ELSEVIER

Biophysical Chemistry 89 (2001) 253–264

Biophysical Chemistry

www.elsevier.nl/locate/bpc

Group additivity calculations of the thermodynamic properties of unfolded proteins in aqueous solution: a critical comparison of peptide-based and HKF models

Andrew W. Hakin[a], Gavin R. Hedwig[b,*]

[a] *Department of Chemistry and Biochemistry, University of Lethbridge, Lethbridge, Alberta T1K 3M4, Canada*
[b] *Institute of Fundamental Sciences − Chemistry, Massey University, Private Bag 11222, Palmerston North, New Zealand*

Received 7 August 2000; received in revised form 17 October 2000; accepted 17 October 2000

Abstract

A recent paper in this journal [Amend and Helgeson, Biophys. Chem. 84 (2000) 105] presented a new group additivity model to calculate various thermodynamic properties of unfolded proteins in aqueous solution. The parameters given for the revised Helgeson–Kirkham–Flowers (HKF) equations of state for all the constituent groups of unfolded proteins can be used, in principle, to calculate the partial molar heat capacity, $C^{o}_{p,2}$, and volume, V^{0}_{2}, at infinite dilution of any polypeptide. Calculations of the values of $C^{o}_{p,2}$ and V^{0}_{2} for several polypeptides have been carried out to test the predictive utility of the HKF group additivity model. The results obtained are in very poor agreement with experimental data, and also with results calculated using a peptide-based group additivity model. A critical assessment of these two additivity models is presented.

Keywords: Group additivity; Unfolded proteins; Heat capacity; Volume; Helgeson–Kirkham–Flowers (HKF) model

1. Introduction

The random coil form of an unfolded protein in aqueous solution is one which is completely devoid of any structure [1,2]. It is, in effect, at one extreme of a range of different structural conformations that have been observed for proteins in their denatured states [2]. This random coil form of a denatured protein is of particular interest because it is the ideal reference state to use in discussions of the thermodynamic stability of proteins in aqueous solution [3]. Since some residual structure is retained in the denatured states for

* Corresponding author.
E-mail addresses: hakin@uleth.ca (A.W. Hakin), g.hedwig@massey.ac.nz (G.R. Hedwig).

0301-4622/01/$ - see front matter
PII: S0301-4622(00)00229-5

many proteins [2,4], the random coil form is not widely accessible by experiment. It is useful, therefore, to have a procedure by which its thermodynamic properties can be obtained by some indirect method. One such method is to use the principle of group additivity. This procedure assumes that any thermodynamic property of a fully unfolded protein, e.g. the partial molar volume or heat capacity, is given by a simple summation of the thermodynamic properties of the various constituent groups of the protein molecule. It is convenient in the case of proteins to choose as the constituent groups the various amino acid side-chains, the repeating unit of the backbone chain, and the ionic amino and carboxyl terminal groups. Assuming the principle of group additivity applies to oligopeptides, then any partial molar thermodynamic property for a protein at infinite dilution, Y_2^0, is given by the expression:

$$Y_2^0 = Y_2^0(NH_3^+ + CHCOO^-) + (N-1)Y_2^0(CHCONH) + \Sigma Y_2^0(R_i) \quad (1)$$

where N is the number of amino acid residues in the protein, R_i is the side-chain of the ith amino acid, CHCONH is the repeating unit of the backbone peptide chain, and NH_3^+ and $CHCOO^-$ are the ionic end groups of the peptide chain. The group contributions to the thermodynamic property are obtained using data for small solutes that are chosen to model the various components of a protein molecule.

In recent years, the amino acid side-chains in a protein have been modeled using a variety of small molecules. The zwitterionic amino acids have been used to derive, for example, the partial molar volume [5–7] and partial molar compressibility [5,6] of the amino acid side-chains in proteins. One of the limitations in using the simple amino acids as model compounds is that the charged NH_3^+ and COO^- groups interfere with the solvation of the adjacent side-chain. The calculation of side-chain contributions based on thermodynamic data for the amino acids will not necessarily be appropriate for a polypeptide. Makhatadze and Privalov [8] used the partial molar heat capacities of a disparate set of small organic molecules to derive the heat capacity contributions of amino acid side-chains. For example, the side-chains of the amino acids leucine and glutamine were modeled using *n*-butane and propionamide, respectively. The use of these solutes as model compounds has severe limitations, as has been discussed in detail elsewhere [9,10]. In previous work over several years [9,11,12], we have used small peptides as compounds to model the amino acid side-chains of proteins. This peptide-based model to evaluate the group contributions for the various constituents of proteins, the details of which are outlined below, has some distinct advantages over previous models that have been used. It also shows considerable promise [9,13] as a means to estimate reliable partial molar heat capacities and volumes of unfolded proteins.

In a recent paper in this journal [14], Amend and Helgeson have proposed yet another method, based on group additivity algorithms, to predict various thermodynamic properties of unfolded proteins over extended ranges of temperature and pressure. Unfortunately, in their paper, no consideration was given to recent work on peptide-based group additivity [9], or to other additivity models reported in the literature [5,8]. It is important, therefore, to examine carefully the predictive utility of the model proposed by Amend and Helegson [14]. In this paper, we report a critical comparison of how reliably their model, and our own peptide-based model [9,13], predict the partial molar volumes and heat capacities of oligopeptides in aqueous solution. As shown herein, the results from this comparison are quite astounding.

2. Peptide-based group additivity model

To obtain reliable estimations of thermodynamic properties of an unfolded protein using Eq. (1), it is important to choose as model compounds those that resemble, as closely as possible, the moieties being modeled. Several years ago we [11] proposed the use of tripeptides of sequence glycyl-X-glycine (gly-X-gly), where X is one of the naturally occurring amino acids, as compounds to

model the amino acid side-chains in polypeptides and proteins. In these tripeptides, the single side-chain on the central amino acid is flanked by two peptide groups, which is structurally analogous to that found in proteins. The retention of this backbone peptide structure, along with the side-chain, makes the tripeptides ideal compounds to model the side-chains of proteins.

The general method used to derive amino acid side-chain contributions is illustrated as follows for the thermodynamic property heat capacity. The partial molar heat capacity of any side-chain R of amino acid X, $C_p^o(R)$, is obtained using the equation:

$$C_p^o(R) = C_{p,2}^o(\text{gly-X-gly}) - C_{p,2}^o(\text{glyglygly}) + C_p^o(H) \quad (2)$$

where $C_{p,2}^o$(gly-X-gly) and $C_{p,2}^o$(glyglygly) are the partial molar heat capacities at infinite dilution for the peptides gly-X-gly and glyglygly, respectively, and $C_p^o(H)$ is the heat capacity of the hydrogen atom of the methylene moiety of triglycine. The assumption inherent in the derivation of $C_p^o(R)$ using Eq. (2) is that any interactions between the side-chain R and the charged amino and carboxyl groups in the tripeptide gly-X-gly make negligible contributions to the side-chain heat capacity. This assumption has been verified in a recent paper [15]. The side-chain heat capacities derived using $C_{p,2}^o$ data for two tetrapeptides with the amino acid sequences gly-X-glygly and glygly-X-gly, and two pentapeptides of sequence glygly-X-glylgly were identical, within the combined experimental uncertainties, with those derived using tripeptides as model compounds. If ionic end-group effects were indeed significant in the tripeptides, then different values for $C_p^o(R)$ would have been obtained using $C_{p,2}^o$ data for tetra- and pentapeptides, since in these peptides each side-chain is at a greater distance from the ionic end-groups. The $C_p^o(H)$ values used in Eq. (2) are taken from the literature [8], although, as shown below, this is of little consequence in additivity calculations. In recent papers [9,10], we derived the partial molar heat capacities over the temperature range 10–100°C for all 20 amino acid side-chains using $C_{p,2}^o$ data for the tripeptides of the sequence gly-X-gly.

The partial molar heat capacity of the backbone peptide group, CHCONH, required for group additivity calculations using Eq. (1) is derived from the partial molar heat capacity of the glycyl group, CH_2CONH, using the equation:

$$C_p^o(\text{CHCONH}) = C_p^o(\text{CH}_2\text{CONH}) - C_p^o(H) \quad (3)$$

where $C_p^o(H)$ is as previously defined for Eq. (2) It is evident from a comparison of Eqs. (2) and (3) that the sum $\{C_p^o(\text{CHCONH}) + C_p^o(R)\}$ is independent of the value chosen for $C_p^o(H)$, and hence is unaffected by any debate there may be over the correct $C_p^o(H)$ values to use for tripeptide model compounds [10]. The partial molar heat capacity of the glycyl group was derived using $C_{p,2}^o$ data over the temperature range 10–100°C for a series of peptides of sequence ala$(\text{gly})_n$, $n = 2$–4 [16]. This procedure is outlined in more detail in Section 4.

The contribution to the heat capacity of a polypeptide of the ionic end-groups NH_3^+ and CH_2COO^- was derived using the equation:

$$C_p^o(NH_3^+ + CH_2COO^-) = C_{p,2}^o(\text{glyglygly}) - 2C_p^o(\text{CH}_2\text{CONH}) \quad (4)$$

where $C_{p,2}^o$(glyglygly) is the partial molar heat capacity of triglycine at infinite dilution. The partial molar heat capacity of the end-groups ($NH_3^+ + CHCOO^-$) is obtained from the quantity $C_p^o(NH_3^+ + CH_2COO^-)$ by subtraction of the heat capacity of the H atom.

Using the estimates of the heat capacities of all 20 amino acid side-chains, the backbone peptide group, and the ionic end-groups, the heat capacity of any random-coil form of a polypeptide or protein of known amino sequence can be obtained at any temperature in the range 10–100°C using Eq. (1).

Essentially the same procedure as outlined for heat capacity may be used to estimate the partial molar volumes of unfolded proteins [13]. The only difference is that the amino acid side-chain par-

tial molar volumes were estimated using the equation:

$$V^0(\mathrm{R}) = V_2^0(\text{gly-X-gly}) - V_2^0(\text{glyglygly}) \quad (5)$$

The quantity $V^0(\mathrm{R})$ gives the contribution to the partial molar volume on replacing the methylene hydrogen atom of the glycyl unit with the side-chain R. These $V^0(\mathrm{R})$ values were then coupled with the partial molar volume for the *glycyl* group $V^0(\mathrm{CH_2CONH})$ in additivity relationships. This approach was used for volumes because comparisons of our side-chain volumes with those obtained using other model systems were not required.

3. HKF group additivity model

The HKF group additivity model, which combines the revised Helgeson–Kirkham–Flowers (HKF) equations of state with group additivity algorithms, has been discussed in considerable detail in the literature [17,18]. However, it is useful to outline briefly the application of the model to partial molar heat capacities and volumes, which are the properties of interest in this paper. The basic premise of the revised HKF equations of state [19] is that any partial molar thermodynamic property of a solute at infinite dilution, Y_2^0 can be represented as the sum of two components, a solvation (or electrostatic) contribution and a non-solvation (or structural) contribution, viz.:

$$Y_2^0 = Y_{2,\mathrm{e}}^0 + Y_{2,\mathrm{s}}^0 \quad (6)$$

The electrostatic contribution, $Y_{2,\mathrm{e}}^0$, is obtained from the Born equation that defines the relationship between the standard molar Gibbs energy of solvation of a species, ΔG_{S}^0, and the dielectric constant of the medium, ε. This relationship is given by the equation:

$$\Delta G_{\mathrm{S}}^{\mathrm{o}} = \omega((1/\varepsilon) - 1) \quad (7)$$

where ω is an effective Born coefficient that is specific to the species of interest. The value of ω for a species is often obtained empirically from correlations of ω with the standard molar entropy of various neutral and charged aqueous species [17,20]. The electrostatic contribution to the partial molar volume at infinite dilution is obtained by differentiating Eq. (7) with respect to pressure at constant temperature:

$$V_{2,\mathrm{e}}^0 = -\omega Q + ((1/\varepsilon) - 1)(\partial\omega/\partial p)_T \quad (8)$$

where the quantity $-Q$ is the derivative of the reciprocal of the dielectric constant of the medium with respect to pressure, at constant temperature. The form of the electrostatic contribution to the partial molar heat capacity at infinite dilution is somewhat more complicated:

$$C_{\mathrm{p,2,e}}^{\mathrm{o}} = \omega T(\partial Y/\partial T)_p + 2TY(\partial\omega/\partial T)_p - T(1/\varepsilon - 1)(\partial^2\omega/\partial T^2)_p \quad (9)$$

In this expression, the Born function $-Y$ is the derivative of the reciprocal of the dielectric constant of the medium with respect to temperature, at constant pressure. For neutral organic species, the somewhat abstruse forms of Eqs. (8) and (9) are overcome by setting the derivatives of ω with respect to temperature and pressure equal to zero [20].

The structural contributions to the partial molar volumes and heat capacities at infinite dilution are represented by the following empirical equations:

$$V_{2,\mathrm{s}}^0 = a_1 + (a_2/(\Psi + p)) + (1/(T - \theta))\{a_3 + (a_4/(\Psi + p))\} \quad (10)$$

$$C_{\mathrm{p,2,s}}^{\mathrm{o}} = c_1 + \left(c_2/(T - \theta)^2\right) - \left\{2T/(T - \theta)^3\right\}[a_3(p - p_{\mathrm{r}}) + a_4 \ln\{(\Psi + p)/(\Psi + p_{\mathrm{r}})\}] \quad (11)$$

In Eqs. (10) and (11), a_1, a_2, a_3, a_4, c_1 and c_2 are solute-dependent coefficients and θ and Ψ are solvent-dependent parameters which, for water,

are assigned values of $\theta = 228$ K and $\Psi = 2600$ bar [17]. The term p_r defines a reference pressure that is set equal to 1 bar. Eq. (10) is often presented in a simplified form as:

$$V^0_{2,s} = \sigma + (\xi/(T-\theta)) \quad (12)$$

Similarly, at $p = p_r$, which is the focus in this study, Eq. (11) reduces to:

$$C^o_{p,2,s} = c_1 + \left(c_2/(T-\theta)^2\right) \quad (13)$$

The final expressions for the partial molar volume and heat capacity for a solute at infinite dilution are then obtained by combining the relevant equations given above to obtain:

$$V^0_2 = \sigma + (\xi/(T-\theta)) - \omega Q \quad (14)$$

and

$$C^o_{p,2} = c_1 + \left(c_2/(T-\theta)^2\right) + \omega T \mathrm{X} \quad (15)$$

In Eq. (15), the symbol X represents the quantity $(\partial Y/\partial T)_p$. Values for the various equation-of-state parameters a_i and c_i are generally obtained [18] from regression analyses using partial molar volume, compressibility and heat capacity data taken from the literature. If $C^o_{p,2}$ data as a function of temperature are unavailable, then the c_1 and c_2 coefficients are evaluated using group additivity methods [14,18].

In the specific model proposed by Amend and Helgeson [14], regression analyses of literature data for the peptides $(gly)_m$, $m = 3$–5, were used to obtain the coefficients of Eqs. (14) and (15). Equation-of-state parameters for the glycyl group were then obtained from correlations of the coefficients with the number of glycyl groups. The coefficients for all the amino acid side-chains were calculated from those of the parent zwitterionic amino acids, reported in earlier work [18], by subtraction of the corresponding coefficients for a group defined as the amino acid backbone (AAB), which has the structure $CHNH_2COOH$. The coefficients for AAB were in turn derived from analyses of the partial molar properties of the AAB group obtained using the equation [18]:

$$Y^0_2(\mathrm{AAB}) = Y^0_2(\mathrm{serine}) - Y^0_2(\mathrm{CH_2OH}) \quad (16)$$

where Y^0_2(serine) and $Y^0_2(CH_2OH)$ are the partial molar properties of the species in parentheses. Equation-of-state parameters for the peptide backbone group (PB), CHCONH, were obtained using literature heat capacity [21] and volumetric data [22] for four unfolded proteins. Firstly, the partial molar heat capacities and volumes of PB were calculated from the experimental protein data using group additivity methods. The contributions of the various protein constituent groups were calculated using Eqs. (14) and (15) and generated coefficients [14]. Having obtained the group partial molar properties for PB, the various coefficients were then obtained by regression analyses [14].

4. Glycyl group contributions

A perusal of the literature results for the heat capacity of the glycyl group at 25°C reported previously [16] indicated considerable disparity amongst the various estimates. Furthermore, in the only study to determine the temperature dependence of $C^o_p(CH_2CONH)$ [8], data were available at only five temperatures in the range 5–100°C. As a result of this survey, we carried out a careful and detailed study [16] to determine reliable values of the partial molar volumes and heat capacities of the glycyl group over a wide temperature range. Amend and Helgeson [16] chose not to use the results obtained in our study and instead, repeated analyses of literature data that had essentially been carried out previously [8]. In view of this oversight, a careful examination of the results reported by Amend and Helgeson is pertinent.

The most reliable of the earlier determinations of $C^o_p(CH_2CONH)$ were obtained using $C^o_{p,2}$ data for the oligoglycines, $(gly)_m$, $m = 1$–5. Plots of $C^o_{p,2}$ against the number of glycyl groups are linear in the range $m = 3$–5, so estimates of

$C_p^o(CH_2CONH)$ are obtained from the slopes of these plots. The main difficulty with this procedure is that the oligoglycines tetra- and pentaglycine are only sparingly soluble in pure water. This means that the $C_{p,2}^o$ values obtained for these compounds have large uncertainties, which in turn leads to a high uncertainty in the estimate for $C_p^o(CH_2CONH)$. To overcome these difficulties, we used an alternative series of peptides, ala(gly)$_n$, $n = 2–4$, the higher members of which are more soluble in water than the corresponding oligoglycines. Values for $C_p^o(CH_2CONH)$ were also estimated [16] using $C_{p,2}^o$ data for the two neutral amino acid and peptide derivatives *N*-acetylglycinamide ($CH_3CONHCH_2CONH_2$, AcglyNH$_2$) and *N*-acetylglycylgylcinamide ($CH_3CONHCH_2CONHCH_2CONH_2$, AcglyglyNH$_2$). The difference between the $C_{p,2}^o$ values for these two compounds is simply $C_p^o(CH_2CONH)$. Plots of $C_p^o(CH_2CONH)$ against temperature for these two sets of model compounds are shown in Fig. 1. The agreement between the temperature dependence of $C_p^o(CH_2CONH)$ derived using the two sets of model compounds is within the combined estimated uncertainties. Over the temperature range used, $C_p^o(CH_2CONH)$ increases with increasing temperature. In contrast, the results calculated using the c_1, c_2 and ω parameters reported by Amend and Helgeson [14] have a significantly different temperature dependence, as shown in Fig. 1. With increasing temperature, the values of $C_p^o(CH_2CONH)$ initially decrease and become approximately constant in the range 80–100°C. Included in Fig. 1 are the results reported by Makhatadze and Privalov, which were obtained using $C_{p,2}^o$ data for the (gly)$_m$ series of peptides [8]. It is interesting to note the different temperature profile for these results, given that these same $C_{p,2}^o$ data were part of the data set used by Amend and Helgeson [14] in regression analyses to generate their c_1 and c_2 coefficients for the glycyl group.

In our earlier paper [16], the partial molar volumes of the glycyl group, $V^0(CH_2CONH)$, over the temperature range 10–80°C were derived from V_2^0 data for the ala(gly)$_n$ peptides and for the compounds AcglyNH$_2$ and AcglyglyNH$_2$. As the

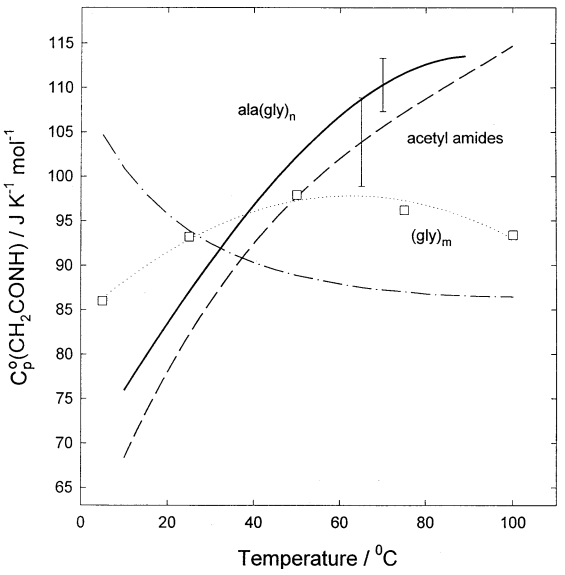

Fig. 1. Temperature dependence of the partial molar heat capacity of the glycyl group: — derived from $C_{p,2}^o$ data for the ala(gly)$_n$ peptides [16]; – – – derived from $C_{p,2}^o$ data for AcglyNH$_2$ and AcglyglyNH$_2$ [16]; □ data from [8]; and – · – · – calculated using Eq. (15) and the parameters from [14].

solution densities were measured using a differential scanning densimetric method (DSD) [23], the uncertainties for the $V^0(CH_2CONH)$ values are high. The results obtained, along with the uncertainties, are shown in Fig. 2. In a recent study [24], more precise partial molar volume data have been determined for the peptides ala(gly)$_n$, $n = 2–4$, at the four temperatures 18, 25, 30 and 40°C. The $V^0(CH_2CONH)$ values derived from these results are also given in Fig. 2. The agreement among the sets of $V^0(CH_2CONH)$ results is within the estimated uncertainties.

As shown in Fig. 2, the temperature dependence of the $V^0(CH_2CONH)$ results reported by Amend and Helgeson [14] is opposite in sign to that derived in our own work. This trend is similar to that which we derived earlier [16] using V_2^0 data for the (gly)$_m$ peptides reported by Makhatadze et al. [22]. It should be noted, however, that if the V_2^0 values for the (gly)$_m$ peptides determined by Chalikian and co-workers [25] had been used in the regression analyses carried out

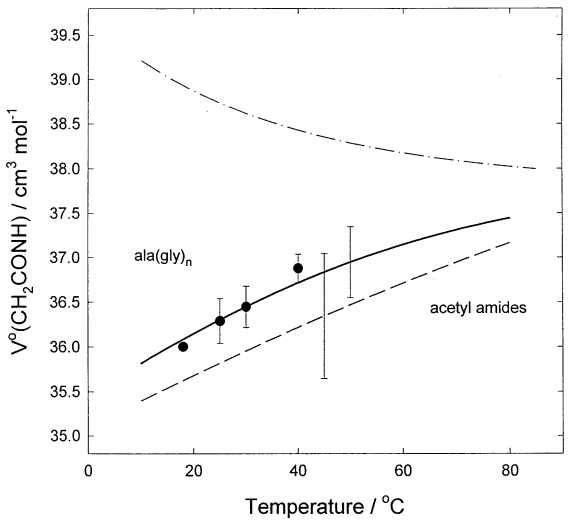

Fig. 2. Temperature dependence of the partial molar volume of the glycyl group. — derived from V_2^0 data for the ala(gly)$_n$ peptides [16]; ● from V_2^0 data for the ala(gly)$_n$ peptides [24]; – – – derived from V_2^0 data for AcglyNH$_2$ and AcglyglyNH$_2$ [16]; and – · – · – calculated using Eq. (14) and the parameters from [14].

by Amend and Helgeson, then the temperature dependence of $V^0(CH_2CONH)$ would have been large and positive [16]. The disparity among the results that can be derived from V_2^0 data for the (gly)$_m$ peptides is an illustration of how difficult it is to obtain reliable results for compounds which are sparingly soluble and also slow to dissolve. For this reason, the earlier recommendation [16] that the partial molar heat capacities and volumes of the glycyl group derived from thermodynamic data for the ala(gly)$_n$ series of peptides better represent the glycyl group contributions in a polypeptide is still valid.

5. Partial molar heat capacities of oligopeptides

In the paper by Amend and Helgeson [14], a summary is given of the parameters for the equations of state for all the amino acid side-chains, the peptide backbone unit (PB), and the unit referred to as the amino acid backbone (AAB). This latter group was used by Amend and Helgeson to model the ionic end-groups of a protein. In principle, these parameters can be used to generate the heat capacity contributions for all of the constituent groups of any fully unfolded polypeptide. Although Amend and Helgeson have used their additivity scheme to evaluate only the thermodynamic properties of some unfolded thermophilic proteins, if their scheme has any general predictive utility, then it should also be applicable to small polypeptides. After all, these smaller molecules have exactly the same types of functional groups exposed to the solvent as for any completely unfolded protein. It is of interest therefore, to see how well these equation-of-state parameters can predict the partial molar heat capacities of some small oligopeptides.

In recent work [26], we synthesized some tetra- and pentapeptides which have different combinations of neutral side-chains. The partial molar heat capacities of these peptides have been determined experimentally over the temperature range 10–100°C [15]. The specific purpose of this work was to test the applicability of our peptide-based group additivity model. The Amend and Helgeson group additivity model can also be tested using these compounds.

A comparison is given in Fig. 3 of the partial molar heat capacities of tetraglycine obtained experimentally by DSC [9] with those calculated using both the peptide-based group additivity scheme [9] and that proposed by Amend and Helgeson [14]. The latter calculation of $C^o_{p,2}$(calc) was carried out using the equation:

$$C^o_{p,2}(\text{calc}) = C^o_p(\text{AAB}) + C^o_p(\text{-Gly-}) + 2C^o_p(\text{PB}) + 3R_{gly} \quad (17)$$

where, in the nomenclature used by Amend and Helgeson [14], C^o_p(AAB) is the heat capacity of the end-groups, C^o_p(-Gly-) is the glycyl-group heat capacity, C^o_p(PB) is the heat capacity of the peptide backbone group, and R_{gly} is the heat capacity of the side-chain for glycine, which is just the H atom. The temperature dependence of the partial molar heat capacity at a pressure of 1 bar for each of the terms in Eq. (17) was calculated

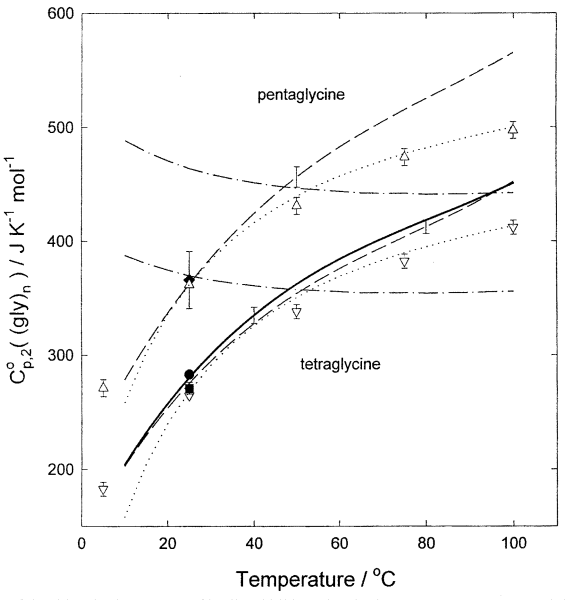

Fig. 3. Temperature dependence of the partial molar heat capacities of tetra- and pentaglycine. $C^o_{p,2}$ data for tetraglycine: — from [9]; ● from [27]; □ from [16]; and ∇ from [8]. $C^o_{p,2}$ data for pentaglycine: ♦ from [27]; and Δ from [8]. Calculated partial molar heat capacities: – – – using peptide-based group additivity [9]; – · – · – using HKF group additivity, Eq. (17); and ..... using Eq. (15) and parameters from [14].

using the equation-of-state parameters given by Amend and Helgeson [14]. For the peptide-based model, the heat capacity of tetraglycine is calculated by adding the heat capacity of the glycyl group [16] to the experimental data for triglycine [10]. This is, in effect, exactly the same as using Eqs. (1)–(4), as outlined in Section 2.

As shown in Fig. 3, the agreement between the experimental data (represented by the full line) and that calculated using peptide-based group additivity (the dashed line) is excellent over the complete temperature range. For the purposes of comparison, other $C^o_{p,2}$ data for tetraglycine taken from the literature [8,16,27] are also included in Fig. 3. At 25°C, all the literature data are concordant, but at higher temperatures the results of Makhatadze and Privalov [8] differ from those we obtained [9]. It is clear from Fig. 3 that the $C^o_{p,2}$(calc)–temperature curve (the dash–dot line) predicted using the Amend and Helgeson group additivity model gives an extremely poor representation of the experimental DSC data. The curve for tetraglycine (the dotted line) derived using the c_1, c_2 and ω parameters reported for tetraglycine [14] follows closely the $C^o_{p,2}$ data determined by Makhatadze and Privalov [8], which is to be expected, as these results were actually used to generate the coefficients. It would appear that some of the coefficients reported for the various constituent groups used in Eq. (6) give a very poor representation of these groups in a small peptide such as tetraglycine.

The only $C^o_{p,2}$ data for pentaglycine determined over a temperature range that have been reported in the literature are those given by Makhatadze and Privalov [8]. As shown in Fig. 3, there is excellent agreement between the experimental data at 25°C [8,27] and those predicted using peptide-based group additivity, but at higher temperatures the experimental data are lower that those predicted. Again, the heat capa-

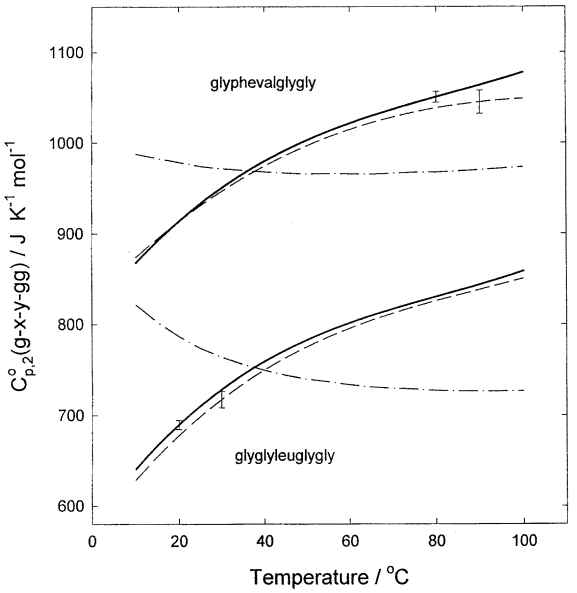

Fig. 4. Temperature dependence of the partial molar heat capacity of glyphevalglygly and glyglyleuglygly: — experimental $C^o_{p,2}$ data [15]; – – – calculated using peptide-based group additivity [9]; and – · – · – calculated using HKF group additivity [14].

city–temperature curve predicted using the Amend and Helgeson parameters for the various constituent groups of pentaglycine does not resemble the experimental data.

Fig. 4 displays the partial molar heat capacities for the two peptides glyglyleuglygly and glyphevalglygly. The experimental data were obtained over the temperature range 10–100°C using DSC [15]. The predicted $C^{o}_{p,2}$(T) curves obtained using the peptide-based additivity model that are also included in Fig. 4 are in excellent agreement with the experimental results. In contrast, the results calculated using the Amend and Helgeson [14] group parameters have temperature dependences that are the inverse of those observed experimentally.

In Fig. 5, the experimental partial molar heat capacities at infinite dilution are shown for a tetrapeptide with two hydrophilic side-chains and for a pentapeptide with one hydrophilic side-chain adjacent to a hydrophobic side-chain. For each peptide, there is excellent agreement over the entire temperature range studied between the experimental $C^{o}_{p,2}$ data and those calculated using the peptide-based group additivity scheme. For these particular peptides, the Amend and Helgeson additivity scheme predicts a temperature dependence which has the same shape as that observed experimentally. For glyserthrgly, the quantitative agreement with the experimental results is quite reasonable across the temperature range from 10 to approximately 60°C.

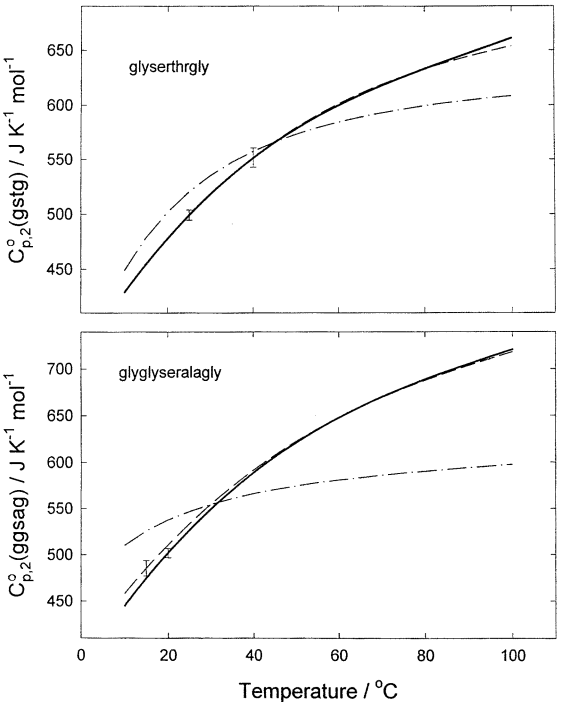

Fig. 5. Temperature dependence of the partial molar heat capacity of glyserthrgly and glyglyseralagly: — experimental $C^{o}_{p,2}$ data [15]; – – – calculated using peptide-based group additivity [9]; and – · – · – calculated using HKF group additivity [14].

6. Partial molar volumes of oligopeptides

The equation-of-state parameters for the various constituent groups of a polypeptide given by Amend and Helgeson [14] can be used to calculate the partial molar volumes of some tetra- and pentapeptides for which experimental data are available over the temperature range 10–90°C [26]. The experimental V^{0}_{2} results for three of these peptides obtained using solution densities measured by DSD, and also using more precise densities determined isothermally at 25°C, are displayed in Fig. 6. For the peptides glyglyleuglygly and glyasnalagly, the agreement between the experimental results and those calculated using peptide-based group additivity [13], viz.:

$$V^{0}_{2}(\text{calc}) = V^{0}_{2}((\text{gly})_3) + \Sigma V^{0}(\text{CH}_2\text{CONH}) + \Sigma V^{0}(\text{R}) \quad (18)$$

is excellent over the complete temperature range studied. Although for glyphealagly there is some deviation between the experimental and predicted curves at high temperature, the agreement is within the combined estimated uncertainties from 10 to a temperature of approximately 80°C.

It is clear from Fig. 6 that the V^{0}_{2}(calc) results predicted using the Amend and Helgeson additivity scheme are in poor agreement with the experimental results, especially for the pentapeptide. For glyglyleuglygly, the differences between the experimental V^{0}_{2} results and V^{0}_{2}(calc) are as large

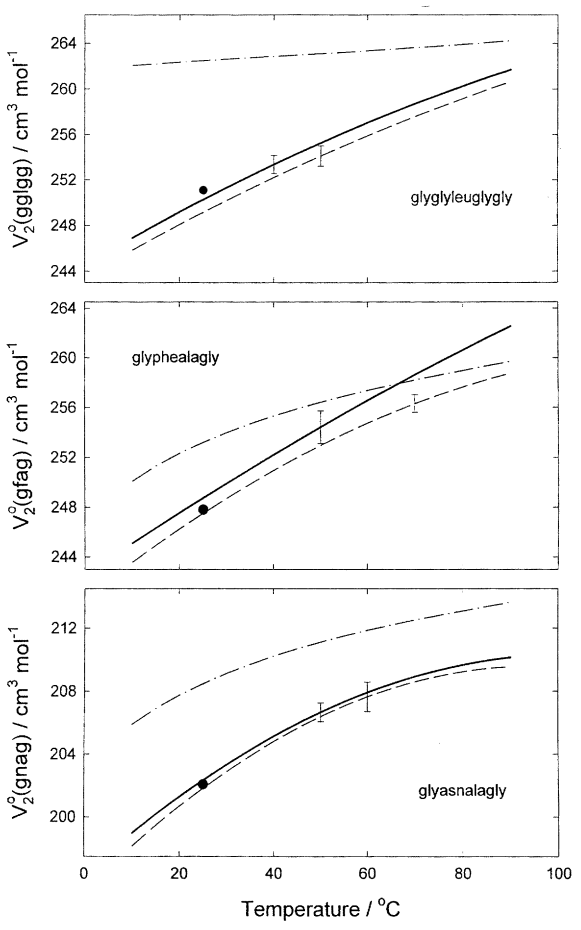

Fig. 6. Temperature dependence of the partial molar volume of the peptides glyglyleuglygly, glyphealagly, and glyasnalagly: — experimental V_2^0 data [26]; – – – calculated using peptide-based group additivity [13]; and – · – · – calculated using HKF group additivity [14].

as 14 cm^3 mol^{-1} at low temperature. The V_2^0(calc)–temperature curve for the peptide glyasnalagly has approximately the same shape as that obtained by experiment, but is displaced to higher volumes by approximately 4–6 cm^3 mol^{-1}.

In earlier work [26], the partial molar volumes at infinite dilution were determined at 25°C for nine peptides which have either one or two side-chains. The differences between the experimental results and those calculated using Eq. (18) were in the range 0.5–1.3 cm^3 mol^{-1} (0.2–0.6% of V_2^0). In contrast, the V_2^0(calc) values for the same peptides calculated using the Amend and Helgeson group additivity scheme [14] differ from the experimental values by 5–14 cm^3 mol^{-1}, i.e. 10-fold larger than those obtained using the peptide-based group additivity scheme.

7. Concluding remarks

There are several factors that contribute to the poor predictive utility of the Amend and Helgeson additivity scheme [14]. Any model that utilizes semi-empirical equations-of-state that are generated using thermodynamic data from the literature is only as reliable as the input data sets used in its construction. In the model proposed by Amend and Helgeson, some rather unreliable data have been utilized, for example those used to generate the parameters for the glycyl group, as discussed in Section 4. Moreover, more reliable c_1 and c_2 equation-of-state parameters would have been attained by using new $C_{p,2}^o$ data for almost all of the amino acid acids over the temperature range 15–55°C that have been reported in several recent papers [28–32].

It is clear from the results presented above in Sections 4–6, that if reliable predictions of the partial molar heat capacities and volumes of unfolded polypeptides in aqueous solution are required, then it would be ill-advised to use the current equation-of-state parameters for the various constituent groups of proteins reported by Amend and Helgeson [14]. On the other hand, the results presented in Sections 5 and 6, and elsewhere [15,26], show that the peptide-based additivity model predicts the partial molar heat capacities and volumes of polypeptides rather successfully, at least for the side-chain types that have been tested to date.

Acknowledgements

AWH gratefully acknowledges the receipt of an operating grant from the Natural Sciences and Engineering Research Council of Canada (NSERC).

References

[1] D. Shortle, The denatured state (the other half of the folding equation) and its role in protein stability, FASEB J. 10 (1996) 27–34.

[2] L.J. Smith, K.M. Fiebig, H. Schwalbe, C.M. Dobson, The concept of a random coil-residual structure in peptides and denatured proteins, Fold. Design 1 (1996) R95–R106.

[3] P.L. Privalov, Stability of small globular proteins, Adv. Protein Chem. 33 (1979) 167–237.

[4] V.P. Denisov, B.-H. Jonsson, B. Halle, Hydration of denatured and molten globule proteins, Nature Struct. Biol. 6 (1999) 253–260.

[5] D.P. Kharakoz, Partial volumes and compressibilities of extended polypeptide chains in aqueous solution — additivity scheme and implication of protein unfolding at normal and high pressure, Biochemistry 36 (1997) 10276–10285.

[6] M. Iqbal, R.E. Verrall, Implications of protein folding. Additivity schemes for volumes and compressibilities, J. Biol. Chem. 263 (1988) 4159–4165.

[7] A.A. Zamyatnin, Amino acid, peptide, and protein volume in solution, Annu. Rev. Biophys. Bioeng. 13 (1984) 145–165.

[8] 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.

[9] M. Häckel, H.-J. Hinz, G.R. Hedwig, A new set of peptide-based group heat capacities for use in protein stability calculations, J. Mol. Biol. 291 (1999) 197–213.

[10] M. Häckel, H.-J. Hinz, G.R. Hedwig, Tripeptides in aqueous solution: model compounds for the evaluation of the partial molar heat capacities of amino acid side-chains in proteins, Thermochim. Acta 308 (1998) 23–34.

[11] J.F. Reading, G.R. Hedwig, Thermodynamic properties of peptide solutions. Part 6. The amino acid side-chain contributions to the partial molar volumes and heat capacities of some tripeptides in aqueous solution, J. Chem. Soc. Faraday Trans. 86 (1990) 3117–3123.

[12] G.R. Hedwig, Partial molar heat capacities and volumes of some peptides that model side-chains of proteins, J. Chem. Soc. Faraday Trans. 89 (1993) 2761–2768.

[13] M. Häckel, H.-J. Hinz, G.R. Hedwig, Partial molar volumes of proteins: amino acid side-chain contributions derived from the partial molar volumes of some tripeptides over the temperature range 10–90°C, Biophys. Chem. 82 (1999) 35–50.

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

[15] M. Häckel, H.-J. Hinz, G.R. Hedwig, Additivity of the partial molar heat capacities of the amino acid side-chains of small peptides: Implications for unfolded proteins, Phys. Chem. Chem. Phys. 2 (2000) 5463–5468.

[16] M. Häckel, G.R. Hedwig, H.-J. Hinz, The partial molar heat capacity and volume of the peptide backbone group of proteins in aqueous solution, Biophys. Chem. 73 (1998) 163–177.

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

[18] 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.

[19] 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.

[20] 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.

[21] P.L. Privalov, E.I. Tiktopulo, S.Y. Venyaminov, Y.V. Griko, G.I. Makhatadze, N.N. Khechinashvili, J. Mol. Biol. 205 (1989) 737–750.

[22] 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.

[23] H.-J. Hinz, T. Vogl, R. Meyer, An alternative interpretation of the heat capacity changes associated with protein unfolding, Biophys. Chem. 52 (1994) 275–285.

[24] A.W. Hakin, H. Høiland, G.R. Hedwig, Volumetric properties of some oligopeptides in aqueous solution: partial molar expansibilities and isothermal compressibilities at 298.15 K for the peptides of sequence $Ala(gly)_n$, $n = 1–4$, Phys. Chem. Chem. Phys. 2 (2000) 4850–4857.

[25] T.V. Chalikian, A.P. Sarvazyan, T. Funck, K.J. Breslauer, Partial molar volumes, expansibilities, and compressibilities of oligoglycines in aqueous solutions at 18–55°C, Biopolymers 34 (1994) 541–553.

[26] M. Häckel, H.-J. Hinz, G.R. Hedwig, The partial molar volumes of some tetra- and pentapeptides in aqueous solution: a test of amino acid side-chain group additivity for unfolded proteins, Phys. Chem. Chem. Phys. 2 (2000) 4843–4849.

[27] 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.

[28] A.W. Hakin, M.M. Duke, S.A. Klassen, R.M. McKay, K.E. Preuss, Apparent molar heat capacities and volumes of some aqueous solutions of aliphatic amino acids at 288.15, 298.15, 313.15, and 328.15 K, Can. J. Chem. 72 (1994) 362–368.

[29] M.M. Duke, A.W. Hakin, R.M. McKay, K.E. Preuss, The volumetric and thermochemical properties of aqueous solutions of L-valine, L-leucine, and L-isoleucine at 288.15, 298.15, 313.15, and 328.15 K, Can. J. Chem. 72 (1994) 1489–1494.

[30] A.W. Hakin, M.M. Duke, J.L. Marty, K.E. Preuss, Some thermodynamic properties of aqueous amino acid systems at 288.15, 298.15, 313.15, and 328.15 K, J. Chem. Soc. Faraday Trans. 90 (1994) 2027–2035.

[31] A.W. Hakin, M.M. Duke, L.L. Groft, J.L. Marty, M.L. Rushfeldt, Calorimetric investigations of aqueous amino acid and dipeptide systems from 288.15 to 328.15 K, Can. J. Chem. 73 (1995) 725–734.

[32] 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) 84–89.