Calculation of the Partial Specific Volumes of Proteins in Concentrated Salt, Sugar, and Amino Acid Solutions

Tsutomu ARAKAWA
Amgen, 1900 Oak Terrace Lane, Thousand Oaks, CA 91320, U.S.A.

Received for publication, June 19, 1986

A method for calculating the isopotential partial specific volumes of proteins in concentrated salt, sugar, and amino acid solutions has been developed. It is based on the finding that the preferential hydration of the protein in these solutions is relatively independent of the concentration of the additive and is proportional to the specific surface area of the proteins, i.e., to the ratio of the total accessible surface area to molecular weight. Agreement between the calculated and experimental values was satisfactory, indicating the reliability of the proposed method. These calculations show that the isopotential partial specific volume increases greatly with the concentration of the additive, in particular in the case of Na₂SO₄, (NH₄)₂SO₄ and sucrose, and for smaller proteins.

The physicochemical characterization of proteins requires, at times, working in solvent systems which contain high concentrations of low-molecular-weight compounds (1-3). The proper performance of a number of physicochemical techniques, such as sedimentation equilibrium, in these systems requires knowledge of the preferential hydration of the proteins and their apparent partial specific volume, \( \phi_{2}^{\circ} \), measured at dialysis equilibrium. This last quantity can be greatly different from the true partial specific volume, \( \phi_{2} \), of the protein, as a result of the preferential interactions of the protein with the solvent components (4-7). Whereas \( \phi_{2}^{\circ} \) can be calculated from the amino acid composition of a protein with reasonable accuracy, the isopotential apparent partial specific volume, \( \phi_{2}^{\circ} \), can be obtained only experimentally, e.g., by densimetry, after equilibration of the protein solution with the reference solvent by dialysis (8). This procedure, however, requires relatively large amounts of protein and extensive dialysis, which may not be practical for many proteins because both of stability problems and availability of the purified material. In view of these difficulties it would seem useful to have a simple method for estimating the isopotential apparent partial specific volume in such concentrated solvents. We have now developed such a procedure for concentrated salt, sugar, and amino acid solutions.

MATERIALS AND METHODS

Defining component 1 = water, component 2 = protein, and component 3 = additive, the isopoten-
tial apparent partial specific volume, $\phi'_2^0$, can be calculated from the isomolal partial specific volume, $\phi_2^0$, and the preferential hydration of the protein in the particular solvent, $\xi_1 = \left( \frac{\partial\rho_i}{\partial g_2} \right) T_{i, p_1, \rho_3}$ by Ref. 8,

$$\phi'_2^0 = \phi_2^0 + g_3 \xi_1 \left( \frac{1}{\rho_0} - \bar{\delta}_3 \right)$$

(1)

where $g_i$ is the concentration of component $i$ in grams per gram of water, $T$ is the thermodynamic (Kelvin) temperature, $\mu_i$ is the chemical potential of component $i$, $\rho_0$ is the density of the solvent, and $\bar{\delta}_3$ is the partial specific volume of the additive. The value of $\bar{\delta}_3$ can be measured by densimetry, and it has been reported for a number of compounds (5-7, 9-14). Alternately, the value can be calculated using published data (15, 16). The value of $\phi'_2^0$ can be either determined by densimetry (8) or calculated from the amino acid composition by the method of Cohn and Edsall (17). The value of $\rho_0$ can be measured by densimetry at the given solvent composition, $g_3$, at which needs to be calculated. For many cases, however, $\rho_0$ is available in the literature as a function of solvent composition (15, 16).

The calculation of $\xi_1$ is based on the observation (5-7, 9-13, 18) that, for a number of additives, this parameter (1) is proportional to the specific surface area of proteins, $s_2$ (except at extremes of pH) and (2) is near to independent of the concentration of component 3. (The specific surface area is defined as $S_2/M_2$ where $S_2$ and $M_2$ are the surface area and molecular weight of the protein.) This relation is true for aqueous solvent systems in which component 3 is one of a number of salts (Na$_2$SO$_4$, NaCl, MgSO$_4$, and CH$_3$COONa), sugars (glucose, sucrose, lactose), or amino acids (glycine, alanine, and betaine). This characteristic property of $\xi_1$ arises from the mechanism of the preferential interactions in these solvent systems, as described below.

The value of $\xi_1$ is related to the absolute bindings of water, $A_1$, and additive, $A_3$, by

$$\xi_1 = A_1 - \frac{A_3}{g_3}$$

The protein hydration, $A_1$, depends little on $g_3$. For those cases where the Donnan effect dominates, $A_3$ is negative, in particular at lower $g_3$. When the additive binding to the proteins dominates, $A_3$ is positive. However, it has been shown that for those additives described above, $A_3$ is negligibly small or negative and that neither mechanism dominates in their interactions with the proteins at higher $g_3$ (5-7, 9-13, 19, 20). Instead, those compounds were shown to be excluded from the protein surface due to the surface tension effect (5-7, 9-13, 21). Thus, $\xi_1$ depends on the surface tension increment of the additive, $\left( \frac{\partial\sigma}{\partial m_3} \right)$ and the surface area of the protein, but not on $g_3$, since

$$\xi_1 \propto s_2 \left( \frac{\partial\sigma}{\partial m_3} \right)$$

This equation is derived from the Gibbs adsorption isotherm.

As a result, for these systems the value of $\xi_1$ for any protein B can be calculated from that determined for a specific protein A by

$$\xi_1^B = \xi_1^A \left( \frac{s_2^B}{s_2^A} \right)$$

(2)

where $s_2^B$ and $s_2^A$ are the specific surface areas of proteins B and A, respectively. Since for globular proteins, the surface area of a protein molecule is proportional, within a close approximation, to $M_2^{2/3}$ (22), we obtain $(s_2^B/s_2^A) = (M_2^A/M_2^B)^{1/3}$ and, hence,

$$\xi_1^B = \xi_1^A \left( \frac{M_2^A}{M_2^B} \right)^{1/3}$$

(3)

As is evident, $\xi_1$ is not very sensitive to molecular weight, permitting the use of approximate values of molecular weights in this calculation. Such approximate values may be obtained from, e.g., a sedimentation equilibrium with the use of $\phi'_2^0$, instead of $\phi'_2^0$, or a gel permeation chromatography.

Preferential interactions have been measured for a number of salts, sugars, and amino acids. Typical values are listed in Table I. Values of $\xi_1$ for other compounds can be found in the literature (5-7, 9-13, 18-21, 23). For KBr and KCl, the $\xi_1$ values determined for NaBr and NaCl can be used. Using the data of Table I, the value of $\xi_1$ for other proteins can be calculated by Eq. 3.
### RESULTS

To test the reliability of this method, calculations of $\xi_1$ and from it $\phi_2^\circ$ were carried out for a variety of proteins. The results are summarized in Tables II and III where they are compared with the experimental data. The experimental errors in determination of $\phi_2^\circ$ are usually about $\pm 0.005 \text{ ml/g}$ (24) when protein concentrations of the samples are low (<3 mg/ml). Taking this into account, it is evident that the agreement is good.

Let us carry out a sample calculation on lysozyme in 1 M NaCl. The values of $g_3$ and $\rho_0$ can be determined experimentally for 1 M NaCl in dilute buffer, or they can be taken from the literature (15, 16), since the contribution of dilute buffer components to $g_3$ and $\rho_0$ are negligible. In this calculation, these values were taken from the literature (16). The value of $\xi_1$ for lysozyme was calculated with Eq. 3 from that value listed in Table I for bovine serum albumin (BSA) ($M_2 = \ldots$).

### TABLE II. Calculation of $\Delta\phi_2^\circ$ and comparison with experimental values for various systems.

<table>
<thead>
<tr>
<th>Additive</th>
<th>Protein</th>
<th>$g_3$ (g/g)</th>
<th>$\rho_0$ (g/ml)</th>
<th>$\xi_1$ (g/g)</th>
<th>$\Delta\phi_2^\circ$ (ml/g)</th>
<th>$\Delta\phi_2^\circ$ (ml/g)</th>
<th>(g3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>NaCl</td>
<td>Lysozyme</td>
<td>0.0638</td>
<td>1.0413</td>
<td>0.42</td>
<td>0.016</td>
<td>0.016</td>
<td>0.0597</td>
</tr>
<tr>
<td>MgSO_4</td>
<td>β-LG</td>
<td>0.111</td>
<td>1.1034</td>
<td>0.46</td>
<td>0.040</td>
<td>0.040</td>
<td>0.121</td>
</tr>
<tr>
<td>CH_3COONa</td>
<td>BSA</td>
<td>0.0869</td>
<td>1.0386</td>
<td>0.33</td>
<td>0.013</td>
<td>0.013</td>
<td>0.865</td>
</tr>
<tr>
<td>β-Alanine</td>
<td>BSA</td>
<td>0.0949</td>
<td>1.0294</td>
<td>0.32</td>
<td>0.009</td>
<td>0.009</td>
<td>0.0949</td>
</tr>
<tr>
<td>α-Alanine</td>
<td>BSA</td>
<td>0.0951</td>
<td>1.0275</td>
<td>0.30</td>
<td>0.009</td>
<td>0.009</td>
<td>0.0951</td>
</tr>
<tr>
<td>Sucrose</td>
<td>Tubulin</td>
<td>0.190</td>
<td>1.0635</td>
<td>0.20</td>
<td>0.012</td>
<td>0.012</td>
<td>0.192</td>
</tr>
<tr>
<td>Sucrose</td>
<td>Tubulin</td>
<td>0.282</td>
<td>1.1366</td>
<td>0.20</td>
<td>0.016</td>
<td>0.016</td>
<td>0.282</td>
</tr>
<tr>
<td>Sucrose</td>
<td>Tubulin</td>
<td>0.470</td>
<td>1.1366</td>
<td>0.20</td>
<td>0.022</td>
<td>0.022</td>
<td>0.435</td>
</tr>
</tbody>
</table>

* From Arakawa and Timasheff (6, 7, 10).  

### TABLE III. Calculation of $\Delta\phi_2^\circ$ for pig heart lactate dehydrogenase in (NH_4)_2SO_4.

<table>
<thead>
<tr>
<th>(M)</th>
<th>$g_3$ (g/g)</th>
<th>$\rho_0$ (g/ml)</th>
<th>$\Delta\phi_2^\circ$ (ml/g)</th>
<th>$\Delta\phi_2^\circ$ (ml/g)</th>
<th>(M)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.309</td>
<td>0.0416</td>
<td>1.0220</td>
<td>0.010</td>
<td>0.006</td>
<td>0.3</td>
</tr>
<tr>
<td>0.633</td>
<td>0.0869</td>
<td>1.0456</td>
<td>0.020</td>
<td>0.015</td>
<td>0.6</td>
</tr>
<tr>
<td>0.971</td>
<td>0.136</td>
<td>1.0691</td>
<td>0.030</td>
<td>0.022</td>
<td>0.9</td>
</tr>
<tr>
<td>1.145</td>
<td>0.163</td>
<td>1.0808</td>
<td>0.035</td>
<td>0.030</td>
<td>1.2</td>
</tr>
</tbody>
</table>

* From Tuengler et al. (24).
68,000). The value of $\varphi_3$ is listed in Table I as a constant. Although for some additives, $\varphi_3$ can have a dependence on the concentration of the additive, the variation is in general less than 10% and can be neglected. Introduction of these values into Eq. 1 leads to

$$\xi_1 g_3 \left( \frac{1}{\rho_0} - \tilde{\epsilon}_3 \right) = 0.016 = \phi'_2 - \phi_2 = \Lambda \phi_2$$

The experimentally determined value of $\Lambda \phi_2$ at $g_3=0.0597$ is 0.016 ml/g, identical to the calculated value. Now, $\phi'_2$ can be obtained simply by adding the calculated $\Lambda \phi_2$ to $\phi_2$ either determined experimentally or calculated. The experimental value of $\phi_2$ for lysozyme is 0.698 ml/g and, hence, $\phi'_2$ in 1 M NaCl should be 0.714 ml/g.

The value of $\xi_1$ for the (NH$_4$)$_2$SO$_4$ system is also given in Table I, since this salt is used as a protein precipitant and also in the medium for X-ray crystallography. This value was calculated from the isopotential partial specific volume data for lobster tail lactate dehydrogenase ($M_2=136,000$), reported by Tuengler et al. (24). The isomolal partial specific volume required to calculate $\xi_2$ was assumed to be equal to the isopotential partial specific volume obtained in the absence of (NH$_4$)$_2$SO$_4$. This is based on the fact that the isomolal partial specific volume is nearly constant in the absence and presence of the additives described here and in the absence of the additives the isopotential value is identical or close to the isomolal value, since in the dilute buffer the preferential interactions are negligible, as has been found for many systems (5-7, 10-12). The values of $\tilde{\epsilon}_3$ and $\rho_0$ were calculated from the literature (16, 24). This calculated value of $\xi_2$ is of the same order of magnitude as that for Na$_2$SO$_4$ measured with BSA (6), in agreement with the similar effectiveness of the two salts as protein precipitant. From the obtained $\xi_1$, $\Lambda \phi_2$ was calculated for pig heart lactate dehydrogenase ($M_2=145,300$). The results are shown in Table III, along with the experimental values (24). In all cases, the calculated values agreed with the measured ones within experimental error.

Calculations were carried out also for BSA and RNase A in KBr, CsCl, and sucrose, since these additives are typical solutes used in density gradient sedimentation. The values of $\phi_2$ are 0.726 for BSA and 0.693 for RNase A. Figure 1 shows dependence of the calculated $\phi'_2$ on the density of the solvent. The $\xi_1$ value for KBr was set equal to that for NaBr. It is seen that, in all cases, the calculated $\phi'_2$ increases greatly with the additive concentration. The increase is greatest for sucrose and smallest for KBr, which is consistent with the extent of preferential hydration in these systems. The effect is stronger for RNase A than for BSA, reflecting the greater interaction for RNase A which has a larger specific surface area, $s_2$, than BSA. The arrows shown in Fig. 1 indicate the density at which these two proteins would form bands if there were no preferential solvent interactions. At these densities and additive concentrations, the expected values of $\phi'_2$ are calculated to be 0.745 ml/g in KBr, 0.774 in CsCl, and 0.794 in sucrose for BSA and 0.730 in KBr, 0.784 in CsCl, and 0.828 in sucrose for RNase A. From this calculation it is clear that these proteins would not band at the densities shown in Fig. 1 but at positions of lower solvent density. The expected banding positions for RNase A are at $\rho_0=1.38$ in KBr, $\rho_0=1.31$ in CsCl, and $\rho_0=1.28$ in sucrose and those for BSA are at $\rho_0=1.34$ in
KBr, \( \rho_0 = 1.31 \) in CsCl, and \( \rho_0 = 1.29 \) in sucrose, i.e., positions at which the reciprocal of \( \phi_2^\circ \) is equal to the density of the solvent.

**DISCUSSION**

The results described here show that the preferential hydration of proteins can be very large in concentrated solutions, in particular for smaller proteins and strongly interacting solutes, such as sucrose. Aune and Timasheff (4) and Tuengler et al. (24) have shown that use of \( \phi_2^\circ \) instead of \( \phi_2^\prime \) can lead to erroneous estimate of the molecular weight of proteins in the analysis of sedimentation equilibrium data obtained in the presence of concentrated salts and hence to incorrect numbers of subunit in oligomeric proteins. This is also true of other physicochemical techniques, e.g., density gradient sedimentation. Therefore, in these techniques it is absolutely required to use \( \phi_2^\circ \).

Although, in general, good agreement is found between the values calculated according to the procedure of this paper and experimental values, deviations can be expected when the proteins have highly asymmetric shapes or in the case of lipoproteins and glycoproteins. In addition, certain enzymes from the extremely halophilic bacteria have shown negative, concentration-dependent values of \( \xi_1 \) (26). These enzymes are unique in that they are stable only in the presence of extremely high salts (26), in contrast to many proteins which are in the native state in dilute salts. It should be pointed out also that the effect of the concentrated salts on \( \phi_2^\circ \) is not due simply to the high ionic strength, since salts such as KBr have much smaller effect on \( \phi_2^\circ \). Therefore, it is totally incorrect to attempt such calculations for one salt from the interaction data measured with another salt.

I thank Dr. S.N. Timasheff for critically reading the manuscript and J. Bennett for typing.

**REFERENCES**