Group additivity calculation of the standard molal thermodynamic properties of aqueous amino acids, polypeptides and unfolded proteins as a function of temperature, pressure and ionization state

J. M. Dick, D. E. Larowe, H. C. Helgeson

Biogeosciences Discussions, 2, 1515–1615, 2005
www.biogeosciences.net/bgd/2/1515/
SRef-ID: 1810-6285/bgd/2005-2-1515
European Geosciences Union

Biogeosciences Discussions is the access reviewed discussion forum of *Biogeosciences*

Group additivity calculation of the standard molal thermodynamic properties of aqueous amino acids, polypeptides and unfolded proteins as a function of temperature, pressure and ionization state

J. M. Dick, D. E. LaRowe, and H. C. Helgeson

Department of Earth and Planetary Science, University of California, Berkeley, CA 94720-4767, USA

Received: 5 September 2005 – Accepted: 18 September 2005 – Published: 11 October 2005

Correspondence to: J. M. Dick (jedick@berkeley.edu)

Abstract

Thermodynamic calculation of the chemical speciation of proteins and the limits of protein metastability affords a quantitative understanding of the biogeochemical constraints on the distribution of proteins within and among different organisms and chemical environments. These calculations depend on accurate determination of the ionization states and standard molal Gibbs free energies of proteins as a function of temperature and pressure, which are not generally available. Hence, to aid predictions of the standard molal thermodynamic properties of ionized proteins as a function of temperature and pressure, calculated values are given below of the standard molal thermodynamic properties at 25°C and 1 bar and the revised Helgeson-Kirkham-Flowers equations of state parameters of the structural groups comprising amino acids, polypeptides and unfolded proteins. Group additivity and correlation algorithms were used to calculate contributions by ionized and neutral sidechain and backbone groups to the standard molal Gibbs free energy (ΔG°), enthalpy (ΔH°), entropy (S°), isobaric heat capacity (C°_P), volume (V°) and isothermal compressibility (κ°_T) of multiple reference model compounds. Experimental values of C°_P, V° and κ°_T at high temperature were taken from the recent literature, which ensures an internally consistent revision of the thermodynamic properties and equations of state parameters of the sidechain and backbone groups of proteins, as well as organic groups. As a result, ΔG°, ΔH°, S°, C°_P, V° and κ°_T of unfolded proteins in any ionization state can be calculated up to $T \simeq 300$°C and $P \simeq 5000$ bars. In addition, the ionization states of unfolded proteins as a function of not only pH, but also temperature and pressure can be calculated by taking account of the degree of ionization of the sidechain and backbone groups present in the sequence. Calculations of this kind represent a first step in the prediction of chemical affinities of many biogeochemical reactions, as well as of the relative stabilities of proteins as a function of temperature, pressure, composition and intra- and extracellular chemical potentials of O_2, H_2, NH_3, H_2PO_4 and CO_2.

BGD

2, 1515–1615, 2005

1. Introduction

The interplay between genomes and their chemical environments, as well as the formation and relative stabilities of proteomes that support life in vastly different parts of the biosphere have received considerable attention in recent years (Elser et al., 2000; Gasch et al., 2000; Kato et al., 2004; Tyson et al., 2004; Schulze, 2005; Boonyaratanakornkit et al., 2005). A thermodynamic approach offers a quantitative understanding of consequences of biogeochemical reactions. For example, the generation of pe-pH diagrams of proteins facilitates the understanding of their relative stabilities in natural gradients of oxidation state and pH, which have been documented in biogeochemical and physiological settings (Schafer and Buettner, 2001; Ding et al., 2001). The computerized calculation of these and other types of speciation diagrams, however, depends on knowing the temperature-, pressure- and pH-dependence of the ionization states and values of the standard molal Gibbs free energies of proteins, which are generally unavailable.

The purpose of the present study is to develop the ability to calculate the mean net charge and values of the standard molal Gibbs free energy ($\overline{Z}$ and ΔG°, respectively) of unfolded proteins as a function of temperature, pressure and pH. Values of ΔG° for proteins and their constituent groups as a function of temperature (T) and pressure (P) are calculated from the revised Helgeson-Kirkham-Flowers (HKF) equation of state (Helgeson et al., 1981; Tanger and Helgeson, 1988) summarized in Appendix A.

In the revised HKF equations of state, P_r and T_r represent the reference temperature and pressure of 1 bar and 298.15 K, Ψ and Θ denote solvent parameters equal to 2600 bar and 228 K, and ϵ, ϵ_{P_r,T_r} and Y_{P_r,T_r} stand for dielectric properties of pure water (these and other symbols and abbreviations used in the text are summarized in Table 1). The species-dependent variables in the revised HKF equations of state include ΔG°_f and $S^\circ_{P_r,T_r}$ (which along with ΔH°_f are referred to as the *standard molal thermodynamic properties at 25°C and 1 bar*), and a_1, a_2, a_3, a_4, c_1, c_2 and ω, known as the *revised HKF parameters*. Values of ΔG°_f, ΔH°_f and S° of the *reference model*

compounds can be taken from the literature or are calculated from those of the corresponding interphase transfer properties. On the other hand, the values of the revised HKF parameters are obtained from experimental values of C_P°, V° and κ_T°, or estimated from correlations with the standard molal thermodynamic properties where experimental measurements are not available. For example, improved accuracy is acheived in estimating the solvation HKF parameter, ω, by retrieving values from calorimetric measurements at *high temperature* (>100°C) and from correlations with the standard molal entropy of hydration (ΔS_{hyd}°).

In the present study, we adopt the hypothesis that the properties and parameters discussed above can be calculated for neutral and ionized unfolded proteins by summing those of the constituent sidechain and backbone groups. These group contributions can be calculated from experimental data for reference model compounds including amino acids, Gly–X–Gly tripeptides and other polypeptides, and unfolded proteins. Earlier calculations of this kind were precluded by the dearth of available experimental data for amino acids (Amend and Helgeson, 1997a). Since then, high-temperature experimental data for amino acids (Hakin et al., 1998; Clarke and Tremaine, 1999; Clarke et al., 2000) and calorimetric and volumetric data for Gly–X–Gly tripeptides (Downes and Hedwig, 1995; Vogl et al., 1995; Häckel et al., 1998, 1999a,b) have been published. As a consequence, improved correlations and group additivity model compound approximations can be used to estimate the temperature and pressure dependence of the standard molal thermodynamic properties of amino acids, polypeptides and unfolded proteins.

2. Summary of group additivity equations

Group additivty algorithms have been used to calculate the thermodynamic properties of aqueous organic species both at 25°C and 1 bar (Cabani et al., 1981) and as a function of temperature and pressure (Amend and Helgeson, 1997b). Amino acids have also been the subject of group additivity analyses of the contributions by backbone

and sidechain groups (Amend and Helgeson, 1997a), as well as the contributions by the smaller organic groups that constitute the sidechain groups (Marriott et al., 1998). In addition, the additivity of standard molal volumes of ionized aqueous organic compounds has been demonstrated (Lepori and Gianni, 2000). Group additivity algorithms for calculating the thermodynamic properties of neutral proteins have been in place for at least 60 years, since Cohn and Edsall (1943) calculated group contributions to the volumes of aqueous proteins at or near 25°C. More recent group additivity calculations of C_P° of unfolded proteins were carried out for temperature increments ranging up to 125°C (Makhatadze and Privalov, 1990; Privalov and Makhatadze, 1990; Makhatadze et al., 1990). The standard molal thermodyanmic properties and revised HKF equations of state parameters of nonionized unfolded proteins and those in selected ionization states were first calculated by Amend and Helgeson (2000). However, they did not include contributions by the ionized amino acid backbone or the cysteine or tyrosine sidechain groups. A group additivity scheme that involved a more complete treatment of ionization was proposed by Kharakoz (1997) for the calculation of volume and compressibility of ionized polypeptides and proteins at 25°C and 1 bar. In the discussion that follows, the ionization-specific contributions to these and other standard molal thermodynamic properties as a function of temperature and pressure will be considered.

In the present study, the standard molal thermodynamic properties of selected reference sidechain groups are first calculated according to

$$\Xi_{[SC]} = \sum_{i=1} n_i \Xi_i \,, \tag{1}$$

where Ξ_i stands for any standard molal thermodynamic property or equation of state parameter of the ith organic group, n_i indicates the number of occurrences of the ith organic group in the reference sidechain group, and $\Xi_{[SC]}$ represents the calculated sidechain group contribution. (The definitions of these symbols and the others used in the text are summarized in Table 1.) The values of $\Xi_{[AABB]}$ and $\Xi_{[GXGBB]}$ are then calculated from those of the reference sidechain groups and those of amino acids and

Gly–X–Gly tripeptides using

$$\Xi_{AA} = \Xi_{[AABB]} + \Xi_{[SC]} \quad (2)$$

and

$$\Xi_{GXG} = \Xi_{[GXGBB]} + \Xi_{[SC]}\,, \quad (3)$$

where [AABB] represents the neutral zwitterionic amino acid backbone group, which has a formula of $C_2H_4NO_2$, [GXGBB] represents the backbone group of Gly–X–Gly tripeptides, with a formula of $C_6H_{10}N_3O_4$, and *AA*, *GXG* and [*SC*] represent, respectively, any amino acid, Gly–X–Gly tripeptide, or sidechain group (the italic text is used to generically identify these species).

Simultaneous consideration of Eqs. (1)–(3) and high-temperature calorimetric data for diols and diamines has lead to a revision of the contributions by organic groups given by Amend and Helgeson (1997b). This revision is summarized in Appendix B. The primary feature of this revision is the addition of a correction term for bifunctional compounds such as diols and diamines. In consequence, the contributions by terminal groups such as [$-CH_3$] and [$-CH_2OH$] have been updated, which permits the accurate calculation of the contributions by [Ala], [Ser] and the other reference model sidechain groups to the properties and parameters of both amino acids and Gly–X–Gly tripeptides.

Using the values of $\Xi_{[AABB]}$, Eq. (2) is used to calculate the contributions by all the neutral sidechain groups except for [Lys] and [Arg]. This statement can be validated by inspection of Fig. 1, which shows the p*K*a values of the ionization reactions of sidechain and backbone groups in amino acids as a function of temperature. At 25°C and 1 bar, the p*K*a values can be calculated from the values of ΔG_f° of amino acids taken from the literature and summarized in Table 2, but at other temperatures, they are calculated from values of ΔG° predicted from the revised HKF equations of state and the equations of state parameters of backbone and sidechain groups derived below. It is apparent from Fig. 1 that, for the ionizable amino acids other than Lys and Arg, the

BGD
2, 1515–1615, 2005

ionization of the sidechain group occurs in conditions where [AABB] is stable relative to either of its charged counterparts. Consequently, Eq. (2) can be used to retrieve the thermodynamic properties of most neutral sidechain groups. However, because the ionizations of [Lys] and [Arg] occur at pHs corresponding to stability of [AABB$^-$], their properties may be retrieved from

$$\Xi_{\mathrm{Lys}^-} = \Xi_{[\mathrm{AABB}^-]} + \Xi_{[\mathrm{Lys}]} \tag{4}$$

and

$$\Xi_{\mathrm{Arg}^-} = \Xi_{[\mathrm{AABB}^-]} + \Xi_{[\mathrm{Arg}]} \,. \tag{5}$$

In order to retrieve the thermodynamic properties of charged sidechain groups, we use charge-explicit versions of Eq. (2), which can be written as

$$\Xi_{AA^-} = \Xi_{[\mathrm{AABB}]} + \Xi_{[SC^-]} \tag{6}$$

for [Glu$^-$], [Asp$^-$], [Cys$^-$] and [Tyr$^-$],[1] and

$$\Xi_{AA^+} = \Xi_{[\mathrm{AABB}]} + \Xi_{[SC^+]} \tag{7}$$

for [His$^+$], [Lys$^+$] and [Arg$^+$].

In order to calculate the group additivity contributions to the standard molal thermodynamic properties of unfolded proteins, we first choose a hypothetical reference state corresponding to a nonionized protein, represented by UP^0. Accordingly, the calculation of $\Xi_{[\mathrm{UPBB}]}$ described in Sect. 4 takes account of the properties of unfolded proteins or polypeptides referenced to their nonionized form, together with the group additivity scheme adopted by Amend and Helgeson (1997b), which can be written as

$$\Xi_{UP^0} = \Xi_{[\mathrm{AABB}]} + (n-1)\,\Xi_{[\mathrm{PBB}]} + \sum_{i=1}^{\hat{i}} n_{[SC]_i}\Xi_{[SC]_i} \,, \tag{8}$$

[1] Standard three-letter abbreviations are used in the text to denote amino acids and sidechain groups; these or the standard one-letter abbreviations are used in the figures.

where [UPBB], and $[SC]_i$ represent the protein backbone group (C_2H_2NO) and the ith sidechain group, respectively, n stands for the number of residues (length) of the protein, $n_{[SC]_i}$ designates the number of occurrences of the ith sidechain group in the protein, and $i = 1, 2, \ldots, \hat{i}$, where $\hat{i}$ denotes the number of different kinds of sidechain groups in the protein, which is equal to 20 for proteins consisting of the common sidechain groups. It should be noted that [AABB] in Eq. (8) is used to represent the contribution by the terminal groups to the standard molal thermodynamic properties and equations of state parameters of unfolded proteins.

Although the application of the group additivity algorithm represented by Eq. (8) is restricted to neutral unfolded proteins with nonionized sidechain groups, actual proteins in aqueous solution are ionized, often with large amounts of both positively and negatively charged groups. Therefore, the comparison in Sect. 5 of experimental and predicted thermodynamic properties of unfolded proteins takes account of a simple model of ionization of unfolded proteins.

3. Regression and correlation of the revised HKF equations of state parameters of amino acids and Gly–X–Gly tripeptides

The revised HKF equations of state parameters of amino acids and Gly–X–Gly tripeptides are summarized in Tables 3 and 4 and are derived from experimental measurements available in the literature of C_P°, V° and κ_T° of amino acids and C_P° and V° of Gly–X–Gly tripeptides as a function of temperature and pressure. Because high-temperature (to ~250°C) calorimetric data are available for only three amino acids, a correlation algorithm has been developed to estimate the values of ω of the remainder of the amino acids. This correlation of ω with the entropy of hydration (ΔS_{hyd}°) of neutral amino acids closely represents the values of the ω parameter in the revised HKF equations of state. Insufficient high-temperature calorimetric data are available to generate a similar regression for charged amino acids, but a correlation between ΔS_{hyd}° and ω of alkali- and flouride- group metal ions can be used to calculate provisional estimates

of ω of charged amino acids. Calculation of ΔS°_{hyd} of both neutral and charged amino acids depends on the standard molal thermodynamic properties of both the gaseous and aqueous species. Because the values of ΔG°_f, ΔH°_f and $S^\circ_{P_r,T_r}$ of gaseous amino acids are generally not available in the literature, a group additivity algorithm is used in the present study to calculate them, in the manner described in Appendix C.

3.1. Thermodynamic properties of aqueous amino acids at high temperature

Values of ω of neutral and charged amino acids can be estimated from correlations with ΔS°_{hyd} in the following manner. Values of ω of Gly, Ala and Pro are independently regressed using Eq. (A30) and C°_P data at temperatures >200°C. A correlation between ω and ΔS°_{hyd} can be observed for these amino acids, and is depicted in Fig. 5. The corresponding equations are given in Table 6. The close correspondence apparent for His and Val in Fig. 2 between the straight lines and the experimental points supports the extrapolation of the correlation to predict ω of amino acids that have values of ΔS°_{hyd} beyond the range represented by Gly, Ala and Pro.

For metal anions and cations, similar correlations can be documented. These correlations are illustrated in Fig. 5 and are represented by Eqs. (6.2) and (6.3) in Table 6.

Insufficient experimental data at high temperatures and uncertainty of the applicability of the revised HKF equations of state (Schulte et al., 2001) lead to a high uncertainty in the regression of ω from high temperature values of V°. It can be seen in Fig. 3 that V° of Gly, Ala, Pro and Ser at high temperatures predicted using Eqs. (A30) and (A11) is consistent with the overall trend of these data, with the exception of Ala, which exhibits a decrease in V° at elevated temperatures that is larger than predicted. Nevertheless, an uncertainy in V° of even this magnitude (~10 cm^3 mol^{-1}) corresponds to a relatively small energetic uncertainty (~0.25 kcal mol^{-1} at 25°C), so the value of ω retrieved from experimental C°_P data is preferred. It should perhaps be noted that a constant error V° of 10 cm^{-3} mol^{-1} would contribute an uncertainty of of 1.2 kcal mol^{-1} in ΔG° at 300°C and 5000 kbar, which is considerably less than our estimate for the total uncertainty in

BGD
2, 1515–1615, 2005

the calculation of protein free energies (see Sect. 6).

Experimental heat capacities to 125°C of Gly–Gly–Gly, Gly–Ala–Gly and Gly–Thr–Gly have been reported (Makhatadze and Privalov, 1990; Downes and Hedwig, 1995). The results of iterative estimations of ω of these tripeptides indicate that the experimental heat capacity data can be closely represented by using values of ω given by

$$\omega_{Gly-X-Gly} \times 10^{-5} = \omega_{AA} \times 10^{-5} - 1.77\,, \tag{9}$$

where ω is in cal mol^{-1} and ω_{AA} is the value for the corresponding amino acid. Using Eq. (9), it is possible to calculate the values of ω of the remaining Gly–X–Gly tripeptides, which are given in Table 4 and which are used to calculate the values of $\Delta C^\circ_{P,\mathrm{n}}$ and $\Delta V^\circ_\mathrm{n}$ plotted in Figs. 7–8. In contrast to those of the amino acids, the values of ω of the Gly–X–Gly tripeptides are negative. If these values are representative of the behavior of amino acids and tripeptides at high temperature, measurements near the critical point of water might indicate that the standard molal thermodynamic properties of the Gly–X–Gly tripeptides diverge in the opposite direction from those of amino acids.

The regression lines in the plots of $\Delta V^\circ_\mathrm{n}$ vs. $1/(T-\Theta)$ depicted in Fig. 8 for the Gly–X–Gly tripeptides generally fall within the experimental uncertainty represented by the brackets, supporting the applicability of the revised HKF equations of state and the accuracy of the values of ω derived from C°_P data. Nevertheless, the experimentally derived values of V°_n at high temperatures are generally lower than those predicted by the equations of state.

3.2. Regression retrieval of the non-solvation parameters c_1, c_2, a_1, a_2, a_3 and a_4

The experimental or calculated values of C°_P, V° and κ°_T are combined with values of ω to give $\Delta C^\circ_{P,\mathrm{n}}$, ΔV_n and $\Delta\kappa^\circ_{T,\mathrm{n}}$. The intercepts and slopes of regression lines on plots of $\Delta C^\circ_{P,\mathrm{n}}$ vs. $1/(T-\Theta)^2$ correspond to c_1 and c_2, respectively, in the revised HKF

equations of state. On plots vs. $1/(T-\Theta)$, regression lines of ΔV°_n have intercepts of σ and slopes of ξ, while the intercepts and slopes of regression lines for $\Delta\kappa^\circ_{T,n}$ correspond to $-\left(\partial\sigma/\partial P\right)_T$ and $-\left(\partial\xi/\partial P\right)_T$, respectively. These regression plots are shown in Figs. (2)–(4) for amino acids and in Figs. (7)–(8) for Gly–X–Gly tripeptides. Values of C°_P and V° that are estimated in the present study are indicated by open diamonds. For Gly^+, values of C°_P are calculated by combining values of $\Delta C^\circ_{P,ion}$ taken from Wang et al. (1996) with those of C°_P of Gly calculated using Eq. (A30) and the revised HKF parameters taken from Table 3. The polynomial fits of the experimental scanning measurements of C°_P and V° given by Hedwig et al. have been used to generate values at 15, 25, 40, 55, 70, 80, 90 and 100°C.

The values of c_1, c_2, σ and ξ of the Gly–X–Gly tripeptides retrieved from the regression plots are given in Table 4. The values of a_2 and a_4 of the amino acids are retrieved from those of $-\left(\partial\sigma/\partial P\right)_T$ and $-\left(\partial\xi/\partial P\right)_T$ using Eqs. (A27) and (A28), respectively. The values of a_1 and a_3 are then calculated from those of σ and a_2, and ξ and a_4 using Eqs. (A24) and (A25), respectively. The values of a_1, a_2, a_3 and a_4, together with those of c_1 and c_2, are given in Table 3.

In general, experimental compressibility data for the amino acids refer to isentropic compressibilities (κ°_S). Accordingly, values of the isothermal compressibilities (κ°_T) of the amino acids are calculated from (Desnoyers and Philip, 1972; Amend and Helgeson, 1997b)

$$\kappa^\circ_T = \kappa^\circ_S + \frac{TV^\circ_{H_2O}\alpha^\circ_{H_2O}\left(2E^\circ - V^\circ_{H_2O}\alpha^\circ_{H_2O}C^\circ_P/C^\circ_{P,H_2O}\right)}{C^\circ_{P,H_2O}}, \tag{10}$$

where E° represents the isobaric exapansibility of the amino acid and $C^\circ_{H_2O}$, $V^\circ_{H_2O}$ and α_{H_2O} represent the standard molal heat capacity, volume and coefficient of isobaric thermal expansion of H_2O, respectively. Values of C°_P, V° and E° are calculated from the values of c_1, c_2, σ and ξ determined in the manner described above, and $C^\circ_{H_2O}$, $V^\circ_{H_2O}$ and α_{H_2O} are taken from SUPCRT92.

Experimental data for the charged amino acids and Gly–X–Gly tripeptides are derived from measurements of the corresponding aqueous Na- or Cl- electrolytes. Because these organic salts can be considered to be completely dissociated under the conditions of the experiments, their measured properties can be related to those of the corresponding charged amino acid or Gly–X–Gly tripeptide by subtracting the property of Na^+ or Cl^- at each experimental temperature, which can be calculated using SUPCRT92 and the parameters taken from Shock et al. (1997).

Judging from the scatter of the data points, especially those representing values from different laboratories, the representative uncertainties in the experimental data for amino acids and Gly–X–Gly tripeptides are equal to 2.5% for C_P°, 1% for V°, and (for amino acids) 2.5% for κ_T°. In general, the figures indicate that the revised HKF equations of state closely represent the bulk of the experimental calorimetric, volumetric, and compressional data for both the neutral and charged reference model compounds for protein sidechains.

It should perhaps be noted that the [His] sidechain group, and not the amino acid backbone group, is charged when HisHCl dissociates to give His^+. Consequently, ΔC_P° and V° of HisHCl reported by Jardine et al. (2001) are included in the regression calculations. However, the experimental values of ΔC_P° and V° of NaHis (Jardine et al., 2001), NaPro and ProHCl (Sorenson et al., 2003) and NaVal and ValHCl (Price et al., 2003b), along with the measurements of Wang et al. (1996), are not included in the regression calculations. Instead, they are reserved to be used to test the model compound prediction of ΔC_P° of [AABB$^+$] and [AABB$^-$] (see Sect. 5).

3.3. Model compound and correlation algorithms for non-solvation parameters

Where experimental data are insufficient to regress the revised HKF parameters of amino acids, model compound algorithms can be used to estimate values of C_P° and V° as a function of temperature. This approach is used to generate values of of C_P° of Asp^-, Glu^-, Cys, Tyr, Cys^- and Tyr^- and of V° of Cys^- and Tyr^- at 15, 25, 40 and 55 °C, as well as values of C_P° and V° of Lys^- at 25, 50 and 75°C. The close correspondence

between the experimental values of C_P° of Tyr at 15 and 25°C and those estimated using the equations in Table 5 provides support for the use of these equations for estimating C_P° at higher temperatures. The values calculated in this manner are included in the subsequent regression analysis of the revised HKF parameters, described above, and are identified in Figs. 2 and 3 by open diamonds.

Experimentally unconstrained parameters of amino acids with charged counterparts, i.e. Arg^- and Lys, can be estimated by taking account of the corresponding ionizations of other amino acids. For example, c_2 and ξ of Lys are calculated from

$$\Xi_{Lys} = \Xi_{Lys^+} + \Xi_{Arg} - \Xi_{Arg^+} . \quad (11)$$

The values of c_1 and σ of Lys are then fixed by the available values of C_P° and V° of Lys at 25 °C and 1 bar (Jolicoeur et al., 1986). However, because no experimental data can be found for C_P° and V° of Arg^-, the values of c_1, c_2, σ and ξ of Arg^- are all estimated from

$$\Xi_{Arg^-} = \Xi_{Arg} + \Xi_{Lys^-} - \Xi_{Lys} . \quad (12)$$

Finally, correlations between known values of V°, a_2 and a_4 of amino acids are used to predict values of a_2 and a_4 of Arg, Arg^-, Asp, Cys^-, Lys, Lys^- and Tyr^-. Such an approach was already used for amino acids (Amend and Helgeson, 1997a) and for other organic species (Plyasunov and Shock, 2001). Updated correlations for the amino acids are depicted in Fig. 6, and the corresponding equations are given in Table 6. The intercepts of correlation lines in this figure are charge-dependent, but the slopes in a first approximation can be considered to be independent of charge. This interpretation is consistent with the provisional correlations adopted by Amend and Helgeson (1997a), which were restricted at that time by the available experimental data for ionized amino acids.

4. Group additivity calculation of the contributions by sidechain and backbone groups to the properties and parameters of unfolded proteins

The thermodynamic properties of [AABB] can be assessed by from the group additivity regression plots of ΔH_f°, ΔS°, C_P°, V°, c_2 and ξ shown in Figs. 9 and 10. In the latter figure are also depicted regression plots of C_P°, V°, c_2 and ξ of [GXGBB]. Because fewer reference model sidechain groups or amino acids have regressed values of a_2, a_4 and ω, these parameters are correlated in the manner described below. The sidechain groups represented by [Ala], [Ser], [Glu], [Gln], [Leu] and [Phe] are chosen as the reference model sidechain groups. Respectively, they contain the terminal and branched organic groups represented by [$-CH_3$], [$-CH_2OH$], [$-COOH$], [$-CONH_2$], [$-CHCH_3-$], and [$-C_6H_5$]. In these sidechain groups, $n_{[-CH_2-]}$ is equal to 0 in [Ala] and [Ser], 1 in [Leu] and [Phe], and 2 in [Glu] and [Gln]. In selecting this representative set, Glu and Gln are chosen over their counterparts with shorter chains (Asp and Asn). Also, Leu is included instead of the shorter-chain Val and the isomer Ile, in which the branch occurs closer to the backbone. Although [Lys] contains the terminal group represented by [$-CH_2NH_2$], Lys is not included as a model compound for $\Xi_{[AABB]}$ because the sidechain group of neutral Lys is charged (see Table 2).

4.1. ΔG_f°, ΔH_f° and $S^\circ_{P_r,T_r}$

It can be deduced from Fig. 9 that the symbols representing ΔH_f° and $S^\circ_{P_r,T_r}$ of the neutral reference amino acids as a function of ΔH_f° and $S^\circ_{P_r,T_r}$ of the sidechain groups are consistent with regression lines of unit slope. It follows from Eq. (2) that the intercepts of these regression lines correspond to ΔH_f° and $S^\circ_{P_r,T_r}$ of [AABB]. Likewise, the intercepts of the regression lines in Fig. 9 for ΔH_f° and $S^\circ_{P_r,T_r}$ of the charged reference amino acids as a function of ΔH_f° and $S^\circ_{P_r,T_r}$ of the sidechain groups correspond to ΔH_f° and $S^\circ_{P_r,T_r}$ of [$AABB^+$] and [$AABB^-$]. Glu^- is excluded from consideration in the calculations of ΔH_f° and $S^\circ_{P_r,T_r}$ of [$AABB^-$] because the negative charge of Glu^- occurs on its sidechain

group, rather than on the amino acid backbone group (see Table 2). The values of ΔH_f° and $S^\circ_{P_r,T_r}$ of [AABB], [AABB$^+$] and [AABB$^-$] retrieved from Fig. 9 are given in Table 8, together with corresponding values of ΔG_f° calculated from $\Delta G_f^\circ = \Delta H_f^\circ - T\Delta S_f^\circ$ and S° of the elements.

4.2. C_P°, V° and κ_T°

The values of c_1, c_2 and ω of [AABB$^+$] are calculated from

$$\Xi_{[\mathrm{AABB}^+]} = \Xi_{[\mathrm{AABB}]} + \Xi_{\mathrm{Gly}^+} - \Xi_{\mathrm{Gly}} \,. \tag{13}$$

In contrast, the remaining equations of state parameters of [AABB$^+$], and those of [AABB$^-$], are calculated from those of Asp and Lys. These amino acids with ionizable sidechain groups are chosen to model the properties of ionized amino acid backbone groups because their sidechain groups most closely resemble the structure of the ionizable groups in the amino acid backbone group. Accordingly, the parameters of [AABB$^+$] and [AABB$^-$] in the revised HKF equations of state can be calculated from

$$\Xi_{[\mathrm{AABB}^+]} = \Xi_{[\mathrm{AABB}]} + \Xi_{\mathrm{Asp}} - \Xi_{\mathrm{Asp}^-} \tag{14}$$

and

$$\Xi_{[\mathrm{AABB}^-]} = \Xi_{[\mathrm{AABB}]} + \Xi_{\mathrm{Lys}^-} - \Xi_{\mathrm{Lys}} \,, \tag{15}$$

where Ξ stands for a_1, a_2, a_3 or a_4 of [AABB$^+$] or c_1, c_2, a_1, a_2, a_3, a_4 or ω of [AABB$^-$]. Compared to the zwitterionic [AABB], the net charge on [AABB$^+$] arises from neutralization of the negatively charged carboxylic acid group; this explains why Ξ_{Asp} and Ξ_{Asp^-} are used to calculate its parameters. Similarly, Ξ_{Lys} and Ξ_{Lys^+} are used to derive the parameters of [AABB$^-$].

4.3. Sidechain groups and ionization

The standard molal thermodynamic properties and equations of state parameters of the neutral sidechain groups except [Lys] and [Arg] are calculated by taking account

BGD
2, 1515–1615, 2005

of the properties of the amino acids given in Tables 2 and 3, and the group additivity algorithm represented by Eq. (1). Because the pKs of [Lys] and [Arg] lie within the stability field for [AABB$^-$] over most of the experimental temperature range (see Fig. 1), the values of the standard molal thermodynamic properties and equations of state parameters of [Lys] and [Arg] are calculated from Eqs. (4) and (5).

Comparative calculations summarized in the subsequent pages strongly support the hypothesis adopted above that the standard molal thermodynamic properties and equations of state parameters of neutral protein sidechain groups are equal in a first approximation to those of the neutral sidechain groups in amino acids. In contrast, equilibrium constants of the ionization reactions of sidechain and backbone groups in unfolded aqueous proteins may differ considerably from those of the corresponding amino acid ionization reactions (Nozaki and Tanford, 1967a). Accordingly, calculated and experimental values of the standard molal Gibbs free energies of ionization (ΔG°_{ion}) at 25°C and 1 bar and the pKs of deprotonation reactions of sidechains and [AABB] are listed in Table 9. The experimental values shown in this table are consistent with observations of the ionization behavior of a number of polypeptides and unfolded proteins, so in a first approximation they can be considered representative of the ionization of sidechain and terminal groups in unfolded proteins. This observation is further supported by independent evidence that specific types of ionizable sidechain groups in unfolded proteins tend to have unique equilibrium constants, while those in folded proteins are much more variable (Tollinger et al., 2003). It can be seen in Table 9 that the pKs at 25°C and 1 bar of the ionization reactions of [Asp], [Glu], [His], [Cys], and [Tyr] in amino acids and proteins differ from those in unfolded proteins by 0.50 or less. However, the pKs of [AABB] in amino acids and [AABB] in unfolded proteins differ by more than a log unit.

4.4. ΔG°_f, ΔH°_f and S° of [UPBB]

Amend and Helgeson (2000) estimated values of ΔH° and S° from those of [UPBB]

and

$$\Xi_{[\mathrm{UPBB}]} + \Xi_{[\mathrm{Gly}]} = \Xi_{\mathrm{diketopiperazine}}/2\,, \tag{16}$$

using the properties of diketopiperazine given by Shock (1992). Because ΔH°_f and S° estimated in this manner are used to calculate ΔG°_f of [UPBB], the uncertainty in such a characterization contributes considerably in the calculation of Gibbs free energy at elevated temperatures and pressures. Diketopiperazine is a circular dipeptide that might be expected to behave differently than polypeptide chains or unfolded proteins. Therefore, linear dipeptides for which ΔH° and S° are known are included in the present analysis. The properties of [UPBB] can be estimated from the model compound and group additivity statements represented by

$$\Xi_{[\mathrm{UPBB}]} = \Xi_{[\mathrm{AABB}]} + \Xi_{\mathrm{GlyGly}} - \left(\Xi_{\mathrm{Gly}} + \Xi_{\mathrm{Gly}}\right)\,, \tag{17}$$

$$\Xi_{[\mathrm{UPBB}]} = \Xi_{[\mathrm{AABB}]} + \Xi_{\mathrm{AlaGly}} - \left(\Xi_{\mathrm{Ala}} + \Xi_{\mathrm{Gly}}\right)\,, \tag{18}$$

$$\Xi_{[\mathrm{UPBB}]} = \Xi_{[\mathrm{AABB}]} + \Xi_{\mathrm{LeuGly}} - \left(\Xi_{\mathrm{Leu}} + \Xi_{\mathrm{Gly}}\right)\,, \tag{19}$$

Values of ΔG°_f, ΔH°_f and S° of the dipeptides can be found in Shock (1992). Unlike the present study, the entropies of the elements used by Shock (1992) are from Wagman et al. (1982). Nevertheless, differences between these and the values used in the present study lead to variations in the calculation of ΔG°_f that are less than an order of magnitude smaller than the uncertainties associated with the model compound analysis. These four equations yield $\Delta H^\circ_{f,[\mathrm{UPBB}]}$= −44.08, −48.04, −43.49 and −45.29 kcal mol^{-1} and $S^\circ_{[\mathrm{UPBB}]}$= 9.44, −5.44, 4.88 and −2.40 cal mol^{-1} K^{-1}, respectively. The averages of these values are given in Table 8, as are the uncertainties involved, which are estimated from the differences between the values calculated from Eqs. (16)–(19). The estimated uncertainty in ΔG°_f is smaller than that in ΔH°_f, which is consistent with the scatter of the values of ΔG°_f (−22.64, −21.84, −20.34, −19.99 kcal mol^{-1}) that can

be calculated from the above equations. However, to ensure consistency, the value of ΔG_f° given in Table 8 is not the average of the afore-mentioned values, but instead is calculated from $\Delta H^\circ_{f,[\mathrm{UPBB}]}$, $S^\circ_{[\mathrm{UPBB}]}$, and S° of the elements.

The refinement of ΔG_f°, ΔH_f° and S° is possible by considering equilibrium constants and enthalpies of dissolution of crystalline proteins, or by taking into account the measured enthalpies and Gibbs free energies of polypeptides, or preferably, entire proteins. Such a characterization, however, depends on the quantification of the activity coefficients of folded proteins (because the dissolution experiments take place in relatively high ionic strength solutions), as well as the measurement or prediction of the properties of protein unfolding (because the dissolution experiments yield folded proteins in solution). However, this calculation outside the scope of the current investigation.

4.5. C_P° and V° of [UPBB] and [PPBB]

The reference values of C_P° and V° of [UPBB] and [PPBB] can be calculated from the corresponding experimental measurements of nonionized unfolded proteins (Ξ_{UP^0}) using the group additivity algorithm represented by Eq. (8). The reference proteins chosen for this calculation include the eight proteins listed in Table 11 for which Privalov et al. have measured C_P° and/or V° from 25 to 100°C. Their reported values at 125°C, which are actually derived using group additivity, are included in Figs. 11 and 12, and generally are consistent with the trend of the regression lines. It has been shown in calorimetric experiments that the measured C_P° and V° of protein ionization reactions are essentially compensated by corresponding reactions with the denaturing buffer. In addition, experimental values of the standard molal heat capacities as a function of temperature of four unfolded proteins in the reference set were found to approach common values despite variations in solution pH from 2 to 6 (Privalov et al., 1989), supporting the notion that there is a compensating effect and that the measured values reflect the properties of nonionized proteins. Consequently, we consider the data from calorimetric measurements of V° and C_P° of unfolded proteins in a denaturing buffer to represent reference values for nonionized unfolded proteins.

BGD
2, 1515–1615, 2005

Even in a strong denaturant such as 8*M* GuHCl, unfolded proteins in aqueous solution may not adopt completely unstructured conformations (Georgescu et al., 2001). The degree of formation of secondary structure in unfolded proteins also depends on other solution conditions including temperature and pH (Guzman-Casado et al., 2003). The formation of secondary structure, and decrease of exposure to the solvent, may result in deviations of the standard molal thermodynamic properties from the values predicted by group additivity. Such a deviation has been noted for heat capacity by Georgescu et al. (2001) and Guzman-Casado et al. (2003), among others. In contrast to proteins, polypeptides have more random conformations in aqueous solutions. Accordingly, pentapeptides are used as the reference model compounds to calculate the values of C_P° and V° of [PPBB]. Because the values of C_P° and V° of [UPBB] and [PPBB] discussed below are derived from experimental properties of unfolded proteins and pentapeptides, respectively, they implicitly reflect the consequences of the actual conformation of the molecules, such as the effects of the ϕ and ψ angles on thermodynamic properties. These angles are not fixed in unfolded proteins, and they may be especially variable for small sidechain groups such as [Gly]. Consequently, a higher uncertainty would be expected with the calculation of the properties and parameters of [PPBB] from polyglycine peptides. In contrast with other studies, therefore, polyglycines are not used here as model compounds for the properties of [PPBB].

Regression lines for c_1, c_2, σ and ξ of [UPBB] and [PPBB] are shown in Figs. 11 and 12. The values of c_2 represented by the slopes of the lines are common to both [UPBB] and [PPBB]. In contrast, the intercepts of the correlation lines, representing values of c_1, are different for [UPBB] and [PPBB]. Note that, within the scatter of the experimental data points represented by the symbols, the contributions by the sidechain groups derived from amino acids and those derived from experimental data for the Gly–X–Gly tripeptides (see above) results in the same values of $\Delta C^\circ_{P,\mathrm{n}}$ and $\Delta V^\circ_\mathrm{n}$ of [UPBB] and [PPBB]. As in the case of c_2, the values of σ and ξ are respectively identical for [UPBB] and [PPBB]. The latter values are combined with estimated values of a_2 and a_4 derived in the manner described below, to calculate the values of a_1 and a_3 of [UPBB].

4.6. κ_T°, a_2 and a_4 of [UPBB]

The isothermal compressibility (κ_T°) at 25°C and 1 bar of [UPBB] can be calculated from the measured compressibilities of poly-*d*,*l*-alanine and poly-glutamic acid and κ_T° of [Ala] and [Glu] using the relations

$$\kappa^\circ_{T,[\mathrm{UPBB}]} = \kappa^\circ_{T,\mathrm{poly-}d,l\mathrm{-alanine}} - \kappa^\circ_{T,[\mathrm{Ala}]} \tag{20}$$

and

$$\kappa^\circ_{T,[\mathrm{UPBB}]} = \kappa^\circ_{T,\mathrm{poly-glutamic\,acid}} - \kappa^\circ_{T,[\mathrm{Glu}]}\,, \tag{21}$$

which yield values of the isothermal compressibility of [UPBB] of -13.40×10^{-4} and -13.98×10^{-4} cm^3 bar^{-1} mol^{-1}, respectively. In Eqs. (20) and (21), the isothermal compressibilities of the polypeptides are calculated using Eq. (10) and the values of κ_S° measured by Kharakoz (1997). The values of C_P°, V° and E° of the polypeptides in Eq. (10) are estimated using the revised HKF equation of state and the properties and parameters of the sidechain and backbone groups taken from Table 10, along with the values of σ and ξ of [UPBB] calculated above.

The group additivity algorithm used by Amend and Helgeson (2000) to calculate values of a_2 and a_4 of [UPBB] is given by

$$\Xi_{[\mathrm{UPBB}]} = \Xi_{[\mathrm{AABB}]} + 2\Xi_{[-\mathrm{CH_2}]} - \Xi_{[-\mathrm{CH_2OH}]} - \Xi_{[-\mathrm{CH_3}]}\,, \tag{22}$$

where Ξ represents a_2 or a_4 of the subscripted group. Using this equation and the group contributions given in Table 10, results in $a_{2,[\mathrm{UPBB}]} \times 10^{-2} = -3.75$ cal mol^{-1} and $a_{4,[\mathrm{UPBB}]} \times 10^{-4} = -1.53$ cal mol^{-1} K. In the revised HKF equations of state, these values, when taken together with the value of $\omega_{[\mathrm{UPBB}]}$ give $\kappa^\circ_{T,[\mathrm{UPBB}]} = -37.15 \times 10^{-4}$ cm^3 bar^{-1} mol^{-1}. This is considerably lower than the value obtained above for the polypeptides. Accordingly, to minimize uncertainty in the prediction of κ_T° at 25°C and 1 bar, the value of a_2 of [UPBB] adopted in the present study is calculated using Eq. (22).

BGD
2, 1515–1615, 2005

In contrast, the value of a_4 of [UPBB] is selected so that the additivity calculation of $\kappa^\circ_{T,[UPBB]}$ at 25° and 1 bar is equal to the average of the results of Eqs. (20) and (21).

5. Comparision of calculated and experimental heat capacities, ionization constants and charges of amino acids and unfolded proteins

Experimental data not included in the regression calculations can be used to test the accuracy of the predictions of thermodynamic properties of ionized amino acids and proteins. Comparative calculations of this kind are presented below for the heat capacities and ionization constants of amino acids at high temperature, for the ionization states and heat capacities of unfolded proteins as a function of pH and temperature, and for the enthalpies of solution of proteins at 25°C and 1 bar.

5.1. C°_P and pK of amino acid ionization at high temperature

Experimental values of $\Delta C^\circ_{P,r}$ of the ionization reaction represented by [AABB$^+$] ⇌ [AABB] + H$^+$ are available to ≥120°C for Gly, Ala, Pro and Val. These are shown in Fig. 13 along with the values of $\Delta C^\circ_{P,r}$ as a function of temperature calculated from the group contributions to $\Xi_{[AABB]}$ and $\Xi_{[AABB^+]}$ given in Tables 8 and 10. The calculated values of $\Delta C^\circ_{P,r}$ shown in Fig. 13, track closely with the experimental data for Gly ionization, from which the group contributions were derived. At low temperatures, they are also consistent with experimental values of $\Delta C^\circ_{P,r}$ of the other amino acids shown in the figure (Gly, Pro and Val), but at temperatures >80°C, the experimental values diverge, by as much as 30 kcal mol^{-1}. These differences in the ionization properties of the amino acid backbone group may arise from interactions with the sidechain group, as well as from differences in experimental conditions, possibly leading to degradation of the amino acids at high temperatures. The latter explanation seems reasonable, as the $\Delta C^\circ_{P,r}$ data below 80°C seem to indicate a unique backbone contribution, regardless of the sidechain group present. The present calculations closely represent these

BGD
2, 1515–1615, 2005

low-temperature measurements, and split the difference between the measurements at high temperatures. The coincidence of predicted and measured values of pK at even higher temperatures, to 250 °C (lower panel of Fig. 13) further supports the applicability of the group contributions at hydrothermal conditions.

It can be deduced from Fig. 13 that *n*-carboxylic acids do not appear to be suitable model compounds for the estimation of C_P° of [AABB$^+$]. This might be because the ionization of [AABB] involves the destructrion of a zwitterion. On the other hand, propanoic and butanoic acids remain appropriate model compounds for [Asp$^+$] and [Glu$^+$], because no zwitterion is destroyed in their ionization.

5.2. Mean net charge of unfolded proteins as a function of temperature and pH

Assuming no interaction among the ionization groups of unfolded proteins, the degree of formation of the ith ionized group (α_i) can be calculated from

$$\alpha_i = \frac{1}{1 + 10^{Z_i(\mathrm{pH}-\mathrm{p}K)}}, \tag{23}$$

where Z_i is the formal charge on the ith ionized group. pK is equal to the negative logarithm of the deprotonation reaction involving any of the charged groups, and can be calculated using the revised HKF equations of state and the group contributions given in Tables 8 and 10. Because there may be many charged groups in a protein molecule, the mean net charge ($\overline{Z}$) is defined; it reflects the contributions by all the ionizable groups in the protein molecules to the total charge (Edsall and Wyman, 1958). It follows that $\overline{Z}$ of any unfolded protein can be calculated from

$$\overline{Z}_{UP} = \sum_i n_i \alpha_i Z_i \,. \tag{24}$$

Eq. (24) permits the calculation of protein charge as a function of not only pH, but also temperature and pressure. The results of this calculation can be compared with the experimental values at 25°C represented in Fig. 14. The calculations reveal that,

BGD
2, 1515–1615, 2005

compared with the measurements of $\overline{Z}$ in 6 M guanidinium hydrochloride (GuHCl, a common laboratory denaturant), the calculations underestimate the $\overline{Z}$ at pHs $\gtrsim$ 8. This is most likely a consequence of disulfide bonds that are retained in the unfolded proteins under oxidizing experimental conditions. Although proteins adopt an unfolded conformation in aqueous solutions of 6 M guanidinium hydrochloride (GuHCl), any disulfide bonds that are present in the native state may remain intact, depending on the oxidation state of the system. Support for this hypothesis has been provided by spectroscopic measurements on RNAS1_BOVIN and LYC_CHICK in 6 M GuHCl, with or without a reducing agent (dithiothreitol) in the solution (Hu and Zou, 1992, 1993). Each of these proteins contain 8 Cys residues that form 4 disulfide bonds in the native protein (Neumann et al., 1964). If the experimental conditions under which the titration data were obtained were sufficiently oxidizing (which is the most likely case, since no attempt was made to provide reducing conditions in the measurements), there is no stability field for Cys^-, relative to Cys, no matter how high the pH. Under these conditions, therefore, accurate calculation of the ionization state of disulfide-containing unfolded proteins can be accomplished by removing from consideration the ionization of the Cys sidechains.

Experimental measurements of the temperature dependence of charge of unfolded proteins is minimal, but Cohn and Edsall's (1943) titrations of crystalline horse serum albumin over the temperature range 5–25°C indicate that protein net charge decreases considerably with increasing temperature at high pH, but varies much less at pHs $\lesssim$6. This observation is in qualitative agreement with the results shown in Fig. 14.

5.3. C_P° and G° of unfolded proteins as a function of temperature, pH, and oxidation state of Cys sidechain groups

The ionization of groups in proteins contributes not only to $\overline{Z}$, but also to the standard molal thermodynamic properties of unfolded proteins. Equation (24) can be rewritten to account for the ionization contribution to the net protein property, represented by

$\Delta\Xi_{ion}$:

$$\Delta\Xi_{ion} = \sum_i n_i \alpha_i \Delta\Xi_{ion,i} \,, \tag{25}$$

where $\Delta\Xi_{ion,i,j}$ represents $\Delta C^\circ_{P,ion}$ or ΔG°_{ion} of the ionization reaction of the ith charged sidechain or backbone group. Any standard molal thermodynamic property of an ionized unfolded protein ($\Xi_{UP^{\bar{z}_{UP}}}$) can then be calculated by summing Eqs. (8) and (25) to give

$$\Xi_{UP^{\bar{z}_{UP}}} = \Xi_{UP^0} + \sum_i n_i \alpha_i \Delta\Xi_{ion,i} \,. \tag{26}$$

Values of ΔG° calculated using Eq. (26) as a function of ionization state and temperature are shown in Fig. 14. For 25°C, the results of the same calculation, but with the exclusion of [Cys] ionization, are shown by the dashed curves. It is evident that the ionization contributions by even a single type of reside can contribute substatially to the thermodynamic properties of proteins. In the future, the contributions by the disulfide group should also be quantified.

6. Estimated uncertainties in the calculation of the standard molal thermodynamic properties of amino acids, sidechain groups and unfolded proteins as a function of temperature and pressure

These uncertainties are representative of amino acids, sidechain groups, and unfolded protein backbone groups. In some cases, discussed above, the trend of experimental data suggest that there is greater uncertainty in the calculation of these properties at high temperature. The following discussion of uncertainty is limited to amino acids, unfolded proteins and their constituent groups. However, a similar interpretation may be made for Gly–X–Gly tripeptides, and polypeptides. It can be seen by comparing the number of significant figures given for the properties and parameters in Tables 8,

BGD
2, 1515–1615, 2005

10 and 3 that the precisional uncertainty is generally at least an order of magnitude smaller than the estimated uncertainties in *AA* and [*SC*] described below.

6.1. Amino acids

The uncertainties, as well as the values of ΔG_f°, ΔH_f° and $S^\circ_{P_r,T_r}$ of the aqueous amino acids given in Table 2 are taken from Amend and Helgeson (1997a), who assigned the uncertainties on the basis of experimental error apparent in the measurements of the solubilities and enthalpies of solution of amino acids. The uncertainties in the non-solvation revised HKF regression parameters are estimated from the scatter of the symbols shown in the regression plots in Figs. 2–4. The consequent uncertainties of c_1 and c_2 are given in Table 3. The uncertainties of the other non-solvation regression parameters are ± 1.2 cm^3 mol^{-1} for σ, $\pm 1.5 \times 10^2$ cm^3 mol^{-1} K^{-1} for ξ, $\pm 3 \times 10^{-4}$ cm^3 bar^{-1} mol^{-1} for $\left(\partial\sigma/\partial P\right)_T$ and $\pm 5 \times 10^{-2}$ cm^3 bar^{-1} mol^{-1} for $\left(\partial\xi/\partial P\right)_T$. From these values, the uncertainties in a_1, a_2, a_3 and a_4 can be calculated by taking the absolute value of the revised HKF equations of state expression containing known parameters. For example, δa_3 (the estimated uncertainty in a_3) can be calculated using a rearrangement of Eq. (A25):

$$\delta a_3 = \pm \left| \delta\xi - \frac{\delta a_4}{\Psi + P} \right| , \tag{27}$$

where $\delta\xi$ and δa_4 represent the uncertainties in the corresponding parameters.

The uncertainties in ΔG_f°, ΔH_f° and S° of amino acids may in fact be higher than those we report here. These values from Amend and Helgeson (1997a) and Shock (1992) for Gly, Ala and Leu differ by as much as 1.93 kcal mol^{-1} (ΔH_f°) and 1.45 cal mol^{-1} K^{-1} ($S^\circ_{P_r,T_r}$).

6.2. Sidechain groups

The total uncertainties in the group additivity calculation of ΔH°_f and $S^\circ_{P_r,T_r}$ and the revised HKF parameters of [AABB] can be inferred from the scatter of data points in the regression plots shown in Figs. 9 and 10. To estimate the contributions by the sidechain groups to the total uncertainties in these properties and parameters, we adopt in a first approximation the hypothesis that [AABB] and [*SC*] contribute equally to the total group additivity uncertainty. Therefore, we can write

$$\delta\Xi_{[SC]} = \delta\Xi_{\text{group additivity}}/2\,, \tag{28}$$

where $\delta\Xi_{\text{group additivity}}$ stands for the total group additivity uncertainty, which is represented in Figs. 9 and 10 by the length of the error bars. Eq. (28) is used to estimate the uncertainties of ΔH°_f, $S^\circ_{P_r,T_r}$, ΔG°_f, c_1, c_2, a_1 and a_3 of [*SC*] listed in Tables 8 and 10. Although the uncertainties in the values of a_2, a_4 and ω of [*SC*] can not be estimated in this manner, we take provisional values that are equal to twice the corresponding uncertainties of *AA*.

There may be compensating uncertainties in ΔH°_f and $S^\circ_{P_r,T_r}$ of the species and the reference model compounds used in the calculations that lead to an uncertainty in ΔG°_f. Accordingly, although the value of ΔG°_f of [AABB] is calculated from the values of ΔH°_f and $S^\circ_{P_r,T_r}$ derived from the regression plots shown in Fig. 9, a group additivity regression analysis of the calculated values of ΔG°_f – similar to that shown for ΔH° and S° in Fig. 9 – is needed to compute a regression uncertainty in this property. The result of this analysis indicates that the uncertainty in ΔG°_f is smaller than the uncertainty in ΔH°_f.

6.3. Protein backbone groups

The uncertainties in ΔG°_f, ΔH°_f and $S^\circ_{P_r,T_r}$ of [UPBB] are discussed in Sect. 4.4. Consideration of the scatter of points apparent in the amino acid sidechain plots of Figs. 11 and 12 indicates the uncertainties in c_1, c_2, σ and ξ of the group additivity analysis,

which can be expressed as

$$\delta\Xi_{\text{group additivity}} = \delta\Xi_{[SC]} + \delta\Xi_{[\text{UPBB}]} \,. \tag{29}$$

For the group additivity analysis, [*SC*] and [UPBB], respectively, the uncertainties in c_1 and c_2 are (3.0, 1.8, 1.2) cal mol^{-1} and (2.4, 0.8, 1.6) cal mol^{-1} K^{-1}. Likewise, the uncertainties in σ and ξ are (2.0, 0.8, 1.2) cm^3 mol^{-1}, and in ξ are (240, 40, 200) cm^3 mol^{-1} K^{-1} (for [*SC*], the uncertainties in σ and ξ are roughly equal to those in a_1 and a_3, respectively, after conversion from caloric to volumetric units). Because a_2 and a_4 of [UPBB] are calculated from Eq. (22) their uncertainties can be considered comparable to those of *AA*. Finally, the uncertainties in a_1 and a_3 of [UPBB] are calculated using Eq. (27) and its counterpart for σ. The uncertainty in ω of [UPBB] is probably of the order of that of [*SC*].

6.4. Uncertainties of standard molal thermodynamic properties as a function of temperature and pressure

Propagated uncertainties in ΔG°, ΔH°, S°, C_P°, V° and κ_T° calculated as a function of temperature and pressure are given in Table 12. Because the uncertainties are expressed as average errors, the propgated uncertainties are computed by summing the absolute value of each term in the corresponding equation. (If the uncertainties were expressed as standard deviations, this calculation would be performed by taking the square root of the sum of squares of each term.) For example, the uncertainty propagated to V° can be calculated using a rearrangement of Eq. (A31):

$$\delta V^\circ = |\delta\sigma| + \left|\frac{\delta\xi}{T-\Theta}\right| + |\delta\omega Q| \,, \tag{30}$$

where δV°, $\delta\sigma$, $\delta\xi$ and $\delta\omega$ represent the uncertainties in the values of V°, σ, ξ and ω. Because the absolute value of each term is summed, regardless of the sign present in the original equation, propagated uncertainties represent estimated maximum uncertainties, and the actual uncertainty will most likely be much smaller. The propagated

BGD
2, 1515–1615, 2005

uncertainties in ΔG° and C_P° are depicted in Fig. 16 over the temperature range 25–250°C at P_{SAT} as percentages of the average value of the 20 neutral amino acids. It can be concluded that, although the equation of state parameters, particularly ω, govern the uncertainty apparent in the calculation of ΔC_P° at high and low temperatures, the largest source of error in the calculation of ΔG° as a function of temperature is not uncertainty in the equations of state parameters, but instead the uncertainty in ΔG_f° and $S^\circ_{P_r,T_r}$.

It might be noted that the estimated uncertainties in ΔH°, S°, C_P°, V° and κ_T° actually decrease with increasing pressure. For C_P°, V° and κ_T°, this is a consequence of the smaller values of uncertainty contributions by the non-solvation terms, as well as the smaller values of X, Q and N – and therefore of the solvation terms – at elevated pressures. In addition, the estimated uncertainties in the calculation of ΔG° of amino acids and sidechain and protein backbone groups shown in Table 12 increase only moderately over this range of temperature and pressure. However, above 300°C and 5000 bar, the predictions of the derivative properties – particularly C_P° – are much more uncertain; in general, then, these might be considered upper limits for the applicability of the group additivity and equation of state calculations in the current context.

6.5. Comparison with other estimates of uncertainty

Our estimated uncertainty of $V^\circ_{P_r,T_r}$ of [*SC*] (±7.2 cm^3 mol^{-1}) is comparable to the largest differences between experimental and calculated volumes observed by Hnědkovský and Cibulka (2004), who used group additivity to model the volumes of benzene and aliphatic hydroxyl derivatives including carboxylic acids and alcohols. However, it should perhaps be emphasized that these high deviations are not representative of group additivity in general. For example, other group additivity schemes have been shown to reproduce the volumes of the reference compounds with accuracies approaching ±1 cm^3 mol^{-1} or less (Lepori and Gianni, 2000; Hnědkovský and Cibulka, 2004). Also in the pursuit of predicting the thermodynamic properties of organic so-

lutes, Cabani et al. (1981) developed a group additivity scheme that includes a constant contribution to each species. Their estimated uncertainties in C°_P and V° at 25°C and 1 bar are ±3.2 cal mol^{-1} and ±0.70 cm^3 mol^{-1}, respectively. We expect that the actual uncertainties in the current calculations to be of this order.

The combined uncertainties in the calculation of ΔG° of each protein residue (i.e. [*SC*] and [UPBB]) are considerably greater than the target of ±0.1 kcal mol^{-1} recommended by Dill (1997) for group additivity calculations of the energetics of conformational interactions of proteins such as protein unfolding and other non-covalent structural changes. Nevertheless, the current method affords a close approximation of the standard molal thermodynamic properties of proteins of differing covalent (or primary) structure, which can be used a reference point for beginning more refined calculations of the properties of folded proteins and their natural complexes.

7. Discussion

7.1. Comparison with other estimates of group additivity uncertainty

Although there are no experimental values of ΔG° of ionized unfolded proteins to compare with the predictions, experimental values of C°_P as a function of temperature and pH are available for at least three proteins not included in the regression calculations (Laderman et al., 1993; Guzman-Casado et al., 2003). A comparison between the experimental values and the predicted values of C°_P of nonionized proteins as a function of temperature and of C°_P of ionized proteins as a function of temperature and pH is depicted in Fig. 15. It can be seen in this figure that predicted values for nonionized unfolded proteins are in general higher than the measured values. (However, it should perhaps be noted that our predictions for the nonioinzed unfolded proteins are comparable to values of C°_P of unfolded proteins calculated by Guzman-Casado et al. (2003), using the group contributions given by Makhatadze and Privalov (1990).) In contrast, the predicted values of C°_P of ionized proteins at experimental pHs much more

closely represent the experimental data. Such a consideration of ionization contributions to the standard molal thermodynamic properties of unfolded proteins may lead to an altered interpretation of the differences between experimental and additive heat capacities, which have been commonly attributed to the formation of residual structure (Georgescu et al., 2001; Guzman-Casado et al., 2003).

7.2. Departures from ideality

Activity coefficients of ionizable groups in proteins would be expected to depart significantly from unity in high ionic strength solutions such as 6 *M* GuHCl. Also, interactions between the ionized groups in unfolded proteins might contribute to their thermodynamic properties (Whitten and García-Moreno E., 2000). A number of models have been proposed that take these considerations into account, such as the Linderstrom-Lang model (Nozaki and Tanford, 1967b). A limitation of such models is their applicability at present to only ambient conditions, but in the future it may be possible to extend them to higher temperatures and pressures.

It should be apparent from Fig. 10 that the both amino acids and Gly–X–Gly tripeptides are well suited to group additivity analysis, which was not the conclusion reached by Hedwig and Hinz (2003). They, and Hakin and Hedwig (2001a), hypothesized that the zwitterionic backbone interacts significantly with the sidechain groups of amino acids. However, the terminal groups of tripeptides, and of unfolded proteins themselves, behave as zwitterions over much of the pH range, and it is probably the case that there are significant interactions between them and the sidechain groups.

Either amino acids or Gly–X–Gly can serve as model compounds for the sidechains; indeed, some of the Gly–X–Gly data are used to estimate the properties of, for example, [Cys] where amino acid data is lacking. Because the data available for Gly–X–Gly include ΔC_P° and V°, but do not include ΔG_f°, ΔH_f° or S° or κ_T°, the experimental data available for amino acids seems the most appropriate source of primary experimental data for this purpose. In summary, the experiments on Gly–X–Gly tripeptides represents a complementary (and indispensible) source of data, and seem to indicate that

the properties of sidechain groups in amino acids and polypeptides are comparable.

7.3. Oxidation-reduction buffers and the stability of amino acids in hydrothermal experiments

The calorimetric measurements might be able to take advantage of oxidation-reduction buffered systems, which have been shown to slow or eliminate the decomposition of amino acids at 200°C and 50 bar (Andersson and Holm, 2000). The preservation of alanine at high temperatures may also favored by ionic strengths of KCl up to 2 *M* (Li et al., 2002). Thermodynamic predictions of the oxidation-reduction stabilization of amino acids were reported by Amend and Helgeson (1997b), but the prediction of stabilities of amino acids relative to their decarboxylation or other decomposition products will require more detailed knowledge of the activity coefficients of all the species involved.

7.4. Electrostatic interactions in unfolded proteins

Several experiments (Tan et al., 1995) and models relying on electrostatic calculations (Elcock, 1999; Kundrotas and Karshikoff, 2002) more sophisticated than those presented here indicate that the individual values of p*K*s of ionizable groups in unfolded proteins may differ substantial from those of model compounds. Although it may be less accurate than such calculations at 25°C, the present method permits the estimation of the titration curves of unfolded proteins at other temperatures and pressures. We are not aware of other titration models with this provision. (Although, however, the results of our calculations as a function of temperature might be able to provide the null, or baseline p*K* values which form the starting point of the electrostatic simulations.)

7.5. Extension of the group additivity database: [Cys−] and [H_2O]

Although recent experimental work on the properties of thiols and polysulfides has appeared (Schulte and Rogers, 2004; Plyasunova et al., 2005), we don't yet have many

measurements on disulfides, at least not enough to easily retrieve their equations of state parameters. Because disulfides are present not only in proteins, but also as a part of the glutathione oxidation-reduction systems in cells, the experimental and theoretical determination of their thermodynamic properties may play an important role in understanding cellular metabolism. This will aid not only the prediction of the thermodynamic properties of proteins, but also the interpretation oxidation-reduction conditions responsible for the stability of disulfide-bonded proteins prevalent different biogeochemical environments, including the intracellular environments of archaeal organisms (Mallick et al., 2002).

8. Conclusions

The group contributions to the standard molal thermodynamic properties at 25°C and 1 bar and the revised HKF equations of state parameters of neutral and charged aqueous sidechain and backbone groups generated above permit calculation of the thermodynamic properties of unfolded proteins with any amino acid sequence in any ionization state as a function of temperature and pressure. It has been demonstrated above that the availability of recent high- and low-temperature calorimetric and experimental data, combined with a group additivity approach using multiple reference model compounds, facilitates calculation with unprecedented accuracy of the the standard molal thermodynamic properties of ionized amino acids, Gly–X–Gly tripeptides, polypeptides and unfolded proteins.

The standard molal thermodynamic properties generated in the present study can be combined with computer codes that perform Gibbs free energy minimization for specific bulk composition. These codes include GEMS-PSI (Karpov et al., 2001; Kulik, 2004), HCh (Shvarov and Bastrakov, 1999) and visualization software capable of generating two- and three-dimensional projections of metastable equilibrium phase relations in compositional hyperspace (Connolly, 1990). Such software facilitates a global interpretation of the consequences of chemical reactions among proteins in biogeochem-

ical systems. The consequences of evolutionary adaption of microbial communities to changes in these systems include the differences in protein composition in thermophilic and mesophilic organisms noted by Fukuchi and Nishikawa (2001) and Kreil and Ouzounis (2001). Even on cellular length and time scales, there is mounting evidence (Conour et al., 2004) that changes of protein composition during cell growth may be linked to oxidation-reduction gradients between different subcellular units (Al-Habori, 1995). The physical chemical basis for such proteomic variation is amenable to thermodynamic assessment of the biogeochemical constraints on protein speciation using the properties, parameters and equations discussed above.

A. Summary of the revised HKF equations of state

A.1. Thermodynamic conventions

According to convention, the standard molal free energy of the hydrogen ion in aqueous solution is zero at all temperatures and pressures. For other charged aqueous, the standard molal thermodynamic properties are given by

$$\Xi = \Xi^{abs} - Z\Xi^{abs}_{H^+}, \quad \text{(A1)}$$

where Ξ and Ξ^{abs} stand for any given conventional and absolute standard molal property, or equation of state coefficients of the aqueous species of interest, $\Xi^{abs}_{H^+}$ denotes the corresponding absolute standard molal property of the hydrogen ion, and Z represents the charge of the aqueous species of interest. It can be seen that $\Xi_{H^+} = 0$ at all temperatures and pressures. This convention establishes the scale by which the standard molal thermodynamic properties of all other species are reported. If needed, calculations that take account of the biochemical standard state (defined at pH=7 and usually at 37°C) can be referenced to the standard state adopted here by taking ac-

count of the appropriate equations. [2]

The standard molal thermodynamic properties of aqueous species other than H_2O are for a hypothetical one molal solution referenced to infinite dilution at 25°C and 1 bar. At other temperatures and pressures, apparent standard molal thermodynamic enthalpies (ΔH°) and Gibbs free energies (ΔG°) are given by

$$\Delta H^\circ = \Delta H^\circ_f + \left(H^\circ_{P,T} - H_{P_r,T_r}\right) \tag{A2}$$

and

$$\Delta G^\circ = \Delta G^\circ_f + \left(G^\circ_{P,T} - G_{P_r,T_r}\right) , \tag{A3}$$

where ΔH°_f and ΔG°_f represent the standard molal enthalpy and Gibbs free energy of formation of the species from the elements at the reference temperature (T_r) and pressure (P_r) of 25°C and 1 bar, respectively, and $H_{P,T} - H_{P_r,T_r}$ and $G_{P,T} - G_{P_r,T_r}$ denote the differences between the standard molal enthalpy and Gibbs free energy, respectively, at the temperature (T) and pressure (P) of interest, and those at T_r and P_r.

The values of ΔG°_f, ΔH°_f and $S^\circ_{P_r,T_r}$ are related by

$$\Delta G^\circ_f = \Delta H^\circ_f - T_r \left(S^\circ_{P_r,T_r} - S^\circ_{P_r,T_r,\,\text{elements}}\right) , \tag{A4}$$

where $S^\circ_{P_r,T_r,\,\text{elements}}$ represents the sum of the standard molal entropies of the elements in the species of interest at 25°C and 1 bar. The values of the entropies of the elements at 25°C and 1 bar are taken Cox et al. (1989).

Where needed, conversions between volumetric and energetic units are calculated from the relation 1 cal = 41.84 cm^3 bar.

[2] See, for example, LaRowe, D. E. and Helgeson, H. C.: Biomolecules in hydrothermal systems: Calculation of the standard molal thermodynamic properties of nucleic-acid bases, nucleosides, and nucleotides at elevated temperatures and pressures, Geochim. Cosmochim. Acta, submitted.

BGD
2, 1515–1615, 2005

A.2. Solvation and non-solvation contributions

The revised HKF equations of state (Helgeson et al., 1981; Tanger and Helgeson, 1988) permit calculation of the standard molal thermodynamic properties of aqueous species as continuous functions of temperature and pressure. They have been used successfully used to represent experimental values of C°_P, V° and κ°_T of a wide variety of organic and inorganic aqueous species, including amino acids (Amend and Helgeson, 1997a; Marriott et al., 1998). The revised HKF equations of state are derived from the separation of variables represented by

$$\Xi = \Delta\Xi_{\mathrm{n}} + \Delta\Xi_{\mathrm{s}}\,, \tag{A5}$$

where Ξ stands for any standard molal thermodynamic property or equations of state parameter of an aqueous species, and $\Delta\Xi_{\mathrm{n}}$ and $\Delta\Xi_{\mathrm{s}}$ refer, respectively, to the nonsolvation and solvation contributions to that property or parameter.

The non-solvation contributions to C°_P, V°, κ°_T and E° of a given species can be expressed as

$$\Delta C^\circ_{P,\mathrm{n}} = c_1 + \frac{c_2}{(T-\Theta)^2} - \left(\frac{2T}{(T-\Theta)^3}\right)\left(a_3\,(P-P_r) + a_4 \ln\left(\frac{\Psi+P}{\Psi+P_r}\right)\right)\,, \tag{A6}$$

$$\Delta V^\circ_{\mathrm{n}} = a_1 + \frac{a_2}{\Psi+P} + \left(a_3 + \frac{a_4}{\Psi+P}\right)\left(\frac{1}{T-\Theta}\right)\,, \tag{A7}$$

$$-\Delta\kappa^\circ_{T,\mathrm{n}} = \left(a_2 + \frac{a_4}{T-\Theta}\right)\left(\frac{1}{\Psi+P}\right)^2\,, \tag{A8}$$

and

$$\Delta E^\circ = -\left(a_3 + \frac{a_4}{\Psi+P}\right)\left(\frac{1}{T-\Theta}\right)^2\,, \tag{A9}$$

BGD
2, 1515–1615, 2005

where a_1, a_2, a_3, a_4, c_1 and c_2 represent temperature- and pressure-independent coefficients for the species of interest, Θ and Ψ represent solvent parameters equal to 228 K and 2600 bars, respectively, T and P refer to the temperature and pressure of interest, and P_r stands for the reference pressure of 1 bar. The solvation contributions to C°_P, V°, κ°_T and E° in the revised HKF equations of state can be expressed as

$$\Delta C^\circ_{P,\mathrm{s}} = \omega T X + 2TY \left(\frac{\partial \omega}{\partial T}\right)_P - T\left(\frac{1}{\epsilon} - 1\right)\left(\frac{\partial^2 \omega}{\partial T^2}\right)_P , \tag{A10}$$

$$\Delta V^\circ_{\mathrm{s}} = -\omega Q + \left(\frac{1}{\epsilon} - 1\right)\left(\frac{\partial \omega}{\partial P}\right)_T , \tag{A11}$$

$$\Delta \kappa^\circ_{T,\mathrm{s}} = \omega N + 2Q\left(\frac{\partial \omega}{\partial P}\right)_T - \left(\frac{1}{\epsilon} - 1\right)\left(\frac{\partial^2 \omega}{\partial P^2}\right)_T , \tag{A12}$$

and

$$\Delta E^\circ = -\omega U , \tag{A13}$$

where ω is an equation of state parameter for the species of interest and Q, N, X, Y and U stand for the partial derivatives of the reciprocal dielectric constant of H_2O $(1/\epsilon)$ given by

$$Q \equiv -\left(\frac{\partial\,(1/\epsilon)}{\partial P}\right)_T , \tag{A14}$$

$$N \equiv \left(\frac{\partial Q}{\partial P}\right)_T , \tag{A15}$$

$$Y \equiv -\left(\frac{\partial\,(1/\epsilon)}{\partial T}\right)_P , \tag{A16}$$

BGD
2, 1515–1615, 2005

$$X \equiv \left(\frac{\partial Y}{\partial T}\right)_P , \tag{A17}$$

and

$$U \equiv \left(\frac{\partial Q}{\partial T}\right)_P . \tag{A18}$$

The values of Q, X, Y, N and U are calculated in the present study using SUPCRT92 (Johnson et al., 1992), which relies on interpolation of the values given by Shock et al. (1992) and Tanger and Helgeson (1988).

The solvation contributions in Eqs. (A10)–(A12) are represented by terms that contain ω and its first- and second- partial derivatives with respect to temperature and pressure. In the revised HKF equations of state, these partial derivates are set to zero for neutral species. In contrast, charged species generally are characterized by partial derivatives of ω that are calculated from charge-dependent correlations (Tanger and Helgeson, 1988). The use of group additivity for charged species, however, must be able to extend to neutral zwitterions, for which the partial derivatives of ω with respect to temperature and pressure are zero. This implies the partial derivatives of ω must be zero also for all charged groups, i.e.

$$\left(\frac{\partial \omega}{\partial T}\right)_P = \left(\frac{\partial \omega}{\partial P}\right)_T = \left(\frac{\partial^2 \omega}{\partial T^2}\right)_P = \left(\frac{\partial^2 \omega}{\partial T^2}\right)_P = 0 . \tag{A19}$$

For neutral species, values of the partial derivatives of ω are set equal to zero. However, group additivity involves charged groups whose partial derivatives of ω do not sum to zero. Nevertheless, comparative calculations indicate that the uncertainty introduced by this approximation is negligible in the context of the present study. Therefore, the partial derivatives of ω for neutral species, as well charged groups are taken to be zero.

BGD
2, 1515–1615, 2005

A.3. ΔG°, ΔH° and S° at T and P

Taking into account Eq. (A19), the revised HKF equations of state for ΔG°, ΔH° and S° can be written as

$$\Delta G^\circ = \Delta G^\circ_f - S^\circ_{P_r,T_r}(T - T_r) - c_1[T \ln\left(\frac{T}{T_r}\right) - T + T_r]$$
$$- c_2\left\{\left[\left(\frac{1}{T-\Theta}\right) - \left(\frac{1}{T_r-\Theta}\right)\right]\left(\frac{\Theta - T}{\Theta}\right) - \frac{T}{\Theta^2}\ln\left[\frac{T_r(T-\Theta)}{T(T_r-\Theta)}\right]\right\}$$
$$+ a_1(P - P_r) + a_2 \ln\left(\frac{\Psi + P}{\Psi + P_r}\right)$$
$$+ \left(\frac{1}{T+\Theta}\right)\left[a_3(P - P_r) + a_4 \ln\left(\frac{\Psi + P}{\Psi + P_r}\right)\right]$$
$$+ \omega\left[Y_{P_r,T_r}(T - T_r) + \frac{1}{\epsilon} - \frac{1}{\epsilon_{P_r,T_r}}\right], \quad \text{(A20)}$$

$$\Delta H^\circ = \Delta H^\circ_f + c_1(T - T_r) - c_2\left[\left(\frac{1}{T-\Theta}\right) - \left(\frac{1}{T_r-\Theta}\right)\right]$$
$$+ a_1(P - P_r) + a_2 \ln\left(\frac{\Psi + P}{\Psi + P_r}\right)$$
$$+ \left(\frac{2T-\Theta}{(T-\Theta)^2}\right)\left[a_3(P - P_r) + a_4 \ln\left(\frac{\Psi + P}{\Psi + P_r}\right)\right]$$
$$+ \omega\left[TY - T_r Y_{P_r,T_r} + \left(\frac{1}{\epsilon} - 1\right) - \left(\frac{1}{\epsilon_{P_r,T_r} - 1}\right)\right], \quad \text{(A21)}$$

BGD
2, 1515–1615, 2005

and

$$S^\circ = S^\circ_{P_r,T_r} + c_1 \ln\left(\frac{T}{T_r}\right) - \frac{c_2}{\Theta}\left\{\left(\frac{1}{T-\Theta}\right) - \left(\frac{1}{T_r-\Theta}\right) + \frac{1}{\Theta}\ln\left[\frac{T_r\,(T-\Theta)}{T\,(T_r-\Theta)}\right]\right\} + \left(\frac{1}{T-\Theta}\right)^2\left[a_3\,(P-P_r) + a_4\ln\left(\frac{\Psi+P}{\Psi+P_r}\right)\right] + \omega\left(Y - Y_{P_r,T_r}\right) . \quad \text{(A22)}$$

A.4. Regression equations

With the aid of the equations described below, values of the non-solvation revised HKF parameters can be regressed from experimental and estimated values of C°_P, V° and κ°_S as a function of temperature. At low pressure, the non-solvation contributions to each of these properties can be expressed as linear functions of two non-solvation revised HKF parameters. The contribution by the pressure-dependent terms is minimal even to a few hundred bars, so all the experimental data considered in the present study, whether measured at ambient or elevated pressures, can be treated in the following manner. For example, at $P = P_r$, Eq. (A6) reduces to

$$\Delta C^\circ_{P,\mathrm{n}} = c_1 + \frac{c_2}{(T-\Theta)^2}\,, \quad \text{(A23)}$$

which can be used to retrieve values of c_1 and c_2 from experimental values of C°_P. Similar regression calculations can be carried out for volumetric data by first defining the parameters σ and ξ as

$$\sigma \equiv a_1 + \frac{a_2}{\Psi+P} \quad \text{(A24)}$$

BGD
2, 1515–1615, 2005

and

$$\xi \equiv a_3 + \frac{a_4}{\Psi + P} \,. \tag{A25}$$

Combining Eqs. (A24) and (A25) with Eq. (A7) yields

$$V^\circ_{\mathrm{n}} = \sigma + \frac{\xi}{T - \Theta} \,, \tag{A26}$$

which can be used to retrieve values of σ and ξ from experimental values of V°. The partial derivatives with respect to pressure of Eqs. (A24) and (A25) are given by

$$\left(\frac{\partial \sigma}{\partial P}\right)_T = \frac{-a_2}{(\Psi + P)^2} \tag{A27}$$

and

$$\left(\frac{\partial \xi}{\partial P}\right)_T = \frac{-a_4}{(\Psi + P)^2} \,, \tag{A28}$$

which can be combined with Eq. A8) to give

$$-\Delta\kappa^\circ_{T,\mathrm{n}} = \left(\frac{\partial \sigma}{\partial P}\right)_T + \left(\frac{\partial \xi}{\partial P}\right)_T \left(\frac{1}{T - \Theta}\right) \,. \tag{A29}$$

This expression can be used to derive values of a_2 and a_4 from experimental values of κ°_T.

The non-solvation term can be included in the regression equations by taking account of Eqs. (A5) and (A19) along with Eqs. (A10) and (A23) for $\Delta C^\circ_{P,\mathrm{n}}$ or Eqs. (A11) and (A26) for $\Delta V^\circ_{\mathrm{n}}$. Thus, it is possible to write

$$C^\circ_P = c_1 + \frac{c_2}{(T - \Theta)^2} + \omega T X \tag{A30}$$

and

$$V^\circ = \sigma + \frac{\xi}{T - \Theta} - \omega Q\,, \tag{A31}$$

which can be used to retrieve values of ω where sufficient high-temperature experimental data are available. Because this is not possible for all of the amino acids, provisional values of ω can be estimated from correlations among the solvation or hydration properties of the species.

B. Calculation of updated values of the revised HKF parameters of organic groups

The equations of state parameters given by Amend and Helgeson (1997b) for aqueous organic groups were derived in part from those for aqueous alcohols and diols, and amines and diamines. Because homologous series were used in their analysis, the non-solvation equations of state parameters of the [$-CH_2-$] group were accurately determined. However, little or no calorimetric data were available at the time for diols and diamines at temperatures other than 25°C. As a consequence, large uncertainties are inherent in their group additivity predictions of the equations of state parameters of molecules such as *n*-alkanes and amino acids. Owing to the availability of more recent experimentdal data reported in the literature, these uncertainties can now be reduced considerably by revising the organic group contirbutions derived from diols and diamines. This can be done by introducing a correction term, [di.corr], to modify the group additivity contributions dervied from diols and diamines Doing so results in revised values of the parameters of the terminal groups used by Amend and Helgeson (1997b), which are given in Table 10.

The values of [di.corr] for ΔG_f°, ΔH_f° and $S^\circ_{P_r,T_r}$ are taken to be zero, because the group additivity calculations using the unmodified group contributions to these properties given by Amend and Helgeson (1997b) accurately predict the properties of amino acids and many other organic compounds, including diols and diamines.

The group equations of state parameters given by Amend and Helgeson (1997b) are indicated below by a superscript AH97b. In contrast, the symbols representing the revised parameters carry no superscript. For propanol, no correction term is needed, so the revised group contributions do not change the group additivity predictions. Consequently, one can write Eq. (B1) where Ξ represents any equation of state parameter. For butane-1,4-diol, a correction term in the revised group additivity algorithm is needed, but in calculating a trial value the revision should not alter the results of the group additivity analysis. Therefore, one can write Eq. (B2) where $\Xi_{[\text{di.corr}]}$ denotes the correction term. The value of this term can be calculated from amino acid and Gly–X–Gly group additivity algorithms. For example, combining Eqs. (B1) and (B2), and taking $\Xi_{[-CH_2-]} = \Xi^{AH97b}_{[-CH_2-]}$, one obtains Eqs. (B3) and (B4). Generalizing Eq. (B4) permits the revised contributions by the other terminal groups (excluding $[-CH_3]$) to be calculated from Eq. (B5) where $[term]$ stands for $[-COOH]$, $[-CH_2OH]$, $[-CONH_2]$ or $[-CH_2NH_2]$. Because $[-CHCH_3-]$ is not a terminal group, its revised group contributions are given by Eq. (B6).

The constraints given by the above equations ensure that the revised group contributions are consistent not only with amino acid properties, but also with the properties of the compounds from which they were originally derived. Taking into account these equations, the value of $\Xi_{[\text{di.corr}]}$ for $\Xi = C^\circ_P$, c_2, V°, ξ, a_2 and a_4, is selected such that the revised group contributions minimize the uncertainty in the calculation of [AABB] and [GXGBB] from the reference amino acids and Gly–X–Gly tripeptides identified above. For example, an initial value of $C^\circ_{P,[\text{di.corr}]} = -22.5$ yields the revised values of $C^\circ_{P,[-CH_3]}$ and $C^\circ_{P,[-CH_2OH]}$ given in Table 10. These revised values yield $C^\circ_{P,[\text{AABB}]}$ of −2.2 and −0.7 cal mol^{-1} K^{-1} from Ala and Ser, respectively, which indicates a much smaller uncertainty than that inherent in the original group contributions adopted by Amend and Helgeson (1997b).

The values of C°_P, c_2, V° and ξ of [AABB] and [GXGBB] correspond to the values of the intercepts of the regression lines on the plots of Ξ_{AA} and Ξ_{GXG} vs. $\Xi_{[SC]}$ depicted in Fig. 10. For the most part, the symbols shown in this figure lie within a representative

total group additivity uncertainty. In contrast, the symbols corresponding to C_P°, c_2, V° and ξ of [$-C_6H_5$] and ξ of [$-$COOH] lie outside the representative uncertainties and are not shown. Consequently, the revised contributions by these groups to the properties are not calculated from Eq. (B5) but are computed from the respective properties of the corresponding amino acids.

Of the groups considered by Amend and Helgeson (1997b), only $[-CH_2]^{AH97b}$ and $[-CH_2OH]^{AH97b}$ have values of a_2 and a_4 that were derived from non-amino acid compounds. Therefore, the values of a_2 and a_4 of [di.corr] and [AABB] can be calculated by solving the system of equations represented by Eq. (B7) using the values of a_2 and a_4 of Ala and Ser given in Table 3.

The values of V° and ξ of [$-CHCH_3-$] and those of a_2 and a_4 of [$-CHCH_3-$], [$-$COOH], [$-CONH_2$], [$-C_6H_5$] and [$-CH_2NH_2$] can be computed by combining appropriate statements of Eqs. (1) and (2) with the equations of state parameters of amino acids taken from Table 3 and those of [AABB], [$-CH_3$] and [$-CH_2-$] given in Table 10. The corresponding properties of of [di.corr], as well as values of ω of all the groups can be calculated from the regression of the experimental high-temperature heat capacities of mono- and di-alcohols and amines described below.

Experimental values of C_P° at 280 bar and temperatures to ~250°C along the vapor-liquid saturation curve of propanol, butane-1,4-diol, hexane-1,6-diol, propylamine, butane-1,4-diamine and hexane-1,6-diamine have been reported in the literature (Inglese and Wood, 1996; Inglese et al., 1997). These high-temperature data are suitable for determining the values of ω of [di.corr], [$-CH_3$], [$-CH_2-$], [$-CH_2OH$], [$-CH_2NH_2$], as well as those of c_1 and c_2 of [di.corr] by taking account the systems of equations represented by Eqs. (B8) and (B9) for alcohols and amines, respectively. Iterative regression of the high-temperature experimental data of the diols and diamines shown as symbols in Fig. 17 yields trial values of ω of [$-CH_2$], which can be substituted into Eqs. (B8) and (B9) to retrieve values of ω of the four remaining groups by iterative regression to obtain the best fits of the data. Note in Fig. 17 that the the revised group contributions to ω result in regression lines which are consistent within experimental

uncertainty with both the high-temperature data for the mono- and di- alcohols and amines. The dearth of such high-temperature data for these types of compounds at the time prevented Amend and Helgeson (1997b) from performing a similar calculation.

C. Calculation of ΔG_f°, ΔH_f° and $S^\circ_{P_r,T_R}$ of gaseous amino acids

Group additivity is used in the following manner to calculate the values of ΔG_f°, ΔH_f° and S° of the gaseous amino acids – except for Ala, Gly and Pro – given in Table 2. The properties of the sidechain groups in the amino acids can be calculated from the group additivity algorithm represented by Eq. (1), where $\Xi_{[SC]}$ and Ξ_i stand for any property or parameter of the sidechain group of interest and the ith group contribution, respectively. The gaseous group contributions to ΔH_f° and S° at 25°C and 1 bar are summarized in Table 15 and are taken from Domalski and Hearing (1993), with the exception of a few groups for which the properties can be estimated using the strategies described in the footnotes to the table. In particular, the properties of [AABB] given in Table 15 are calculated using Eq. (2) and the experimental properties of Ala. For gaseous amino acids, we adopt a modification of the additivity scheme developed by Benson and Buss (1958) and Domalski and Hearing (1993). Specific statements of Eq. (1) are given in Table 14 for all of the gaseous amino acid sidechain groups except [Gly] and [Pro]. These sidechain groups are not readily modeled by the additivity scheme ([Gly] because it has no atoms other than H; [Pro] because it is bonded twice to [AABB]). In addition, ΔH_f° and $S^\circ_{P_r,T_r}$ of Gly and Pro have been measured or can be calculated.

BGD
2, 1515–1615, 2005

D. Group additivity calculation of the standard molal thermodynamic properties of crystalline proteins as a function of temperature and pressure

The available high-temperature experimental values of C°_P (up to 146.85°C) of crystalline amino acids and proteins indicate that the trend of C°_P of these compounds can be modeled with a linear approximation. This is not the generally the case for crystalline compounds, nor is it true of proteins at temperatures below 100 K. Nevertheless, in the context of the present study, the values of c in the Maier-Kelley equation can be taken to be zero for the amino acids, polypeptides, proteins, and their consituent groups.

D.1. Revision of the Maier-Kelley parameters of crystalline Leu

Unlike C°_P of the other amino acids, the heat capacity of Leu exhibits an upward curvature near the upper temperature limit of the experimental measurements (Hutchens et al., 1963). This observation lead Helgeson et al. (1998) to fit the C°_P of Leu with a non-zero value of c using the Maier-Kelley equation, represented by

$$C^\circ_P = a + bT + cT^{-2}\,, \tag{D1}$$

where a, b and c are temperature-independent coefficients for the species of interest. However, to maintin compatibility with the linear trend of C°_P of crystalline proteins, and because the observed curvature in C°_P of Leu is characteristic of the anticipation of a structural transition in the crystal, the Maier-Kelley parameters of Leu are reevaluated in the present study with the aid of the regression plot shown in Fig. 18. It is evident from this figure that the first two terms of Eq. (D1) accurately fit the experimental values of C°_P of Leu below the onset of transition. The values of the Maier-Kelley parameters of the remaining amino acids are taken from Helgeson et al. (1998).

D.2. Calculation of the standard molal thermodynamic properties at 25 °C and 1 bar and the Maier-Kelley parameters of crystalline amino acid backbone and sidechain groups

The values of ΔH_f°, $S^\circ_{P_r,T_r}$, V° and C_P° at 25°C and 1 bar, and b of crystalline Ala, Val, Leu and Ile taken from Helgeson et al. (1998) can combined with the group contributions by [$-CH_3$], [$-CH_2-$] and [> CH−] to these properties taken from Richard and Helgeson (1998) to generate the regression plots shown in Fig. 19. The intercepts of the regression lines in this figure represent the contributions by the crystalline [AABB] to each of these properties, the values of which are given in Table 16. The values of ΔG_f° and a of [AABB] given in Table 16 are generated using Eqs. (A4) or (D1), combined with the values of ΔH_f° and $S^\circ_{P_r,T_r}$, or C_P°, b and c, respectively, of [AABB].

D.3. Calculation of $S^\circ_{P_r,T_r}$ of the crystalline protein backbone group

The group additivity calculation of ΔG_f°, ΔH_f° and $S^\circ_{P_r,T_r}$ of crystalline proteins, which is required for the additive calculation of ΔH°_{sol} of proteins, can be performed in the manner described below. $S^\circ_{P_r,T_r}$ of the crystalline protein backbone can be calculated from values of $S^\circ_{P_r,T_r}$ reported by Hutchens et al. (1969) for crystalline insulin and chymotrypsinogen A (0.3144 and 0.3227 cal K^{-1} g^{-1}, respectively) and by Mrevlishvili (1986) for bovine albumin (0.3207 cal K^{-1} g^{-1}). Converting these values to units of cal K^{-1} mol^{-1} using masses of 5773, 25666, and 66433 g, respectively, for one mole of the insulin monomer with one-half mole zinc, chymotrypsinogen A, and bovine albumin, and substituting the contributions to S° given in Table 16 from the crystalline sidechain groups and amino acid backbone into Eq. 8, along with the numer of amino acid residues of each protein, gives 11.18, 11.56, and 11.13 cal K^{-1} mol^{-1} for the values of $S^\circ_{P_r,T_r}$ of the crystalline protein backbone in insulin, chymotrypsinogen A, and bovine albumin, respectively. Because of the uncertain contribution by zinc to $S^\circ_{P_r,T_r}$ of insulin, the mean of only the latter two values, 11.34, is shown in Table 16. It can

BGD
2, 1515–1615, 2005

be combined with $S^\circ_{P_r,T_r}$ of [AABB] from the same table and that of H_2O (16.71 cal K^{-1} mol^{-1}; calculated using SUPCRT92) to give ΔS° of the reaction represented by [AABB] $\rightleftharpoons$ [PBB] + H_2O. This value, 8.83 cal K^{-1} mol^{-1}, is about 0.2 to 0.5 cal K^{-1} mol^{-1} less than those calculated by Hutchens et al. (1969).

D.4. Calculation of ΔG°_f and ΔH°_f of the crystalline protein backbone group

An experimental value of ΔH° of combustion of a crystalline protein of known sequence is available in the literature for only one protein (Tsuzuki et al., 1958). Although measurements of ΔH°_f of combustion of two other crystalline proteins have been reported (Kienzle et al., 2001), they are referenced to proteins of ambiguous composition. Therefore, a provisional value of ΔH°_f of [PBB], given in Table 16, is calculated from Eq. (8) using the enthalpy of combustion of INS_PIG (5382.2 cal g^{-1}) reported by Tsuzuki et al., 1958 combined with the values of ΔH°_f of the crystalline amino acid backbone groups summarized above and ΔH°_f of the gaseous combustion products. Because only one reference model protein is included in this calculation, the estimated uncertainty is 1 kcal mol^{-1}, which is large compared to the uncertainty estimated for ΔH°_f of unfolded proteins. Reconnaisance calculations using ΔH° of combustion of crystalline tri- and tetrapeptides (Chemical Rubber Company, 1975) yield values of of ΔH°_f of [UPBB] that are 3 kcal mol^{-1} smaller than that calculated from insulin, which differ from the value obtained from the calculation with INS_PIG by more than the estimated uncertainty, supporting the notion that crystalline polypeptides of this length are perhaps poor candidates for modeling the thermodynamic properties of crystalline proteins in the reference state adopted here.

A comparison of calculated and experimental values of ΔH°_{sol} of four proteins is shown in Fig. 20. The calculated values of ΔH°_{sol} of the four proteins represented in this figure are calculated from Eq. (8) and the respective group contributions to ΔH°_f of nonionized unfolded and crystalline proteins taken from Tables 8 and 16. It can be seen in Fig. 20 that, in general, the additive values of ΔH°_{sol} deviate from the experimen-

tal values taken from Souillac et al. (2002) by less than the experimental uncertainty, which we estimate to be ±20% of the experimental value. This degree of uncertainty is comparable to the differences between the experimental ΔH°_{sol} of lyophilized and non-lyophilized (as-is) samples reported by Souillac et al. (2002). It should perhaps be noted, however, that the aqueous proteins in these experiments are most likely not nonionized unfolded proteins, but are instead both ionized and folded. The ionization contribution can be approximated by taking account of Eqs. (26) and (23) for arbitrary values of pH. The values of ΔH°_{sol} of ionized unfolded proteins calculated in this manner are represented by the position of the text-labeled symbols shown in Fig. 20 for pHs of 3, 7 and 10. It is apparent from the trend of these values that the experimental data are in some cases more closely represented by ionized unfolded proteins present in a solutions between pH 7 and 10. The question remains of what is the experimental pH of the protein solution, which is not reported by Souillac et al. (2002). In the future, it is evident that protein ionization may contribute substatially to the observed ΔH°_{sol}. A possible compensating factor may arise from the ionization state of the crystalline proteins. Because they contain at least some water, the crystalline proteins themselves might be ionized, with perhaps similar enthalpic consequences as for aqueous proteins. Such an interpretation would suggest a lessening the effects of ionization upon dissolution of the proteins.

Another source of uncertainty in these comparative calculations is the difference between the enthalpies of ionized unfolded proteins and those of ionized folded proteins. Because the measurements of ΔH°_{sol} probably involved folded proteins, and because protein unfolding is generally an endothermic process, the calculated values of ΔH°_{sol} may be overestimated by 10 to 50 kcal mol^{-1}. Consideration of folded proteins would lead to shifts in the calculated values of ΔH°_{sol} shown in Fig. 20 in the same direction as shown for the ionization effects.

BGD
2, 1515–1615, 2005

D.5. Calculation of $V^{\circ}_{P_r,T_r}$ of the crystalline protein backbone group

As with heat capacity, the volume of protein crystals changes considerably with the degree of hydration. For example, consideration of the unit cell parameters of crystals of RNP_BOVIN with Protein Data Bank IDs 1BEL and 1C0C indicates that the unit cell volume decreases from 109 000 to 72 000 Å^3 upon desiccation by exposure to $CaSO_4$. It has also been noted that temperature (Kurinov and Harrison, 1995) and pressure (Vant et al., 2002) may affect the volumetric properties of protein crystals. Nevertheless, it might prove useful for geochemical calculations to adopt a reference state for calculating the contributions by the protein backbone group to $V^{\circ}_{P_r,T_r}$ of crystalline proteins.

D.6. Regression of the Maier-Kelley parameters of the crystalline protein backbone group

Figure 21 shows the results of an iterative fitting procedure to retrieve values of *a* and *b* of [UPBB], which are reported in Table 16. The result is a close correspondence between the calculated and experimental heat capacities as a function of temperature. In terms of the reproduction of experimental data, our results are similar to those of Bakk's 2002 calculations of the contributions by the vibrational modes of the molecules. But the group additivity concept adopted here has the added advantage of using just the amino acid sequence to predict the heat capacities of crystalline proteins.

Acknowledgements. This work was performed as a part of J.M.D.'s Ph.D. thesis project. Preliminary results of the calculations of protein ionization were presented at the 14th V. M. Goldschmidt conference in Copenhagen in June 2004. Funding for this project was provided by the US National Science Foundation (Grants EAR-96-13753 and EAR-03-09829), the US Department of Energy (Continuing Grants FG03-85ER-13419 and FG02-03ER-15418), the American Chemical Society Petroleum Research Fund (Grant 36467-AC2) and the University of California Committee on Research. We are grateful to F. MacKenzie and the staff, students and faculty of the University of Hawai'i for their hospitality during H.C.H.'s sabbatical in 2002–2003. This work benefitted from J. Amend's comments on an earlier version of the manuscript.

References

Ahluwalia, J. C., Ostiguy, C., Perron, G., and Desnoyers, J. E.: Volumes and heat capacities of some amino acids in water at 25°C, Can. J. Chem., 55, 3364–3367, 1977. 1597

Al-Habori, M.: Microcompartmentation, metabolic channeling and carbohydrate metabolism, Int. J. Biochem. Cell Biol., 27, 123–132, 1995.

Amend, J. P. and Helgeson, H. C.: Calculation of the standard molal thermodynamic properties of aqueous biomolecules at elevated temperatures and pressures. Part 1. L-α-amino acids, J. Chem. Soc., Faraday Trans., 93, 1927–1941, 1997a.

Amend, J. P. and Helgeson, H. C.: Group additivity equations of state for calculating the standard molal thermodynamic properties of aqueous organic species at elevated temperatures and pressures, Geochim. Cosmochim. Acta, 61, 11–46, 1997b.

Amend, J. P. and Helgeson, H. C.: Calculation of the standard molal thermodynamic properties of aqueous biomolecules at elevated temperatures and pressures. II. Unfolded proteins, Biophys. Chem., 84, 105–136, 2000.

Andersson, E. and Holm, N. G.: The stability of some selected amino acids under attempted redox constrained hydrothermal conditions, Orig. Life Evol. Biosph., 30, 9–23, 2000.

Bakk, A.: Heat capacities of solid state proteins: implications for protein stability in solution, Physica A, 313, 540–548, 2002. 1547 1518, 1519, 1527, 1539, 1549, 1578 1518, 1520, 1521, 1525, 1545, 1555, 1556, 1557, 1558, 1585, 1591 1519, 1530, 1534 1545 1563

Partial molal volumes and expansibilities of some amino acids in aqueous solutions, J. Indian Chem. Soc., 76, 431–437, 1999. 1597

Benson, S. W. and Buss, J. H.: Additivity rules for the estimation of molecular properties – Thermodynamic properties, J. Chem. Phys., 29, 546–572, 1958. 1558, 1593

Boeckmann, B., Bairoch, A., Apweiler, R., Blatter, M. C., Estreicher, A., Gasteiger, E., Martin, M. J., Michoud, K., O'Donovan, C.. Phan, I., Pilbout, S., and Schneider, M.: The SWISS-PROT protein knowledgebase and its supplement TrEMBL in 2003, Nucleic Acids Res., 31, 365–370, 2003. 1589

Boonyaratanakornkit, B. B., Simpson, A. J., Whitehead, T. A., Fraser, C. M., El-Sayed, N. M. A., and Clark, D. S.: Transcriptional profiling of the hyperthermophilic methanarchaeon *Methanococcus jannaschii* in response to lethal heat and non-lethal cold shock, Environ. Microbiol., 7, 789–797, 2005.

Cabani, S., Gianni, P., Mollica, V., and Lepori, L.: Group contributions to the thermodynamic

properties of non-ionic organic solutes in dilute aqueous solution, J. Solution Chem., 10, 563–595, 1981.
Chemical Rubber Company: Handbook of biochemistry and molecular biology, CRC Press, 1975.
Clarke, R. G. and Tremaine, P. R.: Amino acids under hydrothermal conditions: Apparent molar volumes of aqueous α-alanine, β-alanine, and proline at temperatures from 298 to 523 K and pressures up to 20.0 MPa, J. Phys. Chem. B, 103, 5131–5144, 1999.
Clarke, R. G., Hnědkovský, L., Tremaine, P. R., and Majer, V.: Amino acids under hydrothermal conditions: Apparent molar heat capacities of aqueous α-alanine, β-alanine, glycine, and proline at temperatures from 298 to 500 K and pressures up to 30.0 MPa, J. Phys. Chem. B, 104, 11 781–11 793, 2000. 1517 1518, 1543, 1611 1561 1518, 1597 1518, 1596
Clarke, R. G. F., Collins, C. M., Roberts, J. C., Trevani, L. N., Bartholomew, R. J., and Tremaine, P. R.: Ionization constants of aqueous amino acids at temperatures up to 250°C using hydrothermal pH indicators and UV-visible spectroscopy: Glycine, α-alanine, and proline, Geochim. Cosmochim. Acta, 69, 3029–3043, 2005. 1610
Cohn, E. J. and Edsall, J. T.: Proteins, amino acids and peptides as ions and dipolar ions, Reinhold Publishing Corporation, New York, 1943.
Connolly, J. A. D.: Multivariable phase diagrams: An algorithm based on generalized thermodynamics, Am. J. Sci., 290, 666–718, 1990.
Conour, J. E., Graham, W. V., and Gaskins, H. R.: A combined in vitro/bioinformatic investigation of redox regulatory mechanisms governing cell cycle progression, Physiol. Genomics, 18, 196–205, 2004.
Cox, J. D., Wagman, D. D., and Medvedev, V. A., eds.: CODATA Key Values for Thermodynamics, Hemisphere Publishing Corporation, New York, 1989.
Desnoyers, J. E. and Philip, P. R.: Isothermal compressibilities of aqueous solutions of tetraalkylammonium bromides, Can. J. Chem., 50, 1094–1096, 1972. 1519, 1537 1546 1547 1548 1525
Di Lorenzo, M. L., Zhang, G., Pyda, M., Lebedev, B. V., and Wunderlich, B.: Heat capacity of solid-state biopolymers by thermal analysis, J. Polym. Sci. Pt. B-Polym. Phys., 37, 16, 2093–2102, 1999. 1615
Dill, K. A.: Additivity principles in biochemistry, J. Biol. Chem., 272, 701–704, 1997.
Ding, K., Seyfried, W. E., Tivey, M. K., and Bradley, A. M.: In situ measurement of dissolved H_2 and H_2S in high-temperature hydrothermal vent fluids at the Main Endeavour Field, Juan de

Fuca Ridge, Earth Planet. Sci. Lett., 186, 417–425, 2001. 1543 1517
DiPaola, G. and Belleau, B.: Apparent molal heat capacities and volumes of amino acids in aqueous polyol solutions, Can. J. Chem., 56, 1827–1831, 1978. 1596
Domalski, E. S. and Hearing, E. D.: Estimation of the thermodynamic properties of C-H-N-O-S-Halogen compounds at 298.15 K, J. Phys. Chem. Ref. Data, 22, 805–1159, 1993.
Downes, C. J. and Hedwig, G. R.: Partial molar heat capacities of the peptides glycylglycylglycine, glycyl-L-alanylglycine and glycyl-DL-threonylglycine in aqueous solution over the temperature range 50 to 125 ˚C, Biophys. Chem., 55, 279–288, 1995. 1558, 1593 1518, 1524, 1601
Downes, C. J., Hakin, A. W., and Hedwig, G. R.: The partial molar heat capacities of glycine and glycylglycine in aqueous solution at elevated temperatures and at p=10.0 MPa, J. Chem. Thermodyn., 33, 873–890, 2001. 1596
Duke, M. M., Hakin, A. W., McKay, R. M., and Preuss, K. E.: The volumetric and thermochemical properties of aqueous solutions of L-valine, L-leucine, and L-isoleucine at 288.15 K, 298.15 K, and 328.15 K, Can. J. Chem., 72, 1489–1494, 1994. 1596, 1597
Edsall, J. T. and Wyman, J.: Biophysical Chemistry, vol. 1, Academic Press, New York, 1958.
Elcock, A.: Realistic modeling of the denatured states of proteins allows accurate calculations of the pH dependence of protein stability, J. Mol. Biol., 294, 1051–1062, 1999.
Elser, J. J., Sterner, R. W., Gorokhova, E., Fagan, W. F., Markow, T. A., Cotner, J. B., Harrison, J. F., Hobbie, S. E., Odell, G. M., and Weider, L.: Biological stoichiometry from genes to ecosystems, Ecol. Lett., 3, 540–550, 2000.
Fukuchi, S. and Nishikawa, K.: Protein surface amino acid compositions distinctively differ between thermophilic and mesophilic bacteria, J. Mol. Biol., 309, 835–843, 2001.
Gasch, A. P., Spellman, P. T., Kao, C. M., Carmel-Harel, O., Eisen, M. B., Storz, G., Botstein, D., and Brown, P. O.: Genomic expression programs in the response of yeast cells to environmental changes, Mol. Biol. Cell, 11, 4241–4257, 2000.
Georgescu, R. E., García-Mira, M. M., Tasayco, M. L., and Sánchez-Ruiz, J. M.: Heat capacity analysis of oxidized *Escherichia coli* thioredoxin fragments (1-73,74-108) and their noncovalent complex – Evidence for the burial of apolar surface in protein unfolded states, Eur. J. Biochem., 268, 1477–1485, 2001. 1536 1545 1517 1547 1517 1533, 1544
Griko, Yu. V., Makhatadze, G. I., Privalov, P. L., and Hartley, R. W.: Thermodynamics of barnase unfolding, Protein Sci., 3, 669–676, 1994. 1589
Guzman-Casado, M., Parody-Morreale, A., Robic, S., Marqusee, S., and Sanchez-Ruiz, J. M.:

Energetic evidence for formation of a pH-dependent hydrophobic cluster in the denatured state of *Thermus thermophilus* ribonuclease H, J. Mol. Biol., 329, 731–743, 2003.
Häckel, M., Hinz, H.-J., and Hedwig, G. R.: Tripeptides in aqueous solution: Model compounds for the evaluation of the partial molar heat capacities of amino acid side-chains in proteins, Thermochim. Acta, 308, 23–34, 1998.
Häckel, M., Hinz, H.-J., and Hedwig, G. R.: A new set of peptide-based group heat capacities for use in protein stability calculations, J. Mol. Biol., 291, 197–213, 1999a.
Häckel, M., Hinz, H.-J., and Hedwig, G. R.: Partial molar volumes of proteins: amino acid side-chain contributions derived from the partial molar volumes of some tripeptides over the temperature range 10–90°C, Biophys. Chem, 82, 35–50, 1999b. 1533, 1543, 1544, 1589, 1609 1518, 1582, 1601, 1602 1518, 1582, 1601, 1602 1518
Hakin, A. W. and Hedwig, G. R.: Group additivity calculations of the thermodynamic properties of unfolded proteins in aqueous solution: a critical comparison of peptide-based and HKF models, Biophys. Chem., 89, 253–264, 2001a. 1544
Hakin, A. W. and Hedwig, G. R.: Apparent and partial molar heat capacities and volumes of the amino acids L-lysine monohydrochloride and L-arginine monohydrochloride in aqueous solution at temperatures from T=288.15 K to T=328.15 K, J. Chem. Thermodyn., 33, 1709–1723, 2001b 1596, 1597
Hakin, A. W., Duke, M. M., Klassen, S. A., McKay, R. M., and Preuss, K. E.: Apparent molar heat capacities and volumes of some aqueous solutions of aliphatic amino acids at 288.15, 298.15, 313.15, and 328.15 K, Can. J. Chem, 72, 362–368, 1994a. 1582, 1596, 1597
Hakin, A. W., Duke, M. M., Marty, J. L., and Preuss, K. E.: Some thermodynamic properties of aqueous amino acid systems at 288.15, 298.15, 313.15 and 328.15 K: Group additivity analyses of standard-state volumes and heat capacities, J. Chem. Soc., Faraday Trans., 90, 2027–2035, 1994b. 1596, 1597
Hakin, A. W., Duke, M. M., Groft, L. L., Marty, J. L., and Rushfeldt, M. L.: Calorimetric investigations of aqueous amino acid and dipeptide systems from 288.15 to 328.15 K, Can. J. Chem., 73, 725–734, 1995. 1596, 1597
Hakin, A. W., Copeland, A. K., Liu, J. L., Marriott, R. A., and Preuss, K. E.: Densities, apparent molar volumes, and apparent molar heat capacities of *l*-arginine, *l*-proline and *d,l*-methionine in water at 288.15, 298.15, 313.15, and 328.15 K, J. Chem. Eng. Data, 42, 84–89, 1997a. 1596, 1597
Hakin, A. W., Groft, L. L., Marty, J. L., and Rushfeldt, M. L.: Modeling of transfer properties

using revised HKF theory: Thermodynamics of transfer of an amino acid and small peptides from water to urea-water solutions at 298.15 K, Can. J. Chem., 75, 456–464, 1997b.
Hakin, A. W., Daisley, D. C., Delgado, L., Liu, J. L., Marriott, R. A., Marty, J. L., and Tompkins, G.: Volumetric properties of glycine in water at elevated temperatures and pressures measured with a new optically driven vibrating-tube densimeter, J. Chem. Thermodyn., 30, 583–606, 1998.
Hedwig, G. R. and Hinz, H.-J.: Group additivity schemes for the calculation of the partial molar heat capacities and volumes of unfolded proteins in aqueous solution, Biophys. Chem., 100, 239–260, 2003.
Helgeson, H. C., Kirkham, D. H., and Flowers, G. C.: Theoretical prediction of the thermodynamic behavior of aqueous electrolytes at high pressures and temperatures. 4. Calculation of activity coefficients, osmotic coefficients, and apparent molal and standard and relative partial molal properties to 600 °C and 5 Kb, Am. J. Sci., 281, 1249–1516, 1981.
Helgeson, H. C., Owens, C. E., Knox, A. M., and Richard, L.: Calculation of the standard molal thermodynamic properties of crystalline, liquid, and gas organic molecules at high temperatures and pressures, Geochim. Cosmochim. Acta, 62, 985–1081, 1998.
Hnědkovský, L. and Cibulka, I.: Group contributions for an estimation of partial molar volumes at infinite dilution for aqueous organic solutes at extended ranges of temperature and pressure, Int. J. Thermophys., 25, 387–395, 2004. 1518 1544 1517, 1549 1559, 1560 1542
Howard, S. B., Twigg, P. J., Baird, J. K., and Meehan, E. J.: The solubility of hen-egg white lysozyme, J. Cryst. Growth, 90, 94–104, 1988. 1614
Hu, C.-H. and Zou, C.-L.: Disulfide containing proteins denatured in 6 mol/L guanidinium chloride are not completely unfolded, Sci. China, Ser. B, 35, 1215–1221, 1992.
Hu, C.-H. and Zou, C.-L.: Effect of native disulfide bonds on the denaturation of proteins in 6 mol/L guanidinium hydrochloride, Sci. China, Ser. B, 36, 568–574, 1993.
Hutchens, J. O., Cole, A. G., and Stout, J. W.: Heat capacities from 11 to 305 °K, entropies and free energies of formation of L-valine, L-isoleucine, and L-leucine, J. Phys. Chem., 67, 1128–1130, 1963.
Hutchens, J. O., Cole, A. G., and Stout, J. W.: Heat capacities from 11 to 305 °K and entropies of hydrated and anhydrous bovine zinc insulin and bovine chymotrypsinogen A, J. Biol. Chem., 244, 26–32, 1969.
Inglese, A. and Wood, R. H.: Apparent molar heat capacities of aqueous solutions of 1-propanol, butane-1,4-diol, and hexane-1,6-diol at temperatures from 300 K to 525 K and

a pressure of 28 MPa, J. Chem. Thermodyn., 28, 1059–1070, 1996.
Inglese, A., Sedlbauer, J., Yezdimer, E. M., and Wood, R. H.: Apparent molar heat capacities of aqueous solutions of propylamine, 1,4-butanediamine, 1,6-hexanediamine, propylamine hydrochloride, propionamide, pyridine, and sodium benzenesulfonate at temperatures from 300 K to 525 K and a pressure of 28 MPa, J. Chem. Thermodyn., 29, 517–531, 1997.
Jardine, J. J., Call, T. G., Patterson, B. A., Origlia-Luster, M. L., and Woolley, E. M.: Thermodynamics for proton dissociations from aqueous L-histidine at temperatures from 278.15 K to 393.15 K and at the pressure 0.35 MPa: apparent molar volumes and apparent molar heat capacities of the protonated cationic, neutral zwitterionic, and deprotonated anionic forms, J. Chem. Thermodyn., 33, 1419–1440, 2001. 1537 1537 1559, 1612 1560, 1561, 1615 1557, 1611 1557, 1611 1526, 1596, 1597
Jimenez, P., Roux, M. V., and Turrion, C.: Thermochemical properties of N-heterocyclic compounds. 4. Enthalpies of combustion, vapor pressures and enthalpies of sublimation, and enthalpies of formation of 2-methylimidazole and 2-ethylimidazole, J. Chem. Thermodyn., 24, 1145–1149, 1992. 1593
Johnson, J. W., Oelkers, E. H., and Helgeson, H. C.: SUPCRT92: A software package for calculating the standard molal thermodynamic properties of minerals, gases, aqueous species, and reactions from 1 to 5000 bar and 0 to 1000 ˚C, Comp. Geosci., 18, 899–947, 1992. 1551
Jolicoeur, C. and Lacroix, G.: Thermodynamic properties of aqueous organic solutes in relation to their structure, 3. Apparent molal volumes and heat capacities of low molecular weight alcohols and polyols at 25˚C, Can. J. Chem., 54, 4, 624–631, 1976. 1611
Jolicoeur, C., Riedl, B., Desrochers, D., Lemelin, L. L., Zamojska, R., and Enea, O.: Solvation of amino acid residues in water and urea-water mixtures: Volumes and heat capacities of 20 amino acids in water and in 8 molar urea at 25 ˚C, J. Solution Chem., 15, 109–128, 1986.
Karpov, I. K., Chudnenko, K. V., Kulik, D. A., Avchenko, O. V., and Bychinskii, V. A.: Minimization of gibbs free energy in geochemical systems by convex programming, Geochem. Int., 39, 1108–1119, 2001.
Kato, K., Kobayashi, T., Yamamoto, H., Nakagawa, T., Maki, Y., and Hoaki, T.: Microbial mat boundaries between chemolithotrophs and phototrophs in geothermal hot spring effluents, Geomicrobiol. J., 21, 91–98, 2004. 1527, 1596, 1597 1546 1517
Kharakoz, D. P.: Volumetric properties of proteins and their analogues in diluted water solutions, 2. Partial adiabatic compressibilities of amino acids at 15–70˚C, J. Phys. Chem., 95, 5634–

5642, 1991.
Kharakoz, D. P.: Partial volumes and compressibilities of extended polypeptide chains in aqueous solution: Additivity scheme and implication of protein unfolding at normal and high pressure, Biochemistry, 36, 10 276–10 285, 1997. 1598 1519, 1534
Kharakoz, D. P.: Volumetric properties of proteins and their analogs in diluted water solutions. 1. Partial volumes of amino acids at 15–55°C, Biophys. Chem., 34, 115–125, 1989. 1597
Kienzle, E., Schrag, I., Butterwick, R., and Opitz, B.: Calculation of gross energy in pet foods: new data on heat combustion and fibre analysis in a selection of foods for dogs and cats, J. Anim. Physiol. a. Anim. Nutr., 85, 148–157, 2001. 1561
Kikuchi, M., Sakurai, M., and Nitta, K.: Partial molar volumes and adiabatic compressibilites of amino acids in dilute aqueous solutions at 5, 15, 25, 35, and 45°C, J. Chem. Eng. Data, 40, 935–942, 1995. 1597, 1598
Kreil, D. P. and Ouzounis, C. A.: Identification of thermophilic species by the amino acid compositions deduced from their genomes, Nuc. Acids Res., 29, 1608–1615, 2001. 1547
Kulagina, T. G., Bykova, T. A., and Lebedev, B. V.: Thermodynamic functions of lysozyme, myoglobin, beta-lactoglobulin, albumin, and poly-L-tryptophan in the temperature range 0–330 K, Biofizika, 46, 210–215, 2001. 1615
Kulik, D. A.: GEM-Selektor version 2-PSI, http://les.web.psi.ch/Software/GEMS-PSI/index.html, 2004.
Kundrotas, P. J. and Karshikoff, A.: Model for calculation of electrostatic interactions in unfolded proteins, Phys. Rev. E, 65, 2002.
Kurinov, I. V. and Harrison, R. W.: The influence of temperature on lysozyme crystals – Structure and dynamics of protein and water, Acta Crystallogr. Sect. D-Biol. Crystallogr., 51, 98–109, 1995.
Laderman, K. A., Davis, B. R., Krutzsch, H. C., Lewis, M. S., Griko, Y. V., Privalov, P. L., and Avnfinsen, C. B.: The purification and characterization of an extremeley thermostable α-amylase from the hypterthermophilic archaebacterium *Pyrococcus furiosus*, J. Biol. Chem., 268, 24 394–24 401, 1993.
Lepori, L. and Gianni, P.: Partial molar volumes of ionic and nonionic organic solutes in water: A simple additivity scheme based on the intrinsic volume approach, J. Solution Chem., 29, 405–447, 2000.
Li, J., Wang, X. G., Klein, M. T., and Brill, T. B.: Spectroscopy of hydrothermal reactions, 19: pH and salt dependence of decarboxylation of alpha-alanine at 280-330 degrees C in an FT-IR

spectroscopy flow reactor, Int. J. Chem. Kinet., 34, 271–277, 2002. 1546 1545 1563 1543, 1589, 1609 1519, 1542 1545

Makhatadze, G. I. and Privalov, P. L.: Heat capacity of alcohols in aqueous solutions in the temperature range from 5° to 125°C, J. Solution Chem., 18, 927–936, 1989. 1611

Makhatadze, G. I. and Privalov, P. L.: Heat capacity of proteins. 1. Partial molar heat capacity of individual amino acid residues in aqueous solution: Hydration effect, J. Mol. Biol., 213, 375–384, 1990.

Makhatadze, G. I., Medvedkin, V. N., and Privalov, P. L.: Partial molar volumes of polypeptides and their constituent groups in aqueous solution over a broad temperature range, Biopolymers, 30, 1001–1010, 1990. 1519, 1524, 1543, 1582, 1601 1519, 1582, 1589

Makhatadze, G. I., Kim, K. S., Woodward, C., and Privalov, P. L.: Thermodynamics of BPTI folding, Protein Sci., 2, 2028–2036, 1993. 1589

Makhatadze, G. I., Clore, G. M., Gronenborn, A. M., and Privalov, P. L.: Thermodynamics of unfolding of the all β-sheet protein interleukin-1β, Biochemistry, 33, 9327–9332, 1994. 1589

Mallick, P., Boutz, D. R., Eisenberg, D., and Yeates, T. O.: Genomic evidence that the intracellular proteins of archaeal microbes contain disulfide bonds, Proc. Natl. Acad. Sci. U.S.A., 99, 9679–9684, 2002.

Marriott, R. A., Hakin, A. W., and Liu, J. L.: Modeling of thermodynamic properties of amino acids and peptides using additivity and HKF theory, J. Solution Chem., 27, 771–802, 1998. 1546 1519, 1549, 1582, 1596, 1597

Marriott, R. A., Hakin, A. W., Liu, J. L., and Lutter, E.: The volumetric properties of aqueous solutions of glycylglycine and L-serine at elevated temperatures and pressures, J. Chem. Thermodyn., 33, 959–982, 2001. 1597

Millero, F. J., Lo Surdo, A., and Shin, C.: Apparent molal volumes and adiabatic compressibilities of aqueous amino acids at 25°C, J. Phys. Chem., 82, 784–792, 1999. 1597, 1598

Mizuguchi, M., Sakurai, M., and Nitta, K.: Partial molar volumes and adiabatic compressibilities of *N*-acetyl-*DL*-serinamide and *N*-acetyl-*L*-threoninamide in dilute aqueous solutions, J. Solution Chem., 26, 579–594, 1997. 1597, 1598

Mrevlishvili, G. M.: Thermodynamic data for biochemistry and biotechnology, chap. 5, pp. 148–176, Springer-Verlag, 1986.

Neumann, H., Goldberger, R. F., and Sela, M.: Interaction of phosphorothioate with the disulfide bonds of ribonuclease and lysozyme, J. Biol. Chem., 239, 1536–1540, 1964. 1560, 1615 1537

Nichols, N., Sköld, R., Spink, C., and Wadsö, I.: Thermochemistry of solutions of biochemical model, 6. α,ω-dicarboxylic acids, -diamines, and -diols in aqueous solution, J. Chem. Thermodyn., 8, 993–999, 1976. 1611

Ngauv, S. N., Sabbah, R., and Laffitte, M.: Thermodynamics of nitrogen compounds, 3. Thermochemical study of glycine and L-alpha-alanine, Thermochim. Acta, 20, 371–380, 1977. 1579, 1593

Nozaki, Y. and Tanford, C.: Intrinsic dissociation constants of aspartyl and glutamyl carboxyl groups, J. Biol. Chem., 242, 4731–4735, 1967a.

Nozaki, Y. and Tanford, C.: Proteins as random coils. II. Hydrogen ion titration curve of ribonuclease in 6M guanidine hydrochloride, J. Am. Chem. Soc., 89, 742–749, 1967b.

Plyasunov, A. V. and Shock, E. L.: Correlation strategy for determining the parameters of the revised Helgeson-Kirkham-Flowers model for aqueous nonelectrolytes, Geochim. Cosmochim. Acta, 65, 3879–3900, 2001.

Plyasunova, N., Plyasunov, A., and Shock, E.: Group contribution values for the thermodynamic functions of hydration at 298.15 K, 0.1 MPa. 2. Aliphatic thiols, alkyl sulfides, and polysulfides, J. Chem. Eng. Dat., 50, 246–253, 2005. 1530 1544, 1608 1527, 1579, 1593 1545

Price, J. L., Jardine, J .J., Call, T. G., Patterson, B. A., Origlia-Luster, M. L., and Woolley, E. M.: Thermodynamics for proton dissociations from aqueous L- histidine at temperatures from 278.15 to 393.15 K and at the pressure 0.35 MPa: apparent molar volumes and apparent molar heat capacities of the protonated cationic, neutral zwitterionic, and deprotonated anionic forms, J. Chem. Thermodyn., 35, 195–198, 2003a. 1596, 1597

Price, J. L., Sorenson, E. C., Merkley, E. D., McRae, B. R., and Woolley, E. M.: Thermodynamics of proton dissociations from aqueous L-valine and L-2-amino-*n*-butanoic acid: apparent molar volumes and apparent molar heat capacities of the protonated cationic, neutral zwitterionic, and deprotonated anionic species at temperatures from $278.15 \leq T/\mathrm{K} \leq 393.15$, at molalities $0.015 \leq m/\mathrm{mol \cdot kg^{-1}} \leq 0.67$, and pressure p=0.35 MPa, J. Chem. Thermodyn., 35, 1425–1467, 2003b.

Privalov, P. L. and Makhatadze, G. I.: Heat capacity of proteins. 2. Partial molar heat capacity of the unfolded polypeptide chain of proteins – protein unfolding effects, J. Mol. Biol., 213, 385–391, 1990.

Privalov, P. L., Tiktopulo, E. I., Venyaminov, S. Y., Griko, Y. V., Makhatadze, G. I., and Khechinashvili, N. N.: Heat capacity and conformation of proteins in the denatured state, J. Mol.

Biol., 205, 737–750, 1989.
Richard, L. and Helgeson, H. C.: Calculation of the thermodynamic properties at elevated temperatures and pressures of saturated and aromatic high molecular weight solid and liquid hydrocarbons in kerogen, bitumen, petroleum, and other organic matter of biogeochemical interest, Geochim. Cosmochim. Acta, 62, 3591–3636, 1998. 1526, 1596, 1597, 1610 1519, 1589 1532 1560, 1613
Roxby, R. and Tanford, C.: Hydrogen ion titration curve of lysozyme in 6 M guanidine hydrochloride, Biochemistry, 10, 3348–3352, 1971. 1608
Sabbah, R. and Laffitte, M.: Thermodynamique de composés azotes. IV. Étude thermochimique de la sarcosine et de la L-proline, Bull. Soc. Chim. Fr., 1, 50–52, 1978. 1579
Schafer, F. Q. and Buettner, G. R.: Redox environment of the cell as viewed through the redox state of the glutathione disulfide/glutathione couple, Free Radic. Biol. Med., 30, 1191–1212, 2001.
Schulte, M. D. and Rogers, K. L.: Thiols in hydrothermal solution: standard partial molal properties and their role in the organic geochemistry of hydrothermal environments, Geochim. Cosmochim. Acta, 68, 1087–1097, 2004.
Schulte, M. D., Shock, E. L., and Wood, R. H.: The temperature dependence of the standard-state thermodynamic properties of aqueous nonelectrolytes, Geochim. Cosmochim. Acta, 65, 3919–3930, 2001.
Schulze, W. X.: Protein analysis in dissolved organic matter: What proteins from organic debris, soil lechate and surface water can tell us – a perspective, Biogeosciences, 2, 75–86, 2005, SRef-ID: 1726-4189/bg/2005-2-75. 1517 1545 1523 1517
Scott, D. W., Berg, W. T., Hossenlop, I. A., Hubbard, W. N., Messerly, J. F., Todd, S. S., Douslin, D. R., McCullough, J. P., and Waddington, G.: Pyrrole: Chemical thermodynamic properties, J. Phys. Chem., 71, 2263–2270, 1967. 1593
Shock, E. L.: Stability of peptides in high-temperature aqueous solutions, Geochim. Cosmochim. Acta, 56, 3481–3491, 1992. 1531, 1539
Shock, E. L.: Organic acids in hydrothermal solutions – standard molal thermodynamic properties of carboxylic acids and estimates of dissociation constants at high temperatures and pressures, Am. J. Sci., 295, 496–580, 1995.
Shock, E. L., Oelkers, E. H., Johnson, J. W., Sverjensky, D. A., and Helgeson, H. C.: Calculation of the thermodynamic properties of aqueous speceis at high pressures and temperatures: Effective electrostatic radii, dissociation constants, and standard partial molal properties to

1000°C and 5 kb., J. Chem. Soc., Faraday Trans., 88, 803–826, 1992.
Shock, E. L., Sassani, D. C., Willis, M., and Sverjensky, D. A.: Inorganic species in geologic fluids: Correlations among standard molal thermodynamic properties of aqueous ions and hydroxide complexes, Geochim. Cosmochim. Acta, 61, 907–950, 1997.
Shvarov, Y. V. and Bastrakov, E. N.: HCh: A software package for geochemical equilibrium modelling, Australian Geological Survey Organisation, record 1999/25, 1999.
Sorenson, E. C., Price, J. L., McRae, B. R., and Woolley, E. M.: Thermodynamics of proton dissociations from aqueous L-proline: apparent molar volumes and apparent molar heat capacities of the protonated cationic, zwitterionic, and deprotonated anionic forms at temperatures from 278.15 K to 393.15 K and at the pressure 0.35 MPa, J. Chem. Thermodyn., 35, 529–553, 2003.
Souillac, P. O., Costantino, H. R., Middaugh, C. R., and Rytting, J. H.: Investigation of protein/carbohydrate interactions in the dried state. 1. Calorimetric studies, J. Pharm. Sci., 91, 206–216, 2002. 1582, 1610 1551 1526 1546 1526, 1597, 1610 1562, 1614
Spink, C. H. and Wädso, I.: Thermochemistry of solutions of biochemical model compounds, 4. Partial molar heat capacities of some amino acids in aqueous solution, J. Chem. Thermodyn., 7, 561–572, 1969. 1596
Steinhardt, J. and Reynolds, J. A.: Multiple Equilibria in Proteins, Academic Press, New York, 1969. 1586
Tan, Y. J., Oliveberg, M., Davis, B., and Fersht, A. R.: Perturbed pK(a) values in the denatured states of proteins, J. Mol. Biol., 254, 980–992, 1995.
Tanger, J. C. I. and Helgeson, H. C.: Calculation of the thermodynamic and transport properties of aqueous species at high pressures and temperatures: Revised equations of state for the standard partial molal properties of ions and electrolytes, Am. J. Sci., 288, 19–98, 1988.
Tollinger, M., Crowhurst, K. A., Kay, L. E., and Forman-Kay, J. D.: Site-specific contributions to the pH dependence of protein stability, Proc. Natl. Acad. Sci. USA, 100, 4545–4550, 2003.
Tsuzuki, T., Harper, D. O., and Hunt, H.: Heats of combustion. VII. The heats of combustion of some amino acids, J. Phys. Chem., 62, 1594, 1958.
Tyson, G. W., Chapman, J., Hugenholtz, P., Allen, E. E., Ram, R. J., Richardson, P. M., Solovyev, V. V., Rubin, E. M., Rokhsar, D. S., and Banfield, J. F.: Community structure and metabolism through reconstruction of microbial genomes from the environment, Nature, 428, 37–43, 2004.
Vant, S. C., Glen, N. F., Kontopidis, G., Sawyer, L., and Schaschke, C. J.: Volumetric changes

to the molecular structure of beta-lactoglobulin processed at high pressure, High Temp.-High Press., 34, 705–712, 2002.

Vogl, T., Hinz, H.-J., and Hedwig, G. R.: Partial molar heat capacities and volumes of Gly–X–Gly tripeptides in aqueous solution: model studies for the rationalization of thermodynamic parameters of proteins, Biophys. Chem., 54, 261–269, 1995.

Wagman, D. D., Evans, W. H., Parker, V. B., Schumm, R. H., Halow, I., Bailey, S. M., Churney, K. L., and Nuttall, R. L.: The NBS tables of chemical thermodynamic properties. Selected values for inorganic and C_1 and C_2 organic substances in SI units, J. Phys. Chem. Ref. Data, 11, 2–1–2–392, 1982.

Wang, P., Oscarson, J. L., Gillespie, S. E., Izatt, R. M., and Cao, H.: Thermodynamics of protonation of amino acid carboxylate groups from 50 to 125 ° C, J. Solution Chem., 25, 243–266, 1996.

Whitten, S. T. and García-Moreno E., B.: pH dependence of stability of staphylococcal nuclease: evidence of substantial electrostatic interactions in the denatured state, Biochemistry, 39, 14 292–14 304, 2000. 1545 1517, 1549, 1551 1530, 1586 1561 1517 1563 1518, 1601, 1602 1531 1525, 1526, 1610 1544, 1608

Yasuda, Y., Tochio, N., Sakurai, M., and Nitta, K.: Partial molar volumes and isentropic compressibilities of amino acids in dilute aqueous solutions, J. Chem. Eng. Data, 43, 205–214, 1998. 1597, 1598

Yu, Y., Makhatadze, G. I., Pace, C. N., and Privalov, P. L.: Energetics of ribonuclease T1 structure, Biochemistry, 33, 3312–3319, 1994. 1589

Zhang, G., Gerdes, S., and Wunderlich, B.: Heat capacities of solid, globular proteins, Macromol. Chem. Phys., 197, 3791–3806, 1996. 1615

Table 1. Symbols and abbreviations used in the text.

Symbol(s)	Definition
[*SC*]	Abbreviation for any sidechain group.
[AABB], [AABB$^+$], [AABB$^-$]	Abbreviations for the neutral zwitterionic, positively charged and negatively charged amino acid backbone group.
[Ala], [Arg], [Arg$^+$], . . . , [Val]	3-letter abbreviations, including formal charge, for sidechain groups.
[GXGBB]	Abbreviation for the Gly–X–Gly backbone group, which is used here to refer to what remains upon removal of the middle (X) sidechain group.
[UPBB], [PPBB]	Abbreviations for the protein backbone group of crystalline proteins and the aqueous unfolded protein and polypeptide backbone groups, all of which share the formula C_2H_2NO.
$\Delta C^\circ_{P,\mathrm{n}}$, $\Delta V^\circ_\mathrm{n}$, $\Delta \kappa^\circ_{T,\mathrm{n}}$	Standard molal non-solvation isobaric heat capacity, volume and isothermal compressibility.
$\Delta C^\circ_{P,\mathrm{s}}$, $\Delta V^\circ_\mathrm{s}$, $\Delta \kappa^\circ_{T,s}$	Standard molal solvation isobaric heat capacity, volume and isothermal compressibility.
ΔG°, ΔH°, S°	Standard molal Gibbs free energy, enthalpy, and entropy at P and T.
ΔG°_f, ΔH°_f, $S^\circ_{P_r,T_r}$	Standard molal Gibbs free energy and enthalpy of formation, and third law entropy, at P_r and T_r.
ΔS°_{hyd}	Standard molal entropy of hydration.
ΔG°_{ion}	Standard molal Gibbs free energy of ionization.
Ξ	Any standard molal thermodynamic property or revised HKF parameter of a given species.
AA	Abbreviation for any amino acid.
A, C, D, . . . ,X, Y.	Conventional 1-letter abbreviations for amino acids or sidechain groups.
Ala, Arg, Asn, . . . , Val	Conventional 3-letter abbreviations for amino acids.
Ala, Arg, Arg$^+$, Arg$^-$, . . .	3-letter abbreviations for the amino acids, including formal charge.
GXG	Abbreviation for any Gly–X–Gly tripeptide (X represents any sidechain group).
P_r, T_r	Reference pressure (1 bar) and temperature (25°C).

BGD
2, 1515–1615, 2005

Table 1. Continued.

Symbol(s)	Definition
P_{sat}	Reference geothermal gradient of pressure, corresponding to P_r at temperatures < 100°C and the saturation vapor pressure of pure H_2O at temperatures ≥ 100°C.
σ, ξ	Volumetric non-solvation parameters in the revised HKF equations of state.
ω	Solvation parameter in the revised HKF equations of state.
c_1, c_2, a_1, a_2, a_3, a_4	Caloric non-solvation parameters in the revised HKF equations of state.
n, n_i	Total number of sidechain groups (length) of a peptide; number of occurrences of the ith sidechain or organic group in a group additivity equation.
pK	Negative logarithm of the equilibrium constant of the deprotonation reaction of a group or species.

BGD
2, 1515–1615, 2005

Table 2. ΔG_f°, ΔH_f° and $S^\circ_{P_r,T_r}$ of aqueous and gaseous amino acids. Values of ΔG_f°, ΔH_f° and $S^\circ_{P_r,T_r}$ of aqueous amino acids in different ionization states are taken from Amend and Helgeson (1997a). Unles otherwise noted, the values for gaseous amino acids are from additivity calculations of ΔH_f° and $S^\circ_{P_r,T_r}$, which are combined with $S^\circ_{P_r,T_r}$ of the elements to give ΔG_f°.

	ΔG_f° kcal mol^{-1}	ΔH_f° kcal mol^{-1}	$S^\circ_{P_r,T_r}$ cal mol^{-1} K^{-1}		ΔG_f° kcal mol^{-1}	ΔH_f° kcal mol^{-1}	$S^\circ_{P_r,T_r}$ cal mol^{-1} K^{-1}		ΔG_f° kcal mol^{-1}	ΔH_f° kcal mol^{-1}	$S^\circ_{P_r,T_r}$ cal mol^{-1} K^{-1}
Aqueous amino acids				Aqueous amino acids				Aqueous amino acids			
Ala^+	-92.01	-133.08	47.61	His^{2+}	-58.89	-115.90	75.30	Trp^+	-30.07	-97.25	72.00
Ala	-88.81	-132.50	38.83	His^+	-56.57	-115.20	69.86	Trp	-26.82	-97.59	59.96
Ala^-	-75.36	-121.47	30.71	His	-48.42	-108.20	66.00	Trp^-	-14.01	-85.66	57.00
				His^-	-35.76	-97.70	58.76				
Arg^{2+}	-72.41	-155.86	87.60					Tyr^+	-94.80	-157.45	70.45
Arg^+	-69.93	-154.88	82.57	Ile^+	-85.15	-151.58	60.37	Tyr	-91.80	-157.74	59.41
Arg	-57.36	-143.06	80.06	Ile	-81.99	-151.60	49.70	Tyr^-	-79.38	-149.42	45.66
Arg^-	-40.34	-129.42	68.72	Ile^-	-68.68	-140.52	42.23	Tyr^{2-}	-65.65	-140.18	30.60
Asn^+	-128.25	-183.44	64.66	Leu^+	-87.38	-153.96	59.87	Val^+	-88.51	-146.59	55.77
Asn	-125.49	-182.70	57.88	Leu	-84.20	-153.60	50.41	Val	-85.33	-146.42	45.68
Asn^-	-114.47	-172.85	53.96	Leu^-	-70.90	-142.27	43.80	Val^-	-72.08	-138.36	28.27

BGD
2, 1515–1615, 2005

Table 2. Continued.

	ΔG_f° kcal mol^{-1}	ΔH_f° kcal mol^{-1}	$S^\circ_{P_r,T_r}$ cal mol^{-1} K^{-1}
Aqueous amino acids			
Asp^+	-175.23	-227.14	61.65
Asp	-172.51	-226.34	55.22
Asp^-	-167.17	-224.54	43.34
Asp^{2-}	-153.52	-209.10	49.34
Cys^+	-82.66	-124.61	52.32
Cys	-80.33	-124.07	46.32
Cys^-	-68.97	-117.08	31.66
Cys^{2-}	-54.26	-108.04	12.65
Gln^+	-129.24	-192.53	70.10
Gln	-126.28	-191.86	62.41
Gln^-	-113.82	-182.11	53.33
Glu^+	-176.00	-234.99	70.53
Glu	-173.05	-234.82	61.20
Glu^-	-167.21	-234.15	43.86
Glu^{2-}	-154.44	-217.28	57.61
Gly^+	-94.16	-125.72	46.91
Gly	-90.95	-124.78	39.29
Gly^-	-77.61	-114.19	30.07
Aqueous amino acids			
Lys^{2+}	-95.86	-170.40	71.68
Lys^+	-92.89	-170.37	61.82
Lys	-80.68	-159.33	57.91
Lys^-	-66.31	-146.34	53.27
Met^+	-123.23	-179.14	70.71
Met	-120.12	-178.52	62.36
Met^-	-107.55	-167.98	55.55
Phe^+	-52.95	-109.04	67.93
Phe	-49.43	-108.90	56.60
Phe^-	-36.76	-98.01	50.63
Pro^+	-76.22	-124.17	58.52
Pro	-73.56	-124.10	49.83
Pro^-	-59.05	-113.84	35.58
Ser^+	-126.90	-172.74	56.16
Ser	-123.92	-172.42	47.24
Ser^-	-111.36	-162.71	37.68
Thr^+	-122.69	-179.31	52.58
Thr	-119.83	-178.94	44.23
Thr^-	-107.41	-169.11	35.54
Gaseous amino acids			
Ala[a]	-71.91	-99.10	94.17
Arg	2.14	-69.05	128.73
Asn	-98.81	-141.13	107.84
Asp	-151.04	-187.80	112.47
Cys	-64.33	-90.08	106.65
Gln	-96.81	-146.06	117.20
Glu	-149.04	-192.73	121.83
Gly[a]	-73.35	-93.30	85.84
His	-14.31	-62.60	104.53
Ile	-66.10	-114.41	121.15
Leu	-66.10	-114.41	121.15
Lys	-49.47	-105.95	132.25
Met	-59.13	-98.34	126.73
Phe	-32.33	-71.95	123.16
Pro[a]	-55.66	-87.52	112.48
Ser	-103.12	-134.94	103.16
Thr	-104.79	-143.42	112.92
Trp	1.48	-51.13	120.88
Tyr	-69.87	-114.70	130.20
Val	-68.10	-109.48	111.79

[a] Values of ΔG_f° of these amino acids are calculated from ΔG°_{hyd} (Plyasunov and Shock, 2001) and ΔG_f° of the corresponding aqueous amino acids. Values of ΔH_f° are taken from Ngauv et al. (1977) and Sabbah and Laffitte (1978). Values of $S^\circ_{P_r,T_r}$ are calculated from from ΔG_f°, ΔH°, and $S^\circ_{P_r,T_r}$ of the elements.

BGD
2, 1515–1615, 2005

Table 3. Values and estimated uncertainties of revised HKF equations of state parameters of amino acids.

Amino Acid	$\omega \times 10^{-5}$ cal mol^{-1}	c_1 cal mol^{-1} K^{-1}	$c_2 \times 10^{-4}$ cal mol^{-1} K	$a_1 \times 10$ cal mol^{-1} bar^{-1}	$a_2 \times 10^{-2}$ cal mol^{-1}	a_3 cal K mol^{-1} bar^{-1}	$a_4 \times 10^{-4}$ cal mol^{-1} K
Ala	0.18	49.5	−7.00	14.90	1.74	7.16	−3.69
Arg	0.22	94.8	−12.50	28.83	8.21	7.20	−5.95
Arg$^+$	0.73	103.8	−9.60	31.72	0.22	5.46	−5.15
Arg$^-$	2.30	141.9	−19.60	31.28	13.76	32.12	−9.94
Asn	0.21	56.5	−11.70	19.83	2.37	3.76	−4.81
Asp	2.47	56.3	−15.30	16.96	5.77	10.11	−6.37
Asp$^-$	0.17	58.1	−11.90	18.96	3.05	−6.62	−2.23
Cys	0.15	59.8	−5.60	18.14	2.40	9.31	−4.97
Cys$^-$	2.59	61.0	−9.50	14.44	5.08	12.24	−3.68
Glu	0.15	65.2	−10.00	22.30	4.23	6.55	−5.84
Glu$^-$	2.65	47.6	−7.80	20.49	6.96	10.86	−7.02
Gln	0.18	68.6	−11.40	23.22	3.58	5.44	−5.23
Gly	0.23	28.5	−8.40	11.30	0.71	3.99	−3.04
His	0.27	81.6	−11.30	24.32	4.80	7.52	−6.05
His$^+$	−0.50	78.0	−12.00	26.42	−4.06	−0.88	−3.07
Ile	0.09	99.7	−3.60	24.49	6.55	18.16	−7.76
Leu	0.09	102.7	−3.30	24.68	7.51	19.93	−8.37
Lys	0.07	89.5	−10.90	24.56	8.30	30.81	−9.53
Lys$^+$	1.21	95.2	−8.00	28.29	−1.51	4.47	−3.99
Lys$^-$	2.67	133.7	−18.00	28.29	10.53	2.97	−7.61
Met	0.13	85.3	−6.60	24.95	6.90	13.59	−7.77
Phe	0.12	108.1	−6.50	28.27	8.80	19.33	−9.38
Pro	0.14	63.6	−9.50	19.39	4.87	11.88	−5.68
Ser	0.18	48.5	−9.40	15.69	0.73	3.87	−3.49
Thr	0.11	65.5	−7.10	18.94	2.83	8.87	−4.87
Trp	0.15	116.8	−7.80	33.91	9.21	14.78	−9.36
Tyr	0.09	106.2	−12.40	30.06	8.49	8.36	−8.58
Tyr$^-$	2.78	109.0	−16.00	26.34	10.49	16.42	−7.59
Val	0.12	83.8	−4.30	21.35	4.48	14.58	−6.12
δ_{AA}	±0.1	±1.0	±1.20	±0.1	±0.48	±0.47	±0.81

BGD
2, 1515–1615, 2005

Table 4. Revised HKF equations of state parameters of Gly–X–Gly tripeptides and [GXGBB].

Tripeptide	$\omega^a \times 10^{-5}$	c_1^b	$c_2^c \times 10^{-4}$	σ^e	$\xi^f \times 10^{-2}$
Ala	−1.59	85.8	−15.80	130.7	−3.20
Arg^+	−1.04	116.0	−20.00	192.0	−7.00
Asn	−1.56	89.0	−18.00	157.9	−8.52
Asp	−1.60	78.9	−13.80	147.9	−6.15
Cys	−1.62	89.0	−11.80	147.3	−5.71
Glu	−1.62	94.6	−15.50	164.5	−6.39
Gln	−1.59	100.5	−18.80	167.3	−6.45
Gly	−1.54	63.9	−16.60	112.2	−3.10
His	−1.50	151.9	−36.10	178.8	−7.69
Ile	−1.68	122.6	−6.80	188.4	−11.34
Leu	−1.68	129.4	−9.60	184.8	−9.18
Lys^+	−0.56	180.0	−36.00	189.0	−9.50
Met	−1.64	105.9	−9.40	190.2	−12.43
Phe	−1.65	128.4	−11.30	204.1	−10.90
Pro	−1.63	91.5	−17.80	149.1	−5.93
Ser	−1.59	77.5	−15.70	130.3	−3.14
Thr	−1.66	99.3	−16.50	149.0	−5.45
Tyr	−1.68	126.8	−17.20	203.9	−9.58
Val	−1.65	111.8	−10.80	164.4	−6.08
δ	±0.10	±1.0	±1.20	±1.2	±1.50
[GXGBB]	−1.54	52.6	−17.6	101.2	−1.59

[a] cal mol^{-1}.
[b] cal mol^{-1} K^{-1}.
[c] cal mol^{-1} K.
[d] 25°C and 1 bar.
[e] cm^3 mol^{-1}.
[f] cm^3 mol^{-1} K.

BGD
2, 1515–1615, 2005

Table 5. Reference model compound algorithms for estimating C_P° and V° of amino acids as a function of temperature.

Amino Acid	Property or Parameter	Reference Model Compound Algorithm
Asp^-	C_P°	Asp[a] + propanoate[b] – propanoic acid[b]
Cys	C_P°	Ser[a] + Gly–Cys–Gly[c] – Gly–Ser–Gly[c]
Cys^-	C_P°, V°	Cys + propanoate[b] – propanoic acid[b]
Glu^-	C_P°	Glu[a] + butanoate[b] – butanoic acid[b]
Lys^-	C_P°, V°	Lys + *n*-butanamine[d] – *n*-butanamine^{+} [d]
Tyr	C_P°	Phe[e] + Gly–Tyr–Gly[f] – Gly–Phe–Gly[f]
Tyr^-	C_P°, V°	Tyr + propanoate[b] – propanoic acid[b]

[a] Hakin et al. (1994a).
[b] Shock (1995).
[c] Häckel et al. (1998).
[d] C_P°: Makhatadze and Privalov (1990); V°: Makhatadze et al. (1990).
[e] Marriott et al. (1998).
[f] Häckel et al. (1999a).

Table 6. Correlations among the properties of amino acids: summary of equations for calculating ω , a_2 (cal mol^{-1}) and a_4 (cal mol^{-1} K) from ΔS°_{hyd} (cal mol^{-1} K^{-1}) and V° (cm^3 mol^{-1}) at 25°C and 1 bar.

	Equation	Z_{AA}
6.1	$\omega \times 10^{-5} = 0.49 + 0.0056 \times \Delta S^\circ_{hyd}$	0
6.2	$\omega \times 10^{-5} = 1.12 - 0.0196 \times \Delta S^\circ_{hyd}$	−1
6.3	$\omega \times 10^{-5} = -0.17 - 0.0196 \times \Delta S^\circ_{hyd}$	+1
6.4	$a_2 \times 10^{-2} = -4.80 + 0.105 \times V^\circ$	0
6.5	$a_2 \times 10^{-2} = -1.25 + 0.105 \times V^\circ$	−1
6.6	$a_2 \times 10^{-2} = -13.50 + 0.105 \times V^\circ$	+1
6.7	$a_4 \times 10^{-4} = -2.78 - 0.721 a_2 \times 10^{-2}$	0
6.8	$a_4 \times 10^{-4} = -2.18 - 0.721 a_2 \times 10^{-2}$	−1
6.9	$a_4 \times 10^{-4} = -5.48 - 0.721 a_2 \times 10^{-2}$	+1

BGD
2, 1515–1615, 2005

Table 7. C_P°, V° and κ_T° of amino acids and C_P° and V° of Gly–X–Gly tripeptides at 25°C and 1 bar. Values are not given for tripeptides for which no calorimetric or volumetric data are available in the literature.

	GXG tripeptides			Amino Acids		
	C_P° [a]	V° [b]		C_P° [a]	V° [b]	κ_T° [c]
Ala	68.2	130.6		33.6	60.4	−23.46
Arg	–	–		67.4	123.7	−3.74
Arg^+	84.8	184.9		77.6	122.4	−50.88
Arg^-	–	–		81.1	143.0	−24.08
Asn	66.6	150.1		30.8	77.3	−29.72
Asp	65.4	143.6		2.7	65.0	−3.39
Asp^-	–	–		32.4	74.9	−43.62
Cys	79.8	143.7		47.1	73.3	−30.38
Cys^-	–	–		18.1	60.3	−25.29
Glu	77.9	159.9		43.5	90.2	−26.73
Glu^-	–	–		7.6	80.0	−43.68
Gln	76.8	162.5		43.8	93.6	−25.65
Gly	44.2	112.1		9.3	43.2	−24.56
His	92.2	172.0		56.2	99.3	−26.18
His^+	–	–		58.2	97.7	−56.95
Ile	124.1	176.9		91.6	105.8	−28.75
Leu	125.2	176.4		95.2	107.9	−28.19
Lys	–	–		66.7	108.8	−33.34
Lys^+	111.9	177.0		67.9	107.4	−55.85
Lys^-	–	–		72.8	112.2	−26.94
Met	101.7	177.0		70.7	105.6	−27.05
Phe	120.5	193.1		93.8	122.2	−29.40
Pro	70.2	145.2		43.0	82.7	−21.27
Ser	60.1	130.2		27.8	60.7	−27.94
Thr	80.9	145.8		50.1	77.4	−26.46
Trp	–	–		99.6	143.6	−26.96
Tyr	107.2	194.9		80.2	124.7	−23.98
Tyr^-	–	–		51.2	111.7	−28.09
Val	104.9	160.3		74.0	90.7	−27.37
δ_{GXG}	±4.4	±3.6	δ_{AA}	±4.4	±3.6	±11.05

[a] cal mol^{-1} K^{-1}.
[b] cm^3 mol^{-1}.
[c] cm^3 bar^{-1} mol^{-1}.

BGD
2, 1515–1615, 2005

Table 8. ΔG_f°, ΔH_f° and S° of aqueous amino acid sidechain and backbone groups and organic groups and ionized protein sidechain groups and unfolded protein backbone groups.

Group	$\Delta G_f^{\circ a}$	$\Delta H_f^{\circ a}$	$S^{\circ b}$	Group	$\Delta G_f^{\circ a}$	$\Delta H_f^{\circ a}$	$S^{\circ b}$	Group	$\Delta G_f^{\circ a}$	$\Delta H_f^{\circ a}$	$S^{\circ b}$
Amino Acid Sidechain Groups				**Amino Acid Sidechain Groups**				**Protein Backbone and Terminal Groups**			
[Ala]	−3.94	−13.29	16.85	[Ser]	−39.05	−53.21	25.26	[AABB$^+$]	−89.94	−119.58	37.74
[Arg]	31.63	−20.73	54.74	[Thr]	−34.96	−59.73	22.25	[AABB$^-$]	−75.07	−108.69	24.38
[Arg$^+$]	14.94	−35.67	60.59	[Trp]	58.05	21.62	37.98	[UPBB]	−21.45	−45.22	1.62
[Asn]	−40.61	−63.49	35.9	[Tyr]	−6.92	−38.53	37.43				
[Asp]	−87.64	−107.13	33.24	[Tyr$^-$]	5.50	−30.21	23.68	**Organic Groups**[d]			
[Asp$^-$]	−82.29	−105.33	21.36	[Val]	−0.46	−27.21	23.7	[−CH$_2$−]	2.24	−5.67	6.07
[Cys]	4.54	−4.86	24.34					[−CH$_3$]	−2.19	−12.46	13.78
[Cys$^-$]	15.91	2.13	9.68	**Amino Acid Backbone Groups**				[−CHCH$_3$−]	3.47	−12.48	11.71
[Gln]	−41.40	−72.65	40.43	[AABB]	−84.87	−119.21	21.98	[−CH$_2$NH$_2$]	7.04	−12.43	21.44
[Glu]	−88.18	−115.61	39.22	[AABB$^+$]	−88.00	−119.58	31.23	[−CH$_2$OH]	−41.53	−56.79	21.57
[Glu$^-$]	−82.34	−114.94	21.88	[AABB$^-$]	−71.97	−108.69	13.97	[−CONH$_2$]	−46.55	−61.42	30.14
[Gly]	−6.07	−5.57	17.31					[−COOH]	−92.57	−104.38	26.41
[His]	36.46	11.01	44.02	**Protein Sidechain Groups**[c]				[−C$_6$H$_5$]	31.66	15.17	31.02
[His$^+$]	28.31	4.01	47.88	[Arg$^+$]	15.59	−35.67	58.41	[di.corr]	0.00	0.00	0.00
[Ile]	2.89	−32.39	27.72	[Asp$^-$]	−82.06	−105.33	20.59				
[Leu]	0.67	−34.39	28.43	[Cys$^-$]	15.88	2.13	9.78	**Representative Uncertainties**			
[Lys]	5.66	−37.65	39.29	[Glu$^-$]	−82.04	−114.94	20.87	δ_{AA}	±0.50	±0.50	±0.50
[Lys$^+$]	−8.01	−51.16	39.84	[His$^+$]	27.86	4.01	49.39	$\delta_{[SC]}$	±1.45	±1.89	±1.85
[Met]	−35.25	−59.31	40.38	[Lys$^+$]	−7.84	−51.16	39.24	$\delta_{[UPBB]}$	±1.45	±2.82	±7.82
[Phe]	35.44	10.31	34.62	[Tyr$^-$]	6.18	−30.21	21.4				
[Pro]	11.31	−4.89	27.85								

[a] kcal mol^{-1}.

[b] cal mol^{-1} K^{-1}.

[c] The values of the properties of any neutral protein sidechain group are taken to be equal to those of the corresponding amino acid sidechain group.

[d] Amend and Helgeson (1997b) (except [di.corr]).

BGD

2, 1515–1615, 2005

Table 9. ΔG°_{ion} of ionization reactions and pK at 25°C and 1 bar of sidechain and amino acid backbone groups in amino acids and unfolded proteins.

	Amino Acids		Unfolded Proteins	
Ionized group	ΔG°_{ion} [a,b]	pK [b]	ΔG°_{ion} [a,c]	pK [c]
[Asp$^-$]	5.34	3.91	5.57	4.08
[Glu$^-$]	5.84	4.28	6.14	4.50
[His$^+$]	−8.15	5.97	−8.60	6.30
[Lys$^+$]	−14.37	10.53	−14.19	10.40
[Arg$^+$]	−17.02	12.47	−16.37	12.00
[Cys$^-$]	11.36	8.33	11.33	8.30
[Tyr$^-$]	12.42	9.10	13.10	9.60
[AABB$^+$]	−2.70	1.98	−4.64	3.40
[AABB$^-$]	13.33	9.77	10.23	7.50
δ	±0.50	±0.37	±0.26	±0.19

[a] kcal mol^{-1}.

[b] Values of ΔG°_{ion} of the sidechain and backbone groups of amino acids are calculated from corresponding values of ΔG°_f taken from Table 2. The uncertainty of ΔG°_{ion} of amino acids is estimated to be equal to the uncertainty in the value of ΔG°_f. The values and uncertainty of pK of the corresponding deprotonation reactions of amino acids are calculated from those of ΔG°_{ion} and $\Delta G^\circ_{ion} = -2.303RT \log K$.

[c] pKs of reference model compounds for the deprotonation reactions of sidechain and amino acid backbone groups in aqueous proteins are taken from Steinhardt and Reynolds (1969). The estimated uncertainty of pKa arising from interactions in the unfolded protein is calculated by halving the range of values measured by Tollinger et al. (2003) for deprotonation reactions of [Asp] and [Glu] in. Using the values and uncertainty of pKa of groups in unfolded proteins, those of ΔG°_{ion} are calculated from $\Delta G^\circ_{ion} = -2.303RT \log K$.

Table 10. Revised HKF equations of state parameters and C_P°, V° and κ_T° at 25 °C and 1 bar of aqueous sidechain, backbone and organic groups, and their estimated uncertainties.

Group	c_1^a	$c_2^b \times 10^{-4}$	$a_1^c \times 10$	$a_2^d \times 10^{-2}$	a_3^e	$a_4^b \times 10^{-4}$	$\omega^d \times 10^{-5}$	$C_P^{\circ a}$	$V^{\circ f}$	$\kappa_T^\circ \times 10^4$
Amino Acid and Protein Sidechain Groups										
[Ala]	27.9	3.60	5.40	3.60	4.39	-1.92	-0.05	35.7	26.7	5.81
[Arg]	73.2	-1.90	18.10	13.39	50.81	-10.09	-0.53	74.2	105.9	-1.18
[Arg^+]	82.2	1.00	22.20	2.08	2.69	-3.38	0.50	79.7	88.7	-21.62
[Asn]	34.9	-1.10	10.30	4.23	0.99	-3.04	-0.02	32.8	43.6	-0.45
[Asp]	36.5	-1.30	9.50	4.91	-9.39	-0.46	-0.06	34.4	41.2	26.87
[Asp^-]	34.7	-4.70	7.50	7.63	7.34	-4.60	2.24	4.7	31.3	-14.35
[Cys]	38.2	5.00	8.60	4.26	6.54	-3.20	-0.08	49.1	39.6	-1.12
[Cys^-]	39.4	1.10	4.90	6.94	9.47	-1.91	2.36	20.1	26.4	3.97
[Gln]	47.0	-0.80	13.70	5.44	2.67	-3.46	-0.05	45.8	59.9	3.61
[Glu]	43.6	0.60	12.80	6.09	3.78	-4.07	-0.08	45.5	56.5	2.53
[Glu^-]	26.0	2.80	11.00	8.82	8.09	-5.25	2.42	9.6	46.3	-14.41
[Gly]	6.9	2.20	1.80	2.57	1.22	-1.27	0.00	11.4	9.5	4.70
[His]	60.0	-0.70	14.80	6.66	4.75	-4.28	0.04	58.2	65.5	3.08
[His^+]	56.4	-1.40	16.90	-2.20	-3.65	-1.30	0.28	51.0	61.2	-27.69
[Ile]	78.1	7.00	15.00	8.41	15.39	-5.99	-0.14	93.6	72.1	0.51
[Leu]	81.1	7.30	15.20	9.37	17.16	-6.60	-0.14	97.2	74.2	1.07
[Lys]	65.0	-0.30	15.10	10.16	21.66	-7.76	-0.16	65.8	75.1	-4.08
[Lys^+]	73.6	2.60	19.10	0.35	1.70	-2.22	0.98	70.0	73.7	-26.59
[Met]	63.7	4.00	15.50	8.76	10.82	-6.00	-0.10	72.7	71.9	2.22
[Phe]	86.5	4.10	18.80	10.66	16.56	-7.61	-0.11	95.8	88.5	-0.13
[Pro]	42.0	1.10	9.90	6.73	9.11	-3.91	-0.09	45.1	49.0	7.99
[Ser]	26.9	1.20	6.20	2.59	1.10	-1.72	-0.05	29.8	27.0	1.32
[Thr]	43.9	3.50	9.40	4.69	6.10	-3.10	-0.12	52.1	43.7	2.80
[Trp]	95.2	2.80	24.40	11.07	12.01	-7.59	-0.08	101.6	109.9	2.30
[Tyr]	84.6	-1.80	20.60	10.35	5.59	-6.81	-0.14	82.2	90.9	5.28
[Tyr^-]	87.4	-5.40	16.80	12.35	13.65	-5.82	2.55	53.2	77.9	1.18
[Val]	62.2	6.30	11.80	6.34	11.81	-4.35	-0.11	76.0	56.9	1.89

BGD
2, 1515–1615, 2005

Table 10. Continued.

Group	c_1^a	$c_2^b \times 10^{-4}$	$a_1^c \times 10$	$a_2^d \times 10^{-2}$	a_3^e	$a_4^b \times 10^{-4}$	$\omega^d \times 10^{-5}$	$C_P^{\circ a}$	$V^{\circ f}$	$\kappa_T^\circ \times 10^4$
Amino Acid Backbone and Protein Backbone Groups										
[AABB]	21.6	-10.6	0.947	-1.86	2.77	-1.77	0.23	-2.0	33.7	-29.26
[AABB$^+$]	23.4	-7.2	1.15	-4.58	-13.96	2.37	-2.07	27.6	43.6	11.96
[AABB$^-$]	68.7	-17.7	1.32	0.37	-18.69	0.15	2.83	6.9	37.2	-22.91
[UPBB]	11.2	-7.5	0.805	-3.75	-14.38	1.13	0.05	-4.5	21.3	-13.70
[PPBB]	21.2	-7.5	0.805	-3.75	-14.38	1.13	0.05	5.5	21.5	-13.70
Organic Groups										
[$-CH_2-$]	16.0	2.3	0.349	2.15	3.09	-1.70	0.00	20.7	16.0	-1.69
[$-CH_3$]	31.5	1.9	0.555	3.60	3.96	-1.92	-0.05	35.8	26.9	5.81
[$-CHCH_3-$]	30.1	4.4	0.62	3.62	10.12	-2.98	-0.09	39.9	31.2	-3.04
[$-CH_2NH_2$]	25.9	-2.5	0.698	0.00	-0.74	0.00	-0.19	22.6	29.4	1.78
[$-CH_2OH$]	26.2	0.9	0.586	2.59	3.03	-1.72	-0.05	28.5	26.9	1.32
[$-CONH_2$]	14.5	-4.9	0.702	1.14	-5.03	-0.06	-0.05	5.0	28.1	6.99
[$-COOH$]	12.9	-4.1	0.582	1.79	-2.40	-0.67	-0.08	5.3	24.4	5.91
[$-C_6H_5$]	70.4	1.8	1.533	8.51	13.47	-5.91	-0.11	75.1	72.5	1.56
[di.corr]	0.0	-6.7	0.006	-1.48	-0.05	2.32	-0.02	-13.5	3.4	11.49
Representative Uncertainties										
$\delta_{[SC]}$	1.8	0.8	0.2	0.96	0.97	1.62	0.20	±5.2	±7.2	±22.09
$\delta_{[UPBB]}$	1.2	1.6	0.4	0.50	0.93	1.00	0.20	±6.3	±5.9	±13.78

[a] cal mol^{-1} K^{-1}.
[b] cal mol^{-1} K.
[c] cal mol^{-1} bar^{-1}.
[d] cal mol^{-1}.
[e] cal K mol^{-1} bar^{-1}.
[f] cm^3 mol^{-1}.

BGD
2, 1515–1615, 2005

Table 11. Protein abbreviations, identification, length, mass, formula and sources of C°_P and V° data as a function of temperature for the reference set of proteins.

Abbreviation[a]	Protein	Organism	n	Mass[b,c]	Formula[b]	Property[d]
ALBU_BOVIN	Serum albumin	*Bos taurus*	583	66433	$C_{2934}H_{4615}N_{781}O_{897}S_{39}$	$C^\circ_{P,(cr)}$
AMYA_PYRFU	Alpha amylase	*Pyrococcus furiosus*	648	76178	$C_{3526}H_{5325}N_{885}O_{976}S_{14}$	$C^{\circ e}_{P,(aq)}$
BPT1_BOVIN	Pancreatic trypsin inhibitor	*Bos taurus*	58	6518	$C_{284}H_{438}N_{84}O_{79}S_7$	$C^{\circ f}_{P,(aq)}$
CTRA_BOVIN	Chymotrypsinogen A	*Bos taurus*	245	25665	$C_{1127}H_{1784}N_{308}O_{352}S_{12}$	$C^\circ_{P,(cr)}$
CYC_BOVIN	Cytochrome C	*Bos taurus*	104	11572	$C_{517}H_{825}N_{143}O_{150}S_4$	$C^{\circ g}_{P,(aq)}$, $V^{\circ h}_{(aq)}$
IL1B_HUMAN	Interleukin 1-Beta	*Homo sapiens*	153	17377	$C_{773}H_{1219}N_{201}O_{237}S_8$	$C^{\circ i}_{P,(cr)}$
INS_BOVIN	Insulin	*Bos taurus*	51	5722	$C_{254}H_{381}N_{65}O_{74}S_6$	$C^\circ_{P,(cr)}$
LACB_BOVIN	Beta-Lactoglobulin	*Bos taurus*	162	18367	$C_{821}H_{1322}N_{206}O_{250}S_9$	$C^\circ_{P,(cr)}$
LYC_CHICK	Lysozyme C	*Gallus gallus*	129	14313	$C_{613}H_{959}N_{193}O_{185}S_{10}$	$C^{\circ g}_{P,(aq)}$, $V^{\circ h}_{(aq)}$, $C^\circ_{P,(cr)}$
MYG_HORSE	Myoglobin	*Equus caballus*	153	16952	$C_{769}H_{1212}N_{210}O_{218}S_2$	$C^\circ_{P,(cr)}$
MYG_PHYCA	Myoglobin	*Physeter catodon*	153	17200	$C_{783}H_{1240}N_{216}O_{216}S_2$	$C^{\circ g}_{P,(aq)}$, $V^{\circ h}_{(aq)}$
OVAL_CHICK	Ovalbumin	*Gallus gallus*	385	42750	$C_{1901}H_{2999}N_{499}O_{575}S_{22}$	$C^\circ_{P,(cr)}$
RNBR_BACAM	Ribonuclease	*Bacillus amyloliquefaciens*	110	12383	$C_{555}H_{847}N_{153}O_{170}$	$C^{\circ j}_{P,(aq)}$
RNH_ECOLI	Ribonuclease H	*Escherichia coli*	155	17597	$C_{776}H_{1215}N_{227}O_{228}S_7$	$C^{\circ k}_{P,(aq)}$
RNH_THET8	Ribonuclease H	*Thermus themophilus*	166	18728	$C_{829}H_{1297}N_{253}O_{231}S_7$	$C^{\circ k}_{P,(aq)}$
RNP_BOVIN	Pancreatic ribonuclease	*Bos taurus*	124	13690	$C_{575}H_{909}N_{171}O_{193}S_{12}$	$C^{\circ g}_{P,(aq)}$, $V^{\circ h}_{(aq)}$, $C^\circ_{P,(cr)}$
RNT1_ASPOR	Ribonuclease T1	*Aspergillus oryzae*	104	11089	$C_{479}H_{687}N_{127}O_{171}S_4$	$C^{\circ l}_{P,(aq)}$

[a] Abbreviation used in the SWISS-PROT database (Boeckmann et al., 2003).
[b] For the neutral, disulfide-free and metal-free protein.
[c] g mol^{-1}.
[d] Values of $C^\circ_{P,(cr)}$ are taken from the sources indicated in the caption to Fig. 21.
Others are indicated by:
[e] Laderman et al. (1993).
[f] Makhatadze et al. (1993).
[g] Privalov and Makhatadze (1990).
[h] Makhatadze et al. (1990)
[i] Makhatadze et al. (1994).
[j] Griko et al. (1994).
[k] Guzman-Casado et al. (2003)
[l] Yu et al. (1994).

BGD
2, 1515–1615, 2005

Table 12. Estimated uncertainties in the calculation of the standard molal thermodynamic properties of amino acids and sidechain and unfolded protein backbone groups at 100 and 200°C and P_{SAT} and at 200 and 300°C at 5000 bar. Estimated uncertainties at 25°C and 1 bar are given in Tables 2 and 7 for *AA* and Tables 8 and 10 for [*SC*] and [UPBB].

	100°C, P_{SAT}			200°C, P_{SAT}		
Property	*AA*	[*SC*]	[UPBB]	*AA*	[*SC*]	[UPBB]
$\Delta G°$[a]	0.57	1.63	2.08	0.74	2.00	3.05
$\Delta H°$[a]	0.75	2.26	3.20	1.13	2.96	3.87
$S°$[b]	1.26	2.95	8.97	2.15	4.59	10.52
$C_P°$[b]	3.0	5.0	4.8	5.3	10.1	9.6
$V°$[c]	2.7	5.4	4.8	3.3	6.6	6.3
$\kappa_T°$[d]	8.72	17.43	11.94	21.79	43.58	39.17
	200°C, 5000 bar			300°C, 5000 bar		
Property	*AA*	[*SC*]	[UPBB]	*AA*	[*SC*]	[UPBB]
$\Delta G°$[a]	0.90	2.33	3.40	1.19	2.94	4.60
$\Delta H°$[a]	1.07	2.84	3.71	1.10	2.85	3.70
$S°$[b]	1.74	3.78	9.59	2.04	4.30	10.07
$C_P°$[b]	1.9	3.3	2.8	2.3	4.3	3.8
$V°$[c]	2.1	4.1	3.8	2.0	4.1	3.9
$\kappa_T°$[d]	5.56	11.13	6.72	5.62	11.24	7.29

[a] kcal mol^{-1}.
[b] cal mol^{-1} K^{-1}.
[c] cm^3 mol^{-1}.
[d] 10^{-4} cm^{-3} bar^{-1} mol^{-1}.

BGD
2, 1515–1615, 2005

Table 13. Group additivity equations for calculating the parameters of [di.corr] and revising the parameters of the organic groups given by Amend and Helgeson (1997b).

$$\Xi_{[-CH_3]} + \Xi_{[-CH_2-]} + \Xi_{[-COOH]} = \Xi^{AH97b}_{[-CH_3]} + \Xi^{AH97b}_{[-CH_2-]} + \Xi^{AH97b}_{[-COOH]} \quad \text{(B1)}$$

$$2\Xi_{[-COOH]} + 2\Xi_{[-CH_2-]} + \Xi_{[di.corr]} = 2\Xi^{AH97b}_{[-COOH]} + 2\Xi^{AH97b}_{[-CH_2-]} \quad \text{(B2)}$$

$$\Xi_{[-CH_3]} = \Xi^{AH97b}_{[-CH_3]} + \frac{1}{2}\Xi_{[di.corr]} \quad \text{(B3)}$$

$$\Xi_{[-COOH]} = \Xi^{AH97b}_{[-COOH]} - \frac{1}{2}\Xi_{[di.corr]} \quad \text{(B4)}$$

$$\Xi_{[term]} = \Xi^{AH97b}_{[term]} - \frac{1}{2}\Xi_{[di.corr]} \quad \text{(B5)}$$

$$\Xi_{[-CHCH_3-]} = \Xi^{AH97b}_{[-CHCH_3-]} \quad \text{(B6)}$$

(B7)

$$\Xi_{[AABB]} = \Xi_{Ala} - \left(\Xi^{AH97b}_{[-CH_3]} + \frac{1}{2}\Xi_{[di.corr]}\right)$$

$$\Xi_{[AABB]} = \Xi_{Ser} - \left(\Xi^{AH97b}_{[-CH_2OH]} - \frac{1}{2}\Xi_{[di.corr]}\right) \quad \text{(B8)}$$

$$\Xi_{propanol} = \Xi_{[-CH_3]} + \Xi_{[-CH_2-]} + \Xi_{[-CH_2OH]}$$

$$\Xi_{butane\text{-}1,4\text{-}diol} = 2\Xi_{[-CH_2-]} + 2\Xi_{[-CH_2OH]} + \Xi_{[di.corr]}$$

$$\Xi_{hexane\text{-}1,6\text{-}diol} = 4\Xi_{[-CH_2-]} + 2\Xi_{[-CH_2OH]} + \Xi_{[di.corr]} \quad \text{(B9)}$$

$$\Xi_{propylamine} = \Xi_{[-CH_3]} + \Xi_{[-CH_2-]} + \Xi_{[-CH_2NH_2]}$$

$$\Xi_{butane\text{-}1,4\text{-}diamine} = 2\Xi_{[-CH_2-]} + 2\Xi_{[-CH_2NH_2]} + \Xi_{[di.corr]}$$

$$\Xi_{hexane\text{-}1,6\text{-}diamine} = 4\Xi_{[-CH_2-]} + 2\Xi_{[-CH_2NH_2]} + \Xi_{[di.corr]}$$

BGD
2, 1515–1615, 2005

Table 14. Group additivity algorithms for calculating ΔH_f° and $S^\circ_{P_r,T_r}$ of gaseous amino acid sidechains.

Group	Group additivity algorithm
[Ala]	$[C-(H)_3(C)]$
[Arg]	$2[C-(H)_2(C)_2] + [C-(H)_2(C)(N)]+[N-(H)(C)_2] + [C-(N)_2(N_A)] + [N-(H)_2(C)]_{second,amino\ acids} + [N_A-(H)(C)]$
[Asn]	$[C-(H)_2(CO)(C)] + [CO-(C)(N)] + [N-(H)_2(CO)]_{amides,\ ureas}$
[Asp]	$[C-(H)_2(CO)(C)] + [CO-(C)(O)] + [O-(H)(CO)]$
[Cys]	$[C-(H)_2(C)(S)] + [S-(C)(H)]$
[Gln]	$[C-(H)_2(C)_2] + [C-(H)_2(CO)(C)] + [CO-(C)(N)] + [N-(H)_2(CO)]_{amides,\ ureas}$
[Glu]	$[C-(H)_2(C)_2] + [C-(H)_2(CO)(C)] + [CO-(C)(O)] + [O-(H)(CO)]$
[His]	S°: $[C-(H)_2(C)(C_B)] + [C_B-(N)(C_B)(C)] + [C_B-(H)(C_B)(N_A)] + [N_I-(C_B)] + [C_B-(H)(N_I)(N)] + [N-(H)(C_B)_2]$ ΔH_f°: [2 – ethylimidazole] – $[C-(H)_3(C)]$
[Ile]	$[C-(H)_2(C)_2] + 2[C-(H)_3(C)] + [C-(H)(C)_3]$
[Leu]	$[C-(H)_2(C)_2] + 2[C-(H)_3(C)] + [C-(H)(C)_3]$
[Lys]	$3[C-(H)_2(C)_2] + [C-(H)_2(C)(N)] + [N-(H)_2(C)]_{second}$
[Met]	$[C-(H)_2(C)_2] + [C-(H)_2(C)(S)]+[S-(C)_2]+[C-(H)_3(S)]$
[Phe]	$[C-(H)_2(C)(C_B)] + [C_B-(C)(C_B)_2] + 5[C_B-(H)(C_B)_2]$
[Ser]	$[C-(H)_2(O)(C)] + [O-(H)(C)]$
[Thr]	$[C-(H)(O)(C)_2]_{alcohols,peroxides} + [C-(H)_3(C)] + [O-(H)(C)]$
[Trp]	$[C-(H)_2(C)(C_B)] + [C_B-(C)(C_B)_2] + [C_B-(H)(C_B)(N)] + [N-(H)(C_B)_2] + [C_B-(N)(C_B)_2] + 4[C_B-(H)(C_B)_2] + [C_{BF}-(C_{BF})(C_B)_2]$
[Tyr]	$[C-(H)_2(C)(C_B)] + [C_B-(C)(C_B)_2] + 4[C_B-(H)(C_B)_2] + [C_B-(O)(C_B)_2] + [O-(H)(C_B)]$
[Val]	$2[C-(H)_3(C)] + [C-(H)(C)_3]$

BGD
2, 1515–1615, 2005

Table 15. Group contributions to ΔH_f° and S° of gaseous amino acids.[a]

	ΔH_f° kcal mol^{-1}	$S^\circ_{P_r,T_r}$ cal mol^{-1} K^{-1}		ΔH_f° kcal mol^{-1}	$S^\circ_{P_r,T_r}$ cal mol^{-1} K^{-1}		ΔH_f° kcal mol^{-1}	$S^\circ_{P_r,T_r}$ cal mol^{-1} K^{-1}
Groups			**Groups**			**Groups**		
[AABB][c]	-89.00	63.74	[C – (H) (O) $(C)_2$][b1]	-6.24	-10.29	[CO–(C)(N)]	-31.85	13.55
[C – $(H)_3$ (C)]	-10.10	30.43	[C – $(H)_2$ (O) (C)]	-7.86	10.38	[N – $(H)_2$ (CO)][b3]	-15.06	21.09
[C – $(H)_2$ $(C)_2$]	-4.93	9.36	[C – $(H)_2$ (C) (N)]	-6.76	10.10	[C – $(H)_3$ (S)]	-10.10	30.43
[C – (H) $(C)_3$]	-0.28	-12.81	[C_B – (H) (N_I) (N)][d]	-6.76	10.10	[C – $(H)_2$ (C) (S)]	-5.54	10.01
[C_B – (H) $(C_B)_2$]	3.30	11.55	[N – $(H)_2$ (C)][b2]	4.60	30.33	[S – (H)(C)]	4.46	32.90
[C_B – (C) $(C_B)_2$]	5.65	-8.51	[N-(H)$(C)_2$]	16.14	8.12	[S – $(C)_2$]	11.23	13.19
[C – $(H)_2$ (C) (C_B)]	-5.10	10.18	[N_A – (H)(C)][e]	16.14	8.12			
[C_{BF} – (C_{BF})$(C_B)_2$]	4.80	0.00	[N – (H)$(CB)_2$][h]	19.97	8.12	**Model compounds**		
[CO – (C) (O)]	-32.80	14.96	[N_I – (C_B)]	16.49	11.24	2-ethylimidazole[j]	16.30	—
[O – (H)(CO)]	-60.78	24.31	[C_B – (N) $(C_B)_2$]	-0.31	-10.40	pyrrole[k]	25.88	—
[O – (H)(C_B)]	-38.31	29.04	[C_B – (N) (C_B) (C)][f]	-0.31	-10.40			
[O – (H)(C)]	-38.08	29.04	[C – $(N)_2$ (N_A)[f]]	-0.31	-10.40	**Estimated Uncertainty**[j]		
[C_B – (O)$(C_B)_2$]	-1.14	-10.45	[C_B – (H) (C_B) (N_A)][g]	3.30	11.55	δ	0.60	0.50
[C – $(H)_2$ (CO) (C)]	-5.22	9.46	[C_B – (H) (C_B) (N)][i]	-0.34	11.55			

[a] Unless otherwise noted, values of ΔH_f° and $S^\circ_{P_r,T_r}$ are from Domalski and Hearing (1993). Specific designations are represented by [b1] (alcohols, peroxides), [b2] (second, amino acids) and [b3] (amides, ureas). [c] ΔH_f° of gaseous Ala is taken from Ngauv et al. (1977). $S^\circ_{P_r,T_r}$ of Ala is calculated from ΔH_f°, the entropies of the elements, and ΔG_f° taken from Plyasunov and Shock (2001). [d–g] ΔH_f° and $S^\circ_{P_r,T_r}$ of the indicated groups are made equivalent to those of ([d]) [C – $(H)_2$ (C) (N)], ([e]) [N – (H) $(C)_2$], ([f]) [C_B – (N) $(C_B)_2$] or ([g]) [C_B – (H) $(C_B)_2$]. [h] $S^\circ_{[N-(H)(CB)_2]} = S^\circ_{[N-(H)(C)_2]}$. [i] $\Delta H^\circ_{f,[C_B-(H)(C_B)(N)]} = \frac{1}{2}\Delta H^\circ_{f,pyrrole} - \Delta H^\circ_{f,[C_B-(H)(C_B)_2]} - \frac{1}{2}\Delta H^\circ_{f,[N-(H)(CB)_2]}$; $S^\circ_{[C_B-(H)(C_B)(N)]} = S^\circ_{[C_B-(H)(C_B)_2]}$. [j] Jimenez et al. (1992). [k] Scott et al. (1967). [l] These values are Benson and Buss' (1958) estimates of the uncertainty in the additive prediction of ΔH_f° and $S^\circ_{P_r,T_r}$ of gaseous species.

BGD
2, 1515–1615, 2005

Table 16. ΔG°_f, ΔH°_f, $S^\circ_{P_r,T_r}$, Maier-Kelley parameters and C°_P and V° at 25°C and 1 bar of sidechain and backbone groups of crystalline amino acids and proteins. Values of the Maier-Kelley parameter c are taken to be zero.

Group	ΔG°_f [a]	ΔH°_f [a]	$S^\circ_{P_r,T_r}$ [b]	a [b]	b [c] $\times 10^3$	C°_P [b]	V° [d]
[Ala]	−1.22	−12.12	11.66	1.9	42.1	14.4	26.7
[Arg]	29.88	−26.67	40.68	1.5	132.3	41.0	105.9
[Asn]	−39.25	−66.12	22.49	3.0	68.9	23.6	43.6
[Asp]	−87.25	−110.26	21.44	3.1	64.4	22.3	41.2
[Cys]	5.97	−4.32	21.38	−2.8	89.7	24.0	39.6
[Gln]	−39.79	−74.92	27.4	4.4	83.3	29.2	59.9
[Glu]	−87.49	−118.94	25.76	3.1	80.1	27.0	59.9
[Gly]	−3.01	−6.02	5.52	−0.4	31.1	8.9	9.5
[His]	38.05	10.88	38.24	0.4	126.6	38.2	65.5
[Ile]	4.33	−30.12	30.49	3.3	90.2	30.2	72.1
[Leu]	1.96	−32.22	31.4	2.6	97.0	31.5	74.2
[Lys]	4.62	−39.83	35.46	−0.6	121.4	35.6	75.1
[Met]	−33.48	−58.82	36.1	5.4	90.4	32.3	71.9
[Phe]	36.74	10.78	31.84	−1.8	119.2	33.7	88.5
[Pro]	15.23	−3.32	19.99	−0.3	72.7	21.4	49.0
[Ser]	−35.96	−52.75	16.43	3.1	48.7	17.6	27.0
[Thr]	−32.72	−58.98	17.28	0.0	76.7	22.9	43.7
[Trp]	58.78	23.18	40.78	−1.7	147	42.1	109.9
[Tyr]	−8.11	−41.36	31.93	−1.7	129.6	37.0	0.0
[Val]	1.48	−25.32	23.53	3.6	73.8	25.6	56.9
[AABB]	−87.22	−122.38	19.22	3.9	36.5	14.8	33.7
[PBB]	−22.33	−43.22	11.34	1.3	25.0	8.8	—
$\delta_{[SC]}$	±0.50	±0.50	±0.10	±2.0	±5.0	±2.0	±3.0
$\delta_{[PBB]}$	±1.00	±1.00	±0.12	±2.0	±5.0	±2.0	±10.0

[a] kcal mol^{-1}. [b] cal mol^{-1} K^{-1}. [c] cal mol^{-1} K^{-2}. [d] cm^3 mol^{-1}.

BGD
2, 1515–1615, 2005

EGU

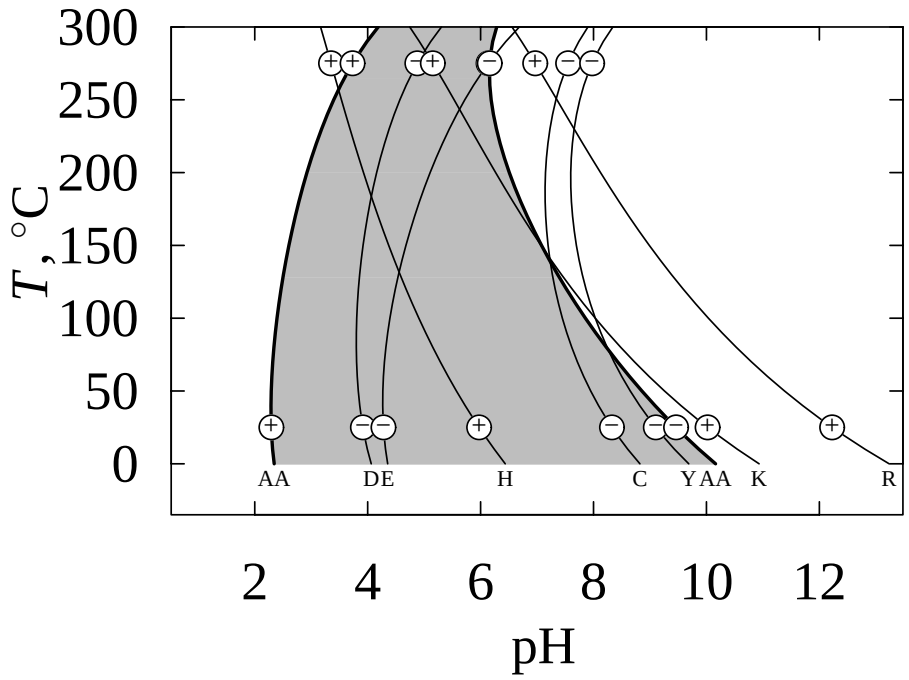

Fig. 1. p*K*s of sidechain and backbone groups in amino acids as a function of temperature at P_{sat}, calculated at 25°C using values of the standard molal thermodynamic properties taken from Table 8 and at other temperatures using these values and the values of the equations of state parameters taken from Table 10. Thin curves represent equal activity of neutral and ionized forms of the sidechain groups, which are identified by the one-letter abbreviations shown beneath the curves. The bold curves represent equal activities of either [AABB$^+$] or [AABB$^-$] relative to [AABB], and therefore delimit the stability field of [AABB], represented by grey shading. Charges of the ionized groups are indicated by "+" and "–" symbols.

BGD
2, 1515–1615, 2005

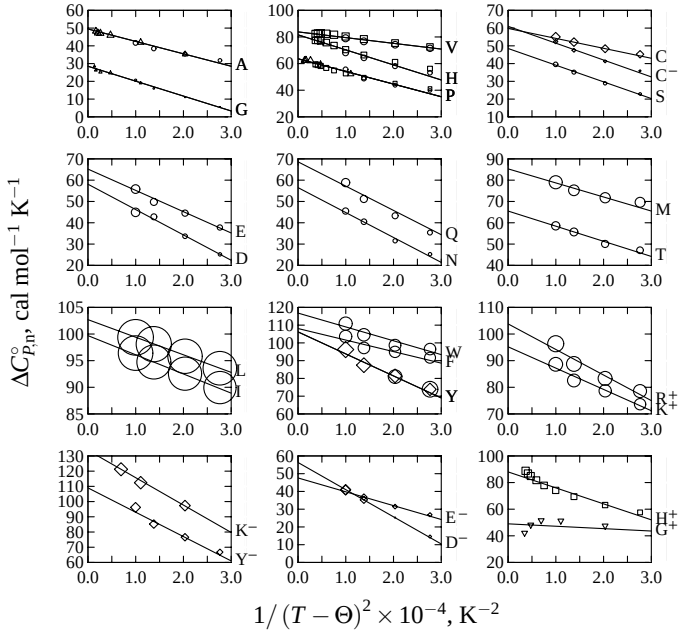

Fig. 2. $\Delta C^{\circ}_{P,n}$ of amino acids as a function of $1/(T-\Theta)^2$. The intercepts and slopes of the regression lines correspond to c_1 and c_2, respectively, in the revised HKF equations of state. Values of $\Delta C^{\circ}_{P,n}$ are calculated from experimental values of C°_P taken from (◯) Hakin et al. (1994a) (Ala, Gly, Ser, Thr), Hakin et al. (1994b) (Asp, Glu), Duke et al. (1940) (Ile, Leu, Val), Hakin et al. (1995) (Asn, Gln), Hakin et al. (1997a) (Arg, Met, Pro), Marriott et al. (1998) (His, Phe, Trp, Tyr), Hakin and Hedwid (2001b) (Lys^+, Arg^+); (□) Price et al. (2003a) (His), Jardine et al. (2001) (His^+), Price et al. (2003b) (Val); (△) Clarke et al. (2000); (▽) Downes et al. (2001). Values of C°_P estimated in the present study are represented by open diamonds (◇). The diameters of the symbols represent an estimated average uncertainty of ±2.5% in the experimental value of C°_P. The data at 25° and 1 bar given by Jolicoeur et al. (1986), Spink and Wädso (1969) and DiPaola and Belleau (1978) are not shown.

BGD
2, 1515–1615, 2005

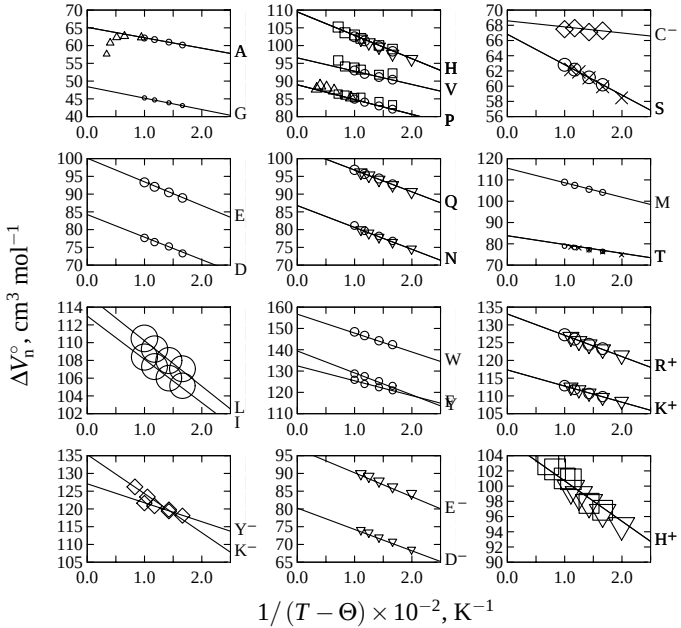

Fig. 3. ΔV_n° of amino acids as a function of $1/(T - \Theta)$. The intercepts and slopes of the regression lines correspond to σ and ξ, respectively, in the revised HKF equations of state. Values of ΔV_n° are calculated from experimental values of V° from (◯) Hakin et al. (1994a) (Ala, Gly, Ser, Thr), Hakin et al. (1994b) (Asp, Glu), Duke et al. (1940) (Ile, Leu, Val), Hakin et al. (1995) (Asn, Gln), Hakin et al. (1997a) (Arg, Met, Pro), Marriott et al. (1998) (His, Phe, Trp, Tyr), Hakin and Hedwid (2001b) (Lys$^+$, Arg$^+$); (□) Price et al. (2003a) (His), Jardine et al. (2001) (His$^+$ and His$^-$), Sorenson et al. (2003) (Pro), Price et al. (2003b) (Val); (△) Clarke and Tremaine (1999); (▽) Yasuda et al. (1998); (×) Mizuguchi et al. (1997). Not shown are low-temperature data given by Marriott et al. (2001), Kharakoz (1998), Kikuchi et al. (1995), Banipal and Kapoor (1999), Jolicoeur et al. (1986), Millero et al. (1998), Ahluwalia et al. (1977). The diameters of the symbols represent an estimated average uncertainty of ±1% in the experimental value of V°. Values of C_P° estimated in the present study are represented by open diamonds (◇).

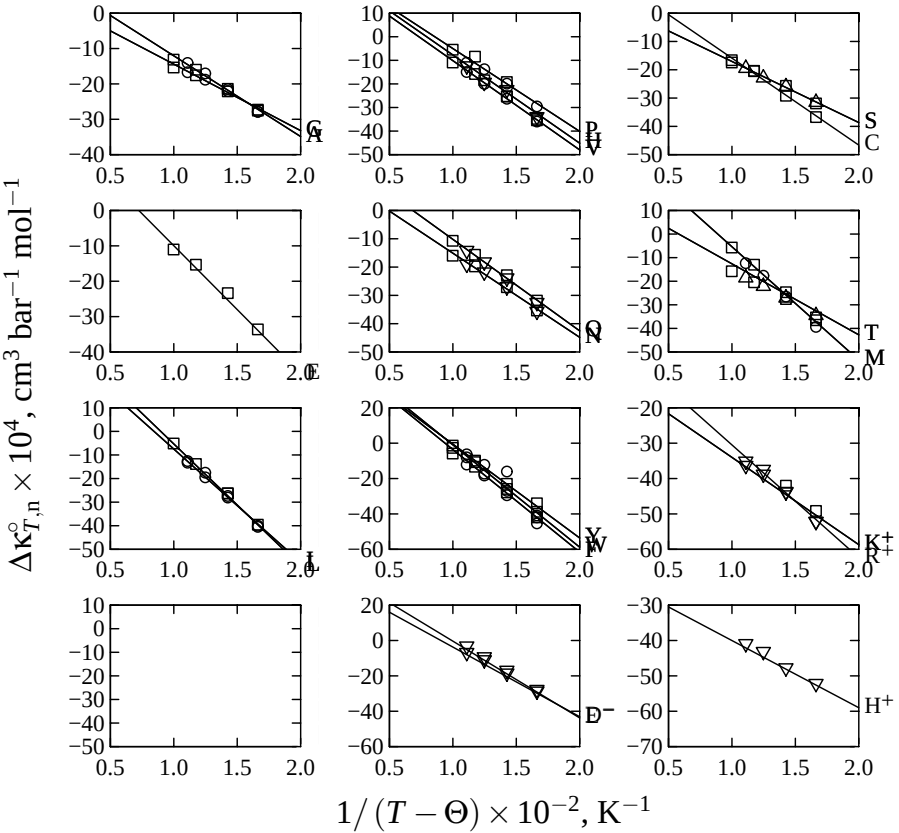

Fig. 4. $\Delta\kappa^{\circ}_{T,\mathrm{n}}$ of amino acids as a function of $1/(T-\Theta)$. The slopes and intercepts of the regression lines correspond to $-\left(\partial\sigma/\partial P\right)_T$ and $-\left(\partial\xi/\partial P\right)_T$, respectively, in the revised HKF equations of state. Values of $\Delta\kappa^{\circ}_{T,\mathrm{n}}$ are calculated from experimental values of κ°_S from (○) Kikuchi et al., 1995; (□) Kharakoz, 1991; (△) Mizuguchi et al., 1997; (▽) Yasuda et al., 1998; and (×) Millero et al., 1998. The diameters of the symbols represent an estimated average uncertainty of ±2.5% in the experimental value of κ°_T.

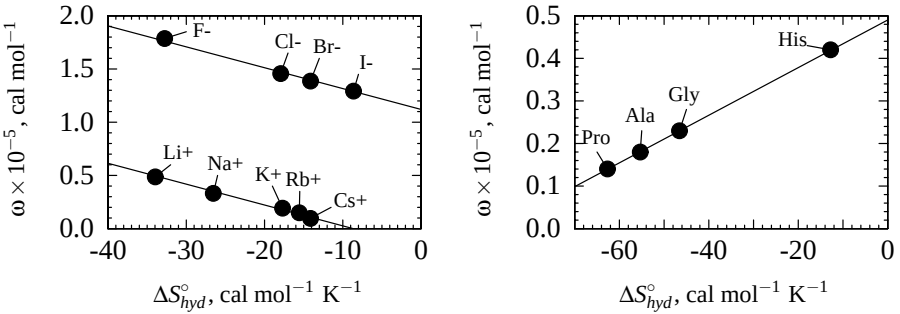

Fig. 5. ω as a function of ΔS°_{hyd} at 25°C and 1 bar for *left*, monovalent metal anions and cations and *right*, neutral amino acids. The symbols represent experimental data (see text), but the lines represent correlations corresponding to Eqs. (6.1)–(6.3).

BGD
2, 1515–1615, 2005

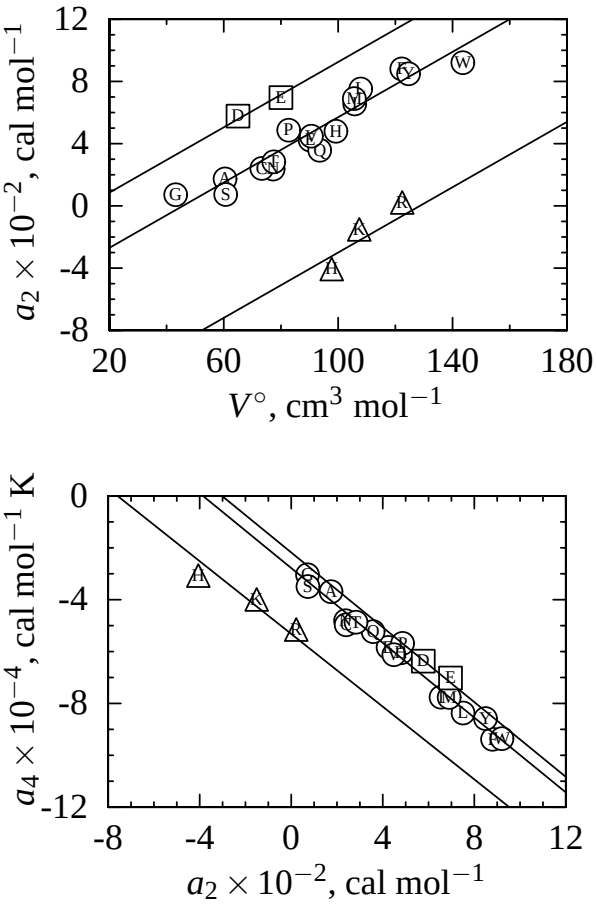

Fig. 6. Correlations between V° and a_2 (top), and between a_2 and a_4 (bottom) for neutral (◯), positively charged (△), and negatively charged (□) amino acids.

BGD
2, 1515–1615, 2005

Thermodynamic properties of ionized proteins

J. M. Dick et al.

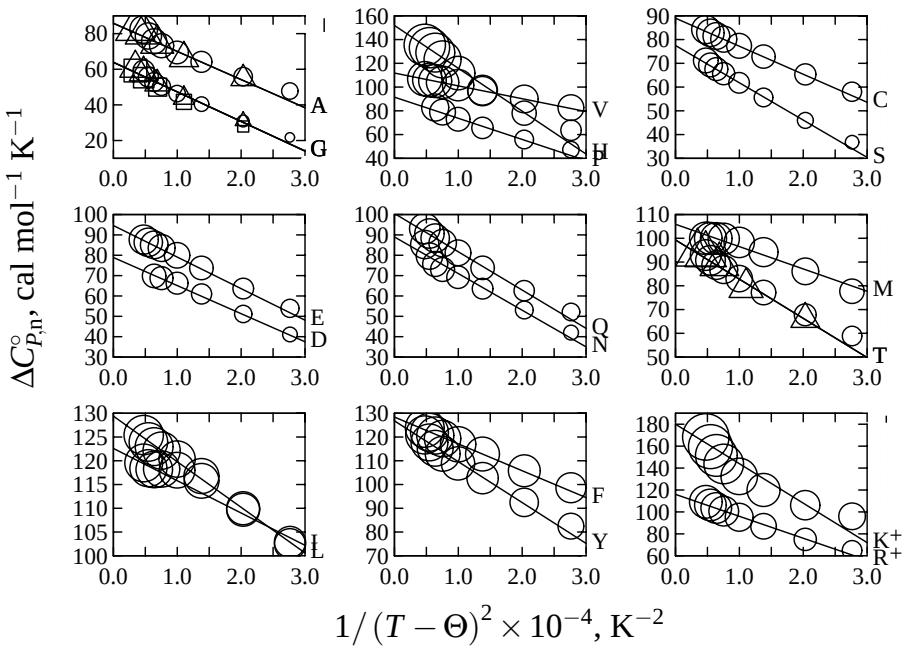

Fig. 7. $\Delta C^{\circ}_{P,n}$ as a function of $1/(T - \Theta)^2$ for Gly–X–Gly tripeptides. The intercepts and slopes of the regression lines correspond to c_1 and c_2, respectively, in the revised HKF equations of state. Values of $\Delta C^{\circ}_{P,n}$ are calculated from experimental values of C°_P from (◯) (Häckel et al., 1999a, 1998; Vogl et al., 1995), (□) Downes and Hedwig, 1995 and (Δ) Makhatadze and Privalov, 1990. The diameters of the symbols represent an estimated average uncertainty of ±2.5% in the experimental value of C°_P.

BGD
2, 1515–1615, 2005

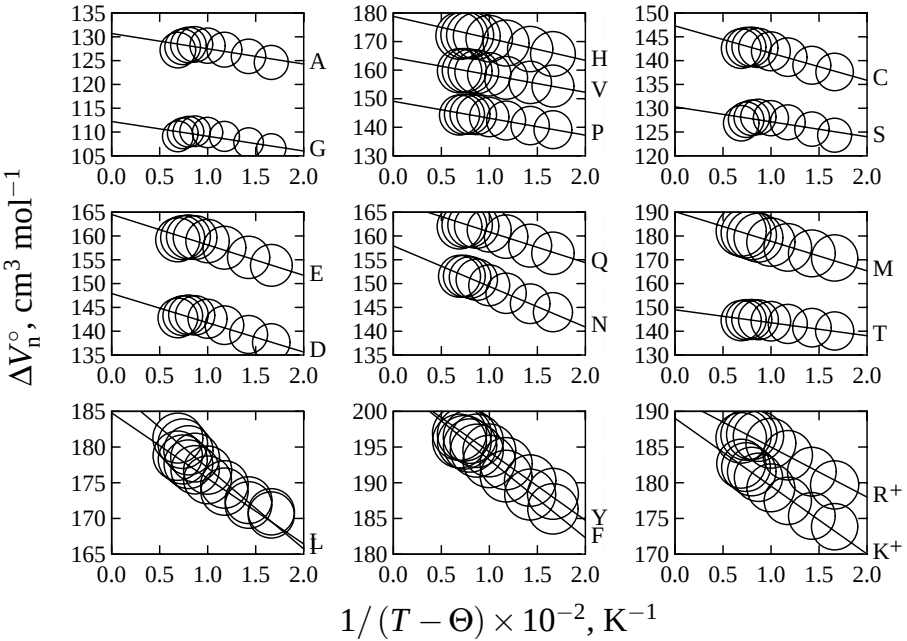

Fig. 8. ΔV_n° as a function of $1/(T - \Theta)$ for Gly–X–Gly tripeptides. The regression lines have intercepts and slopes that correspond, respectively, to σ and ξ in the revised HKF equations of state. Values of ΔV_n° are calculated from experimental values of V° from (◯) (Häckel et al., 1999a, 1998; Vogl et al., 1995). The diameters of the symbols represent an estimated average uncertainty of ±2.5% in the experimental value of C_P°.

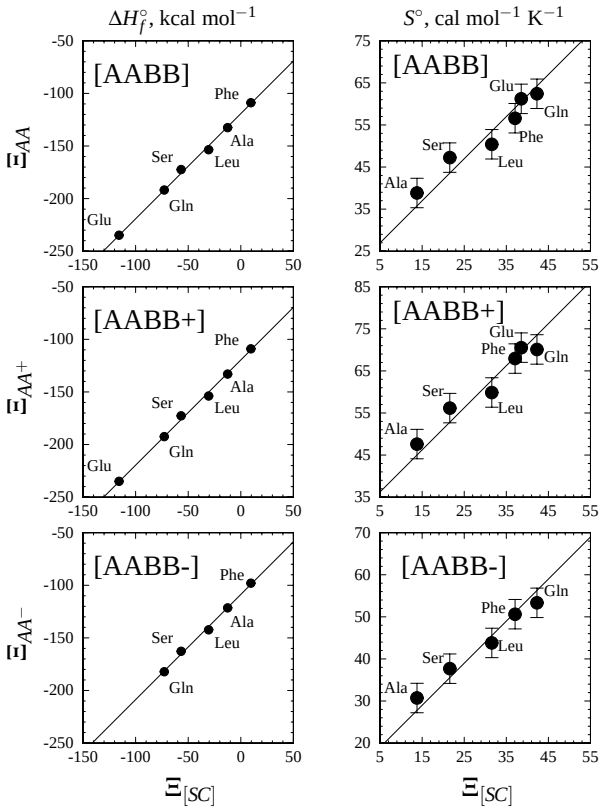

Fig. 9. Ξ_{AA} (top), Ξ_{AA^+} (middle), and Ξ_{AA^-} (bottom) as a function of $\Xi_{[SC]}$ for $\Xi = \Delta H_f^\circ$ (left) and $\Xi = S^\circ$ (right). The symbols are labeled by the reference sidechain groups and represent values of Ξ_{AA}, Ξ_{AA^+} and Ξ_{AA^-} taken from Table 2, and those of $\Xi_{[SC]}$ calculated using Eq. (1) and the group contributions given in Table 8. The regression lines have unit slopes and intercepts that correspond to ΔH_f° or S° of [AABB], [AABB$^+$] and [AABB$^-$]. The uncertainty in the group additivity calculation of ΔH_f° is represented by the diameters of the symbols, and that of S° by the lengths of the error bars.

BGD
2, 1515–1615, 2005

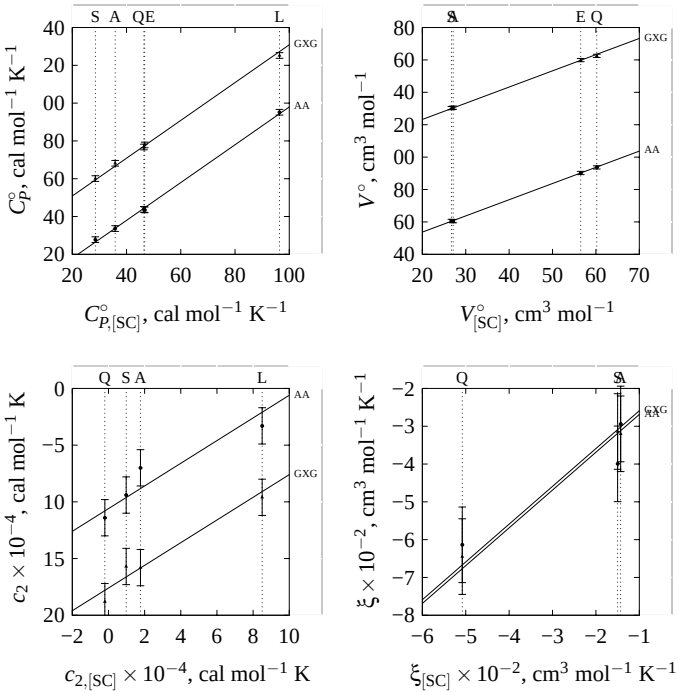

Fig. 10. C_P°, V°, c_2 and ξ of amino acids and Gly–X–Gly tripeptides as a function of the corresponding properties of reference model sidechain groups. Values of C_P°, V°, c_2 and ξ of *AA* and *GXG* are taken from Tables 3, 4 and 7, except those of ξ of *AA* which are calculated using Eq. (A25) and values of a_3 and a_4 taken from Table 3. Values of C_P°, V°, c_2 and ξ of the sidechain groups, corresponding to the positions of the droplines, are calculated using appropriate statements of Eq. (1) and the group contributions given in Table 10. The regression lines have unit slope and intercepts that correspond to C_P°, V°, c_2 and ξ of [AABB] and [GXGBB]. The lengths of the error bars indicate the uncertainty of the group additivity calculation of the properties of AA and GXGBB, estimated from the scatter of the symbols around the regression lines.

BGD
2, 1515–1615, 2005

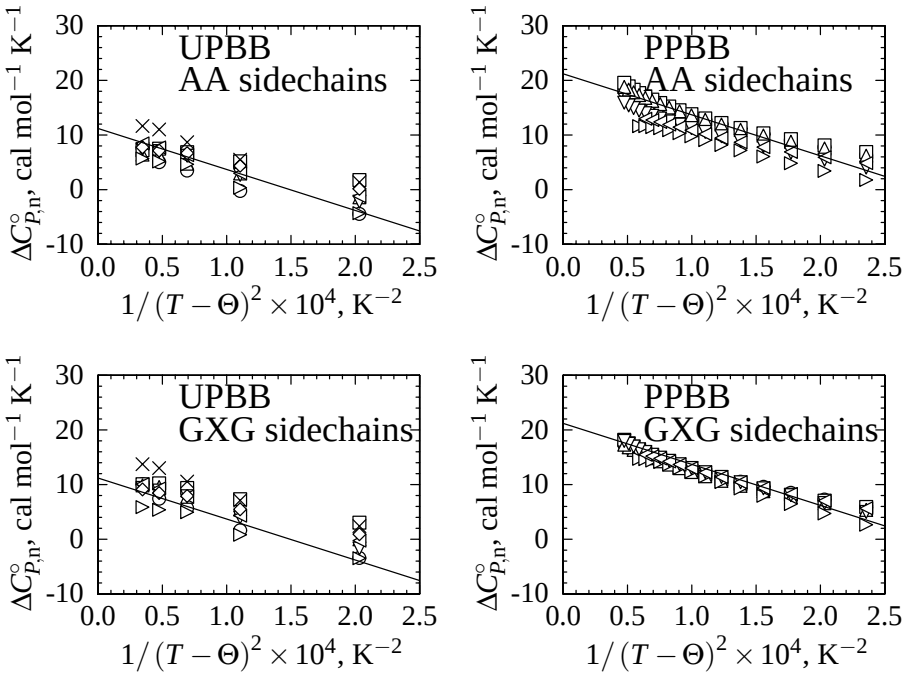

Fig. 11. $\Delta C^\circ_{P,n}$ as a function of $1/(T-\Theta)^2$ of the aqueous unfolded protein backbone and polypeptide backbone. The intercepts and slopes of the regression lines correspond to c_1 and c_2, respectively in Eq. (A23), and are consistent with the group contributions by [UPBB] (*left*) or [PPBB] (*right*) given in Table (10). The symbols represent values of $\Delta C^\circ_{P,n}$ calculated using Eqs. (A5), (A10) and (8), the group contributions by protein sidechains generated from amino acid reference model compounds (*upper*; Table 10) or Gly–X–Gly reference model compounds (*lower*), and experimental values of C°_P of the unfolded proteins CYC_BOVIN (◇), MYG_PHYCA (▷), LYC_CHICK (▽), RNP_BOVIN (◁), BPT1_BOVIN (□), IL1B_HUMAN (△), RNBR_BACAM (◯) and RNT1_ASPOR (×), and of the polypeptides GGLGG (◯) and GGSAG (△). The reported values of C°_P of unfolded proteins at 125°C are derived using estimated values of V° at this temperature.

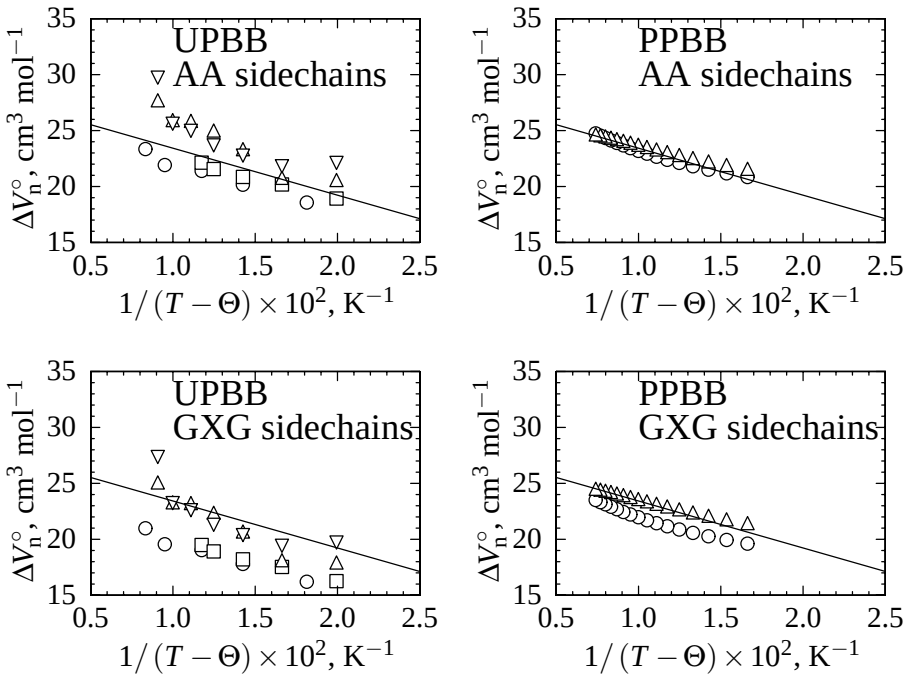

Fig. 12. ΔV_n° as a function of $1/(T - \Theta)$ of the aqueous unfolded protein backbone and polypeptide backbone. The intercepts and slopes of the regression lines correspond to σ and ξ, respectively in Eq. (A7), and are consistent with the contributions by [UPBB] (*left*) or [PPBB] (*right*) given in Table (10). The symbols represent values of ΔV_n° calculated using Eqs. (A5), (A11) and (8), the group contributions by protein sidechains generated from amino acid model compounds (*upper*; Table 10) or Gly–X–Gly model compounds (*lower*), and experimental values of C_p° of the unfolded proteins CYC_BOVIN (◯), LYC_CHICK (△), MYG_PHYCA (□), and RNP_BOVIN (▽), and of the polypeptides GGLGG (◯) and GGSAG (△). The reported values of V° of unfolded proteins at 125°C are generated from group additivity.

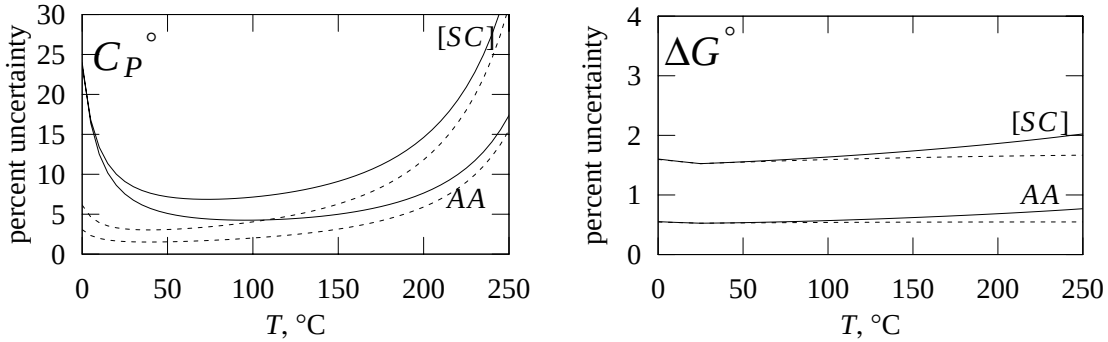

Fig. 13. Propagated amino acid (*AA*) and sidechain group ([*SC*]) uncertainties in C_P° (*top*) and ΔG° (*bottom*) as a function of temperature at P_{sat}. Calculated total uncertainties are converted to percentages of the average value of the property of the 20 neutral amino acids. Solid curves represent total propagated uncertainties; the dashed curves represent either the contributions by $\delta\omega$ to $\delta\Delta C_P^\circ$ or the contributions by $\delta\Delta G_f^\circ$ and δS° to $\delta\Delta G^\circ$.

BGD
2, 1515–1615, 2005

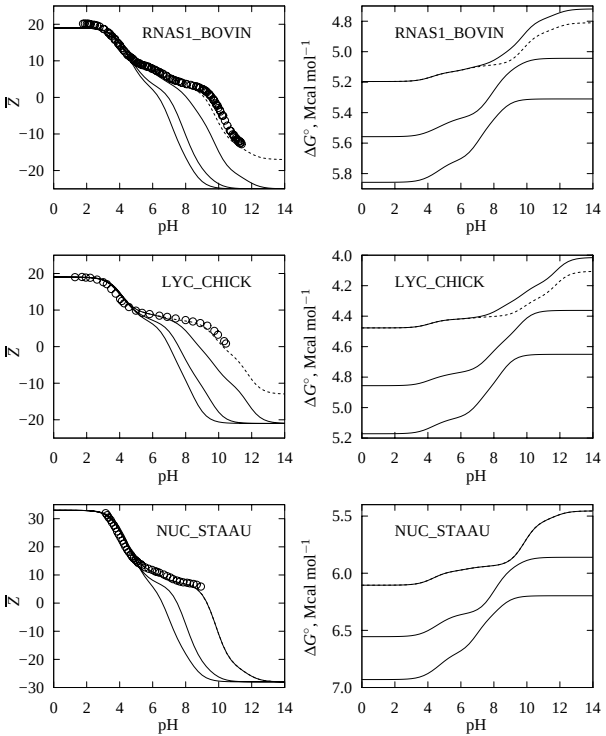

Fig. 14. Calculated and experimental values of the mean net charge ($\overline{Z}$; *left*) and calculated values of the standard molal Gibbs free energy (ΔG°; *right*) of unfolded RNAS1_BOVIN (*upper*), LYC_CHICK (*middle*) and NUC_STAAU (*lower*) as a function of pH and temperature. Open circles represent values of $\overline{Z}$ reported from titration experiments in 6.0 M GuHCL at 25°C (Nozaki and Tanford, 1967b; Roxby and Tanford, 1971; Whitten and García-Moreno E., 2000), and solid curves represent values of $\overline{Z}$ and ΔG° calculated using Eqs. (23) and (24) at 25, 100 and 150°C. Dashed curves – not shown for NUC_STAAU – represent the exclusion of the [Cys] ionization reaction in the calculation of $\overline{Z}$ and ΔG° at 25°C.

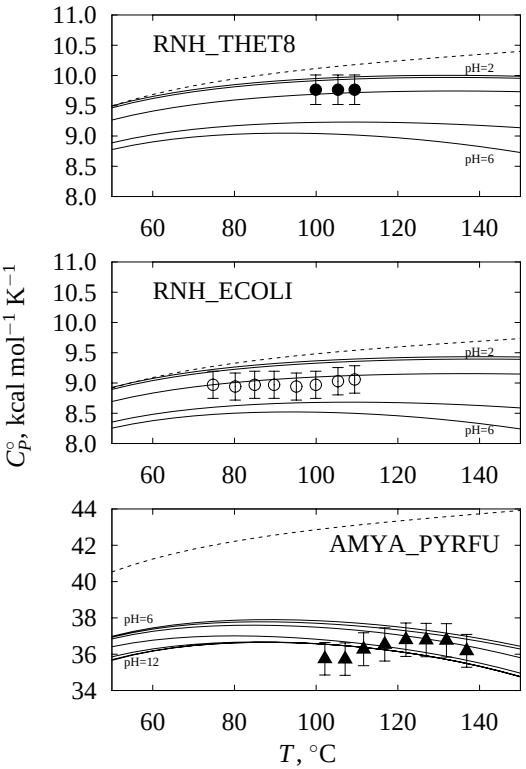

Fig. 15. C_P° to temperatures > 100 °C at P_{sat} of heat-denatured thermophilic proteins. Experimental values taken from Figure 6 of Guzman-Casado et al. (2003) are shown for ribonuclease H from *T. thermophilus* (RNH_THET8) at pHs of 4.0–5.0 (◯) and for ribonuclease H from *E. coli* (RNH_ECOLI) at pHs of 2.0–3.5 (◯). Experimental values of C_P° of unfolded AMYA_PYRFU (▲) are taken from the curve for pH=10.3 in Fig. 4 of Laderman et al. (1993), from 122 to 137°C. At temperatures greater than 125°C, the pH=10.3 curve in Fig. 4 of Laderman et al. (1993) coincides with those for pH=7.0 and pH=8.0. Values predicted in the present study for nonionized unfolded proteins are shown by the dashed curves.

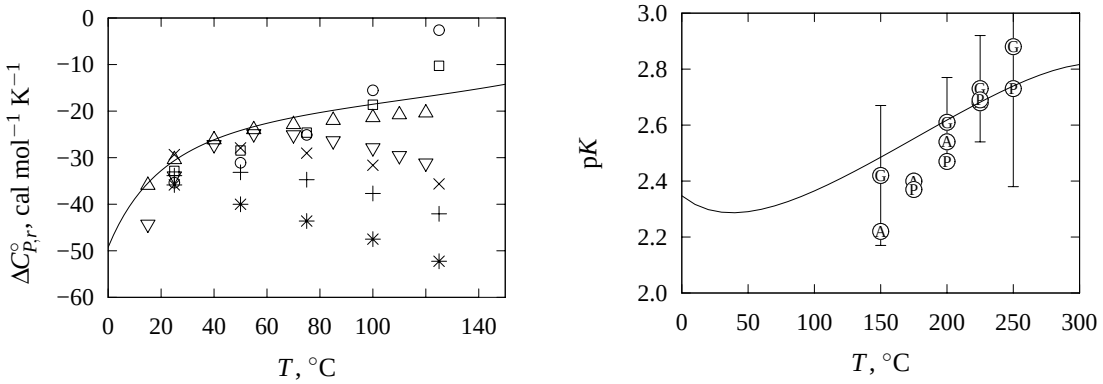

Fig. 16. Comparison of experimental and calculated values of $\Delta C^\circ_{P,r}$ (*top*) and pK (*bottom*) as a function of temperature at P_{SAT} of the reaction represented by [COOH] $\rightleftharpoons$ [COO$^-$] + H$^+$, corresponding to [AABB$^+$] $\rightleftharpoons$ [AABB] + H$^+$ for amino acids with nonionizable sidechain groups. The curve in the top panel represent values of $\Delta C^\circ_{P,r}$ of the latter reaction calculated from Eq. (A30) and the values of c_1, c_2 and ω of [AABB] and [AABB$^+$] taken from Table 10. Experimental values are taken from Wang et al. (1996) for Ala (◯) and Gly (□), from Price et al. (2003b) for Val (△), and from Sorenson et al. (2003) for Pro (▽). Experimental values of $\Delta C^\circ_{P,r}$ of [COOH] $\rightleftharpoons$ [COO$^-$] + H$^+$ are calculated using the revised HKF parameters taken from Shock (1995) for acetic (+), propanoic (×) and butanoic (∗) acids and their ionized counterparts. Experimental values of pK, shown in the lower panel, are represented by circles with a one-letter amino acid label, and were measured at 45 bar by Clarke et al. (2005). These authors' estimated experimental errors for Gly are represented by the error bars shown in the plot. (For clarity, the error bars for Ala and Pro are omitted. Clarke et al. (2005) report experimental uncertainties for Ala of nearly ±1 pK unit at 150°C, but less than ±0.1 at 225°C. All reported experimental errors for Pro are less than ±0.05 pK unit.) The solid curve represents values of pK calculated at P_{SAT} using the values of the revised HKF parameters and ΔG°_f, ΔH°_f and $S^\circ_{P_r,T_r}$ of [AABB$^+$] and [AABB].

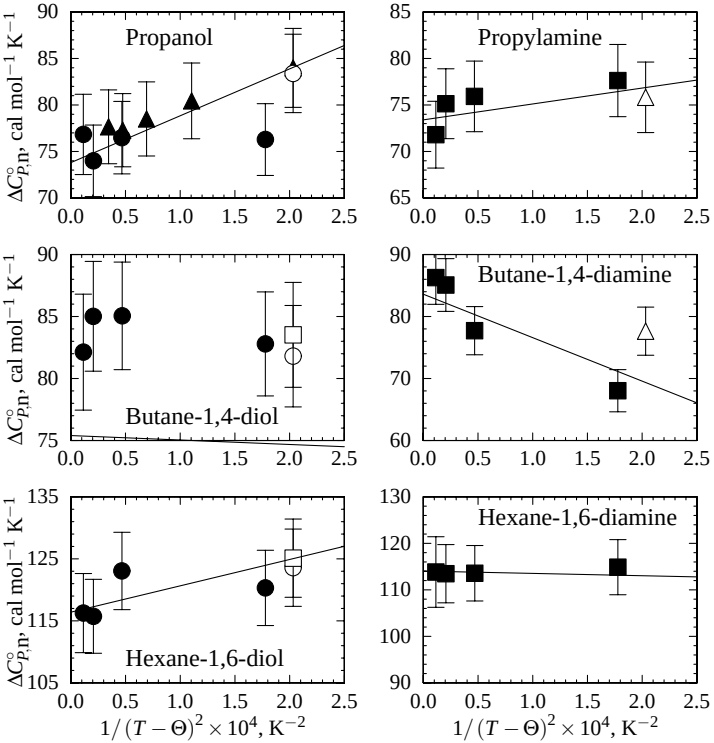

Fig. 17. $\Delta C^{\circ}_{P,n}$ as a function of $1/(T-\Theta)^2$ of aqueous propanol, butane-1,4-diol, hexane-1,6-diol, propylamine, butane-1,4-diamine, and hexane-1,6-diamine. The symbols represent values of $\Delta C^{\circ}_{P,n}$ calculated from Eqs. (A5) and (A10) using values of ω estimated with the group contributions given in Table 10 and experimental values of C°_P from (◯) Inglese and Wood, 1996; (■) Inglese et al., 1997; (▲) Makhatadze and Privalov, 1989; (◯) Jolicoeur and Lacroix, 1976; (□) Nichols et al., 1976; (△) Cabani et al., 1981. The error bars represent an estimated 5% uncertainty in the value of C°_P. The intercepts and slopes of the lines correspond to values of c_1 and c_2 that can be calculated with the organic group contributions given in Table 10.

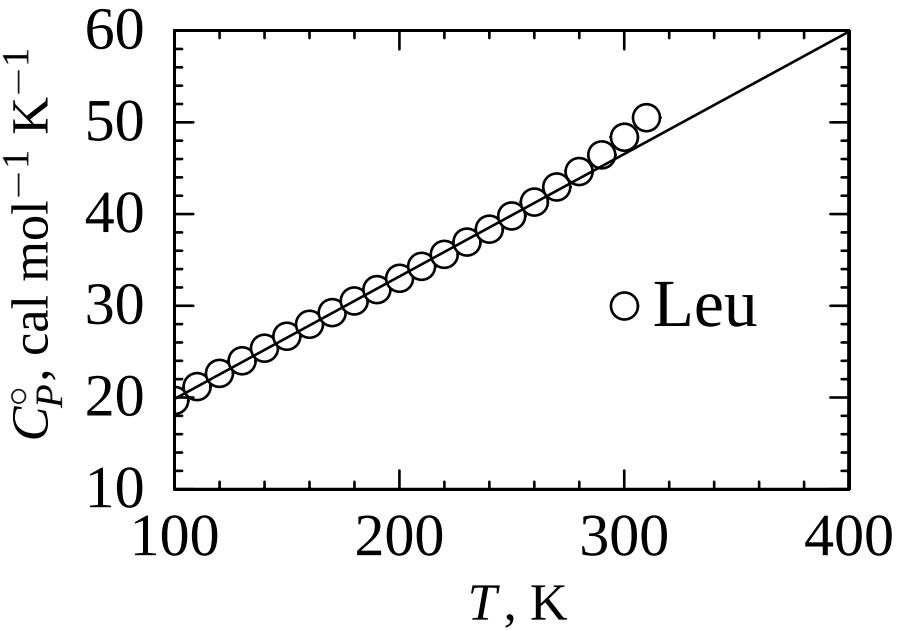

Fig. 18. C_P° of crystalline Leu as a function of temperature. The symbols represent experimental data taken from Hutchens et al. (1963), but the regression line corresponds to Eq. (D1) with a = 6.5 cal mol^{-1} K^{-1}, b = 0.1335 cal mol^{-1} K^{-2} and c = 0 cal mol^{-1} K.

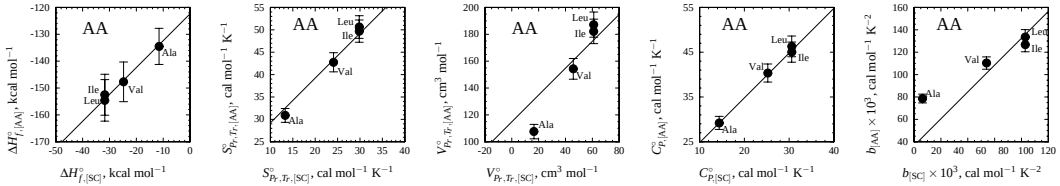

Fig. 19. ΔH°_f, $S^{\circ}_{P_r,T_r}$, V° and C°_P at 25°C and 1 bar and b of crystalline amino acids as a function of the corresponding property of the resepective sidechain groups, calculated using the group contributions given by Richard and Helgeson (1998).

BGD
2, 1515–1615, 2005

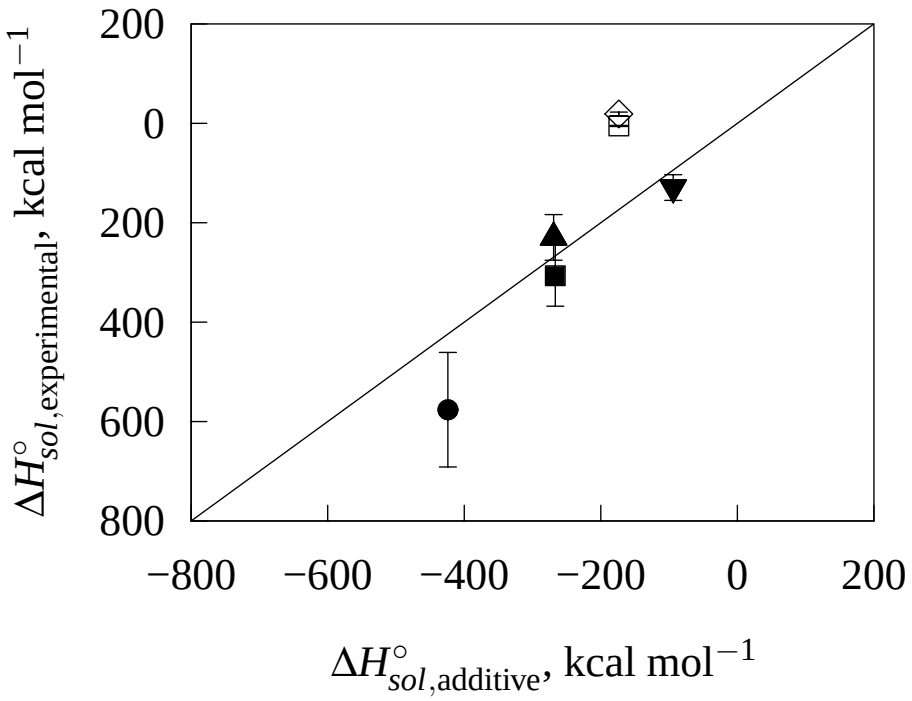

Fig. 20. Comparison of calculated and experimental values of ΔH°_{sol} of DRN1_HUMAN (◯), SOMA_HUMAN (■), Protropin_HUMAN (▲) and IGFA_HUMAN (▼). Error bars represent ±20% uncertainty in the experimental values of ΔH°_{sol} taken from Souillac et al. (2002). Symbols represent values of ΔH°_{sol} calculated for nonionized unfolded proteins. Also shown for LYC_CHICK are calculated values of ΔH°_{sol} and the experimental values of $-\Delta H^{\circ}_{cr}$ of the orthorhombic (◊) and tetragonal (□) crystal forms taken from Howard et al. (1988).

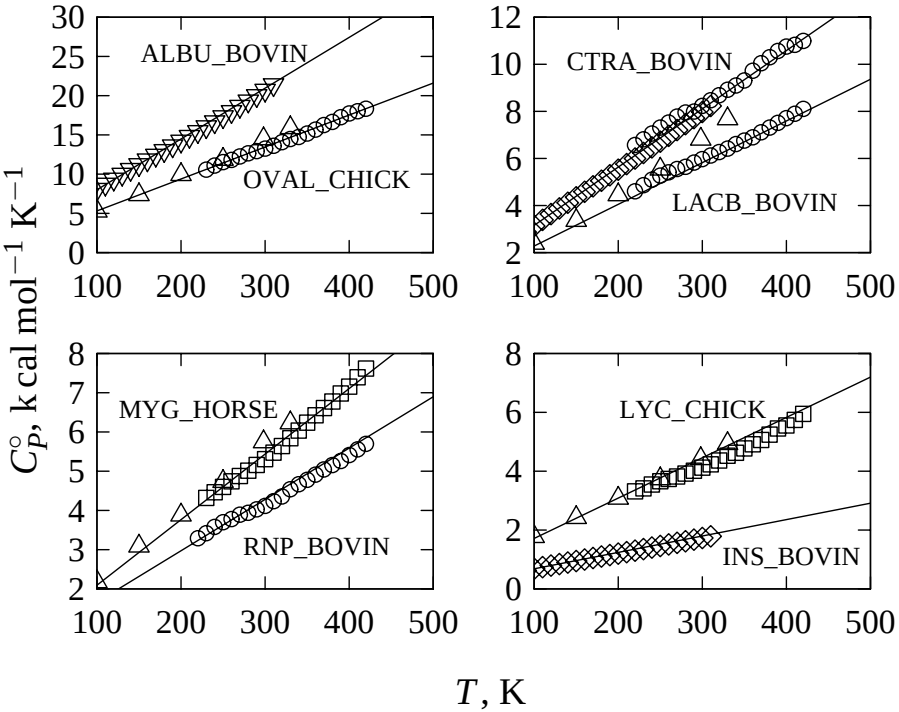

Fig. 21. C_P° of crystalline proteins as a function of temperature. The symbols represent experimental values of C_P° taken from (◯) Zhang et al. (1996), (□) Di Lorenzo et al. (1999), (△) Kulagina et al. (2001), (▽) Mrevlishvili (1986) and (◇) Hutchens et al. (1969). The regression curves are consistent with Eq. (D1) and the values of a and b of the crystalline sidechain and backbone groups given in Table 16.

BGD
2, 1515–1615, 2005