Estimating hydration changes upon biomolecular reactions from osmotic stress, high pressure, and preferential hydration experiments

Seishi Shimizu[†]

York Structural Biology Laboratory, Department of Chemistry, University of York, York YO10 5YW, United Kingdom

Edited by George N. Somero, Stanford University, Pacific Grove, CA, and approved December 1, 2003 (received for review September 11, 2003)

How do we estimate, from thermodynamic measurements, the number of water molecules adsorbed or released from biomolecules as a result of a biochemical process such as binding and allosteric effects? Volumetric and osmotic stress analyses are established methods for estimating water numbers; however, these techniques often yield conflicting results. In contrast, Kirkwood–Buff theory offers a novel way to calculate excess hydration number from volumetric data, provides a quantitative condition to gauge the accuracy of osmotic stress analysis, and clarifies the relationship between osmotic and volumetric analyses. I have applied Kirkwood–Buff theory to calculate water numbers for two processes: (*i*) the allosteric transition of hemoglobin and (*ii*) the binding of camphor to cytochrome P450. I show that osmotic stress analysis may overestimate hydration number changes for these processes.

water | partial molar volume | preferential interaction | hemoglobin | cytochrome P450

Water plays a central role in a wide range of biomolecular processes, from protein folding, stability, and denaturation (1–3) to physiological regulation and allosteric effects (4, 5). Water is involved in these processes in a variety of ways, ranging from direct bridging to collective effects (such as hydrophoic effects) (1–5). The enumeration of water molecules is crucial in order to understand how biomolecular processes work.

Osmotic stress analysis (OSA) aims to estimate the number of water molecules adsorbed (or released) as a result of biomolecular processes (4–10). To do so, osmolytes (such as glycerol and polyethylene glycol, known as protein stabilisers) are added to the system (4–6). Because protein-stabilizing osmolytes, preferentially excluded from protein surfaces (11–15), are not accessible to cavities, grooves, channels, or pockets formed by biomolecules, these regions are subject to osmotic stress (4–6). Osmotic stress and the accompanied change of water activity modulate the equilibrium of the process, and the number of waters adsorbed upon the reaction *in the absence of osmolytes* are enumerated by measuring the change of equilibrium constant with respect to osmotic pressure (4–6). The underlying assumption is that osmolytes are "inert": they neither interact nor act directly on macromolecules because they are excluded (4–6). OSA was first applied to haemoglobin: ≈65 water molecules were inferred to be adsorbed upon the transition from the T state to the R state. This estimation was suggested to be consistent with the change in buried surface area (4, 16). Since then, OSA has been applied to various biomolecular processes, including ion channels, DNA–protein, and carbohydrate–protein interactions (5–10).

In spite of its popularity, the validity of OSA has been debated (6, 11–15). Timasheff (11) pointed out that a cosolvent cannot be both excluded and inert at the same time, because exclusion requires a positive free energy change. He emphasised that this free energy upon exclusion, which is related to the experimentally measurable preferential hydration parameter, is the origin of the osmolyte-induced equilibrium shift (11). It is totally unrelated to the osmotic stress, because osmotic pressure is a colligative property (11). In reply to this criticism, Parsegian *et al.* (6) demonstrated the equivalence between the equations used in OSA and in the preferential hydration analysis. However, this demonstrated equivalence was later questioned (12). It was also suggested that OSA is based on a misinterpretation of the equation of preferential hydration (12, 15). In addition, the number of water molecules enumerated by OSA often depends on the choice of osmolytes (9, 10, 12, 13, 17). This suggests that osmolytes do not always behave as assumed. Moreover, it was suggested that OSA may underestimate the number of water molecules released upon reaction, because OSA ignores the osmolytes present with hydrated water (13). The accuracy, applicability, and validity of OSA have been thus questioned.

Volumetric analysis is another method that estimates the number of water molecules involved in reaction by measuring the changes in partial molar volumes of biomolecules (18, 19). Partial molar volumes are obtained from the change of equilibrium when hydrostatic pressures is applied, in contrast to OSA's "volume of water" determined by the application of osmotic pressure (19). Partial molar volumes can be measured from high pressure experiments (19, 20), as well as densimetry (18). The number of water molecules may be inferred by the use of a two-state model for water (21), or through the estimated water density increment in the vicinity of biomolecules (10, 19).

The number of waters estimated from OSA and from volumetric analyses is often inconsistent (10, 18, 19, 22). For example, a 4-fold difference was seen in the equilibrium dissociation of human IFN-γ (10) and up to a 3-fold difference in the camphor binding of cytochrome P-450 (19). These discrepancies were considered to reflect the different aspects of biomoleculer hydration that osmotic and hydrostatic pressures modulate (19, 20). The lack of a rigorous theoretical framework is the major hindrance in clarifying the underlying relationship.

In this paper, I propose a method to enumerate the number of water molecules taken up by biomolecular processes. The foundation of this method is Kirkwood–Buff (KB) theory (23, 24). Using KB theory offers the unique advantage of determining microscopic characteristics of the solution from measurable thermodynamic quantities. In addition, KB theory connects the preferential hydration parameter and the partial molar volume to spatial integrals that involve the distribution of solvent molecules around the solute. This advantage has been appreciated mainly in the study of small molecules, where solvation (24–26) and volumetric properties (27–29) in multiple-component solutions were explained microscopically. Because KB theory is applicable to molecules of any size or shape (24), it can be applied to any protein and to any osmolyte. I, therefore, combine this theory with the structural information of biomol-

This paper was submitted directly (Track II) to the PNAS office.

Abbreviations: KB, Kirkwood–Buff; OSA, osmotic stress analysis; SASA, solvent-accessible surface area.

[†]E-mail: shimizu@ysbl.york.ac.uk.

ecules to clarify the foundation, validity, and applicability of OSA. An application of KB theory to determine the water numbers associated with (*i*) the allosteric transition of haemoglobin and (*ii*) the binding of camphor to cytochrome P450 indicates that OSA may overestimate the number of water molecules taken up during these processes.

Preferential Hydration Parameter and Partial Molar Volume of Biomolecules

Consider a system that consists of water ($i = 1$), biomolecule ($i = 2$), and osmolyte ($i = 3$) molecules. The preferential hydration parameter, ν_{21}, is expressed by (11, 12)

$$\nu_{21} = \left(\frac{\partial n_1}{\partial n_2}\right)_{T,P,\mu_1} = -\left(\frac{\partial \mu_2}{\partial \mu_1}\right)_{T,P,n_2}, \qquad [1]$$

where n_i and μ_i represent the density (molarity) and chemical potential of species i. ν_{21} can be measured from dialysis equilibrium (11, 12) and sedimentation equilibrium (30).

When 2 undergoes a reaction such as a conformational change, binding, etc., the change of ν_{21}, $\Delta\nu_{21}$,[‡] is directly related to the effect of water activity a_1, or the osmotic pressure Π, on the equilibrium constant, K, of the reaction (11, 12)

$$\left(\frac{\partial \ln K}{\partial \ln a_1}\right)_{T,P,n_2} = -\frac{RT}{V_1}\left(\frac{\partial \ln K}{\partial \Pi}\right)_{T,P,n_2} = \Delta\nu_{21}, \qquad [2]$$

where V_i is partial molar volume of species i.

OSA proposes that the number of water molecules adsorbed or released upon the conformational change is equal to $\Delta\nu_{21}^0$[§] (5, 6). Is this valid? What is its accuracy and applicability? To answer these questions, KB theory is employed to seek the relationship between ν_{21} and the structure of the solution. ν_{21} at infinite dilution of biomolecules is given by KB theory as (24, 25)[¶]

$$\nu_{21} = N_{21} - \frac{n_1}{n_3} N_{23}, \qquad [3]$$

where N_{2i} is the excess number of component i around the biomolecule, defined by (23, 24)

$$N_{2i} = n_i N_A G_{2i} = n_i N_A \int d\vec{r}[g_{2i}(\vec{r}) - 1], \qquad [4]$$

where $g_{2i}(\vec{r})$ is the correlation function between the components 2 and i when they are separated by $\vec{r}$, G_{2i} is often referred to as the KB parameter, and N_A is Avogadros number. Eq. **3** is a rigorous result applicable to molecules of any size or shape (23, 24).

N_{2i} signifies the change in number of component i when a biomolecule is introduced into the system. Two factors contribute to this change. The first is the inaccessibility of solvent molecules to biomolecules to intrinsic (core) volume V_I and thermal volume (volume inaccessible due to thermal motion) V_T (18). This contribution is $-n_iV_E = -n_i(V_I + V_T)$, where V_E is called excluded volume hereafter. The second is the change of solvent–solvent interactions when a biomolecule is introduced into the system. This contribution is often called solvent-reorganization (21, 28). Because OSA aims to measure the change in the number of water in the hydration shell (5, 6, 11), the first contribution that is irrelevant should be subtracted out. This gives the following expression for the excess solvation numbers in the shell[‖]:

$$N'_{2i} = N_{2i} + n_iV_E. \qquad [5]$$

Eqs. **3** and **5** lead to

$$\Delta\nu_{21} = \Delta N'_{21} - \frac{n_1}{n_3}\Delta N'_{23}. \qquad [6]$$

What is the relationship between Eq. **6** and previous theories? Equations identical in form to Eqs. **3** and **6** have been derived previously by Tanford (31), Timasheff (11, 12), Eisenberg (30), Schellman (32), and Record and coworkers (13). In these theories, parameters that correspond to $\Delta N'_{21}$ and $\Delta N'_{23}$ express the changes in effective occupancy of the binding sites (which were assumed to exist on protein surfaces) by water and osmolyte molecules (11, 12, 32). In contrast, KB (24, 25) has clarified the physical meaning of these parameters through Eq. **4**.

Eq. **6** is the foundation of OSA (6). The derivation of this equation by Parsegian *et al.* (5, 6) was based on the Gibbs–Duhem equation: in their notation, the chemical potential of a biomolecule, μ_M, is related to the chemical potential and excess numbers of water (W) and cosolvent (S) according to $d\mu_M = -N_W d\mu_W - N_S d\mu_S$ (5, 6). Timasheff (11, 12) emphasised that N_S and N_W are purely phenomenological parameters representing "site occupancy" (11, 12), which have "no real physical meaning" (12). In order to reconfirm that their N_S and N_W are indeed excess numbers (6), I show that the Gibbs–Duhem equation of Parsegian *et al.*, if slightly corrected, can be derived from a thermodynamic rederivation of KB theory proposed by Hall (33). This rederivation involves reasonable thermodynamic assumptions as outlined below. Consider two parts of a biomolecular solution at infinite dilution: the first part contains a biomolecule, the other part is infinitely far from the biomolecule. The Gibbs–Duhem equations for each part under a constant temperature are

$$0 = dP + d\Pi - n_2 d\mu_2 - n_1^* d\mu_1 - n_3^* d\mu_3 \qquad [7]$$

$$0 = dP - n_1 d\mu_1 - n_3 d\mu_3, \qquad [8]$$

where * represents the average densities around the biomolecules. The densities can be related as (33)

$$n_i^* - n_i = n_2 N_{2i}. \qquad [9]$$

Under the condition of constant pressure ($dP = 0$), Eqs. **5** and **7–9** can be combined with van't Hoff's law ($\Pi = RTn_2$),[††] which leads to

$$d(\mu_2 - RT \ln n_2) = -N'_{21} d\mu_1 - N'_{23} d\mu_3. \qquad [10]$$

The only difference between Eq. **10** and the equation from Parsegian *et al.* is their omission of the $RT \ln n_2$ term. Because Eq. **10** can easily be obtained from Eq. **6** (5, 6), I conclude that the intuitive formulation of Parsegian *et al.* (5, 6) was justified with a slight correction: $d\mu_M$ should be interpreted as the change of the standard chemical potential ($d\mu_2 - RTd \ln n_2$) (33) rather

[‡]Throughout this paper Δ refers to the change upon this reaction.

[§]The superscript 0 represents hereafter the value at $n_3 \to 0$.

[¶]See equation 7 of ref. 25. Although ref. 25 has calculated the quantity that corresponds to ν_{23} rather than ν_{21}, the exchange of subscripts 1 and 3 can transform ν_{23} to ν_{21}, and vice versa. The first two terms of the right hand side of Eq. **3** can be obtained by transforming the subscripts of ref. 25 as $s \to 2$ $c \to 1$ and $w \to 3$. Note that ρ in ref. 25 is equivalent to n of this paper. The approximation seen in equation 7 of ref. 25 refers to the omission of $-RT(\partial \ln n_2/\partial\mu_1)_{T,P,n_2}$, which is, in fact, exactly zero. Therefore, equation 7 of ref. 25 is a rigorous result. Note that specie 2 is at infinitely dilution.

[‖]Similar approaches based on different definitions of V_E have been proposed by Matubayasi *et al.* (28) and Ziekielwicz (26).

[††]Because the protein is at infinite dilution, the dominant contribution to Π is from the first term of osmotic virial expansion, which indeed gives van't Hoff's law. See ref. 15.

Table 1. Solvent-excluded volume and solvent-accessible surface area (SASA) of hemoglobin and P450cam

Structures	R*	R2*	T*	T-LS*	P450cam†	P450†	cam†
PDB ID code‡	1HHO§	1BBB	4HHB§	1HBB	1PHC	2CPP	
V_I¶	68,102	68,598	68,901	68,778	50,210	49,893	134
SASA¶	23,848	24,638	25,005	23,849	16,623	16,822	292

*Conformational states of human hemoglobin. T-LS is the low-salt T state.
†P450cam, camphor-bound cytochrome P450; P450, camphor-free cytochrome P450; cam, camphor.
‡ID code of the Protein Data Bank at Brookhaven National Laboratory.
§Rerefined by modern method from the published structures (G. N. Murshudov, A. A. Vagin, and E. J. Dodson, unpublished work).
¶SASA and intrinsic volume V_I [also known as solvent excluded volume (34)] calculated by MOLECULAR SURFACE PACKAGE (34), using the van der Waals parameters of Chalikian *et al.* (35). The units are Å^3 for V_I and Å^2 for SASA.

than of the chemical potential ($d\mu_2$). However, this correction does not affect Eq. **6** or Eq. **3**: the terms originated from RT ln n_2 in these equations are exactly zero, because $(\partial \ln n_2/\partial\mu_1)_{T,P,n_2} = 0$.

In addition to preferential hydration parameter, partial molar volume of biomolecules, V_2, has given many insights into the hydration of biomolecules (18, 19). V_2 is related to how the hydrostatic pressure modulates the equilibrium:

$$-RT\left(\frac{\partial \ln K}{\partial P}\right)_{T,N_i} = \Delta V_2. \qquad [11]$$

KB theory gives the following expression for V_2 (24)‡‡:

$$V_2 = -V_1 N_{21} - V_3 N_{23} + RT\kappa_T, \qquad [12]$$

where the last term (which contains κ_T, isothermal compressibility) is negligible for biological macromolecules§§ (18). The change of V_2 is expressed by using Eq. **5** as

$$\Delta V_2 = -V_1 \Delta N'_{21} - V_2 \Delta N'_{23} + (n_1 V_1 + n_3 V_3)\Delta V_E. \qquad [13]$$

Eqs. **6** and **13** constitute two independent relationships between $\Delta N'_{21}$ and $\Delta N'_{23}$. Previously, Timasheff emphasised that $\Delta N'_{21}$ and $\Delta N'_{23}$ were indeterminates, because they were coupled to conform to Eq. **6** (11) and Eq. **6** was the only relationship known that connects $\Delta N'_{21}$ and $\Delta N'_{23}$. In contrast, now I have an additional relationship between $\Delta N'_{21}$ and $\Delta N'_{23}$ (Eq. **13**) through another measurable quantity ΔV_2. Therefore, $\Delta N'_{21}$ and $\Delta N'_{23}$ are no longer indeterminates. They can now be determined. They have clear physical meaning as excess solvation numbers defined through Eqs. **4** and **5**.

The Relationship Between Volumetric and Osmotic Stress Analyses

Here, I focus on the limit of zero osmolyte concentration, because the aim of OSA is to measure $\Delta N'^0_{21}$ (4, 5).

At this limit, Eq. **13** reduces to

$$\Delta N'^0_{21} = -n^0_1(\Delta V^0_2 - \Delta V_E), \qquad [14]$$

which is equivalent to two-component KB theory (24, 28, 29). Here, the calculation of $\Delta N'^0_{21}$ requires only the volumetric data because the osmotic data is irrelevant. Note that bulk density n^0_1 is used here, in contrast to the use of water density increments assumed around biomolecules (10, 19) or the two-state model of water (21). This difference arises from having a different goal: previous investigators intended to calculate the coordination number of water (through various models and assumptions), whereas I aim to calculate *excess* hydration number (in which case, the only assumption is in the calculation of ΔV_E).

OSA aims to measure the same $\Delta N'^0_{21}$, as above, by assuming in Eq. **6**

$$\Delta N'^0_{21} = \Delta \nu^0_{21} \qquad [15]$$

for strongly excluded osmolytes (11, 12). Eq. **15** is merely an assumption, in contrast to Eq. **14**, which is rigorous other than the calculation of V_E via volumetric analysis (18).

What does this assumption imply? Because the assumption immediately leads to $((n_1/n_3)\Delta N'_{23})'^0 = 0$ in Eq. **6**, the following condition is derived:

$$\Delta G^0_{23} = -\Delta V^0_E. \qquad [16]$$

Molecular crowding studies have shown that second cross virial coefficient $B_{23} = -G^0_{23}$ (24) signifies the volume inaccessible to the osmolytes (15), which is larger than the volume inaccessible to water V^0_E. Eq. **16** implies that OSA is valid only when the changes of these two volumes are equal.

Another implication is derived by combining Eqs. **14** and **15**, which leads to

$$V^0_1 \Delta \nu^0_{21} = -(\Delta V^0_2 - \Delta V_E) \qquad [17]$$

and should hold when OSA is accurate.¶¶ This poses a restriction on $\Delta \nu^0_{21}$, which suggests that OSA is indeed a restricted case of preferential hydration (11).

Osmotic Stress Analysis May Overestimate the Number of Water Molecules Taken Up by Biomolecules

As was shown by a rigorous analysis based on KB theory, the estimation of the number of water molecules requires the information on the change of biomolecular structure (see Eq. **14**). OSA, on the other hand, does not require any structural information. However, as was shown in the previous section, OSA contains a restriction on $\Delta \nu_{21}$ (Eqs. **16** or **17**) as a direct consequence of its basic assumption (Eq. **15**): OSA contains an implicit relationship between the change of biomolecular structure (ΔV_E) and thermodynamic quantities ($\Delta \nu_{21}$ and ΔV_2). Now, I apply KB theory to biochemical processes and investigate whether the assumption of OSA (Eq. **15**) is accurate. I reexamine the number of water molecules involved in oxygen uptake by hemoglobin and in the binding of camphor to cytochrome P450. These cases were selected because osmotic and high pressure data are available for both processes.

‡‡See equation 6.7.22 of ref. 24. Subscripts *A*, *B*, *S* in ref. 24 should respectively be transformed to 3, 1, and 2 in order to obtain Eq. **12**. ρ in ref. 24 is equivalent to n in this paper.

§§Eq. **12** can also be derived by combining Eqs. **7** and **8**, and differentiating with respect to *P*.

¶¶OSA therefore assumes that ΔV_E should be equal to the sum of the changes of osmotic volume ($V^0_1\Delta\nu^0_{21}$) (19) and hydrostatic (partial molar) volume (ΔV^0_2).

Table 2. The number of water involved in reactions

Reaction	$\Delta N'^0_{21}$*	$\Delta\nu^0_{21}$†
Hemoglobin: T → R‡	−65§	65 ± 4
Hemoglobin: T → R2‡	−22§	65 ± 4
Hemoglobin: T-LS → R‡	−23§	65 ± 4
Hemoglobin: T-LS → R2‡	20§	65 ± 4
P450cam → P450 + cam¶	7.6 ± 1.8‖	19**

*Excess hydration number change calculated by using Eq. **14**, based on the values tabulated in Table 1, and ΔV^0_2 from high-pressure experiments. ΔV_E is the sum of ΔV_I (from Table 1) and thermal volume ΔV_T (35), where the empirical relation $\Delta V_T = \delta \times \Delta$SASA ($\delta$ = 1.0 Å for proteins and 0.56 Å for camphor) was used (18, 35). The units are mol/mol. The error bar was evaluated following Filfil and Chalikian (36).

†From osmotic stress measurements. Hemoglobin by Colombo *et al.* (4) at 37°C and P450cam by Di Primo *et al.* (22) at 20°C.

‡Candidates for the conformational changes upon hemoglobin's oxygen uptake.

§$\Delta V^0_2 \simeq 0$ (19) is used. Assuming the error of ±5 ml/mol, the error bar to ΔN^0_{21} is no larger than ±4.

¶Camphor release from camphor-bound cytochrome P450.

‖$\Delta V^0_2 = -29 \pm 3$ ml/mol (22) is used.

**Error bar not given in ref. 22. It is expected to have accuracy similar to other data in the table.

Tables 1 and 2 summarise my calculations based on Eq. **14**. The calculation of ΔV_E requires ΔV_I and ΔSASA (i.e., the change of solvent-accessible surface area) calculated from structural data (see Table 2). They are summarised in Table 1. For hemoglobin, the conformations for deoxygenated and oxygenated structures are still under debate (37, 38). Therefore, I have followed LiCata and Allewell (17) and calculated ΔV_Es under all possible conformational changes between the deoxygenated and oxygenated structures determined thus far (see Table 1). The calculated number of adsorbed water molecules upon haemoglobin's oxygen uptake varies from −65 to +20 (Table 2) smaller than +65 ± 4 (4) inferred from OSA. It is emphasised that the multiple results of $\Delta N'_{21}$ obtained from my theory is due to the uncertainty regarding the structures of deoxygenated and oxygenated states (37, 38). The expulsion of camphor from cytochrome P450cam, according to Eq. **14** and structural information (Table 1), is accompanied by the absorption of 7.6 ± 1.8 water molecules, smaller than the OSA's estimation of 19 (22). My estimation is much closer than OSA's to the inferrence from the cavity structure, about 6 (19).

In both cases, OSA overestimated the number of water molecules taken up during the reaction. From a traditional solvent binding perspective, one may expect that OSA-based estimation would provide a lower bound of the number of waters, considering the estimated release of 1 osmolyte molecule [from $\Delta\nu_{23} \equiv (\partial n_3/\partial n_2)_{T,P,\mu_3}$] upon the oxygen uptake of hemoglobin (4). However, this expectation involves several assumptions that require careful examination. The first assumption is $\Delta N'_{23} = \Delta\nu_{23}$, whose consequence is similar to that of the assumption (Eq. **15**) for $\Delta\nu_{21}$ (which leads to the restriction on $\Delta\nu_{23}$). The second assumption is that N'_{23} is the number of osmolytes bound to protein surfaces. It will be shown below that N'_{23} may be negative for strongly excluded osmolytes. Therefore, the effect of $\Delta N'_{23}$ on $\Delta\nu_{21}$ requires a careful treatment.

Molecular crowding may be used to explain the reason why OSA may overestimate the number of water molecules adsorbed during biochemical processes (15). For cytochrome P450, Eq. **6** gives $((n_1/n_3)\Delta N'_{23})^0 = n^0_1(\Delta G^0_{23} + \Delta V^0_E) = -11.4$ (from $\Delta\nu_{21} = 19$ and $\Delta N'_{21} = 7.6$ of Table 2), which is indeed far from the OSA assumption, $((n_1/n_3)\Delta N'_{23})^0 = 0$. Why is the increase of hydration upon camphor release accompanied by a decrease in osmolyte solvation? As in the previous section, $-G^0_{23}$ for osmolytes can be approximated by the effective volume (V^{osm}_{exc}) around the protein from which osmolytes are excluded. When protein hydration increases upon reaction ($\Delta N'_{21}$ is positive), the osmolytes are more excluded from the proteins. Therefore, $V^0_E - V^{osm}_{exc}$ becomes more negative upon reaction, and thus $\Delta V^0_E - \Delta V^{osm}_{exc} < 0$. Although this increase of exclusion may be expressed phenomenologically by the local-bulk partition coefficient (13), the above argument is more realistic because it is based on the direct consideration of solution structure, which does not require postulates used in the local-bulk partition model such as local and bulk domains, partition coefficients, and SASA (13).

Combination of preferential hydration, volumetric, and structural data gives a further support to the above argument. Let us consider ribonuclease A in an aqueous trehalose solution for an example. Trehalose is an osmolyte whose degree of "exclusion" from protein surfaces lies between glycerol and betaine (13). Solving Eqs. **3**, **5**, and **12** with experimental values (ν_{21} = 863.1 mol/mol in molarity scale, V_1 = 18.02, V_2 = 9,558, V_3 = 206.2 ml/mol) at n_3 = 0.2 mol/liter [i.e., (n_1/n_3) = 266.1] (39) and structural information (V_E = 13,016 ml/mol) (34), I obtain $N'_{21} = 218$, $(n_1/n_3)N'_{23} = -645.3$ mol/mol (N'_{21} and N'_{23} are the excess numbers defined through Eqs. **4** and **5**). Because proteins take more compact structures in osmolyte solutions (40), the estimated $(n_1/n_3)N'_{23}$ constitutes an upper bound. Therefore, this example demonstrates that $(n_1/n_3)N'_{23}$ may indeed be large and negative, consistent with the discussion based on molecular crowding (15) presented above. A similar trend is seen for proteins in other osmolytes.

The above discussions shed light on a stumblingblock of the simple solvent binding perspective: even when the osmolytes are "strongly excluded" from a protein surface, $((n_1/n_3)\Delta N'_{23})^0$ may take a large negative value. Therefore, exclusion of osmolytes may be reflected in a large negative excess solvation number indicating that the exclusion is not merely a lack of binding as has been assumed in OSA.

Conclusion

I have demonstrated in this paper that KB theory gives a clear, unified perspective on osmotic and volumetric experiments. These experiments provide complementary information on excess solvation numbers of osmolytes and water. In the absence of osmolytes, volumetric data alone (with structural data) can determine excess hydration number in the shell. The condition of OSA's applicability was derived. A relationship between preferential hydration parameter, partial molar volume, and excluded volume was clarified. Moreover, the analysis of this paper gave clear support to the suggestion that OSA is a restricted case of preferential hydration (11, 12). I have shown that OSA may overestimate the number of water molecules taken up, because the effect due to strong exclusion of osmolytes from protein surfaces is not negligible. In conclusion, I have proposed a clear, general method to calculate the number of water molecules taken up during a biochemical reaction from the structural and partial molar volume changes. This paves the way toward the clarification of the long-standing debate about the role of water on a wide variety of molecular recognitions and allosteric regulations.

Note Added in Proof. A recent volumetric study on the dissociation of glucose-hexokinase complex (41) gives further support to the conclusion of this article that OSA may, through a large negative $(n_1/n_3)\ \Delta N^0_{23}$, overestimate the number of waters adsorbed. In this example, $\Delta\nu^0_{21} = 326$ via OSA (42), whereas ΔN^0_{21} (calculated via Eq. **14**) is 33.

This paper is dedicated to the memory of Professor Tatsuo Ooi. I am grateful to Tigran Chalikian, Eleanor Dodson, Guy Dodson, and Sheena Radford for stimulating discussions, and to Miguel Ortiz-Lombardia and Tim Kirk for help with computation. I am indebted to Chandra Boon, Colin Kleanthous, and David Goodall for numerous detailed suggestions. This work was supported in part by the Innovation and Research Priming Fund of the University of York.

1. Kauzmann, W. (1959) *Adv. Protein Chem.* **14,** 1–63.
2. Dill, K. A. (1990) *Biochemistry* **29,** 7133–7155.
3. Shimizu, S. & Chan, H. S. (2001) *J. Am. Chem. Soc.* **123,** 2083–2084.
4. Colombo, M. F., Rau, D. C. & Parsegian, V. A. (1992) *Science* **256,** 655–659.
5. Parsegian, V. A., Rand, R. P. & Rau, D. C. (1995) *Methods Enzymol.* **259,** 43–94.
6. Parsegian, V. A., Rand, R. P. & Rau, D. C. (2000) *Proc. Natl. Acad. Sci. USA* **97,** 3987–3992.
7. Sidorova, N. Y. & Rau, D. C. (1996) *Proc. Natl. Acad. Sci. USA* **93,** 12272–12277.
8. Robinson, C. R. & Sligar, S. G. (1998) *Proc. Natl. Acad. Sci. USA* **95,** 2186–2191.
9. Brown, M. P., Grillo, A. O., Boyer, M. & Royer, C. A. (1999) *Protein Sci.* **8,** 1276–1285.
10. Webb, J. N., Webb, S. D., Cleland, J. L., Carpenter, J. F. & Randolph, T. W. (2001) *Proc. Natl. Acad. Sci. USA* **98,** 7259–7264.
11. Timasheff, S. N. (1998) *Proc. Natl. Acad. Sci. USA* **95,** 7363–7367.
12. Timasheff, S. N. (2002) *Proc. Natl. Acad. Sci. USA* **99,** 9721–9726.
13. Courtenay, E. S., Capp, M. W., Anderson, C. F. & Record, M. T., Jr. (2000) *Biochemistry* **39,** 4455–4471.
14. Tang, K. E. S. & Bloomfield, V. A. (2002) *Biophys. J.* **82,** 2876–2891.
15. Davis-Searles, P. R., Saunders, A. J., Erie, D. A., Windsor, D. J. & Pielak, G. J. (2001) *Annu. Rev. Biophys. Biomol. Struct.* **30,** 271–306.
16. Lesk, A. M., Janin, J., Wodak, S. & Chothia, C. (1985) *J. Mol. Biol.* **183,** 267–270.
17. LiCata, V. J. & Allewell, N. M. (1998) *Methods Enzymol.* **295,** 42–63.
18. Chalikian, T. V. (2003) *Annu. Rev. Biophys. Biomol. Struct.* **32,** 207–235.
19. Kornblatt, J. A. & Kornblatt, M. J. (2002) *Biochim. Biophys. Acta* **1595,** 30–47.
20. Robinson, C. R. & Sligar, S. G. (1995) *Method Enzymol.* **259,** 395–427.
21. Chalikian, T. V. (2001) *J. Phys. Chem. B* **105,** 12566–12578.
22. Di Primo, C., Deprez, E., Hui Bon Hoa, G. & Douzou, P. (1995) *Biophys J.* **68,** 2056–2061.
23. Kirkwood, J. G. & Buff, F. P. (1951) *J. Chem. Phys.* **19,** 774–777.
24. Ben-Naim, A. (1992) *Statistical Thermodynamics for Chemists and Biochemists* (Plenum, New York), pp. 372–382, 448–457.
25. Chitra, R. & Smith, P. E. (2001) *J. Phys. Chem. B* **105,** 11513–11522.
26. Zielkiewicz, J. (1995) *J. Phys. Chem.* **99,** 4787–4793.
27. Harano, Y., Imai, T., Kovalenko, A., Kinoshita, M. & Hirata, F. (2001) *J. Chem. Phys.* **114,** 9506–9511.
28. Matubayasi, N. & Levy, R. M. (1996) *J. Phys. Chem.* **100,** 2681–2688.
29. Paulaitis, M. E., Ashbaugh, H. S. & Garde, S. (1994) *Biophys. Chem.* **51,** 349–357.
30. Ebel, C., Eisenberg, H. & Ghirlando, R. (2000) *Biophys. J.* **78,** 385–393.
31. Tanford, C. (1969) *J. Mol. Biol.* **39,** 539–544.
32. Schellman, J. A. (1987) *Biopolymers* **26,** 549–559.
33. Hall, D. G. (1971) *Trans. Faraday Soc.* **67,** 2516–2524.
34. Connolly, M. L. (1993) *J. Mol. Graphics* **11,** 139–141.
35. Chalikian, T. V., Totrov, M., Abagyan, R. & Breslauer, K. J. (1996) *J. Mol. Biol.* **260,** 588–603.
36. Filfil, R. & Chalikian, T. V. (2003) *J. Mol. Biol.* **326,** 1271–1288.
37. Tame, J. R. H. (1999) *Trends Biochem. Sci.* **24,** 372–377.
38. Lukin, J. A., Kontaxis, G., Simplaceanu, V., Bax, A. & Ho, C. (2003) *Proc. Natl. Acad. Sci. USA* **100,** 517–520.
39. Xie, G. & Timasheff, S. N. (1997) *Biophys. Chem.* **64,** 25–43.
40. Kim, Y. S., Jones, L. S., Dong, A., Kendrick, B. S., Chang, B. S., Manning, M. C., Randolph, T. & Carpenter, J. F. (2003) *Protein Sci.* **12,** 1252–1261.
41. Filfil, R. & Chalikian, T. V. (2003) *FEBS Lett.* **554,** 351–356.
42. Reid, C. & Rand, R. P. (1997) *Biophys. J.* **72,** 1022–1030.