Phase diagram for a model of urate oxidase

N. Wentzel, D. L. Pagan, J. D. Gunton

Department of Physics, Lehigh University, Bethlehem, Pennsylvania, 18015

(Dated: May 31, 2007)

Abstract

Urate Oxidase from *Asperigillus flavus* has been shown to be a model protein for studying the effects of polyethylene glycol (PEG) on the crystallization of large proteins. Extensive experimental studies based on SAXS [Vivares et al, J. Phys. Chem. B 108, 6498 (2004)] have determined the effects of salt, pH, temperature, and most importantly PEG, on the crystallization of this protein. Recently, some aspects of the phase diagram have also been determined experimentally. In this paper we use Monte Carlo techniques to predict the phase diagram for urate oxidase in solution with PEG, including the liquid-liquid and liquid-solid coexistence curves. The model used includes an electrostatic interaction, van der Waals attraction, and a polymer-induced depletion interaction [Vivares et al, Eur. Phys. J. E 9, 15 (2002)]. Results from the simulation are compared with experimental results.
I. INTRODUCTION

In studies of phases of protein solutions, polymer reagents, especially polyethylene glycol (PEG), are commonly used as crystallizing agents\textsuperscript{1,2}. Aggregation of proteins in the presence of PEG or other polymers is usually attributed to so-called depletion effects, in which the overlap of excluded volumes around the proteins increases the volume of the solution available to the polymers, thereby increasing the total entropy of the polymer-protein mixture.

An analytical form to describe the depletion effect via an effective attraction between the colloidal particles in the limit $q \equiv \sigma_{\text{pol}} / \sigma_{\text{col}} \ll 1$, where the colloidal particle diameter $\sigma_{\text{col}}$ is much greater than the polymer diameter $\sigma_{\text{pol}}$, was first developed by Asakura and Oosawa\textsuperscript{3,4}, who assumed that both the colloids and polymers are hard spheres. Because polymers may not behave like hard spheres, other studies have tried to develop more realistic models for an effective pair interaction, including the analytical PRI SM model\textsuperscript{5–8} that accounts for internal degrees of freedom of the polymers. Numerical forms for the depletion interaction include density functional calculations for hard spheres\textsuperscript{9} and generalized non-additive hard-sphere models\textsuperscript{10}. The Asakura-Oosawa (AO) model, however, has been shown to produce a reasonable phase diagram for hard sphere solutions\textsuperscript{11–13}.

Measurements of the pair potential between colloids mixed with small polymers in the dilute and semi-dilute regimes\textsuperscript{14,15} has shown that the AO model works well in the dilute regime; a rescaled version\textsuperscript{16} works well in the semi-dilute regime. For solutions of particles with $q \geq 1$, the role of polymer in precipitation is somewhat more complicated. Theoretical investigation has found that in this limit the depletion interaction is well described by accounting for many-body interactions, but it is poorly described by effective pair potentials\textsuperscript{17}. Experimental results have shown that adding PEG can induce repulsion between lysozyme molecules, rather than the attraction predicted by depletion theory\textsuperscript{18}.

Larger proteins like urate oxidase or glucose isomerase, while smaller than many colloidal particles, have been studied for polymer sizes where $q < 1$ and $q > 1$. These proteins do not
aggregate from adding salt alone, but they have been shown to exhibit attractive behavior with the addition of PEG\textsuperscript{19–22}. This attraction is attributed to depletion effects. Salts that are commonly used to aid in forming aggregates of some globular proteins do not aid in forming aggregations of urate oxidase. PEG, however, will cause urate oxidase to aggregate even in the absence of salts.

In this paper we study the effect of PEG on phase diagrams for proteins in solution using model protein-protein interactions based on the experimental and theoretical work on urate oxidase\textsuperscript{19–21}. This model includes three components: an electrostatic repulsion, a van der Waals attraction, and a depletion interaction. The first two effects are considered using the so-called DLVO theory for colloidal interactions\textsuperscript{23,24}, a commonly studied model. The depletion interaction is modeled using either the Asakura-Oosawa (AO) or PRISM models. The depletion term is such that when the PEG concentrations are low, the DLVO-like terms dominate the interaction. Adding PEG increases the effect of the depletion interaction, overcoming even weakly repulsive conditions in the DLVO model and making the net protein-protein interaction attractive. For urate oxidase, either the AO or the PRISM interaction meets these requirements. In our work we use the AO model. Small angle x-ray scattering results (SAXS), together with an approximate theory, have been used to determine the depletion potential for urate oxidase molecules in solution with PEG\textsuperscript{20}. In that work, which did not assume any depletion interaction between urate oxidase molecules but rather measured the depletion interaction, the AO interaction was found to be slightly stronger than the measured depletion. A model using DLVO plus the PRISM interaction has been used to study the effects of PEG on second virial coefficients for lysozyme and bovine serum albumin\textsuperscript{25,26}. Other studies have used DLVO and AO to study crystallization of proteins such as apoferritin\textsuperscript{27} and cytochrome BC1\textsuperscript{28}.

In this paper, we calculate the equilibrium phase diagram for the model, finding the coexistence curves for both the liquid-liquid and liquid-solid phases, giving the protein concentration as a function of the PEG concentration for various temperatures. For the liquid-liquid
phase separation we use Gibbs Ensemble Monte Carlo (GEMC) simulations. We also provide a semi-quantitative calculation of the liquid-solid and liquid-liquid coexistence curves using first order thermodynamic perturbation theory.

In section 2 we discuss the model we use for the interaction between urate oxidase proteins. In section 3 we discuss the methods we use to analyze the model. We present our results in section 4 and a conclusion in section 5.

II. MODEL

The pair potential we study is the sum of two direct potentials, the DLVO potential and the AO depletion potential, that describe the interaction between two spherical protein molecules. We use the values of the parameters in the model that were estimated by Vivarès et al.\textsuperscript{20} The DLVO potential is itself the sum of three distinct potentials— a hard sphere potential, a screened Coulomb repulsion, and a van der Waals attraction. The Coulomb potential represents the charge-induced repulsive force between two spherical particles. The van der Waals attraction represents the dispersion interaction between two spherical particles.

The hard sphere potential is written as

\[
V_{HS}(r) = \begin{cases} 
\infty, & r < \sigma \\
0, & r \geq \sigma,
\end{cases}
\]

where \(\sigma\) is the diameter of the particles. The repulsive Coulomb potential, valid outside the hard core diameter \((r > \sigma)\) is written in \(k_B T\) units as

\[
V_{Coul}(r) = \frac{Z_{\text{eff}}^2 L_B}{\sigma (1 + 0.5 \sigma / \lambda_D)^2} \frac{\sigma}{r} \exp\left(-\frac{r - \sigma}{\lambda_D}\right)
\]

where \(Z_{\text{eff}}\) is the effective charge on the particles, \(\lambda_D\) is the Debye screening length, and \(L_B = e^2 / (4\pi \epsilon_s \epsilon_0 k_B T)\) is the Bjerrum length. We write the van der Waals attraction in \(k_B T\) units as...
units, also valid only when \( r > \sigma \), as

\[
V_{vdW} = -J_{vdW} \frac{\sigma}{r} \exp\left(\frac{-(r - \sigma)}{d_{vdW}}\right)
\]

where \( J_{vdW} \) (in \( k_B T \) units) and \( d_{vdW} \) are, respectively, a characteristic depth and range for the interaction. Note that this Yukawa form for the van der Waals interaction has been shown to be sufficient to describe pair potentials in protein solutions\(^{29,30}\). For \( r > \sigma \) the depletion interaction is represented by the Asakura-Oosawa potential, written in \( k_B T \) units as

\[
V_{AO}(r) = \begin{cases} 
-\frac{4}{3} \pi d^3 n_p \left[ 1 - \frac{3r}{4d} + \frac{r^3}{16d^3} \right], & \sigma \leq r \leq 2d \\
0, & r > 2d,
\end{cases}
\]

where \( n_p \) is the number density of the polymer and \( d = \left( r_g + \sigma / 2 \right) \) is the closest center-to-center distance between a polymer with radius of gyration \( r_g \) and a protein of diameter \( \sigma \). The total interaction model is the sum of the expressions in equations 1 through 4

\[
V(r) = V_{HS}(r) + V_{coul}(r) + V_{vdW}(r) + V_{AO}(r)
\]

The values that we use to calculate the potential for the urate oxidase and 8000Da PEG system studied by Vivarès et al\(^{20}\) are given in Table I. The shape of \( V(r) \) using these parameters is shown in figure 1.

The osmotic second virial coefficient \( B_2 \) for the model can be found using the equation\(^{31}\)

\[
B_2 = 2\pi \int_0^{\infty} r^2 \left( 1 - e^{-V(r)/k_BT} \right) dr
\]

Figure 2 shows \( B_2 / B_{2HS} \) as a function of PEG concentration for \( T = 293.15 \), where \( B_{2HS} = 2\pi/3 \).
III. METHODS

A. Simulations

The bulk of our work involves the use of Monte Carlo methods to study the metastable liquid-liquid coexistence for the model. We use the standard Gibbs Ensemble Monte Carlo (GEMC) method\textsuperscript{32–34} for a system of $N = 512$ particles. The runs equilibrated for $1.3 \times 10^9$ Monte Carlo steps; after equilibration we obtained statistics for $0.7 \times 10^9$ Monte Carlo steps. Because of the deep contact energy for two particles given the parameters in our model, our equilibration times were quite long due to the interim formation of non-equilibrated clusters. (An alternative method that would break up these clusters and hence equilibrate faster, such as parallel tempering\textsuperscript{33,34} would reduce the number of steps required.)

B. Perturbation Theory

An alternative, quicker, but less accurate approach to the study of the phase diagram is thermodynamic perturbation theory\textsuperscript{35–37}. In our case we treat the net attractive pair potential potential as a perturbation to a hard sphere system. Following previous studies\textsuperscript{35–38} we expand the free energy per particle as

$$\beta f = \beta f_0 + \beta f_1 + \beta O (\beta^2)$$  \hspace{1cm} (7)

where $f_0$ is the free energy of a reference system and $\beta f_1$ is the first perturbation term. We will calculate $f$ only to first order, although higher order calculations are possible. The first order term $\beta f_1$ is given by

$$\beta f_1 = 2\pi \rho \beta \int_0^\infty r^2 V(r) g_0(r) \, dr$$  \hspace{1cm} (8)
where \( V(r) \) is the protein-protein interaction potential and \( g_0(r) \) is the radial correlation function for the reference system.

For our liquid reference system we use the well known hard sphere system; for the hard sphere fluid, the free energy per particle and the equation of state are given by the Carnahan-Starling expressions\(^{39}\)

\[
\beta f_0^{\text{liq}} = \frac{\eta(4 - 3\eta)}{(1 - \eta)^2} \ln(\eta) \tag{9}
\]

\[
\frac{\beta p_0^{\text{liq}}}{\rho} = 1 + \eta + \eta^2 - \eta^3 \frac{1 - \eta}{(1 - \eta)^3}. \tag{10}
\]

Here \( \eta = \pi\rho/6 \) is the packing fraction for spheres of unit diameter. Along with these equations for the reference system, we use Boublık’s form\(^{40}\) for the fluid phase radial distribution function \( g_0^{\text{liq}}(r) \) to calculate the first perturbation term for the liquid, \( \beta f_1^{\text{liq}} \), using equation 8.

To calculate the free energy per particle and equation of state for the hard sphere solid we use as the reference system the expressions for the fcc solid given by Alder et al\(^{41}\), namely

\[
\beta f_0^{\text{sol}} = \ln \rho_0 + 3 \ln \left( \frac{3}{2\alpha} \right) - S_0 - S_1\alpha - S_2\alpha^2 \tag{11}
\]

\[
\frac{\beta p_0^{\text{sol}}}{\rho} = \frac{3}{\alpha} + C_0 + C_1\alpha. \tag{12}
\]

The coefficients \( S_0, S_1, S_2, C_0, \) and \( C_1, \) which depend in part on the dimensionality \( D \) of the system, are listed in table II; \( \rho_0 = \sqrt{2} \) is the close-packed density, and \( \alpha = (\rho_0 - \rho)/\rho. \) Along with these equations for the reference systems, we use the expression for the solid phase radial distribution function \( g_0^{\text{sol}}(r) \) from Choi et al\(^{42}\) to calculate the first perturbation term for the solid, \( \beta f_1^{\text{sol}} \), using equation 8.

Knowing the free energy, other thermodynamic functions can be determined. The pres-
sure, \( p \), and chemical potential, \( \mu \), can be found using the following standard thermodynamic relationships:

\[
\frac{\beta p}{\rho} = \rho \left( \frac{\partial \beta f}{\partial \rho} \right)
\]

(13)

\[
\beta \mu = \beta f + \frac{\beta p}{\rho}
\]

(14)

The conditions for coexistence between two phases are equal temperature, equal pressure, and equal chemical potential between each phase, respectively.

IV. RESULTS

Figure 3 shows the simulation results and the perturbation calculation for \( T=293.15K \), with the protein density \( \rho \) plotted versus the PEG concentration \( c_{\text{PEG}} \). We used GEMC to simulate the liquid-liquid coexistence at \( T = 293.15K \). We fit the simulation data to the standard form 

\[
\rho \pm -\rho^* = A|c_{\text{PEG}} - c_{\text{PEG}}^*| \pm B|c_{\text{PEG}} - c_{\text{PEG}}^*|^\beta
\]

where \( \rho^* \) and \( c_{\text{PEG}}^* \) are the critical protein density and PEG concentration, respectively, and \( \beta = 0.3258 \) is the Ising exponent. The values \( A, B, \rho^* \), and \( c_{\text{PEG}}^* \) were free parameters for the fit. The resulting fit parameters were \( A = 0.04(7), B = 0.40(1), \rho^* = 0.4(1), \) and \( c_{\text{PEG}}^* = 4.150(0) \). We also used perturbation theory to find both liquid-liquid and liquid-solid phase coexistence. The two methods give quantitatively similar results for the liquid-liquid curve. For low density side of the liquid-liquid curve, the lines converge to zero as \( c_{\text{PEG}} \) increases.

To test the accuracy of our model we compare our results with the experimental results of Vivarès et al for the phase boundaries for urate oxidase in PEG at \( T = 293.15K \). We note that in figure 3 our perturbation results agree with the experimental results showing the liquidus curve below the liquid-liquid curve. In figure 4 we show our liquid-liquid results from GEMC and perturbation theory along with the experimental values. Our results from
both methods are below the values from experiment. Figure 5 shows our results for the liquidus line from perturbation theory along with the liquidus and liquid-liquid data from the experiment. Here our liquidus line is below the measured liquid phase boundary for PEG concentration greater than 4% PEG. In both figures we have excluded measured data points above 12% PEG. Above 12% PEG the liquidus and liquid-liquid data from experiment are indistinguishable from one another and have protein concentration \( \rho \approx 0 \).

V. CONCLUSION

We have calculated the phase coexistence surface for a model of urate oxidase, as a function of the PEG concentration and temperature. To our knowledge this is the first such calculation for urate oxidase. We note that the parameters we use in our model for urate oxidase in PEG 8000Da (table I) give \( q = 0.8 \). Previous work using different models for the direct potential between the colloidal particles and experimental observations of different systems has suggested that for \( q \gtrsim 0.3 \) the liquid-liquid phase separation should be stable\(^\text{12,13} \). Perturbation results for our model and parameters show that the liquid-liquid phase separation is metastable and that the liquid-liquid phase line falls between the stable liquid-solid phase boundaries. These results are in qualitative agreement with experiments for urate oxidase in PEG 8000Da. Both perturbation theory and simulation results for the liquid-liquid phase separation curve are in qualitative agreement with the experimental phase diagram but fall below the experimental data.

There are several reasons one can suggest for the quantitative differences. The first deal with the depletion potential. For low PEG concentrations the Asakura-Oosawa potential used in the model is more strongly attractive than the measured depletion potential for urate oxidase at low PEG concentrations\(^\text{20} \). If the depletion were modeled using a weaker interaction, the phase separation would begin to appear at higher PEG concentrations than for the AO model. This will increase the critical values for \( c_{PEG}^* \). Also, the low and high
density branches would remain closer to the critical protein concentrations $\rho^*$. As a result, a model for depletion that is weaker than AO would give liquid-liquid curves in figures 3 and 4 positioned to the right (higher PEG concentrations) of our current results obtained using AO for depletion. Another possible reason for the quantitative differences is that we have not accounted for the non-hard-sphere behavior of the polymers. Accounting for the non-hard-sphere behavior of the polymers above their critical concentration using a semi-dilute AO model could improve the model results. Another possible reason for the quantitative differences is that there might be an as yet unknown interaction between PEG and the urate oxidase molecules not accounted for by simple depletion models. This has been proposed recently\textsuperscript{18} as a possible reason for the failure of the depletion models to describe the effect of PEG on lysozyme.

Another consideration is that the parameters used in our description of the DLVO potential might be inaccurate. These interactions were obtained by a use of the hypernetted chain approximation to obtain agreement with the measured scattering intensity for urate oxidase in solution without PEG\textsuperscript{20}. It is possible that the approximations involved could lead to an inaccurate potential, particularly in the case of higher PEG concentrations.

VI. ACKNOWLEDGMENTS

The authors wish to thank Denis Vivarès for his helpful comments. This work was supported by a grant from the G Harold and Leila Y. Mathers foundation.

\begin{thebibliography}{99}

\bibitem{1} A. McPherson.  \textit{Crystallization of Biological Macromolecules} (Cold Spring Harbor Laboratory Press, 1999).

\bibitem{2} J. D. Gunton, A. Shiryayev, and D. L. Pagan.  \textit{Protein Condensation: Kinetic Pathways to}

<table>
<thead>
<tr>
<th>symbol</th>
<th>value</th>
<th>name</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\sigma$</td>
<td>7.0 nm</td>
<td>protein diameter</td>
</tr>
<tr>
<td>$Z_{\text{eff}}$</td>
<td>-4</td>
<td>protein charge</td>
</tr>
<tr>
<td>$\lambda_D$</td>
<td>18.98 Å</td>
<td>Debye length</td>
</tr>
<tr>
<td>$\epsilon_s$</td>
<td>$80\epsilon_0$</td>
<td>$H_2O$ permittivity</td>
</tr>
<tr>
<td>$J_{\text{vdW}}$</td>
<td>2.5</td>
<td>van der Waals depth</td>
</tr>
<tr>
<td>$d_{\text{vdW}}$</td>
<td>3 Å</td>
<td>van der Waals length</td>
</tr>
<tr>
<td>$r_g$</td>
<td>2.8 nm</td>
<td>PEG 8000Da diameter</td>
</tr>
</tbody>
</table>

**TABLE II: Hard Sphere Solid Free Energy Coefficients**

<table>
<thead>
<tr>
<th>$S_0$</th>
<th>$S_1$ (from $C_0 - D$)</th>
<th>$S_2$ ($\frac{(C_1 - C_0)}{2}$)</th>
<th>$C_1$</th>
<th>$D$</th>
</tr>
</thead>
<tbody>
<tr>
<td>-0.24</td>
<td>$C_0 - D$</td>
<td>$(C_1 - C_0)/2$</td>
<td>2.56</td>
<td>0.56</td>
</tr>
</tbody>
</table>
List of Figures

1. Shape of the model potential for T=293.15K for four PEG concentrations. From top to bottom, the concentrations are 0%, 3%, 5%, and 7% PEG in solution. ................................................................. 15

2. Second virial coefficient $B_2/B_2^{HS}$ for the model at T=293.15K, investigated. ........................ 16

3. Comparison of perturbation calculations (lines) to the GEMC results (diamonds) for T=293.15K. Solid lines show the liquidus (lower line) and solid (upper line) phase boundaries and dotted lines show the liquid-liquid phases. 17

4. Comparison of liquid-liquid curves at T=293.15K for urate oxidase from Monte Carlo simulation (diamonds) and perturbation calculation (dashed line) with experimental results (open circles). ........................................... 18

5. Comparison of liquidus curves at T=293.15K for urate oxidase from perturbation calculation (solid line) and experimental results (filled circles). Also shown are liquid-liquid results from experiment (open circles). ............. 19
FIG. 1: Shape of the model potential for $T=293.15\,K$ for four PEG concentrations. From top to bottom, the concentrations are 0%, 3%, 5%, and 7% PEG in solution.
FIG. 2: Second virial coefficient $B_2 / B_2^{HS}$ for the model at $T=293.15\text{K}$, investigated.
FIG. 3: Comparison of perturbation calculations (lines) to the GEMC results (diamonds) for T=293.15K. Solid lines show the liquidus (lower line) and solid (upper line) phase boundaries and dotted lines show the liquid-liquid phases.
FIG. 4: Comparison of liquid-liquid curves at $T=293.15\text{K}$ for urate oxidase from Monte Carlo simulation (diamonds) and perturbation calculation (dashed line) with experimental results (open circles).
FIG. 5: Comparison of liquidus curves at T=293.15K for urate oxidase from perturbation calculation (solid line) and experimental results (filled circles). Also shown are liquid-liquid results from experiment (open circles).