ELSEVIER

Biophysical Chemistry 84 (2000) 105–136

Biophysical Chemistry

www.elsevier.nl/locate/bpc

Calculation of the standard molal thermodynamic properties of aqueous biomolecules at elevated temperatures and pressures II. Unfolded proteins

Jan P. Amend[a,*], Harold C. Helgeson[b]

[a] *Department of Earth and Planetary Sciences, Washington University, St. Louis, MO 63130, USA*
[b] *Department of Geology and Geophysics, University of California, Berkeley, CA 94720, USA*

Received 8 November 1999; received in revised form 16 December 1999; accepted 16 December 1999

Abstract

Equations of state for completely unfolded proteins have been generated from group additivity algorithms and the revised Helgeson–Kirkham–Flowers (HKF) equations of state to compute the standard molal thermodynamic properties of these molecules at elevated temperatures and pressures. The requisite equations of state parameters were computed from those of groups retrieved by regression of experimental calorimetric and densimetric data reported in the literature. This approach permits calculation of the standard molal thermodynamic properties as a function of temperature and pressure for any completely unfolded protein for which the amino acid sequence is known. Calculations of this kind have been carried out for 11 thermophilic proteins. The thermodynamic properties reported below can be combined with those for protein unfolding to compute the corresponding properties of completely folded (i.e. native) proteins.

Keywords: Unfolded protein; Thermodynamics; Temperature; Pressure; Group additivity; Equations of state

* Corresponding author. Tel.: +1-314-935-8651; fax: +1-314-935-7361.
E-mail address: amend@levee.wustl.edu (J.P. Amend)

0301-4622/00/$ - see front matter
PII: S0301-4622(00)00116-2

1. Introduction

Interest in the thermodynamic behavior of thermophilic enzymes has increased dramatically in the last decade, largely because they have come to be recognized as effective biocatalysts with greater efficiencies and half-lives than their low-temperature counterparts [1]. The thermal stability of enzymes isolated from thermophilic microorganisms are controlled to a large extent by the relative standard molal enthalpies and heat capacities of the folded and unfolded protein species, which can be calculated from equations of state [2–4]. The parameters in these equations can be quantified with the aid of group additivity algorithms by fitting them to experimental data reported in the literature.

Enthalpies and heat capacities of denaturation as a function of temperature for dozens of proteins have been measured calorimetrically over the last several decades [5–11]. In addition, heat capacities and volumes of four denatured proteins (cytochrome *c*, lysozyme, myoglobin, and ribonuclease) have been determined experimentally at various temperatures between 5 and 125°C [12,13]. Makhatadze and Privalov [14], Makhatadze et al. [13], and Privalov and Makhatadze [15] used group contributions and reported amino acid sequences of cytochrome *c*, lysozyme, myoglobin, and ribonuclease, together with experimental heat capacities and volumes of various aqueous organic model compounds at temperatures between 5 and 125°C to predict values of these properties for the completely unfolded proteins. These predicted properties were then compared with the experimental heat capacities and volumes of the corresponding acid- and heat-denatured proteins. The differences between the calculated values and the bulk of the experimental data over this temperature range are less than 5 and 3% for the heat capacities and volumes, respectively, which are within the experimental uncertainty [13,15]. This observation led to the conclusion that the structural state of a protein resulting from heat-denaturation or exposure to strong denaturants (such as 8 M urea or guanidinium hydrochloride) corresponds to that of the completely unfolded protein, void of residual structure and fully exposed to the solvent [12,15]. Privalov, Makhatadze, and their students have since used their additivity algorithm and model compounds, together with a linear approximation of the dependence of the heat capacity of the folded protein on temperature, to extrapolate calorimetric unfolding data for a variety of proteins to temperatures as high as 125°C.

The purpose of the present communication is to explore, within the context of classical solution chemistry, an alternate approach using group additivity algorithms to generate equations of state parameters for unfolded proteins. These parameters can be used in conjunction with corresponding equations of state for their folded counterparts to predict accurately the relative stabilities of thermophilic and hyperthermophilic enzymes [4]. Only the completely unfolded state is considered in the present communication, which provides a reference frame for further thermodynamic investigation of protein stability. The term unfolded protein is used below to refer to a polypeptide chain which is completely exposed to the solvent and in a random-coil formation with no residual structure. This use is consistent with the definition of unfolded proteins given by Dill and Shortle [16].

2. Standard state conventions

Standard molal Gibbs free energies and enthalpies of aqueous species are expressed in this communication as *apparent* standard molal Gibbs free energies and enthalpies of formation (ΔG° and ΔH°, respectively) which are defined by [17,18]

$$\Delta G^{\circ} \equiv \Delta G_f^{\circ} + \left(G_{P,T}^{\circ} - G_{P_r,T_r}^{\circ}\right) \quad (1)$$

and

$$\Delta H^{\circ} \equiv \Delta H_f^{\circ} + \left(H_{P,T}^{\circ} - H_{P_r,T_r}^{\circ}\right) \quad (2)$$

where ΔG_f° and ΔH_f° refer to the standard molal Gibbs free energy and enthalpy of formation of the species from the elements in their stable form

at the reference pressure (P_r) and temperature (T_r) of 1 bar and 298.15 K, and $G^o_{P,T} - G^o_{P_r,T_r}$ and $H^o_{P,T} - H^o_{P_r,T_r}$ stand for the differences in the standard molal Gibbs free energy and enthalpy, respectively, of the aqueous species at the pressure (P) and temperature (T) of interest, and those at P_r and T_r. The standard state convention adopted for H_2O is one of unit activity of pure water at any temperature and pressure, but that for aqueous species other than H_2O corresponds to unit activity of the species in a hypothetical 1 molal solution referenced to infinite dilution at any temperature and pressure. Note that the latter convention differs from the widely used biological standard state for aqueous species, which is restricted to neutral pH. The standard state adopted in the present study can be converted to its biological counterpart by taking explicit account of the contribution of neutral pH and/or its partial derivatives with respect to temperature or pressure [19,20].[1]

The standard molal thermodynamic properties of charged aqueous species discussed below are conventional properties consistent with

$$\Xi_j = \Xi^{abs}_j - Z_j \Xi^{abs}_{H^+} \quad (3)$$

where Ξ_j and Ξ^{abs}_j represent any given conventional and absolute standard molal property of the jth aqueous species, $\Xi^{abs}_{H^+}$ stands for the corresponding absolute standard molal property of the hydrogen ion, and Z_j denotes the charge on the subscripted species. The equations used in the present study to compute the standard molal properties of aqueous biomolecules at elevated temperatures and pressures are summarized in Appendix A. These equations are referred to as the revised HKF equations of state, which are based on electrostatic theory [18] and scaling laws [21] for solute–solvent interaction (see Appendix A). They have been widely used to represent the standard molal properties of a variety of charged and neutral aqueous organic, inorganic, and biochemical species over broad ranges of temperature and pressure [4,20,22–48]. The parameters required to calculate the standard molal properties of unfolded proteins as a function of temperature and pressure using the revised HKF equations of state can be generated from group additivity relations in the manner described below.

3. Group additivity relations for unfolded proteins

Group additivity algorithms were first used to estimate thermodynamic properties of proteins at ambient conditions by Cohn and Edsall [49]. Jacobsen [50] later employed the group additivity approach to estimate the hydrational contribution to the compressibilities of proteins using group contributions computed from the adiabatic compressibilities of amino acids and alcohols in various solvents. In recent years, similar strategies have been used to calculate the thermodynamic properties of protein unfolding at temperatures ranging from 5 to 125°C at P_{SAT} [15,51–56].[2] In contrast to the additivity algorithms employed in these various studies, the group additivity approach described in the following pages was used to generate equations of state parameters for unfolded proteins which can be combined with the equations of state summarized in Appendix A to calculate the thermodynamic properties of these proteins as a function of temperature and pressure.

The group inventory adopted in the present study for unfolded proteins can be depicted schematically by writing

$$H_2N\text{–}AA_1\text{–}AA_2\text{–}\ldots\text{–}AA_{n-1}\text{–}CHRCOOH$$

where $-AA_1-$, $-AA_2-$, $-AA_{n-1}-$ represent the internal amino acid residues,[3] n designates the total number of residues in the protein, $-NH_2$

[1] It should perhaps be emphasized in this regard that the value of neutral pH changes appreciably with increasing temperature and/or pressure from 7.0 at 25°C and 1 bar to ~6.0 at 100°C and 100 bars.

[2] P_{SAT} is used throughout this study to represent pressures corresponding to liquid–vapor equilibrium for the system H_2O, except at temperatures < 100°C where it denotes the reference pressure (P_r) of 1 bar.

and −CHRCOOH stand for the N- and C-terminal groups, respectively, and in this case R corresponds to the side chain in the C-terminal group. It follows from this schematic and the group additivity hypothesis that any standard molal thermodynamic property and equation of state parameter (see below) for an unfolded protein (Ξ_P) can be expressed as

$$\Xi_P = \Xi_{-NH_2} + \Xi_{-AA_1-} + \Xi_{-AA_2-} + \ldots + \Xi_{-AA_{n-1}-} + \Xi_{-CHRCOOH} \quad (4)$$

where Ξ_{-AA_1-}, Ξ_{-AA_2-}, $\ldots\Xi_{-AA_{n-1}-}$, Ξ_{-NH_2}, and $\Xi_{-CHRCOOH}$ denote corresponding standard molal properties or equations of state parameters for the subscripted groups in the protein.

The −CHRCOOH group in Eq. (4) consists of the two sub-groups R and >CHCOOH, the latter of which can be combined with $-NH_2$ to give $-CHNH_2COOH$, which is equivalent in composition and idealized structure to the amino acid backbone, denoted below as AAB. The equations of state parameters and standard molal thermodynamic properties at 25°C and 1 bar for this group were determined in the first part of this series of papers [20]. Corresponding values of Ξ for the amino acid residues shown in Eq. (4) can be calculated from the group additivity relation represented by

$$\Xi_{-AA-} = \Xi_{PB} + \Xi_R \quad (5)$$

where Ξ_{PB} and Ξ_R denote the property or parameter of the protein backbone (see footnote 2) and R groups, respectively. It follows from Eqs. (4) and (5) that

$$\Xi_P = \Xi_{AAB} + (n-1)\Xi_{PB} + \sum_{i=1}^{\hat{i}} n_i \Xi_i \quad (6)$$

where n_i stands for the number of moles of the ith side chain in one mole of the protein, ($i = 1, 2, \ldots \hat{i}$), Ξ_i refers to any equation of state parameter or standard molal property at 25°C and 1 bar for the subscripted side chain, and $\hat{i}$ represents the total number of different kinds of R groups in the protein, including both neutral and ionized side chains.

In a first approximation, the hypothesis is adopted that the thermodynamic properties of corresponding side chains, other than that of glycine (see below), in amino acids and unfolded proteins are essentially equivalent. This hypothesis is based on the observation that the structures of the R groups in the amino acids and unfolded proteins are similar, and that the Zwitterion contribution to the values of Ξ_{AAB} given by Amend and Helgeson [20] are negligible. Values of Ξ_i in Eq. (6) can be taken from the tabulation of equations of state parameters and standard molal properties at 25°C and 1 bar for the R groups given by Amend and Helgeson [20]. The revised HKF equations of state parameters generated by Marriot et al. [48] from fits of the equations of state to their experimental values of the heat capacities and volumes of 17 aqueous amino acids at 15, 25, 40, and 55°C were not used in the present study because they were derived in part by using the effective Born coefficient (ω) as an adjustable fit parameter. Because C_P^o and V^o of aqueous species at these low temperatures are relatively insensitive to the magnitude of ω, the regression procedure used by Marriot et al. [48] may lead to substantial error in predicted values of the standard molal thermodynamic properties of the amino acids at elevated temperatures. It can be shown that in most cases the predicted values of C_P^o and V^o computed using the revised HKF equation of state parameters given by Amend and Helgeson [20] compare favorably with the corresponding experimental values at 15, 25, 40, and 55°C reported by Marriot et al. [48]. The approach described here almost certainly affords more realistic values of Ξ_i in Eq. (6) than those

[3]The term 'amino acid residue' (−AA−) is used in the present communication to refer to the group in a peptide or protein consisting of a protein backbone (PB) and a side chain (R). The idealized structures and chemical formulas of −AA− and PB can be represented by −CHRCONH− and >CHCONH−, respectively, where > denotes single bonds to two different groups and − depicts a single bond between any two groups.

postulating additivity of the properties and parameters of smaller groups to generate those of the R groups in unfolded proteins. Only as a last resort was the latter approach used by Amend and Helgeson [20] to generate first approximations of c_1 and c_2 for certain R groups.

The tabulation of equations of state parameters and standard molal properties at 25°C and 1 bar given by Amend and Helgeson [20] was adopted in the present study for all but C_P^o and V^o of R_{Gly} because all of the other side chains are polyatomic and the values are internally consistent and compatible with the thermodynamic data for the other amino acids and similar species reported in the literature. In the case of R_{Gly}, the fact that the side chain consists only of −H renders the group additivity approach to calculating values of C_P^o and V^o of R_{Gly} from those of Gly highly problematical. Fortunately, adequate experimental data are available in the literature for glycyl peptides to permit an alternate approach to be used to generate more reliable values of C_P^o and V^o for the −H side chains in proteins. Accordingly, these data were regressed as a function of temperature at P_{SAT} with the revised HKF equations of state given in Appendix A to retrieve values of C_P^o, V^o, c_1, c_2, σ, and ξ for the Gly residue (−Gly−), which were then combined in a group additivity algorithm with those for PB computed below to generate values of these properties and parameters for R_{Gly} in peptides, and by inference in unfolded proteins.

4. Calculation of the standard molal thermodynamic properties of –Gly– and PB as a function of temperature at P_{SAT}

The standard molal thermodynamic properties of −Gly− and PB can be estimated as a function of temperature and pressure from Eqs. (1) and (2), and those summarized in Appendix A by taking account of experimental data reported in the literature for aqueous diketopiperazine (cyclic diglycine), glycyl peptides, and unfolded proteins.

4.1. Standard molal Gibbs free energies and enthalpies of formation (ΔG_f^o and ΔH_f^o) and standard molal entropies (S^o) at 25°C and 1 bar

Values of ΔG_f^o and ΔH_f^o of PB were determined from experimental values of the corresponding properties of diketopiperazine taken from Shock [33] by first computing values of ΔG_f^o and ΔH_f^o of −Gly− from the relation

$$\Xi_{-\mathrm{Gly}-} = \frac{\Xi_{\mathrm{Diketopiperazine}}}{2} \tag{7}$$

where $\Xi_{-\mathrm{Gly}-}$ and $\Xi_{\mathrm{Diketopiperazine}}$ designate ΔG_f^o or ΔH_f^o of the subscripted species. The computed values of ΔG_f^o and ΔH_f^o of −Gly− were then combined with those of R_{Gly} given by Amend and Helgeson [20] to retrieve values of ΔG_f^o and ΔH_f^o of PB from a modified statement of Eq. (5) corresponding to

$$\Xi_{\mathrm{PB}} = \Xi_{-\mathrm{Gly}-} - \Xi_{\mathrm{R_{Gly}}} \tag{8}$$

where Ξ_{PB}, $\Xi_{-\mathrm{Gly}-}$, and $\Xi_{\mathrm{R_{Gly}}}$ denote ΔG_f^o or ΔH_f^o of the subscripted species. The values of S^o of −Gly− and PB were computed from those of ΔG_f^o and ΔH_f^o using values of S^o of the elements taken from Cox et al. [57]. The calculated values of ΔG_f^o, ΔH_f^o, and S^o of −Gly− and PB are given in Table 1.

4.2. Standard molal heat capacities and volumes as a function of temperature at P_{SAT}

Experimental values of C_P^o and V^o reported in the literature for glycine and glycyl peptides at

Table 1
Values of ΔG_f^o, ΔH_f^o, and S^o at 25°C and 1 bar for −Gly− and PB computed from Eqs. (7) and (8) (see text)

Group	ΔG_f^o (kcal mol^{-1})	ΔH_f^o (kcal mol^{-1})	S^o (cal mol^{-1} K^{-1})
−Gly−	−28.72	−49.65	26.81
PB	−20.15	−40.49	13.18

Table 2
Values of C_P^o and V^o reported in the literature for glycine and glycyl peptides at 25°C and 1 bar[a]

Compound	C_P^o (cal mol^{-1} K^{-1})	V^o (cm^3 mol^{-1})
Glycine	9.4[1], 8.8[2], 10.9[3], 7.4[4], 14.6[5], 11.2[6]	43.19[9,10], 43.25[1], 43.33[11], 43.3[12,13], 43.2[14], 43.09[15], 43.22[16], 43.26[17]
Diglycine	25.1[1], 24.1[3], 38.0[4], 39.9[5]	76.23[9], 76.27[1], 76.34[18], 77.2[19], 76.76[20]
Triglycine	44.4[1], 52.1[5], 41.9[7], 45.0[8]	112.11[9], 111.81[1], 113.5[19], 112.9[21], 112.51[20]
Tetraglycine	67.6[1], 63.2[7]	149.6[9], 149.7[1], 149.98[20] 151.1[21]
Pentaglycine	89.1[1], 86.4[7]	187.1[1], 189.5[21]

[a]*Notes*. 1, Jolicoeur and Boileau [58]; 2, Gucker et al. [61]; 3, Hedwig et al. [62]; 4, Kresheck and Benjamin [63]; 5, Cabani et al. [64]; 6, Spink and Wadsö [65]; 7, Makhatadze and Privalov [14]; 8, Downes and Hedwig [66]; 9, Mishra and Ahluwalia [67]; 10, Millero et al. [68]; 11, Cabani et al. [69]; 12, Kharakoz [70]; 13, Kirchnerova et al. [71]; 14, Chalikian et al. [72]; 15, Bülent Belibagli and Ayranci [73]; 16, Lark and Bala [74]; 17, Wadi and Goyal [75]; 18, Ellerton et al. [76]; 19, Cohn et al. [77]; 20, Iqbal and Verrall [78]; 21, Makhatadze et al. [53].

various temperatures are given in Tables 2–4. The values for 25°C in Table 2 are plotted as symbols in Fig. 1, where it can be seen that the values of C_P^o and V^o of the glycyl peptides behave linearly with glycine number (n_{Gly}) for $n_{Gly} \geq 3$. Deviations from linearity at low values of n_{Gly} are an expected consequence of the effect of terminal groups on the thermodynamic and electrostatic properties of short-chain species [24,35,45].

The reported uncertainties associated with the experimental data plotted in Fig. 1 are approximately consistent with the size of the symbols. However, it can be deduced from the distribution of the C_P^o and V^o data taken from different sources for the same value of n_{Gly} that these uncertainties are considerably less than those represented by the scatter of the data. For example, reported experimental uncertainties associated with the values of C_P^o of diglycine (C_{P,Gly_2}^o) range from ± 0.1 to ± 5.0 cal mol^{-1} K^{-1}, but the values of C_{P,Gly_2}^o at 25°C and 1 bar reported by different investigators differ by as much as 15.8 cal mol^{-1} K^{-1}. Although this large discrepancy was acknowledged by Jolicoeur and Boileau [58], no adequate explanation of it was provided. It should perhaps be noted that the two most recent values of C_{P,Gly_2}^o given in Table 2 are in close agreement with each other and more or less consistent with the values of C_P^o for the homologous series of glycylpeptides shown in Fig. 1a. Relative to the experimental values of C_P^o, those

Table 3
Experimental standard molal heat capacities (C_P^o, cal mol^{-1} K^{-1}) reported in the literature for glycyl peptides at 1 bar and various temperatures other than 25°C (see Table 1)

Glycyl peptide	Temperature (°C)					
	30	50	75	100	125	Ref.
Diglycine	33.0					[79]
Triglycine		57.2	68.1	77.1	86.2	[14]
		59.3	70.1	80.0	87.6	[66]
	57.1					[79]
Tetraglycine		80.8	91.4	98.5	104.8	[14]
Pentaglycine		103.0	113.2	118.8	124.4	[14]

Table 4
Experimental standard molal volumes (V^o, cm^3 mol^{-1}) reported in the literature for glycine and glycyl peptides at 1 bar and various temperatures other than 25°C (see Table 1)

Compound	Temperature (°C)									
	15	18	30	35	40	45	50	55	75	Ref.
Glycine	42.24			44.26		45.09		45.91		[69]
	42.4			43.8	43.9			44.3		[70]
		42.7			44.0			44.2		[72]
			43.59	43.79	44.15					[74]
	42.35			44.12						[75]
Triglycine							114.8		116.7	[53]
Tetraglycine							153.3		155.2	[53]
Pentaglycine							191.2		192.9	[53]

of V^o shown in Fig. 1b for a given value of n_{Gly} differ by only 2.4 cm^3 mol^{-1} or less.

The equations of the regression lines in Fig. 1a,b for $n_{\mathrm{Gly}} \geq 3$ are given by

$$C^{\circ}_{P,\mathrm{Gly}_n} = -25.72 + 22.43 n_{\mathrm{Gly}} \quad (9)$$

and

$$V^{\circ}_{\mathrm{Gly}_n} = -4.10 + 38.73 n_{\mathrm{Gly}} \quad (10)$$

where $C_P,^{\circ}_{\mathrm{Gly_n}}$ and $V^{\circ}_{\mathrm{Gly}_n}$ denote C°_P and V°, respectively, of the subscripted species, and the slopes represent the values of $C_P,^{\circ}_{-\mathrm{Gly}-}$ and V°_{-Gly-}. Values of C°_P and V° of Gly and Gly_n at 25°C and 1 bar computed from Eqs. (9) and (10) are given in Table 5 for $3 \leq n_{\mathrm{Gly}} \leq 5$, and values of $C^{\circ}_{P,-\mathrm{Gly}-}$ and $V^{\circ}_{-\mathrm{Gly}-}$ are given in Table 6. The values of the equations of state parameters required to calculate the standard molal properties of −Gly− at higher temperatures and pressures were generated from those of Gly_3, Gly_4, and Gly_5 in the manner described below.

The values of V°_{Gly} and those of C°_P and V^o of Gly_3, Gly_4, and Gly_5 given in Tables 2–4 were combined with corresponding values of $\Delta C^{\circ}_{P,s}$ and ΔV°_s to determine values of $\Delta C^{\circ}_{P,n}$ and ΔV°_n from the first identities in Eqs. (A2) and (A3), respectively. The values of $\Delta C^{\circ}_{P,s}$ and ΔV°_s used in the calculations were determined from Eqs. (A10) and (A11), respectively, using values of ω taken

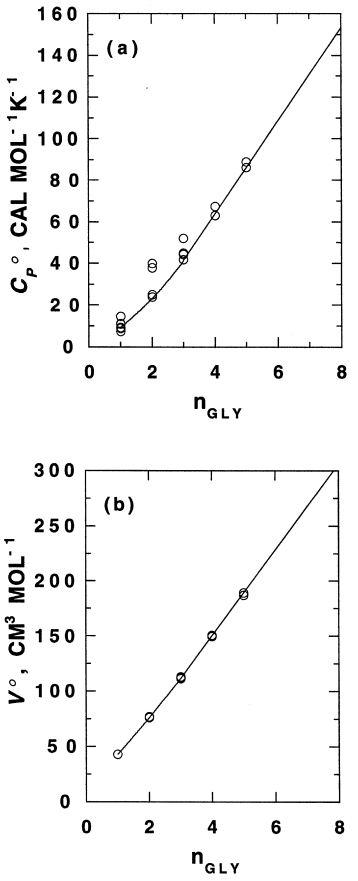

Fig. 1. Correlation of C°_P (a) and V° (b) with glycine number (n_{Gly}) at 25°C and 1 bar. The open circles represent experimental values taken from Table 1. The regression lines for $n_{\mathrm{Gly}} \geq 3$ are consistent with Eqs. (9) and (10).

Table 5
Values of C_P^o, c_1, c_2, V^o, σ, ξ, and ω for glycine and glycyl peptides consistent with the curves and regression lines shown in Figs. 1–4 (see text)

Species	C_P^o (cal mol^{-1} K^{-1})	c_1 (cal mol^{-1} K^{-1})	$c_2 \times 10^{-4}$ (cal K mol^{-1})	V^o (cm^3 mol^{-1})	σ (cm^3 mol^{-1} at 1 bar)	$\xi \times 10^{-2}$ (cm^3 K mol^{-1} at 1 bar)	$\omega \times 10^{-5}$[a] (cal mol^{-1})
Gly	9.4	16.9122[b]	−4.8400[b]	43.19	47.6140[b]	−3.6000[b]	−0.2550
Gly_2	24.1			76.23			
Gly_3	41.57	83.1693	−22.2061	112.09	121.2630	−7.1863	−0.3870
Gly_4	64.00	102.2541	−20.8558	150.82	158.0143	−5.9273	−0.4530
Gly_5	86.43	121.3389	−19.5055	189.55	194.7656	−4.6683	−0.5190

[a] Computed in Appendix A.
[b] Amend and Helgeson [20].

Table 6
Summary of the standard molal thermodynamic properties at 25°C and 1 bar of PB, −Gly−, and R_{Gly} generated in the present study (see text) and those of AAB and other R groups given by Amend and Helgeson [20]

Group	ΔG_f^o (kcal mol^{-1})	ΔH_f^o (kcal mol^{-1})	S^o (cal mol^{-1} K^{-1})	C_P^o (cal mol^{-1} K^{-1})	V^o (cm^3 mol^{-1})
AAB	−82.38	−115.62	25.67	10.82	32.26
PB	−20.15	−40.49	13.18	12.24	25.61
−Gly−	−28.72	−49.65	26.81	22.43	38.73
R_{Gly}	−8.57	−9.16	13.62	10.19	13.12
R_{Ala}	−6.43	−16.88	13.16	22.98	28.21
R_{Val}	−2.95	−30.80	20.01	61.38	58.65
R_{Leu}	−1.82	−37.98	24.74	84.28	75.31
R_{Ile}	0.39	−35.98	24.04	80.78	73.19
R_{Ser}	−41.54	−56.80	21.57	17.28	28.31
R_{Thr}	−37.45	−63.32	18.56	39.18	44.60
R_{Asp}	−89.87	−110.72	29.55	19.58	41.57
R_{Asp-}	−84.79	−108.92	17.67	−10.02	28.62
R_{Glu}	−90.67	−119.20	35.53	31.48	57.10
R_{Glu-}	−84.83	−118.53	18.19	−4.52	42.79
R_{Asn}	−43.11	−67.08	32.21	19.08	44.94
R_{Gln}	−43.90	−76.24	36.74	33.88	61.54
R_{His}	33.96	7.42	40.34	46.78	66.53
R_{His+}	25.81	0.42	44.19	40.85	66.00
R_{Lys}	1.70	−43.71	32.24	52.98	76.45
R_{Lys+}	−10.51	−54.75	36.15	47.05	76.34
R_{Arg}	25.06	−27.44	54.25	55.88	91.44
R_{Arg+}	12.49	−39.26	56.76	49.95	90.50
R_{Cys}	2.05	−8.45	20.65	34.08	41.18
R_{Met}	−37.74	−62.90	36.69	59.18	73.09
R_{Phe}	32.95	6.72	30.93	80.98	89.66
R_{Tyr}	−9.42	−42.12	33.74	60.68	90.74
R_{Trp}	55.56	18.03	34.29	89.58	111.54
R_{Pro}[a]	8.82	−10.82	16.31	30.38	50.24

[a] Because proline is an imino acid, its R group contains two unsatisfied bonds, rather than one, which is the case for all of the other R groups shown in the table. The equations of state parameters and standard molal properties at 25°C and 1 bar for R_{Pro} were computed from, $\Xi_{R_{Pro}} = \Xi_{Pro} - \Xi_{AAB}$, where $\Xi_{R_{Pro}}$, Ξ_{Pro}, and Ξ_{AAB} stand for the corresponding parameters and properties of the subscripted species. Values of Ξ_{Pro} and Ξ_{AAB} are given in Amend and Helgeson [20].

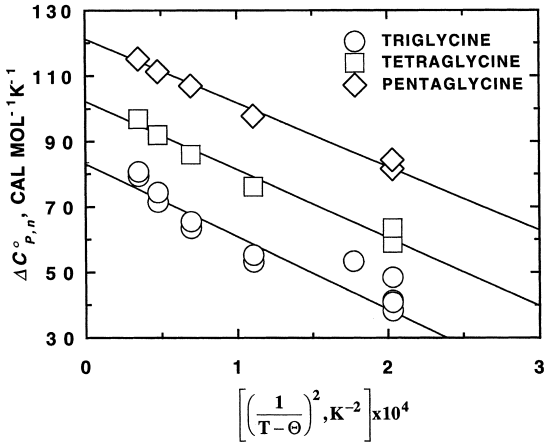

Fig. 2. $\Delta C^{\circ}_{P,n}$ for Gly_3, Gly_4, and Gly_5 as a function of $1/(T-\Theta)^2$ at P_{SAT}. The symbols represent values of $\Delta C^{\circ}_{P,n}$ computed from Eqs. (A2) and (A10) using values of C°_P and ω taken from Tables 2 and 3, and Table A1. The regression lines represent fits of Eq. (A9) to the data.

from Table A1 in Appendix A and those of the Born functions tabulated by Shock et al. [26]. Values of the non-solvation parameters c_1 and c_2 for Gly_3, Gly_4, and Gly_5 were obtained by regression with Eq. (A9) of the experimental values of $\Delta C^{\circ}_{P,n}$ represented by the symbols in Fig. 2. In an analogous manner, values of the non-solvation parameters σ and ξ for Gly, Gly_3, Gly_4, and Gly_5 were generated by regression with Eq. (A8) of the experimental values of ΔV°_n depicted by the symbols in Fig. 3. It can be seen in Figs. 2 and 3 that the regression lines representing the equations of state are in close agreement with the bulk of the experimental data. Unless otherwise denoted by error bars, the reported uncertainties associated with these data are approximately consistent with the size of the symbols. However, it can be deduced from the distribution of the symbols in

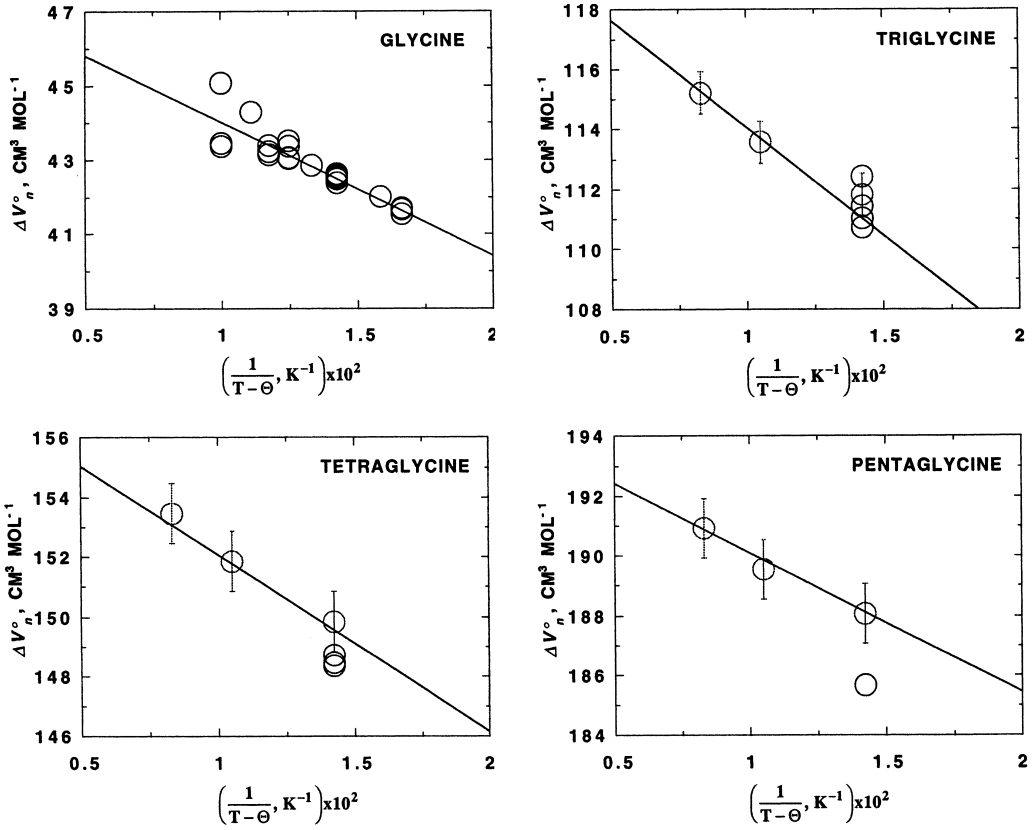

Fig. 3. ΔV°_n as a function of $1/(T-\Theta)$ for Gly (a), Gly_3 (b), Gly_4 (c), and Gly_5 (d). The symbols represent values of ΔV°_n computed from Eqs. (A3) and (A11) using values of V° and ω taken from Tables 2, 4, and A1. The regression lines represent fits of Eq. (A8) to the data.

Table 7
Summary of equation of state parameters for PB, −Gly−, and R_{Gly} generated in the present study (see text) and those for AAB and other R groups given by Amend and Helgeson [20]

Group	σ ($cm^3\ mol^{-1}$ at 1 bar)	$\xi \times 10^{-2}$ ($cm^3\ K\ mol^{-1}$ at 1 bar)	$a_1 \times 10$ (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 K mol^{-1})	c_1 (cal mol^{-1} K^{-1})	$c_2 \times 10^{-4}$ (cal K mol^{-1})	$\omega \times 10^{-5}$ (cal mol^{-1})
AAB	38.2095	−4.2650	9.9970	−2.2486	−1.0953	−2.3665	26.1454	−7.7600	−0.0487
PB	33.6126	−5.3413	9.6313	−4.1553	−4.6186	−2.1192	25.5240	−5.9094	0.1403
−Gly−	36.7513	1.2590	9.7377	−2.4807	11.9023	−2.3131	19.0840	1.3503	−0.0660
R_{Gly}	3.1387	6.6003	0.1064	1.6746	16.5209	−0.1939	−6.4400	7.2597	−0.2063
R_{Ala}	25.5268	1.4908	5.3484	1.9749	5.0072	−0.3756	10.5608	5.2180	−0.1993
R_{Val}	57.9238	−0.0840	11.4296	6.2815	14.4232	−3.8037	39.3764	9.4696	−0.3030
R_{Leu}	76.6808	−1.6947	14.9109	8.9067	18.4097	−5.8419	56.9534	11.7676	−0.3747
R_{Ile}	74.5905	−1.6947	14.3452	9.0698	18.4097	−5.8419	53.5509	11.7676	−0.3640
R_{Ser}	27.8363	0.3549	5.9405	1.8536	3.0120	−0.5628	22.4267	−2.4726	0.0134
R_{Thr}	44.5702	0.1319	9.2035	3.7694	7.9250	−1.9793	40.0714	−0.1746	0.0589
R_{Asp}	42.8821	−1.1332	9.1493	2.8957	4.2842	−1.8188	30.5480	−5.8793	−0.1075
R_{Asp-}	27.5170	3.4212	8.1337	−4.0494	0.7483	1.9322	21.5210	−9.4180	1.3614
R_{Glu}	58.5804	−1.4277	12.4529	4.0285	4.8156	−2.1401	36.9528	−3.5813	−0.1981
R_{Glu-}	43.1232	2.3986	11.2791	−2.5289	2.3960	0.8679	25.6629	−8.7846	1.3536
R_{Asn}	45.9615	−1.0082	9.9803	2.6139	3.4261	−1.5179	29.8295	−5.9529	−0.1479
R_{Gln}	62.7913	−1.3027	13.5672	3.7467	3.9575	−1.8392	39.3348	−3.6549	−0.2165
R_{His}	68.4904	−1.9061	14.3351	5.2924	8.1205	−3.2971	46.7049	−1.1775	−0.2709
R_{His+}	63.7982	1.6677	15.2632	−0.0385	7.3515	−0.8754	50.0042	−4.2134	0.0650
R_{Lys}	79.9946	−2.7786	16.2846	7.3736	10.7444	−4.5220	49.4263	1.0844	−0.1482
R_{Lys+}	76.1200	0.3362	17.7691	1.1036	7.2429	−1.6749	51.4635	−1.7460	0.0950
R_{Arg}	102.3673	−8.6068	21.0135	8.9820	1.9044	−5.8459	54.2131	−1.3388	−0.4816
R_{Arg+}	97.1540	−4.5739	22.2249	2.5903	−0.4915	−2.7156	59.2954	−4.3747	0.0500
R_{Cys}	40.8089	0.3098	8.1060	4.2858	10.6912	−2.5882	34.8088	−0.2367	0.0272
R_{Met}	74.9826	−1.7484	15.0953	7.3511	13.0402	−4.4787	43.6866	6.6573	−0.2157
R_{Phe}	91.5560	−2.2456	18.3752	9.1987	18.8750	−6.3054	70.5153	3.0498	−0.4684
R_{Tyr}	94.1361	−2.7190	19.0527	9.0729	14.0876	−5.3545	56.2289	1.4233	−0.1711
R_{Trp}	117.3652	−4.4370	24.2634	9.8526	12.6873	−6.0583	116.9001	−14.2481	−0.1793
R_{Pro}[a]	49.6130	−0.0447	9.3532	6.5149	11.8589	−3.1123	20.0533	3.9740	−0.2471

[a] See footnote in Table 6.

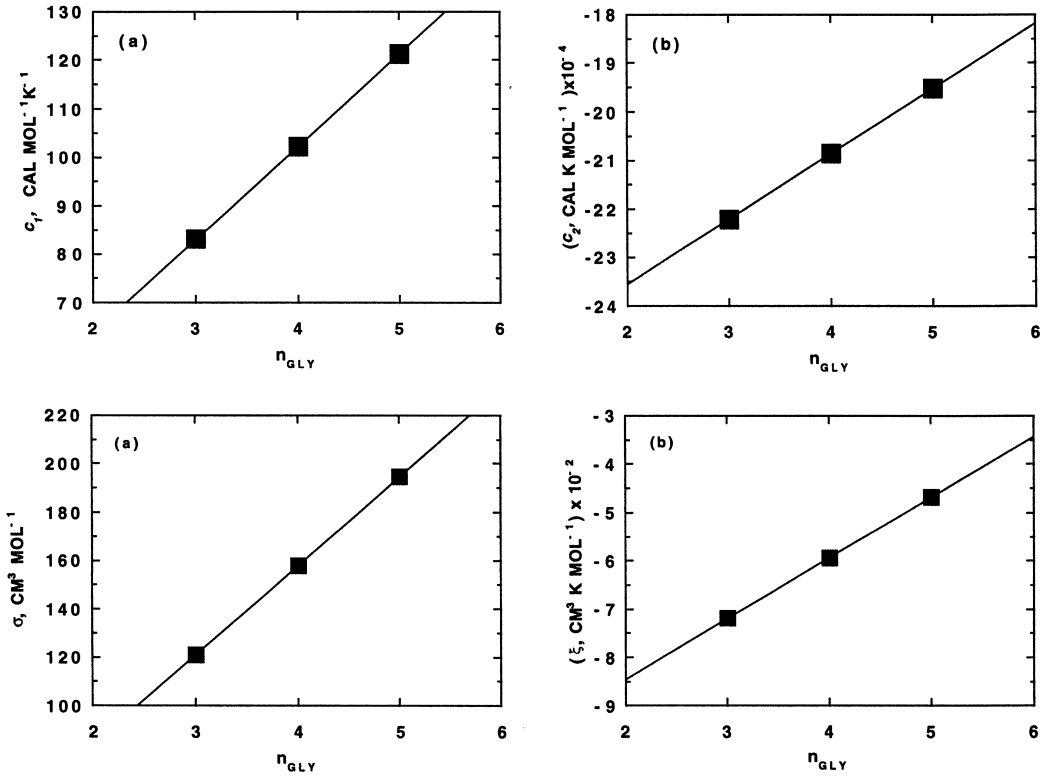

Fig. 4. Correlation of c_1 (a), c_2 (b), σ (c), and ξ (d) with n_{Gly} for glycyl peptides. The symbols depict the intercepts and slopes of the lines shown in Figs. 2 and 3. The slopes of the lines shown above correspond to the values of these parameters for −Gly− given in Table 7.

Table 8
Experimental standard molal heat capacities (C_P^o, cal mol^{-1} K^{-1}) and volumes (V^o, cm^3 mol^{-1}) of denatured cytochrome *c*, lysozyme, myoglobin, and ribonuclease at various temperatures reported by Makhatadze et al. [53] and Privalov and Makhatadze [15]

T (°C)	Cytochrome *c*		Lysozyme		Myoglobin		Ribonuclease	
	C_P^o	V^o	C_P^o	V^o	C_P^o	V^o	C_P^o	V^o
5	5115		6023	10270	7624	12520	5521	9860
10		8230						
15				10270		12680		9800
25	5688	8370	6573	10580	8341	12760	6214	9900
35				10780		12850		10000
40		8480				12930		
45				10880				10150
50	6023		7051		8819		6620	
55				10880				10220
60		8520						
65				11100				10720
75	6214	8660	7433		9369		6979	
100	6238		7505		9393		7122	
125	6286		7529		9441		7194	

these figures that in many cases the reported uncertainties are much smaller than those implied by the scatter of the data, which exceeds by far the discrepancies between the symbols and regression lines shown in these figures.

The values of c_1 and c_2 corresponding to the intercepts and slopes, respectively, of the regression lines in Fig. 2 and those of σ and ξ represented by the intercepts and slopes, respectively, of the regression lines in Fig. 3 are given in Table 5. Corresponding values of C_P^o and V^o calculated from Eqs. (A2) and (A3) using values of these parameters and those of ω taken from Table A1 in Appendix A are also given in Table 5. The computed values of c_1, c_2, σ, and ξ for Gly_3, Gly_4, and Gly_5 are plotted as symbols in Fig. 4. The respective slopes of the regression lines shown in this figure correspond to the values of c_1, c_2, σ, and ξ for −Gly− given in Table 7, which are consistent with the regression lines in Figs. 2 and 3.

Values of C_P^o, c_1, c_2, V^o, σ, and ξ for PB were generated by regression analysis of the experimental values of C_P^o and V^o for the four denatured proteins given in Table 8 using computed values of C_P^o and V^o for the R groups and −Gly− units in these proteins taken from Tables 9 and 10.

The regression calculations were carried out by first computing values of $C^o_{P,PB}$ and V^o_{PB} at the temperatures shown in Tables 9 and 10 from the additivity algorithms represented by

$$C^o_{P,PB} = \left(C^o_{P,P} - C^o_{P,AAB} - n_{Gly} C^o_{P,-Gly-} - \sum_{i=1}^{\hat{i}-n_{Gly}} n_i C^o_{P,i} \right) / n_{PB} \quad (11)$$

and

$$V^o_{PB} = \left(V^o_P - V^o_{AAB} - n_{Gly} V_{-Gly-} - \sum_{i=1}^{\hat{i}-n_{Gly}} n_i V_i^o \right) / n_{PB} \quad (12)$$

Table 9
Standard molal heat capacities (C_P^o, cal mol^{-1} K^{-1}) of AAB, −Gly−, and R groups other than R_{Gly} at P_{SAT} and various temperatures computed from Eq. (A2) using parameters given in Table 6

Group	Temperature (°C)				
	25	50	75	100	125
AAB	10.82	18.06	21.34	23.15	24.32
−Gly−	22.43	21.24	20.80	20.66	20.70
R_{Ala}	22.98	18.34	16.53	15.86	15.84
R_{Val}	61.38	52.89	49.52	48.17	47.94
R_{Leu}	84.28	73.73	69.53	67.85	67.56
R_{Ile}	80.78	70.22	66.00	64.30	63.97
R_{Ser}	17.28	19.56	20.56	21.06	21.34
R_{Thr}	39.18	39.28	39.25	39.15	38.98
R_{Asp}	19.58	25.14	27.75	29.28	30.40
R_{Asp-}	−10.02	−2.62	−1.09	−2.22	−5.31
R_{Glu}	31.48	35.00	36.81	38.06	39.18
R_{Glu-}	−4.52	2.30	3.59	2.33	−0.81
R_{Asn}	19.08	24.75	27.45	29.10	30.36
R_{Gln}	33.88	37.48	39.36	40.67	41.85
R_{His}	46.78	48.14	49.09	49.99	51.03
R_{His+}	40.85	44.69	46.32	47.08	47.43
R_{Lys}	52.98	52.12	51.93	52.05	52.39
R_{Lys+}	47.05	48.58	49.13	49.29	49.22
R_{Arg}	55.88	57.59	58.98	60.42	62.17
R_{Arg+}	49.95	53.96	55.67	56.52	56.93
R_{Cys}	34.08	34.27	34.32	34.31	34.25
R_{CysX}[a]	26.86	35.31	39.07	41.06	42.25
R_{Met}	59.18	53.22	50.85	49.91	49.75
R_{Phe}	80.98	78.61	78.16	78.61	79.75
R_{Tyr}	60.68	59.53	59.24	59.34	59.71
R_{Trp}	89.58	102.97	109.15	112.69	115.12
R_{Pro}	30.38	26.94	25.73	25.44	25.74

[a] This notation refers to the carboxyamido-methylated cysteine side chain ($-CH_2SCH_2CONH_2$). The values of C_P^o for this R group were calculated from Eq. (A2) using equations of state parameters computed from $\Xi_{R_{CysX}} = \Xi_{R_{Cys}} + \Xi_{R_{Asn}} + \Xi_{-CH_2-} - \Xi_{-CH_3}$ where $\Xi_{R_{CysX}}$, $\Xi_{R_{Cys}}$, $\Xi_{R_{Asn}}$, Ξ_{-CH_2-}, and Ξ_{-CH_3} denote c_1, c_2, or ω for the subscripted species, which are tabulated in Amend and Helgeson [20,45].

where $C^o_{P,i}$ and V_i^o represent C_P^o and V^o, respectively, of the ith side chain, and n_{PB} denotes the number of PB groups in the protein, excluding those accounted for by −Gly− (n_{Gly}) (see above), which can be calculated from

$$n_{PB} = n - n_{Gly} - 1 \quad (13)$$

The values of $C^o_{P,i}$, $C^o_{P,-Gly-}$, $C^o_{P,AAB}$, V_i^o, V^o_{-Gly-}, and V^o_{AAB} in Tables 9 and 10 required to compute

Table 10
Standard molal volumes (V^o, $cm^3\ mol^{-1}$) of AAB, −Gly−, and R groups other than R_{Gly} at P_{SAT} and various temperatures computed from Eq. (A3) using parameters from Table 6

Group	Temperature (°C)									
	15	25	35	40	45	50	55	60	65	75
AAB	31.25	32.27	33.03	33.34	33.63	33.88	34.11	34.32	34.51	34.85
−Gly−	39.02	38.73	38.51	38.43	38.35	38.28	38.22	38.17	38.13	38.05
R_{Ala}	28.57	28.23	28.00	27.90	27.82	27.76	27.70	27.65	27.61	27.56
R_{Val}	58.60	58.65	58.70	58.73	58.77	58.80	58.84	58.88	58.92	59.02
R_{Leu}	74.90	75.34	75.69	75.85	75.99	76.13	76.26	76.38	76.50	76.74
R_{Ile}	72.77	73.20	73.56	73.71	73.85	73.99	74.12	74.24	74.36	74.59
R_{Ser}	28.39	28.31	28.24	28.21	28.19	28.17	28.15	28.13	28.11	28.08
R_{Thr}	44.63	44.60	44.56	44.55	44.54	44.52	44.51	44.50	44.48	44.46
R_{Asp}	41.34	41.62	41.84	41.93	42.01	42.09	42.16	42.23	42.29	42.41
R_{Asp-}	29.62	28.70	27.92	27.56	27.22	26.89	26.57	26.24	25.92	25.28
R_{Glu}	56.74	57.10	57.38	57.50	57.61	57.71	57.81	57.90	57.99	58.15
R_{Glu-}	43.49	42.82	42.22	41.94	41.66	41.39	41.12	40.85	40.57	40.01
R_{Asn}	44.68	44.94	45.14	45.22	45.30	45.37	45.44	45.51	45.57	45.69
R_{Gln}	61.21	61.54	61.80	61.91	62.02	62.11	62.20	62.29	62.38	62.54
R_{His}	66.05	66.53	66.90	67.06	67.21	67.35	67.48	67.60	67.72	67.94
R_{His+}	66.62	66.02	65.58	65.39	65.22	65.07	64.94	64.81	64.70	64.49
R_{Lys}	75.77	76.45	76.96	77.18	77.37	77.55	77.71	77.86	78.00	78.25
R_{Lys+}	76.42	76.34	76.26	76.23	76.20	76.17	76.14	76.12	76.09	76.04
R_{Arg}	89.35	91.44	93.03	93.70	94.31	94.86	95.36	95.83	96.26	97.05
R_{Arg+}	89.43	90.53	91.34	91.68	91.98	92.24	92.48	92.70	92.89	93.23
R_{Cys}	41.25	41.18	41.12	41.09	41.07	41.05	41.03	41.01	41.00	40.96
R_{CysX} [a]	76.88	76.61	76.41	76.33	76.25	76.19	76.13	76.07	76.02	75.94
R_{Met}	72.66	73.09	73.43	73.58	73.71	73.83	73.95	74.06	74.16	74.36
R_{Phe}	89.20	89.78	90.24	90.44	90.63	90.81	90.98	91.14	91.30	91.60
R_{Tyr}	90.25	90.91	91.42	91.63	91.82	92.00	92.16	92.31	92.45	92.70
R_{Trp}	110.47	111.54	112.35	112.69	113.00	113.28	113.53	113.76	113.98	114.36
R_{Pro}	50.20	50.24	50.28	50.30	50.33	50.35	50.38	50.42	50.45	50.52

[a] This notation refers to the carboxyamido-methylated cysteine side chain ($-CH_2SCH_2CONH_2$). The values of V^o for this group were calculated from Eq. (A3) using equation of state parameters computed from $\Xi_{R_{CysX}} = \Xi_{R_{Cys}} + \Xi_{R_{Asn}} + \Xi_{-CH_2-} - \Xi_{-CH_3}$ where $\Xi_{R_{CysX}}$, $\Xi_{R_{Cys}}$, $\Xi_{R_{Asn}}$, Ξ_{-CH_2-}, and Ξ_{-CH_3} denote σ, ξ, or ω for the subscripted species, which are tabulated in Amend and Helgeson [20,45].

corresponding values of $C^o_{P,PB}$ and V^o_{PB} from Eqs. (11) and (12) for the temperatures shown in Tables 9 and 10 were calculated from Eqs. (A2) and (A3) using equations of state parameters given in Table 7. The computed values of $C^o_{P,PB}$ and V^o_{PB} are given in Table 11.

Because the experiments responsible for the values of C^o_P and V^o given in Table 8 were carried out using acid solutions [15,53], the side chains of the His, Lys, and Arg residues were taken to be fully protonated in the calculations described above. In contrast, those of the Asp and Glu residues were regarded in a first approximation as half ionized and half neutral. These approximations are based on consideration of the dissociation constants of His, Lys, Arg, Asp, and Glu at elevated temperatures [20]. The Cys side chains used in the calculations to determine values of C^o_P and V^o of PB were taken to be carboxyamidomethylated (see references in Tables 9 and 10) in order to be consistent with the corresponding experimental values for cytochrome *c*, lysozyme, myoglobin, and ribonuclease. The values of n_{Gly}, n_i, n_{PB}, and n for these proteins were taken from Dayhoff et al. [59].

The values of $C^o_{P,PB}$ and V^o_{PB} given in Table 11 were regressed with Eqs. (A9) and (A8), respectively, in the same manner as those of C^o_P and V^o

Table 11
Standard molal heat capacities (C_P^o, cal mol^{-1} K^{-1}) and volumes (V^o, cm^3 mol^{-1}) of PB calculated from the experimental values of C_P^o and V^o for cytochrome *c*, lysozyme, myoglobin, and ribonuclease given in Table 8 and those of the constituent groups in Tables 9 and 10 using Eqs. (11) and (12) in conjunction with amino acid sequences reported by Dayhoff et al. [59]

T (°C)	PB (cytochrome *c*)		PB (lysozyme)		PB (myoglobin)		PB (ribonuclease)	
	C_P^o	V^o	C_P^o	V^o	C_P^o	V^o	C_P^o	V^o
5	6.87		8.98	24.31	2.70	24.34	9.25	25.07
10		21.34						
15				24.10		25.43		24.52
25	14.30	22.78	12.42	26.60	11.08	25.96	14.25	25.31
35				28.19		26.57		26.11
40		23.92				27.12		
45				28.94				27.31
50	18.48		15.82		15.96		17.19	
55				28.85				27.87
60		24.30						
65				30.67				31.98
75	20.75	25.79	18.68		20.44		19.91	
100	21.05		18.97		20.84		20.88	
125	21.54		18.85		21.21		21.26	

of Gly, Gly_3, Gly_4, and Gly_5 discussed above. Values of $\Delta C_{P,s}^o$ and ΔV_s^o for PB were computed from Eqs. (A10) and (A11) using values of ω given in Table A1 and those of the Born functions tabulated in Shock et al. [26]. It can be seen in Fig. 5 that the values of $\Delta C_{P,n}^o$ for PB derived from the calorimetric data given in Table 8 for cytochrome *c*, lysozyme, myoglobin, and ribonuclease differ from one another by an average of approximately ± 1 cal (mol of backbone)$^{-1}$ K^{-1}. The largest discrepancies in this figure among the experimental data represented by the symbols occur at the highest value of $1/(T-\Theta)^2$ which corresponds to the lowest temperature (25°C). In contrast, at values of $1/(T-\Theta)^2$ corresponding to 75, 100, and 125°C, the values of $\Delta C_{P,n}^o$ for PB computed from the experimental data for three of the four proteins are almost coincident, differing by approximately ± 0.3 cal (mol of backbone)$^{-1}$ K^{-1}.

Because this remarkable coincidence is almost certainly not fortuitous, it strongly supports the group additivity equation of state approach adopted in the present study to compute C_P^o for unfolded proteins as a function of temperature. Even the difference between the values of $\Delta C_{P,n}^o$ for PB in Fig. 5 derived from the experimental data for lysozyme in Table 8 and those represented by the regression line introduce a difference of only ~300 cal mol^{-1} or less between the experimental heat capacities of lysozyme in Fig. 6 and the computed values represented by the curve shown in this figure, which is consistent with the regression line in Fig. 5 (see below). It can be

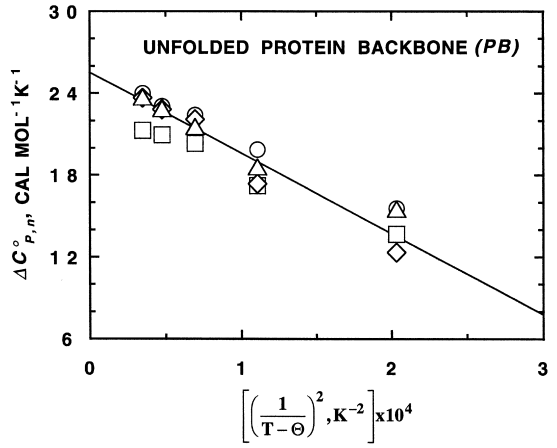

Fig. 5. $\Delta C_{P,n}^o$ for PB as a function of $1/(T-\Theta)^2$ at P_{SAT}. The symbols represent values of $\Delta C_{P,n}^o$ computed from Eqs. (11), (A2) and (A10) using values of ω for the constituent groups taken from Table A1 and those of C_P^o of unfolded cytochrome *c* (○), lysozyme (□), myoglobin (◇), and ribonuclease (△) given in Table 8. The regression lines correspond to fits of Eq. (A9) to the data.

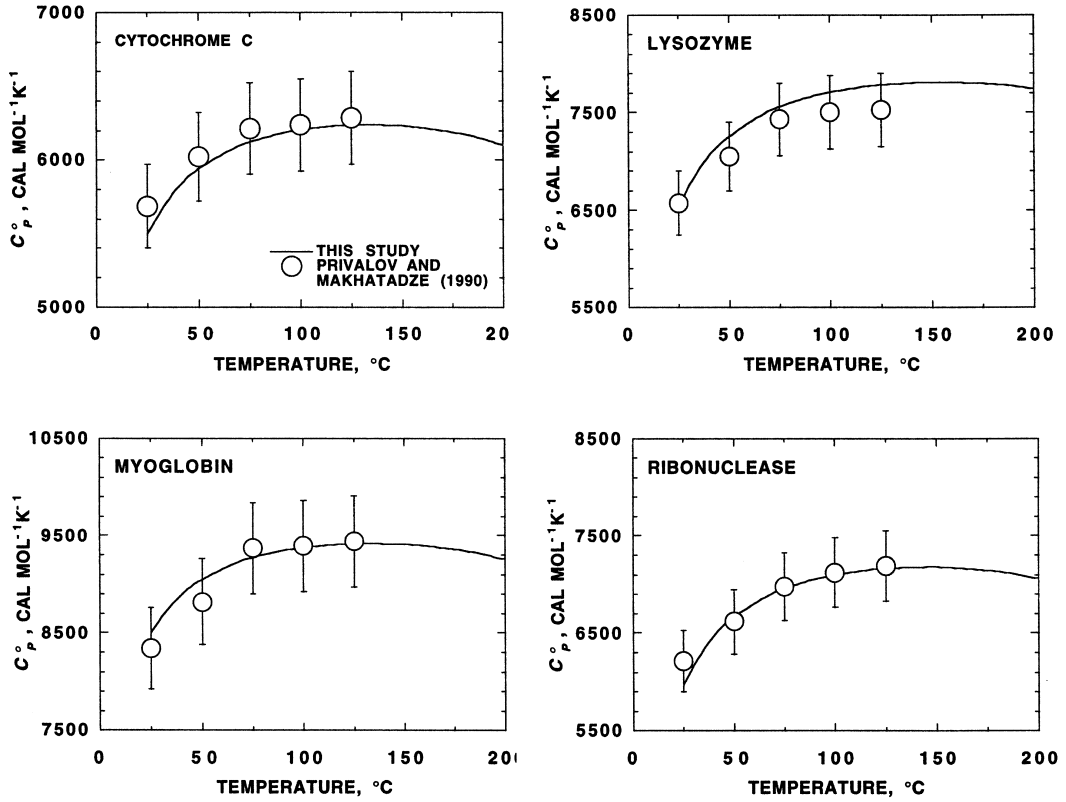

Fig. 6. Comparison of experimental and calculated values of C_P° as a function of temperature at P_{SAT} for cytochrome *c*, lysozyme, myoglobin, and ribonuclease. The symbols depict experimental data taken from Table 8, but the curves were generated from Eq. (A2) using the group additivity algorithm represented by Eq. (6) and parameters taken from Table 7.

deduced from Fig. 6 that the regression uncertainties represented by the differences between the symbols and the regression line in Fig. 5 introduce corresponding uncertainties in C_P° for the four proteins computed in the present study that are within the 5% experimental uncertainty assigned by Privalov and Makhatadze [15]. In fact, the curves in Fig. 6 (all of which are consistent with the regression lines in Fig. 5) are within 1% of the bulk of the experimental values and within 3% of the rest. It can be seen in Fig. 7 that the values of ΔV_n° for PB derived from the experimental values of V° given in Table 8 exhibit considerable scatter, but for the most part, the values of ΔV_n° for PB computed from the data for any of the four proteins vary smoothly with temperature. The scatter of the symbols in Fig. 7 can be attributed to the relatively large uncertainties associated with calculations of apparent molal volumes from experimental density measurements.

Makhatadze et al. [53] report experimental uncertainties in the values of V° given in Table 8 to be on the order of $\pm 3\%$. Uncertainties associated with values of V° for aqueous species derived from density measurements may arise from several sources, including less than perfect experimental reproducibility, magnification of uncertainties resulting from low concentrations of the solute, and errors in extrapolating to infinite dilution apparent molal volumes (ϕ_V) computed from the densities. Makhatadze et al. [53] recognized explicitly the first of these sources of uncertainty, but not the latter two, which are discussed below.

Values of ϕ_V of the four unfolded proteins listed in Table 8 were computed from density

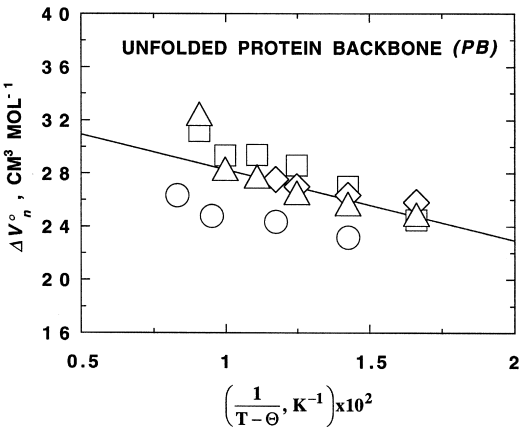

Fig. 7. ΔV_n^o for PB as a function of $1/(T-\Theta)$. The symbols represent values of ΔV_n^o computed from Eqs. (12), (A3) and (A11) using values of ω for the constituent groups taken from Table A1 and those of V^o of unfolded cytochrome c (○), lysozyme (□), myoglobin (◇), and ribonuclease (△) given in Table 8. However, the regression lines correspond to fits of Eq. (A8) to the data.

measurements of the solution using the relation

$$\phi_V = \frac{M}{W}\left(1 - \frac{\rho_{soln}}{\rho_w} + \frac{W}{\rho_{soln}}\right) \quad (14)$$

where M represents the molecular weight of the solute, W stands for the concentration of the solute in grams per liter of solution, and ρ_{soln} and ρ_w denote the densities of the solution and pure water, respectively [53].[4] Uncertainties in ϕ_V arising from those in ρ_{soln} can be assessed from the finite difference derivative (Δ) of Eq. (14), which can be written as

$$\Delta\phi_V = -\left(\frac{M}{\rho_w W} + \frac{M}{\rho_{soln}^2}\right)\Delta\rho_{soln} \quad (15)$$

It can be deduced from Eq. (15) that uncertainties in density measurements are magnified considerably in computing values of ϕ_V for low concentrations of the solute [60]. For example, Makhatadze et al. [53] report densities to three decimals, which is consistent with an uncertainty in ρ_{soln} of 5×10^{-4} g cm^{-3}. In the case of lysozyme (for which $M = 14\,300$ g mol^{-1}), if $W = 0.001$, $\Delta\phi_V$ computed from Eq. (15) is ~ 7150 cm^3 mol^{-1}. In contrast, if $W = 0.01$, ρ_{soln} reduces to ~ 720 cm^3 mol^{-1}. It follows that ρ_{soln} for dilute solutions must be accurate to $\sim 10^{-5}$ g cm^{-3} or better to yield reliable values of ϕ_V.

The experimental values of V^o for unfolded cytochrome c, lysozyme, myoglobin, and ribonuclease shown in Table 8 were generated by linear extrapolation of values of ϕ_V to infinite dilution from concentrations $> \sim 3$ g l^{-1} [53], which contravenes electrostatic theory [60]. However, these extrapolation errors are far smaller than the uncertainties associated with computing values of ϕ_V from ρ_{soln} discussed above. The discrepancies between the regression line and the symbols representing experimental values of V^o for the proteins in Fig. 6 introduce corresponding uncertainties in computed values of V^o for the four proteins. However, it can be seen in Fig. 8 that the values of V^o for the proteins represented by the curves generated in the present study are for the most part within the 3% experimental uncertainty reported by Makhatadze et al. [53].

The intercepts and slopes of the regression lines shown in Figs. 5 and 7 represent weighted estimates. The respective values of c_1 and c_2 for Eq. (A8) that correspond to the intercept and slope of the lines in Fig. 5 are shown in Table 7 together with those of σ and ξ for Eq. (A9), which are consistent with the intercept and slope of the regression line in Fig. 7. The values of $C_{P,PB}^o$ and V_{PB}^o at 25°C in Table 6 were computed from Eqs. (A2) and (A3), respectively, using parameters given in Table 7. These values are in close agreement with values obtained using several alternate approaches. For example, V_{PB}^o at 25°C estimated in accord with Eqs. (7) and (8) differs from that adopted in this study by less than 2 cm^3 mol^{-1}. Similarly, $C_{P,PB}^o$ at 25°C estimated with this alternate approach differs by only slightly more than 2 cal mol^{-1} K^{-1}.

The values of C_P^o, V^o, c_1, c_2, σ, and ξ for PB and −Gly− in Tables 6 and 7 can be used to calculate values of the corresponding properties

[4] A typographical error in the corresponding equation published by Makhatadze et al. [53] should be corrected. This can be done by replacing ρ of the solvent in the denominator of the third term within the parentheses in their equation with ρ of the solution.

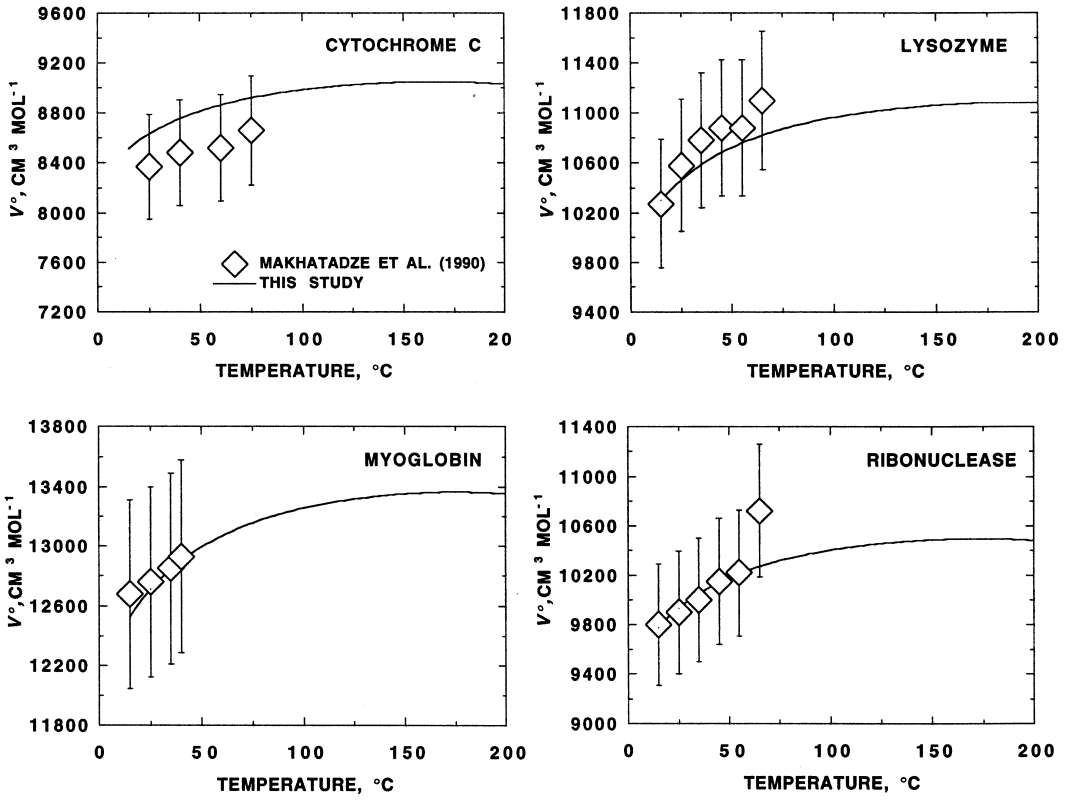

Fig. 8. Comparison of experimental and computed values of V^{o} as a function of temperature at P_{SAT} for cytochrome *c*, lysozyme, myoglobin, and ribonuclease. The symbols depict experimental data taken from Table 8, but the curves were generated from Eq. (A3) using the group additivity algorithm represented by Eq. (6) and parameters taken from Table 7.

and parameters for R_{Gly} from an appropriate statement of Eq. (8) in which $\Xi_{R_{Gly}}$, Ξ_{-Gly-} and Ξ_{PB} represent C_P^o, V^o, c_1, c_2, σ, or ξ for the subscripted species. The results of these calculations for R_{Gly} are given in Tables 6 and 7. Owing to the dearth of experimental compressibility data in the literature for glycyl peptides and denatured proteins, values of the equations of state parameters represented by a_2 and a_4 for −Gly− and PB were calculated from group additivity relations. These values were then combined with those of σ and ξ given in Table 7 to obtain values of a_1 and a_3 in the manner described below.

4.3. Calculation of equations of state parameters a_1, a_2, a_3, and a_4

The values of a_2 and a_4 for −Gly− and PB in Table 7 were calculated from those of their constituent groups tabulated in Amend and Helgeson [20] using Eq. (8) and the relation

$$\Xi_{PB} = \Xi_{AAB} + 2\Xi_{-CH_2-} - \Xi_{-CH_2OH} - \Xi_{-CH_3} \quad (16)$$

respectively, where Ξ_{-Gly-}, $\Xi_{R_{Gly}}$, Ξ_{AAB}, Ξ_{-CH2-}, and Ξ_{-CH3} denote a_2 or a_4 of the subscripted groups. The values of a_1 and a_3 for PB and −Gly− in Table 7 were then computed from Eq. (A3) using the values of a_2, a_4, σ, and ξ for these groups given in the table. The equation of state parameters and standard molal properties at 25°C and 1 bar given in Tables 7 and 6, respectively, permit calculation of these properties of unfolded proteins and their constituent groups as a function of temperature and pressure from the equations summarized in Appendix A.

5. Calculation of the standard molal thermodynamic properties of thermophilic proteins at elevated temperatures

The complete amino acid sequences are now known for a number of thermophilic proteins. To illustrate the application and utility of the group additivity equations of state derived above, these sequences have been used together with Eq. (6) and the equations of state parameters and standard molal properties at 25°C and 1 bar of the groups given in Tables 6 and 7 to compute corresponding parameters and properties for 11 unfolded thermophilic proteins. These parameters and properties (summarized in Table 12) were then used in conjunction with the equations of state given in Appendix A to calculate the standard molal properties of these unfolded proteins as a function of temperature and pressure. Values of ΔG^o, ΔH^o, S^o, C_P^o and V^o computed in this manner at P_{SAT} and temperatures between 25 and 250°C are given in Tables 13–17, respectively. The unfolded thermophilic proteins considered are rubredoxin and ferredoxin from *Pyrococcus furiosus* (*Pf*), ferredoxin from *Thermococcus litoralis* (*Tl*), elongation factor Tu (EF-Tu) from *Thermotoga maritima* (*Tm*) and *Thermus thermophilus* (*Tt*), Ompα from *Tm*, glutamine synthase from *Tm* and *Pyrococcus woesei* (*Pw*), glyceraldehyde-3-phosphate dehydrogenase (GAPDH) from *Tm*, and β-galactosidase and Indole-3-glycerol-phosphate synthase (IndGroP synthase) from *Sulfolobus solfataricus* (*Ss*). The computed values of ΔG^o, ΔH^o, S^o, C_P^o and V^o of

Table 12
Standard molal thermodynamic properties and equations of state parameters for 11 unfolded thermophilic proteins computed from Eq. (6) using properties and parameters for groups taken from Tables 6 and 7

Protein	ΔG_f^o (kcal mol^{-1})	ΔH_f^o (kcal mol^{-1})	S^o (cal mol^{-1} K^{-1})	C_P^o (cal mol^{-1} K^{-1})	V^o (cm^3 mol^{-1})	$a_1 \times 10$ (cal mol^{-1} bar^{-1})
Rubredoxin (*Pf*)	−2308	−4713	1874	2363	4059	1084.36
Ferredoxin (*Pf*)	−3219	−6218	2315	2905	4932	1330.80
Ferredoxin (*Tl*)	−2768	−5320	2000	2520	4187	1133.29
EF-Tu (*Tm*)	−16043	−35102	15241	20363	32132	8491.19
EF-Tu (*Tt*)	−15856	−35002	15420	20638	32328	8545.99
Ompα (*Tm*)	−17622	−36313	14508	19213	30702	8185.44
Gln synthase (*Tm*)	−16891	−37996	16934	23153	36188	9536.93
Gln synthase (*Pw*)	−16828	−37753	16793	22845	36101	9507.73
GAPDH (*Tm*)	−13020	−28821	12443	17470	26488	6991.28
β-Galactosidase (*Ss*)	−17694	−40807	19348	25585	40487	10653.03
IndGroP synthase (*Ss*)	−10292	−22750	9870	13484	20799	5495.01

Protein	$a_2 \times 10^{-2}$ (cal mol^{-1})	a_3 (cal K mol^{-1} bar^{-1})	$a_4 \times 10^{-4}$ (cal K mol^{-1})	c_1 (cal mol^{-1} K^{-1})	$c_2 \times 10^{-4}$ (cal K mol^{-1})	$\omega \times 10^{-5}$ (cal mol^{-1})
Rubredoxin (*Pf*)	−59.32	262.23	−212.94	3163.23	−312.61	18.11
Ferredoxin (*Pf*)	−85.92	250.05	−255.34	3827.88	−355.24	22.14
Ferredoxin (*Tl*)	−94.56	242.22	−215.75	3242.93	−271.26	18.92
EF-Tu (*Tm*)	−230.05	2052.71	−1753.21	24287.34	−1560.48	82.84
EF-Tu (*Tt*)	−243.45	2069.69	−1767.66	24593.12	−1594.11	78.66
Ompα (*Tm*)	−326.09	1683.97	−1634.82	23892.46	−1854.78	100.10
Gln synthase (*Tm*)	−133.10	2269.24	−2011.99	27818.54	−1906.02	87.07
Gln synthase (*Pw*)	−175.24	2351.82	−1985.77	27618.28	−1948.08	89.20
GAPDH (*Tm*)	−149.86	1725.39	−1479.43	20518.65	−1242.74	57.51
β-Galactosidase (*Ss*)	−99.57	2390.96	−2238.62	31387.20	−2503.12	78.73
IndGroP synthase (*Ss*)	−51.18	1172.71	−1167.44	16042.28	−1044.58	47.85

Table 13
Standard molal Gibbs free energies (ΔG^o, kcal mol^{-1}) of 11 unfolded thermophilic proteins as a function of temperature at P_{SAT} computed from Eqs. (6) and (A14) using equations of state parameters and standard molal properties at 25°C and 1 bar taken from Tables 6 and 7, together with the amino acid sequences reported in the references listed below

Protein	Temperature (°C)						
	25	50	75	100	150	200	250
Rubredoxin (*Pf*)[a]	−2308	−2358	−2412	−2472	−2603	−2750	−2909
Ferredoxin (*Pf*)[b]	−3219	−3280	−3347	−3420	−3582	−3762	−3956
Ferredoxin (*Tl*)[b]	−2768	−2820	−2878	−2941	−3081	−3236	−3403
EF-Tu (*Tm*)[c]	−16043	−16445	−16890	−17374	−18449	−19649	−20947
EF-Tu (*Tt*)[c]	−15856	−16263	−16712	−17202	−18291	−19507	−20824
Ompα (*Tm*)[d]	−17622	−18005	−18428	−18889	−19916	−21062	−22301
Gln synthase (*Tm*)[e]	−16891	−17339	−17835	−18376	−19581	−20930	−22394
Gln synthase (*Pw*)[e]	−16828	−17272	−17763	−18299	−19494	−20831	−22280
GAPDH (*Tm*)[f]	−13020	−13349	−13715	−14114	−15004	−16001	−17084
β-Galactosidase (*Ss*)[g]	−17694	−18204	−18769	−19385	−20756	−22291	−23959
IndGroP synthase (*Ss*)[h]	−10292	−10553	−10842	−11157	−11858	−12643	−13494

[a] Blake et al. [80].
[b] Busse et al. [81].
[c] Bachleitner et al. [82].
[d] Engel et al. [83].
[e] Sanangelantoni et al. [84].
[f] Schultes et al. [85].
[g] Cubellis et al. [86].
[h] Andreotti et al. [87].

four of these proteins [rubredoxin (*Pf*), IndGroP synthase (*Ss*), GAPDH (*Tm*), and β-galactosidase (*Ss*)] are plotted as examples in Figs. 9–13. These proteins were chosen because they are representative of the composition, size, and thermodynamic behavior as a function of temperature and pressure, of the group of thermophilic enzymes considered in this study.

It can be seen in Figs. 9–11 that the curves representing ΔG^o, ΔH^o, and S^o as a function of temperature are monotonic, decreasing with increasing temperature for ΔG^o, but increasing over

Table 14
Standard molal enthalpies (ΔH^o, kcal mol^{-1}) of 11 unfolded thermophilic proteins as a function of temperature at P_{SAT} computed from Eqs. (6) and (A13) using equations of state parameters and standard molal properties at 25°C and 1 bar taken from Tables 6 and 7, together with the amino acid sequences reported in the references given in the footnotes for Table 13

Protein	Temperature (°C)						
	25	50	75	100	150	200	250
Rubredoxin (*Pf*)	−4713	−4650	−4583	−4514	−4377	−4248	−4144
Ferredoxin (*Pf*)	−6218	−6141	−6059	−5975	−5809	−5653	−5528
Ferredoxin (*Tl*)	−5320	−5254	−5184	−5113	−4972	−4840	−4734
EF-Tu (*Tm*)	−35102	−34573	−34022	−33464	−32348	−31264	−30291
EF-Tu (*Tt*)	−35002	−34466	−33907	−33341	−32206	−31101	−30100
Ompα (*Tm*)	−36313	−35809	−35279	−34741	−33665	−32630	−31731
Gln synthase (*Tm*)	−37996	−37392	−36762	−36122	−34839	−33588	−32451
Gln synthase (*Pw*)	−37753	−37157	−36533	−35900	−34629	−33391	−32271
GAPDH (*Tm*)	−28821	−28368	−27898	−27421	−26466	−25531	−24671
β-Galactosidase (*Ss*)	−40807	−40135	−39426	−38703	−37244	−35807	−34467
IndGroP synthase (*Ss*)	−22750	−22399	−22034	−21663	−20920	−20195	−19531

Table 15
Standard molal entropies (S^o, cal mol^{-1} K^{-1}) of 11 unfolded thermophilic proteins as a function of temperature at P_{SAT} computed from Eqs. (6) and (A12) using equations of state parameters and standard molal properties at 25°C and 1 bar taken from Tables 6 and 7, together with the amino acid sequences reported in the references given in the footnotes for Table 13

Protein	Temperature (°C)						
	25	50	75	100	150	200	250
Rubredoxin (*Pf*)	1874	2077	2278	2468	2812	3098	3303
Ferredoxin (*Pf*)	2315	2563	2807	3039	3455	3801	4048
Ferredoxin (*Tl*)	2000	2214	2422	2619	2973	3266	3475
EF-Tu (*Tm*)	15 241	16 944	18 585	20 133	22 933	25 335	27 257
EF-Tu (*Tt*)	15 420	17 147	18 822	20 383	23 230	25 679	27 656
Ompα (*Tm*)	14 508	16 130	17 708	19 202	21 901	24 195	25 970
Gln synthase (*Tm*)	16 934	18 876	20 754	22 529	25 749	28 523	30 769
Gln synthase (*Pw*)	16 793	18 713	20 571	22 329	25 517	28 261	30 473
GAPDH (*Tm*)	12 443	13 901	15 302	16 623	19 020	21 093	22 792
β-Galactosidase (*Ss*)	19 348	21 511	23 623	25 629	29 289	32 475	35 121
IndGroP synthase (*Ss*)	9870	10 999	12 087	13 115	14 979	16 588	17 898

the same temperature range for ΔH^o and S^o. The solid curves in these figures represent values of the standard molal properties of unfolded proteins at pHs for which the His, Lys, Arg, Asp, and Glu residues are all ionized, but the remaining residues are neutral. In contrast, the dashed curves denote corresponding values of ΔG^o, ΔH^o, and S^o for the proteins at pHs for which only the His, Lys, and Arg residues are ionized. Note that the dashed and corresponding solid curves are demonstrably different, not only in position, but also in configuration. Differences of this kind are even more pronounced in the sets of solid and dashed curves shown in Figs. 12 and 13, which correspond to C_P^o and V^o analogs of those for ΔG^o, ΔH^o, and S^o in Figs. 9–11. It can be seen that the solid curves representing C_P^o and V^o in these figures maximize as a function of temperature in all four cases between ~100 and 150°C. In contrast, all of the dashed curves in Figs. 12 and 13 exhibit a reverse sigmoid configuration and increase with increasing temperature. The difference in the configurations of the solid and dashed curves at high temperatures results from the difference in sign of ω for the R groups of the Asp and Glu residues relative to those of

Table 16
Standard molal heat capacities (C_P^o, cal mol^{-1} K^{-1}) of 11 unfolded thermophilic proteins as a function of temperature at P_{SAT} computed from Eqs. (6) and (A2) using equations of state parameters taken from Table 7, together with the amino acid sequences reported in the references given in the footnotes for Table 13

Protein	Temperature (°C)						
	25	50	75	100	150	200	250
Rubredoxin (*Pf*)	2363	2635	2733	2758	2679	2376	1413
Ferredoxin (*Pf*)	2905	3212	3320	3345	3243	2870	1692
Ferredoxin (*Tl*)	2520	2753	2832	2846	2752	2429	1421
EF-Tu (*Tm*)	20 363	21 728	22 228	22 371	22 039	20 665	16 271
EF-Tu (*Tt*)	20 638	22 039	22 560	22 720	22 429	21 135	16 969
Ompα (*Tm*)	19 213	20 834	21 425	21 591	21 184	19 520	14 208
Gln synthase (*Tm*)	23 153	24 835	25 470	25 678	25 386	23 968	19 363
Gln synthase (*Pw*)	22 845	24 567	25 215	25 428	25 127	23 674	18 956
GAPDH (*Tm*)	17 470	18 566	18 978	19 113	18 916	17 978	14 936
β-Galactosidase (*Ss*)	25 585	27 829	28 723	29 082	28 984	27 777	23 654
IndGroP synthase (*Ss*)	13 484	14 406	14 753	14 867	14 706	13 927	11 396

Table 17
Standard molal volumes (V^o, $cm^3\ mol^{-1}$) of 11 unfolded thermophilic proteins as a function of temperature at P_{SAT} computed from Eqs. (6) and (A3) using equations of state parameters taken from Table 7, together with the amino acid sequences reported in the references given in the footnotes for Table 13

Protein	Temperature (°C)						
	25	50	75	100	150	200	250
Rubredoxin (*Pf*)	4059	4139	4179	4195	4178	4077	3772
Ferredoxin (*Pf*)	4932	5038	5090	5114	5097	4976	4605
Ferredoxin (*Tl*)	4187	4271	4313	4330	4313	4208	3890
EF-Tu (*Tm*)	32 132	32 833	33 208	33 412	33 494	33 128	31 797
EF-Tu (*Tt*)	32 328	33 037	33 418	33 628	33 727	33 392	32 137
Ompα (*Tm*)	30 702	31 382	31 738	31 921	31 941	31 451	29 812
Gln synthase (*Tm*)	35 434	36 255	36 699	36 946	37 072	36 710	35 326
Gln synthase (*Pw*)	36 101	36 892	37 318	37 551	37 657	37 273	35 846
GAPDH (*Tm*)	26 488	27 085	27 411	27 595	27 704	27 482	26 580
β-Galactosidase (*Ss*)	40 487	41 429	41 947	42 246	42 454	42 184	40 971
IndGroP synthase (*Ss*)	20 799	21 299	21 571	21 726	21 818	21 633	20 881

Asp^- and Glu^-. Large positive values of ω are characteristic of monovalent anions such as R_{Asp-} and R_{Glu-}. In contrast, relatively small negative values of ω are representative of neutral polar species such as R_{Asp} and R_{Glu}. The standard molal thermodynamic properties of unfolded proteins with neutral His, Lys, and Arg residues differ only slightly from the corresponding properties of the proteins with these side chains ionized because the relative differences in the values of ω for cationic and neutral R groups are significantly less than the corresponding differences in ω for anionic and neutral groups.

It can be deduced from the curves shown in Figs. 12 and 13, and to a lesser degree from those depicted in Figs. 9–11, that solution pH and the relative abundances of neutral and charged Asp and Glu residues play a major role in determining the thermodynamic behavior of unfolded proteins. Furthermore, it can be shown that by taking account of the equations of state parameters for the R groups in Table 7, the relative abundances of residues other than Asp and Glu may affect appreciably the thermodynamic behavior of unfolded proteins as a function of temperature and/or pressure. For example, proteins with relatively large numbers of Val, Leu, Ile, Lys, Met, Phe, and Tyr residues exhibit substantially lower first derivatives of C_P^o with respect to temperature at low temperatures. For proteins with even larger numbers of these residues, the sign of this derivative may be negative. It follows that curves representing C_P^o of unfolded proteins as a function of temperature, and as a consequence, curves corresponding to the other standard molal properties may exhibit any of four possible configurations: they may maximize, minimize, or exhibit sigmoid or reverse sigmoid configurations, depending on the relative abundances of different amino acid residues and their ionization states in the unfolded protein [20,30,45]. The same observations apply to curves representing the standard molal thermodynamic properties of native proteins, as well as those for the unfolding process [4].

6. Computational uncertainties

Representative uncertainties associated with the values of ΔG_f^o, ΔH_f^o, S^o, C_P^o, and V^o at 25°C and 1 bar for the unfolded proteins computed in the present study can be assessed from estimates of the uncertainties in these properties for the amino acid residues. The latter uncertainties are essentially equivalent to those associated with the corresponding properties of the amino acids [20], which are of the order of ± 0.02 kcal mol^{-1}, ± 0.20 kcal mol^{-1}, ± 0.74 cal mol^{-1} K^{-1}, ± 1–2 cal mol^{-1} K^{-1}, and ± 0.5–1 cm^3 mol^{-1} for ΔG_f^o,

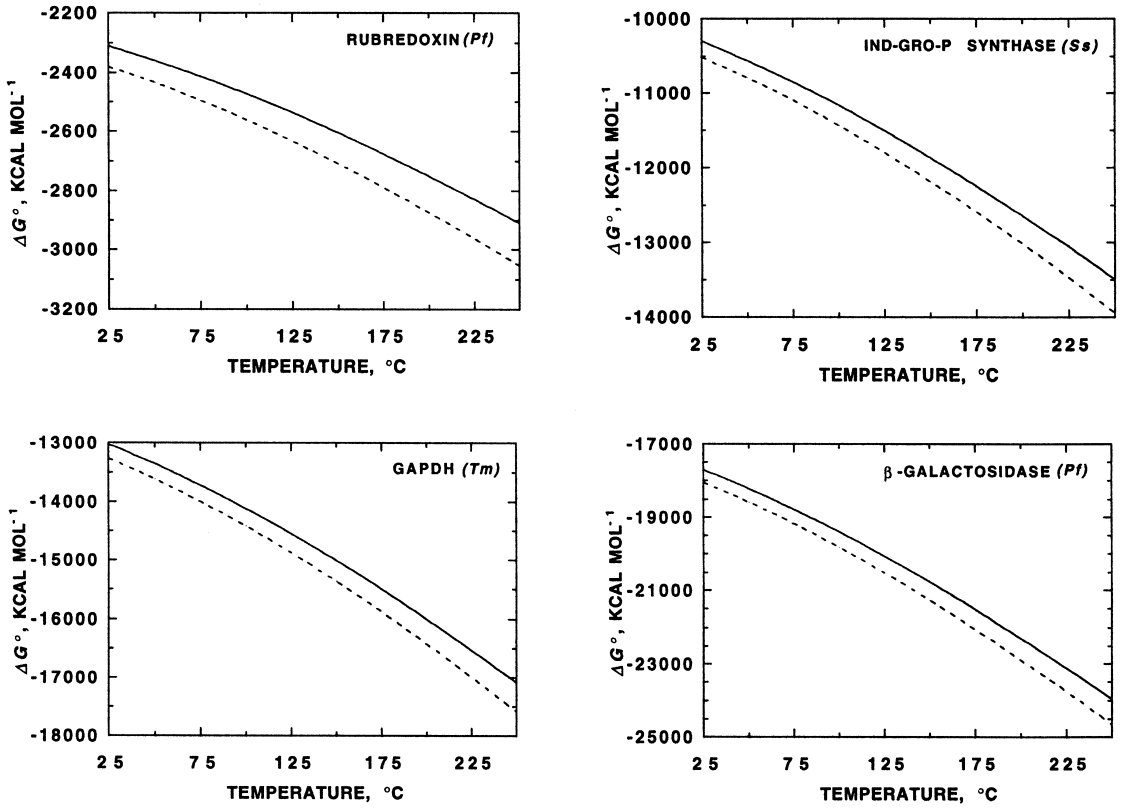

Fig. 9. ΔG^{o} as a function of temperature at P_{SAT} for rubredoxin (*Pf*), IndGroP synthase (*Ss*), GAPDH (*Tm*), and β-galactosidase (*Ss*). The curves were generated from Eq. (A14) using the group additivity algorithm represented by Eq. (6) and values of ΔG_f^{o} and equations of state parameters for the constituent groups given in Tables 6 and 7. The solid curves represent the proteins with fully ionized R groups in the His, Lys, Arg, Asp, and Glu residues. In contrast, only the R groups of the His, Lys, and Arg residues are ionized in the proteins represented by the dashed curves.

ΔH_f^{o}, S^{o}, C_P^{o}, and V^{o} at 25°C and 1 bar, respectively (Table 18). In order to estimate representative uncertainties in calculated values of ΔG^{o}, ΔH^{o}, S^{o}, C_P^{o}, and V^{o} of unfolded proteins at elevated temperatures and pressures, the uncertainties in the equations of state parameters used in the calculations need to be assessed. Representative uncertainties associated with these parameters can be taken to be equivalent to those for the amino acids [20]. Combining the contributions by the representative uncertainties associated with the standard molal thermodynamic properties at 25°C and 1 bar with those arising from the uncertainties in the equations of state parameters yields the total representative uncertainties in computed values of ΔG^{o}, ΔH^{o}, S^{o}, C_P^{o}, and V^{o} of amino acid residues at elevated temperatures and pressures. As examples, calculated maximum representative uncertainties of standard molal properties of amino acid residues at 100°C and P_{SAT} and at 250°C and 1 kbar are given in Table 18. The contributions of the uncertainties in equations of state parameters and standard molal properties at 25°C and 1 bar to the uncertainties in standard molal properties at elevated temperatures and pressures are also listed in Table 18. It should be emphasized that the representative uncertainties in the standard molal properties of unfolded proteins are not obtained by adding the corresponding uncertainties of the appropriate residues. Owing to probable cancellation of the uncertainties associated with the properties and

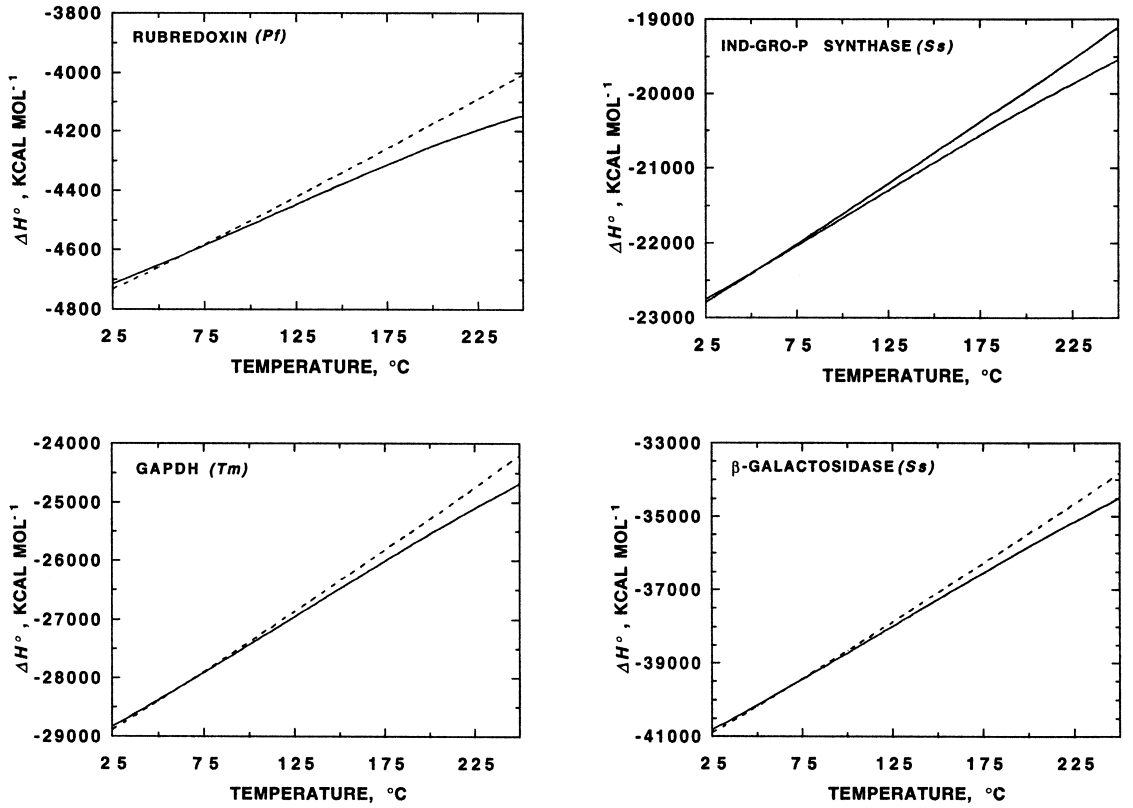

Fig. 10. ΔH° as a function of temperature at P_{SAT} for rubredoxin (*Pf*), IndGroP synthase (*Ss*), GAPDH (*Tm*), and β-galactosidase (*Ss*). The curves were generated from Eq. (A13) using the group additivity algorithm represented by Eq. (6) and values of ΔH_f° and equations of state parameters for the constituent groups taken from Tables 6 and 7 (see caption of Fig. 9).

parameters of the different kinds of residues in each protein, the representative uncertainties of the standard molal properties of unfolded proteins are likely to be considerably less.

7. Concluding remarks

The group additivity equation of state approach adopted in this communication for computing values of the standard molal thermodynamic properties of −Gly− and PB as a function of temperature and pressure, together with those of AAB and the R groups tabulated in the first of this series of papers [20] permits prediction of the thermodynamic behavior as a function of temperature and pressure of any unfolded protein for which the amino acid sequence is known. Calculations of this kind for C_P° can be combined with those of the corresponding heat capacities of folded proteins to quantify the effects of temperature on the standard molal thermodynamic properties of unfolding [4]. Nevertheless, additional experimental calorimetry, density, and sound velocity data for peptides and unfolded proteins at elevated temperatures are needed to verify and/or refine the group additivity approach and values of the equations of state parameters and standard molal properties at 25°C and 1 bar adopted in the present study. Until further refinements can be made, the properties and parameters tabulated above afford the means to generate close approximations of ΔG°, ΔH°, S°, C_P°, and V° for unfolded proteins as a function of temperature and pressure to at least 250°C and 1 kbar.

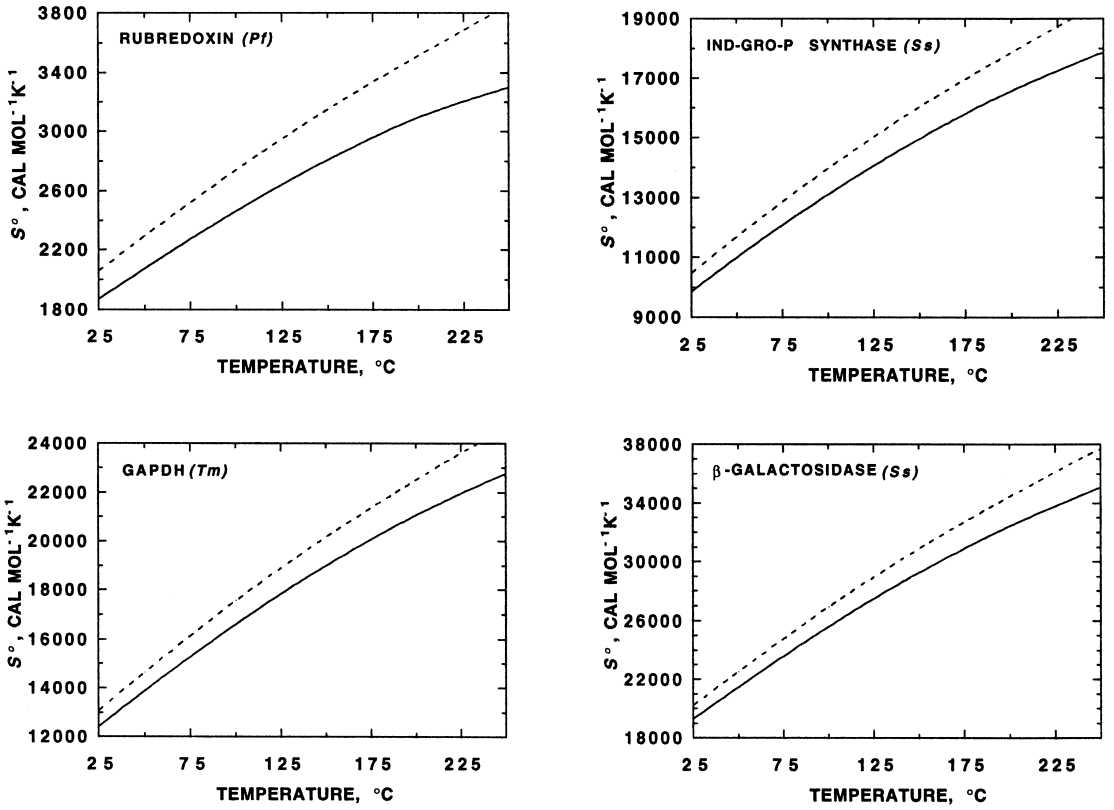

Fig. 11. S^{o} as a function of temperature at P_{SAT} for rubredoxin (*Pf*), IndGroP synthase (*Ss*), GAPDH (*Tm*), and β-galactosidase (*Ss*). The curves were generated from Eq. (A12) using the group additivity algorithm represented by Eq. (6) and values of S^{o} at 25°C and 1 bar and equations of state parameters for the constituent groups taken from Tables 6 and 7 (see caption of Fig. 9).

The equations of state and group additivity algorithms adopted in the present study have been incorporated into a software package called ORGANOBIOGEOTHERM (or O-BiG-T for short), which can be obtained at no cost from the Laboratory of Theoretical Geochemistry and Biogeochemistry at the University of California at Berkeley, otherwise known as Prediction Central.

Acknowledgements

The research reported above was supported by the National Science Foundation (NSF Grants EAR-8606052, EAR-9117393, and EAR-9613753), the Department of Energy (DOE Grant DE-FG03-85ER-13419), and the Committee on Research at the University of California, Berkeley. We are indebted to Christine Owens, Laurent Richard, Barbara Ransom, Eric Oelkers, Everett Shock, Vitalii Pokrovskii, Doug Clark, John Baross, Sergio Cabani, George Makhatadze, and Peter Privalov for helpful discussions, encouragement, and assistance during the course of this study. Thanks are also due Ken Dill for a constructive review of an earlier version of this manuscript. Finally, we would like to thank Gianni Colonna for his warm hospitality in providing space and support for us in the Research Center of Computational Science, Department of Biochemistry and Biophysics, Medical School at the Second University of Naples during HCH's 1995/96 sabbatical year in Italy.

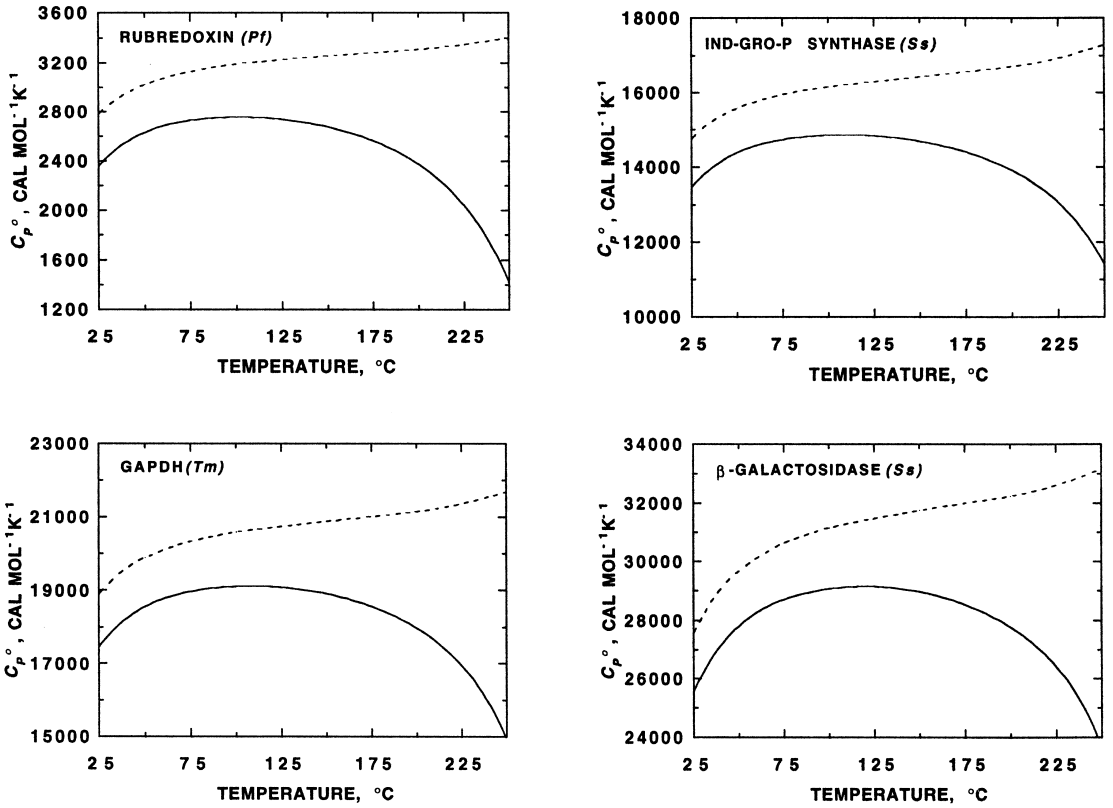

Fig. 12. C_P° as a function of temperature at P_{SAT} for rubredoxin (*Pf*), IndGroP synthase (*Ss*), GAPDH (*Tm*), and β-galactosidase (*Ss*). The curves were generated from Eq. (A2) using the group additivity algorithm represented by Eq. (6) and values of equations of state parameters for the constituent groups taken from Table 7 (see caption of Fig. 9).

Appendix A
Summary of the revised HKF equations of state

The standard molal thermodynamic properties of aqueous species in the revised HKF equations of state are expressed as the sum of structural and solvation contributions. A general statement of this separation of variables is represented by

$$\Xi = \Delta\Xi_n + \Delta\Xi_s \tag{A1}$$

where Ξ denotes any standard molal property, and $\Delta\Xi_n$ and $\Delta\Xi_s$ stand for the non-solvation (or structural) and solvation (or electrostatic) contributions to Ξ, respectively. The revised HKF equations for the standard molal heat capacity (C_P°) and volume (V°) of an aqueous species are given by [21]

$$\begin{aligned} C_P^\circ &= \Delta C_{P,n}^\circ + \Delta C_{P,s}^\circ \\ &= c_1 + \frac{c_2}{(T-\Theta)^2} - \left[\frac{2T}{(T-\Theta)^3}\right]\left[a_3(P-P_r) + a_4\ln\left(\frac{\Psi+P}{\Psi+P_r}\right)\right] + \omega TX + 2TY\left(\frac{\partial\omega}{\partial T}\right)_P \\ &\quad - T\left(\frac{1}{\varepsilon}-1\right)\left(\frac{\partial^2\omega}{\partial T^2}\right)_P \end{aligned} \tag{A2}$$

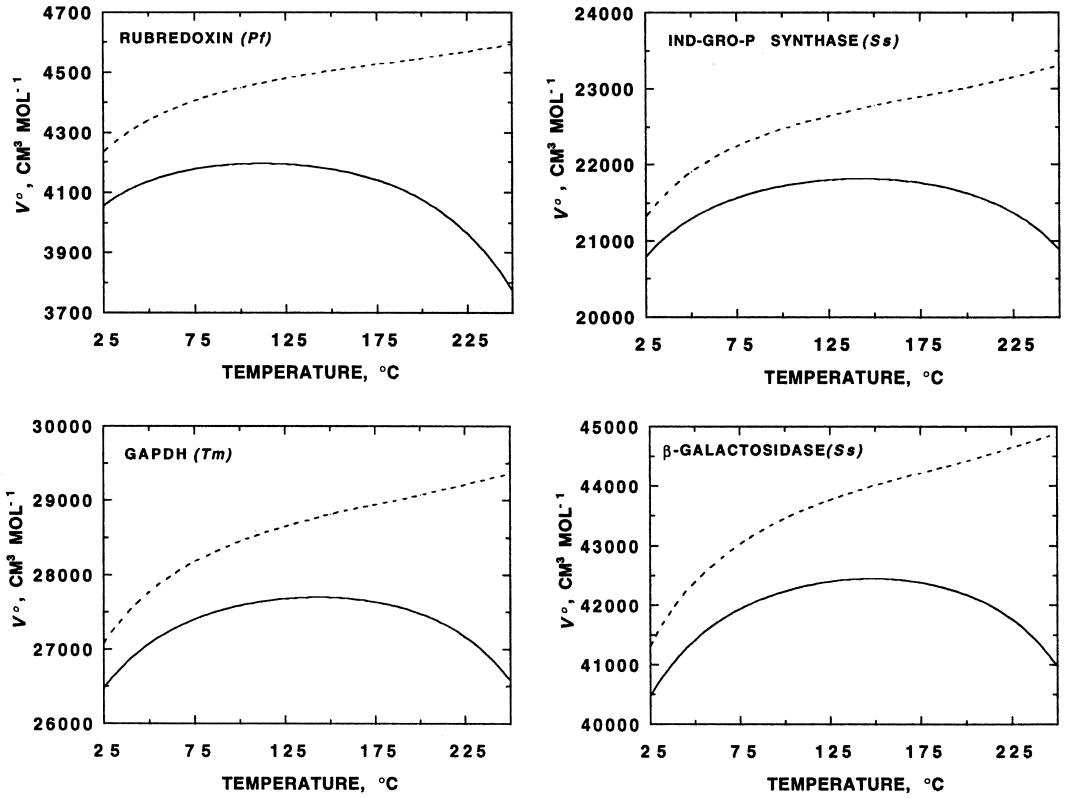

Fig. 13. V° as a function of temperature at P_{SAT} for rubredoxin (*Pf*), IndGroP synthase (*Ss*), GAPDH (*Tm*), and β-galactosidase (*Ss*). The curves were generated from Eq. (A3) using the group additivity algorithm represented by Eq. (6) and values of equations of state parameters for the constituent groups taken from Table 7 (see caption of Fig. 9).

and

$$V^{\circ} = \Delta V_n^{\circ} + \Delta V_s^{\circ}$$
$$= \sigma + \frac{\xi}{T-\Theta} - \omega Q + \left(\frac{1}{\varepsilon} - 1\right)\left(\frac{\partial \omega}{\partial P}\right)_T$$
$$= a_1 + \frac{a_2}{\Psi + P} + \left(a_3 + \frac{a_4}{\Psi + P}\right)\left(\frac{1}{T-\Theta}\right) - \omega Q + \left(\frac{1}{\varepsilon} - 1\right)\left(\frac{\partial \omega}{\partial P}\right)_T \quad \text{(A3)}$$

where c_1, c_2, σ, ξ, a_1, a_2, a_3, a_4, and ω stand for equation of state parameters for these species, Ψ and Θ denote solvent parameters equal to 2600 bar and 228 K, T and P represent the temperature and pressure of interest, P_r signifies the reference pressure of 1 bar, and X, Y, and Q represent the Born functions defined by

$$X \equiv \left(\frac{\partial Y}{\partial T}\right)_P \quad \text{(A4)}$$

$$Y \equiv -\left[\frac{\partial (1/\varepsilon)}{\partial T}\right]_P \quad \text{(A5)}$$

and

$$Q \equiv -\left[\frac{\partial (1/\varepsilon)}{\partial P}\right]_T \quad \text{(A6)}$$

where ε stands for the dielectric constant of pure water. The values of X, Y, and Q used in the present study were taken from Shock et al. [26]. The non-solvation contributions to C_P° and V°

Table 18
Calculated maximum representative uncertainties associated with computed values of the standard molal properties of amino acid residues at 100°C and P_{SAT} and at 250°C and 1 kbar arising from estimated uncertainties in the equation of state parameters and the values of the standard molal properties at 25°C and 1 bar

Property or parameter	Estimated uncertainty	ΔG^o (kcal mol^{-1})	ΔH^o (kcal mol^{-1})	S^o (cal mol^{-1} K^{-1})	C_P^o (cal mol^{-1} K^{-1})	V^o (cm^3 mol^{-1})
Uncertainty at 100°C and P_{SAT}						
ΔG_f^o (kcal mol^{-1})	0.02[a]	0.02				
ΔH_f^o (kcal mol^{-1})	0.20[a]		0.20			
S^o (cal mol^{-1} K^{-1})	0.74[a]	0.06		0.74		
C_P^o (cal mol^{-1} K^{-1})	1.00[a]				1.00	
V^o (cm^3 mol^{-1})	0.50[a]					0.50
a_1 (cal mol^{-1} bar^{-1} × 10)	0.08	0	0			0.33
a_2 (cal mol^{-1} × 10^{-2})	0.25	0	0			0.40
a_3 (cal K mol^{-1} bar^{-1})	0.75		0	0	0	0.22
a_4 (cal K mol^{-1}) × 10^{-4}	0.20	0	0	0	0	0.22
c_1(cal mol^{-1} K^{-1})	0.50	0	0.04	0.11	0.50	
c_2 (cal K mol^{-1}) × 10^{-4}	0.50	0.01	0.04	0.11	0.24	
ω (kcal mol^{-1})	5.00	0.01	0.04	0.12	0.71	0.24
Total uncertainty		0.10	0.32	1.08	2.45	1.91
Uncertainty at 250°C and 1 kbar						
ΔG_f^o (kcal mol^{-1})	0.02[a]	0.02				
ΔH_f^o (kcal mol^{-1})	0.20[a]		0.20			
S^o (cal mol^{-1} K^{-1})	0.74[a]	0.17		0.74		
C_P^o (cal mol^{-1} K^{-1})	1.00[a]					
V^o (cm^3 mol^{-1})	0.50[a]					0.50
a_1 (cal mol^{-1} bar^{-1} × 10)	0.08	0.01	0.01			0.33
a_2 (cal mol^{-1} × 10^{-2})	0.25	0.01	0.01			0.29
a_3 (cal K mol^{-1} bar^{-1})	0.75		0.01	0.01	0.03	0.11
a_4 (cal K mol^{-1}) × 10^{-4}	0.20	0	0.01	0.01	0.03	0.08
c_1(cal mol^{-1} K^{-1})	0.50	0.03	0.11	0.28	0.50	
c_2 (cal K mol^{-1}) × 10^{-4}	0.50	0.03	0.05	0.15	0.06	
ω (kcal mol^{-1})	5.00	0.03	0.16	0.37	1.50	0.66
Total uncertainty		0.30	0.56	1.56	3.12	1.97

[a]At 25°C and 1 bar.

can be expressed as

$$\Delta C^o_{P,n} = c_1 + c_2\frac{1}{(T-\Theta)^2} - \left[\frac{2T}{(T-\Theta)^3}\right] \times \left[a_3(P-P_r) + a_4\ln\left(\frac{\Psi+P}{\Psi+P_r}\right)\right] \quad \text{(A7)}$$

and

$$\Delta V^o_n = \sigma + \xi\frac{1}{T-\Theta} \quad \text{(A8)}$$

At $P = P_r$, Eq. (A7) reduces to

$$\Delta C^o_{P,n} = c_1 + c_2\frac{1}{(T-\Theta)^2} \quad \text{(A9)}$$

The solvation contributions to C_P^o and V^o can be written as

$$\Delta C^o_{P,s} = \omega TX + 2TY\left(\frac{\partial\omega}{\partial T}\right)_P T\left(\frac{1}{\varepsilon}-1\right)\left(\frac{\partial^2\omega}{\partial T^2}\right)_P \quad \text{(A10)}$$

and

$$\Delta V_s^{\circ} = -\omega Q + \left(\frac{1}{\varepsilon} - 1\right)\left(\frac{\partial \omega}{\partial P}\right)_T \quad \text{(A11)}$$

The standard molal entropy of an aqueous species as a function of temperature and pressure can be computed from

$$\begin{aligned} S^{\circ} = {} & S^{\circ}_{P_r,T_r} + c_1 \ln\left(\frac{T}{T_r}\right) - \frac{c_2}{\Theta} \\ & \times \left\{ \left(\frac{1}{T-\Theta}\right) - \left(\frac{1}{T_r-\Theta}\right) \right. \\ & \left. + \frac{1}{\Theta} \ln\left[\frac{T_r(T-\Theta)}{T(T_r-\Theta)}\right] \right\} + \left(\frac{1}{T-\Theta}\right)^2 \\ & \times \left[a_3(P-P_r) + a_4 \ln\left(\frac{\Psi+P}{\Psi+P_r}\right) \right] + \omega Y \\ & - \left(\frac{1}{\varepsilon} - 1\right)\left(\frac{\partial \omega}{\partial T}\right)_P - \omega_{P_r,T_r} Y_{P_r,T_r} \end{aligned} \quad \text{(A12)}$$

where T_r represents the reference temperature of 298.15 K. The apparent standard molal enthalpy (ΔH°) and Gibbs free energy (ΔG°) of formation are defined by

$$\begin{aligned} \Delta H^{\circ} \equiv {} & \Delta H_f^{\circ} + \left(H^{\circ}_{P,T} - H^{\circ}_{P_r,T_r}\right) \\ = {} & \Delta H_f^{\circ} + c_1(T-T_r) - c_2 \\ & \times \left[\left(\frac{1}{T-\Theta}\right) - \left(\frac{1}{T_r-\Theta}\right) \right] + a_1(P-P_r) \\ & + a_2 \ln\left(\frac{\Psi+P}{\Psi+P_r}\right) + \left[\frac{2T-\Theta}{(T-\Theta)^2}\right] \\ & \times \left[a_3(P-P_r) + a_4 \ln\left(\frac{\Psi+P}{\Psi+P_r}\right) \right] \\ & + \omega\left(\frac{1}{\varepsilon} - 1\right) + \omega T Y - T\left(\frac{1}{\varepsilon} - 1\right)\left(\frac{\partial \omega}{\partial T}\right)_P \\ & - \omega_{P_r,T_r}\left(\frac{1}{\varepsilon_{P_r,T_r}} - 1\right) - \omega_{P_r,T_r} T_r Y_{P_r,T_r} \end{aligned} \quad \text{(A13)}$$

and

$$\begin{aligned} \Delta G^{\circ} \equiv {} & \Delta G_f^{\circ} + \left(G^{\circ}_{P,T} - G^{\circ}_{P_r,T_r}\right) \\ = {} & \Delta G_f^{\circ} - S^{\circ}_{P_r,T_r}(T-T_r) \\ & - c_1\left[T \ln\left(\frac{T}{T_r}\right) - T + T_r \right] \\ & - c_2\left\{ \left[\left(\frac{1}{T-\Theta}\right) - \left(\frac{1}{T_r-\Theta}\right) \right]\left(\frac{\Theta-T}{\Theta}\right) \right. \\ & \left. - \frac{T}{\Theta^2} \ln\left[\frac{T_r(T-\Theta)}{T(T_r-\Theta)}\right] \right\} + a_1(P-P_r) \\ & + a_2 \ln\left(\frac{\Psi+P}{\Psi+P_r}\right) + \left(\frac{1}{T-\Theta}\right) \\ & \times \left[a_3(P-P_r) + a_4 \ln\left(\frac{\Psi+P}{\Psi+P_r}\right) \right] \\ & + \omega\left(\frac{1}{\varepsilon} - 1\right) - \omega_{P_r,T_r}\left(\frac{1}{\varepsilon_{P_r,T_r}} - 1\right) \\ & - \omega_{P_r,T_r} Y_{P_r,T_r}(T-T_r) \end{aligned} \quad \text{(A14)}$$

where ΔH_f° and ΔG_f° refer to the standard molal enthalpy and Gibbs free energy of formation from the elements in their stable form at 298.15 K and 1 bar, respectively. For neutral species $(\partial\omega/\partial P)_T$, $(\partial\omega/\partial T)_P$, $(\partial^2\omega/\partial P^2)_T$, and $(\partial^2\omega/\partial T^2)_P$ in Eqs. (A2) and (A3) and Eqs. (A10), (A11), (A12), (A13) and (A14) are taken to be zero [24,25,35].

Values of ω for Gly, Gly_n ($n = 2$–5), –Gly–, AAB, and PB given in Table A1 were estimated from group additivity algorithms adopted in this study, together with correlations of ω with S° for neutral aqueous species discussed in detail in Amend and Helgeson [45]. The value of ω for PB was generated with the value of S° given in Table 6 from the correlation shown in Fig. 14, which represents a modification of the one shown in

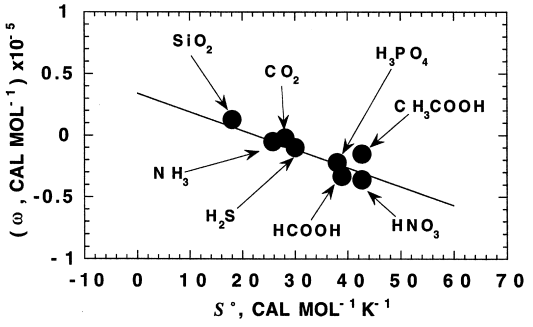

Fig. 14. Correlation of ω with S° at 25°C and 1 bar for neutral aqueous organic and inorganic species with at least one hetero-atom. The regression line corresponds to that generated by Shock et al. [25], which is consistent with Eq. (A15) (see text).

Amend and Helgeson [20]. The equation of the correlation line in this figure is given by

$$\omega = -1514.4S^{o} + 0.34 \times 10^{5} \quad \text{(A15)}$$

and can be used to estimate values of ω for groups with at least one hetero-atom. The values of ω for Gly and AAB were taken from Amend and Helgeson [20] and those for −Gly− and Gly_n were calculated with Eqs. (5) and (6), respectively, using values of the requisite groups given in Table 7 and Table A1.

Table A1
Values of S^{o} and ω for Gly [20] and those for Gly_n, −Gly−, AAB, and PB generated in the present study (see text)

Group	S^{oa}	$\omega \times 10^{-5b}$
Gly	39.29	−0.2550
Gly_2	66.10	−0.3210
Gly_3	92.91	−0.3870
Gly_4	119.72	−0.4530
Gly_5	146.53	−0.5190
−Gly−	26.81	−0.0660
AAB	25.67	−0.0487
PB	13.18	0.1403

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

References

[1] M.W.W. Adams, R.M. Kelly, Biocatalysis at extreme temperatures. Enzyme systems near and above 100°C, ACS Symposium Series 498 (1992) 215.
[2] P.L. Privalov, Stability of proteins. Small globular proteins, Adv. Protein Chem. 33 (1979) 167–241.
[3] G.I. Makhatadze, Measuring protein thermostability by differential scanning calorimetry, Current Protocols Protein Sci. (1998) 7.9.1–7.9.14.
[4] H.C. Helgeson, C.E. Owens, J.P. Amend, Toward a global equation of state for protein unfolding as a function of temperature, pressure, and solution composition, Protein Sci. (2000) (in prep.).
[5] P.L. Privalov, N.N. Khechinashvili, A thermodynamic approach to the problem of stabilization of globular protein structure: a calorimetric study, J. Mol. Biol. 86 (1974) 665–684.
[6] E.I. Tiktopulo, P.L. Privalov, Papain denaturation is not a two-state transition, FEBS Lett. 91 (1978) 57–58.
[7] P.L. Mateo, P.L. Privalov, Pepsinogen denaturation is not a two-state transition, FEBS Lett. 123 (1981) 189–192.
[8] Y.V. Griko, P.L. Privalov, S.Y. Venyaminov, Thermodynamic study of the apomyoglobin structure, J. Mol. Biol. 202 (1988) 127–138.
[9] Y.V. Griko, G.I. Makhatadze, P.L. Privalov, R.W. Hartley, Thermodynamics of barnase unfolding, Protein Sci. 3 (1994) 669–676.
[10] P.L. Wintrode, G.I. Makhatadze, P.L. Privalov, Thermodynamics of ubiquitin unfolding, Proteins Struct. Funct. Genet. 18 (1994) 246–253.
[11] B.S. McCrary, S.P. Edmondson, J.W. Shriver, Hyperthermophile protein folding thermodynamics: differential scanning calorimetry and chemical denaturation of Sac7d, J. Mol. Biol. 264 (1996) 784–805.
[12] P.L. Privalov, E.I. Tiktopulo, S.Y. Venyaminov, Y.V. Griko, G.I. Makhatadze, N.N. Khechinashvili, Heat capacity and conformation of proteins in the denatured state, J. Mol. Biol. 205 (1989) 737–750.
[13] G.I. Makhatadze, S.J. Gill, P.L. Privalov, Partial molar heat capacities of the side chains of some amino acid residues in aqueous solution, Biophys. Chem. 38 (1990) 33–37.
[14] G.I. Makhatadze, P.L. Privalov, Heat capacity of proteins. I. Partial molar heat capacity of individual amino acid residues in aqueous solution: hydration effect, J. Mol. Biol. 213 (1990) 375–384.
[15] P.L. Privalov, G.I. Makhatadze, Heat capacity of proteins. II. Partial molar heat capacity of the unfolded polypeptide chain of proteins: protein unfolding effects, J. Mol. Biol. 213 (1990) 385–391.
[16] K.A. Dill, D. Shortle, Denatured states of proteins, Ann. Rev. Biochem. 60 (1991) 795–825.
[17] S.W. Benson, Thermochemical Kinetics, J. Wiley & Sons, New York, 1968.
[18] H.C. Helgeson, D.H. Kirkham, G.C. Flowers, Theoretical prediction of the thermodynamic behavior of aqueous electrolytes at high pressures and temperatures: IV. 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 (1981) 1249–1516.
[19] P.W. Atkins, Physical Chemistry, W.H. Freeman and Co, New York, 1986.
[20] J.P. Amend, H.C. Helgeson, 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 (1997) 1927–1941.
[21] J.C. Tanger, H.C. Helgeson, 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 (1988) 19–98.
[22] D.A. Sverjensky, Calculation of the thermodynamic properties of aqueous species and the solubilities of minerals in supercritical electrolyte solutions. In: I.S.E. Carmichael, H.P. Eugster (Eds.), Reviews in Mineralogy, Mineralogical Society of America, 1987.
[23] E.L. Shock, H.C. Helgeson, Calculation of the thermodynamic and transport properties of aqueous species at

high pressures and temperatures: correlation algorithms for ionic species and equation of state predictions to 5 kb and 1000°C, Geochim. Cosmochim. Acta 52 (1988) 2009–2036.

[24] E.L. Shock, H.C. Helgeson, Calculation of the thermodynamic and transport properties of aqueous species at high pressures and temperatures: standard partial molal properties of organic species, Geochim. Cosmochim. Acta 54 (1990) 915–945.

[25] E.L. Shock, H.C. Helgeson, D.A. Sverjensky, Calculation of the thermodynamic and transport properties of aqueous species at high pressures and temperatures: standard partial molal properties of inorganic neutral species, Geochim. Cosmochim. Acta 53 (1989) 2157–2183.

[26] E.L. Shock, E.H. Oelkers, J.W. Johnson, D.A. Sverjensky, H.C. Helgeson, Calculation of the thermodynamic properties of aqueous species at high pressures and temperatures: effective electrostatic radii, dissociation constants and standard partial molal properties to 1000°C and 5 kbar, J. Chem. Soc., Faraday Trans. 88 (1992) 803–826.

[27] E.L. Shock, D.C. Sassani, M. Willis, D.A. Sverjensky, Inorganic species in geologic fluids: correlations among standard molal thermodynamic properties of aqueous ions and hydroxide complexes, Geochim. Cosmochim. Acta 61 (1997) 907–950.

[28] H.C. Helgeson, Organic/inorganic reactions in metamorphic processes, Can. Mineral. 29 (1991) 707–739.

[29] H.C. Helgeson, Effects of complex formation in flowing fluids on the hydrothermal solubilities of minerals as a function of fluid pressure and temperature in the critical and supercritical regions of the system H_2O, Geochim. Cosmochim. Acta 56 (1992) 3191–3207.

[30] H.C. Helgeson, Calculation of the thermodynamic properties and relative stabilities of aqueous acetic and chloroacetic acids, acetate and chloroacetates, and acetyl and chloroacetyl chlorides at high and low temperatures and pressures, Appl. Geochem. 7 (1992) 291–308.

[31] H.C. Helgeson, Thermodynamic prediction of the relative stabilities of hyperthermophilic enzymes, Chemical Thermodynamics, T.M. Letcher Blackwell Science Ltd, Oxford, 1999.

[32] J.W. Johnson, E.H. Oelkers, H.C. Helgeson, SUPCRT92: a software package for calculating the standard molal properties of minerals, gases, aqueous species, and reactions from 1 to 5000 bar and 0 to 1000°C, Comp. Geosci. 18 (1992) 899–947.

[33] E.L. Shock, Stability of peptides in high-temperature aqueous solutions, Geochim. Cosmochim. Acta 56 (1992) 3481–3491.

[34] E.L. Shock, Hydrothermal dehydration of aqueous organic compounds, Geochim. Cosmochim. Acta 57 (1993) 3341–3349.

[35] E.L. Shock, 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 (1995) 496–580.

[36] H.C. Helgeson, A.M. Knox, C.E. Owens, E.L. Shock, Petroleum, oil field waters, and authigenic mineral assemblages: are they in metastable equilibrium in hydrocarbon reservoirs? Geochim. Cosmochim. Acta 57 (1993) 3295–3339.

[37] H.C. Helgeson, L. Richard, C.E. Owens, J.P. Amend, The organic–inorganic interface in hydrothermal systems, In: Y.K. Kharaka, O. Chudaev, Water–Rock Interaction, Proceedings of the 8th International Symposium on Water–Rock Interaction, Vladivostok, 1995.

[38] E.H. Oelkers, H.C. Helgeson, Multiple ion association in supercritical aqueous solutions of single electrolytes, Science 261 (1993) 888–891.

[39] M.D. Schulte, E.L. Shock, Aldehydes in hydrothermal solution: standard partial molal thermodynamic properties and relative stabilities at high temperatures and pressures, Geochim. Cosmochim. Acta 57 (1993) 3835–3846.

[40] E.L. Shock, C.M. Koretsky, Metal–organic complexes in geochemical processes — calculation of standard partial molal thermodynamic properties of aqueous acetate complexes at high pressures and temperatures, Geochim. Cosmochim. Acta 57 (1993) 4899–4922.

[41] H.C. Helgeson, J.P. Amend, Relative stabilities of biomolecules at high temperatures and pressures, Thermochim. Acta 245 (1994) 89–119.

[42] J.R. Haas, E.L. Shock, D.C. Sassani, Rare earth elements in hydrothermal systems: estimates of standard partial molal thermodynamic properties of aqueous complexes of the REE at high pressures and temperatures, Geochim. Cosmochim. Acta 59 (1995) 4329–4350.

[43] V.A. Pokrovskii, H.C. Helgeson, Thermodynamic properties of aqueous species and the solubilities of minerals at high pressures and temperatures: the system Al_2O_3–H_2O–NaCl, Am. J. Sci. 295 (1995) 1255–1342.

[44] C.B. Xiao, P.R. Tremaine, Apparent molar heat capacities and volumes of $LaCl_3$(aq), $La(ClO_4)_3$(aq), and $Gd(ClO_4)_3$(aq) between the temperatures 283 K and 338 K, J. Chem. Thermodyn. 28 (1996) 43–66.

[45] J.P. Amend, H.C. Helgeson, Group additivity equations of state for calculating the standard molal thermodynamic properties of aqueous organic molecules at elevated temperatures and pressures, Geochim. Cosmochim. Acta 61 (1997) 11–46.

[46] A.W. Hakin, A.K. Copeland, J.L. Liu, R.A. Marriott, K.E. Preuss, 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 (1997) 84.

[47] D.A. Sverjensky, E.L. Shock, H.C. Helgeson, Prediction of the thermodynamic and transport properties of aqueous metal complexes to 1000°C and 5 kb, Geochim. Cosmochim. Acta 61 (1997) 1359–1412.

[48] R.A. Marriott, A.W. Hakin, J.L. Liu, Modeling of thermodynamic properties of amino acids and peptides using

additivity and HKF theory, J. Solution Chem. 27 (1998) 771–802.
[49] E.J. Cohn, J.T. Edsall, Proteins Amino Acids and Peptides as Ions and Dipolar Ions, Reinhold Publishing Corporation, 1943.
[50] B. Jacobsen, On the adiabatic compressibility of aqueous solutions, Ark. Kemi. 2 (1950) 177–210.
[51] K.P. Murphy, S.J. Gill, Solid model compounds and the thermodynamics of protein unfolding, J. Mol. Biol. 222 (1991) 699–709.
[52] D.P. Kharakoz, A.P. Sarvazyan, Hydrational and intrinsic compressibilities of globular proteins, Biopolymers 33 (1993) 11–26.
[53] G.I. Makhatadze, V.N. Medvedkin, P.L. Privalov, Partial molar volumes of polypeptides and their constituent groups in aqueous solution over a broad temperature range, Biopolymers 30 (1990) 1001–1010.
[54] P.L. Privalov, G.I. Makhatadze, Contribution of hydration to protein folding thermodynamics. II. The entropy and Gibbs free energy of hydration, J. Mol. Biol. 232 (1993) 660–679.
[55] G.I. Makhatadze, P.L. Privalov, Contribution of hydration to the protein folding thermodynamics. I. The enthalpy of hydration, J. Mol. Biol. 232 (1993) 639–659.
[56] G.I. Makhatadze, P.L. Privalov, Energetics of protein structure, Adv. Protein Chem. 47 (1995) 307–425.
[57] J.D. Cox, D.D. Wagman, V.A. Medvedev, CODATA Key Values for Thermodynamics, Hemisphere Publishing Corporation, New York, 1989.
[58] C. Jolicoeur, J. Boileau, Apparent molal volumes and heat capacities of low molecular weight peptides in water at 25°C, Can. J. Chem. 56 (1978) 2707–2713.
[59] M.O. Dayhoff, L.T. Hunt, W.C. Barker, R.M. Schwartz, M.C. Blomquist, Atlas of Protein Sequence and Structure, National Biomedical Research Foundation, Washington, DC, 1972.
[60] H.C. Helgeson, D.H. Kirkham, Theoretical prediction of the thermodynamic properties of aqueous electrolytes at high pressures and temperatures. III. Equation of state for aqueous species at infinite dilution, Am. J. Sci. 276 (1976) 97–240.
[61] F.T. Gucker, W.L. Ford, C.E. Moser, The apparent and partial molal heat capacities and volumes of glycine and glycolamide, J. Phys. Chem. 43 (1939) 153–168.
[62] G.R. Hedwig, J.F. Reading, T.H. Lilley, Aqueous solutions containing amino acids and peptides. Part 27. Partial molar heat capacities and partial molar volumes of some *n*-acetyl amino acid amides, some *n*-acetyl peptide amides and two peptides at 25°C, J. Chem. Soc., Faraday Trans. 87 (1991) 1751–1758.
[63] G.C. Kresheck, L. Benjamin, Calorimetric studies of the hydrophobic nature of several protein constituents and ovalbumin in water and in aqueous urea, J. Phys. Chem. 68 (1964) 2476–2486.
[64] S. Cabani, G. Conti, E. Matteoli, Apparent molal heat capacities in aqueous solution of molecules containing the peptide linkage: cyclic and open chain dipeptides, Biopolymers 16 (1977) 465–467.
[65] C.H. Spink, I. Wadsö, Thermochemistry of solutions of biochemical model compounds. 4. The partial molar heat capacities of some amino acids in aqueous solution, J. Chem. Thermodyn. 7 (1975) 561–572.
[66] C.J. Downes, G.R. Hedwig, Partial molar heat capacities of the peptides glycylglycylglycine, glycyl-L-ananylglycine and glycyl-DL-threonylglycine in aqueous solution over the temperature range 50°C to 125°C, Biophys. Chem. 55 (1995) 279–288.
[67] A.K. Mishra, J.C. Ahluwalia, Apparent molal volumes of amino acids, *n*-acetylamino acids and peptides in aqueous solutions, J. Phys. Chem. 88 (1984) 86–92.
[68] F.J. Millero, A. LoSurdo, C. Shin, The apparent molal volumes and adiabatic compressibilities of aqueous amino acids at 25°C, J. Phys. Chem. 82 (1978) 784–792.
[69] S. Cabani, P. Gianni, V. Mollica, L. Lepori, Group contributions to the thermodynamic properties of nonionic organic solutes in dilute aqueous solutions, J. Solution Chem. 10 (1981) 563–595.
[70] D.P. Kharakoz, Volumetric properties of proteins and their analogs in diluted water solutions. 1. Partial volumes of amino acids at 15–55°C, Biophys. Chem. 34 (1989) 115–125.
[71] J. Kirchnerova, P.G. Farrell, J.T. Edward, Dielectric increments and the conformations of amino acids and betaines in water, J. Phys. Chem. 80 (1976) 1974–1980.
[72] T.V. Chalikian, A.P. Sarvazyan, K.J. Breslauer, Partial molar volumes, expansibilities, and compressibilities of α,ω-aminocarboxylic acids in aqueous solutions between 18 and 55°C, J. Phys. Chem. 97 (1993) 13017–13026.
[73] K. Bülent Belibagli, E. Ayranci, Viscosities and apparent molar volumes of some amino acids in water and in 6 M guanidine hydrochloride at 25°C, J. Solution Chem. 19 (1990) 867–882.
[74] B.S. Lark, K. Bala, Apparent molal volumes and apparent molal expansibilities of glycine in aqueous solutions, Indian J. Chem. 22A (1983) 192–194.
[75] R.K. Wadi, R.K. Goyal, Temperature dependence of apparent molar volumes and viscosity B-coefficients of amino acids in aqueous potassium thyocyanate solutions from 15 to 35°C, J. Solution Chem. 21 (1992) 163–170.
[76] H.D. Ellerton, G. Reinfelds, D.E. Mulcahy, J.P. Dunlop, Activity, density, and relative viscosity data for several amino acids lactamide and raffinose in aqueous solution at 25°C, J. Phys. Chem. 68 (1964) 398–402.
[77] E.J. Cohn, T.L. McMeekin, J.T. Edsall, M.H. Blanchard, Studies in the physical chemistry of amino acids, peptides and related substances. I. The apparent molal volume and the electrostriction of the solvent, J. Am. Chem. Soc. 56 (1934) 784–794.
[78] M. Iqbal, R.E. Verrall, Partial molar volumes and adiabatic compressibilities of glycyl peptides, J. Phys. Chem. 91 (1987) 967–971.
[79] K.P. Prasad, J.C. Ahluwalia, Heat capacity changes and

partial molal heat capacities of some di- and tripeptides in water, Biopolymers 19 (1980) 263–272.

[80] P.R. Blake, J.B. Park, F.O. Bryant et al., Determinants of protein hyperthermostability: purification and amino acid sequence of rubredoxin from the hyperthermophilic archaebacterium *Pyrococcus furiosus* and secondary structure of the zinc adduct by NMR, Biochemistry 30 (1991) 10885–10895.

[81] S.C. Busse, G.N. LaMar, L.P. Yu et al., Proton NMR investigation of the oxidized three-iron clusters in the ferredoxins from the hyperthermophilic archae *Pyrococcus furiosus* and *Thermococcus litoralis*, Biochemistry 31 (1992) 11952–11962.

[82] M. Bachleitner, W. Ludwig, K.O. Stetter, K.H. Schleifer, Nucleotide sequence of the gene coding for the elongation factor Tu from the extremely thermophilic eubacterium *Thermotoga maritima*, FEMS Microbiol. Lett. 57 (1989) 115–120.

[83] A.M. Engel, Z. Cejka, A. Lupas, F. Lottspeich, W. Baumeister, Isolation and cloning of Ompα, a coiled-coil protein spanning the periplasmic space of the ancestral eubacterium *Thermotoga maritima*, EMBO J. 11 (1992) 4369–4378.

[84] A.M. Sanangelantoni, G. Forlani, F. Ambroselli, P. Cammarano, O. Tiboni, The glnA gene of the extremely thermophilic eubacterium *Thermotoga maritima*: cloning primary structure and expression in *Escherichia coli*, J. Gen. Microbiol. 138 (1992) 383–393.

[85] V. Schultes, R. Deutzmann, R. Jaenicke, Complete amino-acid sequence of glyceraldehyde-3-phosphate dehydrogenase from the hyperthermophilic eubacterium *Thermotoga maritima*, Eur. J. Biochem. 192 (1990) 25–31.

[86] M.V. Cubellis, C. Rozzo, P. Montecucchi, M. Rossi, Isolation and sequencing of a new β-galactosidase-encoding archaebacterial gene, Gene 94 (1990) 89–94.

[87] G. Andreotti, M.L. Tutino, G. Sannia, G. Marino, M.V. Cubellis, Indole-3-glycerol-phosphate synthase from *Sulfolobus solfataricus* as a model for studying thermostable TIM-barrel enzymes, Biochim. Biophys. Acta 1208 (1994) 310–315.