The Stabilities of Protein Crystals

Jeremy D. Schmit and Ken A. Dill*

Department of Pharmaceutical Chemistry, University of California at San Francisco, San Francisco, California 94158

Received: November 11, 2009; Revised Manuscript Received: February 2, 2010

We describe a model for protein crystallization equilibria. The model includes four terms, (1) protein translational entropy opposes crystallization, (2) proteins are attracted to each other by a nonelectrostatic contact free energy favoring crystallization, (3) proteins in the crystal repel each other but, to a greater extent, attract counterions sequestered in the crystal, which favors crystallization, and (4) the translational entropy of the counterions opposes their sequestration into the crystal, opposing crystallization. We treat the electrostatics using the nonlinear Poisson−Boltzmann equation, and we use unit cell information from native protein crystals to determine the boundary conditions. This model predicts the stabilities of protein crystals as functions of temperature, pH, and salt concentrations, in good agreement with the data of Pusey et al. on tetragonal and orthorhombic crystal forms of lysozyme. The experiments show a weak dependence of crystal solubility on pH. According to the model, this is because the entropic cost to neutralize the crystal is compensated by favorable protein−salt interactions. Experiments also show that adding salt stabilizes the crystal. Cohn's empirical law predicts that the logarithm of solubility should be a linear function of salt. The present theory predicts nonlinearity, in better agreement with the experiments. The model shows that the salting out phenomena is not due to more counterion shielding but to lowered counterion translational entropy. Models of this type may help guide faster and better ways to crystallize proteins.

1. Introduction

We are interested in the forces that stabilize protein crystals. One motivation is that crystallization is often the rate-limiting step in determining a protein's structure. While technology for protein crystallization has seen big advances,[1,2] including the recent application of microfluidic chips,[3,4] the underlying physical principles are not yet very well understood, and crystallization still remains largely a business of trial and error. Crystallization has two components, nucleation and growth.[5] Nucleation is driven and favored by supersaturating the solution with high protein concentrations, but growth is favored by then reducing the protein concentration to prevent further nucleation from competing with the growth of established nuclei. We focus here on growth and on equilibrium protein crystal solubilities.

Both nucleation and growth processes are commonly treated using two-body models. George and Wilson have shown the power of two-body models. They showed that the "crystallization slot", which is the set of conditions that favor crystal nucleation, corresponds to mildly attractive values of the second virial coefficient B_2.[6] B_2 is a quantity that characterizes the interaction of two proteins with each other in solution.[6–9] The starting point for modeling both B_2 and protein aggregation has traditionally been the DLVO theory for colloid−colloid interactions. DLVO theory describes the interprotein potential as the sum of a screened electrostatic repulsion, a short-range van der Waals attraction, and a hard-core excluded volume term between two protein molecules. Two-body-based approaches have yielded tremendous insight into the behavior of B_2 and into nucleation behavior.[8–14]

However, here, we take the view that in a crystal, the electrostatics are different than those in pairwise protein−protein situations, and it is better treated as a multibody interaction. In particular, two-body models such as DLVO assume that counterions primarily serve to partially shield the protein−protein electrostatic repulsions. However, we believe that within a protein crystal, the high density of charge on the protein molecules must be fully neutralized by counterions in order for the crystal to be electrostatically neutral. The key consequence, described in the model below, is that the crystal contains an internal concentration of counterions that is much higher than the bulk solution concentration. This enrichment of counterions inside of the protein crystal, and the concomitant depletion of co-ions, comes at a considerable entropic cost that opposes crystal formation.[15,16] On the other hand, the counterions within the crystal fully screen the repulsion between proteins in the crystal and ensure that the primary contribution to the electrostatic enthalpy is the favorable protein−salt interaction. We find that in the lysozyme model system, the temperature dependence of the crystal solubility arises from the competition of these two terms.

2. Model

We present a model that describes the solubility of protein crystals. We define the solubility as the concentration of soluble protein in equilibrium with the crystal phase. At equilibrium, the chemical potential, μ_s, of the protein in the bulk solution must be the same as the chemical potential, μ_c, of the protein in the crystal

$$\mu_s = \mu_c \tag{1}$$

For the bulk phase, we assume that the protein is sufficiently dilute that its activity can be approximated by its concentration c, giving the standard expression

* To whom correspondence should be addressed. E-mail: dill@maxwell.compbio.ucsf.edu.

10.1021/jp9107188 © 2010 American Chemical Society
Published on Web 03/03/2010

$$\mu_s = k_B T \ln\left(\frac{c}{c^\circ}\right) \quad (2)$$

Here, $c^\circ = 1$ M is the standard state concentration, and $k_B T \ln(c/c^\circ)$ describes the translational entropy of the proteins in the bulk. To facilitate comparison with experiments, we adopt the same concentration units as those in refs 17 and 18, where $c^\circ = 1.47 \times 10^4$ mg/mL for lysozyme, and salt concentrations are expressed as a percent weight/volume, 1% NaCl w/v $\simeq$ 0.17 M.

Combining eqs 1 and 2 gives the equilibrium solubility of the protein crystal in terms of μ_c

$$\frac{c}{c^\circ} = e^{\mu_c/k_B T} \quad (3)$$

Now, we describe our model of the crystal and μ_c. We are interested in how solubility depends on pH, the concentration of salt c_{salt}, and temperature T; therefore, we decompose the free energy into its enthalpy and entropy components

$$\mu_c = \Delta h(\mathrm{pH}, c_{salt}, T) - T\Delta s(\mathrm{pH}, c_{salt}, T) \quad (4)$$

We assume that the crystal is large enough that we can neglect surface effects and that the enthalpy and entropy are composed of an electrostatic component that depends on the pH and salt and of a nonelectrostatic component that is temperature-independent

$$\Delta h(\mathrm{pH}, [\mathrm{NaCl}], T) = \Delta h_{ES}(\mathrm{pH}, [\mathrm{NaCl}]) + \Delta h_0 \quad (5)$$

$$\Delta s(\mathrm{pH}, [\mathrm{NaCl}], T) = \Delta s_{ES}(\mathrm{pH}, [\mathrm{NaCl}]) + \Delta s_0 \quad (6)$$

The approximation that Δh_0 and Δs_0 are temperature-independent is discussed in the Supporting Information.

We first consider the electrostatic terms. There are three species involved in the electrostatic interactions, the protein macroion, the counterions, and the co-ions. As with the protein, the chemical potentials of the ion species must be equal in the crystal and in the bulk solution, $\ln c_{bulk} = \ln c_\pm \pm e\phi/k_B T$. This leads to expressions for the concentrations of positive and negative counterions, $c_\pm = c_{bulk} e^{\mp e\psi/k_B T}$, which then lead to the Poisson−Boltzmann equation

$$\epsilon_w \epsilon_0 \nabla^2 \psi(r) = -\rho_f(r) - eN_A(c_{bulk} e^{-e\psi(r)/k_B T} - c_{bulk} e^{e\psi(r)/k_B T}) \quad (7)$$

Here, ψ is the electrostatic potential which we take to be zero in the bulk solution, ρ_f is the fixed charge distribution on the surface of the protein, c_{bulk} is the bulk salt concentration, N_A is Avogadro's constant, ϵ_0 is the vacuum permeability, $\epsilon_w \simeq 80$ is the dielectric constant of water, and r is the distance from the center of the protein, which we take here to be a sphere of radius $r = a$. Solving eq 7 with the appropriate boundary conditions gives the electrostatic potential ψ, which encodes the electrostatic influence of each component in the crystal.

We calculate the net signed number of charges as

$$Z = -\sum_i^{\text{acid residues}} \frac{10^{\mathrm{pH}-\mathrm{p}Ka_i}}{1 + 10^{\mathrm{pH}-\mathrm{p}Ka_i}} + \sum_i^{\text{basic residues}} \frac{10^{\mathrm{p}Ka_i-\mathrm{pH}}}{1 + 10^{\mathrm{p}Ka_i-\mathrm{pH}}} \quad (8)$$

where pH represents the solution pH, the pKa_i values are the pK's of the titratable side chains on the given protein, and we take the pK_a's from ref 19. We do not employ lysozyme-specific pK_a values as these are only representative of the solution state and are likely to be perturbed in the crystal. We have found that the use of other pK_a sets, such as those of Grimsley et al.,[20] have minimal effect on the model results, and shifts of several charge units are readily compensated for in the fitting procedure. Given the valency Z of charge on the protein obtained in this way, the electrostatic potential due to each protein alone will be $\psi_p(r) = Ze/4\pi\epsilon_0\epsilon_w r$.

Now, the task of specifying the boundary conditions for this PB equation is greatly simplified by recognizing that the crystal, which is macroscopic, must be electrostatically neutral. We associate with each protein an amount of aqueous volume that correctly accounts for the relative volume fractions of protein molecules plus water. We call this a "cell" (it is different than the actual crystallographic unit cell). We consider the total volume (water plus protein) per protein molecule to be approximately spherical with radius $r = b (b > a)$; see Figure 1. More accurate approximations could also be envisioned, but this is simple and appears to be sufficient for our purposes.

We solve for the interaction between each protein molecule and its ion cloud. We take the protein charge to be at the origin, $\rho_f = Ze\delta(r)$, where e is the elementary charge, and the free ions are confined between the protein surface at $r = a$ and the cavity boundary at $r = b$. The radii are chosen so that the sphere volumes match the appropriate crystal volumes. (The tetragonal crystal of lysozyme (PDB code 193L) has a unit cell with dimensions of $V_{uc} = 78.5 \times 78.5 \times 37.8$ Å^3 and contains eight lysozyme molecules.[21] Since the solvent content is 39.5%, we have $a = (0.605 \times 3V_{uc}/32\pi)^{1/3} = 16.1$ Å and $b = (3V_{uc}/32\pi)^{1/3} = 19.1$ Å. For the orthorhombic crystals of lysozyme (PDB code 1AKI), the unit cell contains four proteins, has a volume $V_{uc} = 59.1 \times 68.5 \times 30.5$ Å^3, and has a solvent content of 42.8%.[22] Therefore, $a = (0.572 \times 3V_{uc}/16\pi)^{1/3} = 16.1$ Å and $b = (3V_{uc}/16\pi)^{1/3} = 19.5$ Å.)

There are two boundary conditions

$$\psi'(r = a) = -\frac{Ze}{4\pi\epsilon_w\epsilon_0 a^2} \quad (9)$$

$$\psi'(r = b) = 0 \quad (10)$$

First, at the protein surface, the electric field must be equal to that of the unscreened macroion. Second, at the outer sphere of the cell, the charge neutrality condition requires that the electric field vanish. The latter boundary condition enforces the Donnan equilibrium between the ions in the interior of the crystal and in the solution.[15,16]

The vanishing field at $r = b$ implies that the cells are noninteracting. While each protein "feels" a repulsive force due to each neighboring protein, there is an equal and opposite attraction to the counterions associated with those proteins. Therefore, the charges on the proteins are fully screened by the salt. Note that the vanishing field approximation is not valid

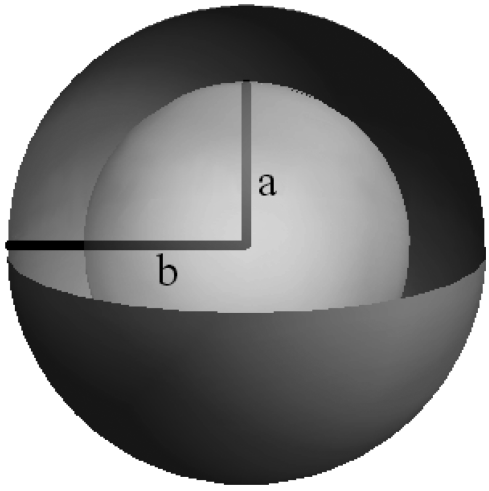

Figure 1. Cell geometry used to solve eq 7. The inner sphere $r = a$ represents the protein with a charge $q = Ze$ located at the origin. The outer sphere $r = b$ represents the total volume (protein + solvent) per protein in the crystal.

for proteins with highly asymmetric charge distributions (i.e., strong dipole or quadrupole moments). This is not a concern in the present work where we study lysozyme, which has a dipole moment of ~250 D or ~50 e Å in the pH range of interest. This is to be compared with a total charge of ~10 e, which is spread over a diameter of ~30 Å.

We now compute the electrostatic potential, ψ_s, of the salt ions in the cell surrounding the protein

$$\psi_s(\vec{x}) = c_{bulk} \int (e^{\psi(\vec{x}')e/k_BT} - e^{-\psi(\vec{x}')e/k_BT})/|\vec{x} - \vec{x}|d^3\vec{x}' \quad (11)$$

where the integration is over the solvent volume in the cell. While ψ_s may be calculated directly from a numerical integration of eq 11, it is more readily obtained as follows. First, we note that because of the geometric symmetry and the charge neutrality of the cell, the sum of the salt and protein terms must vanish at the cell boundary due to Gauss' Law

$$\psi_p(b) + \psi_s(b) = 0 \quad (12)$$

However, the overall potential will be nonzero at $r = b$, leading to a constant of integration that we call $\psi(b) \equiv \psi_0^{crys}$. The constant ψ_0^{crys} is a shift in the potential at all points within the crystal. It arises due to the ionization at the surface of the macroscopic crystal. However, ψ_0^{crys} does not contribute to the protein−salt potential energy within the interior of the crystal because the electric field of the surface charge vanishes within the crystal due to Gauss' Law. The complete potential within a cell is

$$\psi^{crys}(r) = \psi_p(r) + \psi_s^{crys}(r) + \psi_0^{crys} \quad (13)$$

The protein−salt interaction is $Ze\psi_s^{crys}(a)$, which we calculate from the numerical solution to eq 7 using

$$\psi_s^{crys}(r) = \psi^{crys}(r) - \psi_p(r) - \psi^{crys}(b) \quad (14)$$

Now, we need the protein−salt attraction in the bulk state. Here, $\psi_0^{sol} = 0$; therefore, the electrostatic potential is

$$\psi^{sol}(r) = \psi_s^{sol}(r) + \psi_p(r) \quad (15)$$

and the protein−salt interaction is $Ze(\psi^{sol}(a) - \psi_p(a))$. Hence, the difference in electrostatic potential that drives crystal formation is

$$\Delta h_{ES} = Ze(\psi_s^{crys}(a) - \psi_s^{sol}(a)) \quad (16)$$

$$= Ze(\psi^{crys}(a) - \psi^{crys}(b) - \psi^{sol}(a)) \quad (17)$$

We calculate the terms in eq 17 by numerically solving the Poisson−Boltzmann equation for the geometry shown in Figure 1. ψ^{sol} is computed using $b = 50$ Å. This is sufficient to avoid boundary effects as the Debye length at the lowest salt concentration that we consider is ~5 Å. We restrict our analysis to salt concentrations ≤5% w/v in order for $c_{bulk} \leq 1$ M.

Because the protein has high charge density, $e\psi/k_BT \gg 0$, we must use the nonlinear version of the PB equation. However, in the Supporting Information, we give a parallel treatment using the linearized (Debye−Huckel) approximation because this analytical form gives useful insights into dependencies on variables.

Next, we consider the electrostatic translational entropy, the cost of sequestering the counterions at high concentration in the crystal relative to the lower concentration in the bulk. In general, the entropic cost $T\Delta s$ for moving a single ion from a medium 1 with concentration c_1 to a medium 2 with concentration c_2 will be

$$T\Delta s = -k_BT \ln\left(\frac{c_2}{c_1}\right) \quad (18)$$

If the bulk ion concentration is c_{bulk} and if V is the aqueous cavity volume in the cell, then the expected number of ions in the cell if the protein were uncharged would be $n = c_{bulk}V$. If m is the number of excess ions per protein when the protein is charged, the entropy of sequestering and concentrating the ions in the crystal can be computed from a process of "charging up" the crystal with its salt

$$\Delta s_{ES}/k_B = -\int_0^{m_f} \ln\left(\frac{n+m}{n}\right)dm \quad (19)$$

$$= -(n + m_f)\ln\left(1 + \frac{m_f}{n}\right) + m_f \quad (20)$$

To evaluate this entropy, we need to know the numbers, m_f, of both the counterions and the co-ions in the protein crystal. Because of charge neutrality, we must have $m_f^- - m_f^+ = Z$ or

$$4\pi c_{bulk} \int_a^b (e^{\psi(r)e/k_BT} - e^{-\psi(r)e/k_BT})r^2dr = Z \quad (21)$$

where the integral is over the aqueous cavity volume in the cell. The counterion numbers may be determined by assuming that the potential varies slowly within the aqueous cavity and thus may be replaced with an average potential $\psi(r) \rightarrow \bar{\psi}$, yielding

$$Z = \frac{8\pi c_{bulk}}{3}(b^3 - a^3)\sinh\frac{\bar{\psi}e}{k_BT} \quad (22)$$

This equation can be solved for the constant potential $\bar{\psi}$, which can then be used to determine the counterion enrichment numbers

$$m_f^{\pm} = 4\pi c_{bulk} \int_a^b (e^{\mp \bar{\psi} e/k_B T} - 1) r^2 dr \quad (23)$$

$$= \frac{4\pi c_{bulk}}{3}(b^3 - a^3)(e^{\mp \bar{\psi} e/k_B T} - 1) \quad (24)$$

$$= \frac{4\pi c_{bulk}}{3}(b^3 - a^3) \times \left(\exp\left[\mp \sinh^{-1} \frac{3Z}{8\pi c_{bulk}(b^3 - a^3)}\right] - 1\right) \quad (25)$$

Finally, we assemble all of these terms to get μ_c. The total chemical potential for the crystalline state is given by combining eqs 4, 5, 6, 17, and 20

$$\mu_c = \Delta h_{ES} + \Delta h_0 - T\Delta s_{ES} - T\Delta s_0 \quad (26)$$

$$= Ze(\psi^{crys}(a) - \psi^{crys}(b) - \psi^{sol}(a)) + k_B T\left((n + m_f^+)\ln\left(1 + \frac{m_f^+}{n}\right) + m_f^+\right) + k_B T\left((n + m_f^-)\ln\left(1 + \frac{m_f^-}{n}\right) + m_f^-\right) + \Delta h_0 - T\Delta s_0 \quad (27)$$

To compare to experiments, we convert this expression from μ_c to the protein solubility c using eq 3. Of course, each protein crystal differs in detail from the next. One type of protein will contact its neighbors through different contact interactions than another type of protein or even than another crystal lattice of the same protein. Therefore, in the absence of more microscopic knowledge, we currently treat the two quantities Δh_0 and Δs_0 as adjustable parameters that we fit for any given protein and lattice. We note

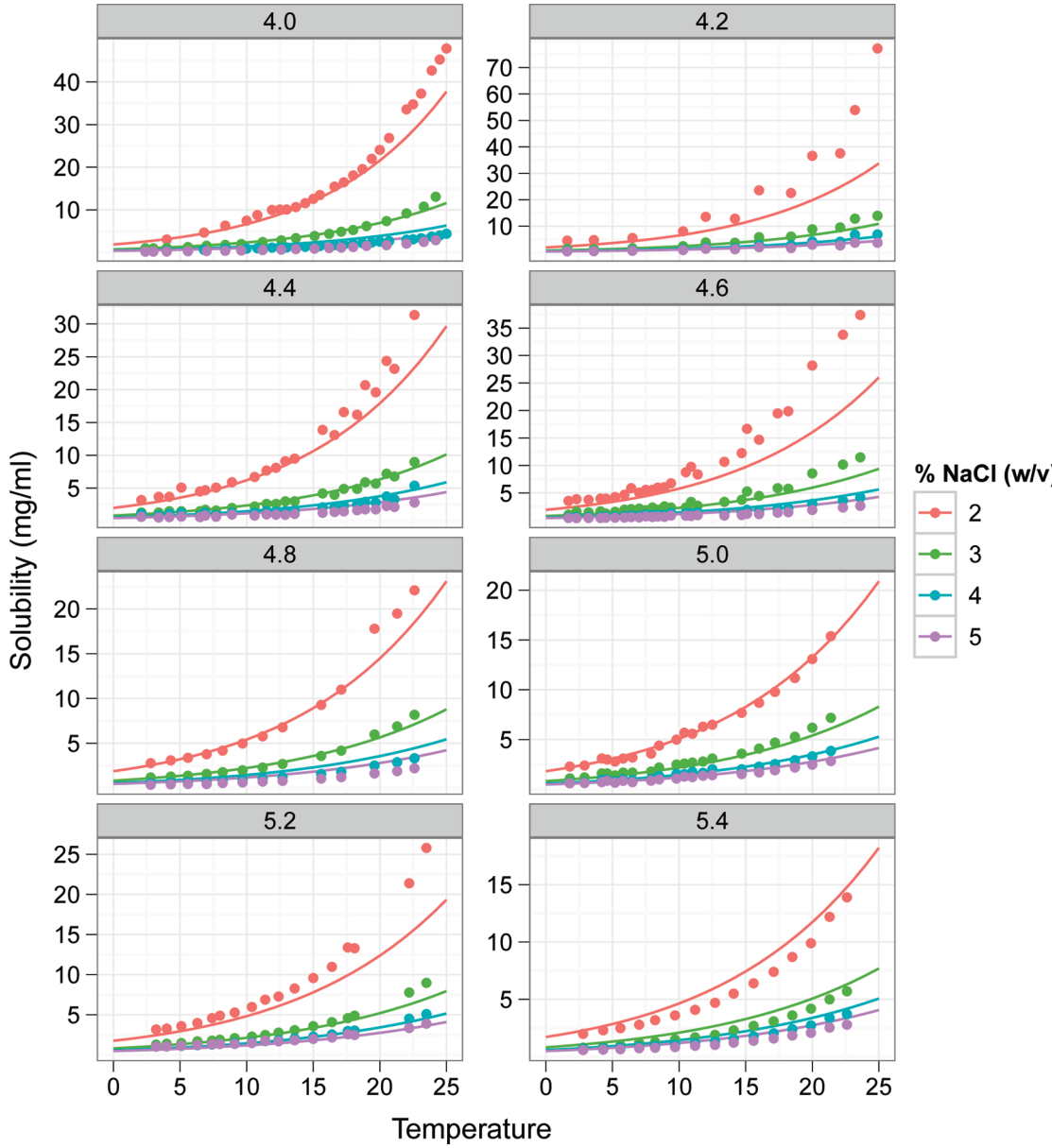

Figure 2. Comparison of the experimental solubility of tetragonal lysozyme crystals (dots)[17] to the predicted solubility, eqs 3 and 27 (lines). Numbers at the top of each panel indicate solution pH.

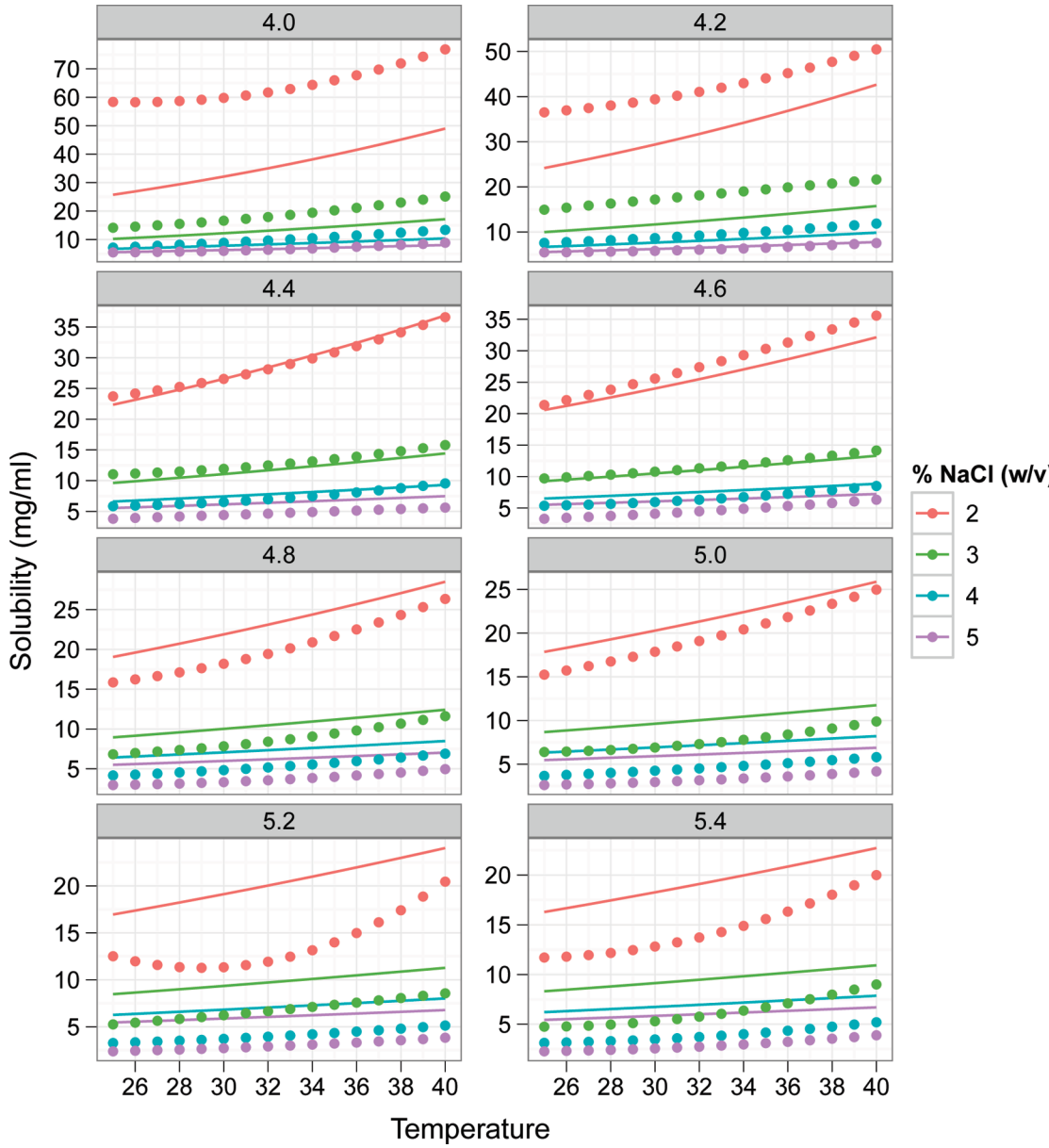

Figure 3. Comparison of the experimental solubility of orthorhombic lysozyme crystals to the predicted solubility, eqs 3 and 27 (lines). The points are plotted from quadratic fits to the solubility as determined in ref 18. Numbers at the top of each panel indicate solution pH.

in the Supporting Information that these two quantities give useful information about the contact interactions for the given protein in its crystal.

3. Results

3.1. Tetragonal Crystals of Lysozyme. For lysozyme in tetragonal crystals, we use $\Delta h_0 = -49.4$ kJ/mol and $\Delta s_0 = -95.7$ J/K/mol, so the contact free energy is $\Delta G = \Delta h_0 - T\Delta s_0 = -22.1$ kJ/mol at $T = 285$ K. Details of the fitting procedure are presented in the Supporting Information. Figure 2 compares the predictions of the theory with the data on the solubilities of lysozyme crystals of Forsythe et al.[17] The theory is in good general agreement with the dependence of the crystal solubility versus temperature, pH, and salt concentrations. Here are the main results. First, not surprisingly, heating dissolves the crystal. Since $\Delta s_0 > 0$ in this case, this effect of temperature is explained by the translational entropies (of both the protein and the salt), corresponding to canonical effects of melting in many simpler systems. Second, adding salt stabilizes the crystalline state. This has traditionally been attributed to the screening of the protein−protein electrostatic repulsions by the salt. However, our model indicates instead that this is mainly due to the reduced penalty in translational entropy of sequestering salt as the salt concentration increases. Third, changing the pH has only relatively small effects on the crystal solubility. In the model, this is because of the electrostatic compensation that is required by charge neutrality. Increasing the charge on the protein leads to a compensating increase in salt sequestered by the protein. This increased sequestration of counterions results in increasingly favorable protein−salt interactions for strongly charged proteins but is opposed by the counterion translational entropy. These two effects are offsetting and result in the weak pH dependence.

3.2. Orthorhombic Crystals. We also tested our model on the extensive data of Ewing et al.[18] on the solubilities of orthorhombic crystals of lysozyme. For these crystals, we found that $\Delta h_0 = -3.6$ kJ/mol and $\Delta s_0 = 54.3$ J/K/mol provide the best fits, resulting in $\Delta G = \Delta h_0 - T\Delta s_0 = -20.2$ kJ/mol at T

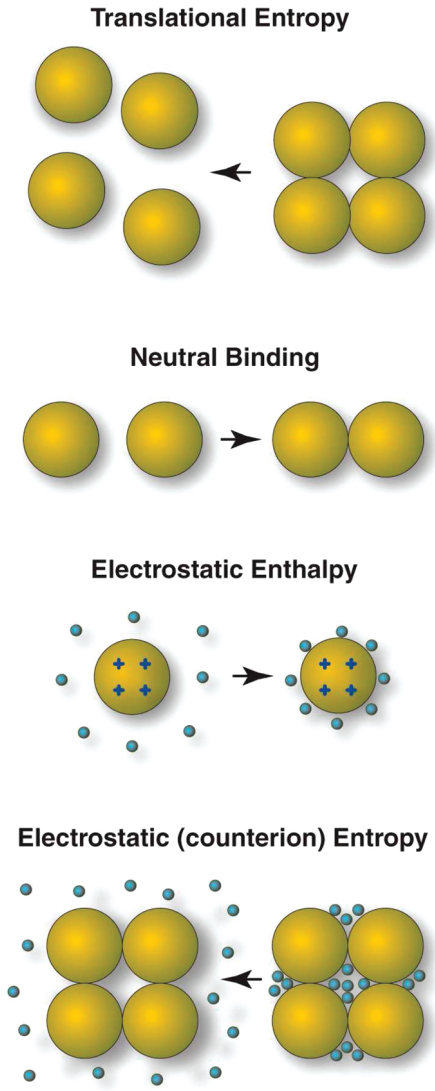

Figure 4. List of the terms included in the free energy for the binding of a protein to the crystal given by eq 27. Terms favoring aggregation are indicated with an arrow pointing to the right, while those opposing aggregation have a left arrow.

= 305 K. These contact parameters are quite different than those for the tetragonal crystals. We discuss these differences in more detail in the Supporting Information.

Figure 3 compares the predictions with experiments. The same overall temperature, pH, and salt trends can be seen here as in the tetragonal crystals. In the orthorhombic crystals, we find that the solubility is more sensitive to pH than our model predicts. However, the maximum deviations are less than a factor of 3, so the free energy errors are only on order of $\sim 1 k_B T$, a satisfactory result given the small number of free parameters.

4. Component Contributions

Figure 4 shows the four component contributions in our model of crystal solubility. There are two nonelectrostatic terms, (1) the translational entropy of the protein opposes crystallization, and (2) the contact free energy of binding, $\Delta h_0 - T\Delta s_0$, drives crystal formation. There are two electrostatic terms, (3) the enthalpy, Δh_{ES} of the protein−counterion charge stabilization and (4) the translational entropy of sequestering the salt ions into the crystal, which opposes crystallization.

Here, we look at the behavior of these component terms. First, Figure 5a shows that the electrostatic potential is relatively

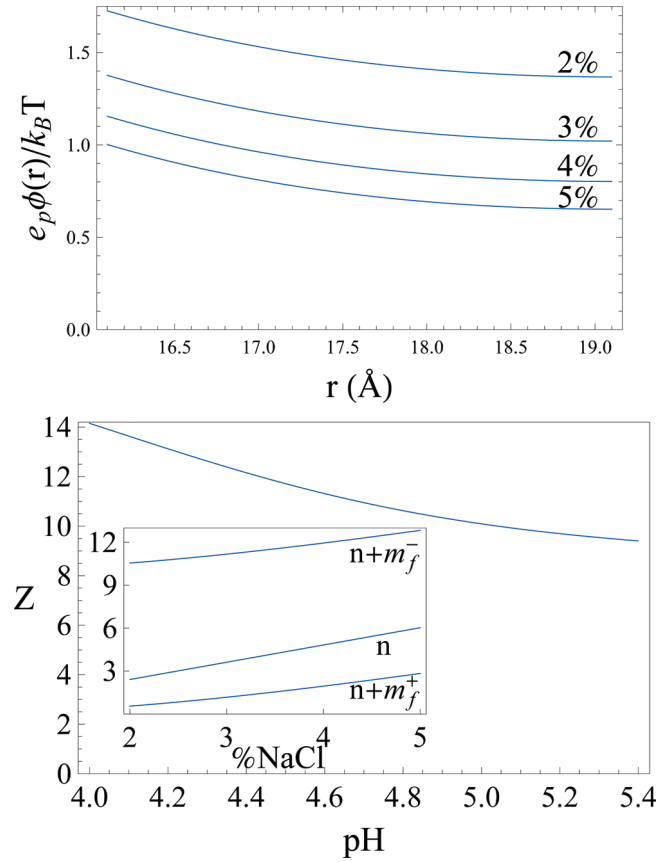

Figure 5. (top) Solutions for the electrostatic potential ψ for a protein charge $Z = +10$ at various salt concentrations. (bottom) Calculated charge from lysozyme using eq 8. Inset: plot of the expected counterion and co-ion numbers per protein for a protein of charge $Z = 10$. Also shown are the expected ion numbers n for a neutral protein. All figures were computed using $a = 16.1$ Å, $b = 19.1$ Å.

constant throughout each protein cell's water volume. It also shows that even at the highest salt studied here, 5%, the electrostatic potential in the cell $e_p\psi/k_BT \approx 1$ is too high to allow for using the linearized PB approach, which requires $e_p\psi/k_BT \ll 1$. Figure 5b shows the calculated protein charge as a function of pH, as well as the ion enrichment as a function of salt concentration.

Second, Figure 6 shows the electrostatic enthalpy. Comparison of the black with the red curves confirms the point noted above that using the linearized form of PB would lead to large errors here.

Third, Figure 7 shows that the salt counterion entropies constitute a large driving force opposing crystallization. This entropy diminishes with increasing bulk salt concentrations.

4.1. pH Dependence. A surprising trend is the weak dependence of the solubility on the solution pH shown in Figure 8. This stands in contrast to the strong dependence on pH that is observed for the nucleation step.[1] We believe that this difference is explained as follows. Crystal nucleation appears to be governed heavily by the two-body interactions of a protein as it lands on the nucleating surface. These nucleation events are too small to condense the counterions into a captured and sequestered phase. Hence, nucleation depends strongly on the charge−charge repulsion between two protein molecules. However, as we noted above, we believe that the crystal electrostatics are different; the protein−protein repulsions are neutralized by the sequestered salt ions, leading to a much weaker dependence of solubility on protein charge than that of

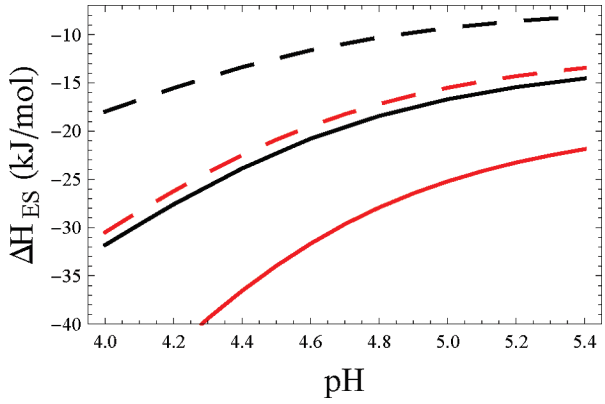

Figure 6. Change in the protein−counterion interaction upon the transfer from the solution to the crystal state, eq 17, as a function of pH as calculated numerically from eq 7 (black) and from the linearized approximation eq S17 (red). The solid lines are computed for 2% and the dashes are computed for 5%.

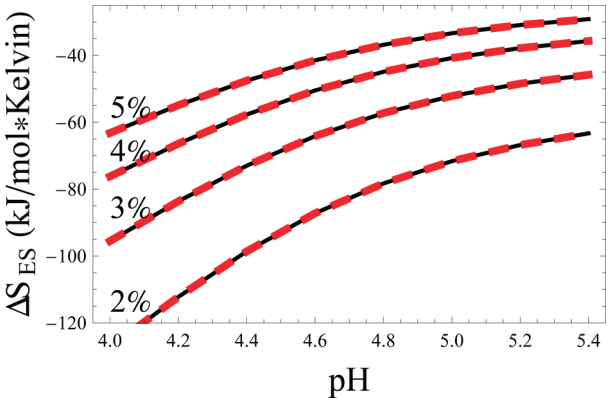

Figure 7. Entropy cost to neutralize the lysozyme crystal as a function of salt concentration (w/v NaCl) and pH. This entropy is calculated using the constant potential approximation (eq 22, red dashes) and numerically from eq 7 (black lines) and includes contributions from both the counterions and co-ions as calculated using eq 20.

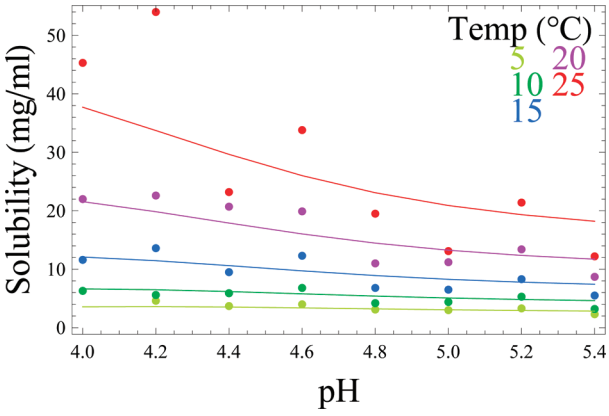

Figure 8. Predicted solubility at 2% NaCl for tetragonal lysozyme crystals as a function of pH.

nucleation on protein charge. In particular, the pH dependence is weakest at low temperatures when the counterion confinement free energy $-T\Delta s_{ES}$ is minimized.

4.2. Salt Dependence. In Figure 9, we plot the log of the protein crystal solubility as a function of the salt concentration. An early empirical formula that has been called Cohn's law[23,24] suggested that this should be a linear function. However, the newer data shown in Figure 9 indicates that instead, this dependence is nonlinear. Our theory captures this nonlinearity.

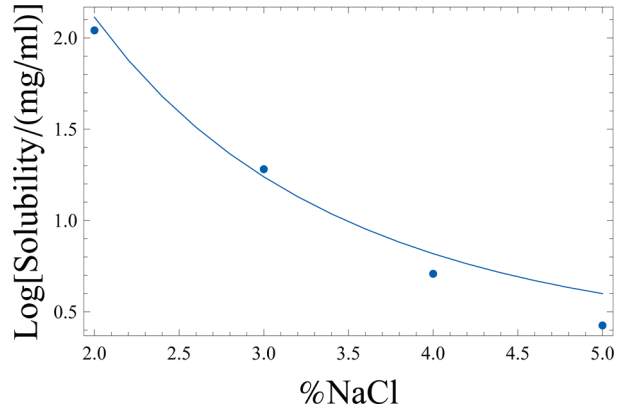

Figure 9. Plot of the logarithm of solubility for tetragonal crystals versus NaCl concentration at $T = 15$ °C and pH 5.0.

We find that neither H_{ES} nor S_{ES} approach linear behavior at small c_{bulk}. At sufficiently dilute salt, the effect of co-ions may be neglected, and the sinh term in eq 22 can be approximated with an exponential. Then, $m_f^- = Z - n$, and the entropy penalty, eq 6, diverges as $-Z \ln(Z/n)$. The enthalpy term, given approximately by eq S17 (Supporting Information), has a leading behavior

$$\Delta h_{ES} \simeq \frac{Z^2 e^2}{4\pi\epsilon_w\epsilon_0}\left(-\frac{3(a+b)}{2(a^2+ab+b^2)} + \kappa + \mathcal{O}(\kappa^2)\right) \tag{28}$$

at low salt concentrations. Since $\kappa \propto c_{bulk}^{1/2}$, this is nonlinear at all values of c_{bulk}. We conclude that Cohn's formula is only likely to apply in cases where the solvent cavities in the aggregate are large enough that Δh_{ES} and Δs_{ES} are slowly varying functions of the salt concentration.

5. Conclusion

We have described a model for the solubilities of protein crystals as a function of temperature, pH, and salt. Our principal finding is that the charge neutrality condition for the crystal eliminates the screened monopole interaction between proteins. This effectively isolates each protein as a local electrostatic entity. Crystal formation is favored by increasing concentrations of protein and salt, which reduce the unfavorable translational entropies. We find that pH, which changes the charge on the protein, has relatively little effect on crystal solubility because it leads to a compensating increase in counterion concentration inside of the crystal. We hope that such models of protein crystallization will be useful for better understanding the physics and for making crystallization technology more efficient and rational.

Acknowledgment. This work was supported by NIH Grant GM34993.

Supporting Information Available: Discussion of temperature effects, linearized approximation for the enthalpy change, and details of the fitting procedure. This material is available free of charge via the Internet at http://pubs.acs.org.

References and Notes

(1) McPherson, A. *Crystallization of biological macromolecules*; Cold Spring Harbor Laboratory Press: Cold Spring Harbor, NY, 1999.

(2) Rieskautt, M.; Ducruix, A. *Macromol. Crystallogr., Part A* **1997**, *276*, 23–59.

(3) Anderson, M.; Hansen, C.; Quake, S. *Biophys. J.* **2005**, *88*, 55A–55A.

(4) Anderson, M.; Hansen, C.; Quake, S. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 16746–16751.

(5) Chayen, N.; Saridakis, E. *Nat. Methods* **2008**, *5*, 147–153.

(6) George, A.; Wilson, W. *Acta Crystallogr.* **1994**, *D50*, 361–5.

(7) Velev, O.; Kaler, E.; Lenhoff, A. *Biophys. J.* **1998**, *75*, 2682–2697.

(8) Elcock, A.; Mccammon, J. *Biophys. J.* **2001**, *80*, 613–625.

(9) Lima, E.; Biscaia, E.; Bostrom, M.; Tavares, F.; Prausnitz, J. *J. Phys. Chem. C* **2007**, *111*, 16055–16059.

(10) Kuehner, D.; Blanch, H.; Prausnitz, J. *Fluid Phase Equilib.* **1996**, *116*, 140–147.

(11) Soumpasis, D.; Georgalis, Y. *Biophys. J.* **1997**, *72*, 2770–2774.

(12) Curtis, R.; Newman, J.; Blanch, H.; Prausnitz, J. *Fluid Phase Equilib.* **2001**, *192*, 131–153.

(13) Tavares, F.; Bratko, D.; Prausnitz, J. *Curr. Opin. Colloid Interface Sci.* **2004**, *9*, 81–86.

(14) Tavares, F.; Bratko, D.; Striolo, A.; Blanch, H.; Prausnitz, J. *J. Chem. Phys.* **2004**, *120*, 9859–9869.

(15) Warren, P. *J. Phys.: Condens. Matter* **2002**, *14*, 7617–7629.

(16) Prinsen, P.; Odijk, T. *J. Chem. Phys.* **2006**, *125*, 074903.

(17) Forsythe, E.; Judge, R.; Pusey, M. *J. Chem. Eng. Data* **1999**, *44*, 637–640.

(18) Ewing, F.; Forsythe, E.; Pusey, M. *Acta Crystallogr., Sect. D* **1994**, *50*, 424–428.

(19) Sillero, A.; Ribeiro, J. *Anal. Biochem.* **1989**, *179*, 319–325.

(20) Grimsley, G.; Scholtz, J.; Pace, C. *Protein Sci.* **2009**, *18*, 247–251.

(21) Vaney, M.; Maignan, S.; Ries-Kautt, M.; Ducriux, A. *Acta Crystallogr.* **1996**, *52*, 505–517.

(22) Artymiuk, P.; Blake, C.; Rice, D.; Wilson, K. *Acta Crystallogr.* **1982**, *38*, 778.

(23) Cohn, E. *Physiol. Rev.* **1925**, *5*, 349.

(24) Cohn, E.; Ferry, J. The interactions of proteins with ions and dipolar ions. In *Proteins, amino acids and peptides as ions and dipolar ions*; Cohn, E., Edsall, J., Eds.; Reinhold: New York, 1943.

JP9107188