Nucleation of the crystalline phase of proteins in the presence of semidilute nonadsorbing polymer

Richard P. Sear
Department of Physics, University of Surrey, Guildford, Surrey GU2 7XH, United Kingdom

(Received 23 February 2001; accepted 6 April 2001)

Starting from a protein solution which is metastable with respect to the crystalline phase, the effect of adding semidilute nonadsorbing polymer is considered. It is found to increase the chemical potential of the protein by a few tenths of kT, which may be enough to lower the barrier to nucleation of the crystalline phase by enough to allow crystallization. It is also shown that assuming that the polymer induces a pairwise additive attraction leads to qualitatively incorrect results.


I. INTRODUCTION

Crystallizing proteins from dilute solution is in general rather difficult, yet the crystalline form is required in order to perform x-ray crystallography. X-ray crystallography is currently the best method of determining the all-important structure of a protein in its native state. This has led to considerable interest in how proteins crystallize; see Ref. 3 for a recent review from a colloidal physics perspective and Refs. 4–8 for recent work. In a dilute solution there may be a large attraction which is sufficiently strong to cause the equilibrium behavior to be coexistence between a dense crystal and a dilute fluid phase. We do not need to specify the potential explicitly as we will require only the fluid-crystal interfacial tension $\gamma_0$ and the difference in chemical potential between that of the crystal and that of the fluid, $\Delta \mu_0$. We use the subscript 0 to indicate quantities before the polymer is added. Our results should therefore apply to a very wide range of proteins despite potentially very large variations in the structure and interactions of different proteins. The polymer is nonadsorbing: the interaction between the protein molecule and a monomer is purely repulsive.

There has been extensive work done on systems of small colloidal particles immersed in polymer solutions of larger polymer coils which do not adsorb on the particles. Scaling theory for these systems was pioneered by de Gennes; see also Refs. 12–15. Field-theoretic methods have yielded results which are exact for small hard particles and ideal polymers, as well as approximate results for polymers in good solvents. Work has also been done using integral equations; see Ref. 19 and references therein. It is well understood that nonadsorbing polymer increases the chemical potential of particles and induces an attraction between these particles which is not pairwise additive: the interaction of three close-by protein molecules is not the sum of three pair potentials. However, most of this work (exceptions are Refs. 20,21) only considered a few particles, two or three, whereas a nucleus of a crystalline phase contains tens of tightly packed particles. Because of this, the deviations from pairwise additivity observed here are much stronger than found in previous work.

Section II briefly outlines classical nucleation theory for the solution before the polymer is added. Section III then describes the effect of the polymer and Sec. IV is a conclusion.

II. CLASSICAL NUCLEATION THEORY

The interaction between protein molecules is taken to include a steeply repulsive core and an attraction. This attraction may be isotropic or highly anisotropic or a mixture of the two, if isotropic it may or may not be short ranged; all we require is that it render a dilute solution of the protein metastable with respect to a crystalline phase of the protein. By metastable we mean that the fluid phase is not the equilibrium phase, it does not correspond to the absolute minimum in the appropriate free energy, but the fluid phase is dynamically stable with respect to crystallization for times...
much larger than the characteristic time scale of the solution. See Ref. 23 for a general introduction to the properties and behavior of metastable fluids.

Classical nucleation theory,23–25 assumes that the nucleus of the crystalline phase has a free energy $\Delta F$ which is the sum of a bulk term and surface term. See the book of Debye,

23 for an excellent introduction to classical nucleation theory. The protein molecules are modeled by spherical particles of diameter $D$. The bulk term is equal to the number of molecules in the nucleus, $n$, times the chemical potential difference $\Delta \mu = \mu_c - \mu$, where $\mu_c$ is the chemical potential of the crystalline phase and $\mu$ is the chemical potential of the fluid phase which contains the nucleus. The fluid phase we are considering is not the true equilibrium phase, nucleation occurs from a fluid which is at a higher chemical potential than the crystal. Thus $\Delta \mu$ is negative. The surface term is the surface area of the nucleus times the surface tension $\gamma$ of the bulk interface between the coexisting crystalline and fluid phases. The surface area of the nucleus is obtained by assuming that the nucleus is a sphere of radius $R$ which is related to $n$ by assuming that the density of spheres within the nucleus is equal to the bulk density of the crystalline phase, taken to be $D^{-3}$: the density of a close-packed simple cubic lattice. Then $n = (4/3) \pi R^3 D^{-3}$. Putting this all together,

$$\Delta F = \frac{1}{3} \pi R^3 D^{-3} \Delta \mu + 4 \pi R^3 \gamma.$$  
(1)

The first term in $\Delta F$ is the bulk term, which is negative and decreases linearly with $n$, and the second term is the surface term which is positive and increases as $n^2$. Thus $\Delta F$ passes through a maximum, denoted by $\Delta F^*$, at some value of $R$, denoted by $R^*$. This maximum is the top of the free energy barrier to nucleation. The nucleus with radius $R^*$ and $n^*$ spheres is called the critical nucleus. Taking the derivative of Eq. (1) with respect to $R$ and equating to zero,

$$R^* = \frac{2 \gamma D^3}{\Delta \mu},$$  
(2)

and inserting this value of $n$ in Eq. (1),

$$\Delta F^* = \frac{16 \pi \gamma^3 D^6}{3 \Delta \mu^2}.$$  
(3)

Nucleation is an activated process with $\Delta F^*$ as the barrier height and so the nucleation rate varies as $\exp(-\Delta F^*/kT)$, where $k$ and $T$ are Boltzmann’s constant and the temperature $T$, respectively. To increase the rate of nucleation of the crystalline phase of the protein we need to reduce the barrier $\Delta F^*$. In the absence of a polymer $\Delta F^*$ and $R^*$ are obtained from Eqs. (2) and (3) by setting $\Delta \mu = \Delta \mu_0$ and $\gamma = \gamma_0$.

### III. NUCLEATION IN THE PRESENCE OF A SEMIDILUTE POLYMER

Now we consider a solution of both protein molecules and flexible polymer molecules. The interaction between a protein molecule and a monomer is taken to be strictly repulsive. Figure 1 is a schematic of the nucleus in a solution of the polymer. We remark that this simple model is not adequate for all polymers, e.g., it is not adequate for PEG [poly(ethylene glycol)]. We consider the case where the polymer molecules are larger than the protein molecules and the polymer solution is semidilute. As the protein molecules are only a few nanometers in diameter this regime is easily accessible in experiment. In the other limit, where the polymer’s radius is smaller than that of the colloidal particles, the polymer induces a short-ranged depletion attraction between the particles. Such attractions are to a good approximation pairwise additive and therefore straightforward to treat. See Refs. 22,26–29 for work where the polymer is at most as large as the colloid.

See the book of de Gennes30 for an introduction to semidilute polymer solutions. They are characterized by a correlation length $\xi$ which scales as $c_M^{-1/4}$, where $c_M$ is the monomer concentration. The correlation length $\xi$ is essentially the distance between collisions between one polymer chain and another. Thus there are $\xi^{-3}$ interchain collisions per unit volume which leads to an osmotic pressure exerted by the polymer $\Pi = kT \xi^{-3}$.

The crystalline phase of the protein is dense, the gaps between protein molecules available to polymer are around $0.4D$ or less across. Forcing a polymer molecule into a tube of diameter $0.4D$ costs roughly $kT$ in free energy per $0.4D$ of the tube’s length as the polymer collides with the walls of the tube at intervals of roughly its diameter. Thus, the free energy density inside such a crystal, either a bulk crystal or a crystalline nucleus, will be of order $kT$ per $0.06D^3$ volume31 whereas in solution it is $kT$ per $\xi^3$ volume. For $\xi$ larger than $D$, which we assume throughout, the free energy density inside a crystal is so much higher than outside that to a good approximation there is no polymer inside the crystal. This agrees with calculations using the theory of Ref. 20. So, as the crystalline nucleus is impermeable to polymer it interacts with the polymer, as a hard particle of radius $R$.

The interaction $w$ of a hard spherical particle of diameter $D$ with a semidilute solution of polymer in a good solvent with a correlation length $\xi$ is

$$w \sim kT \left( \frac{D}{\xi} \right)^{4/3}, \quad D \ll \xi,$$  
(4)
where we took the exponent $\nu$ to have its Flory value of $3/5$. The exponent $\nu$ is defined by $R_g \sim aN^\nu$, which relates the radius of gyration $R_g$ of a single isolated polymer chain in a good solvent to its number $N$ of monomers, each of which is of length $a$. It is the work done ($= \text{the difference in excess chemical potential}$) in taking a particle from a pure solvent and inserting it into the polymer solution. This assumes that the interaction between a protein and the monomers of a polymer is purely repulsive and that there is not much more than one protein molecule per $\xi^3$ volume. The notation $\sim$ indicates that we have neglected a coefficient of order unity; here we derive only the scaling behavior of the interactions with respect to the relevant length scales. Equation (4) can be derived in a couple of ways (see Refs. 11–14). One way is to note that when $D\ll \xi$ the interaction between a particle and polymer must scale as the density of monomers, which scales as $\xi^{-4/5}$ when $\nu=3/5$. As $w/kT$ is dimensionless, we require that $\xi^{-4/5}$ appears as the ratio $(D/\xi)^{4/5}$ which gives Eq. (4).

Straightaway we can derive an approximation for the contribution of polymer to the difference in excess chemical potential $\Delta \mu$ between the dense crystalline phase of the protein and the semidilute polymer solution. It is

$$\Delta \mu \sim \Delta \mu_0 - kT \left( \frac{D}{\xi} \right)^{4/3} + \Pi D^3,$$

where the second term on the right-hand side is the change in free energy of the polymer when a single protein molecule is removed from the semidilute polymer solution and the last term is the change in free energy of this solution when it is compressed by the crystalline phase of the protein expanding by the volume of one protein. The osmotic pressure of a polymer solution $\Pi$ is related to $\xi$ by $\Pi \sim kT \xi^{-3}$. Thus, the last term is equal to $kT(D/\xi)^3$ and so is smaller than the second term, indeed it is of order terms we have dropped and so we drop it and obtain our final expression for $\Delta \mu$ in the presence of a polymer,

$$\Delta \mu \sim \Delta \mu_0 - kT \left( \frac{D}{\xi} \right)^{4/3}.$$

(6)

For a large nucleus, radius $R \gg \xi$, the width of the nucleus–polymer-solution interface, $\xi$, is small and so the capillarity approximation which underlies classical nucleation theory still holds: $\Delta F$ is still a sum of bulk and interfacial terms. The bulk term is given by the $\Delta \mu$ of Eq. (6) times the number of protein molecules in the nucleus. In the presence of semidilute polymer the interfacial tension of a flat interface, $\gamma$, is $^{12}$

$$\gamma \sim \gamma_0 + \frac{kT}{\xi^2},$$

(7)

where $\gamma_0$ is the interface tension in the absence of a polymer. The contribution of the polymer comes from a layer of thickness $\xi$ on top of the surface within which the polymer concentration is depleted. The free energy cost of this depleted layer is $\sim I \xi$, which leads to the second term of Eq. (7). Bearing in mind that $\gamma_0$ will be at least of order $kT/D^2$ we see that the fractional modification to $\gamma_0$ due to polymer is of order $(D/\xi)^2$. For a $\Delta \mu_0$ of order $kT$, the fractional change to $\Delta \mu$ due to polymer is, see Eq. (6), of order $(D/\xi)^{4/3}$. Within our crude scaling theory we only consider leading order contributions and so neglect the contribution of the polymer to $\gamma$, leaving the only contribution of the polymer as the change in $\Delta \mu$ given by Eq. (6).

In the opposite limit for $R$, $R \ll \xi$, then the contribution of the polymer to the free energy of the nucleus is no longer the sum of bulk and interfacial terms, i.e., it is not the sum of terms scaling as $R^2$ and as $R^2$. For $R \ll \xi$ the contribution of the polymer to the free energy of formation of the nucleus $\Delta F^*$ is the sum of two terms: the first is the number of protein molecules in the nucleus times minus the increase in chemical potential in solution due to polymer, Eq. (4), and the second is the free energy cost of inserting the nucleus into the polymer solution. This second term is the free energy cost due to the nucleus excluding polymer from a sphere of radius $R$. Because the polymer cannot penetrate into the nucleus it interacts with the polymer as a single particle, and so this free energy cost is just Eq. (4) with $D$ replaced by $2R$. Adding the two terms together we have the contribution of polymer to $\Delta F$,

$$\Delta F = \Delta F_0 - nkT \left( \frac{D}{\xi} \right)^{4/3} + kT \left( \frac{2R}{\xi} \right)^{4/3}, \quad R \ll \xi,$$

(8)

where $\Delta F_0$ is the free energy of formation of the nucleus before polymer is added. For $R$ a few times $D$ and $R,D \ll \xi$ the second term in Eq. (8) dominates the last term. So, we neglect the last term in Eq. (8). Then as the remaining term is linear in $n$ it is in effect a shift in $\Delta \mu$ and indeed is the same shift in $\Delta \mu$ we found in the opposite limit, $R \gg \xi$. Thus we conclude that as $R \gg \xi$ the leading order contribution of a semidilute polymer to the nucleation barrier is simply to shift $\Delta \mu$ by the amount given in Eq. (6).

So, we have shown that in both limits, $R \gg \xi$ and $R \ll \xi$, the effect of adding a polymer in the semidilute regime is to shift the chemical potential difference $\Delta \mu$ by an amount of order $-(D/\xi)^{4/3}kT$, which will be a few tenths of $kT$. Having shown that it holds in both limits we assume that it also holds for $R/\xi = O(1)$ as well. The shift varies as $\xi^{-4/5}$ and so increases linearly with the density of polymer. The size of the critical nucleus also decreases, Eq. (2). Thus, if decreasing $\Delta \mu$ by a few tenths of $kT$ is enough to reduce the nucleation barrier sufficiently to make the nucleation rate significant, then adding a semidilute polymer is a viable way of inducing nucleation of the crystalline phase of small colloidal particles such as proteins and micelles.

A. Comparison with assumption of pairwise additivity

A pair of particles in a semidilute polymer solution separated by a distance $x$ of order $\xi$ or less feel a polymer-induced attraction toward each other, which we denote by $w_2$. $^{14,17,18}$ For small particles, $D \ll \xi$, and for separations $r$ much smaller than the correlation length, $^{15,17}$

$$w_2 \sim -kT \left( \frac{D}{\xi} \right)^{4/3} \left( \frac{D}{r} \right)^{4/3}, \quad r \ll \xi.$$

(9)

See the above-mentioned references for a derivation of this equation but in short: the strength of the potential $(D/\xi)^{4/3}$ is
the polymer’s contribution to the chemical potential, the scaling \((D/r)^{4/3}\) comes from the decay of correlations in a semidilute solution and the range is \(\xi\) as beyond this the two particles are essentially independent of each other. Naively it might be thought that the contribution of the polymer to the free energy change \(\Delta \mu\) could be estimated using this pair attraction plus a mean-field approximation, which would correspond to the approximation,\(^{33}\)

\[
\Delta \mu \sim \Delta \mu_0 + D^{-3} \int_{r > D} dw_2(r) \quad \text{wrong},
\]

which is just the integral over the number density, \(D^{-3}\) in the crystalline phase, times the potential. As usual with a mean-field approximation, correlations in the positions of particles are neglected and one protein is considered to interact with surrounding particles which are at the mean density. The above-mentioned integral is \(D^{-3}\) times the contribution of the polymer to the second virial coefficient. See Eisenriegler\(^{18}\) for a more accurate treatment of the contribution of polymer to the second virial coefficient. So, inserting Eq. (9), into Eq. (10), we obtain

\[
\Delta \mu \sim \Delta \mu_0 - kT \left(\frac{\xi}{D}\right)^{1/3} \quad \text{wrong}.
\]

If we compare Eqs. (6) and (11) we see that the assumption of a pairwise additive potential plus the mean-field approximation predicts the wrong scaling with \((D/\xi)\) of the polymer contribution to \(\Delta \mu\). It greatly overestimates, by a factor of order \((\xi/D)^{5/3}\), the effect of adding polymer whose correlation length is large with respect to the protein diameter. Naively using the pair potential between an isolated pair of particles (a pair of particles with no others within a distance \(\xi\) of them), as the pair interaction within a dense phase, number density \(\gg \xi^{-3}\) (here the crystal), is qualitatively wrong.

**IV. CONCLUSION**

We have estimated the change in the barrier to nucleation when the polymer is added to a metastable dilute protein solution. The polymer is semidilute, and has a correlation length \(\xi\) larger than the diameter of the protein \(D\); the polymer does not adsorb onto the surface of the protein. When the polymer is added, the barrier \(\Delta F^*\) is reduced due to the increase in the chemical potential of the protein in solution, possibly by enough to allow nucleation. However, the effect is not large, \(\Delta \mu\) decreases by only a few tenths of \(kT\). Also, a naive use of the polymer induced attraction between a pair of proteins yields results which are qualitatively wrong: They grossly overestimate the effect of adding the polymer. This does not mean that an effective potential approach to the effect of the polymer cannot reliably estimate the effect of large polymer molecules, just that this effective potential is many-body in nature and cannot simply be approximated by just a pair potential. This conclusion also applies to the phase behavior of a system of hard-sphere-like colloidal particles and larger polymer molecules.\(^{20}\)

Finally, we note that the second virial coefficient \(B_2\) is often used to measure the strength of the attractions in a protein solution. \(B_2\) is given by

\[
B_2 = B_{2,0} + \frac{1}{2} \int_{r > D} \frac{w_2(r)}{kT}, \quad w_2 \ll kT
\]

\[
\sim B_{2,0} - D^{8/3} \xi^{1/3},
\]

where \(B_{2,0}\) is the second virial coefficient in the absence of the polymer and we have assumed not only that \(w_2/kT\) is small but that \(B_{2,0}\) is dominated by interactions with a range \(\ll \xi\); then the second virial coefficient is the sum of the two terms in Eq. (12). So as the polymer density increases, \(\xi\) decreases and the contribution of polymer to \(B_2\) decreases.

But from Eq. (6) we see that as the polymer density increases it has a larger and larger effect on the nucleation barrier. So, the effect of polymer on \(B_2\) is **anticorrelated** with its effect on \(\Delta F^*\), i.e., when one goes up the other goes down, not correlated as we might naively have expected.

---

9. R. P. Sear, cond-mat/9912199.
10. As emphasized by Evans and co-workers, see, e.g., Refs. 26 and 27, the interaction is only strictly pairwise additive if the polymer is both ideal and does not interact with a particle more than \(\sim 0.15\) times its radius away. However, if the polymer is larger than this but still small with respect to the protein molecule, the potential will be close to pairwise additive. Then the simultaneous interaction of a polymer molecule with more than two protein molecules (which is the origin of deviations from pairwise additivity) will be weak. In the other limit, which we consider here, where the polymer is much larger than the protein, the polymer molecule can potentially interact with many protein molecules, inducing an effective interaction which is very far from being pairwise additive.
29 A. A. Louis, cond-mat/0102220.
31 See pp. 50–51 of Ref. 30.
33 This is a widely used approximation for long-range, strictly pairwise additive attractions, see pp. 148–155 especially Eq. (6.2.15) of J. P. Hansen, and I. R. McDonald, Theory of Simple Liquids, 2nd ed. (Academic, London, 1986).