Microscopic Determination of the Phase Diagrams of Lysozyme and γ-Crystallin Solutions

Giuseppe Pellicane, Dino Costa, and Carlo Caccamo*

Istituto Nazionale per la Fisica della Materia (INFM) and Dipartimento di Fisica, Università degli Studi di Messina, Contrada Papardo, C.P. 50−98166 Messina, Italy

Received: March 26, 2004; In Final Form: April 20, 2004

We determine the phase diagrams of two widely studied protein solutions, namely, lysozyme and γ-crystallin in water and added salt, within a short-range, pair-potential representation of macroparticle interactions. A one-parameter fit of the model is performed on the basis of static or dynamic light scattering and self-interaction chromatography data. Demixing and crystallization curves, as calculated through numerical approaches, turn out to be in good agreement with experimental protein-rich−protein-poor and solubility envelopes. Our findings highlight the minimal assumptions about the nature of the microscopic potential allowing a satisfactory reproduction of the phase-diagram topology of globular protein solutions.

The most powerful technique used in biological and medical sciences to explore the 3D structure of proteins is represented by X-ray crystallography.[1] This experimental method requires highly ordered crystals, whereas the control of protein crystallization from mother solutions is frequently hampered by the adoption of trial-and-error procedures.[1,2] The diffuse resort to empirical protocols is also due to the lack of a microscopic description of the system phase behavior. For instance, understanding how interactions among proteins are altered by the solvent is a difficult task, also considering that specific salt effects determine a ranking of various ions, which is well known as the Hofmeister series.[3]

Despite the apparently enormous diversity, a number of crystallizing globular proteins share some common features. In particular, the second virial coefficient B_2 of such systems falls into a narrow window of small negative values.[4] Closely related to this aspect, the phase diagram is characterized by a metastable protein−protein demixing region just below the solubility line,[5] a peculiar feature that is of crucial interest for different reasons. In fact, experimental and numerical studies[5−7] have shown that the proximity of the critical point to the solubility line should favor the formation of good crystals. Besides the crystallization properties, the protein demixing binodal also plays an important role in several human diseases.[8]

The aforementioned evidences have been rationalized in terms of effective protein−protein interactions described through short-range, centrosymmetric potentials.[9−14] However, such models have hitherto yielded only a qualitative reproduction of the experimental phase diagram and have allowed at best to fit the demixing line,[10,11] implicitly suggesting their minor predictive power. In particular, they fail to capture the sensitivity of the metastable demixing line to solution conditions.[14]

It is evident that a detailed analysis of the complex behavior of protein solutions would require a refined, highly specific treatment of microscopic interactions; in particular, recent studies have pointed out that a model description should embody at least the asymmetry of the molecular shape[15] or the distributed-site nature of interactions among proteins.[16−18] However, we report evidence in this letter that a reasonable reproduction of the phase diagram of prototype globular protein solutions may be offered by a central, pair-potential representation of macromolecular interactions. We consider in this letter lysozyme solutions in water and sodium chloride added salt and γ_{II}-crystallin calf lens in sodium phosphate, for which extended experimental results have been reported.[9,19−22] .

We assume that the effective potential is given by the sum of a repulsive, electrostatic part in the form of a Debye−Hückel contribution,[23]

$$v_{DH}(r) = \frac{1}{4\pi\epsilon_0\epsilon_r}\left[\frac{z_p e}{1+\frac{\chi_{DH}\sigma}{2}}\right]^2 \frac{\exp[-\chi_{DH}(r-\sigma)]}{r} \qquad (1)$$

and an attractive part represented by a short-range, van der Waals term,[24]

$$v_{HA}(r) = -\frac{A_H}{12}\left[\frac{\sigma^2}{r^2} + \frac{\sigma^2}{r^2-\sigma^2} + 2\ln\frac{r^2-\sigma^2}{r^2}\right] \qquad (2)$$

The resulting total interaction is

$$v_{EFF}(r) = \begin{cases} \infty & r < \sigma + \delta \\ v_{HA}(r) + v_{DH}(r) & r \geq \sigma + \delta \end{cases} \qquad (3)$$

Here σ represents the effective protein diameter, A_H is the Hamaker constant, $Q = z_p e$ is the net charge on the protein in electron units, ϵ_r and ϵ_0 are respectively the (solution) relative and vacuum dielectric constants. $\chi_{DH} = [4\pi L_b 1000 N_a I_s]^{1/2}$ is the inverse Debye screening length, where I_s is the ionic strength, $L_b = e^2/(2\pi\epsilon_0\epsilon_r k_B T_{20})$ is the Bjerrum length, k_B is the Boltzmann constant, $T_{20} \equiv 293.15$ K, and N_a is Avogadro's number. The Stern layer thickness δ, related to the intrinsic size of counterions that condense on the macromolecular surface, is introduced in eq 3 to circumvent the singularity of the attractive term at $r = \sigma$.[21] We shall further comment on the functional form of the potential of mean force reported in eq 3.

In our calculations, we fix all parameters entering eq 3, except the Hamaker constant, to values commonly accepted on the basis of several pieces of experimental evidence.[20,21,25,26] As for

* Corresponding author. E-mail: carlo.caccamo@unime.it.

10.1021/jp048651v CCC: $27.50
Published on Web 05/14/2004

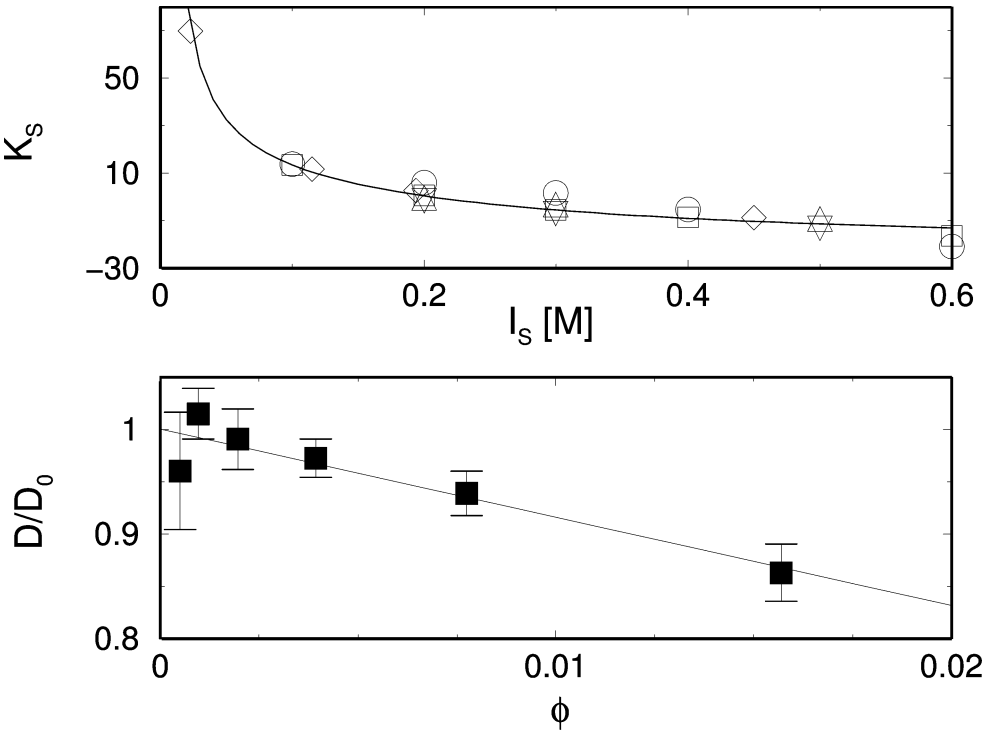

Figure 1. (Top) interaction factor vs ionic strength for lysozyme and NaCl. (○) SIC;[22] (□) SLS;[9,14] (◊) SLS;[21] (△) SLS[27] at $T = 303.15$ K; (▽) SLS[27] at $T = 293.15$ K. (—) DLVO potential with $A_H = 8.61k_BT_{20}$. (Bottom, ■) scaled diffusion coefficient vs protein volume fraction with error bars for γ-crystallin and sodium phosphate.[20] The linear fit $D/D_0 = 1 + K_D\phi$ is also shown.

lysozyme, the Hamaker constant is used as a free parameter to best fit the experimental interaction factor, defined as

$$K_S = K_S^{HS} + 24\int_{\delta/\sigma}^{\infty} x^2\{1 - \exp[-\beta v_{EFF}(x)]\}\, dx \qquad (4)$$

where $x = ((r/\sigma) - 1)$ and $K_S^{HS} = 8$ is the corresponding interaction factor of a fluid of hard spheres. The interaction factor is proportional to the second virial coefficient, $B_2 = K_S\nu/2M$, where $\nu = 0.703$ mL/g and $M = 14300$ Da are respectively the specific volume and the atomic mass of lysozyme. We specifically best fit the self-interaction chromatography (SIC) results[22] as well as the static light scattering (SLS) data from various sources[9,14,21,27] for a lysozyme solution in the range [0.1−0.6]M NaCl at a pH close to that of the experimental phase diagram[5] (Figure 1). An average Hamaker constant is then obtained, $A_H \simeq 8.61$ (in k_BT_{20} units), with a maximum standard deviation equal to 0.26. Other quantities of eq 1−3 for lysozyme are $\sigma = 3.6$ nm, $Q = 10$e, and $\delta = 0.18$ nm.[21,25] Our average is comparable with $A_H = 8.1k_BT$, a value obtained by Muschol and Rosenberger from a similar fit of their experimental data (with $Q = 10.7$e),[21] also displayed in Figure 1.

As for γ-crystallin, the slope of the collective diffusion coefficient, $K_D = -8.43$ for 0.24 M sodium phosphate, has been calculated from experimental data in a very dilute regime[20] (bottom panel of Figure 1). We are in fact unaware of B_2 measurements under the same conditions of the experimental phase diagram. K_D is related to the effective potential by

$$K_D = K_D^{HS} + \sigma\int_{\delta/\sigma}^{\infty}\{1 - g(r)\}F(x)\, dx \qquad (5)$$

where $K_D^{HS} = 1.56$ is the slope of the diffusion coefficient of a fluid of hard spheres. In eq 5, we have approximated the radial distribution function by its dilute counterpart $g(r) = \exp[-\beta v_{EFF}(r)]$ and have used for the hydrodynamic function $F(r)$ the expression given by Batchelor.[28] To calculate the Hamaker constant, we set the net charge Q equal to zero because experimental data were taken close to the isoelectric point of the solution, and the diameter $\sigma \simeq 3.7$ nm, according to the dry-volume estimate of ref 11. We obtain in this way $A_H \simeq 6.68k_BT_{20}$. To check both the robustness and the flexibility of our procedure, we have verified that the shape of the potential, resulting from a given fit, is weakly dependent on the choice of free parameters as well as on their variation inside a reasonable physical range.

On the basis of the above models, we have calculated the phase diagrams of lysozyme and γ-crystallin solutions at $I_s = 0.51$ M (well inside the molarity range fitted in Figure 1) and 0.24 M, respectively. Our results are compared with the experimental demixing and solubility lines reported in refs 5 and 19. Gibbs ensemble Monte Carlo simulations[29] are carried out on a system composed of 1024 particles (with a 3σ interaction cutoff) to characterize the protein-rich−protein-poor phase coexistence. As for the solubility line, the free energy of the fluid phase is calculated in the framework provided by the refined hybrid mean spherical approximation theory,[30,31] whereas for the solid phase, we use the expression for the chemical potential:

$$\mu_{CRY} = \mu_0 - \frac{n_s\epsilon_{EFF}}{2} - 3k_BT\ln(\lambda^*) \qquad (6)$$

Equation 6 constitutes a simplified version of the cell theory[32] that has been extensively applied to systems characterized by short-range interactions.[16,33−35] Although the cell theory for isotropic models constitutes an approximate representation of the free energy of the solid phase, by underestimating the entropic contribution to the free energy,[33] it provides a direct link to the essential properties of the protein crystals, namely, the average number of contacts, n_s, and the translational freedom along one axis, λ (with $\lambda^* = \lambda/(\sigma + \delta)$), of the protein

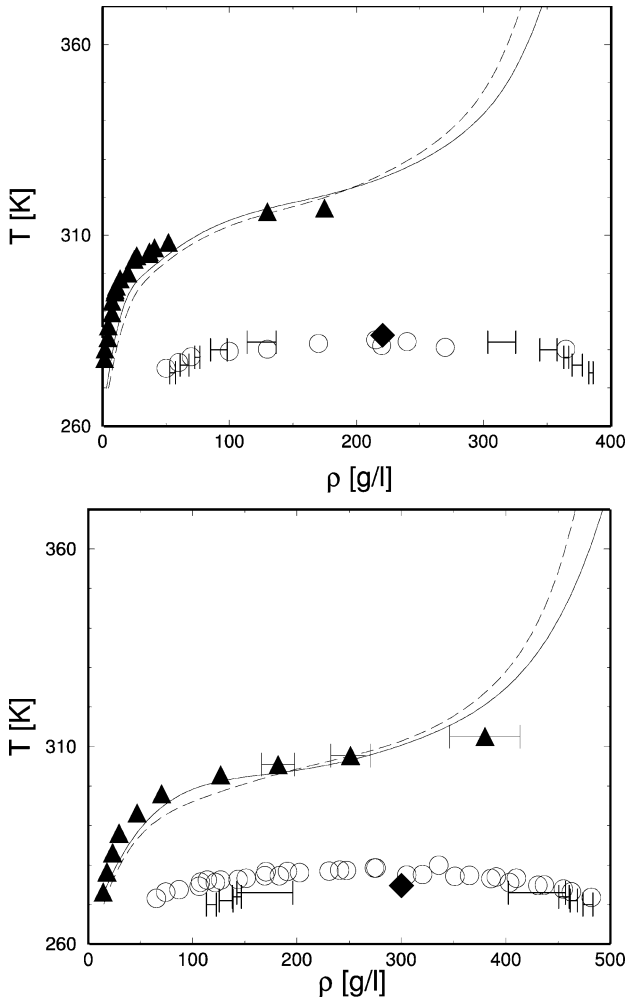

Figure 2. (Top) phase diagram of the lysozyme solution in NaCl at $I_s = 0.51$ M. (○, ▲) Cloud and solubility points, respectively.[5] (Bottom) phase diagram of the γ-crystallin solution in sodium phosphate at $I_s = 0.24$ M. (○, ▲) Cloud and solubility points.[19] In both panels, bars with diamonds are theoretical predictions for the demixing curves and corresponding critical points; full and dashed lines are theoretical crystallization curves with $n_s = 10$ and 9, respectively (eq 6).

inside the unit cell.[36] Other parameters entering eq 6 are μ_0, the standard part of the chemical potential, and ϵ_{EFF}, the minimum of the effective potential (eq 3). A constant value of $\log[V_p/(\sigma + \delta)^3]$ has been included in μ_0, where V_p is the protein volume incorporating the condensed counterion layer.[35] The coexisting fluid branch is then found by equating the chemical potentials of both phases.[16]

We show in Figure 2 that the theoretical demixing envelope for lysozyme reasonably coincides with the experimental cloud points. The critical-point temperature and density have been estimated by fitting the simulation coexistence points through the scaling law for the densities (with an effective nonclassical exponent $\beta = 0.325$) and the law of rectilinear diameters.[37] The flatness of the coexistence curve allows for the fine determination of the critical temperature $T_{cr} = 284 \pm 2$ K, with a broader estimate of the critical density $\rho_{cr} = 221 \pm 12$ g/L. Our critical parameters closely agree with the experimental outcome $T_{cr} = 282.5$ K and $\rho_{cr} = 255 \pm 30$ g/L reported in ref 5. We find a reasonable reproduction of the demixing line also for γ-crystallin (Figure 2), along with a good prediction of the critical temperature and density, $T_{cr} = 275 \pm 3$ K, $\rho_{cr} = 300 \pm 30$ g/L, to be compared with the experimental values $T_{cr} \simeq 278$ K, $\rho_{cr} = 269 \pm 20$ g/L.[19] The availability of additional data for γ-crystallin might lead to even a better agreement with experimental results; for the present, we predict almost quantitatively the critical parameters and the high-density portion of the demixing envelopes, thus improving upon previous results based on other pair, central potential models.[10,11]

As far as the solubility line for lysozyme is concerned, the theoretical crystallization boundary, calculated with $n_s = 9$ and $\lambda = 0.03$ nm in eq 6, is in good agreement with the experimental curve. We obtain similar results if another plausible number of contacts, namely, $n_s = 10$ (with $\lambda = 0.017$ nm), is used (Figure 2). Crystallization curves for lysozyme corresponding to other parameter pairs, such as [$\lambda = 0.02$ nm, $n_s = 10$] or [$\lambda = 0.025$ nm, $n_s = 9$], closely bracket the experimental line, with a temperature discrepancy that does not exceed ~6% at the critical density. All these values of λ are compatible with the temperature factors of the Bragg intensities in lysozyme crystals, where the typical range for λ is [0.014−0.033] nm;[38] a number of contacts between 8 and 10 has been reported for a number of protein crystals in ref 39 and specifically for lysozyme in ref 40. Also, six contacts have been reported for lysozyme.[35] In this case, close agreement with the experimental solubility data around the critical density is obtained for a value of λ not lower than ≃0.1 nm. Such an overestimate of λ is clearly related to the aforementioned underestimate of the entropic contribution in eq 6. Obviously, the role of the anisotropic interactions should be explicitly taken into account for a quantitative description of the crystal phase. Similar accuracy is obtained for calf lens proteins (Figure 2) for comparable values of λ, 0.036 and 0.020 nm for $n_s = 9$ and 10, respectively. In conclusion, it appears that a fairly accurate description of the solubility boundaries is achieved when reasonable values for n_s and λ are inserted into the approximate but simple formula (eq 6) of the chemical potential of the solid phase.

The contributions to the potential of mean force envisaged in eq 3 possess the same functional form of the DLVO theory of the stability of charged colloidal suspensions.[23] The extension of a "colloid physics approach" to biological systems, such as protein solutions, has long been debated. (See, for example, refs 13, 25, and 41−45 and references therein.) In agreement with previous studies, the validation of the theoretical procedure here exploited is mainly derived from a positive, nontrivial parametrization of phenomenological observations. In particular, we envisage the functional form of the attractive potential in eq 2 as being flexible enough to accommodate not only van der Waals forces but also other non-DLVO interactions such as hydration effects.[25,41,46] We assume in fact that these forces act to inflate artificially the value of the Hamaker constant, which plays a different role from the original one, related only to dispersive attractions. The level of description here adopted cannot account for several, highly specific effects related, for instance, to the identity of the salt used to prepare the protein solution. Within our framework, such effects are taken into account only indirectly, through the fit of the experimental data. In other modeling cases, an additional short-range attractive potential is introduced to embody the effects of salt type, pH, and temperature on the protein−protein interaction.[21,27,47] As for the Debye−Hückel-like treatment in eq 1, we note that electrostatic interactions are almost screened out at biological salt molarities, constituting a minor correction to the leading attractive contribution of eq 2.[3,21,26] Interestingly enough, the electrostatic term yields the correct dependence of the potential of mean force on the ionic strength for lysozyme (top panel of Figure 1) on the condition that the order of magnitude of effective protein interactions is properly set by the fit of the Hamaker constant.

As far as a comparison with a more general, multisite representation of protein interactions is concerned, our results support the observation that in the fluid phase, close to the critical point, a distributed approach can be well approximated by its spherically averaged counterpart.[17] Moreover, it has been recently underlined that a nonmonotonic variation of B_2 as a function of salt concentration, as reported in ref 44, can be related to the discreteness of the charge pattern on the protein surface.[18] Our study confirms the substantial adequacy of the uniform smeared-out approximation for the charge distribution[18] when B_2 is monotonically decreasing with the salt molarity as in Figure 1, whereas, obviously, the Debye−Hückel approach is unable to cope with a nonmonotonic dependence on ionic strength.

We have assessed in this letter that basic assumptions about the effective interaction among globular proteins in solution are sufficient to reproduce reasonably the phase boundaries of such systems. In particular, the parametrization of a central, pair potential based on the experimental second virial coefficient or on the collective diffusion coefficient allows for the confident prediction of true cloud points and solubility lines. Because much less effort is required to determine the experimental B_2 or k_D, the advantage of the procedure here exploited becomes evident. Our predictions are also accurate for high protein concentrations, although experimental measurements of B_2 and k_D are usually carried out in a dilute regime. These findings agree with the linearity observed in several Debye plots of scattering data,[21] suggesting that B_2 values measured in undersaturated solutions remain essentially unaltered in the supersaturated regime.

Acknowledgment. We thank Professors J. Janin, J. Doucet, C. Redfield, and B. W. Sigurskjold for useful correspondence and materials provided.

References and Notes

(1) Chayen, N. E. *Trends Biotechnol.* **2002**, *20*, 98.
(2) McPherson, A. *Preparation and Analysis of Protein Crystals*; Krieger: Malabar, India, 1982.
(3) Broide, M.; Tominc, T. M.; Saxowsky, M. D. *Phys. Rev. E* **1996**, *53*, 6325. Grigsby, J. J.; Blanch, H. W.; Prausnitz, J. M. *Biophys. Chem.* **2001**, *91*, 231.
(4) George, A.; Wilson, W. *Acta Crystallogr., Sect. D* **1994**, *50*, 361.
(5) Muschol, M.; Rosenberger, F. *J. Chem. Phys.* **1997**, *107*, 1953.
(6) ten Wolde, P. R.; Frenkel, D. *Science* **1997**, *277*, 1975.
(7) Haas, C.; Drenth, J. *J. Phys. Chem. B* **2000**, *104*, 368.
(8) Eaton, W. A.; Hofrichter, J. *Advances in Protein Chemistry*; Academic Press: San Diego, CA, 1990; Vol. 40, p 63.
(9) Rosenbaum, D. F.; Zamora, P. C.; Zukoski, C. F. *Phys. Rev. Lett.* **1996**, *76*, 150.
(10) Lomakin, A.; Asherie, N.; Benedek, G. B. *J. Chem. Phys.* **1996**, *104*, 1646.
(11) Malfois, M.; Bonnete, F.; Belloni, L.; Tardieu, A. *J. Chem. Phys.* **1996**, *105*, 3290.
(12) Piazza, R.; Peyre, V.; Degiorgio, V. *Phys. Rev. E* **1998**, *58*, R2733.
(13) Poon, W. C. K. *Phys. Rev.* E **1997**, *55*, 3762. Poon, W. C. K.; Egelhaaf, S. U.; Beales, P. A.; Salonen, A.; Sawyer, L. *J. Phys.: Condens. Matter* **2000**, *12*, L569.
(14) Rosenbaum, D. F.; Kulkarni, A.; Ramakrishnan, S.; Zukoski, C. F. *J. Chem. Phys.* **1999**, *111*, 9882.
(15) Neal, B. L.; Asthagiri, D.; Velev, O. D.; Lenhoff, A. M.; Keler, E. W. *J. Cryst. Growth* **1999**, *196*, 377. Sun, N.; Waltz, J. Y. *J. Colloid Interface Sci.* **2001**, *234*, 90.
(16) Sear, R. P. *J. Chem. Phys.* **1999**, *111*, 4800.
(17) Lomakin, A.; Asherie, N.; Benedek, G. B. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 9465.
(18) Allahyarov, E.; Lowen, H.; Louis, A. A.; Hansen, J. P. *Europhys. Lett.* **2002**, *57*, 731. Allahyarov, E.; Lowen, H.; Hansen, J. P.; Louis, A. A. *Phys. Rev. E* **2003**, *67*, 051404.
(19) Broide, M. L.; Berland, C. R.; Pande, J.; Ogun, O. O.; Benedek, G. B. *Proc. Natl. Acad. Sci. U.S.A.* **1991**, *88*, 5660. Berland, C. R.; Thurston, G. M.; Kondo, M.; Broide, M. L.; Pande, J.; Ogun, O.; Benedek, G. B. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 1214.
(20) Fine, B. M.; Lomakin, A.; Ogun, O. O.; Benedek, G. B. *J. Chem. Phys.* **1995**, *104*, 326.
(21) Muschol, M.; Rosenberger, F. *J. Chem. Phys.* **1995**, *103*, 10424.
(22) Tessier, P. M.; Vandrey, S. D.; Berger, B. W.; Pazhianur, R.; Sandler, S. I.; Lenhoff, A. M. *Acta Crystallogr., Sect. D* **2002**, *58*, 1531.
(23) Verwey, E. J. W.; Overbeek, J. T. G. *Theory of the Stability of Lyophobic Colloids*; Elsevier: Amsterdam, 1948.
(24) Derjaguin, B. V.; Landau, L. D. *Acta Physicochim. URSS* **1941**, *14*, 633.
(25) Beretta, S.; Chirico, G.; Baldini, G. *Macromolecules* **2000**, *33*, 8663.
(26) Kuehner, D. E.; Heyer, C.; Remsch, C.; Farnefeld, U. M.; Blanch, H. W.; Prausnitz, J. M. *Biophys. J.* **1997**, *73*, 3211. Kulkarni, A. M.; Chatterjee, A. P.; Schweizer, K. S.; Zukoski, C. F. *Phys. Rev. Lett.* **1999**, *83*, 4554.
(27) Bonneté, F.; Finet, S.; Tardieu, A *J. Cryst. Growth* **1999**, *196*, 403.
(28) Batchelor, J. K. *J. Fluid Mech.* **1976**, *74*, 1.
(29) Panagiotopoulos, A. *Mol. Phys.* **1987**, *61*, 813.
(30) Zerah, G.; Hansen, J.-P. *J. Chem. Phys.* **1986**, *84*, 2336.
(31) Caccamo, C. *Phys. Rep.* **1996**, *274*, 1.
(32) Lennard-Jones, J. E.; Devonshire, A. F. *Proc. R. Soc. London, Ser. A* **1937**, *163*, 53.
(33) Curtis, R. A.; Blanch, H. W.; Prausnitz, J. M. *J. Phys. Chem. B* **2001**, *105*, 2445.
(34) Asherie, N.; Lomakin, A.; Benedek, G. B. *Phys. Rev. Lett.* **1996**, *77*, 4832.
(35) Warren, P. B. *J. Phys.: Condens. Matter* **2002**, *14*, 7617.
(36) A fully microscopic determination of the solid free energy has been attempted by us in previous work (Pellicane, G.; Costa, D.; Caccamo, C. *J. Phys.: Condens. Matter* **2003**, *15*, 365) in terms of a perturbation theory based on the reference free energy of a fcc lattice of hard spheres. The latter, however, does not correspond to the structure of real protein crystals, and the experimental solubility envelope of lysozyme turns out to be overestimated.
(37) Frenkel, D.; Smit, B. *Understanding Molecular Simulation*; Academic Press: London, 1996.
(38) Doucet, J.; Benoit, J. *Nature* **1987**, *325*, 643.
(39) Janin, J.; Rodier, F. *Proteins: Struct., Funct., Genet.* **1995**, *23*, 580.
(40) Pedersen, T. G.; Sigurskjold, B. W.; Andersen, K. V.; Kjaer, M.; Poulsen, F. M.; Dobson, C. M.; Redfield, C. *J. Mol. Biol.* **1991**, *218*, 413.
(41) Farnum, M.; Zukoski, C. *Biophys. J.* **1999**, *76*, 2716.
(42) Rowe, A. J. *Biophys. Chem.* **2001**, *93*, 93.
(43) Piazza, R. *J. Cryst. Growth* **1999**, *196*, 415.
(44) Petsev, D. N.; Vekilov, P. G. *Phys. Rev. Lett.* **2000**, *84*, 1339.
(45) Böstrom, M.; Williams, D. R. M.; Ninham, B. W. *Phys. Rev. Lett.* **2001**, *87*, 168103.
(46) Chernov, A. A. *Phys. Rep.* **1997**, *288*, 61.
(47) Kuehner, D. E.; Blanch, H. W.; Prausnitz, J. M. *Fluid Phase Equilib.* **1996**, *116*, 140.