ELSEVIER

Biophysical Chemistry 98 (2002) 249–265

Biophysical Chemistry

www.elsevier.com/locate/bpc

Hydrophobic forces between protein molecules in aqueous solutions of concentrated electrolyte

R.A. Curtis, C. Steinbrecher, M. Heinemann, H.W. Blanch*, J.M. Prausnitz

Chemical Engineering Department, University of California, Berkeley and Chemical Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA

Received 25 January 2002; received in revised form 21 March 2002; accepted 21 March 2002

Abstract

Protein–protein interactions have been measured for a mutant (D101F) lysozyme and for native lysozyme in concentrated solutions of ammonium sulfate at pH 7 and sodium chloride at pH 4.5. In the mutant lysozyme, a surface aspartate residue has been replaced with a hydrophobic phenylalanine residue. The protein–protein interactions of D101F lysozyme are more attractive than those of native lysozyme for all conditions studied. The salt-induced attraction is correlated with a solvation potential of mean force given by the work required to desolvate the part of the protein surfaces that is buried by the protein–protein interaction. This work is proportional to the aqueous surface-tension increment of the salt and the fractional non-polar surface coverage of the protein. Experimental measurements of osmotic second virial coefficients validate a proposed potential of mean force that ascribes the salt-induced attraction between protein molecules to an enhancement of the hydrophobic attraction. This model provides a first approximation for predicting the protein–protein potential of mean force in concentrated aqueous electrolyte solutions; this potential is useful for determining solution conditions favorable for protein crystallization.
© 2002 Elsevier Science B.V. All rights reserved.

Keywords: Intermolecular interactions; Potentials of mean force; Lysozyme; Salting-out; Salts; Hydrophobic effect

1. Introduction

1.1. Overview

Salt-induced precipitation/crystallization is a common initial step to purify target bio-macromolecules from fermentation broths or crude extracts. Protein crystallization is necessary to obtain protein structures from X-ray diffraction. However, it is difficult to select conditions favorable for selective precipitation of a target protein or for obtaining high-quality crystals because there are a variety of factors that control the thermodynamic and kinetic behavior of protein solutions. Toward a better understanding of these factors, the first step is to determine the effective solvent-averaged protein–protein interaction potentials; these are used to predict properties such as protein

*Corresponding author. Tel.: +1-510-642-1387; fax: +1-510-642-1228.
E-mail address: blanch@socrates.berkeley.edu (H.W. Blanch).

0301-4622/02/$ - see front matter © 2002 Elsevier Science B.V. All rights reserved.
PII: S0301-4622(02)00071-6

activity coefficients, protein diffusion coefficients and free energies. However, the interaction between protein molecules in a concentrated salt solution is not well understood. Since protein solubility typically decreases as salt concentration rises, protein–protein interactions become more attractive with increasing salt concentration. Models based on DLVO theory [1] fail to predict this behavior. The only ionic-strength-dependent term in DLVO theory is the electric double-layer potential, which is negligible at the salt concentrations where salting-out is observed. Although protein solubility is a strong function of the type of salt, following the lyotropic series, in the DLVO theory all ions are point charges and the theory does not distinguish between different types of ions of identical charge.

In addition to the forces reflected in the standard DLVO potentials, there are water-mediated forces between protein molecules; these forces, called solvation forces, include both hydration forces and hydrophobic forces. When the solvent structure adjacent to the protein surface is perturbed, the solvation force is determined by the change in free energy associated with this perturbation [2]. Charged or polar surfaces adsorb water molecules; these surfaces are hydrophilic and the forces between them are called hydration forces. Hydration forces are repulsive because as the distance between two hydrated protein molecules decreases, free energy is required to remove the intervening water molecules around the polar or charged surface groups. Hydrophobic interactions occur between non-polar protein surfaces; these forces are attractive because as the distance between two non-polar protein surfaces decreases, the overall free-energy change associated with desolvating non-polar residues is negative.

In concentrated salt solutions, solvation forces are significantly altered from those in dilute electrolyte solutions because salt ions change the structure of the intervening water [3]. From solubility studies of non-polar organic solutes, it is known that the hydrophobic interactions between non-polar residues increase with addition of salt, leading to the well-known salting-out effect [4]. Studies on the solubilities of model peptides in concentrated salt solutions have shown that in addition to the salting-out effect of non-polar groups, there is a salting-in effect due to an electrostatic interaction between the salt ions and the peptide group [5,6]. Here, we determine the salt-induced protein–protein attraction and correlate this attraction with the effect of the salt on the solvation energetics of the non-polar and polar surface of the protein. We compare protein–protein interactions between native lysozyme molecules with those of a mutant lysozyme as a function of salt type and salt concentration. In the mutated lysozyme molecule, a surface aspartate residue is replaced with hydrophobic phenylalanine. The protein–protein interaction is measured by static light scattering to determine an osmotic second virial coefficient, B_{22}, that is related to an integral over separation and angles of the potential of mean force [7]. The potential of mean force is the free energy of the system when two protein molecules are held at a fixed separation relative to when they are infinitely far apart, averaged with respect to orientation of the two protein molecules and with respect to orientation of all molecules in the medium separating the two protein molecules.

1.2. Protein–protein interactions in protein crystallization

George and Wilson [8] have shown the importance of the pair potential of mean force for predicting solution conditions favorable for protein crystallization; they have shown a B_{22} crystallization window for protein solutions. As a necessary (but not sufficient) condition for protein crystallization, B_{22} should lie between -2×10^{-4} and -8×10^{-4} ml mol/g^2. For B_{22} more positive than -2×10^{-4} ml mol/g^2, the protein–protein attraction is usually not sufficiently strong to form stable protein crystals. For solutions where B_{22} is more negative than -8×10^{-4} ml mol/g^2, amorphous precipitation is likely to occur because protein–protein attractions are sufficiently strong so that the protein molecules do not have adequate time to orient themselves into a crystal lattice. Guo et al. [9] have extended the George and Wilson work to show that there is a relationship between protein solubility and B_{22} which has been correlated with a simple phase equilibrium theory [10]. More

advanced potential-of-mean-force models have been used to determine the entire protein phase diagram [11,12]. The key result of these investigations is that the range of the potential is important for crystallization. For significantly short-ranged potentials, the crystallization window corresponds to a meta-stable liquid–liquid critical point in the vicinity of the liquid–solid equilibrium. This observation is important for crystallization because density fluctuations are enhanced in the critical region, lowering the free energy for formation of critically sized nuclei [13]. But unfortunately, the potential of mean force in concentrated salt solutions is not well understood. To take advantage of these crystallization diagnostics, the individual contributions to the total pair potential of mean force (pmf) need to be determined.

1.3. Potential-of-mean-force model

In dilute aqueous electrolyte solutions, the lysozyme–lysozyme interaction has been modeled accurately with the DLVO theory [1] where the protein molecules are modeled as hard, polarizable spheres containing a uniform net charge. The interaction consists of an electric double-layer potential, an attractive Hamaker dispersion potential, and a hard-sphere repulsion [14,15]. The electric double-layer potential is repulsive because the spheres have identical charge, whereas the dispersion interaction is always attractive between two identical polarizable bodies immersed in a medium. The magnitude of the dispersion interaction is given by the Hamaker constant, H. Calculations using Lifshitz theory predict that H is on the order of 3–5 k_BT for protein molecules immersed in water where k_B is Boltzmann's constant and T is temperature [16]. Values of H obtained from fitting experimental data are sometimes greater than 5 k_BT. Usually, extra protein–protein attraction is included in models by using larger values of H.

Here, we use the DLVO potentials in addition to a solvent-accessible surface area (SASA) potential to account for all other interactions not included in DLVO theory. These interactions include hydration and hydrophobic forces, hydrogen-bonding interactions and protein-specific forces. In the SASA potential, the interaction energy is assumed to be proportional to the total area of the two protein molecules that is within a certain surface-to-surface separation, r_c. The form of the potential is given by:

$$W_{SASA}(r) = -\kappa A(r) \quad (1)$$

where κ is a scaling factor with units of a surface energy, r is center-to-center separation, and $A(r)$ is the buried surface area. For the case where the identical proteins are modeled as spheres (see Fig. 1), $A(r)$ is given by:

$$A(r) = \pi d_2^2 - \pi r d_2 + \pi r_c d_2 \text{ for } d_2 < r < d_2 + r_c \quad (2)$$

As r increases, the amount of buried surface area decreases and the magnitude of $W_{solv}(r)$ is given by a monotonically decreasing function that goes to zero at $d_2 + r_c$ where d_2 is the hard-sphere protein diameter.

The SASA potential has been used to determine the energetic contribution of the hydrophobic effect

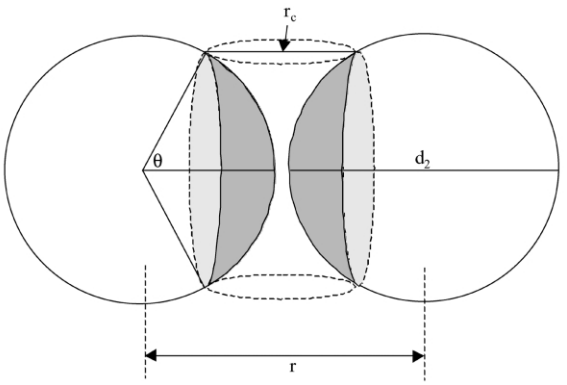

Fig. 1. The solvation force is described by a surface free energy multiplied by the surface area buried by the interaction approximated by the spherical cap area of the proteins colored in dark gray. The surface-to-surface cut-off separation, r_c, is approximated by a solvent diameter. The combined areas of the spherical cap regions enclosed by the box is given by:

$$A(r) = 2\left[\int_0^{\theta} 2\pi(d_2/2)^2 \sin\theta' d\theta'\right] = \pi d_2^2 - \pi r d_2 + \pi r_c d_2$$

where d_2 is the diameter of the protein, θ is the solid angle corresponding to the boundary of the spherical cap and $\cos\theta = \frac{r - r_c}{d_2}$.

in protein folding where the scaling factor is given by the solvation free energy of the part of the surface that is buried during folding. In protein folding studies, the free energy of burying part of the non-polar protein surface is set equal to the free energy to transfer the same surface from water to an organic phase mimicking the protein interior [17–19]. Because the SASA potential we employ here is spherically symmetric, we assume that the protein surface has averaged characteristic non-polar and polar areas. Consequently, the scaling factor includes contributions from the free energy to desolvate the non-polar and the polar part of the protein surface. A positive contribution to the scaling factor results from attractive hydrophobic interactions that are attributable to desolvating the non-polar part of the protein surface. A negative value of the scaling parameter indicates that the protein–protein interaction is repulsive. Repulsive forces may occur because there is an electrostatic desolvation penalty for excluding water in the vicinity of polar or charged surface groups. Forces of this nature are hydration forces.

When the DLVO and the SASA potentials are incorporated into the potential-of-mean-force (pmf) model, the osmotic second virial coefficient is given by:

$$B'_{22} = \frac{2}{3}\pi d_2^3 + \frac{1}{2}\int_{d_2}^{d_2+r_c} (1-e^{-\beta W_{SASA(r)}})4\pi r^2 dr + \frac{1}{2}\int_{d_2+r_c}^{\infty} (1-e^{-\beta W_{DLVO(r)}})4\pi r^2 dr \quad (3)$$

Here B'_{22} is related to the experimentally-determined osmotic second virial coefficient B_{22} by

$$B_{22} = \frac{N_{av}B'_{22}}{M_2^2} \quad (4)$$

where N_{av} is Avogadro's number and $\beta = (k_B T)^{-1}$. W_{DLVO} is the sum of the electric-double-layer repulsion and Hamaker-dispersion potentials. The first term is the contribution to B_{22} from the protein-excluded volume effect. The contribution to B_{22} from the SASA potential and that from the DLVO potentials are separated by recognizing that the SASA potential contains the contributions to the pmf from short-ranged forces acting over a surface-to-surface separation, r_c, which is chosen to be approximately one solvent diameter. A disadvantage of this pmf model is discontinuities at a surface-to-surface separation equal to d_2 and equal to $d_2 + r_c$. However, if desired, the model can be reparameterized so that the sum of the DLVO potentials and the SASA potential is continuous; as shown by Asthagiri et al. [20], a scaling parameter can be used to connect a short-ranged form and a long-ranged form of the dispersion potential. For separations where the solvent is excluded from the region between the two proteins, Asthagiri et al. employ an all-atomistic calculation for the dispersion potential, whereas the longer-range contribution to the dispersion potential was determined using a Lifshitz–Hamaker method.

The main advantage of the model proposed here is that the model provides a simple approximation for separating the salt-dependent potentials in concentrated electrolyte solutions from those in dilute electrolyte solutions, as shown in Fig. 2a,b. In Fig. 2a, the contribution of the protein excluded volume $B_{22,ex}$ and that of the DLVO potential $B_{22,DLVO}$ are plotted vs. ionic strength. In the calculation, the Hamaker constant is set to 8 $k_b T$ which is similar to the value reported for lysozyme [21]. As the ionic strength rises, $B_{22,DLVO}$ as-symptotically approaches a constant value as the electric double layer potential is screened out. In typical salting-out conditions, the ionic strength is greater than 0.5 molal and the proposed crystallization window suggests that B_{22} should be less than -2×10^{-4} ml mol/g^2 to crystallize proteins. For lysozyme, the repulsive part of the potential $B_{22,ex}$ approximately cancels the attractive contribution from the DLVO potential $B_{22,DLVO}$ indicating that protein–protein attraction cannot be described by a model containing only DLVO potentials. For example, Velev et al. [22] have fit lysozyme–lysozyme interactions using a more exact electrostatic model and found that a high Hamaker constant of 13.8 $k_B T$ gives the best fit to data obtained in solutions of 0.005 to 0.5 molar sodium–chloride solutions. Velev et al. proposed that the difference between the fit Hamaker constant and the expected value of 3–5 $k_B T$ was due to the neglect of a small number of highly attractive configurations. It is

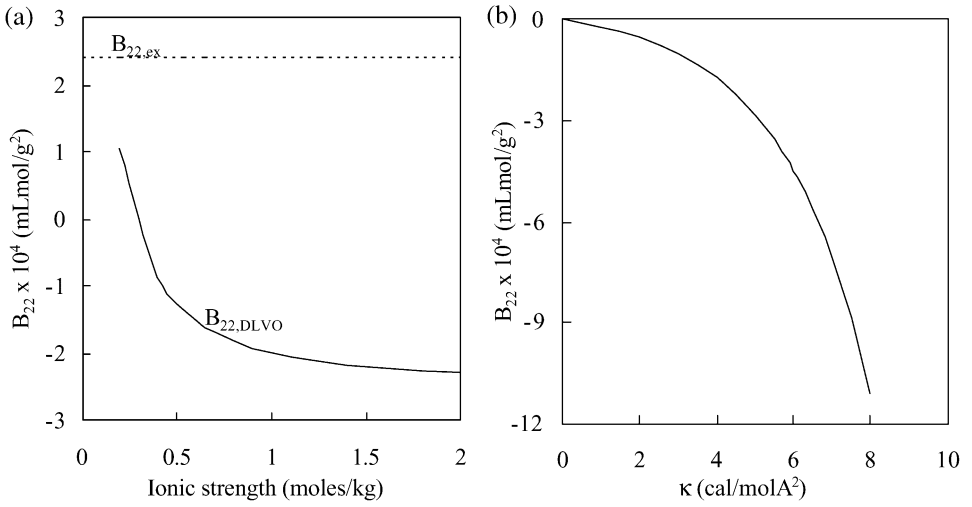

Fig. 2. (a) Contribution of the DLVO potentials $B_{22,\mathrm{DLVO}}$ and the contribution of the protein excluded volume $B_{22,\mathrm{ex}}$ plotted vs. ionic strength (molal). (b) Contribution of the SASA potential $B_{22,\mathrm{SASA}}$ as a function of scaling parameter κ.

these anisotropic interactions that determine the protein orientation in the crystal.

Here, we correlate the SASA potential with the change in the protein–protein attraction as salt concentration rises. In Fig. 2b, the contribution of the SASA potential $B_{22,\mathrm{SASA}}$ is plotted as a function of the scaling parameter, κ. In the model described by Eq. (3), $B_{22,\mathrm{SASA}}$ is given by the difference between the measured B_{22} and $B_{22,\mathrm{DLVO}}$, which would correspond to a range of κ between 5 and 9 cal/molÅ^2 according to Fig. 2b.

Here, we propose that the scaling factor is a linear function of the salt molality, m_3

$$\kappa = \gamma m_3 + \kappa_o \tag{5}$$

where κ_o is a hypothetical value of the scaling factor at a salt molality of zero. γ is a proportionality constant defined by Eq. (5) with units of surface energy per molal. The motivation for using Eq. (5) derives from the classical salting-out equation [23] where

$$\ln\frac{S_2}{S_{2,o}} = \beta - k_s m_3 \tag{6}$$

Here, S_2 is the solubility of a protein in a solution containing salt at a molality of m_3; $S_{2,o}$ is the solubility of the protein in salt-free water; β is a hypothetical protein solubility at zero salt molality; and k_s is the salting-out constant. For conditions where the solubility of the protein is sufficiently low that protein–protein interactions are negligible, it follows from the Cohn equation that the solvation free energy of the protein is proportional to the salt molality if the protein crystal is assumed to be a pure phase [24]. Thus,

$$\frac{g_2}{k_B T} = \frac{g_{2,o}}{k_B T} - \beta + k_s m_3 \tag{7}$$

where g_2 and $g_{2,o}$ are the solvation free energy of the protein in the salt solution and that in salt-free water, respectively. As a first approximation, we propose that the scaling factor, κ, is given by the free energy per unit area to desolvate that part of the protein surface that is made inaccessible to the solvent due to the proximity of the two protein surfaces. Because the solvation free energy is proportional to salt molality according to Eq. (7), we also expect that the scaling parameter is proportional to salt molality.

The approximation that the protein crystal is a pure phase may lead to significant error because it is known that protein crystals contain a significant amount of salt and water [25]. Arakawa and Timasheff [26,27] have determined the solvation free-energy increment, dg_2/dm_3, independent of

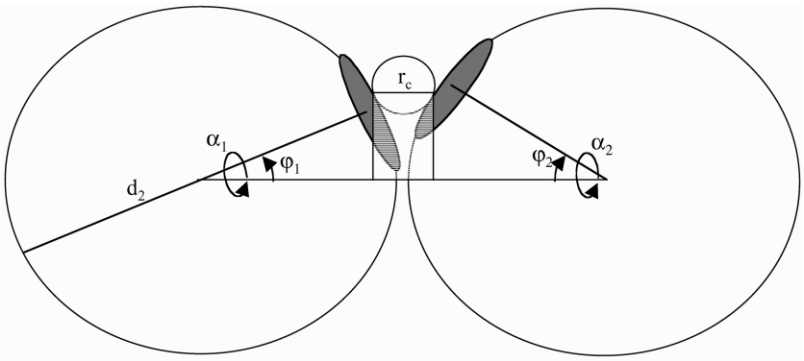

Fig. 3. The mutated residue is modeled as a circular patch on the surface of the sphere. The patch–patch interaction is given by the non-polar surface free energy multiplied by the surface area of the patch that is within a cut-off separation, r_c, of the second sphere. Here, this area is given by the total area of the two patches that is enclosed by the box. See text for description of the angles of rotation, α_1, α_2, φ_1 and φ_2; r_c is the cut-off surface-to-surface separation approximated as a solvent diameter and d_2 is the protein diameter.

the assumption of a pure protein crystal from protein-salt preferential-interaction-parameter measurements indicating that the molal surface-tension increment of the salt provides a first approximation for determining the effect of salt on the solvation free energy. The molal surface-tension increment of the salt is correlated with the position of the salt in the lyotropic series that was originally developed to describe the salting-out effectiveness of various ions for globular proteins [28]. For anions, the series is given in decreasing order of the molal surface-tension increment by $SO_4^{2-} > F^- > Cl^- > Br^- > NO_3^- > CLO_4^- > I^- > SCN^-$; the corresponding series for cations is given by $Mg^{2+} > Na^+ > K^+ > Li^+ > NH_4^+ > Cs^+$. High-lyotropic-series salts are called kosmotropes, whereas low-lyotropic-series salts are called chaotropes. In addition to increasing the surface free energy of the protein, salt ions interact favorably with exposed polar and charged surface groups. Favorable protein-salt interactions are associated with a negative surface-free-energy increment. In this work, we correlate γ with the molal surface-tension increment of the salt.

1.4. Potential-of-mean-force model for anisotropic interactions

Information on the effect of salt on the hydrophobic interaction between protein molecules is provided by comparing the results for native lysozyme with those for D101F lysozyme. The D101F lysozyme is modeled as a sphere with a hydrophobic circular patch on the surface with an area given by the approximate solvent-exposed surface area of the phenylalanine residue. The difference in the B_{22} values of native and of D101F lysozyme is related to an angle-averaged form of the virial coefficient. The angle-dependent solvation free energy of the circular patch mimicking the hydrophobic mutation is given by:

$$\Gamma_p(r,\varphi_1,\varphi_2) = -\kappa_a A_p(r,\varphi_1,\varphi_2) \quad (8)$$

The geometry is shown in Fig. 3. A_p is the combined surface area of the two patches that is within a cut-off surface–surface separation, r_c, of the second sphere; φ_1 and φ_2 are the angles of the spheres with respect to the axis connecting their centers that denote the orientation of the patch, and κ_a is a characteristic scaling parameter for the patch. We denote this parameter with subscript a because the value of this parameter is expected to be given by a non-polar surface free energy. The corresponding form of Eq. (5) is given by:

$$\kappa_a = \gamma_a m_3 + \kappa_{a,o} \quad (9)$$

The patch is approximated as circular such that the orientation of the patch is independent of rotation about the axis connecting the center of the patch and the center of the sphere. The poten-

tial is invariant to rotations of the spheres about the axis connecting the centers of the two spheres, α_1 or α_2. For this case, the angle-averaged potential of mean force is given by:

$$e^{-\beta W_{p(r)}} = \frac{1}{4}\int_0^{\pi} e^{-\beta \Gamma_{p(r,\varphi_1,\varphi_2)}} \sin(\varphi_1)\sin(\varphi_2) d\varphi_1 d\varphi_2 \quad (10)$$

where the orientation dependence of the buried patch surface area is discussed in Appendix A. The contribution from the patch–patch interaction to B'_{22} is:

$$B'_{22,D101F} - B'_{22,native} = \frac{1}{2}\int_{d_2}^{d_2+r_c} (1 - e^{-\beta W_{p(r)}}) 4\pi r^2 dr \quad (11)$$

2. Experimental methods

2.1. Production of D101F lysozyme

The mutant lysozyme was produced using the methylotrophic yeast *Pichia pastoris* expression system following the guidelines outlined in the Invitrogen Pichia Expression Kit. The D101F gene was graciously supplied by Prof. J. Kirsch, Chemistry Department, University of California at Berkeley. Briefly, the gene was inserted into the pPIC9 expression vector which contains the DNA coding for the AOX1 promoter and also the *Pichia* wild-type gene, HIS4. The constructed vector was cloned into the GS115 wild-type strain of *P. pastoris* at the his4 loci from a single crossover event with the HIS4 gene on the vector. The resulting strain of yeast has the same methanol utilization phenotype as the parent strain because the AOX1 gene was kept intact during the insertion of the foreign gene.

The fermentation was performed using a Bioflo III reactor vessel equipped with microprocessor-based PID controllers for pH, temperature and dissolved oxygen control. The fermentation procedure and medium are described in Ref. [29]. The key aspects of the fermentation protocol include a phase of glycerol excess, a glycerol-limited phase and a ramped methanol feed phase during the transition from glycerol to methanol feeds. The cell density initially increases in a repressed excess-glycerol batch phase. After the glycerol is exhausted, a glycerol feed is initiated at a limiting rate further to increase cell density and to allow derepression of the AOX1 promoter. The glycerol-fed batch stage is followed by a methanol-fed batch stage where the flow rate of the methanol is slowly increased by 10% every 30 min until the maximum feed rate rises to 10 ml/h·l. At this time, the fermenter is run in a continuous phase for 24 h with yields of approximately 100 mg/l D101F lysozyme.

2.2. Lysozyme purification

The fermentation broth was centrifuged at $3500 \times g$ for 25 min at room temperature to remove cell debris. The supernatant was than loaded on the weak-cation exchange column (500-ml Pharmacia XK-50 column packed with 400 ml of Pharmacia CM Sepharose fast flow resin) at a flow-rate of 10 ml/min and washed with 600 ml of phosphate buffer at a flow rate of 15 ml/min. The D101F lysozyme was co-eluted with a glycosylated form of lysozyme into a volume of 350 ml using 0.5 M NaCl solution at a flow-rate of 30 ml/min. The eluent was concentrated and diafiltered using a 200-ml Amicon ultrafiltration unit (YM-3 Amicon membrane with a cut-off of 3000 Da) with three volumes of phosphate buffer. The retentate was then diluted with four volumes of phosphate buffer and loaded on the second column (30 ml Amicon C-30 column packed with Pharmacia SP Sepharose fast flow resin) at a flow rate of 2 ml/min. The column was washed with 60 ml phosphate buffer at a flow rate of 3 ml/min. The glycosylated form of D101F lysozyme was eluted using 0.18 M NaCl/phosphate buffer at a flow rate of 3 ml/min and the un-glycosylated form was eluted using the salt solution that was used in the light-scattering experiment or a solution of 0.25 M NaCl, pH 7. The final preparation of the protein required concentrating and washing the eluent with twelve volumes of the salt solution to be used in the experiment using the 60-ml Amicon ultrafiltration unit (YM-3 Amicon membrane with a cut-off of 3000 Da).

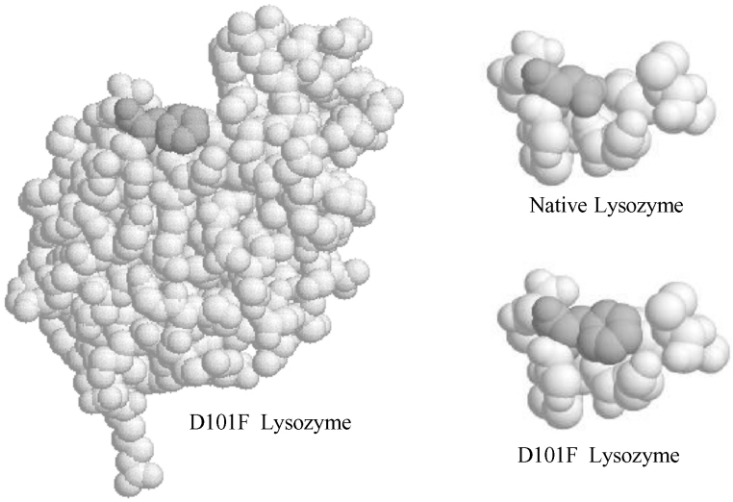

Fig. 4. On the left is the space-filling structure of D101F lysozyme, where the mutated residue is shown in dark gray. The blow-up of the region is shown on the bottom right and the analogous representation for the native form is shown on the top right.

2.3. Light scattering

Light-scattering measurements were made using a Wyatt Technology mini-DAWN static light-scattering detector that uses a 30-mW semiconductor diode laser at a fixed wavelength of 690 nm. A calibration constant is required to relate the voltage output from the 90° detector to the Rayleigh ratio. However, the calibration constant depends on the solvent because the solid angle that is measured by the detectors depends on the refractive index of the solvent. The instrument software accounts for this effect using known relations between calibration constants and the refractive index of the solvent. The calibration was performed with toluene that has a known high R_θ.

For each experiment, R_θ was measured for the solvent and for 10–12 protein samples in order of decreasing concentration. Each sample was pumped through a 0.02-μm Anotop syringe filter at a flow rate of 0.1 ml/min using a Sage-Instruments syringe pump. The lysozyme concentration was measured using a Milton Roy Spectronic 1201 spectrophotometer at a fixed wavelength of 280 nm. A value of 2.635 $cm^3/g{\cdot}cm$ [30] was used as the extinction coefficient. The lysozyme activity was measured using the Micrococcus lysodeikticus technique [31] to assure that the protein folded correctly.

2.4. Light-scattering theory

For the three-component system water (1), protein (2), salt (3) in dilute protein solutions, the light-scattering equation is given by [32]:

$$\frac{Kc_2(\partial n/\partial c_2)^2_{T,\mu_1,\mu_3}}{\bar{R}_\theta} = \frac{1}{RT}\left(\frac{\partial \Pi}{\partial c_2}\right)_{T,\mu_1,\mu_3} \tag{12}$$

where $\bar{R}_\theta$ is the excess Rayleigh scattering of the protein solution over the aqueous salt solution, K is the light-scattering constant, n is the refractive index, c_2 is weight concentration of protein, and Π is the osmotic pressure. The measurements are made in solutions dilute in protein where only two-body interactions are significant. In this case, Eq. (12) reduces to:

$$\frac{Kc_2(\partial n/\partial c_2)^2_{T,\mu_1,\mu_3}}{\bar{R}_\theta} = \frac{1}{M_2} + 2B_{22}c_2 \tag{13}$$

A plot of $Kc_2(\partial n/\partial c_2)_{T,\mu_1,\mu_3}/\bar{R}_\theta$ vs. c_2 can be used to determine the osmotic second virial coefficient of the protein, B_{22}, and the molecular weight of the protein.

Table 1
Parameters used in potential-or-mean-force calculations

Net charge, pH 7	8
Net charge, pH 4.5	12
Effective spherical diameter (Å)	34.4
f	0.49
Area hydrophobic patch (Å^2)	97.0

The net-charge parameters from hydrogen-ion titrations are given in [37]. The effective spherical diameter is determined from lysozyme crystal dimensions. f is the fractional coverage of non-polar groups for native lysozyme.

2.5. Free-energy minimization for D101F

The free-energy minimization of D101F lysozyme was performed using the modules Biopolymer and Discover of the Insight98 software package of Molecular Simulation Inc. Fig. 4 shows a space-filling representation of the mutated residue with its surrounding neighborhood along with the same representation of the full molecule. An analogous drawing for the native form is also shown for comparison. The mutated residue sits at the edge of a cleft where it is accessible to the solvent. The accessible surface areas of the D101F mutant and native lysoyme were calculated with Insight98 using the Michael Connolly Surface package [33]. Results are reported in Table 1.

3. Results and discussion

Infinite-dilution weight-average molecular weights and osmotic second virial coefficients were determined for D101F lysozyme dissolved in solutions of aqueous ammonium sulfate at pH 7 or aqueous sodium chloride with 50-mM sodium acetate buffer, pH 4.5 as a function of salt concentration. The corresponding values for the native form have been reported [34]. We choose to study solutions of ammonium sulfate or solutions of sodium chloride because these solutions are commonly used to precipitate or crystallize proteins. Furthermore, the solubility behavior of lysoyme in concentrated salt solutions is not well understood. Aqueous solutions of ammonium sulfate should have a larger Hofmeister effect than those of sodium chloride because sulfate ion is more kosmotropic than chloride ion. However, lysozyme solubility follows the reverse lyotropic series.

In this work, the salt-induced protein–protein attractions are correlated with the molal surface-tension increment of the salt. For this reason, the correlations are based on a molality scale instead of the traditional ionic-strength scale which is generally used when electrostatic interactions are the dominant intermolecular forces. The surface tension increments are 3.11 and 2.36 cal/mol A^2-kg/mole for ammonium sulfate and for sodium chloride, respectively. In all regressions of the light-scattering data, the surface free-energy parameters are regressed from the data holding all other parameters fixed as listed in Table 1.

The reported error bars are determined from the standard deviation of the slope of the line from the light-scattering data containing 10–12 points. Experiments were repeated if the error was greater than 0.5 ml mol/g^2.

Fig. 5 shows B_{22} for native lysozyme in solutions of sodium chloride or ammonium sulfate plotted vs. salt molality. In both cases, B_{22} decreases with rising salt molality indicating that the protein–protein attraction is enhanced with addition of salt. We propose that this increase in attraction results from the Hofmeister effect where addition of salt increases the surface free energy of the non-polar part of the protein surface, thereby increasing the driving force to bury those residues. However, because lysozyme solubility follows the reverse lyotropic series with respect to the anion, we expect that the protein–protein interactions are more attractive in sodium chloride than in ammonium sulfate. The abnormal solubility behavior is attributed to changes in the protein–protein interactions upon anion binding to the positively-charged protein surface. Interactions between lysozyme-chaotropic anion complexes are more attractive than those between uncomplexed lysozyme molecules [35], whereas the opposite behavior is observed for interactions between lysozyme-kosmotropic anion complexes [34]. Here, the protein–protein interactions in sodium chloride are slightly more attractive than those in ammonium sulfate when compared at constant ionic strength, while the opposite holds if the intermolecular interactions are compared at the

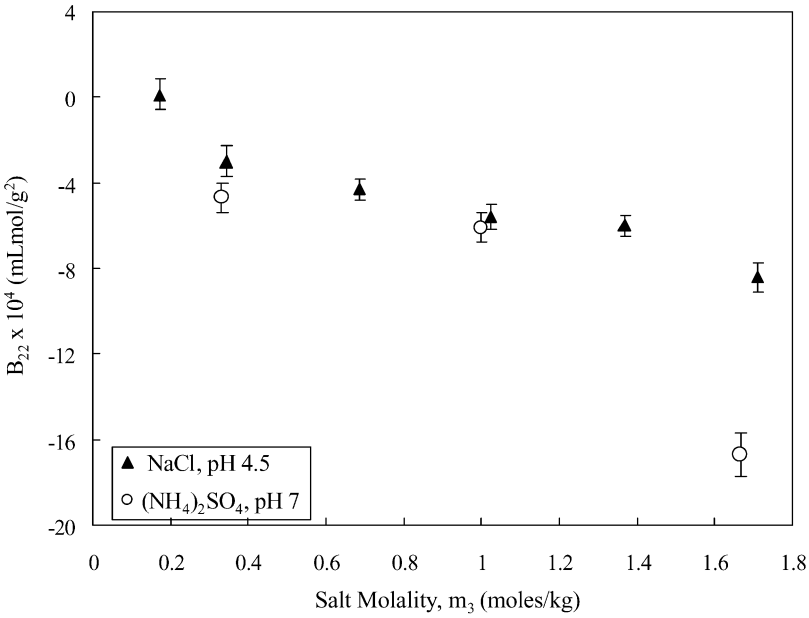

Fig. 5. Experimental B_{22} for lysozyme in solutions of sodium chloride at pH 4.5 or ammonium sulfate at pH 7 vs. salt molality.

same molality. It is possible that the effect of ion binding on the protein–protein interactions is not significant for the conditions studied here. As reported by Curtis et al. [34], B_{22} becomes more positive as pH is lowered in concentrated ammonium–sulfate solutions. This effect is attributed to the formation of soluble lysozyme–sulfate complexes at low pH. In this work, the experiments in ammonium sulfate solutions are at pH 7 where sulfate–ion binding is minimal because lysozyme has a low net positive charge at this pH. In sodium–chloride solutions, because the slightly chaotropic chloride ion binds to the slightly chaotropic imidazolium ion, the protein surface chemistry is not significantly altered; consequently, the protein–protein interactions might not change upon chloride–ion binding.

The surface-free-energy parameter, κ, is fit to the values of B_{22} using the pmf in Eq. (3). In Fig. 6a,b, we plot κ vs. salt molality for the case where the pmf does not include the Hamaker dispersion potential and the case where the Hamaker constant is set to 8 k_BT. The values of γ [see Eq. (5)] and κ_o are obtained from the slopes and the intercepts on the vertical axis of the best-fit lines to the data shown in Fig. 6a,b. For the ammonium–sulfate and sodium–chloride solutions, the fit values of κ and γ are insensitive to the choice of the Hamaker constant. The fit values of κ_o are approximately 6 cal/molÅ^2 for both salt solutions. For the solutions of sodium chloride, the ratio of γ to the molal surface-tension increment is 0.40, similar to the value 0.68 obtained for the ratio of the measured surface free-energy increment to the molal surface-tension increment from experiments in 1.0 molal solutions of sodium chloride [26]. The data obtained for ammonium–sulfate solutions exhibit considerable scatter about the best-fit line. Consequently, the corresponding value of γ is approximate although it is of the same order of magnitude as that obtained from solutions of sodium chloride.

Fig. 7 compares osmotic second virial coefficients for D101F lysozyme to those of native lysozyme. The difference in B_{22} is due to hydrophobic interactions of the mutated residue. The contribution to the potential of mean force from the mutation is reproduced by our model for the patch–patch interaction in Eq. (11). Fig. 8 shows the results of fitting the non-polar surface free-energy parameter, κ_a using this model. The fit

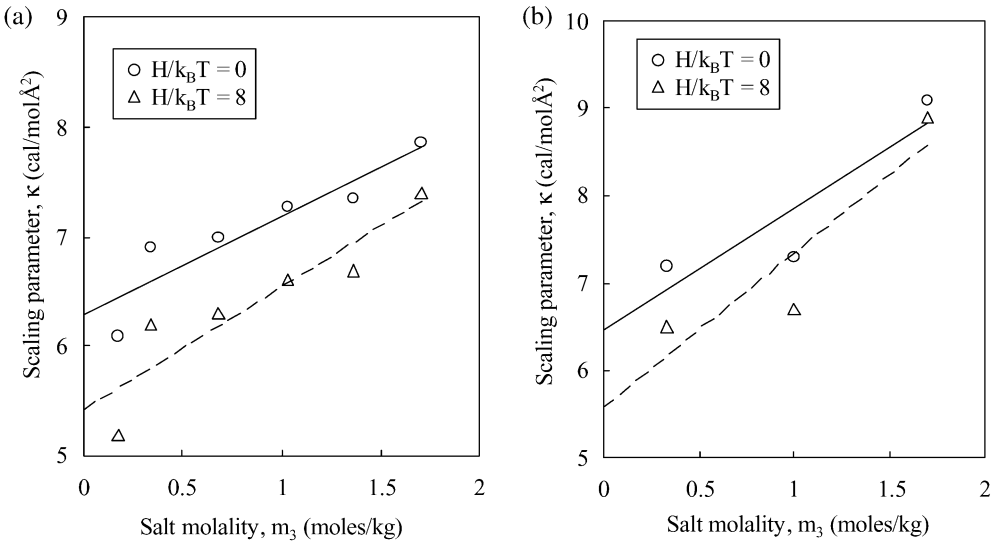

Fig. 6. Surface free energy of lysozyme in solutions of sodium chloride at pH 4.5 determined from fitting B_{22} to the potential-of-mean-force model described by Eq. (3), with $r_c = 3$ Å. For $H = 8.0\ k_BT$, the slope of the plot, $\gamma = 0.93$ (cal/molÅ^2) (kg/mole) or for $H = 0.0\ k_BT$, the slope of the plot, $\gamma = 0.94$ (cal/molÅ^2) (kg/mole). Surface free energy of lysozyme in solutions of ammonium sulfate, pH 7, determined from fitting B_{22} to the potential-of-mean-force model described by Eq. (3), with $r_c = 3$ Å. For H = 8.0 k_BT, the slope of the plot, $\gamma = 1.9$ (cal/molÅ^2) (kg/mole) or for $H = 0.0\ k_BT$, the slope of the plot, $\gamma = 1.5$ (cal/molÅ^2) (kg/mole).

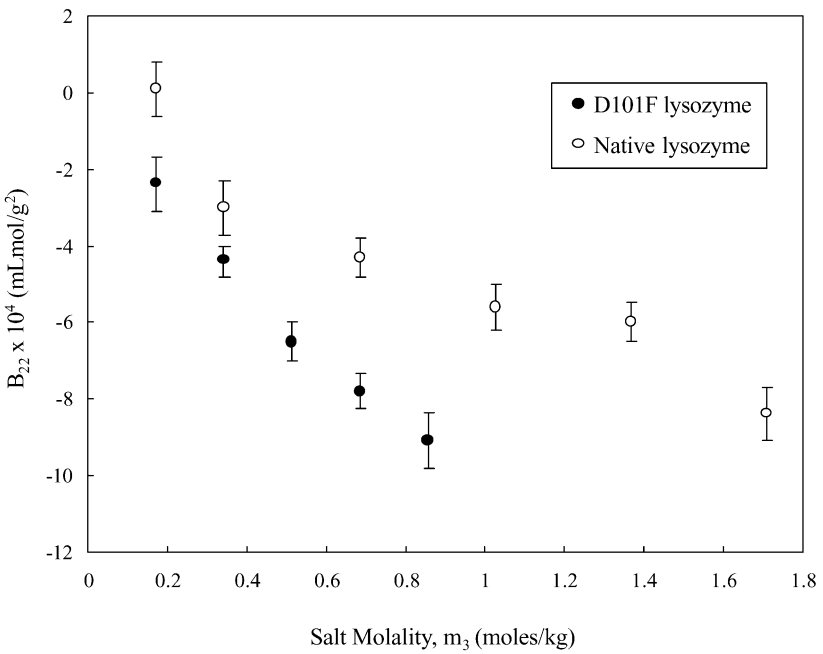

Fig. 7. Experimental B_{22} for D101F lysozyme or for native lysozyme in solutions of sodium chloride, pH 4.5, vs. salt molality.

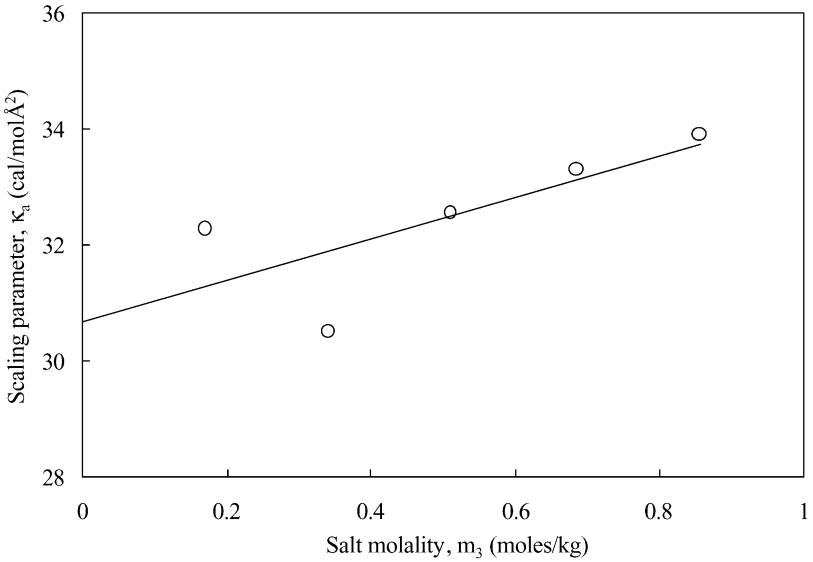

Fig. 8. Hydrophobic surface free energy for D101F lysozyme in solutions of sodium chloride at pH 4.5 calculated from the pmf model [Eq. (11)] vs. salt molality.

values of κ_a vary between 30 and 34 cal/molÅ^2 and increase with rising salt concentration indicating that the non-polar patch is salted-out by addition of sodium chloride. The fit values of κ_a are significantly greater than the corresponding values of κ because the patch is non-polar whereas the protein surface contains polar and charged groups that lead to a negative contribution to the scaling parameter.

The slope of the best-fit line given in the plot, γ_a, is equal to 3.6 cal/molÅ^2-kg/mole, larger than the value of γ (see Fig. 6a). If the primary effect of salt is to raise the surface free energy of the non-polar groups, we expect that the ratio of γ_a to γ is given by the inverse of the ratios of the fractional non-polar surface coverage of the patch, f_a, to that of the protein surface, f. The fraction of the solvent accessible surface area of lysozyme that is carbon is approximately 0.49. Because the patch is predominantly non-polar, $\gamma_a/f_a = 3.6$, in semi-qualitative agreement with the ratio, $\gamma/f = 1.8$.

Given the scatter of the data and the simplifying geometric assumption contained in the pmf model, we cannot draw any conclusions regarding the effect of salt on the solvation free energy of the polar part of the protein surface. Most likely, the interaction between salt ions and proteins is specific for each system; basic proteins have very strong interactions with chaotropic anions, but not with kosmotropic anions. On the other hand, no corresponding trend between acidic proteins and cations is observed.

Fig. 9 gives the apparent molecular weights for D101F lysozyme in solutions of ammonium sulfate or of sodium chloride. The experimental results as-symptotically approach 16 000 g/mol somewhat larger than the molecular weight of the monomer, 14 600 g/mol, indicating that there is either a small degree of pre-aggregation or a high-molecular-weight impurity in all the solutions. High-molecular-weight impurities were not detected in SDS-gel electrophoresis; it is therefore more likely that there is a small degree of irreversible aggregation. For concentrated ammonium–sulfate solutions, the large apparent molecular weights indicate that there is strong association at low protein concentrations. The concentration dependence of the association could not be examined because the sensitivity of the experiment was insufficient at protein concentrations less than 1 g/l for solutions of lysozyme. The apparent molec-

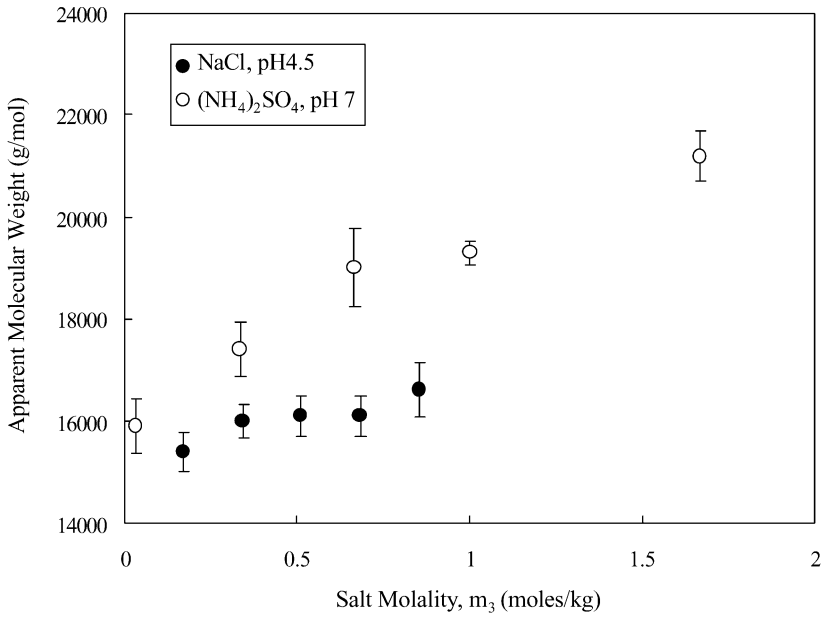

Fig. 9. Apparent molecular weights for D101F lysozyme in solutions of ammonium sulfate, pH 7, or sodium chloride, pH 4.5, vs. salt molality.

ular weights of D101F lysozyme in solutions of ammonium sulfate are greater than those in solutions of sodium chloride indicating that the patch–patch interaction is more attractive in the presence of ammonium sulfate, as expected because the molal surface-tension increment of ammonium sulfate is larger than that of sodium chloride.

Elcock and McCammon [36] use a similar pmf model to describe protein–protein interactions for lysozyme in sodium chloride solutions. These authors use more precise calculations for determining a SASA potential, an electrostatic double-layer potential, and an electrostatic desolvation potential; the solvent-accessible surface area is calculated using a probe molecule rolled over the protein surface using the protein structure determined from crystallization data. The electrostatic desolvation potential or hydration force is included in this model to account for the unfavorable removal of water in the vicinity of charged surface groups. Using that model, Elcock and McCammon obtain a fit value of the scaling parameter for the SASA potential equal to 24 cal/molÅ^2. We assume that the hydration force is contained in the SASA potential. Because decreasing the scaling parameter raises the protein–protein repulsion, we expect that our value for κ should be less than that reported by Elcock and McCammon [36]. According to the model for the patch–patch interaction, κ_a should reflect the solvation energetics of only the mutated residue. The electrostatic desolvation potential should be less for the D101F lysozyme than that for the native form because a charged group has been replaced by a non-polar residue. The fit value of κ_a should reflect the non-polar solvation of the phenylalanine residue minus the unfavorable electrostatic desolvation penalty described above. Thus, we expect that our fit value of κ_a is similar to or slightly larger than the scaling parameter reported by Elcock and McCammon. The difference between these values may also be due to the simplifying assumptions of our model where we have neglected the shape of the protein.

4. Conclusions

The primary goal of this work is to probe the protein–protein interaction in concentrated salt

solutions and to propose a simple potential of mean force that can be used in phase-equilibrium calculations and as a protein-crystallization diagnostic. Previous salting-out studies of model compounds and of proteins have shown that the primary effect of salt is to increase the non-polar surface free energy of the protein. Generally, protein salting-out occurs because the non-polar part of the solvent-accessible surface area is reduced in the crystal due to the formation of protein–protein contacts. To incorporate this effect into our model, we relate the scaling parameter of a SASA potential with the surface free energy of the protein molecule.

The scaling parameter is expressed as a linear function of the salt molality where γ is the linear coefficient that is determined from the slope of a plot of scaling parameter vs. salt molality. We have determined γ from fitting the interaction between two protein surfaces and γ_a from fitting the contribution to the protein–protein interaction from a non-polar patch on the protein surface. For sodium–chloride solutions, the ratio of γ to the molal surface-tension increment of the salt is similar to that reported by Arakawa and Timasheff [26] for the ratio of the surface free energy increment to the same molal surface-tension increment. There is considerable scatter in the data obtained from ammonium–sulfate solutions, although it appears that the fit value of γ for these solutions is similar to that obtained in sodium chloride. In further support of the proposed potential of mean force, the obtained value of γ_a is within a factor of 2 for γ/f where f is the fraction of the protein surface that is non-polar.

The fit values of the scaling parameters for the patch–patch interaction are significantly greater than the corresponding values for the protein–protein interaction indicating that hydration forces make a significant contribution to the overall protein–protein potential of mean force. Because these forces are most likely longer-ranged than a solvent diameter, we do not expect that the SASA potential provides a truly realistic description for these forces. However, recent work has shown that protein–protein interactions under crystallizing conditions are accurately modeled using an adhesive potential [11]. Subsequently, we conclude that the dominant protein–protein interactions are short-ranged, consistent with the proposed potential of mean force. It is possible that the contributions of the non-polar groups and of the polar groups to the SASA potential can be separated in a manner similar to the empirical separation of the dispersion force and the SASA potential given in our model. Thus, we expect that the proposed potential of mean force may be useful for describing the hydrophobic interaction between proteins in concentrated salt solutions.

Acknowledgments

The authors are grateful to the Office for Basic Energy Sciences of the US Department of Energy and to the National Science Foundation (Grant #CTS-9530793) for financial support and to Prof. J. Kirsch for providing the D101F lysozyme and to Professor P. Curtis for help with cloning gene into P. Pastoris.

Appendix A: Calculation of the area of patch buried during protein–protein interaction

The effect of the mutation on protein–protein interactions is included in the model by placing circular hydrophobic patches on the surfaces of the spheres. The two-body potential is obtained from the combined area contained on the patches that is within a separation, r_c, of the opposing sphere multiplied by the negative of the non-polar surface free energy. The interaction is referred to as the patch–patch interaction, although the interaction includes all configurations where either patch is at least partially desolvated. Subsequently, non-local interactions of a patch and the uniform surface are also included in the overall potential.

This appendix derives the area of a patch within a separation, r_c of a second sphere as illustrated in Fig. 10. The area is determined by calculating the common arc length between the patch and concentric circles that are everywhere equidistant from the opposing sphere. The concentric circles are given by the dashed lines in Fig. 10. Integration of the common arc lengths to the cut-off separation gives the buried surface area.

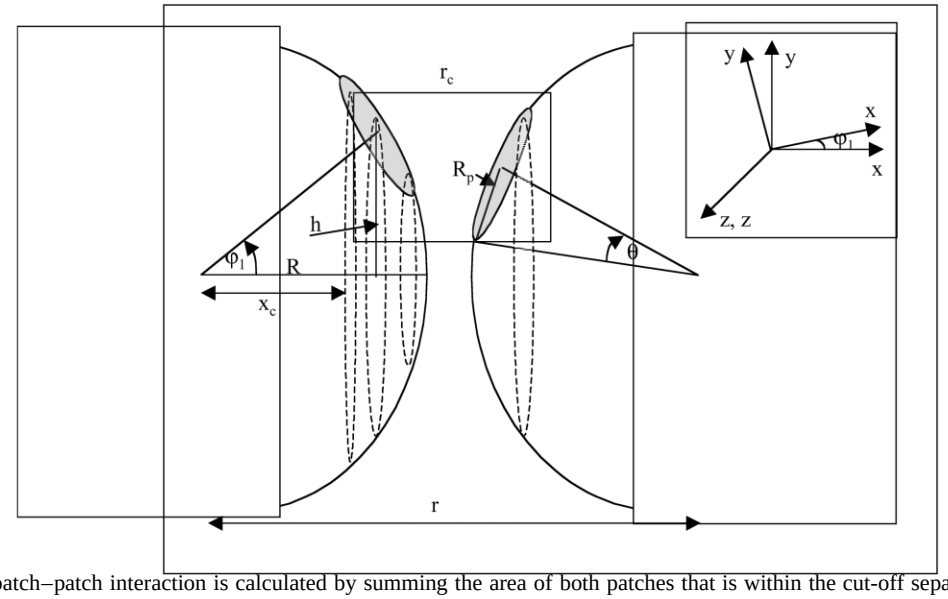

Fig. 10. The patch–patch interaction is calculated by summing the area of both patches that is within the cut-off separation r_c given by the area enclosed in the box. The variables are explained in Appendix A. The inset shows the coordinate system for the calculation along with that denoting the orientation of the patch given by the primed variables.

The centers of the concentric circles are at a distance x_o from the center of the spheres in which they are contained, and a distance $r-2x_o$ from the opposing sphere. The radii, h, are given by the chord theorem:

$$h(x_o)=\left[(R+x_o)(R-x_o)\right]^{1/2} \quad \text{(A.1)}$$

Subscript o refers to the coordinates of the concentric circles. The patch is approximated as circular with an area given by the surface area calculated from the crystal structure given in Table 1. The solid angle of the patch, θ is determined from fitting the area of the patch, A_p

$$A_p=2\pi R^2\int_0^{\theta}\sin\theta' d\theta' \quad \text{(A.2)}$$

The projected radius of the patch, R_p, is:

$$R_p=R\sin\theta \quad \text{(A.3)}$$

The equation of the boundary of the patch needs to be determined to calculate the intersection of the patch with the concentric circles. The equation for the boundary of the patch is given by:

$$y'^2_p+z'^2_p=R^2_p \quad \text{(A.4)}$$

where subscript p refers to the coordinates of the boundary of the patch. The coordinate transformation is calculated from the matrix, C, for a rotation given by φ_1.

$$C=\begin{vmatrix}\cos\varphi_1 & \sin\varphi_1 & 0\\ -\sin\varphi_1 & \cos\varphi_1 & 0\\ 0 & 0 & 1\end{vmatrix} \quad \text{(A.5)}$$

where

$$\begin{vmatrix}x'\\ y'\\ z'\end{vmatrix}=C\begin{vmatrix}x\\ y\\ z\end{vmatrix} \quad \text{(A.6)}$$

The corresponding equation for the boundary of the patch in the x,y,z coordinate system is calculated from Eq. (A.5) and Eq. (A.3):

$$[y_p\cos\varphi_1-x_p\sin\varphi_1]^2+z^2_p=R^2_p \quad \text{(A.7)}$$

The equation for the concentric spheres as a function of x is:

$$y_o^2+z_o^2=h^2=(R-x_o)(R+x_o) \quad \text{(A.8)}$$

Eq. (A.6) and Eq. (A.7) are solved simultaneously for the intersection of the concentric circles with the boundary of the patch as a function of x. The arc-length, δ common to the circle and the patch is determined from the points of intersection given by z_i and y_i:

$$\delta(x)=2h(x)\arctan\left|\frac{z_i(x)}{y_i(x)}\right| \quad \text{if } y_i>0 \quad \text{(A.9)}$$

$$\delta(x)=2h(x)\left[\pi-\arctan\left|\frac{z_i(x)}{y_i(x)}\right|\right] \quad \text{if } y_i<0$$

If the boundary of the patch is contained within the concentric circle, then

$$\delta(x)=2\pi h(x) \quad \text{(A.10)}$$

The area is determined by integration over x where the lower limit of integration is given by:

$$x_c=\frac{r-r_c}{2} \quad \text{(A.11)}$$

Thus, the area is given by:

$$\text{Area}=\int_{x_c}^{R}\delta(x)\left[1+\frac{x}{(R^2-x^2)^{1/2}}\right]dx \quad \text{(A.12)}$$

References

[1] E.J.W. Verwey, J.T.G. Overbeek, Theory of the Stability of Lyophobic Colloids, Elsevier Publishing Company, Amsterdam, 1948.

[2] J. Israelachvili, Intermolecular and Surface Forces: with Applications to Colloidal and Biological Systems, Academic Press, London, 1992.

[3] K.D. Collins, M.W. Washabaugh, The Hofmeister effect and the behavior of water at interfaces, Q. Rev. Biophys. 18 (1985) 323–421.

[4] D.R. Robinson, W.P. Jencks, The effect of concentrated salt solutions on the activity coefficient of acetyltetraglycine ethyl ester, J. Am. Chem. Soc. 87 (1965) 2470–2479.

[5] P.K. Nandi, D.R. Robinson, The effects of salts on the free energy of the peptide group, J. Am. Chem. Soc. 94 (1972) 1299–1307.

[6] P.H. von Hippel, T. Schleich, Ion effects on the solution structure of biological macromolecules, Acc. Chem. Res. 2 (1969) 257–265.

[7] W.G. McMillan, J.E. Mayer, The statistical thermodynamics of multicomponent systems, J. Chem. Phys. 13 (1945) 276–305.

[8] A. George, W.W. Wilson, Predicting protein crystallization from a dilute solution property, Acta Cryst. D51 (1994) 361–365.

[9] B. Guo, S. Kao, H. McDonald, A. Asanov, L.L. Combs, W.W. Wilson, Correlation of second virial coefficients and solubilities useful in protein crystal growth, J. Cryst. Growth 196 (1999) 424–433.

[10] C. Haas, J. Drenth, W.W. Wilson, Relation between the solubility of proteins in aqueous solutions and the second virial coefficient of the solution, J. Phys. Chem. B. 103 (1999) 2808–2811.

[11] D. Rosenbaum, P.C. Zamora, C.F. Zukoski, Phase behavior of small attractive colloidal particles, Phys. Rev. Lett. 76 (1996) 150–153.

[12] C. Haas, J. Drenth, The protein-water phase diagram and the growth of protein crystals from aqueous solution, J. Phys. Chem. B 102 (1998) 4226–4232.

[13] P.R. ten Wolde, D. Frenkel, Enhancement of protein crystal nucleation by critical density fluctuations, Science 277 (1997) 1975–1978.

[14] C.J. Coen, H.W. Blanch, J.M. Prasnitz, Salting-out of aqueous proteins: phase equilibria and intermolecular potentials, AIChE J. 41 (1995) 996–1004.

[15] V.L. Vilker, C.K. Colton, K.A. Smith, Osmotic pressure of concentrated protein solutions: the effect of concentration and pH in saline solutions of bovine serum albumin, J. Colloid Int. Sci. 79 (1981) 548–566.

[16] S. Nir, Van-der-Waals interactions between surfaces of biological interest, Prog. Surf. Sci. 8 (1976) 1–58.

[17] F. Eisenhaber, Hydrophobic regions on protein surfaces, derivation of the solvation energy from their area distribution in crystallographic protein surfaces, Protein Sci. 5 (1996) 1676–1686.

[18] D. Eisenberg, A.D. McLachlan, Solvation energy in protein folding and binding, Nature 319 (1986) 199–203.

[19] R.B. Hermann, Theory of hydrophobic bonding. II. The correlation of hydrocarbon solubility in water with solvent cavity surface area, J. Phys. Chem. 76 (1972) 2754–2759.

[20] D. Asthagiri, B.L. Neal, A.M. Lenhoff, Calculation of short-range interactions between proteins, Biophys. Chem. 78 (1999) 219–231.

[21] W. Eberstein, Y. Georgalis, W. Saenger, Molecular interactions in crystallizing lysozyme solutions studied by photon correlation spectroscopy, J. Cryst. Growth 143 (1994) 71–78.

[22] O.D. Velev, E.W. Kaler, A.M. Lenhoff, Protein interactions in solution characterized by light and neutron scattering: comparison of lysozyme and chymotrypsinogen, Biophys. J. 75 (1998) 2682–2697.

[23] E.J. Cohn, Proteins, amino acids and peptides, in: E.J. Cohn, J.T. Edsall (Eds.), The Solubilty of Proteins, Reinhold, New York, 1943, p. 586.
[24] W. Melander, C. Horvath, Salt effects on hydrophobic interactions in precipitation and chromatography of proteins: An interpretation of the lyotropic series, Arch. Biochem. Biophys. 183 (1977) 200–215.
[25] A. McPherson, Current approaches to macromolecular crystallization, Eur. J. Biochem. 189 (1990) 1–23.
[26] T. Arakawa, S.N. Timasheff, Mechanism of protein salting in and salting out by divalent cation salts: balance between hydration and salt binding, Biochemistry 23 (1984) 5912–5923.
[27] T. Arakawa, S.N. Timasheff, Theory of protein solubility, Meth. Enzymology 114 (1985) 49–77.
[28] F. Hofmeister, Zue Lehre von der Wirkung der Saltze, Arch. Expt. Pathol. Pharmakol. 24 (1888) 247–260.
[29] C. Munshi, H.C. Lee, High-level expression of recombinant Aplysia ADP-Ribosyl Cyclase in *Pichia pastoris* by fermentation, Prot. Express. Purific. 11 (1997) 104–110.
[30] A.J. Sophianopoulus, C.K. Rhodes, D.N. Holcomb, K.E. van Holde, Physical studies of lysozyme. I. Characterization, J. Biol. Chem. 237 (1962) 1107–1112.
[31] D. Shugar, The measurement of lysozyme activity and the ultra-violet inactivation of lysozyme, Biochim. Biophys. Acta 8 (1952) 302–309.
[32] W.H. Stockmayer, Light scattering in multi-component solutions, J. Chem. Phys. 18 (1950) 58–61.
[33] M. Connolly, The molecular surface package, J. Mol. Graph. 11 (1993) 139–143.
[34] R.A. Curtis, A. Montaser, J.M. Prausnitz, H.W. Blanch, Protein–protein and protein–salt interactions in aqueous protein solutions containing concentrated electrolytes, Biotech. Bioeng. 57 (1998) 11–21.
[35] M.A. Ries-Kautt, A.F. Ducruix, Relative effectiveness of various ions on the solubility and crystal growth of proteins, J. Biol. Chem. 264 (1989) 745–750.
[36] A.H. Elcock, J.A. McCammon, Calculation of weak protein–protein interactions: the pH dependence of the second virial coefficient, Biophys. J. 80 (2001) 613–625.
[37] D.E. Kuehner, J. Engmann, F. Fergg, M. Wernick, H.W. Blanch, J.M. Prausnitz, Lysozyme net charge and ion binding in concentrated aqueous electrolyte solutions, J. Phys. Chem. B 103 (1999) 1368–1374.