**Parametrization of Neutral Residues.** As discussed under Methods, a number of different approaches were tried for fitting to the data; for the sake of brevity we focus only on the results obtained at the single temperature of 100 °C. Model 1 seeks to describe the temperature dependence $\Delta G_{hyd}(T)$ in terms only of a single SASA-dependent term. Not surprisingly, given that this approach does not appear to work well in reproducing absolute hydration free energies,[7] it also fails to accurately describe $\Delta G_{hyd}(T)$ (see column 1 of Table 3): the average error per residue amounts to 0.25 kcal/mol, which, given that the average change in $\Delta G_{hyd}$ is 0.90 kcal/mol, corresponds to an approximately 30% error. Matters are slightly improved by the use of model 2, which adds a temperature-dependent electrostatic component $\Delta G_{elec}$ to the SASA-dependent term: column 2 gives the results obtained from $\Delta G_{elec}$ calculations that assumed that the only temperature-dependent variable is the solvent dielectric. Further study showed, however, that better fits to the data could be obtained by assuming that not only the solvent dielectric but also the atomic radii vary with temperature. This was shown by recalculating $\Delta G_{elec}$(100 °C) with all of the atomic radii multiplied by a radius scaling factor (see Methods) and repeating the fits to the data using model 2. Figure 6 shows the average error for $\Delta\Delta G_{hyd}$(25 °C→100 °C) as a function of the RSF. The error shows a roughly parabolic dependence on the RSF, decreasing as it is increased from 1.00 to 1.02 before rising again as it exceeds 1.03. Note that the average error for an RSF value of 1.00 in Figure 2 is the same as that given in column 2 of Table 3; column 3 gives the results obtained with the optimum value of the RSF (1.020). Inclusion of a temperature dependence of the atomic radii into the model therefore results in a decrease in the average error from 0.23 to 0.17 kcal/mol.

A qualitative guide to the behavior of the electrostatic component with changing parameters can be obtained from the Born model[34] used to describe (and strictly only true for) the

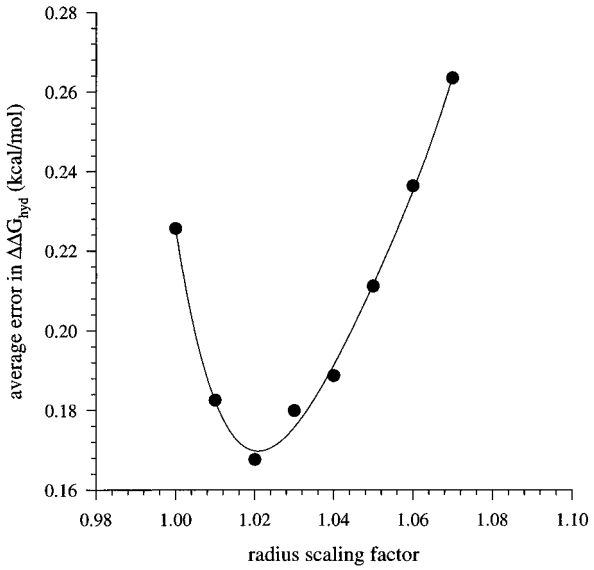


**Figure 6.** Dependence of average error on radius scaling factor (RSF) at 100 °C, obtained using model 2.

electrostatic solvation of spherical ions:

$$\Delta G_{elec} = \frac{166.03q^2}{r}\left\{\frac{1}{\epsilon} - 1\right\}$$

Here $q$ represents the atomic charge, $r$ the atomic radius, $\epsilon$ the solvent dielectric, and $\Delta G_{elec}$, as before, the electrostatic free energy change resulting from transfer of the ion from the gas phase to the solvent. From the above equation it can be seen that since an increase in temperature results in a decrease in the solvent dielectric, the electrostatic solvation energy becomes less negative (i.e. less favorable) at higher temperatures. It is also clear that an increase in the atomic radius, or a decrease in the atomic charge, will cause a similar qualitative change in the electrostatic solvation energy. Now, since better fits to the data are obtained when the atomic radii are increased at higher temperatures, it would appear that changing *only* the solvent dielectric underestimates the unfavorable change in hydration of polar groups that accompanies increasing the temperature.

It would be misleading to suggest that this result indicates that atomic radii do actually increase with temperature: since the radius simply plays the role of an adjustable parameter, it is important not to place too much emphasis on its meaning. It is, however, worth pointing out that a similar need for increasing radii with temperature has been found in studies attempting to use the Born model to describe the hydration entropies of simple monovalent ions.[35] Differentiating the Born equation for the free energy with respect to temperature gives an analytical expression for the entropy that has separate $\partial\epsilon/\partial T$ and $\partial r/\partial T$ terms. Rashin and Bukatin[36] have used data on thermal expansion coefficients to argue that the $\partial r/\partial T$ term is effectively zero and can therefore be neglected. Kono et al.[35] have shown, however, that inclusion of this term is essential if the Born model is to be used to obtain correct values for $\Delta S_{hyd}$. The argument of Rashin and Bukatin[36] assumes (quite reasonably) that $\partial r/\partial T$ should have a physical meaning. One might hope, for example, that a nonzero value for $\partial r/\partial T$ would be manifested in changes in the peak positions for the solute−solvent radial distribution functions: in this regard Kono et al. refer to the RISM calculations of Roux et al.,[37] which show a value of $\partial r/\partial T = 8.4 \times 10^{-4}$ Å/K for $Cl^-$. For comparison, an RSF of 1.030 at 100 °C corresponds to a value of $\partial r/\partial T = 8.0 \times 10^{-4}$ Å/K for an atom of 2 Å radius. The radius increments obtained in our

work are therefore of similar magnitude to those found in more sophisticated theoretical treatments. This is clearly an area that warrants further investigation, however, and since there appears to have been little discussion up to now of this issue in the literature, it is perhaps safest for the moment to view the increase in atomic radius in the present context as simply an empirical adjustment, which may or may not have a physical analogue.

The above discussion refers to results obtained using a single $\gamma$ value applied to the entire SASA. As outlined in the Introduction, however, P&M assigned different values of $\gamma$ to the polar, aliphatic, and aromatic regions.[6,7] They found that the temperature dependence of $\Delta G_{hyd}$ for the nonpolar side chains could be described well using this approach but that extending it to include the polar regions of other residues worked less well because the derived values for $\gamma_{polar}$ varied strongly from residue to residue (see Table 5 of ref 6 and Table 4 of ref 7): it was not possible to arrive at a single average $\gamma_{polar}$ value that worked well for all polar residues. Clearly, the use of three independent $\gamma$ values represents an increase in complexity of the model, but it is nevertheless of interest to see if results can be significantly improved by use of this approach. As mentioned under Methods, we maintain consistency with the original PARSE scheme by ensuring that at 25 °C $\gamma_{aliphatic} = \gamma_{aromatic} = \gamma_{polar}$. Column 4 of Table 3 shows the results obtained using model 3.

Not surprisingly, given the use of more parameters in the fitting procedure, better agreement is obtained when three $\gamma$ values are used instead of a single $\gamma$ (compare columns 1 and 4). However, importantly, agreement is not as good as when a single $\gamma$ is combined with an optimized $\Delta\Delta G_{elec}$ term (compare columns 3 and 4): this result suggests that a properly parametrized electrostatic component is more capable of describing the temperature dependence of polar group hydration than a SASA-dependent component. The point becomes clearer when one considers the results obtained when three $\gamma$ values are combined with an optimized $\Delta\Delta G_{elec}$ (column 5): the agreement with experimental results now becomes excellent, with the average error being reduced to 0.05 kcal/mol, corresponding to a percentage error of only around 6%. Compared to the single $\gamma$ value model (column 3), a rather higher value of the RSF, 1.055, provided the best results. It is worth noting in passing that when the three omitted residues, Arg, Met, and Trp, are reintroduced to the analysis and the fit using model 4 is repeated, the results become much worse (column 6). In particular, the average error per residue more than doubles and the maximum error for any one residue almost quadruples. The change in $\Delta G_{hyd}$ for Met is strongly underestimated by the model (or overestimated by P&M), while the changes for both Arg and Trp appear strongly overestimated.

Two other aspects of the results are especially notable. First, there is a clear linear relationship between the best-fit $\Delta\gamma$ values obtained for a particular RSF and the value of the RSF itself (Figure 7). All three $\Delta\gamma$ values decrease in magnitude as the RSF increases, since the contribution from $\Delta\Delta G_{elec}$ becomes progressively more dominant. Furthermore, the best-fit $\Delta\gamma$ values show a dependence on the RSF in the order $\Delta\gamma_{polar} > \Delta\gamma_{aromatic} > \Delta\gamma_{aliphatic}$. This is not particularly surprising but is, nevertheless, encouraging, since it is to be expected that the SASA contribution from polar groups should be the most affected by changes in the relative contribution of $\Delta\Delta G_{elec}$. What is more interesting, however, is the fact that both the $\Delta\gamma_{polar}$ and $\Delta\gamma_{aromatic}$ plots actually pass through zero: i.e., there exist values of the RSF at which the SASA-dependent polar and aromatic terms contribute nothing to the temperature dependence of the hydration free energy. As parameters for use in the fitting

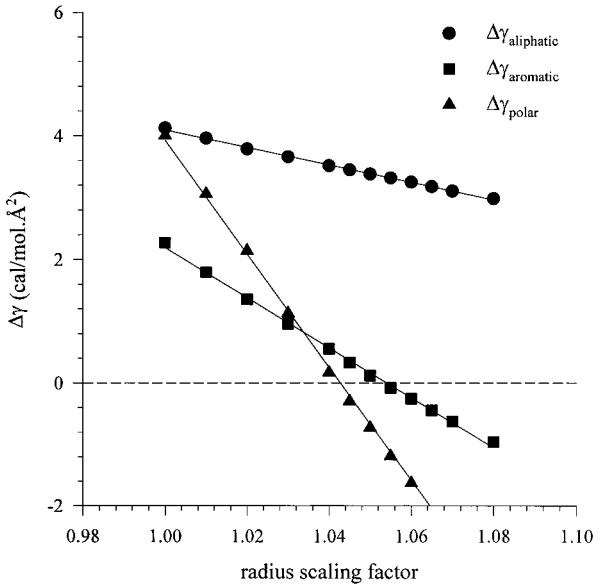


**Figure 7.** Dependence of best-fit $\Delta\gamma$ values on the radius scaling factor at 100 °C.

procedure then, $\Delta\gamma_{polar}$ and $\Delta\gamma_{aromatic}$ are, at these values of the RSF, essentially redundant. It is interesting then that $\Delta\gamma_{polar}$ and $\Delta\gamma_{aromatic}$ reach zero at similar values of the RSF and that these values are in turn similar to the overall optimal RSF value. This suggests that it should be possible to set both $\Delta\gamma_{polar}$ and $\Delta\gamma_{aromatic}$ to zero with little effect on the overall agreement between calculated and experimental hydration free energies: that this is indeed the case can be seen by comparing columns 7 and 5. The small effect on the average error that results from constraining $\Delta\gamma_{aromatic}$ to be zero is perhaps to be expected since only 3 of the 16 residues fitted contain aromatic regions: if a greater proportion of the molecules in the parametrization set were aromatic, this might no longer hold true. The fact that $\Delta\gamma_{polar}$ can be set to zero with little effect on the overall error is more interesting, however, since 11 of the 16 residues contain polar regions: this suggests that the contributions from polar groups to the temperature dependence of hydration free energies can be adequately described by the use of only a properly parameterized electrostatic component.

Although both $\Delta\gamma_{polar}$ and $\Delta\gamma_{aromatic}$ can be omitted from model 4 with little loss in accuracy, we decided to retain $\Delta\gamma_{aromatic}$ to establish a functional form that is more likely to be generalizable in the future to a wider range of molecules. One could argue, of course, that if our purpose is simply to reproduce the temperature dependencies of the amino acid residues as accurately as possible, we should also retain $\Delta\gamma_{polar}$ since inclusion of this term involves essentially no additional expense in the calculations and results in better fits to the data. However, it is considerably less appealing from an intuitive viewpoint to consider polar group hydration as being described by a combination of a $\Delta\gamma_{polar}$ term and a $\Delta\Delta G_{elec}$ term, particularly when, as happens at higher temperatures, the two terms act in opposition to one another. Model 4 was therefore used (with $\Delta\gamma_{polar}$ set to zero) to fit P&M's experimental results at 5, 50, 75, and 100 °C. The complete set of optimized $\gamma_{aliphatic}$, $\gamma_{aromatic}$, and RSF values are given in Table 4 together with the average and maximum errors obtained. A plot of the optimal RSF as a function of temperature $T$ (Figures 8 and 9) can be fitted to a quadratic function:

$$\mathrm{RSF} = 0.991252 + 3.142959 \times 10^{-4} \times T(^{\circ}\mathrm{C}) + 2.354012 \times 10^{-6} \times T(^{\circ}\mathrm{C})^2 \qquad (r^2 = 0.9996)$$

**TABLE 4: Optimized $\gamma$ and RSF Values for Each of the Various Temperatures, Together with the Average and Maximum Errors Obtained**

| temp (°C) | $\gamma_{aliphatic}$ | $\gamma_{aromatic}$ | RSF | av error | av $\lvert\Delta\Delta G_{hyd}\rvert$ | max error | max $\lvert\Delta\Delta G_{hyd}\rvert$ |
|---|---|---|---|---|---|---|---|
| 5 | 3.6 | 4.7 | 0.9932 | 0.03 | 0.34 | 0.06 | 0.47 |
| 25 | 5.4 | 5.4 | 1.000 | | | | |
| 50 | 7.0 | 5.8 | 1.0129 | 0.02 | 0.36 | 0.06 | 0.50 |
| 75 | 8.1 | 5.9 | 1.0285 | 0.04 | 0.66 | 0.08 | 0.93 |
| 100 | 8.7 | 5.7 | 1.0460 | 0.07 | 0.90 | 0.18 | 1.31 |

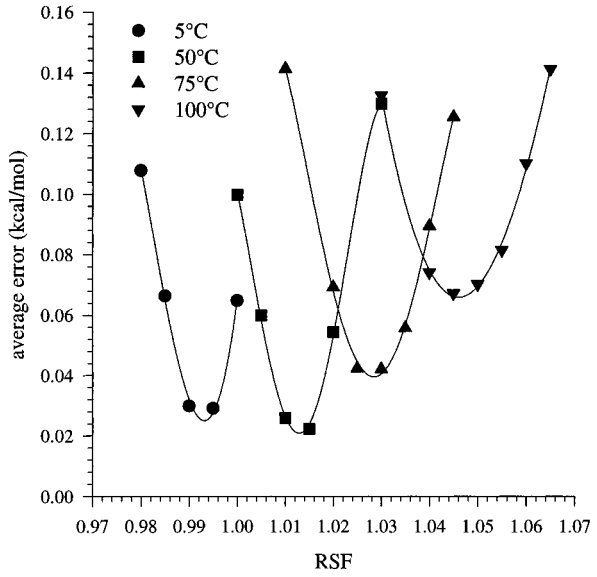


**Figure 8.** Dependencies of average errors on radius scaling factor at 5, 50, 75, and 100 °C obtained using model 4.

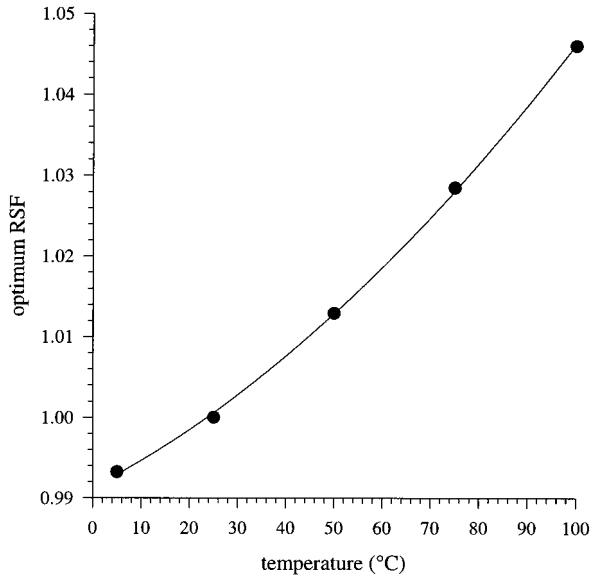


**Figure 9.** Temperature dependence of the optimum RSF value.

The optimized values of $\gamma_{aliphatic}$ and $\gamma_{aromatic}$, shown in Figure 10, can also be fitted by quadratic functions:

$$\gamma_{aliphatic} = 3.134694 + 0.1001063 \times T(°C) - 4.418417 \times 10^{-4} \times T(°C)^2 \quad (r^2 = 0.9998)$$

$$\gamma_{aromatic} = 4.511007 + 0.0407585 \times T(°C) - 2.928069 \times 10^{-4} \times T(°C)^2 \quad (r^2 = 0.9950)$$

Our temperature dependencies of $\gamma_{aliphatic}$ and $\gamma_{aromatic}$ are in reasonable qualitative agreement with those obtained by P&M,

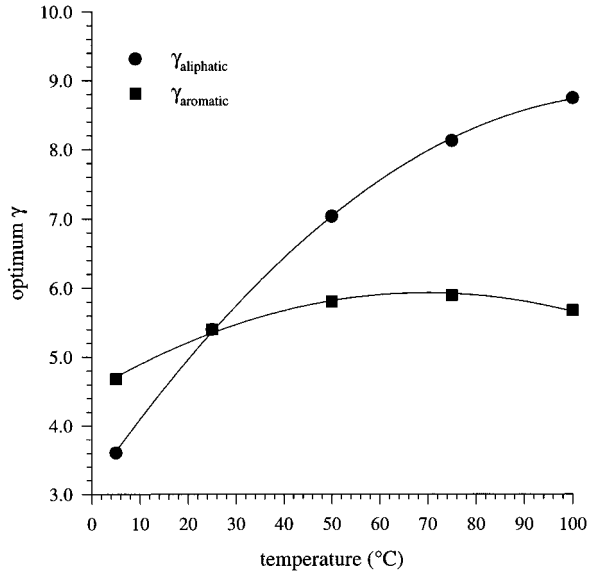


**Figure 10.** Temperature dependence of the optimized $\gamma$ values.

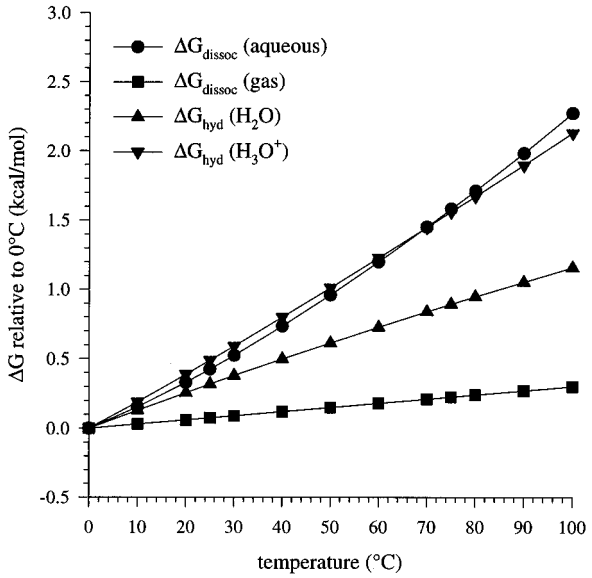


**Figure 11.** Temperature dependence of the processes comprising the thermodynamic cycle for autoionization of water (depicted in Figure 3).

a result that is to be expected since we are essentially analyzing the same data; it is, however, reassuring that our emphasis on using a continuum electrostatics method for dealing with polar group hydration does not lead to drastic changes in the relative contribution of aliphatic and aromatic groups to $\Delta G_{hyd}(T)$.

**Temperature Dependence of Hydration Free Energies for Charged Residues.** The approach that we have adopted in the case of the charged amino acid residues makes use of data on the temperature dependence of the autoionization product of water, $K_w$, to obtain $\Delta G_{hyd}$ as a function of temperature [$\Delta G_{hyd}$(T)] for $H_3O^+$ (see Methods). Figure 11 shows the temperature dependencies of each of the processes present in the $K_w$ thermodynamic cycle (Figure 3), with the exception of $\Delta G_{hyd}(T)$ for $OH^-$, which shows an essentially identical behavior to that of $H_3O^+$. Each process becomes more unfavorable as the temperature increases, although the magnitude of the temperature dependence varies significantly between processes. As described under Methods, we calculated the temperature dependence of the gas-phase dissociation process using the relation $\Delta G = \Delta H - T\Delta S$ and gas-phase acidity and basicity data. Since $\Delta S$ is so small in magnitude (3 cal/mol·K), $\Delta G_{dissoc}$ for the gas-

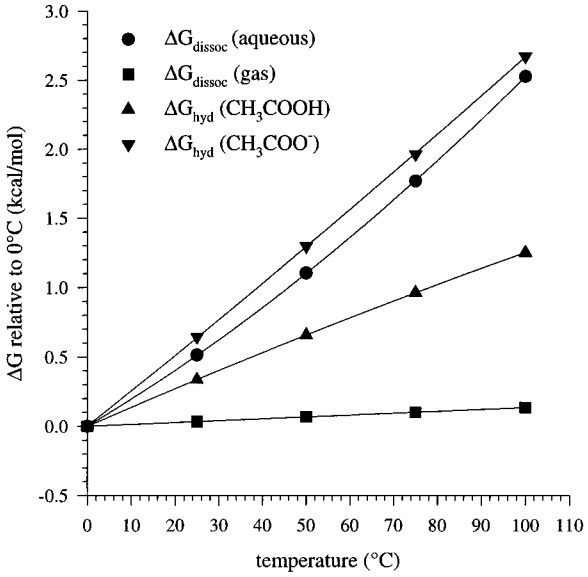


**Figure 12.** Temperature dependence of the processes comprising the thermodynamic cycle for ionization of acetic acid (depicted in Figure 2).

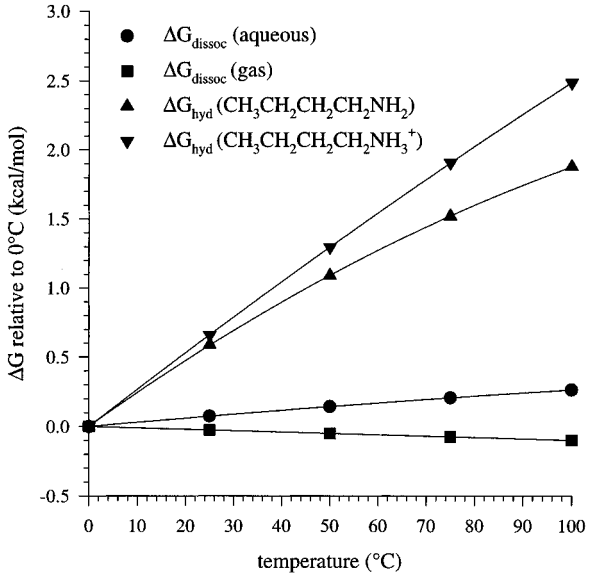


**Figure 13.** Temperature dependence of the processes comprising the thermodynamic cycle for ionization of *n*-butylamine (depicted in Figure 5).

phase process changes by only 0.3 kcal/mol when the temperature is increased from 0 to 100 °C. In contrast, the aqueous-phase dissociation process, as calculated from the experimental $K_w$ values, changes by more than 2 kcal/mol over the same temperature range. To account for this difference there must be significant differences in the temperature dependencies of $\Delta G_{hyd}$ for $H_3O^+$ and $OH^-$ relative to that for $H_2O$. Under the assumption used to extract $\Delta G_{hyd}(T)$ for $H_3O^+$ and $OH^-$, it is clear that both are much more strongly affected by increasing temperature than is $H_2O$. As we shall see below, this greater temperature dependence exhibited by charged molecules relative to their uncharged analogues is a general feature, and one that is in line with the relationship shown earlier (Figure 4) between $\Delta G_{hyd}$ and $\Delta S_{hyd}$ for simple ions.

Having an estimate of $\Delta G_{hyd}(T)$ for $H_3O^+$ allows us to calculate $\Delta G_{hyd}(T)$ for acetate ion using the thermodynamic cycle shown in Figure 2. We show the temperature dependencies of each of the processes in the cycle in Figure 12. Again, $\Delta G_{dissoc}$ in the gas phase shows only a very modest dependence on temperature, since $\Delta S_{dissoc}$ is calculated as being only $-1.333$ cal/mol·K from combining experimental data on the gas-phase acidity of acetic acid[26] and the gas-phase basicity of water[27] (Table 2). As before, $\Delta G_{dissoc}$ in the aqueous phase increases by ~2 kcal/mol over the same temperature range. This might at first sight seem odd since the p$K_a$ of acetic acid actually changes only very slightly from 0 to 100 °C;[31] it should be remembered, however, that the expression used for calculating $\Delta G_{dissoc}$ from p$K_a$ values has the temperature $T$ present as a multiplicative factor (eq 1). A temperature-independent p$K_a$ value does not therefore imply a temperature-independent $\Delta G_{dissoc}$ value; in fact, quite the opposite is the case. $\Delta G_{hyd}(T)$ for $H_2O$ and acetic acid (the latter obtained directly from P&M's results) show a temperature dependence approximately half that of $H_3O^+$ and acetate, respectively.

We use a similar method to obtain $\Delta G_{hyd}(T)$ for protonated *n*-butylamine, which is essentially identical with the side chain of Lys, by use of the thermodynamic cycle of Figure 5. Once again (Figure 13), $\Delta G_{dissoc}$ in the gas phase is only weakly affected by changing temperature ($\Delta S_{dissoc} = 1.0$ cal/mol·K). In contrast to the above results for water and acetic acid, however, $\Delta G_{dissoc}$ in the aqueous phase for *n*-butylamine is *also* largely independent of temperature: this demonstrates further the different behaviors of p$K_a$ and $\Delta G_{dissoc}$ values since the p$K_a$ of diglycolamine changes by more than 2.5 units over the 100 °C temperature range. Nevertheless, the same result is obtained as before: $\Delta G_{hyd}$ for the charged residue is more strongly affected by temperature (Figure 13), although in this case the difference is far less marked.

The much smaller temperature dependence of the aqueous phase $\Delta G_{dissoc}$ process for protonated *n*-butylamine compared to acetic acid (or water) can be understood in terms of the hydration of the species present on either side of the equation.[38] In the case of the dissociation process of the protonated amine, each side of the equation contains one neutral and one charged molecule: such reactions have been termed *isoelectric* or *iso-Coulombic*. In the case of the dissociation process of the carboxylic acid, the left-hand side of the equation contains two neutral species, while the right-hand side contains one cation and one anion. Since the entropy of hydration of charged species is more negative than for neutral analogues, processes in which there is a net increase in the number of charged species are more adversely affected by increasing temperature than those that do not. A large unfavorable change is therefore expected for the carboxylic acid ionization process, while little change is expected for the amine ionization process.

**Parametrization of Charged Residues.** As for the case of the neutral residues, we parametrize the charged residues using model 4. We hoped initially that the same $\gamma$ and RSF values derived for the neutral residues would also provide adequate descriptions of $\Delta G_{hyd}(T)$ for the charged residues. It was found, however, that use of the neutral residue RSF values results in a considerable overestimation of the temperature dependence, so for the $-COO^-$ and $-NH_3^+$ groups separate RSF values were derived. Figure 14 shows the temperature dependence of the optimized RSF values for these two groups; both can be fitted well to quadratic functions, with the $-NH_3^+$ group showing a smaller temperature dependence.

$$\mathrm{RSF}(NH_3^+) = 0.999089 - 1.485714 \times 10^{-6} \times T(°C) + 1.302857 \times 10^{-6} \times T(°C)^2 \qquad (r^2 = 0.9987)$$

$$\mathrm{RSF}(COO^-) = 0.990259 + 3.273143 \times 10^{-4} \times T(°C) + 1.302857 \times 10^{-6} \times T(°C)^2 \qquad (r^2 = 0.9991)$$

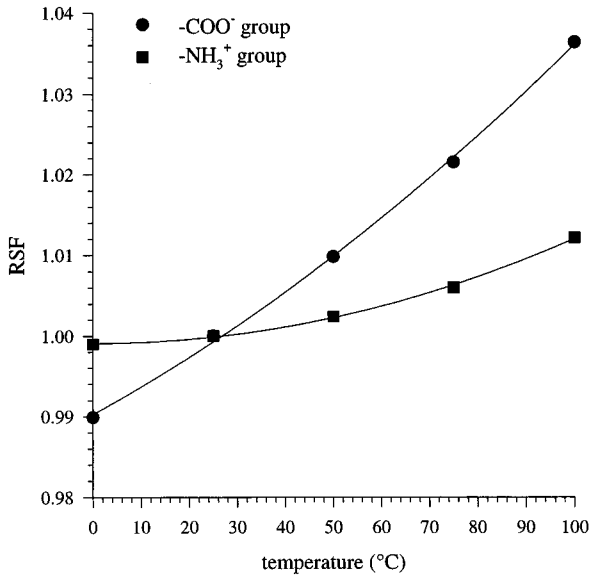


**Figure 14.** Temperature dependence of the optimized RSF values for the $-NH_3^+$ and $COO^-$ groups.

**TABLE 5: Comparison of Calculated and Experimental Entropies of Hydration for a Series of Small Neutral Molecules**[a]

| molecule | $\Delta S_{hyd}$ (calcd) | $\Delta S_{hyd}$ (exptl) | molecule | $\Delta S_{hyd}$ (calcd) | $\Delta S_{hyd}$ (exptl) |
|---|---|---|---|---|---|
| methanol | −14.6 | −17.1 | acetic acid | −13.8 | −18.0 |
| ethanol | −17.3 | −23.5 | propanoic acid | −16.6 | −21.7 |
| 1-propanol | −19.1 | −28.2 | butanoic acid | −18.7 | −24.5 |
| 2-propanol | −19.3 | −28.9 | benzene | −8.9 | −20.7 |
| butanol | −21.5 | −31.8 | methylbenzene | −13.9 | −24.3 |
| propanone | −17.6 | −18.0 | 1,3-dimethylbenzene | −19.0 | −28.6 |
| butanone | −20.4 | −22.6 | phenol | −12.3 | −21.6 |
| ethylamine | −16.3 | −26.4 | 2-methylphenol | −17.0 | −28.7 |
| propylamine | −18.6 | −28.1 | pyridine | −11.3 | −22.4 |
| butylamine | −20.8 | −31.1 | 2-methylpyridine | −16.5 | −26.8 |

[a] All values are given in EU (cal/mol·K). Experimental results are taken from refs 6 and 7.

**Application to Other Molecules.** Although we have concentrated here entirely on the temperature dependence of amino acid hydration free energies, it is of interest to see if the derived parameters can also be carried over to treat other similar molecules. Hydration free energies over an extended temperature range do not appear to have been reported for many molecules, but entropies of hydration at 25 °C are at least available for quite a wide range of neutral compounds. Table 5 shows the correspondence between entropies of hydration calculated by applying the parameters given in Table 4 with the experimental results.[6,7,21] Agreement is only modest: the calculated values are similar in magnitude to the experimental numbers and, within any given series of homologues, show the correct qualitative change as the alkyl chain increases in length. Quantitatively, however, the agreement is often quite poor and is particularly so for aromatic molecules: the calculated $\Delta S_{hyd}$ value for benzene, for example, is −9 cal/mol·K, while the experimental value is −21 cal/mol·K.

These are obviously rather disappointing results and demonstrate that although our parameters describe the amino acid residues well enough, they may not be directly transferable to other molecules. In part, this may result from the fact that in applying our amino acid parameters to small model compounds we are essentially taking a route exactly opposite from that taken by P&M, who sought to describe the amino acids using parameters derived from model molecules. The fact that poor results are obtained with aromatic molecules, for example, is almost certainly due to under-representation of aromatic compounds in the set of molecules used to derive the parameters (i.e. the neutral amino acid side chains). There seems little reason to believe that the basic model that we have used is unsuitable for application to a wider range of small molecules; clearly though, such applications will require efforts to be made to properly parametrize the model.

## Conclusions

We have described here the development of a parameter set that accurately reproduces experimentally based estimates of amino acid hydration free energies in the temperature range 5−100 °C. In the case of the neutral residues we have taken the results derived by P&M[6,7] as our basis; for the charged residues we have combined these results with data on the temperature dependence of p$K_a$ values to arrive at hydration free energies. In extending the PARSE parameter set to account for the temperature dependencies of hydration free energies, we tried a number of approaches: it was found that best results were obtained when a temperature-dependent continuum electrostatics component was combined with separate temperature-dependent SASA components for polar, aliphatic, and aromatic regions. The results were sensitive to the details of the electrostatics calculations and were markedly improved when the atomic radii were scaled by a temperature-dependent factor, or RSF, which increases in magnitude as the temperature increases. A similar need for increasing radii at higher temperatures has been found in applications of the Born model to the hydration entropies of simple ions.[35,37] We also found here that when such an optimized electrostatics component was used, the SASA-dependent contributions from polar and aromatic groups could be neglected with little loss in accuracy. For polar groups this seems to be a general conclusion, but for aromatic groups this result seems more likely a consequence of under-representation of aromatic compounds in the set of molecules used to parametrize the method.

From our analysis of charged residues, the most important result is that their hydration free energies exhibit a considerably greater degree of temperature dependence than their uncharged analogues. We have commented elsewhere on the possible implications of this result for salt bridge stability in proteins;[39] here we have restricted our attention to the development of parameters to reproduce these temperature dependencies. It is worth stressing, of course, that our results for charged residues are dependent on the results obtained previously for neutral residues by P&M. In one respect, this is a clear disadvantage, since errors in the neutral residues will be propagated to the charged residues, giving a measure of uncertainty to the *absolute* values obtained. In another respect however, this is actually an advantage, since the *relative* values of charged and neutral analogues should be quite accurately reproduced. In most applications of interest, for example, calculating p$K_a$ values of ionizable groups as a function of temperature, it is these relative values that are of importance: since the model has been parametrized for this purpose, it should be of use in such applications.

To summarize, the parametrization work reported here allows us to study the temperature dependence of protein and peptide hydration using a Poisson−Boltzmann-based model. This means that continuum electrostatics models can now be used to investigate the effects of each of the most important environmental variables on protein stability, since previous studies have demonstrated the utility of PB-based models in describing the effects of both ionic strength[40] and pH[40−42] on protein thermodynamics.

**Acknowledgment.** A.H.E. is very grateful to the Wellcome Trust for financial support under Grant 041157/Z/94/Z. This work was supported in part by grants from the National Institutes of Health, the National Science Foundation, and the National Science Foundation MetaCenter Program.

## References and Notes

(1) Makhatadze, G. I.; Privalov, P. L. *Adv. Protein Chem.* **1995**, *47*, 307.
(2) Honig, B.; Yang, A. S. *Adv. Protein Chem.* **1995**, *46*, 27.
(3) Oobatake, M.; Ooi, T. *Prog. Biophys. Mol. Biol.* **1993**, *59*, 237.
(4) Chothia, C. *Nature* **1974**, *248*, 338.
(5) Hermann, R. B. *J. Phys. Chem.* **1972**, *76*, 2754.
(6) Makhatadze, G. I.; Privalov, P. L. *J. Mol. Biol.* **1993**, *232*, 639.
(7) Privalov, P. L.; Makhatadze, G. I. *J. Mol. Biol.* **1993**, *232*, 660.
(8) Ooi, T.; Oobatake, M.; Némethy, G.; Scheraga, H. A. *Proc. Natl. Acad. Sci. U.S.A.* **1987**, *84*, 3086.
(9) Lazaridis, T.; Archontis, G.; Karplus, M. *Adv. Protein Chem.* **1995**, *47*, 231.
(10) Honig, B.; Nicholls, A. *Science* **1995**, *268*, 1144.
(11) Sitkoff, D.; Sharp, K. A.; Honig, B. *J. Phys. Chem.* **1994**, *98*, 1978.
(12) Arald, J. C.; Nicholls, A.; Sharp, K. A.; Honig, B.; Tempczyk, A.; Hendrickson, T. F.; Still, W. C. *J. Am. Chem. Soc.* **1991**, *113*, 1454.
(13) Ewing, T. J. A.; Lybrand, T. P. *J. Phys. Chem.* **1994**, *98*, 1748.
(14) Mohan, V.; Davis, M. E.; McCammon, J. A.; Pettitt, B. M. *J. Phys. Chem.* **1992**, *96*, 6428.
(15) Elcock, A. H.; Richards, W. G. *J. Am. Chem. Soc.* **1993**, *115*, 7930.
(16) Sitkoff, D.; Lockhart, D. J.; Sharp, K. A.; Honig, B. *Biophys. J.* **1994**, *67*, 2251.
(17) Yang, A.-S.; Honig, B. *J. Mol. Biol.* **1995**, *252*, 351.
(18) Yang, A.-S.; Honig, B. *J. Mol. Biol.* **1995**, *252*, 366.
(19) Makhatadze, G. I.; Privalov, P. L. *J. Mol. Biol.* **1990**, *213*, 375.
(20) Wolfenden, R.; Andersson, L.; Cullis, P. M.; Southgate, C. C. B. *Biochemistry* **1981**, *20*, 849.
(21) Cabani, S.; Gianni, P.; Mollica, V.; Lepori, L. *J. Solution Chem.* **1981**, *10*, 563.
(22) Sigmaplot, Jandel Scientific Software, San Rafael, CA, 1994.
(23) QUANTA version 4.0, Molecular Simulations Inc., Waltham, MA, 1994.
(24) (a) Madura, J. D.; Davis, M. E.; Gilson, M. K.; Wade, R. C.; Luty, B. A.; McCammon, J. A. *Rev. Comput. Chem.* **1994**, *5*, 229. (b) Madura, J. D.; Briggs, J. M.; Wade, R. C.; Davis, M. E.; Luty, B. A.; Ilin, A.; Antosiewicz, J.; Gilson, M. K.; Bagheri, B.; Scott, L. R.; McCammon, J. A. *Comput. Phys. Comm.* **1995**, *91*, 57.
(25) *CRC Handbook of Chemistry and Physics*, 72nd ed.; CRC Press: Boca Raton, FL 1992.
(26) Bartmess, J. *The 1987 Gas-Phase Acidity Scale*; Department of Chemistry, University of Tennessee: Knoxville, TN, 1987.
(27) Lias, S. G.; Liebman, J. F.; Levin, R. D. *J. Phys. Chem. Ref. Data* **1984**, *13*, 695.
(28) Marcus, Y.; Ben-Naim, A. *J. Chem. Phys.* **1985**, *83*, 4744.
(29) Pearson, R. G. *J. Am. Chem. Soc.* **1986**, *108*, 6109.
(30) Marcus, Y. *Ion Solvation*, Wiley-Interscience: Chichester, U.K., 1985.
(31) Fisher, J. R.; Barnes, H. L. *J. Phys. Chem.* **1972**, *76*, 90.
(32) Oscarson, J. L.; Wu, G.; Faux, P. W.; Izatt, R. M.; Christensen, J. J. *Thermochim. Acta* **1989**, *154*, 119.
(33) Martell, A. E. *Critical Stability Constants*, Plenum Press: New York, 1974.
(34) Born, M. *Z. Phys.* **1920**, *1*, 45.
(35) Kono, H.; Ohtsuki, Y.; Abe, T. *J. Phys. Chem.* **1996**, *100*, 9935.
(36) Rashin, A. A.; Bukatin, M. A. *Biophys. Chem.* **1994**, *51*, 167.
(37) Roux, B.; Yu, H.-A.; Karplus, M. *J. Phys. Chem.* **1990**, *94*, 4683.
(38) Edsall, J. T.; Wyman, J. *Biophysical Chemistry*; Academic Press: New York, 1958.
(39) Elcock, A. H.; McCammon, J. A. Manuscript in preparation.
(40) Yang, A.-S.; Honig, B. *J. Mol. Biol.* **1994**, *237*, 602.
(41) Yang, A.-S.; Honig, B. *J. Mol. Biol.* **1993**, *231*, 459.
(42) Antosiewicz, J.; McCammon, J. A.; Gilson, M. K. *J. Mol. Biol.* **1994**, *238*, 415.