The Combined Simulation Approach of Atomistic and Continuum Models for the Thermodynamics of Lysozyme Crystals

Jaeeon Chang, Abraham M. Lenhoff, and Stanley I. Sandler*

Center for Molecular and Engineering Thermodynamics, Department of Chemical Engineering, University of Delaware, Newark, Delaware 19716

Received: May 17, 2005; In Final Form: August 12, 2005

We have studied the thermodynamic properties of hen egg white lysozyme crystals using a novel simulation method combining atomistic Monte Carlo simulation to calculate van der Waals interactions and the boundary element method to solve the Poisson−Boltzmann equation for the electrostatic interactions. For computational simplicity, we treat the protein as a rigid body, using the crystallographic coordinates of all non-hydrogen atoms of the protein to describe the detailed shape. NVT Monte Carlo simulations are carried out for tetragonal and orthorhombic crystals to obtain the van der Waals energy, incorporating an implicit solvation effect. For crystal phases, an optimally linearized Poisson−Boltzmann equation is used to include the effect of the Donnan equilibrium of the salt ions. The Helmholtz energy is obtained from expanded ensemble Monte Carlo simulations. By using the force field parameters that had previously been tuned for the solution properties, reasonable agreement with experiment is found for the crystallization energy of the tetragonal form. The prediction of the entropy is also reasonable with a slight underestimation suggesting the release of a few water molecules per protein during the crystallization. However, the predictions of the properties of the orthorhombic crystal are poor, probably due to differences in the solvation structure as indicated by experiments, and also as a result of the approximate force field used.

I. Introduction

Protein crystals are used to determine the three-dimensional structure of proteins by X-ray crystallography and to maintain a uniform release rate of pharmaceutical products, for example, in microcrystalline, time-release insulin. Obtaining protein crystals in the laboratory can be a difficult, time-consuming task. Generally, proteins crystallize only in a narrow range of solution conditions that are difficult to predict a priori. At present, determining the solution conditions at which a protein will crystallize is largely done by empirical trial-and-error screening methods, which is a slow and tedious process. An important, largely unsolved problem is being able to predict the crystallization conditions and, more importantly, the resulting crystal form.

There is empirical evidence that crystallization occurs under solution conditions where the osmotic second virial coefficient is slightly negative, indicating a weak net attraction between the proteins in solution; these conditions are referred to as the crystallization slot.[1] The interactions between proteins in solution depend on pH, which affects the net charge of the protein, the ionic strength of the solution, the salt type, the concentration of precipitants (for example, polymers) that may be used, as well as the usual variables such as protein concentration and temperature. If the solution conditions are such that the net interactions between the proteins are repulsive or too weakly attractive, crystallization will not occur. However, if the interaction is too attractive, agglomeration will be too rapid for the proteins to orient properly, and an unstructured precipitate or a gel may form. Further, even when a protein does crystallize, different crystal forms can be obtained with relatively small changes in solution conditions.

A molecular level statistical mechanical study of protein crystallization and precipitation could provide insights into how the details of the protein interactions affect the solubility and the stability of possible crystal phases. There have been several theoretical and simulation studies to explain the phase behavior of protein solutions using simple isotropic models.[2−6] While the fluid phase behavior of solutions of globular proteins can be described qualitatively by using an isotropic potential, it is not applicable for the quantitative prediction of fluid−solid equilibrium. This is because the formation of a protein crystal is the result of specific interactions between local surface regions of proteins. Among many factors, the geometrical compatibility between the docking surfaces plays a significant role in determining the crystal structure. A few theoretical studies have considered the effect of anisotropic interactions on protein solubility,[7−9] and it was found that the protein interactions must be highly anisotropic to explain experimental data. Several modeling studies of the osmotic second virial coefficient considered to some extent the effect of anisotropy in the charge distribution and the shape of proteins.[10−12] However, the prediction of the condensed phase properties of the proteins at a fully atomistic level is very demanding of computational resources, and simulations have often been done using coarse-grained models with an implicit or continuum treatment of solvent water.[6,13,14]

Hen egg white lysozyme has been the subject of many experimental studies, and its thermodynamic properties are fairly well-known.[15−22] Depending upon the solution conditions, lysozyme crystallizes in several forms, including tetragonal, orthorhombic, monoclinic, and triclinic structures, and there is a well-known transition from the tetragonal (low-temperature)

* To whom correspondence should be addressed. E-mail: sandler@che.Udel.edu. Phone: 302-831-2945. Fax: 302-831-3226.

10.1021/jp0525989 CCC: $30.25
Published on Web 09/23/2005

form to the orthorhombic (high-temperature) form near room temperature.[15,18] It is known that the transition to the orthorhombic form also occurs at lower temperatures when the solution is basic (pH >9).[23]

In this study, we have carried out a detailed simulation of lysozyme crystals to gain a better molecular-level understanding of the thermodynamics of protein crystallization. We use Monte Carlo simulations for an atomistically detailed model to calculate the van der Waals (VDW) interactions and the boundary element method to solve the Poisson−Boltzmann (PB) equation for electrostatic interactions. The tetragonal and orthorhombic forms of lysozyme crystals are modeled, accounting for the detailed protein shape using the experimental crystallographic coordinates. For simplicity, the protein is treated as a rigid body, thereby neglecting intramolecular conformational degrees of freedom. The effects of hydrogen bonding and solvation are considered only implicitly using a semiempirical force field.[24]

The electrostatic energy is calculated within the framework of continuum electrostatics by solving the linearized Poisson−Boltzmann equation using the boundary element method.[25] In doing so, we note that the conventional Poisson−Boltzmann equation for electrolyte solutions is inappropriate for protein crystals because the constraint of electroneutrality requires the charge of the protein to be considered. Also, the overall concentration of ionic species within the crystal should differ from that in the solution phase as a result of the Donnan equilibrium effect.[26−28] Recently, Deserno and von Grünberg[29] proposed an optimally linearized form of the Poisson−Boltzmann equation expanded about an average electric potential over the condensed phase incorporating the Donnan equilibria of small ions partitioned between the two phases. We adopt their method to obtain the electrostatic behavior of lysozyme crystals in equilibrium with a bulk solution phase of specified conditions such as pH and ionic strength. Further, we obtain the free energies of the crystals from the expanded ensemble Monte Carlo method.[30,31] A comparison with experiment is made for the energy and entropy change of crystallization to assess the validity and accuracy of the approximations used.

II. Simulation Methods

Hen egg white lysozyme consists of 129 amino acids and 1001 non-hydrogen atoms and has a molecular weight of 14 300. The crystallographic coordinates for lysozyme crystals are taken from the Protein Data Bank.[32] Several sets of pdb coordinates were considered for each crystal structure: 193L,[33] 1JIS,[34] 1LZA[35] for the tetragonal form and 1F0W,[36] 1HSX,[37] 1JJ1,[34] 1JPO,[38] and 1BGI[39] for the orthorhombic form.

A. Monte Carlo Simulation. The VDW energies of the crystals were calculated from NVT MC simulations for systems containing 16 protein molecules (two unit cells for the tetragonal and four unit cells for the orthorhombic crystals, respectively). Simulations were performed at the reported experimental density, with the configurations of the simulation boxes shown in Figure 1. For simplicity, the protein was treated as a rigid body, thereby neglecting intramolecular degrees of freedom.

Intermolecular interactions were explicitly considered at the atomistic level using the force field of Asthagiri et al.[24] This is a semiempirical force field in which a few parameters were tuned to reproduce the experimental binding free energies of several proteins by implicitly accounting for the water-mediated hydrogen-bonding and solvation effects. The VDW interactions were described using the Lennard-Jones potentials of the OPLS force field[40] with a scale factor of 0.5 when the atom−atom distance was less than 6 Å, and were represented by Hamaker

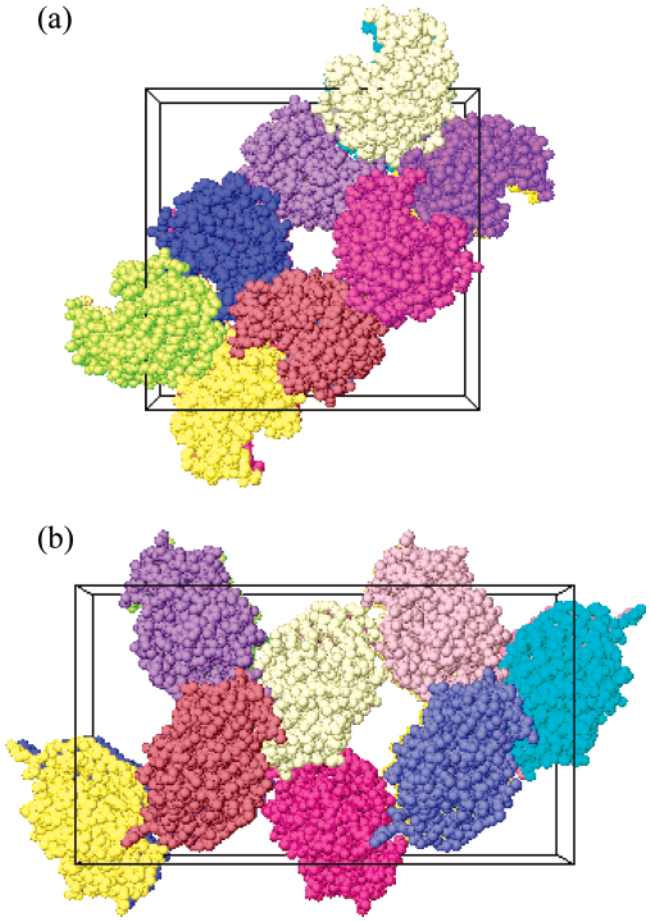

Figure 1. Lysozyme crystals in tetragonal form (PDB 193L) and in orthorhombic form (PDB 1F0W) used in MC simulations. A total of 16 protein molecules are used, comprising 2 unit cells for the tetragonal form (a) and 4 unit cells for the orthorhombic form (b).

interactions with a Hamaker constant of $3.1kT$ at larger distances.[24] When this force field was fitted to the binding free energies of proteins in solution, the entropic contribution resulting from the loss of the translational, rotational, and vibrational degrees of freedom of the proteins was assumed to be constant and treated as an adjustable parameter, estimated to be $13kT$. Removing this entropic contribution, the force field then mainly represents the energetic interactions between proteins incorporating solvation effects implicitly. The energetic contribution of solvent water is roughly reflected in the scale parameter of 0.5 by which the protein−protein VDW interactions are mediated, while the entropic contributions of water and ions of the electrostatic interactions are separately treated by the Poisson−Boltzmann equation.

In this study of lysozyme crystals we use the same force field parameters fitted to the thermodynamic properties of the solution phase. That is, we assume that the VDW interactions between proteins are unchanged despite the phase change. The prediction of the thermodynamic properties with this implicit model would be most reliable if the solvation structure in the crystal is unchanged from that in the solution phase. Conversely, if we find a significant deviation from experiment, the difference could be attributed to a change in solvation between the solution and crystal phase.

To speed up the simulations, the neighbor-list method was used based on the atom−atom distance. The Hamaker interactions were cut off at 12 Å, and the long-range correction, which contributes about 10% to the total VDW energy, was included by integrating the potentials assuming the pair distribution

function to be unity beyond the cutoff distance. The magnitudes of the maximum translation and rotation were adjusted in the equilibration stage so as to give an acceptance ratio of approximately 50%, resulting in values of ~0.1 Å for translation and ~1° for rotation, respectively. The ensemble averages were obtained during 15 000 MC cycles in the production stage.

The free energies of the crystals were obtained from expanded ensemble Monte Carlo simulations,[30,31] in which the energy of the system is varied with a coupling parameter λ (as in the thermodynamic integration method[41]) which ranges from $\lambda = 0$ for the reference Einstein crystal to $\lambda = 1$ for the real crystal,

$$U(\lambda) = \lambda \sum_{i<j} u_{ij} + (1-\lambda) \sum_i u_i^{\mathrm{Ein}} \quad (1)$$

The classical Einstein crystal was used as a reference system with the following external potential imposed on each molecule

$$u^{\mathrm{Ein}}(\mathbf{r}) = \lambda_t(\mathbf{r}-\mathbf{r}_0)^2 + \lambda_\theta(1-\cos\theta) + \lambda_\omega(1-\cos\omega) \quad (2)$$

where each λ_α is a force constant, $\mathbf{r}$ is the position vector of the center of mass of the protein, and $\mathbf{r}_0$ is its fixed position in the lattice obtained as the average position from the previous NVT MC simulation. The orientation of the protein is given by the three Euler angles $\phi\theta\psi$, which measure deviations from the reference orientation, an average orientation that has also been determined from the previous NVT MC simulation. The orientational Einstein field for nonlinear rigid molecules can be represented by two angles θ and ω, where $\omega = \phi + \psi$.[42]

The translational part of the free energy of the Einstein crystal is given by

$$\frac{\beta A_t^{\mathrm{Ein}}}{N} = -\frac{3}{2}\ln\left(\frac{\pi}{\beta\lambda_t}\right) + 3\ln\Lambda \quad (3)$$

where k is the Boltzmann constant, T is temperature, $\beta = 1/kT$, $\Lambda = h/\sqrt{2\pi MkT}$ with M being the mass of a protein molecule, and h is Planck's constant. For the lysozyme crystal, the force constant λ_t was set to 400 kJ/mol/Å^2, and the corresponding value of $\beta A_t^{\mathrm{Ein}}/N$ at room temperature is 5.91. The orientational contribution to the free energy of the Einstein crystal was obtained by integrating the orientational Einstein field, the last two terms of eq 2, over the Euler angles, resulting in the analytical expression[43]

$$\frac{\beta A_o^{\mathrm{Ein}}}{N} = \ln\frac{2\beta\lambda_\theta}{1-\exp(-2\beta\lambda_\theta)} + \beta\lambda_\omega - \ln[I_0(\beta\lambda_\omega)] \quad (4)$$

where I_0 is the Bessel function. Both λ_θ and λ_ω were set to 40 000 kJ/mol, and the corresponding value of $\beta A_o^{\mathrm{Ein}}/N$ at room temperature is 16.15.

The partition function for the expanded ensemble is a weighted sum of the individual partition functions of the subensembles

$$Z_E(N,V,T) = \sum_i \exp(w_i) Z(N,V,T;\lambda_i) \quad (5)$$

Transitions between adjacent subensembles ($\lambda_i \rightarrow \lambda_j$) were accepted using the Metropolis MC scheme with the predetermined weight factor w_i.[30] Equation 5 leads to the free energy difference being given by

$$\beta A_j - \beta A_i = w_j - w_i - \ln(P_j/P_i) \quad (6)$$

where P_i is the probability of finding a system with the parameter value λ_i during the expanded ensemble MC simulation. The details of implementing the expanded ensemble MC method for the solid phase and the procedure for optimizing the weight factors are given in our previous studies of fluid−solid transitions.[14,31]

B. Boundary Element Method. The electrostatic energy of the protein crystal was calculated from the linearized Poisson−Boltzmann equation, which was solved by the boundary element method.[25] This electrostatic energy is a free energy in view of the degrees of freedom of the solvent and the ions. The electrostatic free energy of crystallization was obtained as the difference between the electrostatic free energy of an isolated protein molecule immersed in the bulk solution and that in the crystal phase in Donnan equilibrium with the bulk solution.

For the interior domain of the protein, the electric potential ϕ is governed by the Poisson equation,

$$\nabla^2\phi = -\frac{\rho(\mathbf{r})}{\epsilon_i} \quad (7)$$

In the boundary integral form this is written as

$$\phi(\mathbf{r}) + \int \phi(\mathbf{r}') \frac{\partial G^i(\mathbf{r}'-\mathbf{r})}{\partial n(\mathbf{r}')}\, \mathrm{d}S(\mathbf{r}') - \int \frac{\partial \phi(\mathbf{r}')}{\partial n(\mathbf{r}')} G^i(\mathbf{r}'-\mathbf{r})\, \mathrm{d}S(\mathbf{r}') = \frac{1}{4\pi\epsilon_i}\sum_k \frac{q_k}{|\mathbf{r}-\mathbf{r}_k|} \quad (8)$$

where ϵ_i is the uniform dielectric constant inside the protein, q_k, is a charge placed at $\mathbf{r}_k$, $G^i(\mathbf{r}'-\mathbf{r}) = 1/4\pi|\mathbf{r}'-\mathbf{r}|$, and the derivative is with respect to the outward normal direction on the boundary surface.

For the exterior domain we used the linearized Poisson−Boltzmann equation,

$$\nabla^2\phi = \kappa^2\phi \quad (9)$$

where κ is the Debye screening parameter defined by $\kappa^2 = e^2\sum c_i^0 z_i^2/\epsilon_e kT$, e is the unit charge, c_i^0 is the bulk concentration of ion, z_i is the ion valency, and ϵ_e is the dielectric constant of the aqueous phase. The PB equation can be cast into the boundary integral form

$$\phi(\mathbf{r}) - \int \phi(\mathbf{r}') \frac{\partial G_{\mathrm{p}}^e(\mathbf{r}'-\mathbf{r})}{\partial n(\mathbf{r}')}\, \mathrm{d}S(\mathbf{r}') + \int \frac{\partial \phi(\mathbf{r}')}{\partial n(\mathbf{r}')} G_{\mathrm{p}}^e(\mathbf{r}'-\mathbf{r})\, \mathrm{d}S(\mathbf{r}') = 0 \quad (10)$$

where we used the periodic Green's function[44] for crystals subject to three-dimensional periodicity given by

$$G_{\mathrm{p}}^e(\mathbf{r}'-\mathbf{r}) = \sum_{l=-\infty}^{\infty}\sum_{m=-\infty}^{\infty}\sum_{n=-\infty}^{\infty} G^e(\mathbf{r}'-\mathbf{r}+\mathbf{r}_{lmn}) \quad (11)$$

where $G^e(\mathbf{r}'-\mathbf{r}) = \exp(-\kappa|\mathbf{r}'-\mathbf{r}|)/4\pi|\mathbf{r}'-\mathbf{r}|$, $\mathbf{r}_{lmn} = l\mathbf{a} + m\mathbf{b} + n\mathbf{c}$ with $\mathbf{a}$, $\mathbf{b}$, and $\mathbf{c}$ being the unit cell vectors, and l, m, n being integers. In practice, a finite sum for the right-hand side of eq 11 over 5^3 cells was used as the periodic Green's function converges rapidly in the cases studied here.

To meet the constraint of electroneutrality in the crystal, we used the optimally linearized form of the Poisson−Boltzmann equation proposed by Deserno and von Grünberg,[29] which takes into account the Donnan equilibria of ionic species between

the solution and crystal phases. The PB equation is given in a modified form as

$$\nabla^2\phi(\mathbf{r}) = \bar{\kappa}^2\left\{\phi(\mathbf{r}) - \left[\bar{\phi} - \frac{\tanh(ze\beta\bar{\phi})}{ze\beta}\right]\right\} \qquad (12)$$

where $\bar{\phi}$ is the average potential within the crystal and $\bar{\kappa}$ is the effective Debye screening parameter, both taking into the account the Donnan equilibrium. Compared with the electrostatic screening effect of the ions in the solution, the screening of the electrostatic interactions in the crystal phase is enhanced through the larger value of $\bar{\kappa}$ due to the higher concentration of counterions, i.e., $\bar{\kappa}^2 = \kappa^2 \cosh(ze\beta\bar{\phi}) \geq \kappa^2$. The optimum value of the Donnan potential is self-consistently chosen to be

$$\exp(ze\beta\bar{\phi}) = \frac{-Z_p c_p + \sqrt{(Z_p c_p)^2 + 4(zc_0)^2}}{2zc_0} \qquad (13)$$

where Z_p is the charge of the protein, and c_p and c_0 are the number densities of the protein in the crystal and of the ionic species in the bulk solution, respectively. Equation 12 can be written in the same form as eq 9 by defining a shift potential as

$$\varphi(\mathbf{r}) = \phi(\mathbf{r}) - \left[\bar{\phi} - \frac{\tanh(ze\beta\bar{\phi})}{ze\beta}\right] \qquad (14)$$

so that the same numerical algorithm as used to solve the conventional PB equation can be used.

A unit cell of a protein crystal for the boundary element calculations was constructed with use of the PDB atomic coordinates. A simply connected domain for a protein consisting of triangular elements was obtained with Connolly's program[45] to calculate the solvent accessible surface using a solvent probe radius of 1.4 Å. A further simplification was made by removing triangles with very small edges, while keeping the overall shape unchanged. The removal was done as follows: (1) the edge of smallest size was found, (2) the two triangular elements sharing the edge were identified, (3) the two vertexes of the edge were merged, and (4) the two elements were removed. Care was taken so that each edge was shared by exactly two elements, and then a check was made that the final domain obtained was a simply connected surface without internal cavities. The initial solvent accessible surface consisted of more than 10 000 triangles, and the simplified domains consist of 1 000 to 2 000 elements, which were then used to construct a unit cell. In Figure 2 boundary element domains for the unit cell of the tetragonal structure are shown.

Inside each domain the charges were placed at the positions of the 30 ionizable residues and the C and N termini. At a given pH, the magnitudes of the charges were determined from the Henderson−Hasselbalch equation with use of experimental pK_a data.[46] While a more rigorous approach could be used accounting for the activity coefficient of the hydrogen ion, the effect of the ionic strength on the net charge of lysozyme is rather small compared with the primary effect caused by the change of pH.[47] Also, since the electrostatic interactions are largely screened at most crystallization conditions we consider, the Henderson−Hasselbalch equation suffices to estimate the charges of ionizable groups with reasonable accuracy. For Asp, Glu, and Arg, the average positions of the two resonant atoms were taken as the charge centers. As lysozyme contains many arginine residues, it is positively charged at acidic and neutral conditions, and its isoelectric point is 11.2.[47]

In some cases, calculations were done for two temperatures, in which case the dielectric constants for the aqueous phase

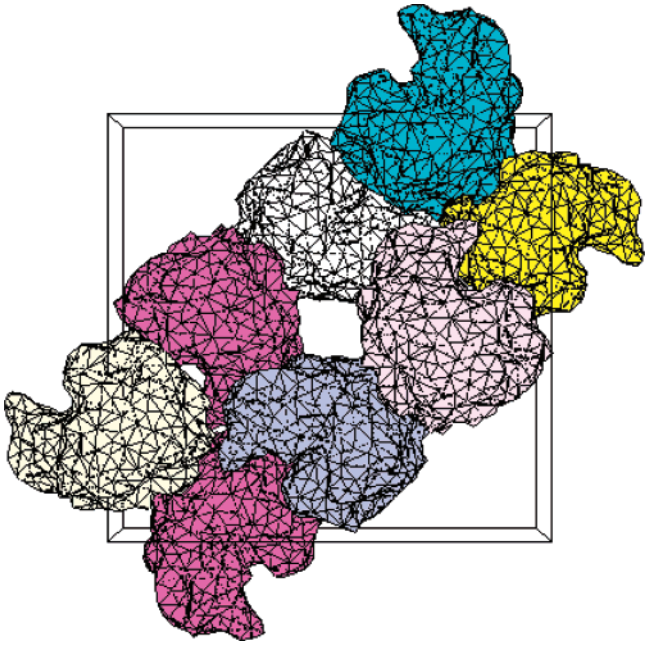

Figure 2. Tessellated domains for the unit cell of tetragonal lysozyme crystal used in the boundary element calculation.

were set to 80 at 298 K and 88 at 273 K, respectively, and 4 inside the protein domain. The solution of the boundary element equation was obtained by using an iterative scheme that is faster and requires fewer computer resources than directly solving the full matrix equation.[48] The validity of the program was tested for a simple two-sphere system, and the numerical solutions were found to be in good agreement with the analytical results.[49]

III. Results and Discussion

The VDW energies (or enthalpies if we ignore the small PV term) of the lysozyme crystals were obtained from the NVT Monte Carlo simulations. The density or the unit cell dimensions was fixed at the experimental values given by each set of PDB structure data. There was virtually no difference in the positions and orientations of a protein molecule between the input PBD coordinates and the final equilibrated ones because of the tightly bound structure. Compared with its large size, a protein in the crystal phase samples only a very narrow range of the configurational space during the simulation, as indicated by the fact that the mean translational displacement is about 0.1 Å and the rotational displacement is about 1° for the acceptance ratio of 50%.

Simulation results for the configurational energies are shown in Table 1 and Figure 3. Most simulations were performed at 298 K, at which the two crystal structures are known to coexist at pH 4.6. There is no appreciable effect of temperature on the VDW energy. A few selected runs at the lower temperature of 273 K show only a slight increase in attraction of about 0.5 kJ/mol due to the reduced thermal fluctuations. The VDW interactions in the tetragonal form are less attractive than those in the orthorhombic form, although the former is of slightly higher density. Noticeable variations in the energy among the PDB structures are observed, with a range of about 5 kJ/mol for the tetragonal form and about 20 kJ/mol for the orthorhombic form. Since the simulation uncertainty of the energy for these crystals is about 0.1 kJ/mol and the protein molecules are treated as rigid bodies, these variations can be ascribed to differences in the input PDB coordinates, particularly for the side chain residues. A recently reported high-temperature orthorhombic

TABLE 1: The van der Waals Energy of Lysozyme Crystals Obtained from NVT Monte Carlo Simulations[a]

	tetragonal			orthorhombic				
	193L	1JIS	1LZA	1F0W	1HSX	1JJ1	1JPO	1BGI
density (mol/L)	0.0570	0.0558	0.0559	0.0534	0.0533	0.0540	0.0533	0.0525
U(total) at 298 K	−112.5	−117.4	−116.1	−134.4	−138.5	−128.3	−139.2	−146.5
U(OPLS)	−88.8	−94.0	−92.8	−109.4	−115.0	−102.6	−115.0	−122.8
U(total) at 273 K	−113.0	−118.0	−116.7	−134.9			−139.7	

[a] The total energy is compared with the short-range contribution described by the scaled OPLS force field for $r < 6$ Å. Values are in units of kJ/mol

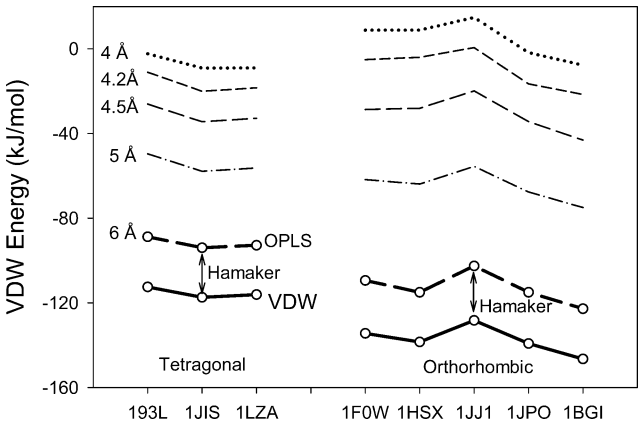

Figure 3. The van der Waals energies of the lysozyme crystals. The thick solid line is for the van der Waals energy, and the thick dashed line is for the short-range contribution ($r < 6$ Å) described by the scaled OPLS force field. The contributions of the atom pairs at shorter ranges are also shown for comparison.

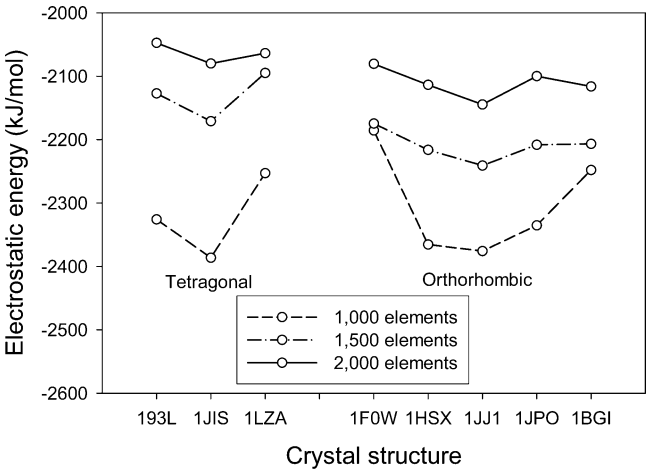

Figure 4. The electrostatic energy of a single lysozyme molecule in solution at pH 4.5 and $I = 0.36$ M calculated from the boundary element method. Lines are results with different numbers of boundary elements.

structure (1BGI) has different unit cell dimensions and a different packing structure from the other orthorhombic forms,[39] and has the largest attractive energy. The calculated VDW energies for the tetragonal form are about −115 kJ/mol, compared to the reported experimental energies of crystallization, which vary over the range −40 to −140 kJ/mol, and are very sensitive to solution conditions such as pH and ionic strength.[15−21] For the orthorhombic form, there is a more pronounced difference: the simulated values are about −135 kJ/mol compared to the experimental values of −18 to −32 kJ/mol.[15,16]

In Figure 3, the total VDW energy and its short-range part represented by the scaled OPLS force field for distances less than 6 Å are shown. The dispersion interactions for individual atom pairs are switched from the scaled OPLS energy to the water-mediated Hamaker interactions at 6 Å, at which a water molecule can intervene between the two atoms to allow solvation.[24] Since there is no rigorous criterion to choose this cutoff distance for the solvation to occur, it should be considered an adjustable parameter, together with the scale factor of 0.5 in the OPLS force field. For the purpose of comparison, the detailed contributions of the short-range dispersion interactions are also shown in Figure 3 for different distances. The variation in the total VDW energy is seen to be almost the same as that of the shorter range interactions near contacts. The difference in the LJ energy between PBD structures is unchanged with the solvation distance, and can be traced back to the contact distance of 4 Å. On the other hand, the relative proportions of the short-range interactions to the total energy vary with the cutoff distance. Above 4.5 Å, the interactions in the orthorhombic form become more attractive, whereas there is an opposite trend below this distance. The more rapid increase in the energy of the orthorhombic form with decreasing distance indicates fewer attractive pairs at contact than in the tetragonal form.[50] Overall, the Lennard-Jones interactions within 6 Å given by the OPLS force field dominate over the water-mediated Hamaker interactions at longer distances. The magnitude of the Hamaker interaction is almost the same for all PDB structures that we studied.

The PB equation for the protein crystals was solved for the simplified solvent accessible surfaces of the proteins by using the boundary element method with simplified domains of from 1000 to 2000 elements. We found that even a representation with 2000 elements requires too much memory to solve the equations for the unit cell of the tetragonal form containing eight protein molecules. Thus, this representation was used only for the calculation of the solution phase to determine the effect of the number of elements. Figure 4 shows the absolute electrostatic free energy of a lysozyme molecule immersed in a solution at 298 K, pH 4.5, and ionic strength 0.36 M. There is a significant variation in the solvation free energy depending on the number of boundary elements. Also, with the same number of elements, different energies result from the different PDB structures. There are several reasons for this sensitivity. Although the simplification process of the boundary elements maintains the overall shape of the protein, it changes the local landscape of the protein surface. The details of the boundary shape can strongly affect the electrostatics, since most of the protein charges lie immediately beneath the boundary surface. Removal of nearby edges can change the distance between the charge and the dielectric boundary. Moreover, removing smaller edges in the domain simplification process makes the protein domain slightly smaller. Although the extent of the shrinkage is only about 0.1 to 0.2 Å, this leads to more stable hydration due to the penetration of the water phase, resulting in a decrease of the electrostatic energy with a smaller number of elements. The calculation with additional boundary elements results in less variation in the electrostatic energy among the different PBD

TABLE 2: The Electrostatic Energies of Lysozyme Crystals at 298 K Calculated from the Boundary Element Method[a]

(a) With 1000 Boundary Elements

		tetragonal			orthorhombic				
pH (net charge)	I (M)	193L	1JIS	1LZA	1F0W	1HSX	1JJ1	1JPO	1BGI
4.5 (11.0)	0.36	17.0	30.6	25.4	16.9	14.9	21.4	42.3	20.3
	0.55	12.1	23.9	20.1	12.2	10.1	15.9	35.5	15.7
	0.88	6.9	16.9	14.3	7.5	5.4	10.4	27.8	11.1
5.2 (10.0)	0.36	15. 6	29.3	24.5	13.1	11.3	17.0	39.1	18.9
	0.55	10.8	22.7	19.3	9.1	7.2	12.3	32.8	14.5
	0.88	5.9	16.0	13.8	5.3	3.4	7.9	26.0	10.1
7.8 (7.6)	0.36	12.7	26.3	22.0	7.9	6.3	11.1	35.7	15.7
	0.55	8.2	20.0	17.0	4.9	3.2	7.6	30.1	11.9
	0.88	4.1	14.2	12.2	2.5	0.7	4.8	24.2	8.4
10.6 (2.3)	0.36	4.0	9.9	10.3	−2.2	−1.8	0.3	26.8	4.4
	0.55	2.8	7.8	8.5	−2.4	−2.4	−0.3	23.5	4.1
	0.88	1.7	6.2	6.8	−2.4	−2.7	−0.5	19.8	3.6

(b) With 1500 Boundary Elements

		tetragonal			orthorhombic				
pH (net charge)	I (M)	193L	1JIS	1LZA	1F0W	1HSX	1JJ1	1JPO	1BGI
4.5 (11.0)	0.36	19.9	19.6	25.2	15.9	17.4	20.1	43.0	18.5
	0.36[b]	(18.5)	(18.2)	(23.3)	(14.8)	(16.1)	(18.7)	(39.7)	(17.2)
	0.55	14.8	15.0	20.4	11.6	12.7	15.4	36.9	14.4
	0.88	9.4	10.2	15.1	7.4	8.1	10.7	29.9	10.3
5.2 (10.0)	0.36	18.2	18.5	24.4	12.3	13.9	16.3	40.0	17.2
	0.55	13.3	14.1	19.7	8.7	9.9	12.4	34.4	13.3
	0.88	8.2	9.5	14.7	5.2	6.1	8.7	28.1	9.5
7.8 (7.6)	0.36	15.3	16.1	22.1	7.3	9.0	11.2	36.9	14.2
	0.55	10.7	11.9	17.7	4.6	5.9	8.4	31.9	10.9
	0.88	6.39	8.09	13.4	2.5	3.3	6.0	26.5	7.9
10.6 (2.3)	0.36	1.20	3.99	12.5	−2.4	0.7	2.4	28.5	3.8
	0.55	0.07	2.91	11.0	−2.5	0.1	1.9	25.6	3.6
	0.88	−0.78	2.04	9.5	−2.3	−0.3	1.7	22.2	3.4

[a] Values are in units of kJ/mol. [b] Numbers in parentheses are results at 273 K.

structures, indicating a reduced effect of the artificial discretization. Despite this sensitivity for the absolute value of the electrostatic energy, the boundary element calculation is still useful for the calculation of the crystallization energy, which is the difference in the absolute energies between the solution and the crystal phases.

In Table 2, the electrostatic parts of the free energy of crystallization of the different lysozyme crystal forms are shown at 298 and 273 K with varying pH and ionic strength. The free energy of the solution phase was obtained by solving the usual linearized PB equation for a single protein molecule, while that of the crystal phase was obtained by solving the optimally linearized PB equation, taking into account the Donnan equilibria of the ionic species. The free energy of crystallization is the difference between the two energies. Two sets of results were obtained for 1000 and 1500 elements on each molecule, respectively. There is only a small difference in energy of less than 3 kJ/mol between the results for the two discretizations, except for the case of the 1JIS structure, for which the free energy reduction is as much as 10 kJ/mol. However, no systematic trend is found in the change of the free energy with the refinement of the boundary elements, though the variation among the different PDB structures is somewhat reduced. Due to the increased effective Debye screening parameters as a result of the Donnan equilibrium, the free energy of crystallization is about half the value of that without the electroneutrality constraint.

Overall, the tetragonal and orthorhombic structures give similar magnitudes of electrostatic free energies, except for the orthorhombic 1JPO structure, which shows the largest repulsions. Lowering the temperature to 273 K (and changing the dielectric constant to 88) leads to the decrease of the crystallization energy by about 1 or 2 kJ/mol due to enhanced charge screening of the ions caused by the reduced thermal fluctuations and increased dielectric constant.

The effects of pH and ionic strength on the electrostatic crystallization energy can be explained in terms of the screened charge−charge interactions, the magnitude of which gradually decreases with increasing pH and ionic strength. At higher pH the charge−charge repulsion is reduced, also an increase in the ionic strength leads to a reduced repulsion as a result of enhanced charge screening. The experimental data of Schall et al.[19] show this sensitivity to the solution conditions. The change from pH 4.6, $I = 0.88$ M to pH 5.2, $I = 0.55$ M decreased the magnitude of the crystallization energy by about 30 kJ/mol, compared with a change of less than 4 kJ/mol in our predictions.

In Table 3, the simulation results of the thermodynamic properties of lysozyme crystals at 298 K, pH 4.5, and $I = 0.36$ M are shown. The VDW contribution to the Helmholtz energy of the lysozyme crystals was calculated from the expanded ensemble simulations, and the total Helmholtz energy was then approximated as the sum of the electrostatic and VDW free energies. Adding the electrostatic free energy obtained for the fixed experimental configuration is reasonable considering that the tightly bound crystal structure results from the MC simulations with only the VDW interactions. The VDW contribution for the entropy of the crystal is determined from the difference between the VDW energy obtained from the NVT MC simulation and the corresponding Helmholtz free energy obtained from the expanded ensemble MC simulation. The electrostatic free energy is separated into energetic and entropic parts by using the Gibbs−Helmholtz equation $\partial(A/T)/\partial T = -U/T^2$ to calculate the electrostatic energy, using the finite difference approximation with respect to the temperature and the results at 298 and 273

TABLE 3: The Thermodynamic Properties of Lysozyme Crystals at 298 K, pH 4.5, $I = 0.36$ M from Simulations[a]

	tetragonal			orthorhombic				
	193L	1JIS	1LZA	1F0W	1HSX	1JJ1	1JPO	1BGI
$A - A°$(VDW)	−43.1	−52.8	−50.1	−64.2	−67.9	−56.6	−69.8	−80.5
A (elec)	19.9	19.6	25.2	15.9	17.0	20.0	43.0	18.5
$A - A°$(total)	−23.2	−33.2	−24.9	−48.3	−50.5	−36.5	−26.8	−62.0
U(VDW)	−112.5	−117.4	−116.1	−134.4	−138.5	−128.3	−139.2	−146.5
U(elec)	2.8	2.8	3.0	3.0	2.1	3.7	4.0	3.3
U(total)	−109.7	−114.6	−113.1	−131.4	−136.4	−124.6	−135.2	−143.2
$S - S°$(VDW)	−233	−217	−221	−235	−237	−240	−233	−222
S(elec)	−57	−56	−74	−43	−51	−55	−131	−51
$S - S°$(total)	−290	−273	−295	−278	−288	−295	−364	−273

[a] The electrostatic free energy is separated into the energy and entropy parts by using the Gibbs−Helmholtz equation. The standard state refers to a hypothetical ideal state at a concentration of 1 mol/L. Energies are in units of kJ/mol, and entropies are in units of J/mol/K.

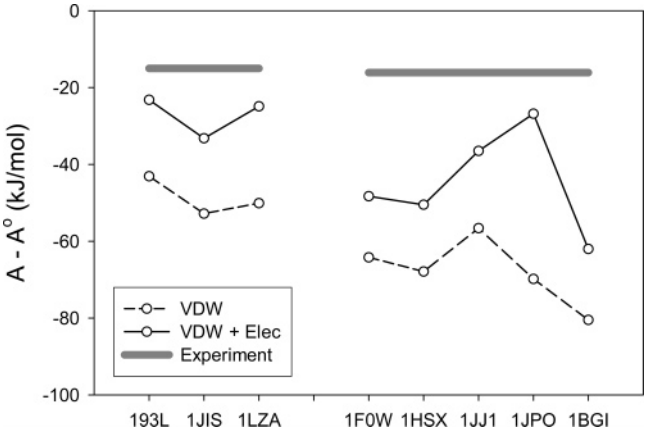

Figure 5. The Helmholtz energies of the lysozyme crystals at pH 4.5 and $I = 0.36$ M. The dashed lines are for the VDW energies, and the solid lines for the total energies including the electrostatic contribution. The thick lines are from experimental data (refs 16 and 18).

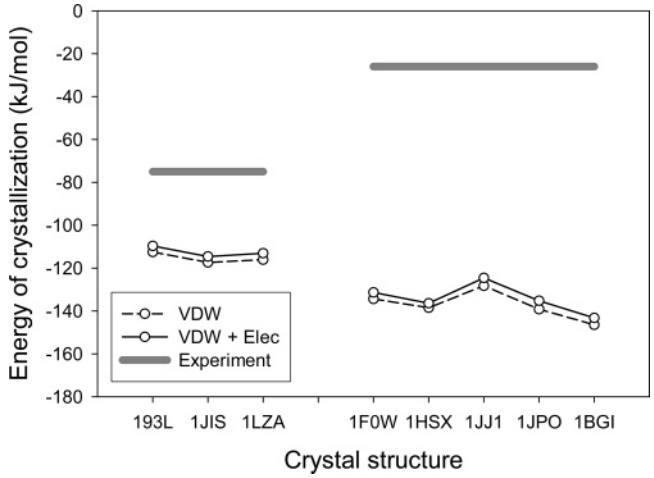

Figure 6. The crystallization energies of lysozyme crystals at pH 4.5 and $I = 0.36$ M. The lines have the same meanings as in Figure 5.

K. The electrostatic entropy is then obtained from the difference between the electrostatic energy and the electrostatic free energy, though an alternative route is possible without requiring temperature derivative.[51]

To make comparisons with the literature data, the Helmholtz energy is reported with respect to a hypothetical standard state at a concentration of 1 mol/L (=1/1660 Å^{-3}). The translational part of the Helmholtz energy at the standard state is $A°/NkT = \ln(\Lambda^3/1660) - 1$. The calculated Helmholtz energies at pH 4.5 and $I = 0.36$ M are shown in Figure 5. The values for the tetragonal form are in the range of −23 to −33 kJ/mol, and are a little higher than those for the orthorhombic form, which vary more widely from −27 to −62 kJ/mol. Solubility data as a function of temperature were used to obtain the Gibbs energy of crystallization using the van't Hoff equation. The experimental value is −15 kJ/mol for the tetragonal form, which is obtained from the solubility data of Cacioppo and Pusey[16] [pH 4.5, $I = 0.38$ M (2% NaCl, 0.1 M NaAc)], and −16 kJ/mol for the orthorhombic form using data of Ewing et al.[18] [pH 4.6, $I = 0.38$ M (2% NaCl, 0.1 M NaAc)]. The two experimental values are close to each other since the crystal structure transition occurs near 298 K. The predicted Helmholtz energies are considerably lower than the experimental values, and a further analysis for individual contributions of the energy and entropy effects is necessary to trace the origin of this difference.

In Figure 6 the crystallization energies are compared with experimental data at the same solution conditions. The VDW contributions play a dominant role in both crystals, and the electrostatic contributions are small and positive. For the tetragonal crystal, the experimental crystallization energy is estimated to be −75 kJ/mol.[16] Our predicted crystallization energy is about 38 kJ/mol larger in magnitude than the experimental value. The difference mainly comes from the VDW part with little contribution from the electrostatics, and is acceptable considering the wide range of reported experimental values and the fact that we have made no parameter adjustments to fit the lysozyme crystal data. On the other hand, this difference reflects the change of solvation environment upon the crystallization. The less attractive energy change suggests energetically unfavorable release of water molecules from crystal contacts. For the orthorhombic crystal, the experimental crystallization energy is −26 kJ/mol,[18] whereas the predicted value is about −134 kJ/mol, even more attractive than for the simulated tetragonal crystal. The larger disagreement for the orthorhombic crystal form suggests that there should be a larger difference in the solvation.

The entropies of the two lysozyme crystals at the same solution conditions are shown in Figure 7. The simulated values for the VDW contributions to the entropy, arising from the translational and rotational confinement of the proteins, are of comparable magnitude for the tetragonal and orthorhombic forms, and vary in the range of −240 to −215 J/mol/K. The predictions for the tetragonal form are in good agreement with the theoretical estimate of −215 J/mol/K obtained from a mean field argument.[52] The experimental values of −201 J/mol/K at the same conditions[16] and −193 J/mol/K obtained at a slightly higher ionic concentration [pH 4.5 and $I = 0.43$ M][53,54] are close to the mean-field value and the VDW entropies from simulations, but the apparently good agreement may be fortuitous since experimental values include electrostatic and solvation effects.

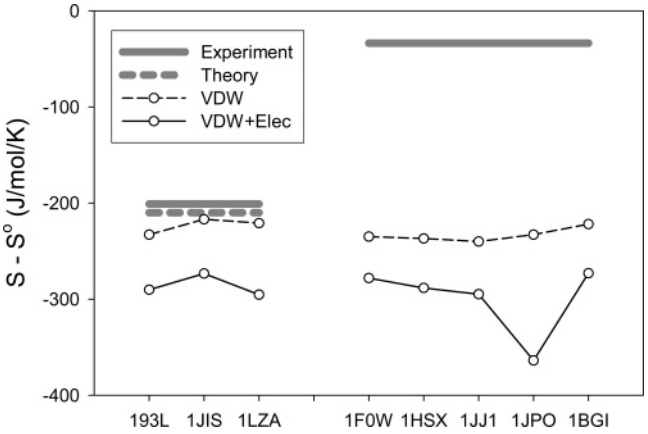

Figure 7. The entropies of the lysozyme crystals at pH 4.5 and $I = 0.36$ M. The lines have the same meanings as in Figure 5. The thick dashed line is the theoretical estimate obtained from the mean field theory (ref 52).

TABLE 4: The Hydrated Water Structures of Lysozyme Crystals

	tetragonal		orthorhombic	
	193L	1JIS	1F0W	1HSX
no. of hydrated water molecules per protein	142	142	137	153
no. of bridging water molecules between proteins				
$r < 3.6$ Å	28	23	12	15
$r < 4.0$ Å	46	41	28	29

The predicted electrostatic contributions to the entropy are negative, and the deviation of the total entropy from experiment is about −85 J/mol/K, arising from the reorganization of water molecules and ions. Assuming this discrepancy is solely due to the change of the solvation structure, which was not considered in our analysis, this implies that there should be a release of about 4 water molecules upon crystallization if the entropy change of the water release is approximated by the entropy change on the melting of ice (∼22 J/mol/K). This is in contrast to the view of Vekilov et al.,[53] who neglected the VDW contribution of proteins and ascribed all the entropy to the change of the solvent effect, and concluded that 8 to 9 water molecules per each lysozyme are trapped in the crystal. Since our simulation results reasonably predict that the entropy change of proteins is significant for the tetragonal lysozyme crystal, it is likely that the solvation structure in the crystal and the solution phase will be more or less the same.

However, a large discrepancy between experiment and simulation results is found for the orthorhombic crystal. The experimental value at pH 4.5 and $I = 0.36$ M is −33.6 J/mol/K,[18] while simulation predicts an entropy of about −300 J/mol/K including a rather minor electrostatic contribution. While the predicted VDW and electrostatic contributions for the orthorhombic form are comparable to those for the tetragonal form, the large deviation from the experimental value suggests a change in the solvation structure different from the tetragonal crystal form and the solution phase. To clarify this, we examined the 140 to 150 hydrated water molecules in the experimental PDB structures as shown in Table 4. In this analysis, a water molecule is counted as a bridging molecule if it is also close to another protein. Two distances of 3.6 and 4.0 Å were used as cutoff distances in ascertaining whether a water molecule is considered to be a bridging molecule. In either case, there is a decrease of about 10 or more bridging water molecules for each protein in the orthorhombic crystal forms, whereas the total number of hydrated water molecules is almost the same. This suggests that additional water molecules are expelled from between the contacting surfaces when a lysozyme molecule becomes part of an orthorhombic crystal, which is an energetically less favorable but entropically more favorable process. As the rigid protein model was used in this work, there might be some underestimate of the VDW entropy due to using structures that are too tightly bound, and that the vibrational entropy change,[55] which leads to a positive contribution, has been neglected.

IV. Conclusions

We have carried out Monte Carlo simulations of the hen egg white lysozyme crystals at the atomistic level to account for the details of the anisotropic shape of the protein and the boundary element calculations to solve the Poisson−Boltzmann equation for the electrostatic interactions. The MC simulation effectively describes the docking patterns between the intermolecular contact regions of the tightly bound crystal structures. Owing to limitation of the implicit solvent model used, indirect reasoning was used to clarify the difference in the solvation structures between the two lysozyme crystals we examined. We first assumed the same extent of solvation effect for the two crystals as in the solution phase, and predicted the changes of energy and entropy upon crystallization. The crystallization energy of the tetragonal structure agrees reasonably well with experimental data, while there is a considerable disagreement for the orthorhombic structure. This suggests that there should be a significant difference in the solvation structure between the two crystal forms. A large difference in the experimental energy of crystallization between the two crystals of close densities indicates that there is energetically unfavorable solvation in the orthorhombic form. This is also supported by the entropy calculation. The entropy of the tetragonal structure is predicted reasonably well, and is mostly determined by its packing structure. The much higher value of the entropy of the orthorhombic crystal can only be explained in terms of the entropy gain of the water molecules released during the crystallization.

Acknowledgment. Financial support of this research was provided from contract DE-FG02-85ER13436 from the division of Basic Energy Sciences of the U.S. Department of Energy and grants CTS-0083709 and BES-0078844 from the U.S. National Science Foundation.

References and Notes

(1) George, A.; Wilson, W. W. *Acta Crystallogr. D* **1994**, *50*, 361−365.
(2) Rosenbaum, D.; Zamora, P. C.; Zukoski, C. F. *Phys. Rev. Lett.* **1996**, *76*, 150−153.
(3) Asherie, N.; Lomakin, A.; Benedek, G. B. *Phys. Rev. Lett.* **1996**, *77*, 4832−4835.
(4) Hagen, M. H. J.; Frenkel, D. *J. Chem. Phys.* **1994**, *101*, 4093−4097.
(5) ten Wolde, P. R.; Frenkel, D. *Science* **1997**, *277*, 1975−1978.
(6) Pellicane G.; Costa, D.; Caccamo, C. *J. Phys. Chem. B* **2004**, *108*, 7538−7541.
(7) Haas, C.; Drenth, J.; Wilson, W. W. *J. Phys. Chem. B* **1999**, *103*, 2808−2811.
(8) Sear, R. *J. Chem. Phys.* **1999**, *111*, 4800−4806.
(9) Curtis, R. A.; Blanch, H. W.; Prausnitz, J. M. *J. Phys. Chem. B* **2001**, *105*, 2445−2452.
(10) Neal, B. L.; Asthagiri, D.; Lenhoff, A. M. *Biophys. J.* **1998**, *75*, 2469−2477.
(11) Elcock, A. H.; McCammon J. A. *Biophys. J.* **2001**, *80*, 613−625.
(12) Lund, M.; Jönsson, B. *Biophys. J.* **2003**, *85*, 2940−2947.
(13) Carlsson, F.; Malmsten, M.; Linse, P. *J. Phys. Chem. B* **2001**, *105*, 12189−12195.

(14) Chang, J.; Lenhoff, A. M.; Sandler, S. I. *J. Chem. Phys.* **2004**, *120*, 3003−3014.
(15) Ataka, M.; Asai, M. *J. Cryst. Growth* **1988**, *90*, 86−93.
(16) Cacioppo E.; Pusey, M. L. *J. Cryst. Growth* **1991**, *114*, 286−292.
(17) Rosenberger, F.; Howard, S. B.; Sowers J. W.; Nyce, T. A. *J. Cryst. Growth* **1993**, *129*, 1−12.
(18) Ewing, F.; Forsythe, E.; Pusey, M. *Acta Crystallogr. D* **1994**, *50*, 424−428.
(19) Schall, C. A.; Arnold, E.; Wiencek, J. M. *J. Cryst. Growth* **1996**, *165*, 293−298.
(20) Broide, M. L.; Tominc, T. M.; Saxowsky, M. D. *Phys. Rev. E* **1996**, *53*, 6325−6335.
(21) Muschol, M.; Rosenberger, F. *J. Chem. Phys.* **1995**, *103*, 10424−10432.
(22) Muschol, M.; Rosenberger, F. *J. Chem. Phys.* **1997**, *107*, 1953−1962.
(23) Velev, O. D.; Kaler, E. W.; Lenhoff, A. M. *Biophys. J.* **1998**, *75*, 2682−2697.
(24) Asthagiri, D.; Neal, B. L.; Lenhoff, A. M. *Biophys. Chem.* **1999**, *78*, 219−231.
(25) Yoon, B. J.; Lenhoff, A. M. *J. Comput. Chem.* **1990**, *11*, 1080−1086.
(26) Donnan, F. G. *Chem. Rev.* **1924**, *1*, 73−90.
(27) Overbeek, J. Th. G. *Prog. Biophys. Biophys. Chem.* **1956**, *6*, 57−84.
(28) Tamashiro, M. N.; Levin, Y.; Barbosa, M. C. *Eur. Phys. J. B* **1998**, *1*, 337−343.
(29) Deserno, M.; von Grünberg, H.-H. *Phys. Rev. E* **2002**, *66*, 011401-1−011401-15.
(30) Lyubartsev, A. P.; Martsinovski, A. A.; Shevkunov, S. V.; Vorontsov-Vel'yaminov, P. N. *J. Chem. Phys.* **1992**, *96*, 1776−1783.
(31) Chang, J.; Sandler, S. I. *J. Chem. Phys.* **2003**, *118*, 8390−8395.
(32) Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. *Nucleic Acids Res.* **2000**, *28*, 235−242.
(33) Vaney, M. C.; Maignam, S.; Ries-Kautt, M.; Ducruix A. *Acta Crystallogr. D* **1996**, *52*, 505−517.
(34) Datta, S.; Biswal, B. K.; Vijayan M. *Acta Crystallogr. D* **2001**, *57*, 1614−1620.
(35) Maenaka, K.; Matsushima, M.; Song, H.; Sunada1, F.; Watanabe, K.; Kumagai, I. *J. Mol. Biol.* **1995**, *247*, 281−293.
(36) Biswal, B. K.; Sukumar, N.; Vijayan M. *Acta Crystallogr. D* **1999**, *55*, 934−937.
(37) Sukumar, N.; Biswal, B. K.; Vijayan M. *Acta Crystallogr. D* **2000**, *56*, 1110−1119.
(38) Bradbrook, S.; Deacon, A.; Habash, J.; Helliwell, J. R.; Helliwell, M.; Nieh, Y. P.; Snell, E. H.; Trapani, G.; Thompson, A. W.; Campbell, J. W.; Allinson, N. M.; Moon, K.; Ursby, T.; Wulff, M. *Proc.* SPIE-*Int. Soc. Opt. Eng.* **1995**, *2421*, 160.
(39) Oki, H.; Natuura, Y.; Komatsu, H.; Chernov, A. A. *Acta Crystallogr. D* **1999**, *55*, 114−121.
(40) Jorgensen, W. L.; Tirado-Rives, J. *J. Am. Chem. Soc.* **1988**, *110*, 1657−1666.
(41) Frenkel, D.; Ladd, A. J. C. *J. Chem. Phys.* **1984**, *81*, 3188−3193.
(42) Báez, L. A.; Clancy, P. *Mol. Phys.* **1995**, *86*, 385−396.
(43) Vlot, M. J.; Huinink, J.; van der Eerden, J. P. *J. Chem. Phys.* **1999**, *110*, 55−61.
(44) Johnson, C. A.; Wu, P.; Lenhoff, A. M. *Langmuir* **1994**, *10*, 3705−3713.
(45) Connolly, M. L. *J. Mol. Graph* **1993**, *11*, 139−141.
(46) Kuramitsu, S.; Hamaguchi, K. *J. Bioichem.* **1980**, *87*, 1215−1219.
(47) Kuehner, D. E.; Engmann, J.; Fergg, F.; Wernick, M.; Blanch, H. W.; Prausnitz, J. M. *J. Phys. Chem. B* **1999**, *103*, 1368−1374.
(48) Zhou, H.-X *Biophys. J.* **1993**, *65*, 955−963.
(49) McClurg, R. B.; Zukoski, C. F. *J. Colloid Interface Sci.* **1998**, *208*, 529−542.
(50) Matsuura, Y.; Chernov, A. A. *Acta Crystallogr. D* **2003**, *59*, 1347−1356.
(51) Roth, C. M.; Sader, J. E.; Lenhoff, A. M. *J. Colloid Interface Sci.* **1998**, *203*, 218−221.
(52) Finkelstein, A.; Janin, J. *Protein Eng.* **1989**, *3*, 1−10.
(53) Vekilov, P. G.; Feeling-Taylor, A. R.; Yau, S.-T.; Petsev, D. *Acta Crystallogr. D* **2002**, *58*, 1611−1616.
(54) Petsev, D. N.; Wu, X.; Galkin, O.; Vekilov, P. G. *J. Phys. Chem. B* **2003**, *107*, 3921−3926.
(55) Tidor, B.; Karplus, M. *J. Mol. Biol.* **1994**, *238*, 405−414.