J. Mol. Biol. (1996) **260**, 588–603

JMB

The Hydration of Globular Proteins as Derived from Volume and Compressibility Measurements: Cross Correlating Thermodynamic and Structural Data

Tigran V. Chalikian[1], Maxim Totrov[2], Ruben Abagyan[2] and Kenneth J. Breslauer[1]*

[1]*Department of Chemistry Rutgers, The State University of New Jersey, Piscataway NJ 08855, USA*

[2]*The Skirball Institute of Biomolecular Medicine New York University Medical Center, 540 1st Ave New York, NY 10016, USA*

We report the first thermodynamic characterization of protein hydration that does not depend on model compound data but rather is based exclusively on macroscopic (volumetric) and microscopic (X-ray) measurements on protein molecules themselves. By combining these macroscopic and microscopic characterizations, we describe a quantitative model that allows one for the first time to predict the partial specific volumes, v°, and the partial specific adiabatic compressibilities, k_S°, of globular proteins from the crystallographic coordinates of the constituent atoms, without using data derived from studies on low-molecular-mass model compounds. Specifically, we have used acoustic and densimetric techniques to determine v° and k_S° for 15 globular proteins over a temperature range from 18 to 55°C. For the subset of the 12 proteins with known three-dimensional structures, we calculated the molecular volumes as well as the solvent-accessible surface areas of the constituent charged, polar and nonpolar atomic groups. By combining these measured and calculated properties and applying linear regression analysis, we determined, as a function of temperature, the average hydration contributions to v° and k_S° of 1 Å^2 of the charged, polar, and nonpolar solvent-accessible protein surfaces. We compared these results with those derived from studies on low-molecular-mass compounds to assess the validity of existing models of protein hydration based on small molecule data. This comparison revealed the following features: the hydration contributions to v° and k_S° of charged protein surface groups are similar to those of charged groups in small organic molecules. By contrast, the hydration contributions to v° and k_S° of polar protein surface groups are qualitatively different from those of polar groups in low-molecular-mass compounds. We suggest that this disparity may reflect the presence of networks of water molecules adjacent to polar protein surface areas, with these networks involving waters from second and third coordination spheres. For nonpolar protein surface groups, we find the ability of low-molecular-mass compounds to model successfully protein properties depends on the temperature domain being examined. Specifically, at room temperatures and below, the hydration contribution to k_S° of protein nonpolar surface atomic groups is close to that of nonpolar groups in small organic molecules. By contrast, at higher temperatures, the hydration contribution to k_S° of protein nonpolar surface groups becomes more negative than that of nonpolar groups in small organic molecules. We suggest that this behaviour may reflect nonpolar groups on protein surfaces being hydrated independently at low temperatures, while at higher temperatures some of the solvating waters become influenced by neighboring polar groups. We discuss the implications of our aggregate results in terms of various approaches currently being used to describe the hydration properties of globular proteins, particularly focusing on the limitations of existing additive models based on small molecule data.

Keywords: globular proteins; structure; hydration; volume; adiabatic compressibility

**Corresponding author*

0022–2836/96/290588–16 $18.00/0

Introduction

The partial specific volume, $v°$, and the partial specific adiabatic compressibility, k_S°, of a protein are macroscopic observables which are particularly sensitive to the hydration properties of solvent-exposed atomic groups, as well as to the structure, dynamics, and conformational properties of the solvent inaccessible protein interior (Gavish *et al*., 1983; Zamyatnin, 1984; Gekko & Hasegawa, 1986; Kharakoz & Sarvazyan, 1993; Chalikian *et al*., 1994a). Recent biochemical and instrumental advances have permitted these observables to be determined for a wide range of protein systems. These data have yielded unique insights into the thermodynamic properties of proteins, as well as protein hydration and conformational transitions (Chalikian *et al*., 1995, 1996; Nölting & Sligar, 1993; Ybe & Kahn, 1994; Foygel *et al*., 1995; Tamura & Gekko, 1995; Chalikian & Breslauer, 1996). A quantitative understanding of the data obtained from such studies, however, requires knowledge of the contributions that specific protein states and components make to the overall measured values of $v°$ and k_S°. In this regard, a characteristic feature of native globular proteins is that, at room temperatures, their partial specific adiabatic compressibilities, k_S°, range from -1×10^{-6} to 10×10^{-6} $cm^3 g^{-1}$ bar^{-1} (Gavish *et al*., 1993; Gekko & Hasegawa, 1986; Kharakoz & Sarvazyan, 1993; Chalikian & Breslauer, 1996; Gekko & Noguchi, 1979; Gekko & Hasegawa, 1989), while their partial specific volumes, $v°$, only range from 0.70 to 0.75 cm^3 g^{-1} (Zamyatnin, 1984; Gekko & Hasegawa, 1986; Kharakoz & Sarvazyan, 1993). Thus, k_S° offers a more sensitive observable than $v°$ for distinguishing between native state protein properties. In recognition of this fact, there have been various attempts to relate k_S° of globular proteins to the total hydrophobicity of the protein (Gekko & Hasegawa, 1986; Gekko & Noguchi, 1979; Iqball & Verrall, 1988), to the fractions of different polar and nonpolar residues (Gekko & Hasegawa, 1986; Gekko & Noguchi, 1979), to the α-helix content (Gekko & Hasegawa, 1986), to differences between the partial specific volume, $v°$, and the sum of the van der Waals volumes of the constituent amino acid residues (Gekko & Noguchi, 1979), and even to a collection of the 27 diverse physico-chemical properties of the free amino acids (Gromiha & Ponnuswamy, 1993). In such efforts, a number of investigators have emphasized that the partial specific adiabatic compressibility, k_S°, of a globular protein can be considered to be the sum of two oppositely contributing major terms: (1) a positive ''intrinsic compressibility'' term, k_M, which arises from the imperfect packing of the amino acid residues within the solvent-inaccessible core of the protein; and (2) a negative ''hydration compressibility'', Δk_h, which represents a decrease in the solvent compressibility due to the interactions of water molecules with the solvent-exposed atomic groups of the protein (Gavish *et al*., 1983; Gekko & Hasegawa, 1986; Kharakoz & Sarvazyan, 1993; Chalikian *et al*., 1994a; Gekko & Noguchi, 1979). We judge the approach of Kharakoz & Sarvazyan (1993) to be the most elaborate effort to date for discriminating between the intrinsic, k_M, and hydration contributions, Δk_h, to k_S°. Their approach is based on additive calculations of Δk_h from the number of solvent accessible polar, nonpolar, and charged atomic groups on the protein surface as determined from X-ray studies and the compressibility contributions of each of these classes of groups as derived from solution studies on low-molecular-mass model compounds such as amino acids (Kharakoz & Sarvazyan, 1993). However, their ''counting'' procedure does not account for partially buried groups and, as just noted, they assume that the compressibility contribution of an atomic group on a protein surface can be modeled by the contribution of the same group within a low-molecular-mass compound.

To circumvent these potential limitations, we describe here a new approach for defining the hydration contributions to protein compressibility and volume which does not rely on low-molecular-mass model compound data and which addresses the counting problem by considering the total accessible surface areas of specific classes of solvent exposed groups, rather than simply the number of the solvent exposed atomic groups. In this approach, we determine the compressibility and/or volume contributions per 1 $Å^2$ of each type of solvent accessible surface (polar, nonpolar, charged) by a linear regression analysis of $v°$, k_S°, and X-ray data derived exclusively on the protein systems themselves. Specifically, we use acoustic and densimetric methods to determine k_S° and $v°$ for the native states of 15 globular proteins at 18, 25, 35, 45, and 55°C, and crystallographic data on 12 of these proteins to calculate the accessible surface areas of polar, nonpolar, and charged groups in each protein. These two data sets allow us to define the average contribution per 1 $Å^2$ of solvent accessible polar, nonpolar, and charged protein groups to the experimental $v°$ and k_S° values. Comparison of these results with those obtained from studies on low molecular weight compounds allows us to assess the limitations of using small molecule data to model the properties of chemically similar components within proteins.

In short, the experimental data and the analytical approach presented and described in this work allows one to interpret protein volume and compressibility data in terms of protein properties, particularly, the hydration of different protein domains, without relying on model compound data. Our results also permit us to define useful general relationships between the molecular mass of a protein, its intrinsic volume, and its total solvent accessible surface area.

Abbreviation used: BSA, bovine serum albumin.

Table 1. Protein molecular masses and relative specific sound velocity increments, $[u]$ ($cm^3\ g^{-1}$), as a function of temperature

Protein	M (kDa)	18°C	25°C	35°C	45°C	55°C
1. Conalbumin	75.5	0.186	0.176	0.163	0.152	0.143
2. BSA	68.0	0.180	0.164	0.144	0.132	0.129
3. Hemoglobin	68.0	0.183	0.168	0.158	0.147	
4. Ovalbumin	46.0	0.181	0.169	0.162	0.154	0.147
5. Pepsin	35.5	0.187	0.173	0.154	0.147	0.139
6. α-Chymotrypsinogen A	25.7	0.202	0.193	0.184		
7. α-Chymotrypsin	25.3	0.210	0.188	0.182	0.172	
8. Trypsin	23.0	0.223	0.199	0.186	0.180	0.175
9. Trypsinogen	23.0	0.190	0.179	0.159	0.156	0.144
10. β-Lactoglobulin	18.4	0.200	0.186	0.167	0.150	
11. Myoglobin	17.8	0.180	0.164	0.147	0.142	0.138
12. α-Lactalbumin	14.3	0.189	0.181	0.155	0.150	
13. Lysozyme	14.3	0.207	0.186	0.172	0.164	0.157
14. Ribonuclease A	13.6	0.202	0.193	0.179	0.173	
15. Cytochrome *c*	12.4	0.215	0.202	0.187	0.173	0.167

Estimated average error of measurements is (± 0.003) $cm^3\ g^{-1}$.

Results

Volume and compressibility data

Tables 1, 2, and 3 list, along with the relevant molecular masses, the relative specific sound velocity increments, $[u]$ (Table 1), the apparent specific volumes, φV (Table 2), and the apparent specific adiabatic compressibilities, φK_S (Table 3), we have measured and/or calculated for the indicated 15 globular proteins in their native states at 18, 25, 35, 45, and 55°C. For such globular proteins, it previously has been shown that the concentration dependences of φV and φK_S are negligible at protein concentrations between 0 and 5 mg/ml (Gekko & Hasegawa, 1986; Gekko & Noguchi, 1979). Consequently, within the limits of the experimental uncertainty, the data determined here at protein concentrations of 3 mg/ml correspond to partial $v°$ and $k_S°$ values obtained by extrapolation to infinite dilution. For this reason, throughout this article, we will not discriminate between partial and apparent values.

Table 4 compares our data on $v°$ and $k_S°$ with literature values at 25°C. This somewhat restrictive comparison is due to the fact that most previous studies only determined volume and compressibility data at this single temperature. Inspection of these data reveals that, for most of the proteins studied, reasonably good agreement is found between our $v°$ and/or $k_S°$ data and the corresponding literature values. Note, however, there are some proteins, for which the agreement is poor. At present, we have no explanation for these few discrepancies.

Temperature dependences of $[u]$, $v°$, and $k_S°$

Inspection of the data in Tables 1, 2, and 3 reveals for all 15 globular proteins that with increasing temperature the values of $[u]$ decrease while the values of $v°$ and $k_S°$ increase. All the temperature dependencies of the relative specific sound velocity increments, $[u]$, can be fit by second order polynomial functions. The average temperature slope, $\Delta[u]/\Delta T$, that results from such fits increases

Table 2. Protein apparent specific volumes, φV ($cm^3\ g^{-1}$), as a function of temperature

Proteins	M (kDa)	18°C	25°C	35°C	45°C	55°C
1. Conalbumin	75.5	0.726	0.729	0.733	0.735	0.739[a]
2. BSA	68.0	0.735	0.739	0.744	0.751	0.756[a]
3. Hemoglobin	68.0	0.743	0.745	0.747	0.750	
4. Ovalbumin	46.0	0.735	0.737	0.740	0.743	0.746[a]
5. Pepsin	35.5	0.730	0.733	0.736	0.738	0.741
6. α-Chymotrypsinogen A	25.7	0.727	0.730	0.733		
7. α-Chymotrypsin	25.3	0.717	0.721	0.724	0.727	
8. Trypsin	23.0	0.718	0.720	0.724	0.727	0.731[a]
9. Trypsinogen	23.0	0.721	0.725	0.730	0.734	0.739
10. β-Lactoglobulin	18.4	0.731	0.734	0.737	0.741	
11. Myoglobin	17.8	0.742	0.745	0.748	0.750	0.753[a]
12. α-Lactalbumin	14.3	0.711	0.713	0.717	0.720	
13. Lysozyme	14.3	0.699	0.702	0.704	0.707	0.710
14. Ribonuclease A	13.6	0.702	0.704	0.707	0.710	
15. Cytochrome *c*	12.4	0.735	0.738	0.742	0.746	0.750

Estimated average error of measurements is (± 0.003) $cm^3\ g^{-1}$.
[a] Extrapolated values.

Table 3. Protein apparent specific adiabatic compressibilities, k_S° (10^{-6} cm^3 g^{-1} bar^{-1}), as a function of temperature

Protein	18°C	25°C	35°C	45°C	55°C
1. Conalbumin	3.6 ± 0.5	4.6 ± 0.5	5.8 ± 0.5	6.7 ± 0.5	7.5 ± 0.5
2. BSA	5.0 ± 0.5	6.6 ± 0.5	8.4 ± 0.5	9.8 ± 0.5	10.1 ± 0.5
3. Hemoglobin	5.5 ± 0.5	6.8 ± 0.5	7.5 ± 0.5	8.4 ± 0.5	
4. Ovalbumin	4.9 ± 0.5	6.0 ± 0.5	6.5 ± 0.5	7.2 ± 0.5	7.8 ± 0.5
5. Pepsin	3.9 ± 0.5	5.2 ± 0.5	6.9 ± 0.5	7.4 ± 0.5	8.0 ± 0.5
6. α-Chymotrypsinogen A	2.2 ± 0.6	3.2 ± 0.5	4.4 ± 0.5		
7. α-Chymotrypsin	0.6 ± 0.6	2.8 ± 0.5	3.4 ± 0.5	4.3 ± 0.5	
8. Trypsin	−0.5 ± 0.5	1.7 ± 0.5	3.0 ± 0.5	3.6 ± 0.5	4.1 ± 0.5
9. Trypsinogen	2.8 ± 0.5	4.0 ± 0.5	5.9 ± 0.5	6.3 ± 0.5	7.4 ± 0.5
10. β-Lactoglobulin	2.8 ± 0.5	4.2 ± 0.5	5.8 ± 0.5	7.4 ± 0.5	
11. Myoglobin	5.6 ± 0.5	7.1 ± 0.5	8.5 ± 0.5	8.8 ± 0.5	9.1 ± 0.5
12. α-Lactalbumin	2.0 ± 0.5	2.7 ± 0.5	5.1 ± 0.5	5.6 ± 0.5	
13. Lysozyme	−0.8 ± 0.5	1.3 ± 0.5	2.5 ± 0.5	3.3 ± 0.5	3.9 ± 0.5
14. Ribonuclease A	−0.6 ± 0.5	0.9 ± 0.5	2.2 ± 0.5	2.7 ± 0.5	
15. Cytochrome *c*	1.8 ± 0.5	3.1 ± 0.5	4.5 ± 0.5	5.8 ± 0.5	6.4 ± 0.5

from a value of $-2.1(\pm 0.5) \times 10^{-3}$ cm^3 g^{-1} K^{-1} at 18°C to a value of $-0.2(\pm 0.4) \times 10^{-3}$ cm^3 g^{-1} K^{-1} at 55°C, with the value at 25°C being equal to $-1.7(\pm 0.4) \times 10^{-3}$ cm^3 g^{-1} K^{-1}. By contrast, the temperature dependencies of the partial specific volumes, v°, are almost linear for all 15 proteins, a result consistent with previous observations (Bull & Breese, 1973; Hinz *et al*., 1994). To be specific, all 15 proteins exhibit similar temperature slopes, $\Delta v^\circ / \Delta T$, with an average value equal to $3.5(\pm 0.8) \times 10^{-4}$ cm^3 g^{-1} K^{-1}.

As with $[u]$, the temperature dependencies of k_S° for all proteins can be approximated by second order polynomial functions, as shown in Figure 1. This Figure is divided into two panels to facilitate visualization of and discrimination between the curves (Figure 1A and B). The average temperature slope, $\Delta k_S^\circ / \Delta T$, that results from such fits decreases from a value of $2.0(\pm 0.5) \times 10^{-7}$ cm^3 g^{-1} bar^{-1} K^{-1} at 18°C to a value of $0.1(\pm 0.4) \times 10^{-7}$ cm^3 g^{-1} bar^{-1} K^{-1} at 55°C, with an average value of $1.7(\pm 0.3) \times 10^{-7}$ cm^3 g^{-1} bar^{-1} K^{-1} at 25°C.

Discussion

Dissecting the partial volume and partial compressibility protein data into contributing components

Any attempt to understand and to interpret macroscopic data in terms of microscopic phenomena requires one to conceptually and mathematically resolve each macroscopic observable into contributing components which can be ascribed to specific molecular ''events''. In the sections which follow, we describe such dissections for the partial volume and the partial compressibility data that we have determined for the globular proteins studied in this work.

The volume data

According to scaled particle theory (Reiss, 1965), the partial molar volume, V°, of a solute can be considered to be equal to the sum of four terms (Pierotti, 1965; Stillinger, 1973; Kharakoz, 1992):

$$V^\circ = V_M + V_T + V_I + \beta_{T0} RT \qquad (1)$$

where V_M, is the intrinsic molar volume of the solute, which corresponds to the domain which water cannot penetrate; V_T is the ''thermal volume'', which corresponds to an ''empty'' domain around the solute molecule which results from the mutual thermal motions of the solute and solvent molecules; V_I is the ''interaction volume'', which represents the change in the solvent volume due to hydration; and $\beta_{T0} RT$ is the ideal term, where β_{T0} is the coefficient of solvent isothermal compressibility and R is the universal gas constant.

The ideal term ($\beta_{T0} RT$) in equation (1) is small (about 1 cm^3 mol^{-1} for aqueous solutions) and therefore can be neglected when considering large macromolecules such as proteins, which usually have partial molar volumes, V°, on the order of $\sim 10^4$ cm^3 mol^{-1} (e.g. see Kuntz & Kauzmann, 1974; Gekko & Hasegawa, 1986). Consequently, the partial specific volume, v°, of a protein of a molecular mass M can be considered equal to the sum of only three terms:

$$v^\circ = v_M + v_T + v_I \qquad (2)$$

where $v_M = V_M / M$; $v_T = V_T / M$; $v_I = V_I / M$.

Equation (2) can be reframed into a form which is useful for characterizing solute hydration properties:

$$v^\circ = v_M + v_T + n_h (V_h^\circ - V_0^\circ) / M \qquad (2a)$$

where n_h is the number of waters involved in the hydration shell of a protein; while V_h° and V_0° are the partial molar volumes of water in the hydration shell and in the bulk state, respectively.

For globular proteins, the intrinsic volume, V_M, is equal to the sum of the van der Waals volumes of the constituent atoms plus the total volume of the voids inside the protein molecule which result from its imperfect packing. The thermal volume, V_T, can be considered to correspond to a layer of ''empty'' space around the solute molecule which results

from the mutual thermal motions of the solute and solvent molecules. For low-molecular-mass compounds, it has been shown that, on average, the thickness, Δ, of such a layer equals 0.50 Å at 25°C and does not significantly depend of the shape and the chemical nature of the solute molecule (Kharakoz, 1992). The interaction volume, V_I, results from the interactions with surrounding water molecules of each atomic group on the protein surface, and, therefore, can be expressed as the following sum:

$$V_I = A_c S_c + A_p S_p + A_n S_n \quad (3)$$

where S_c, S_p, and S_n are the solvent accessible surface areas of the charged, polar, and nonpolar atomic groups, respectively, such that $S_A = S_c + S_p + S_n$; and A_c, A_p, and A_n are the contributions to the interaction volume, V_I, of a unit (1 Å^2) of the solvent accessible surface areas of the charged, polar, and nonpolar atomic groups, respectively. This interaction volume, V_I, is related to the partial molar volumes, V°_{hi}, and to the numbers, n_{hi}, of the waters which solvate the charged, polar, and nonpolar atomic groups. The relevant expression is:

$$V_I = \sum_{i=1}^{3} n_{hi}(V^\circ_{hi} - V^\circ_0) \quad (4)$$

The compressibility data

Analogous to the above analysis of volume, the partial specific adiabatic compressibility, k°_S, of a globular protein can be expressed as the sum of both intrinsic and hydration contributions:

$$k^\circ_S = (K_M + B_c S_c + B_p S_p + B_n S_n) / M \quad (5)$$

Table 4. Comparison of data from this work with literature values for the partial specific volume, v°, and the partial specific adiabatic compressibility, k°_S, at 25°C for 15 globular proteins

Protein	v° (cm^3 g^{-1})		k°_S (10^{-6} cm^3 g^{-1} bar^{-1})	
1. Conalbumin	0.729[a]	0.728[b]	4.6[a]	3.6[b]
		0.732[f]		
2. BSA	0.735[a]	0.735[b]	6.6[a]	7.7[b]
		0.735[c]		5.5[c]
		0.734[f]		
		0.735[g]		
3. Hemoglobin	0.743[a]	0.745[b]	6.8[a]	8.1[b]
		0.751[c]		6.8[c]
		0.746[e]		5.7[e]
		0.750[f]		
4. Ovalbumin	0.735[a]	0.746[b]	6.0[a]	6.8[b]
		0.748[f]		
5. Pepsin	0.730[a]	0.743[b]	5.2[a]	6.4[b]
6. α-Chymotrypsinogen A	0.727[a]	0.717[b]	3.2[a]	2.9[b]
		0.721[f]		
		0.733[g]		
7. α-Chymotrypsin	0.717[a]	0.717[b]	2.8[a]	3.0[b]
		0.738[g]		
8. Trypsin	0.718[a]	0.719[b]	1.7[a]	0.7[b]
9. Trypsinogen	0.721[a]	0.718[b]	4.0[a]	1.0[b]
10. β-Lactoglobulin	0.731[a]	0.751[b]	4.2[a]	6.3[b]
		0.751[f]		
		0.750[g]		
11. Myoglobin	0.742[a]	0.747[b]	7.1[a]	6.7[b]
		0.730[e]		4.4[e]
		0.743[f]		
12. α-Lactalbumin	0.711[a]	0.736[b]	2.7[a]	6.1[b]
		0.704[g]		
13. Lysozyme	0.699[a]	0.712[b]	1.3[a]	3.3[b]
		0.725[e]		4.0[e]
		0.703[f]		
		0.702[g]		
14. Ribonuclease A	0.702[a]	0.704[b]	0.9[a]	0.8[b]
		0.692[f]		
		0.696[g]		
15. Cytochrome *c*	0.735[a]	0.725[b]	3.1[a]	0.1[b]
		0.733[d]		2.9[d]
		0.720[f]		

[a] This study.
[b] Gekko & Hasegawa (1986).
[c] Iqball & Verrall (1987).
[d] Kharakoz & Mkhitaryan (1986).
[e] Gavish *et al*. (1983).
[f] Kuntz & Kauzmann (1974).
[g] Lee *et al*. (1979).

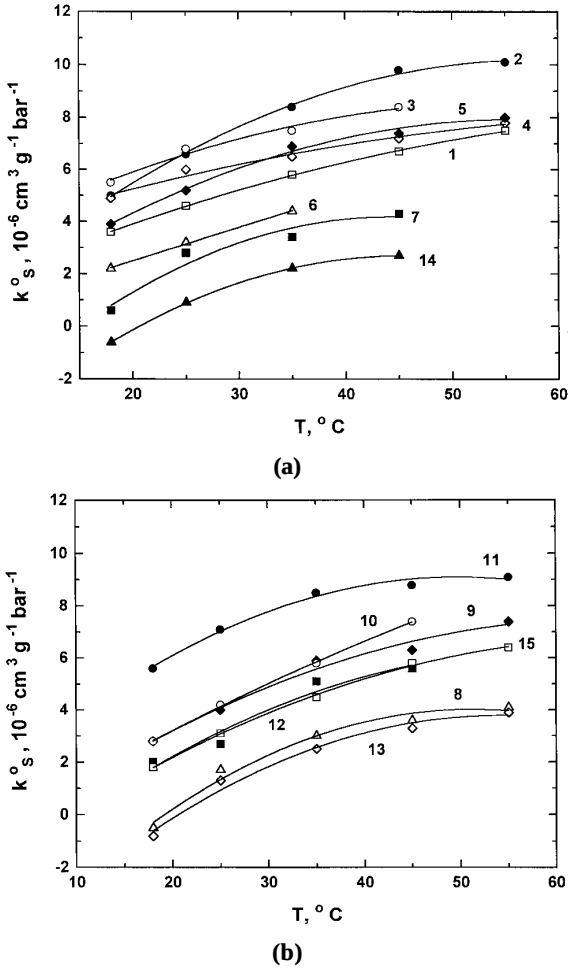

Figure 1. The temperature dependencies of the protein partial specific adiabatic compressibilities (the plot numbers correspond to the corresponding numbers of the proteins in Table 3).

where K_M is the intrinsic molar adiabatic compressibility of the protein; and B_c, B_p, and B_n are the unit (per 1 Å^2) compressibility contributions of the solvent accessible surface areas of the charged, polar, and nonpolar atomic groups, respectively. This partial specific adiabatic compressibility, k_S°, is related to the partial molar adiabatic compressibilities, K_{Shi}°, and to the numbers, n_{hi}, of the waters which solvate the charged, polar, and nonpolar atomic groups by the expression:

$$k_S^\circ = [K_M + \sum_{i=1}^{3} n_{hi}(K_{Shi}^\circ - K_{S0}^\circ)] / M \qquad (6)$$

where K_{S0}° is the partial molar adiabatic compressibility of bulk water.

The intrinsic compressibility, K_M, in equation (6) is equal to $\beta_M V_M$, where β_M is the coefficient of adiabatic compressibility of the protein interior and V_M is the intrinsic molar volume of the protein. As a first approximation, one can assume that β_M is inversely proportional to the packing density, ρ_M, of the protein interior which is equal to the ratio

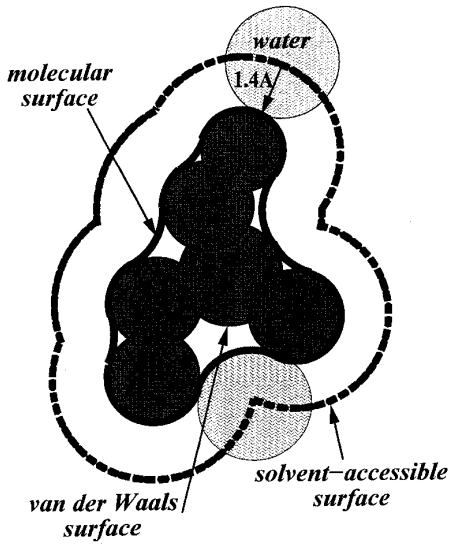

Figure 2. Schemtic diagram of the solvent-accessible, molecular, and van der Waals surfaces of a protein.

V_W / V_M, where V_W is the sum of the van der Waals volumes of the constituent atoms and V_M is the intrinsic volume. This assumption may be an oversimplification, because different elements of secondary structure (α-helices, β-sheets, loops, etc.) may have similar packing densities but still contribute differently to the intrinsic compressibility of a protein. Nevertheless, with this assumption, we can write the following expression for K_M:

$$K_M = B_M V_M^2 / V_W \qquad (7a)$$

where B_M is the coefficient of proportionality. Since $K_M = \beta_M V_M$, the coefficient of adiabatic compressibility of the protein interior, β_M, can be set equal to $B_M V_M / V_W$ to yield the expression:

$$\beta_M = B_M V_M / V_W \qquad (7b)$$

In the above discussion, we have described how our experimental volume and compressibility data on globular proteins can be resolved both conceptually and mathematically into contributing components. In the section which follows, we describe how X-ray data on the same globular proteins can be used to define and to resolve the volumes and solvent accessible surfaces of each protein in terms of contributions from the component atoms.

Use of X-ray data to define intrinsic protein volumes and accessible surface areas: resolving contributions from charged, polar, and nonpolar atoms

In the discussion which follows, we consider the three types of protein surfaces illustrated in Figure 2; namely: the solvent-accessible surface, the molecular surface, and the van der Waals surface. As originally defined by Lee & Richards (1971), the solvent accessible surface is that surface which is traced out by the center of a probe sphere solvent

Table 5. The surface areas (Å^2) of the charged, S_c, polar, S_p, and nonpolar, S_n, atomic groups, the intrinsic volumes, V_M, (Å^3), the specific intrinsic volumes, v_M, (cm^3 g^{-1}), and the sum of van der Waals volumes, V_W, (Å^3) for 12 proteins

Protein	S_c	S_p	S_n	V_M	v_M	V_W
Hemoglobin	4248	5536	15,724	71,113	0.630	52,995
Ovalbumin	3161	5180	8597	48,796	0.639	36,140
Pepsin	1919	5014	6782	39,024	0.662	28,766
α-Chymotrypsinogen A	1165	3936	5714	29,244	0.685	21,573
α-Chymotrypsin	1238	3991	5426	28,353	0.675	20,645
Trypsin	827	3869	4809	26,288	0.688	19,524
Trypsinogen	753	4033	4889	26,120	0.684	19,121
Myoglobin	1242	1702	4772	19,458	0.658	14,380
α-Lactalbumin	1548	1874	3794	15,250	0.642	11,819
Lysozyme	907	2548	3230	15,659	0.659	11,834
Ribonuclease A	927	2401	3462	14,814	0.656	11,428
Cytochrome *c*	1166	1503	3446	12,634	0.614	9717

molecule as it rolls over the surface of a protein. As previously defined by Richards (1977), the molecular surface of a protein has two components: (1) that part of the protein surface which contacts a rolling probe solvent molecule; and (2) a re-entrant surface, which corresponds to a series of patches formed by the interior-facing domain of the probe when it simultaneously contacts more than one atom on the protein surface. The protein van der Waals surface consists of the unoccluded parts of the van der Waals spheres which correspond to each atom of the protein. The atomic van der Waals radii used in our calculations were derived from ECEPP nonbonded parameters (Momany *et al*., 1975) and have the following values: 1.729 Å for an aliphatic carbon atom, 1.604 Å for an aromatic carbon atom, 1.634 Å for a carbonyl carbon atom, 1.634 Å for the γ- and δ2-carbon atoms of the imidazole ring of histidine, 1.729 Å for the δ-carbon atom of proline and hydroxyproline, 1.524 Å for amino and amide nitrogen atoms, 1.373 Å for a carbonyl oxygen atom, 1.402 Å for a hydroxyl oxygen atom, and 1.829 Å for a sulfur atom.

Each of the three surfaces noted above has a corresponding area and volume associated with it. In fact, we have calculated the solvent-accessible surface areas, S_A, the molecular volumes, V_M, and the van der Waals volumes, V_W, of the following proteins using the indicated atomic coordinate sets obtained from the Brookhaven Protein Data Bank (Bernstein *et al*., 1977): 1hda for hemoglobin, 1ova for ovalbumin, 4pep for pepsin, 2cga for α-chymotrypsinogen A, 4cha for α-chymotrypsin, 2ptn for trypsin, 1tgn for trypsinogen, bov for myoglobin, 1alc for α-lactalbumin, 4lyz for lysozyme, 5rsa for ribonuclease A, and 5cyt for cytochrome *c*. For each protein, the total solvent-accessible surface area was subdivided into component solvent-accessible surface areas associated with the different protein atomic groups. The calculations required for these subdivisions used the method originally developed by Shrake & Rupley (1973) in which the surface of each heavy atom (excluding hydrogen atoms) is considered to be a sphere represented by a set of nearly uniformly distributed points. In this work, all calculations were made using a ICM 2.0 program (MOLSOFT, Metuchen, NJ) which employs a recently described dot-exclusion algorithm (Abagyan *et al*., 1994) that significantly increases the computational speed. Table 5 lists the accessible surface areas of the charged, S_c, polar, S_p, and nonpolar, S_n, atomic groups so obtained for the subset of 12 proteins studied here.

The definition of protein molecular surface area, S_M, was first introduced by Richards (1977), with Conolly (1983) developing the mathematical formalism for analytical calculation of S_M. We have calculated S_M using a newly developed ''contour-buildup'' algorithm (Totrov & Abagyan, 1996) which allows one to increase the speed of calculations by an order of magnitude, relative to Conolly's original algorithm. We divided the molecular surface so constructed into triangles (triangulated) with average areas of 0.062 Å^2. The protein molecular volumes, V_M, then were calculated using the expression

$$V_M = (1/3)\int_S \mathbf{r}\,d\mathbf{S_M} \qquad (8)$$

where **r** is a radius-vector from an arbitrary origin to an elementary triangle surface, $\mathbf{S_M}$.

Table 5 lists these molecular (or intrinsic) volumes, V_M (column 5), as well as the sums of the van der Waals volumes, V_W, of the constituent atoms (column 7) for the 12 globular proteins with known crystal structures studied here. The van der Waals volumes, V_W, of the proteins were calculated in two ways: (1) as sums of the standard contributions from the constituent amino acid residues; and (2) as volumes confined within the solvent-accessible surfaces when the radius of the probe is set to 0. Both approaches yield similar V_W values. For this analysis, we calculated and used the following standard contributions in Å^3 of each amino acid residue to the protein van der Waals volume: Ala, 67; Cys, 86; Asp, 91; Glu, 109; Phe, 135; Gly, 48; His, 118; Ile, 124; Lys, 135; Leu, 124; Met, 124; Asn, 96; Pro, 90; Gln, 114; Arg, 148; Ser, 73; Thr, 93; Val, 105; Trp, 163; and Tyr, 141.

Inspection of the data in Table 5 reveal the following general features of globular proteins: (1) their interior packing densities, ρ_M (equal to the ratio V_W / V_M), fall between 0.72 and 0.78; (2) their van der Waals volumes, V_W (Å^3), are proportional to molecular mass, M (Da), with $V_W = (1100 \pm 300) + (0.77 \pm 0.01)\, M$; (3) their intrinsic volumes, V_M (Å^3), are proportional to molecular mass, M (Da), with $V_M = (1200 \pm 500) + (1.04 \pm 0.02)\, M$, while the average specific intrinsic volume, v_M, is 0.658($\pm$0.022) cm^3 g^{-1} (v_M is equal to $N_A V_M / M$, where N_A is the Avogadro's number, M is the protein molecular mass, and V_M is the intrinsic protein volume in Å^3); (4) the total accessible surface area, S_A (Å^2) ($S_c + S_p + S_n$), is proportional to the 2/3 power of the molecular mass, M (Da), with $S_A = -(1800 \pm 200) + (14.5 \pm 0.25)\, M^{2/3}$; (5) the average contributions of charged, polar, and nonpolar atomic groups to S_A are 14($\pm$4)%, 33($\pm$7)%, and 53($\pm$5)%, respectively. These semiempirical relationships should prove useful in a wide range of approximate calculations. Note that these relationships between the molecular mass, M, and S_A and V_M differ somewhat from analogous ones previously reported (Richards, 1977). Such a disparity is to be expected given the somewhat different approaches used and the expanded data base now available for deriving relationships between S_A, V_M, and M.

We also have used the published crystallographic coordinates to define the actual number of solvent-accessible charged, N_c, polar, N_p, and nonpolar, N_n, atoms in each of the 12 globular proteins analyzed in this work. These data are listed in Table 6. In conjunction with the corresponding component surface areas listed in Table 5 (S_c, S_p, S_n), we have calculated the following average values for the accessible surface areas of a charged, a polar, and a nonpolar atoms: 29.5($\pm$2.7) Å^2 (charged atom); 24.0($\pm$1.2) Å^2 (polar atom); and 20.9($\pm$1.2) Å^2 (nonpolar atom).

In the sections which follow, we describe how the X-ray based characterizations discussed above can be combined with the acoustic and densimetric data presented earlier to resolve and to define the contributions that component charged, polar, and nonpolar atoms make to each observable, thereby yielding insight into protein hydration properties.

Table 6. The numbers of charged, N_c, polar, N_p, and nonpolar, N_n, atomic groups in 12 proteins

Protein	N_c	N_p	N_n
Hemoglobin	165	253	730
Ovalbumin	116	202	419
Pepsin	71	203	338
α-Chymotrypsinogen A	39	165	264
α-Chymotrypsin	40	161	268
Trypsin	25	156	221
Trypsinogen	25	159	241
Myoglobin	40	77	203
α-Lactalbumin	52	81	189
Lysozyme	36	103	175
Ribonuclease A	31	101	158
Cytochrome *c*	34	66	171

Resolving intrinsic and hydration contributions to the partial specific protein volumes: defining unit contributions of the charged, polar, and nonpolar constituent atoms

Comparison of the partial (apparent) specific protein volumes, $v°$ (φV), listed in Table 2 with the corresponding specific intrinsic protein volumes, v_M, listed in Table 5, reveals that the average contribution of v_M to $v°$ is about 90%; in other words, $v° - v_M \approx 0.1v°$. The remaining 10% of $v°$ ($0.1v°$) is the sum of a positive contribution from the thermal volume, v_T, and a negative contribution from the interaction volume v_I, as expressed below in a rearranged form of equation (2):

$$0.1v° \approx v° - v_M = v_I + v_T \qquad (9)$$

As emphasized above, the interaction volume, v_I, is the only component of the partial specific volume of a protein that is sensitive to hydration. As reflected in equation (9), the interaction volume, v_I, can be calculated from measured values of the partial specific volume, $v°$, if one can estimate the thermal volume, v_T.

One approach to estimating v_T is to use the average ''thickness'' of the thermal volume, Δ, determined for low-molecular-mass compounds (Kharakoz, 1992), and to multiply this by the total accessible protein surface area, S_A. This approach, however, assumes that the value of Δ determined for low-molecular-mass substances is valid for macromolecules. An alternative approach for estimating v_T assumes that the contribution of nonpolar groups to the interaction volume, v_I, is negligible, since, for the most part, nonpolar groups do not interact directly with solvating water molecules (Kharakoz, 1992). Consequently, according to equation (2), for a nonpolar solvent accessible area, the unit contribution to $v°$ should be determined solely by its contribution to the thermal volume, v_T. This unit contribution in Å^3 to v_T can be approximated by the product of a unit (Å^2) of nonpolar accessible surface and the average thickness of the thermal volume, Δ. If one reasonably assumes that Δ does not depend significantly on the type of the solvent-accessible surface, then, the unit contribution to $v°$ of the nonpolar accessible protein surface should be equal to the unit contribution to v_T of any kind of accessible surface.

To find the unit contribution to $v°$ of the nonpolar surface area, we performed a linear regression analysis of our experimental $v°$ data listed in Table 2 in conjunction with the accessible surface area and intrinsic volume data listed in Table 5 using the equation

$$v° = C_M v_M + [C_{cp}(S_c + S_p) + C_n S_n] / M \qquad (10)$$

Table 7. Values for coefficients C_{cp}, C_n ($cm^3\ mol^{-1}\ Å^{-2}$), and C_M in equation (10), coefficients B_c, B_p, and B_n ($10^{-6}\ cm^3\ mol^{-1}\ bar^{-1}\ Å^{-2}$) in equation (5), and coefficient B_M ($10^{-6}\ cm^3\ mol^{-1}\ bar^{-1}\ Å^{-3}$) in equation (7a) at the indicated temperatures

Coefficients	18°C	25°C	35°C	45°C	55°C
C_{cp}	−0.47 ± 0.05	−0.47 ± 0.05	−0.46 ± 0.05	−0.46 ± 0.05	−0.46 ± 0.05
C_n	0.62 ± 0.05	0.62 ± 0.05	0.63 ± 0.05	0.63 ± 0.05	0.63 ± 0.05
C_M	1.028 ± 0.005	1.033 ± 0.005	1.035 ± 0.005	1.036 ± 0.005	1.041 ± 0.005
B_c	−15 ± 10	−15 ± 10	−15 ± 10	−15 ± 10	−15 ± 10
B_p	−66 ± 3	−62 ± 3	−60 ± 3	−58 ± 3	−57 ± 3
B_n	−13 ± 2	−9 ± 2	−6 ± 2	−5 ± 2	−4 ± 2
B_M	10.9 ± 0.3	11.0 ± 0.3	11.1 ± 0.3	11.2 ± 0.3	11.3 ± 0.3

where C_M is a coefficient which takes into account the volume fluctuations and thermal expansion of the protein interior; C_{cp} is the average unit contribution to $v°$ of the charged or polar surface; and C_n is the unit contribution to $v°$ of the nonpolar surface. Table 7 lists the values we determined for the coefficients in equation (10) by means of a linear regression analysis, while the data in columns 2 and 3 of Table 8 compare the experimental $v°$ values with those calculated at 25°C using equation (10) and the coefficients listed in Table 7. Inspection of these data reveal an average deviation between experimental and calculated $v°$ values of only (± 0.020) $cm^3\ g^{-1}$, which is less than 3% of the absolute value of the partial specific volumes, $v°$, of the proteins studied. Consequently, we consider the agreement between the experimental and calculated $v°$ values at 25°C to be quite good. We have presented this comparison only at a single temperature in order to illustrate the agreement between the calculated and experimental $v°$ values. However, we observe similar agreement at all other temperatures studied.

Note that, in our formulation of equation (10), we intentionally do not discriminate between charged and polar surfaces. This averaging is done because the ''surface'' (nonintrinsic) contribution to $v°$ is relatively small (about 10%) and because the contribution of charged groups to the total accessible protein surfaces also is relatively small (about 14%). In fact, within the limits of our experimental error, it is not possible to estimate reliably the separate unit contributions to $v°$ of the charged and polar accessible surface areas.

Comparison of equation (10) with equations (2) and (3) shows that the thermal volume, v_T, is ''distributed'' evenly between the coefficients C_{cp} and C_n. Inspection of the data in Table 7 reveals, that at 25°C, the contribution of 1 $Å^2$ of a nonpolar surface of a globular protein to $v°$ (coefficient C_n) and, hence, the contribution of 1 $Å^2$ of any surface to v_T is equal to 0.62($\pm$0.05) $cm^3\ mol^{-1}$. Note that this contribution does not significantly depend on temperature. By dividing $0.62\ cm^3\ mol^{-1}\ Å^{-2}$ by Avogadro's number, N_A, one finds that 1.0 $Å^2$ of protein surface contributes 1.0($\pm$0.1) $Å^3$ to the surrounding thermal volume V_T. In other words, the average thickness of the thermal volume, Δ, is about 1.0 Å. Consequently, the thermal volume, V_T, surrounding a protein molecule can be obtained by multiplying its total accessible surface area, S_A, by this value of Δ. Significantly, this Δ value is twice as large as 0.5 Å, the value estimated for low molecular weight compounds (Kharakoz, 1992). This disparity suggests that mutual thermal motions of macromolecules and solvating waters involve additional modes that are absent in small molecules. This distinction underscores one deficiency associated with the direct application of low-molecular-mass model compound data to the analysis of protein molecules.

Further inspection of the data in Table 7 reveals that, at 25°C, the value of C_{cp}, the average unit contribution to $v°$ of the charged or polar surface is negative and equal to −0.47($\pm$0.05) $cm^3\ mol^{-1}\ Å^{-2}$, while not significantly depending on temperature. This value of C_{cp} is a sum of two terms: (1) a positive term equal to the contribution of 1 $Å^2$ of the polar and charged surface to the thermal volume, V_T; and (2) a negative term representing the contribution of 1 $Å^2$ of the polar or charged surface to the interaction volume, V_I. As discussed above, the positive contribution to C_{cp} can be considered to be equal to C_n. Consequently, to find the interaction volume, V_I, one multiplies ($S_c + S_p$), the accessible surface area of the polar and charged atomic groups, by ($C_{cp} - C_n$), which at 25°C equals

Table 8. Comparison of the experimental and calculated values of the partial specific volume, $v°$ ($cm^3\ g^{-1}$), and partial specific adiabatic compressibility, $k_S°$ ($10^{-6}\ cm^3\ g^{-1}\ bar^{-1}$), at 25°C for 12 globular proteins

Protein	$v°$ (exp)	$v°$ (calc)	$k_S°$ (exp)	$k_S°$ (calc)
Hemoglobin	0.745	0.727	6.8	7.4
Ovalbumin	0.737	0.691	6.0	6.1
Pepsin	0.733	0.711	5.2	5.1
α-Chymotrypsinogen A	0.730	0.753	3.2	4.8
α-Chymotrypsin	0.721	0.734	2.8	4.5
Trypsin	0.720	0.745	1.7	4.1
Trypsinogen	0.725	0.741	4.0	3.8
Myoglobin	0.745	0.769	7.1	6.9
α-Lactalbumin	0.713	0.716	2.7	3.0
Lysozyme	0.702	0.708	1.3	1.9
Ribonuclease A	0.704	0.721	0.9	1.3
Cytochrome *c*	0.738	0.706	3.1	3.1

$-1.1(\pm 0.1)$ cm^3 mol^{-1} Å^{-2}. The magnitude and sign of $(C_{cp} - C_n)$ characterizes the contraction of water due to the influence of adjacent charged (electrostriction) or polar (hydrogen bonding) atomic groups.

Resolving intrinsic and hydration contributions to the partial specific adiabatic compressibility data: defining unit contributions of the charged, polar, and nonpolar constituent atoms

To achieve a similar dissection of the compressibility data, we have determined values for the coefficients B_M, B_c, B_p, and B_n in equations (5) and (7a) *via* a linear regression analysis of the experimental k_S° data listed in Table 3, in conjunction with the accessible surface area and intrinsic volume data presented in Table 5. The results of this analysis are given in Table 7, while columns 4 and 5 of Table 8 compare the experimental and calculated values of k_S° at 25°C. Inspection of the data in columns 4 and 5 of Table 8 reveals rather good agreement between the experimental and calculated k_S° values, with the average deviation between the two data sets being only $\pm 0.7 \times 10^{-6}$ cm^3 g^{-1} bar^{-1}, which is very close to the experimental error in the k_S° measurements themselves. Note, however, that for the two pancreatic proteolytic enzymes trypsin and α-chymotrypsin, the difference between the experimental and calculated k_S° values is considerably larger than the average deviation (-2.4×10^{-6} cm^3 g^{-1} bar^{-1} for trypsin and -1.7×10^{-6} cm^3 g^{-1} bar^{-1} for α-chymotrypsin). These larger deviations may reflect partial autoproteolysis of these two proteases which would produce relatively short, unfolded peptide fragments, a possibility not accounted for in our analysis of the data.

Armed with the B_M data listed in Table 7 and the V_M and V_W values listed in Table 5, we used equation (7b) to calculate β_M, the intrinsic coefficients of adiabatic compressibility. On average, we find that β_M at 25°C is equal to $25(\pm 1) \times 10^{-6}$ bar^{-1}. This value, is about twice as high as a previous estimate $13(\pm 3) \times 10^{-6}$ bar^{-1} (Kharakoz & Sarvazyan, 1993), but still characterizes the protein interior as a rigid tightly packed solid-like substance, thereby being conceptually if not quantitatively consistent with previous characterizations. Our data also reveal the temperature dependence of β_M, $\partial(\ln \beta_M)/\partial T$, to be on the order of $1(\pm 1) \times 10^{-3}$ K^{-1}. This value is consistent with one previous estimate (Kharakoz & Sarvazyan, 1993), and its magnitude suggests that the protein interior is solid-like.

Further inspection of the data in Table 7 reveals that B_n, the compressibility contribution of 1 Å^2 of the nonpolar surface, increases with temperature, while manifesting negative values within the entire temperature range studied. By comparison, the compressibility contribution of a —CH_2— group within low-molecular-mass compounds is negative at low temperatures and becomes positive at temperatures above 35°C (Kharakoz, 1991; Chalikian *et al*., 1993, 1994a). Thus, the hydration properties of nonpolar groups on protein surfaces and nonpolar groups within low-molecular-mass compounds are different. For this reason, caution should be exercised when using model compound data to interpret protein hydration properties.

The data in Table 7 also reveal that the compressibility contribution of 1 Å^2 of the charged and polar surfaces (coefficients B_c and B_p) are negative within the temperature range studied. Strictly speaking, waters which solvate positively charged amino groups are distinct, with respect to their volume and compressibility characteristics, from waters which solvate negatively charged carboxyl groups (Chalikian *et al*., 1994b). However, within the limits of our experimental errors and computational assumptions, it is not possible for us to discriminate between the compressibility contributions of positively and negatively charged groups.

Disparities between volume and compressibility data obtained on low-molecular-mass model compounds and on protein systems: implications for protein hydration

Hydration of the polar groups

Kharakoz (1991) has estimated the lowest compressibility contribution of a polar group within an amino acid to be -5.5×10^{-4} cm^3 mol^{-1} bar^{-1} at 25°C. We estimate the average compressibility contribution of a polar group on a protein surface to be $-(15(\pm 1.5)) \times 10^{-4}$ cm^3 mol^{-1} bar^{-1} by multiplying $(62(\pm 3) \times 10^{-6}$ cm^3 mol^{-1} bar^{-1} Å^{-2}, the compressibility contribution at 25°C of 1 Å^2 of a polar surface (coefficient B_p in equation (5)), by $24.0(\pm 1.2)$ Å^2, the average accessible surface area of a polar group. Note that this protein surface polar group compressibility contribution is almost three times more negative than the compressibility contribution of a polar group within an amino acid (-15×10^{-4} cm^3 mol^{-1} bar^{-1} *versus* -5.5×10^{-4} cm^3 mol^{-1} bar^{-1}). This disparity illustrates the potential dangers of directly applying low-molecular-mass model compound data to the analysis of protein systems.

A similar disparity between model compound and protein data emerges upon analysis of the volume. By multiplying $-1.1(\pm 0.1)$ cm^3 mol^{-1} Å^{-2}, our estimated contribution to V_I of 1 Å^2 of a polar and charged surface at 25°C (coefficient C_{cp} in equation (10)), by $24.0(\pm 1.2)$ Å^2, the average accessible surface area of a polar group, we calculate the average contribution to the interaction volume, V_I, at 25°C of a polar group on the protein

surface to be $-26(\pm 4)$ cm^3 mol^{-1}. The lowest V_I contribution at 25°C of a polar group within an amino acid (the hydroxyl group of serine) has been estimated to be –7 cm^3 mol^{-1} (Kharakoz, 1989), which is 3.5 times less negative, thereby once again underscoring the potential deficiency of low-molecular-model compound data.

An examination of equations (3) to (6) suggests that the lower values we find for the contributions to V_I and k_S° of polar groups on the protein surfaces compared with the corresponding values derived from low-molecular-mass compounds could result from either higher numbers of waters in the hydration shell of protein polar groups (n_{hi}) or lower values of the partial molar volume, V_{hp}°, and adiabatic compressibility, K_{Shp}°, of these waters. In low-molecular-mass compounds, the hydration shells of the atomic groups are primarily confined within the first coordination spheres (see Chalikian *et al*. (1994a) and articles cited therein). Consequently, only waters directly contacting such a solute molecule will display altered volume and compressibility characteristics. If this hydration behavior also holds for proteins, then the number of waters in the hydration shell of polar groups, n_{hp}, should correspond to the ratio S_p / S_W, where S_W is equal to 9 Å^2, the effective cross-section of a water molecule. With this assumption, we can use equations (4) and (6) to calculate values of V_{hp}° and K_{Shp}° of 8.8 cm^3 mol^{-1} (49% from the value of bulk water) and 2.5×10^{-4} cm^3 mol^{-1} bar^{-1} (31% from the value of bulk water), respectively. Such low values of V_{hp}° and K_{Shp}° are unrealistic, since even a strong electrolyte such as NaCl exhibits values over twice high (Onori, 1988; Onori & Santucci, 1990). We, therefore, propose that, n_{hp}, the number of waters involved in the hydration shells of protein polar groups, exceeds the number of waters confined within the first hydration layer by a factor of 3 to 4. This enhancement could result from closely located polar groups on a protein surface favoring solvation *via* the formation of water networks which involve water molecules from the second and, possibly, even from the third coordination spheres. In such networks, a water molecule may simultaneously form hydrogen bonds with two or more polar groups on the protein surface, thereby becoming highly immobilized and thus exhibiting unusually low values of V_{hp}° and K_{Shp}°. This description of protein hydration is consistent with previous work on low-molecular-mass compounds in which it was shown that the hydration of a polar group depends on its proximity to other polar groups (Kharakoz, 1991; Chalikian *et al*., 1994a). Specifically, the compressibility contribution of a single polar group (separated from other polar groups by five or more covalent bonds) was found to be positive and equal to $+3.7 \times 10^{-4}$ cm^3 mol^{-1} bar^{-1}, suggesting a ''low'' hydration of single polar groups (Kharakoz, 1991). However, when polar groups within a solute molecule are situated sufficiently close to each other (separated by three or less covalent bonds) then it has been postulated that each adjacent water molecule can, in principle, form simultaneously two hydrogen bonds with the neighboring polar groups (Chalikian *et al*., 1994a; Kharakoz, 1991). Consequently, the hydration of each group within two closely located polar groups ''increases'', and their compressibility becomes negative and equal to -5.5×10^{-4} cm^3 mol^{-1} bar^{-1} (Kharakoz, 1991). By extension, on the protein surface, where there are many closely located polar groups, a water molecule can simultaneously form hydrogen bonds with more than two neighboring polar groups, thereby becoming highly immobilized. As a result, it is reasonable to expect that the compressibility contribution of a polar group on the protein surface could be even more negative. Furthermore, such highly immobilized waters on the protein polar surface may, in turn, form hydrogen bonds with waters from the second hydration shell, thus facilitating formation of water networks. In such a scenario, these highly immobilized waters of the first hydration shell, act, in some respect, as pseudo polar groups on the protein surface, which facilitate the involvement in protein hydration of waters from a second and, even, possibly, a third coordination sphere. Since the concept of immobilized water molecules at the protein surface described here is thermodynamic in nature, it should not be viewed as inconsistent with the less than 500 ps residence times of water molecules at the protein surface that have been kinetically characterized by NMR (Otting *et al*., 1991; Otting & Liepinsh, 1995).

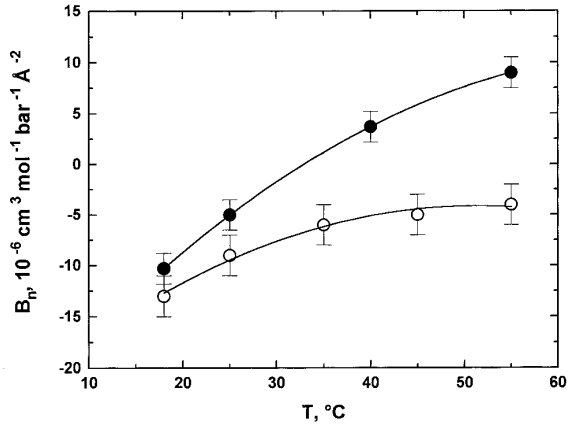

Figure 3. Temperature dependencies of compressibility contributions of 1 Å^2 of protein nonpolar surface (○) and nonbranched chain of —CH_2— groups in α, ω-aminocarboxylic acids (●).

Hydration of nonpolar groups

Inspection of the data in Table 7 reveals that the coefficient B_n, which reflects the unit compressibility contribution of the solvent accessible nonpolar protein surface area, increases over threefold from a value of $-13(\pm 2) \times 10^{-6}$ cm^3 mol^{-1} bar^{-1} Å^{-2}

at 18°C to a value of $-4(\pm 2) \times 10^{-6}$ cm^3 mol^{-1} bar^{-1} Å^{-2} at 55°C. In Figure 3, we compare this temperature dependence of B_n with the temperature dependence of the compressibility contribution per 1 Å^2 of the accessible surface area of an independently hydrated —CH_2— group in the nonbranched chains of α, ω-aminocarboxylic acids (Chalikian *et al*., 1993). Note that, at 18°C (and, perhaps, also at lower temperatures), the lines nearly merge, thereby indicating that the unit compressibility contributions of a hydrophobic domain on a protein surface is close to that of a nonbranched chain of —CH_2— groups in a family of low-molecular-mass compounds. By contrast, at higher temperatures, the lines diverge in a direction which suggests that waters which solvate hydrophobic groups on a protein surface are less compressible than those which solvate a —CH_2— group in small organic molecules. One possible ''explanation'' for this divergence is that at higher temperatures some of the waters solvating nonpolar groups on the protein surface become influenced by neighboring polar groups (possibly, through water networks). Independent of the veracity of this ''explanation'', the direction of the observed divergence is consistent with the conventional wisdom concerning the impact of hydrophobic hydration on compressibility (Chalikian *et al*., 1994a). Specifically, the influence of nonpolar groups on the compressibility of surrounding waters is generally explained by hydrogen bonding between these waters in a restricted space (Chalikian *et al*., 1994a). With an increase in temperature, these hydrogen bonds weaken, thereby allowing some of the water molecules within the original hydration shells of nonpolar groups to form hydrogen bonds with neighboring polar groups on the protein surface.

Hydration of charged groups

Inspection of the data in the fourth row of Table 7 reveals the coefficient B_c, the unit compressibility contribution of the charged protein surface area, to be equal to $-15(\pm 10) \times 10^{-6}$ cm^3 mol^{-1} bar^{-1} Å^{-2}. This value is close to that determined for small charged molecules. For example, the value at 25°C of the partial molar adiabatic compressibility of glycine, -26.6×10^{-4} cm^3 mol^{-1} bar^{-1} (Kharakoz, 1991; Chalikian *et al*., 1993), can be divided by 136 Å^2, the accessible surface area of the charged amino and carboxyl termini of glycine, to calculate a unit compressibility contribution of -20×10^{-6} cm^3 mol^{-1} bar^{-1} Å^{-2} for glycine. The qualitative similarity between this small molecule value and the protein value for B_c listed in Table 7 suggests that charged protein groups and low molecular weight compounds may be solvated in similar ways, with the hydration shells of charged groups, in contrast with polar groups, being confined primarily within the first coordination sphere. We propose that this confinement results from the electrostatic field of the charged groups strongly orienting surrounding water molecules, thereby inhibiting their participation in the water networks that we suggest form around polar groups.

Regression analysis of protein data versus *model compound data for interpreting protein volumetric properties*

For interpreting protein volumetric data, adjustable parameters are included in both the semiempirical ''regression approach'' presented above, and the more conventional ''model compound approach'' discussed above. In our regression approach, the adjustable parameters are the coefficients of the regression analysis, while in the model compound approach, the adjustable parameters are the protein intrinsic volume and intrinsic compressibility. Since both approaches use adjustable parameters, it is reasonable to expect that, for a given family of proteins, calculated values should qualitatively agree with the experimental data. However, due to the presence of the adjustable parameters, such agreement cannot be invoked as proof for the veracity of the model used. In fact, our caution about using small molecule data to model protein hydration is not based on any agreement or disagreement between experimental and calculated data. Instead, our caution is based on our finding that the hydration contributions to volume and compressibility of 1 Å^2 of the protein accessible surface are significantly different from the corresponding contributions derived from studies on low-molecular-mass compounds.

Prediction of the partial specific volume, v°, and the partial specific diabatic compressibility, k_S°, of a globular protein based on X-ray crystallographic data

One of the practical features of the relationships defined in this paper is their predictive power. Specifically, equations (5), (7a), and (10) in conjunction with the data presented in Table 7 allow one to derive the following two expressions at 25°C;

$$v^\circ = [1.033\ 10^{-24} V_M N_A - 0.47(S_c + S_p) + 0.62 S_n] / M \quad (11)$$

$$k_S^\circ = (11 V_M^2 / V_W - 15 S_c - 62 S_p - 9 S_n) / M \quad (12)$$

where v° is in cm^3 g^{-1}; k_S° is in cm^3 g^{-1} bar^{-1}; S_c, S_p, and S_n are in Å^2; and V_M and V_W is in Å^3.

One can use these equations to predict, as a function of temperature, the partial specific volume, v°, and the partial specific adiabatic compressibility, k_S°, of a globular protein based exclusively on crystallographic data (from which one computes values for on the intrinsic volume and the solvent-accessible surface areas). Equations (11) and (12) (or analogous expressions for other

temperatures) allow one to calculate for a native globular protein the ''expected'' values of $v°$ and $k_S°$ at 25°C. Any significant discrepancy between the calculated and measured values of $v°$ and/or $k_S°$ could reflect contributions from a conformational transition or ligand binding event.

For globular proteins with undefined crystallographic structures, rough estimations of $v°$ and $k_S°$, can be obtained by using equations (11) and (12) in conjunction with the following relationships discussed above:

$$V_W = (1100(\pm 300)) + (0.77(\pm 0.01))M$$

$$V_M = (1200(\pm 500)) + (1.04(\pm 0.02))M$$

$$S_A = -(1800(\pm 200)) + (14.5(\pm 0.25))M^{2/3}$$

$$S_c = 0.14S_A;\ S_p = 0.33S_A;\ \text{and}\ S_n = 0.53S_A$$

For example, in the absence of crystallographic data, one can estimate $v°$ and $k_S°$ for a globular protein at 25°C using the following two expressions:

$$v° = 0.647 + 500M^{-1} + 1.56M^{-1/3} \quad (13)$$

$$k_S° = 15.4 \times 10^{-6} + 0.063M^{-1} - 4.0 \times 10^{-4}M^{-1/3} \quad (14)$$

Clearly, analogous relationships can be derived for temperatures other than 25°C.

Concluding Remarks

The study reported here represents the first thermodynamic characterization of protein hydration that does not depend on model compound data but rather is based exclusively on macroscopic ($v°$, $k_S°$) and microscopic (X-ray) measurements on protein molecules themselves. Specifically, we used acoustic and densimetric techniques to determine the partial specific volumes and adiabatic compressibilities of 15 globular proteins in the temperature range between 18 and 55°C. To help interpret these data, we calculated the intrinsic volumes and accessible surface areas for 12 of these proteins from available crystallographic data. By combining these macroscopic and microscopic characterizations, we developed a new quantitative model which allows one to predict the partial specific volume and the partial specific adiabatic compressibility of a globular protein from the crystallographic coordinates of the constituent atoms. We also determined and compared the unit (per 1 Å^2) volume and compressibility contributions of polar, nonpolar, and charged protein surface areas to the corresponding values derived from studies on low-molecular-mass compounds. This comparison revealed the hydration of charged protein surface atomic groups to be very similar to that of charged groups in small organic molecules. By contrast, the hydration of polar protein surface groups was found to be qualitatively different from that of polar groups in small organic molecules. Based on this disparity, we proposed the formation of water networks adjacent to polar protein surfaces, with waters from the second and, perhaps, even, third coordination spheres being involved in such networks. For nonpolar protein surface groups, we found the ability of low-molecular-mass compounds to serve as useful protein models depends on the temperature domain being examined. At low temperatures, nonpolar protein surface groups appear to be hydrated independently, while at higher temperatures, some of the solvating waters become influenced by neighboring polar groups. In the aggregate, our results reveal the limitations of thermodynamic analyses of protein hydration based exclusively on additive models which use data derived from low-molecular-mass compounds.

Materials and Methods

Proteins

All proteins studied in this work were of the highest commercially available purity and were exhaustively dialyzed against water or buffer. The following proteins, with the parenthetically indicated catalogue and lot numbers, were purchased from Sigma Chemical (St Louis, MO): conalbumin from chicken egg white (C-0755; lot 107F8020); bovine hemoglobin (H-2500; lot 110H9315); ovalbumin from chicken egg (A-2512; lot 73H7015); pepsin from porcine stomach mucosa (P-6887; lot 123H8035); α-chymotrypsinogen A from bovine pancreas (C-4879; lot 49F8005); α-chymotrypsin from bovine pancreas (C-4129; lot 61H7175); trypsin from bovine pancreas (T-8642; lot 63H7165); trypsinogen from bovine pancreas (T-1143; lot 32H7140); β-lactoglobulin from bovine milk (L-0130; lot 91H7005); horse heart myoglobin (M-1882; lot 61H7106); α-lactalbumin from bovine milk (L-6010; lot 128F8140); lysozyme from chicken egg white (L-6876; lot 89F8276); ribonuclease A from bovine pancreas (R-5500; lot 104H7110); and the oxidized ferri form of horse heart cytochrome *c* (C-7752; lot 62H7115). The bovine serum albumin (05440; lot 38629/2 891) was obtained from Fluka (Rokonkoma, NJ). For each protein studied, the electrophoretically defined purity according to the manufacturer was no less than 95 to 96%, with the purities for conalbumin, hemoglobin, bovine serum albumin, ovalbumin, myoglobin, and cytochrome *c* being greater than 99%.

Preparation of solutions

Except for β-lactoglobulin, α-lactalbumin, and ribonuclease A, all proteins were dissolved in and dialyzed against doubly distilled water. The β-lactoglobulin protein was dissolved in and dialyzed against a buffer (pH 7.0) consisting of 10 mM monobasic sodium phosphate-dibasic sodium phosphate and 10 mM NaCl. The α-lactalbumin protein was dissolved in and dialyzed against a buffer (pH 6.0) consisting of 10 mM cacodylic acid-sodium cacodylate and 2 mM $CaCl_2$ to ensure the holo-form. Ribonuclease A was dissolved in and dialyzed against a pH 6.0 buffer consisting of 10 mM cacodylic acid-sodium cacodylate and 10 mM NaCl.

Concentration determinations

The concentrations of conalbumin, pepsin, trypsin, trypsinogen, and ribonuclease A were determined by dry weight analysis. The concentrations of the remaining proteins were determined spectrophotometrically using the following specific extinction coefficients: bovine serum albumin, $\epsilon_{278} = 0.658\ lg^{-1}\ cm^{-1}$ (Noelken & Timasheff, 1967); hemoglobin, $\epsilon_{278} = 1.873\ lg^{-1}\ cm^{-1}$ (Iqball & Verrall, 1987); ovalbumin, $\epsilon_{280} = 0.750\ lg^{-1}\ cm^{-1}$ (Katz & Miller, 1971); α-chymotrypsinogen A, $\epsilon_{282} = 1.97\ lg^{-1}\ cm^{-1}$ (Jackson & Brandts, 1970); α-chymotrypsin, $\epsilon_{280} = 2.00\ lg^{-1}\ cm^{-1}$ (Privalov & Khechinashvili, 1974); β-lactoglobulin, $\epsilon_{278} = 0.960\ lg^{-1}\ cm^{-1}$ (Townend *et al*., 1960); myoglobin, $\epsilon_{409} = 8.99\ lg^{-1}\ cm^{-1}$ (Crumpton & Polson, 1965); α-lactalbumin, $\epsilon_{280} = 2.09\ lg^{-1}\ cm^{-1}$ (Wetlaufer, 1967); lysozyme, $\epsilon_{281.5} = 2.635\ lg^{-1}\ cm^{-1}$ (Sophianopulos *et al*., 1962); cytochrome *c*, $\epsilon_{409} = 8.56\ lg^{-1}\ cm^{-1}$ (Robinson *et al*., 1983). We independently verified all of these ϵ values by dry weight analysis. For the densimetric and acoustic experiments, protein concentrations were about 3 mg/ml.

Determination of the temperature range where the proteins remain in their native state

For each protein, we determined the temperature dependencies of the UV absorbance at 275, 280, 285, 290, and 295 nm to evaluate if they remained native within the temperature range of our acoustic and densimetric measurements. When a protein is in its native state, this temperature dependence is weak, monotonic, and practically linear at all five wavelengths. With this criterion, under the solvent conditions used here, we found that α-chymotrypsinogen A remains native between 18 and 35°C, hemoglobin, α-chymotrypsin, β-lactoglobulin, α-lactalbumin, and ribonuclease A, remain native between 18 and 45°C, and the 9 other proteins studied remain native within the entire temperature range between 18 and 55°C.

Methods

For each protein, we performed density and ultrasonic velocity measurements only at temperatures where the protein remains native. Specifically, we determined solution densities and sound velocities for α-chymotrypsinogen A at 18, 25, and 35°C; for hemoglobin, α-chymotrypsin, β-lactoglobulin, α-lactalbumin, and ribonuclease A at 18, 25, 35, and 45°C; and for all other proteins at 18, 25, 35, 45, and 55°C.

Solution sound velocities were measured by a resonator method (Eggers & Funck, 1973) with a relative precision of $\pm 1 \times 10^{-4}$% at a frequency about 7.5 MHz, using a previously described ultrasonic resonator cell with a minimum sample volume of 0.8 ml (Sarvazyan, 1982). The characteristic of a protein directly derived from such ultrasonic measurements is the relative specific sound velocity increment, $[u]$, which is equal to $(U - U_0)/(U_0 c)$. In this expression, c is the specific concentration of a solute (equal to the ratio of the solute molar concentration to its molecular weight); U and U_0 are the sound velocities in the solution and the solvent, respectively.

All densities were measured with a precision of $\pm 1.5 \times 10^{-6}\ g/cm^3$ using a vibrating tube densimeter (DMA-60, Anton Paar, Austria). The apparent specific volume, φV, of each protein was calculated from the density data using the well known relationship $\varphi V = 1/\rho_0 - (\rho - \rho_0)/(\rho_0 c)$, where ρ and ρ_0 are the densities of the solution and the solvent, respectively (Kupke, 1973).

The relative specific sound velocity increments, $[u]$, were used in conjunction with the measured apparent specific volume data, φV, to calculate the apparent specific adiabatic compressibility, φK_S, of each protein using the relationship $\varphi K_S = \beta_{S0}(2\varphi V - 2[u] - 1/\rho_0)$, where β_{S0} is the coefficient of adiabatic compressibility of the solvent (Owen & Simons, 1957).

Each densimetric or ultrasonic velocimetric titration experiment was repeated three to five times, with the average values of $[u]$ and φV being used to calculate φK_S.

Acknowledgements

The authors acknowledge the expert technical assistance of Lilit Chalikian in performing all the density and ultrasonic velocity measurements.

References

Abagyan, R. A., Totrov, M. M. & Kuznetsov, D. A. (1994). ICM: a new method for structure modeling and design: application to docking and structure prediction from the distorted native conformation. *J. Comp. Chem.* **15**, 488–506.

Bernstein, F. C., Koetzle, T. F., Williams, G. J. B., Meyer, E. F., Brice, M. D., Rodgers, J. R., Kennard, O., Shimanouchi, T. & Tasumi, M. (1977). The Protein Data Bank: a computer-based archival file for macromolecular structures. *J. Mol. Biol.* **112**, 535–542.

Bull, H. B. & Breese, K. (1973). Temperature dependence of partial volumes of proteins. *Biopolymers,* **12**, 2351–2358.

Chalikian, T. V. & Breslauer, K. J. (1996). Compressibility as a means to detect and characterize globular protein states. *Proc. Natl Acad. Sci. USA,* **93**, 1012–1014.

Chalikian, T. V., Sarvazyan, A. P. & Breslauer, K. J. (1993). Partial molar volumes, expansibilities, and compressibilities of α, ω-aminocarboxylic acids in aqueous solutions between 18 and 55°C. *J. Phys. Chem.* **97**, 13017–13026.

Chalikian, T. V., Sarvazyan, A. P. & Breslauer, K. J. (1994a). Hydration and partial compressibility of biological compounds. *Biophys. Chem.* **51**, 89–109.

Chalikian, T. V., Sarvazyan, A., Gindikin, V. S. & Breslauer, K. J. (1994b). Hydration of charged atomic groups of biological molecules. *Biophys. J.* **66**, A65.

Chalikian, T. V., Gindikin, V. S. & Breslauer, K. J. (1995). Volumetric characterizations of the native, molten globule, and unfolded states of cytochrome *c* at acidic pH. *J. Mol. Biol.* **250**, 291–306.

Chalikian, T. V., Gindikin, V. S. & Breslauer, K. J. (1996). Spectroscopic and volumetric investigation of cytochrome c unfolding at alkaline pH: Characterization of the base-induced unfolded state at 25°C. *FASEB J.* **10**, 164–170.

Conolly, M. L. (1983). Analytical molecular surface calculation. *J. Appl. Crysallog.* **16**, 548–558.

Crumpton, M. J. & Polson, A. (1965). A comparison of the

conformation of sperm whale metmyoglobin with that of apomyoglobin. *J. Mol. Biol.* **11**, 722–729.
Eggers, F. & Funck, Th. (1973). Ultrasonic measurements with milliliter liquid samples in the 0.5–100 MHz range. *Rev. Sci. Instrum.* **44**, 969–978.
Foygel, K., Spector, S., Chatterjee, S. & Kahn, P. (1995). Volume changes of the molten globule transitions of horse heart ferricytochrome *c*: a thermodynamic cycle. *Protein Sci.* **4**, 1426–1429.
Gavish, B., Gratton, E. & Hardy, C. J. (1983). Adiabatic compressibility of globular proteins. *Proc. Natl Acad. Sci. USA*, **80**, 750–754.
Gekko, K. & Hasegawa, Y. (1986). Compressibility-structure relationship of globular proteins. *Biochemistry*, **25**, 6563–6571.
Gekko, K. & Hasegawa, Y. (1989). Effect of temperature on the compressibility of native globular proteins. *J. Phys. Chem.* **93**, 426–429.
Gekko, K. & Noguchi, H. J. (1979). Compressibility of globular proteins in water at 25°C. *Phys. Chem.* **83**, 2706–2714.
Gromiha, M. M. & Ponnuswamy, P. K. (1993). Relationship between amino acid properties and protein compressibility. *J. Theoret. Biol.* **165**, 87–100.
Hinz, H.-J., Vogl, Th. & Meyer, R. (1994). An alternative interpretation of the heat capacity changes associated with protein unfolding. *Biophys. Chem.* **52**, 275–285.
Iqball, M. & Verrall, R. E. (1987). Volumetric properties of aqueous solutions of bovine serum albumin, human serum albumin, and human hemoglobin. *J. Phys. Chem.* **91**, 1935–1941.
Iqball, M. & Verrall, R. E. (1988). Implications of protein folding. Additivity schemes for volumes and compressibilities. *J. Biol. Chem.* **263**, 4159–4165.
Jackson, W. M. & Brandts, J. F. (1970). Thermodynamics of protein denaturation. A calorimetric study of the reversible denaturation of chymotrypsinogen and conclusions regarding the accuracy of the two-state approximation. *Biochemistry*, **9**, 2294–2301.
Katz, S. & Miller, J. E. (1971). Structure-volume relationships for proteins. Comparative dilatometric study of acid-base reactions of lysozyme and ovalbumin in water and denaturing media. *Biochemistry*, **10**, 3569–3574.
Kharakoz, D. P. (1989). Volumetric properties of proteins and their analogs in diluted water solutions. 1. Partial volumes of amino acids at 15–55°C. *Biophys. Chem.* **34**, 5634–5642.
Kharakoz, D. P. (1991). Volumetric properties of proteins and their analogs in diluted water solutions. 2. Partial adiabatic compressibilities of amino acids at 15–70°C. *J. Phys. Chem.* **95**, 5634–5642.
Kharakoz, D. P. (1992). Partial molar volumes of molecules of arbitrary shape and the effect of hydrogen bonding with water. *J. Sol. Chem.* **21**, 569–595.
Kharakoz, D. P. & Mkhitaryan, A. G. (1986). Change in compressibility of cytochrome *c* during oxidation-reduction transition. *Mol. Biol.* **20**, 312–321.
Kharakoz, D. P. & Sarvazyan, A. P. (1993). Hydrational and intrinsic compressibilities of globular proteins. *Biopolymers*, **33**, 11–26.
Kuntz, I. D., Jr & Kauzmann, W. (1974). Hydration of proteins and polypeptides. *Advan. Protein Chem.* **28**, 239–345.
Kupke, D. W. (1973). Density and volume change measurements. In *Physical Principles and Techniques of Protein Chemistry*. (Leach, S. J., ed.), Part C, pp. 1–75, Academic Press, New York and London.
Lee, B. K. & Richards, F. M. (1971). The interpretation of protein structures: Estimation of static accessibility. *J. Mol. Biol.* **55**, 397–400.
Lee, J. C., Gekko, K. & Timasheff, S. N. (1979). Measurements of preferential solvent interactions by densimetric techniques. *Methods Enzymol.* **61**, 26–49.
Momany, F. A., McGuire, R. F., Burgess, A. W. & Sheraga, H. A. (1975). Energy parameters in polypeptides. VII. Geometric parameters, partial atomic charges, nonbonded interactions, hydrogen bond interactions, and intrinsic torsional potentials for the naturally occurring amino acids. *J. Phys. Chem.* **79**, 2361–2381.
Noelken, M. E. & Timashev, S. N. (1967). Preferential solvation of bovine serum albumin in aqueous guanidine hydrochloride. *J. Biol. Chem.* **242**, 5085–5085.
Nölting, B. & Sligar, S. G. (1993). Adiabatic compressibility of molten globules. *Biochemistry*, **32**, 12319–12323.
Onori, G. (1988). Ionic hydration in sodium chloride solutions. *J. Chem. Phys.* **89**, 510–516.
Onori, G. & Santucci, A. (1990). Volumetric properties of 1-1 aqueous electrolyte solutions and ionic hydration. *J. Chem. Phys.* **93**, 2939–2945.
Otting, G. & Liepinsh, E. (1995). Protein hydration viewed by high-resolution NMR spectroscopy: implications for magnetic resonance image contrast. *Acc. Chem. Res.* **28**, 171–177.
Otting, G., Liepinsh, E. & Wüthrich, K. (1991). Protein hydration in aqueous solution. *Science*, **254**, 974–980.
Owen, B. B. & Simons, H. L. (1957). Standard partial molal compressibilities by ultrasonics. 1. Sodium chloride and potassium chloride at 25°C. *J. Phys. Chem.* **61**, 479–482.
Pierotti, R. A. (1965). Aqueous solutions of nonpolar gases. *J. Phys. Chem.* **69**, 281–288.
Privalov, P. L. & Khechinashvili, N. N. (1974). A thermodynamic approach to the problem of stabilization of globular protein structure: a calorimetric study. *J. Mol. Biol.* **86**, 665–684.
Reiss, H. (1965). Scaled particle methods in the statistical thermodynamics of fluids. *Advan. Chem. Phys.* **9**, 1–84.
Richards, F. M. (1977). Areas, volumes, packing, and protein structure. *Annu. Rev. Biophys. Bioeng.* **6**, 151–176.
Robinson, J. B., Jr, Strottmann, J. M. & Stellwagen, E. (1983). A globular high spin form of ferricytochrome *c*. *J. Biol. Chem.* **258**, 6772–6776.
Sarvazyan, A. P. (1982). Development of methods of precise ultrasonic measurements in small volumes of liquids. *Ultrasonics*, **20**, 151–154.
Shrake, A. & Rupley, J. A. (1973). Environment and exposure to solvent of protein atoms. Lysozyme and insulin. *J. Mol. Biol.* **79**, 351–371.
Sophianopoulos, A. J., Rhodes, C. K., Holcomb, D. N. & van Holde, K. E. (1962). Physical studies of lysozyme. 1. Characterization. *J. Biol. Chem.* **237**, 1107–1112.
Stillinger, F. H. (1973). Structure in aqueous solutions of nonpolar solutes from the standpoint of scaled-particle theory. *J. Sol. Chem.* **2**, 141–158.
Tamura, Y. & Gekko, K. (1995). Compactness of thermally and chemically denatured ribonuclease A as re-

vealed by volume and compressibility. *Biochemistry*, **34**, 1878–1884.

Totrov, M. & Abagyan, R. (1996). Contour-buildup algorithm to calculate the analytical molecular surface of proteins. *J. Struct. Biol.* **116**, 138–143.

Townend, R., Winterbotton, R. J. & Timasheff, S. N. (1960). Molecular interactions in β-lactoglobulin. II. Ultracentrifugal and electrophoretic studies of the association of β-lactoglobulin below its isoelectric point. *J. Am. Chem. Soc.* **82**, 3161–3168.

Wetlaufer, D. B. (1967). Osmometry and general characterization of α-lactalbumin. *C. R. Trav. Lab. Carlsberg*, **32**, 125–138.

Ybe, J. A. & Kahn, P. C. (1994). Slow-folding kinetics of ribonuclease A by volume change and circular dichroism: evidence for two independent reactions. *Protein Sci.* **3**, 638–649.

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

Edited by P. E. Wright

(Received 9 February 1996; received in revised form 12 April 1996; accepted 9 May 1996)