Looking for the best experimental conditions to
detail the protein solvation shell in a binary aqueous
solvent via Small Angle Scattering

Maria Grazia Ortore¹, Raffaele Sinibaldi¹, Francesco Spinozzi¹,
Andrea Carbini¹, Flavio Carsughi¹,² and Paolo Mariani¹

1 Dipartimento di Scienze Applicate ai Sistemi Complessi e Unità CNISM, Università
Politecnica delle Marche, Ancona, Italy
2 Institut für Festkörperforschung, Forschungszentrum, Jülich, Germany
E-mail: mg.ortore@alisf1.univpm.it

Abstract. Protein hydration features attract particular interest in different fields, from
biology up to physics, crossing chemistry and medicine. Particular attention is devoted to
proteins dissolved in binary aqueous mixtures, since the presence of cosolvent can induce
modifications in structural and functional properties. We have recently developed a methodology
to obtain a quantitative description on protein solvation shell by a set of in-solution small
angle scattering experiments, simultaneously analysed by a global-fit approach. In this paper,
numerical simulations of small angle scattering curves are presented to figure out the sensitivity
of the technique to different experimental conditions. Simulations concern two model proteins
of different molecular weights and an unique cosolvent. A reliability test is introduced in order
to find the best experimental conditions to be investigated, together with the most suitable
scattering probe (neutrons or X-rays).

1. Introduction
It is tremendously obvious that hydration, or, more generally speaking, solvation processes play
an essential role in determining structural properties and stability of biological macromolecules
in solution and in controlling their dynamical and functional properties. Indeed, water interfaces
and water interplay with biological material likely retain important keys to understand many
relevant biological processes. However, the whole picture is still largely incomplete from
both biological and physical-chemical points-of-view, and this special issue on water interface
demonstrates the peculiar interest devoted to the subject.

Protein hydration in water and protein preferential solvation in aqueous mixed solutions are
two interesting cases. Considering the importance of this issue, it is pointless to enumerate
the experimental approaches attempting to detail the protein hydration mechanisms or to
describe the characteristics of the hydration shell and the consequent effects on protein stability.
Concerning in particular the water in the first hydration shell, the debated question about
its density was successfully faced in 1998 by Svergun and Zaccai, via a combination of Small
Angle X-Ray and Neutron Scattering techniques [1] (SAXS and SANS, respectively; SAS in
general). The setting of protein first shell density, in experimental conditions quite similar to
those existing in vivo, is without doubts a milestone, although this result had been already predicted by molecular dynamic simulations [2] and crystallographic studies [3].

On the other side, the presence of a cosolvent in the aqueous solution can affect protein stability [4, 5], which can be either increased (see for example the case of glycerol addition in protein aqueous solutions) or reduced (as after the addition of urea). Moreover, the cosolvent can also play an important role in protein-protein association [6], even without modification of their structural characteristics. Pioneering studies related to cosolvent effects on proteins started more than 50 years ago [7, 8], but the issue is still much debated [5, 9, 10] nowadays. In particular, the discussion involves the two main different approaches used to investigate the possible outcome of cosolvent in protein aqueous solutions: the study of the solvent network modifications induced by cosolvents (and their related effect on proteins) [11, 12] and the study of the direct effect of cosolvents on protein surface and structure [13, 14].

Many efforts have been devoted to determine the protein solvation shell composition in the presence of a cosolvent by using different thermodynamic techniques, such as dialysis equilibrium and vapor pressure osmometry [15, 16, 17, 13, 18]. Some thermodynamic inconsistencies in a critique of “osmotic stress” were deeply discussed [9], as well as Kirkwood-Buff integrals and the preferential binding parameter, were recently derived to calculate various properties of infinitely dilute proteins in aqueous mixed solvents [19]. Finally, as cosolvent addition may lead either to protein aggregate formation or to an increased repulsion between particles, protein-protein interactions in the presence of a cosolvent have been also widely investigated [20, 21, 6, 22, 23, 24].

Therefore, the study of protein hydration and protein solvation in aqueous mixed solvent is still requiring new and effective experimental and theoretical efforts. This paper focus on a very recent approach, that we have developed to derive the solvation properties of a protein dissolved in water/cosolvent mixtures by in-solution SAXS and SANS experiments [22, 23]. The method is centered on a global and simultaneous analysis of all the measured SAS curves by using a thermodynamic model for the protein solvation, as the equilibrium exchange between water and cosolvent molecules from the solvation layer to the bulk described by Schellman [25]. Indeed, we have observed that the characteristics of the protein-solvent interface can be described if a sufficiently large number of different experimental conditions (i.e. protein concentration, composition of the aqueous mixed solvent, content of deuterated compounds) are considered. The application of SAS techniques should be commented: the main advantage is the possibility to perform the solvation analysis in a controlled environment, eventually quite similar to that observed in vivo. In fact, the experiments concern proteins dissolved in a solution, at various protein concentrations and in buffer conditions that can approach the cellular environment; by contrast, all the experimental thermodynamic techniques previously considered can be applied mainly at infinite protein dilution. In addition, SAS enables the simultaneous investigation of protein solvation shell composition and of protein-protein interactions.

In this work, we will discuss how the successful application of this SAS “global fitting” strategy to derive quantitative information on the protein solvation shell is strongly related to the number of investigated experimental conditions and to the correct choice of these conditions. The paper is organized in a few sections that, other than summarizing previous results, focus on the transfer of the solvent-exchange model [25] to SAS data analysis. A series of numerical simulations have been performed in order to show a way to select the experimental conditions that, being more sensible to protein solvation features, are more suitable to be investigated by SAS. This point is very important, in particular when experiments are scheduled at neutron or synchrotron facilities in a unique and finite beam-time allocation.
2. In-solution SAS experiments

The macroscopical differential coherent scattering cross section for monodisperse and randomly oriented protein particles dissolved in a solvent is

\[ I(Q) = n_p P(Q) S_{\text{eff}}(Q) \]  

where \( n_p \) is the protein number density, \( S_{\text{eff}}(Q) \) is the effective structure factor and \( P(Q) \) is the protein averaged squared form factor. In this equation, \( Q \) is the modulus of the scattering vector \( (Q = 4\pi \sin \theta / \lambda) \), being \( 2\theta \) the scattering angle and \( \lambda \) the X-rays or neutrons wavelength.

To study protein solvation properties in aqueous mixed solvents, the protein form factor in Eq. 1 should explicitly refer to the protein solvation shell, while \( S_{\text{eff}}(Q) \), which is related to the particle-particle interactions, should be modeled taking into account any possible effect related to the presence of the cosolvent. For example, the cosolvent can change the dielectric properties of the solution, modify the local structure of the protein surface and/or change the amount of counter-ions bound to it [22, 23, 24]. In the next paragraph, we will then show as form and structure factors for the solvated protein can be modeled according to the solvent-exchange thermodynamic process described by Schellman and the protein-protein interaction potential.

3. Thermodynamic model applied to SAS curve analysis

The considered system consists of a globular protein dissolved in an aqueous solution containing a cosolvent. Therefore, its elements are protein \((p)\), water \((w)\) and cosolvent \((o)\) molecules, whose number densities are \( n_p, n_w \) and \( n_o \), respectively. A unique protein conformation is assumed, and the cosolvent molecular volume is thought to be larger than water molecular volume. In the following, we will refer to the binary solvent composition indicating either the water molar fraction in the solvent, \( x_w = n_w / (n_w + n_o) \), or the cosolvent molar concentration, \( C_o = n_o / N_A \) \( (N_A \) is Avogadro’s number). On the other side, the protein concentration is defined by \( c_p = n_p M_p / N_A \), being \( M_p \) the protein molecular weight.

According to Schellman’s approach [5], the protein molecule is surrounded by a solvent region (the solvation layer), whose composition can be different from the one of the bulk solvent (see Fig. 1). Moreover, water and cosolvent molecules are not strongly bound to any specific binding site, so that the exchange between a cosolvent and a water molecule involves a few \( k_B T \) of free energy \( (k_B \) is Boltzmann’s constant, \( T \) the absolute temperature). In this view, a large cosolvent molecule in contact with the protein would preserve its own conformation flexibility, so that multiple contacts with the protein can be considered rare and negligible events [26]. Finally, it is assumed that the number of protein contact sites, here indicated by \( m \), does not depend on the solvent composition and can be simply considered as the number of water molecules that belong to the first hydration shell \((h)\) when \( x_w = 1 \). It is well-established that the thickness of the first hydration shell around a protein is about 3 Å [5]; hence, \( m \) can be calculated by the ratio between the volume of the first hydration shell, \( V_h \), and the volume of the water molecules in contact with the protein, \( v_{w,h} \).

Combining all these assumptions, we can write the “mild” 1:1 cosolvent/water \((O/W)\) exchange equilibrium process over a single unspecific contact site \((P_s)\) and its thermodynamic constant as

\[ P_s O + W \rightleftharpoons P_s W + O \]

\[ K = \frac{x_{w,h}}{1 - x_{w,h}} \frac{1 - x_{w,b}}{x_{w,b}} \frac{f_o}{f_w} \]  

In this equation \( x_{w,h} \) represents the probability that sites, on average, are occupied by water, \( x_{w,b} \) is the water molar fraction in the bulk, and \( f_w \) and \( f_o \) are water and cosolvent activity coefficients in the bulk (in molar fraction units).

Small-angle scattering techniques are particularly suitable to investigate the chemical-physical properties of such a system. In fact, according to the SAS experimental low-resolution, bulk
Figure 1. Scaled representations of a solvated BSA molecule based on PDB structure 1BKE, dissolved in a cosolvent/water mixture (C_o = 1 M) for K = 1/2 (left side) and K = 2 (right side). The light- and dark-blue circles represent water molecules (with diameter 3 Å) in the bulk and at contact with the protein, respectively. Three connected circles represent cosolvent molecules: those in the bulk are yellow, while those in contact with the protein are brown. The red line delimits the local domain shell with thickness δ_l = 4.2 Å.

Solvent and protein molecules can be considered as homogeneous scattering materials, each with its own scattering length density, $\rho_s$ and $\rho_p$, respectively. In addition, a third region, between protein and bulk, may be visible by SAS if the protein is preferentially solvated ($K \neq 1$ in Eq. 2) and/or if solvent molecules in contact with protein undergo a variation of their volume. We refer to this region as local domain ($l$), whereas $\rho_l$ is the corresponding scattering length density (see the red line in Fig. 1). Note that bulk solvent molecules fill the gaps among long cosolvent molecules hanged to the protein; therefore, both thickness and volume of local domain ($\delta_l$ and $V_l$, respectively) are larger than those of the first hydration shell and then, as a consequence, the local domain composition, $x_{w,l}$, is in-between $x_{w,h}$ and $x_{w,b}$. In particular, have we derived the following relations [22],

$$x_{w,l} = \begin{cases} \frac{x_{w,b} \nu_{w,b} - x_{w,h} (\nu_{w,h} - \nu_{w,b}) + x_{w,l} (\nu_{w,h} - \nu_{w,b})}{x_{w,b} (\nu_{w,h} - \nu_{w,b}) + V_l} & \text{for } x_{w,h} > \frac{m \nu_{w,b}}{m \nu_{o} - V_h} \\ \frac{x_{w,l}}{x_{w,b} - x_{w,h}} (\nu_{w,h} - \nu_{w,b}) + \nu_{w,b} & \text{otherwise} \end{cases}$$  \hspace{1cm} (3)

$$\nu_{w,l} = \frac{x_{w,h} (x_{w,b} - x_{w,l})}{x_{w,l} (x_{w,h} - x_{w,b})} (\nu_{w,h} - \nu_{w,b}) + \nu_{w,b}$$  \hspace{1cm} (4)

In these equations, $\nu_o$ is the cosolvent molecular volume, that, according to the single-contact assumption, is considered to be the same both in the bulk and in the local domain, while $\nu_{w,b}$ and $\nu_{w,l}$ are the molecular volume of water in the bulk and in the local domains, respectively. Note that because electrostriction effects occurring at the protein surface, the molecular volume of the water in the first hydration shell can be quite different from the one of water in the bulk phase. Therefore, the water molecular volume in the local domain is included between $\nu_{w,b}$ and $\nu_{w,h}$, as indicated in Eq. 4. It should be also noticed that values of $\nu_o$ and $\nu_{w,b}$ can be found in literature as a function of $x_{w,b}$ for many water/cosolvent mixtures [27].

A further equation that, on the basis of protein number density ($n_p$) and core volume ($V_p$), relates the nominal solvent composition to the composition of the bulk and of the local domain,
is reported here,

\[ x_{w,b} = \frac{[1 - n_p(V_p + V_l)][x_{w,l}(\nu_o - \nu_{w,l}) - \nu_o)x_w - n_pV_o(x_w - x_{w,l})}{[1 - n_p(V_p + V_l)][x_{w,l}(\nu_o - \nu_{w,l}) - \nu_o] + n_pV_l(\nu_{w,b} - \nu_o)(x_w - x_{w,b})}. \] (5)

Once the nominal system conditions have been fixed \((n_p \text{ and } x_w \text{ or, alternatively, } c_p \text{ and } C_o)\), when data or hypothesis regarding the dependence of activity coefficients \(f_w \text{ and } f_o\) on the binary solvent composition are available, the set of Eqs. 2, 3, 4 and 5 can be solved and the bulk and local domain water molar fractions \((x_{w,i}, i = b, l)\) can be calculated as a function of the exchange constant \(K\).

The calculation of the domain scattering length densities is then straightforward,

\[ \rho_i = \frac{x_{w,i}(a_w - a_o) + a_o}{x_{w,l}(\nu_o - \nu_{w,l}) + \nu_{a,i}} \quad \text{with } i = b, l \] (6)

\[ \rho_p = \frac{a_o}{V_p} \] (7)

where \(a_j\) is the scattering length of the \(j\) molecule \((j = w, o, p)\). For SAXS, \(a_j = r_e(Z_j - z_j)\), being \(r_e = 0.28 \cdot 10^{-12} \text{ cm}\) the classical electron radius, \(Z_j\) the sum of atomic numbers of all atoms belonging to the \(j\) molecule and \(z_j\) its charge (in proton charge unit e). For SANS, \(a_j\) is the sum over all atoms of the atomic coherent neutron scattering lengths. It should be noticed that in this work, we considered SANS experiments performed on samples prepared using a binary solvent of deuteration grade \(x_D \equiv n_{D^+}/(n_{H^+} + n_{D^+})\), being \(n_{D^+}\) and \(n_{H^+}\) the number densities of all the exchangeable deuterons and protons present in the system, independently on the molecule they origin from. In this context, the scattering length of the \(j\) molecule results \(a_j = a_{j,D} x_D + a_{j,H} (1 - x_D)\), where \(a_{j,D}\) and \(a_{j,H}\) are the sums of the atomic scattering lengths of all the atoms of the \(j\) molecule when all its exchangeable hydrogens are deuterons or protons, respectively [28, 29]. Notice that both \(a_{j,D}\) and \(a_{j,H}\) are modified when non-exchangeable hydrogens of the \(j\) molecules are substituted with deuteriums.

We are now able to write the averaged form factor, \(P_1(Q)\), and the averaged squared form factor \(P(Q)\) of the solvated protein [30, 31, 32]:

\[ P_1(Q) = (\rho_p - \rho_o)V_pP_{1p}(Q) + (\rho_l - \rho_o)V_lP_{1l}(Q) \] (8)

and

\[ P(Q) = (\rho_p - \rho_o)^2V_p^2P_{pp}(Q) + (\rho_l - \rho_o)^2V_l^2P_{ll}(Q) + 2(\rho_p - \rho_o)(\rho_l - \rho_o)V_pV_lP_{pl}(Q) \] (9)

where the single partial form factor \(P_{1i}(Q)\) is the Fourier transform of the probability that a point at distance \(r\) from the protein center lies in the \(i\) domain \((i = p, l)\), while the pair partial form factors \(P_{ii'}(Q) \quad (i, i' = p, l)\) are the Fourier transform of the probability to find a segment of length \(r\) with an end inside the \(i\) domain and the other end inside the \(i'\) domain.

The volumes \(V_i\) and the functions \(P_{1i}(Q)\) and \(P_{ii'}(Q)\) can be calculated with the Montecarlo method previously described [31, 32], considering the protein structure and the local domain thickness \(\delta_i\). Note that, being both \(P_{1i}(Q)\) and \(P_{i'i'}(Q)\) dimensionless quantities \((P_{1i}(0) = P_{i'i'}(0) \equiv 1)\), if the two constrains \((\rho_p - \rho_o)\) and \((\rho_l - \rho_o)\) are fixed, then the relative weight of the second term in Eq. 8, with respect to the first term, mainly depends on the particular ratio \(V_i/V_p\) of the investigated protein. We have also shown that, if the protein volume \(V_p\) changes by a factor \(\gamma\) close to 1 and if we assume subsequent isotropic changes of the protein shape, then the corresponding modifications of \(P_{1i}(Q)\) and \(P_{i'i'}(Q)\) result in a simple scaling on the \(Q\) axis of a factor \(Q\gamma^{1/3}\) [33].
As indicated by Eq. 1, SAS signal also depends on protein-protein interactions. To account for this effect, we need to calculate the structure factor $S(Q)$, which is the Fourier transform of the protein-protein correlation function $g(r)$. The physical origin of this correlation arises from the two-body interaction potential $u(r)$. Here, following Refs. [34, 35], we have written $u(r)$ as a combination of a classical hard sphere (HS) potential and two Yukawa terms, $u(r) = u_{HS}(r) + u_{D}(r) + u_{A}(r)$. The first term defines the contact radius $R$ of the protein,

$$
u_{HS}(r) = \begin{cases} \infty & r \leq 2R \\ 0 & r > 2R \end{cases} \quad (10)$$

while the general expression of the two Yukawa terms contains a constant $B_j$ and an inverse decay length $\kappa_j$,

$$u_j(r) = B_j \exp\left[-\frac{\kappa_j(r-2R)}{r}\right] \quad j = D, A \quad (11)$$

The $u_{D}(r)$ term represents the coulombic screened potential. Therefore, the constant $B_D$ depends on the protein charge $z_p$ and the bulk dielectric constant $\varepsilon$ (which is a function of the bulk solvent composition $x_{w,b}$):

$$B_D = \frac{z_p^2 e^2}{\varepsilon(1 + \kappa_D R)^2} \quad (12)$$

while $\kappa_D$ is the inverse Debye screening length, which is a function of the ionic strength $I_S$ of the solution (due to all microions, e.g. buffer, added salts and protein counterions):

$$\kappa_D = \left(\frac{8\pi e^2 N_A}{k_B T \varepsilon I_S}\right)^{1/2} \quad (13)$$

The $u_{A}(r)$ Yukawa term represents an attractive potential, whose constant $B_A$ can be written in terms of the value $J$ of the potential at protein contact ($r = 2R$)

$$B_A = -2R J \quad (14)$$

and the decay constant $\kappa_A$ is a phenomenological constant.

The calculation of $S(Q)$ requires the solution of the Ornstein-Zernike (OZ) integral equations of the liquid state theory, within some approximate closure relation. Recently, Liu et al. [35] have proposed an efficient and robust solution in the mean spherical approximation (MSA), whose main result is the explanation for the formation of an extra low-$Q$ peak in the $S(Q)$. Here, we are mainly interested in investigating the SAS signal at moderately low and high $Q$, where $P(Q)$ reflects differences in the local domain structures (see Figs. 2-3). Hence, we can follow the simpler approach of Ref. [34], i.e. the HS potential is the reference system, the two Yukawa terms are treated as a perturbation and the OZ is solved in the frame random phase approximation (RPA),

$$S(Q) = \frac{S_{HS}(Q)}{1 + n_p S_{HS}(Q)[U_C(Q) + U_A(Q)]/(k_B T)} \quad (15)$$

where $S_{HS}(Q)$ is the solution of OZ under MSA of the HS potential

$$S_{HS}(Q) = \left\{ 1 - \frac{12\eta \eta (3 - \eta^2) - 2 j_1(2QR)}{(1 - \eta)^4 2QR} \right\}^{-1} \quad (16)$$
being $\eta = \frac{4}{3} R^3 n_p$ the volume fraction of the hard spheres, and $j_1(x)$ the 1st-order spherical Bessel function. $U_j(Q)$ is the Fourier transform of $u_j(r)$, which reads

$$U_j(Q) = 4\pi B_j \frac{\kappa_j \sin(2QR) + Q \cos(2QR)}{Q(Q^2 + \kappa_j^2)}$$

Finally, the combination of form and structure factors gives the experimental SAS observable (see Eq. 1). It should be noticed that the so-called “effective” structure factor $S_{\text{eff}}(Q)$ is defined by

$$S_{\text{eff}}(Q) = 1 + \left[ \frac{P_1(Q)}{P(Q)} \right]^2 [S(Q) - 1]$$

but the ratio $[P_1(Q)]^2 / P(Q)$ is different from 1 only when $QR \gg 1$: therefore, in the $Q$ range usually investigated it can be considered $S_{\text{eff}}(Q) = S(Q)$.

4. Global Fit Strategy

In our terms, global fitting means the contemporary analysis of several scattering curves relative to different protein samples prepared at different solvent composition (e.g., pH, ionic strength, cosolvent content) or measured at different conditions (e.g., temperature, pressure, time) [6, 36, 22, 23]. If a good model for particle structure and interactions is used, and if the model is in some way sensitive to the experimental conditions, the procedure should be able to evidence even small effects on particle features (e.g., hydration or solvation changes, conformational changes, aggregation state, particle and solvation shell densities). Indeed, we have shown that a global fit strategy can be successfully adopted to evaluate the efficiency of a theoretical model to reproduce all the SAS curves obtained from a wide set of different experimental conditions [6, 36].

Generally speaking, the whole scattering data are analyzed according to a unique model, considering a series of fitting parameters that are expected to be common to all samples (e.g., protein structure factor, thickness of solvation shell, molecular volume of water in the bulk, molecular volume of the protein, association or dissociation constants in the case of aggregation equilibria) and a series of parameters which depend on solvent composition (e.g., solvation shell composition, factors which describe the attractive potential, particle charge, ionic strength of the solution, solvent dielectric constant).

Here, we are considering the use of global fitting procedure to estimate the solvation characteristic of a protein dissolved in water in the presence of a cosolvent, according to the solvent exchange model described in the previous paragraph. The main common fitting parameter is then the equilibrium constant $K$, that describes the exchange equilibrium of water and cosolvent molecules between the protein solvation shell and the bulk solvent phase. As clearly expressed in Eq. 2, the $K$ value determines water and cosolvent molar fractions in each domain and hence contributes to the solvated protein $P(Q)$ estimation. Other common fitting parameters are the molecular volume of water in the local domain, the local domain thickness and the protein volume. The protein structure can be described using the PDB crystallographic coordinates, but it should be clear that any structure modification, eventually induced by the cosolvent, has to be accurately investigated. Concerning the protein form factor, the model takes into account uniquely common parameters and the nominal composition of each investigated experimental condition.

More troublesome is the fit of parameters related to the structure factor. In fact, while the dielectric constant of the mixed solvent and the ionic strenght of solution, which play a role in the screened coulombic potential, are known, the protein charge and the attractive potential parameters, $J$ and $\kappa_A$, require a particular attention. Even when the protein dissolved in the
mixture maintains its structure, the surface charge can be modified: it is known that cosolvent molecules, even in small amounts in solution, can bind to specific sites of the protein backbone, locally reducing the barrier for exposing inner protein residues to the solvent, and hence changing the total protein surface charge [37, 38, 39]. In the choice of the fitting parameter, this issue cannot be neglected and the surface charge determination inside the global fitting procedure can be a direct way to assess local modifications induced by a rearrangement of cosolvent and water molecules around protein surface. Furthermore, cosolvents can also play a role in modifying attractive protein-protein potential [40]. Therefore, in some cases it can be very informative to consider interaction parameters as single curve fit parameters and maybe perform complementary experiments to verify the agreement between the obtained results.

5. Experimental set-up and Numerical Simulations
To perform a successful SAS global fitting experiment, it is necessary to assess for the best number of samples and the best sample conditions to be investigated. It is obvious in fact that the quality of results is related to the wideness of the experimental condition set, as the more scattering curves are globally fitted, the less uncertainty in the different common fitting parameters results. However, when SAS experiments are scheduled at neutron or synchrotron facilities in a unique and finite beam-time allocation, the number of samples that can be investigated is finite, and then the best experimental conditions should be preliminary established. Numerical simulations can be very useful to optimize number and experimental conditions on the basis of the available beam-time.

To elucidate this point, we refer to a “real” case, which concerns the analysis by SAS global fit of the preferential solvation of a protein dissolved in a mixed aqueous solution containing a cosolvent. To monitor eventual effects related to protein dimension, two different proteins have been considered, namely lysozyme, LYS, and bovine serum albumine, BSA. On the contrary, a unique cosolvent, with the molecular formula and the physics-chemical properties of urea, has been considered. However, the solvation properties ($K$) of the model cosolvent are considered unknown, so that it can be either preferentially excluded from or preferentially attached to the protein surface. Hence, the determination of the corresponding thermodynamic constant $K$, which can be larger, smaller or equal to unity, is the scope of the SAS experiment, that should be performed on a number of samples as smaller as possible.

In order to establish the best sample conditions from which the thermodynamic exchange constant can be derived by SAS global fit with the best accuracy, extended simulations of SAS curves in very different experimental conditions have been performed by using GENFIT software [41]. In particular, numerical simulations have been performed in the usually accessible $Q$-range from 0.03 to 0.4 Å$^{-1}$, considering the form factor in Eq. 9, using scattering length densities calculated according to Eqs. 6-7. Water molecular volume in the first hydration shell has been considered to be $\nu_{w,h} = 27.3$ Å$^3$ (10% less than the bulk water volume), because of electrostriction effects at the protein surface [1]. For the sake of simplicity, the activity coefficients $f_w$ and $f_o$ have been fixed to unity. The protein and local domain volumes ($V_p$ and $V_l$), as well as the partial form factors $P_{ii}(Q)$, have been calculated on the basis of the protein data bank (PDB) structure and considering a thickness of the local domain $\delta_l = 4.2$ Å. In the case of LYS the entry 6LYZ [42] was used, while for BSA, whose crystalline structure is not yet available, the entry 1BKE [43, 44], corresponding to human serum albumine, was used. We obtained $V_p = 17,200$ Å$^3$ for LYS and $81,000$ Å$^3$ for BSA, in agreement with literature [1, 23]. To model the structure factor, both protein charge and attractive potential parameters were derived from literature data. Assuming that the solvent pH is 7, we considered LYS and BSA charges as $z_p = 5$ and $z_p = 15$, respectively [45, 46]; the ionic strenght of the solutions was considered to be 50 mM and the dielectric constant to vary according to the cosolvent presence in solution [27]. Moreover, the temperature was fixed to $T = 298$ K. Attractive term parameters...
were fixed to $J = 3k_B T$ and $\kappa_A = 0.2 \text{ Å}^{-1}$, in agreement with previous results [34, 22, 23]. It should be finally noticed that all simulations were repeated considering three different $K$ values of $1/2, 1, 2$, which correspond to a protein local domain enriched, indifferent or pauperized in cosolvent with respect to the bulk. In this way, the set of simulated curves can be analyzed in searching for such experimental conditions (namely, $C_o$, $c_p$ and $x_D$, globally indicated by the symbol $i$) in which the SAS curves, corresponding to different values of $K$, are significantly different.

6. Numerical simulation results

The whole sets of SAXS and SANS simulated curves are reported in Figs. 2 and 3. For both LYS and BSA, we calculated 3 ensembles of curves: SAXS from hydrogenated samples and SANS from samples at two deuteration grades $x_D$ 1 and 0.8 ($x_D = 1$ maximizes the contrast between protein and water; $x_D = 0.8$ is the condition where the contrast between water and protein and between the considered cosolvent and water are equal [28]). Each ensemble consists of 5 different protein concentrations $c_p$ (10, 40, 70, 100 and 130 g/l, e.g. from very dilute to highly concentrate conditions), each one at 5 different cosolvent molar concentrations $C_o$ (0.5, 1, 1.5, 2, 2.5 M), resulting in 25 experimental conditions. As for each experimental condition $i$, three different $K$ values were considered ($K = 1/2, 1, 2$), the number of curves simulated for each ensemble is 75. Data in Figs. 2 and 3 show that SAS profiles vary for different $K$ values, and that the differences are mostly pronounced in certain conditions (concentrations, deuteration grade, type of experiment). Note that such variations are especially evident in the lowest and highest $Q$ ranges, as the local domain features reflect on the effective particle shape and the local domain thickness affects the short-range distances probability, respectively.

For each set $i$, to quantitatively characterize the difference between the two curves with $K = 1/2$ and $K = 2$ and the “reference” curve $I_{1}(Q)$ corresponding to $K = 1$, we have introduced the resemblance factor $R_{f,i}^{(i)}$ [47],

\[
R_{f,i}^{(i)} = R_{f,1/2} + R_{f,2} \tag{19}
\]

\[
R_{f,K} = \int_{Q_{\text{min}}}^{Q_{\text{max}}} dQ \left\{ \frac{I_{K}(Q) - I_{1}(Q)}{[I_{K}(Q_{\text{min}})]^{1-\alpha}[I_{K}(Q)]^{\alpha}} \right\}^2 \tag{20}
\]

where $Q_{\text{min}}$ and $Q_{\text{max}}$ are the minimum and the maximum value of the chosen experimental $Q$-range and $I_{K}(Q)$ is the macroscopical differential scattering cross section (Eq. 1) relative to the $K$ value. In the assumption that errors over the SAS detector counts follow a Poisson statistics, it can be found that the errors of $I(Q)$ can be expressed as proportional to $I(Q)^\alpha$, with $\alpha \approx 0.2 - 0.8$, depending on experimental set-up parameters (transmissions, measurement times, etc.). Then, according to the standard $\chi^2$ definition [48], we have chosen the $R_{f,K}$ definition reported in Eq. 20. Here, we fixed $\alpha = 0.5$. According to this definition, $R_{f,i}^{(i)}$ offers a quantitative estimation of how much the experimental condition $i$ is able to evidence a protein local domain composition different from the bulk one: the bigger is $R_{f,i}^{(i)}$, the more convenient will be the $i$-condition. Values of $R_{f,i}^{(i)}$ calculated for the 25 $i$-conditions of each ensemble are reported in Tab. 1 for BSA and in Tab. 2 for LYS. It is evident that $R_{f,i}^{(i)}$ increases with both $c_p,i$ and $C_o,i$.

The first remarkable result concerns the large $R_{f,i}^{(i)}$ observed in SANS simulations with $x_D = 0.8$, proving that sample deuteration can improve SAS sensibility to different $K$ values. However, further simulations (not shown) indicated that this effect is not linear, so that maximizing the sample deuteration grade does not correspond to maximize $R_{f,i}^{(i)}$. Such a result strongly confirms the importance of performing numerical simulations before defining the experimental conditions to be investigated.
Figure 2. SANS (at $x_D = 1$ and 0.8) and SAXS numerical simulations of BSA. In each panel we report results obtained in water/cosolvent mixtures (from left to right: $C_o = 0.5, 1, 1.5, 2$ and $2.5$ M) at different protein concentrations (from bottom to up: $c_p = 10, 40, 70, 100$ and $130$ g L$^{-1}$). Blue, red and green lines refer to $K = 1/2, 1$ and $2$, respectively. For the sake of clarity, in each panel, curves at the same $c_p$ have been scaled upwards by a factor multiple of 1.5.
Figure 3. SANS (at $x_D = 1$ and 0.8) and SAXS numerical simulations of LYS. All other details as in Fig. 2.
As it is also important to verify the applicability of the thermodynamic model (in this case, the Schellman solvent-exchange equilibrium model) and the accuracy of the protein-protein interaction pattern over a range of protein and cosolvent concentrations as wider as possible, a second factor related to the sampling size and design has been introduced. This factor has been called concentration-distribution factor $R_c^{(j)}$

$$R_c^{(j)} = \left\{ \frac{2}{N_j(N_j-1)} \sum_{i \neq i' \in j} \left\{ \left( \frac{c_{p,i} - c_{p,i'}}{c_{p,max} - c_{p,min}} \right)^2 + \left( \frac{C_{o,i} - C_{o,i'}}{C_{o,max} - C_{o,min}} \right)^2 \right\}^{1/2} \right\}^{1/2}, \quad (21)$$

where $N_j$ is the number of experimental $i$ conditions investigated, $c_{p,i}$ and $C_{o,i}$ are protein and cosolvent concentrations relative to the $i$-th condition, respectively. The protein concentration range $(c_{p,max} - c_{p,min})$ and the cosolvent concentration range $(C_{o,max} - C_{o,min})$ have to be wisely chosen by the experimentalist, from one hand to avoid protein aggregation and/or unfolding, and from the other hand to obtain a satisfactory signal/noise ratio. Notice that the more homogeneously distributed $c_{p,i}$ and $C_{o,i}$ are within their corresponding ranges, the higher is the factor $R_c^{(j)}$.

As an indicative example, we have chosen 5 different sets $j$ within each ensemble, each one consisting of $N_j = 5$ $i$-conditions selected from those reported in Tabs. 1 and 2 following a “Union Jack” flag scheme. The different colors in the Tables indicate the different sets. For each set $j$, we have calculated both $R_c^{(j)}$ and the sum of the corresponding resemblance factors.

| $c_p$ (g L$^{-1}$) | 0.5 | 1 | 1.5 | 2 | 2.5 |
|-------------------|-----|---|-----|---|-----|
| $C_o$ (M)         |     |   |     |   |     |
| $R_f^{(j)}$ (10$^{-3}$ Å$^{-1}$) |     |   |     |   |     |
| BSA SANS $x_D=1$  |     |   |     |   |     |
| 10 0.0663         | ![blue] | 0.269 | 0.616 | ![red] | 1.12 | 1.82 |
| 40 0.0802         | ![green] | 0.325 | 0.747 | ![black] | 1.37 | 2.21 |
| 70 0.0920         | ![yellow] | 0.374 | 0.860 | ![lightblue] | 1.57 | 2.55 |
| 100 0.102         | ![magenta] | 0.417 | 0.96  | ![lightgreen] | 1.76 | 2.86 |
| 130 0.112         | ![cyan] | 0.455 | 1.05  | ![purple] | 1.93 | 3.14 |
| BSA SANS $x_D=0.8$|     |   |     |   |     |
| 10 0.0934         | ![blue] | 0.382 | 0.888 | ![red] | 1.65 | 2.72 |
| 40 0.113          | ![green] | 0.463 | 1.08  | ![black] | 2.00 | 3.31 |
| 70 0.130          | ![yellow] | 0.532 | 1.24  | ![lightblue] | 2.31 | 3.82 |
| 100 0.144         | ![magenta] | 0.593 | 1.38  | ![lightgreen] | 2.58 | 4.28 |
| 130 0.158         | ![cyan] | 0.649 | 1.52  | ![purple] | 2.83 | 4.71 |
| BSA SAXS          |     |   |     |   |     |
| 10 0.0112         | ![blue] | 0.0454 | 0.104 | ![red] | 0.187 | 0.295 |
| 40 0.0137         | ![green] | 0.0557 | 0.127 | ![black] | 0.230 | 0.363 |
| 70 0.0159         | ![yellow] | 0.0647 | 0.148 | ![lightblue] | 0.268 | 0.424 |
| 100 0.0179        | ![magenta] | 0.0731 | 0.167 | ![lightgreen] | 0.303 | 0.480 |
| 130 0.0198        | ![cyan] | 0.0809 | 0.186 | ![purple] | 0.336 | 0.533 |

Table 1. Resemblance factors $R_f^{(j)}$ calculated with Eq. 19 and relative to the three ensembles of numerical simulations of BSA reported in Fig. 2. Colored symbols on the right of each factor identify different sets $j$ inside each ensemble. According to a “Union Jack” flag scheme, the different sets are color-coded and defined as: column, row, backslash, star, slash.
Table 2. Resemblance factors $R_f^{(i)}$ calculated with Eq. 19 and relative to the three ensembles of numerical simulations of LYS reported in Fig. 3. See Tab. 1 for all the other details.

|        | $c_p$ (g L$^{-1}$) | 0.5 | 1.0 | 1.5 | 2.0 | 2.5 |
|--------|--------------------|-----|-----|-----|-----|-----|
|        | $C_o$ (M) | $R_f^{(i)}$ (10$^{-3}$ Å$^{-1}$) |     |     |     |     |     |
| LYS SANS $x_D=1$ | 10.0 | 0.312 | 1.24 | 2.81 | 5.08 | 8.21 |
|       | 40.0 | 0.321 | 1.28 | 2.89 | 5.24 | 8.47 |
|       | 70.0 | 0.329 | 1.31 | 2.97 | 5.39 | 8.71 |
|       | 100.0| 0.337 | 1.34 | 3.05 | 5.52 | 8.95 |
|       | 130.0| 0.344 | 1.37 | 3.12 | 5.65 | 9.17 |
| LYS SANS $x_D=0.8$ | 10.0 | 0.445 | 1.79 | 4.10 | 7.56 | 12.6 |
|       | 40.0 | 0.458 | 1.84 | 4.22 | 7.80 | 13.0 |
|       | 70.0 | 0.470 | 1.89 | 4.34 | 8.02 | 13.4 |
|       | 100.0| 0.481 | 1.94 | 4.46 | 8.23 | 13.6 |
|       | 130.0| 0.491 | 1.98 | 4.55 | 8.43 | 14.1 |
| LYS SAXS | 10.0 | 0.0435 | 0.176 | 0.397 | 0.710 | 1.11 |
|       | 40.0 | 0.0486 | 0.196 | 0.445 | 0.796 | 1.25 |
|       | 70.0 | 0.0525 | 0.212 | 0.483 | 0.866 | 1.36 |
|       | 100.0| 0.0557 | 0.226 | 0.516 | 0.925 | 1.46 |
|       | 130.0| 0.0587 | 0.238 | 0.544 | 0.98 | 1.55 |

Looking over the $\sum_{i,j} R_f^{(i)}$, it is straightforward to note that for each protein and for each set $j$, the best experiment for evidencing differences of composition between bulk and protein local domain is SANS performed on samples at deuteration grade $x_D=0.8$ (green point-line). On the contrary, SAXS is the worse experiment (blue point-line), as already evidenced by the bottom curves of Figs. 2 and 3, which do not present remarkable differences for different $K$ values. Instead, it can be considered unexpected that both SANS and SAXS numerical simulations for the smaller protein, LYS, show the higher sensibility (i.e. higher $\sum_{i,j} R_f^{(i)}$ values) to different Ks with respect to the larger protein, BSA. However, it can be easily shown that, at a fixed thickness of the local domain $\delta_l$, the ratio $V_l/V_p$ decreases by increasing $V_p$: hence, as hinted by Eq. 9, the same variation of local domain composition (i.e. the same difference $\rho_l - \rho_b$) leads to a larger relative variation of $P(Q)$ for small proteins with respect to large proteins.

It is also interesting to note that $\sum_{i,j} R_f^{(i)}$ presents different maxima for the different proteins (Fig. 4, left side). In the BSA case, the maximum is observed for the slash set, while for LYS the maximum is observed when the star set is considered. Nevertheless, for both proteins, the two higher values are observed for backslash and star sets, and their difference is small.

However, as the experimental condition choice should also account for the test of applicability of both the thermodynamic and interaction models, the behavior of $R_f^{(i)}$ factor should be considered: for both proteins, the maximum is observed for the star set. To derive a final parameter, we then take the mathematical product of $\sum_{i,j} R_f^{(i)}$ and $R_c^{(j)}$ factors as a combined reliability test. Calculated values are reported in Fig. 4 (right side): in any case, and if only
5 experiments can be performed, the best experimental choice concerns SANS experiments performed on samples prepared at $x_D = 0.8$ with protein and cosolvent concentrations defined by the \textit{star} set. It should be observed that the obtained results can be different if a larger number of experiments can be scheduled or when the concentration limits (both for protein and cosolvent) are different: however, a good choice of experimental conditions can result in SAS curve differences that can be significantly very large (in the presented case, the resemblance factor changes from 6 to 10 times when the considered set moves from column to \textit{star}).

7. Experimental Results
We briefly describe two successful applications of the methodology presented in Sec. 3 for both LYS and BSA. LYS was investigated both in water-glycerol [22] and in water-urea
mixtures [24] by SANS, whereas BSA was investigated in the presence of urea by both SAXS and SANS techniques [23]. Glycerol and urea are two well-known and largely studied cosolvents. However, the molecular mechanisms that govern the stabilizing and destabilizing effects of these cosolvents are still debated. According to literature, stabilizing cosolvents are preferentially excluded from protein solvation shell, while the denaturing chemical agents are preferentially driven to protein surface [49, 50, 51]. Nonetheless, the quantitative estimation of cosolvent molecules located at protein surface was obtained at infinite protein dilution, or in a crystal. This is the reason why our method can be regarded as efficient in determining the existence of a thermodynamic exchange equilibrium in a wide range of protein concentrations and in experimental conditions closer to those in-vivo.

In the first study [22], we have determined the experimental conditions to be investigated via numerical simulations of the form factor of lysozyme in water-glycerol mixtures, considering the protein solvation shell enriched in water. 35 different experimental conditions (protein concentration, water/glycerol molar fraction in the solvent and deuteration grades) have been then experimentally investigated by SANS and analyzed using GENFIT software [41]. A thermodynamic constant $K = 1.87 \pm 0.03$ has been obtained, confirming that lysozyme dissolved in water/glycerol mixed solvents is preferentially hydrated. This result has been compared with literature data concerning the preferential binding coefficient in the case of Ribonuclease A [52, 19], evidencing a good agreement.

In the second study [23], the thermodynamic exchange constant has been determined by SAXS and SANS experiments on BSA dissolved in water-urea solutions. According to extended simulations, 9 SANS curves and 8 SAXS curves, corresponding to different protein and urea concentrations, were simultaneously fitted using GENFIT software [41]. The common parameter $K$ resulted $0.6 \pm 0.1$, indicating that the protein solvation shell is enriched in urea. In this case, the BSA charge has been detected to slightly increase with addition of urea in solution, suggesting a denaturation mechanism in which the main process could be urea localization on the surface of the protein, producing local structural modifications (see also [24]).

It should be finally noticed that a direct relation between the preferential binding coefficients and the thermodynamic constant $K$ was derived (see Eqs. 14-15 in [22]) and that the agreement with results arising from thermodynamic techniques [15, 49, 53, 52, 54, 55, 56] was very satisfactory.

8. Conclusions
In a binary aqueous solvent, the description of protein solvation shell, both in terms of thickness and composition, can be obtained by SAS experiments considering a convenient combination of experimental conditions. This paper offers the detailed guideline to select the type of experiment (SANS or SAXS) and the set of sample conditions to be investigated, in order to obtain successful results using a global fit analysis. As a thermodynamic relation is necessary to connect the investigated samples, the application of Schellman solvent exchange model to the global fit analysis of SAS experimental curves, has been widely described. Extended SAS curve simulations have been then presented and a reliability test to establish the best sample conditions introduced.

Therefore, this paper can be considered a sort of recipe that points the way to perform SAS experiments when the expected result is the determination of a quantitative protein solvation shell description, together with the estimation of protein-protein interactions parameters. Despite previous and successful approaches via thermodynamic techniques [15, 17, 13], we stress the relevance of this method because in-solution SAS concerns experimental environments even close to the in-vivo ones and includes the analysis of protein-protein interactions. Spreading of this methodology can lead to face other noticeable issues, like the effect of cosolvents on protein thermal stability [57], activity and aggregation.
9. Acknowledgments
MGO thanks European Science Foundation for financial support to participate to ESF-FWF conference “Water Interfaces in Physics, Chemistry and Biology: A Multi-Disciplinary Approach”, Obergurgl, Austria, 8-13 December 2007.

Appendix A. List of symbols

| Symbol | Description |
|--------|-------------|
| $\alpha$ | exponent of the power-law of errors on $I(Q)$ |
| $a_j$ | scattering length of the $j$ molecule ($j = w, o, p$) |
| $a_{j,D}$ | sum of the atomic scattering lengths of all the atoms of the $j$ molecule ($j = w, o, p$) when all its exchangeable hydrogens are deuterons |
| $a_{j,H}$ | sum of the atomic scattering lengths of all the atoms of the $j$ molecule ($j = w, o, p$) when all its exchangeable hydrogens are protons |
| $B_j$ | constant of the $j$ Yukawa potential ($j = D, A$) |
| $C_0$ | nominal cosolvent molar concentration of the binary solvent |
| $c_p$ | nominal protein mass concentration of the three-component solution |
| $\delta_l$ | local domain thickness |
| $e$ | proton charge |
| $\varepsilon$ | bulk dielectric constant |
| $f_o$ | bulk cosolvent activity coefficient |
| $f_w$ | bulk water activity coefficient |
| $g(r)$ | protein-protein radial pair correlation function |
| $\gamma$ | ratio between the core protein volume $V_p$ and its nominal value |
| $I(Q)$ | macroscopical differential scattering cross section |
| $I_K(Q)$ | macroscopical differential scattering cross section of a system with exchange constant $K$ |
| $J$ | attractive potential at the contact |
| $K$ | cosolvent/water exchange equilibrium constant |
| $\kappa_A$ | inverse decay length of the attractive Yukawa potential |
| $\kappa_D$ | inverse Debye screening length |
| $m$ | number of protein contact sites |
| $M_p$ | protein molecular weight |
| $\eta$ | volume fraction of the hard spheres |
| $N_A$ | Avogadro’s number |
| $n_{D+}$ | nominal number concentration of all the exchangeable deuterons |
| $n_{H+}$ | nominal number concentration of all the exchangeable protons |
| $N_j$ | number of conditions of the $j$ set |
| $n_o$ | nominal cosolvent number concentration of the three-component solution |
| $n_p$ | nominal protein number concentration of the three-component solution |
| $n_w$ | nominal water number concentration of the three-component solution |
| $P(Q)$ | average squared form factor of the solvated protein |
| $P_i(Q)$ | average form factor of the solvated protein |
| $P_{i,j}(Q)$ | single partial form factor of the $i$ domain ($i = p, l$) |
| $P_{i,i'}(Q)$ | pair partial form factors of the $i, i'$ domains ($i, i' = p, l$) |
| $Q$ | scattering vector modulus |
| $\rho_b$ | scattering length density of the bulk binary solvent |
| $\rho_l$ | scattering length density of the local domain |
| $\rho_p$ | scattering length density of the core protein |
| $R$ | protein-protein contact radius |
| $r_e$ | classical electron radius |
| $R_c(i)$ | concentration-distribution factor of the $j$ set of experimental conditions |
| $R_r(i)$ | resemblance factor of the $i$ set of curves |
Symbol Description
\[ R_{t,K} \] resemblance factor of a curve associated with the exchange constant \( K \)
\[ S(Q) \] protein-protein structure factor
\[ S_{\text{eff}}(Q) \] “effective” structure factor
\[ S_{\text{HS}}(Q) \] structure factor of the hard sphere potential in the mean spherical approximation
\[ u(r) \] protein-protein interaction potential
\[ u_A(r) \] attractive Yukawa potential
\[ u_D(r) \] coulombic screened Yukawa potential
\[ u_{\text{HS}}(r) \] hard sphere potential
\[ U_j(Q) \] Fourier transform of \( u_j(r) \) (\( j = D, A \))
\[ V_h \] first protein hydration shell volume
\[ V_l \] local domain volume
\[ V_p \] core protein volume
\[ \nu_o \] molecular volume of the cosolvent
\[ \nu_{w,b} \] molecular volume of the water in the bulk binary solvent
\[ \nu_{w,l} \] molecular volume of water in contact with the protein
\[ \nu_{w,1} \] molecular volume of the water in the local domain
\[ x_D \] deuteration grade
\[ x_w \] nominal water molar fraction of the binary solvent
\[ x_{w,b} \] water molar fraction of the bulk binary solvent
\[ x_{w,l} \] water molar fraction of the local domain
\[ Z_j \] sum of atomic numbers of all atoms belonging to the \( j \) molecule (\( j = w, o, p \))
\[ z_j \] charge of the \( j \) molecule (\( j = w, o, p \)) in proton charge unit \( e \)

[1] Svergun D, Richard S, Koch M H J, Sayers Z, Kuprin S and Zaccai G 1998 Proc. Natl. Acad. Sci. USA 95 2267–2272
[2] Levitt M and Sharon R 1988 Proc. Natl. Acad. Sci. 85 7557–7561
[3] Burling F T, Weis W I, Flaherty K M and Brünger A T 1996 Science 271 72–77
[4] Klibanov A M 2001 Nature 409 241–246
[5] Schellmann J A 2003 Biophys. J. 85 108–125
[6] Ortore M G, Spinazzoli F, Carsughi C, Mariani P, Bonetti M and Onori G 2005 Chem. Phys. Lett. 418 338–342
[7] Frensdorf H K, Watson M T and Kauzmann W 1953 J. Am. Chem. Soc. 75 5167–5172
[8] Kirkwood J G and Goldberg R J 1950 J. Chem. Phys. 18 54–57
[9] Timasheff S N 2002 Proc. Natl. Acad. Sci. 99 9721–9726
[10] Purkiss A, Skoilakis S and Goodfellow J M 2001 Phyl. Trans. R. Soc. Lond. A 359 1515–1527
[11] Soper A K, Castner E W and Luzar A 2003 Biophys. Chem. 105 649–666
[12] Stumpe M C and Grubmüller H 2007 J. Phys. Chem. B 111 6220–6228
[13] RE Karen S T and Bloomfield V A 2002 Biophys. J. 82 2876–2891
[14] Pike A and Acharya K R 1994 Protein Sci. 3 706–710
[15] Zhang W, Capp M W, Bond J P, Anderson C F and Jr M T R 1996 Biochem. 35 10506–10516
[16] Casassa E F and Eisenberg H 1964 Adv. Protein Chem. 19 287–395
[17] Noellken M E and Timasheff S N 1967 J. Biol. Chem. 242 5080–5085
[18] Smith P E 2006 Biophys. J. 91 849–856
[19] Shulgin I L and Ruckenstein E 2005 J. Chem. Phys. 123 054909
[20] Javid N, Vogtt K, Krywka C, Tolan M and Winter R 2007 Chem. Phys. Chem. 8 679–689
[21] Niebuhr M and Koch M H J 2005 Biophys. J. 89 1978–1983
[22] Sinibaldi R, Ortore M G, Spinazzoli F, Carsughi F, Frielingshaus H, Cinelli S, Onori G and Mariani P 2007 J. Chem. Phys. 126 235101–9
[23] Sinibaldi R, Ortore M G, Spinazzoli F, de Souza Funari S, Teixeira J and Mariani P 2008 Eur. Biophys. J. 37 795–804
[24] Ortore M, Sinibaldi R, Spinazzoli F, Carsughi F, Clemens D, Bonincontro A and Mariani P 2008 submitted
[25] Schellmann J A 2005 Quart. Rev. Biophys. 38 351–361
[26] Baynes B M and Trout B L 2003 J. Phys. Chem. B 107 14058–14067
[27] Lide D R 1996 Handbook of Chemistry and Physics, 77th ed. (Cleveland, Ohio: CRC Press)
[28] Jacrot B 1976 Reg. Prog. Phys. 39 911
[29] Dianoux A J and Lander G 2003 *Neutron Data Booklet* (Grenoble, France: OCP Science)

[30] Feigin L A and Svergun D I 1987 *Structure analysis by small-angle X-ray, neutron scattering.* (New York, Plenum Press)

[31] Spinozzi F, Carsughi F, Mariani P, Teixeira C V and Amaral L Q 2000 *J. Appl. Cryst.* 33 556–559

[32] Spinozzi F, Gazzillo D, Giacometti A, Mariani P and Carsughi F 2002 *Biophys. J.* 82 2165–2175

[33] Spinozzi F, Carsughi F, Mariani P, Saturni L, Bernstorff S, Cinelli S and Onori G 2007 *J. Phys. Chem. B* 111 3822–3830

[34] Narayanan J and Liu X Y 2003 *Biophys. J.* 84 523–532

[35] Liu Y, Chen W R and Chen S H 2005 *J. Chem. Phys.* 122 044507–13

[36] Spinozzi F, Carsughi F, Mariani P, Saturni L, Bernstorff S, Cinelli S and Onori G 2007 *J. Phys. Chem. B* 111 3822–3830

[37] Caballero-Herrera A, Nordstrand K, Berndt K D and Nilsson L 2005 *Biophys. J.* 89 842–857

[38] Pike A C W and Acharya K R 1994 *Protein Sci.* 3 706–710

[39] Bhattacharya J, Ghoshmoulick R, Choudhuri U, Chakrabarty P, Bhattacharya P K, Lahiri P, Chakraborti B and Dasgupta A K 2004 *Anal. Chim. Acta* 522 207–214

[40] Farnum M and Zukoski C 1999 *Biophys. J.* 76 2716–2726

[41] Spinozzi F 2008 [http://alisf1.univpm.it/biophysics/software.htm](http://alisf1.univpm.it/biophysics/software.htm)

[42] Diamond R 1974 *J. Mol. Biol.* 82 371–391

[43] Curry S, Mandelkow H, Brick P and Franks N 1998 *Nat. Struct. Biol.* 5 827–835

[44] Svergun D I, Petoukhov M V and Koch M H J 2001 *Biophys. J.* 80 2946–2953

[45] Tanford C and Buzzel J G 1956 *J. Phys. Chem.* 60 225–231

[46] Tanford C and Rorby R 1972 *Biochem.* 11 2192–2198

[47] Carbini A *Private communication*

[48] Press W H, Teukolsky S A, Vetterling W T and Flannery B P 1994 *Numerical Recipies. The Art of Scientific Computing* (Cambridge, U.K.: Cambridge University Press)

[49] Timasheff S N 1998 *Adv. Protein Chem.* 51 355–432

[50] Timasheff S N and Xie G 2003 *Biophys. Chem.* 105 421–448

[51] Shimizu S and Smith D J 2004 *J. Chem. Phys.* 121 1148

[52] Gekko K and Timasheff S N 1981 *Biochem.* 20 4667–4676

[53] Shimizu S 2004 *J. Chem. Phys.* 120 4989–4990

[54] Xie G and Timasheff S N 1997 *Biophys. Chem.* 64 25

[55] Shulgin I L and Ruckenstein E 2006 *Biophys. J.* 90 704–707

[56] Shimizu S and Matubayasi N 2006 *Chem. Phys. Lett.* 420 518–522

[57] Spinozzi F, Ortore M G, Sinibaldi R, Mariani P, Esposito A, Cinelli S and Onori G 2008 submitted