Crowding induced entropy-enthalpy compensation in protein association equilibria

Young C. Kim\textsuperscript{1,} and Jeetain Mittal\textsuperscript{2,}†

\textsuperscript{1}Center for Computational Materials Science, Naval Research Laboratory, Washington DC 20375, USA
\textsuperscript{2}Department of Chemical Engineering, Lehigh University, Bethlehem, PA 18015

A statistical mechanical theory is presented to predict the effects of macromolecular crowding on protein association equilibria, accounting for both excluded volume and attractive interactions between proteins and crowding molecules. Predicted binding free energies are in excellent agreement with simulation data over a wide range of crowder sizes and packing fraction. It is shown that attractive interactions between proteins and crowding agents counteract the stabilizing effects of excluded volume interactions. A critical attraction strength, for which there is no net effect of crowding, is approximately independent of the crowder packing fraction.

PACS numbers:

Protein-protein interactions are important in many essential biological functions, such as transcription, translation, and signal transduction \cite{1}. A lot of progress has been made in understanding protein association in dilute solution via experiments and simulations \cite{2–5}. Cells, on the other hand, contain various macromolecules, e.g., DNA, RNA, proteins, organelles, etc., which constitute up to 40\% of the cell volume \cite{6}. It is thus crucial to relate \textit{in vitro} experimental or simulation results to those in a crowded cellular environment \cite{7–13}.

Several experimental studies have been performed to understand protein-protein interactions in a crowded environment \cite{14–24}. Most attention has been paid to the steric excluded volume effects of inert crowding agents on the formation of protein complexes \cite{25–27}. Very recent studies have also started to probe the effects of attractive interactions between proteins and crowders on protein association \cite{28–30, 32}. These studies have highlighted the importance of accounting for enthalpic effects arising from attractive interactions in addition to commonly invoked excluded volume effects. It was found that the enthalpic effects can actually increase the binding free energy (thereby destabilizing the bound complex) in contrast to predictions based on available theoretical models that can only capture entropic effects.

Most theoretical models of crowding are based on scaled particle theory (SPT) of hard-sphere fluids \cite{33} or its modified versions and have been applied to interpret experimental and computational results with varying success. The failure of these models in several situations highlights an important role played by attractive crowder-protein interactions. In our earlier work \cite{30}, we had proposed an \textit{ad hoc} mean-field expression to fit our simulation data to provide some insight into the role of attractive crowder-protein interactions in destabilizing protein association. However, there is a need for comprehensive quantitative theory, that can describe the effects of repulsive as well as attractive crowder-protein interactions on the protein-association equilibria.

In this paper, we present a theory that can quantitatively predict the effects of macromolecular crowding on the protein association equilibria accounting for both repulsive and attractive crowder-protein interactions. The statistical mechanics and thermodynamics of a hard-sphere fluid are adapted to yield an approximate analytical expression for the protein-binding free energy in the presence of spherical crowders. Extensive replica exchange Monte Carlo (REMC) simulations have been performed on two distinct protein complexes to test this theory. We find that the theory is in excellent agreement with simulations over a wide range of crowder packing fractions and crowder-protein interactions. The theory identifies the region in parameter space (entropy-enthalpy compensation line in a two parameter plane) separating entropically stabilized area versus enthalpically destabilized one.

\textbf{Theoretical development.} Figure 1 illustrates a thermodynamic cycle that describes a change in the binding free energy $\Delta F_{\text{bind}}$ of two proteins due to the presence of crowding molecules. This change, $\Delta \Delta F_{\text{bind}}$, can be expressed as the difference in the binding free energy in the absence and presence of crowders and is given by

\begin{equation}
\Delta \Delta F_{\text{bind}}(\phi) = \Delta F_{\text{bind}}(\phi) - \Delta F_{\text{bind}}(\phi = 0) = \Delta F_{\text{AB}}^{\text{crowd}} - \Delta F_{A}^{\text{crowd}} - \Delta F_{B}^{\text{crowd}}, \quad (1)
\end{equation}
where $\Delta F_{\text{crowd}}(\phi), (\alpha \in \{A,B,AB\})$ is the solvation free energy of a protein (or complex) $\alpha$ in a crowded solution with crowding packing fraction $\phi$. [For brevity, we will omit the superscript “crowd” below.]

To obtain an expression for $\Delta F_\alpha(\phi)$ in Eq. (1) for a protein or complex $\alpha$, let $U_\alpha(r, \Omega) = \sum_{i \in \alpha} u_i(r_i)$ be the overall interaction between a protein $\alpha$ and a crowder, where $r$ is the distance between the center of mass of the protein and the crowder and $\Omega$ the orientational degree of freedom, while $u_i$ is the interaction between an atom (or residue) $i$ of the protein $\alpha$ and the crowder. For a general Lennard-Jones(LJ)-type potential for $u_i$, it is reasonable to assume that for given $\Omega$, $U_\alpha(r, \Omega)$ exhibits a minimum, $-\epsilon_\alpha^m(\Omega), r = r_\alpha^m(\Omega)$. Following the Weeks-Chandler-Andersen (WCA) theory, we then decompose $U_\alpha$ into the repulsive and attractive parts as

$$U_{\alpha,\text{rep}}(r, \Omega) = \begin{cases} U_\alpha(r, \Omega) + \epsilon_\alpha^m(\Omega) & r < r_\alpha^m(\Omega), \\ 0 & \text{otherwise}, \end{cases}$$

$$U_{\alpha,\text{att}}(r, \Omega) = \begin{cases} -\epsilon_\alpha^m(\Omega) & r < r_\alpha^m(\Omega), \\ U_\alpha(r, \Omega) & \text{otherwise}. \end{cases} \tag{2}$$

The solvation free energy, $\Delta F_\alpha(\phi)$, of the protein in a crowded solution can then be divided into two parts as,

$$\Delta F_\alpha(\phi) = \Delta F_{\alpha,\text{rep}}(\phi) + \Delta F_{\alpha,\text{att}}(\phi), \tag{3}$$

where $\Delta F_{\alpha,\text{rep(att)}}$ is the contribution from the repulsive (attractive) interaction, respectively.

The repulsive contribution, $\Delta F_{\alpha,\text{rep}}$, is obtained by adopting the SPT. The SPT theory provides the free energy for solvating a hard-sphere of radius $R_\alpha$ in a bath of hard-sphere particles of radius $R_c$ as,

$$\Delta F_{\alpha,\text{rep}} = (3y + 3y^2 + y^3)\tilde{\phi} + (4.5y^2 + 3y^3)\tilde{\phi}^2 + 3y^3\tilde{\phi}^3 - \ln(1 - \phi), \tag{4}$$

where $\tilde{\phi} = \phi/(1 - \phi)$ and $y = R_\alpha/R_c$. But can we represent an anisometric protein with soft-core protein-crowder interactions as a hard sphere with an appropriate radius $R_\alpha$ to capture protein’s solvation behavior accurately? Here we use the Boltzmann criteria to define $R_\alpha$ as,

$$\frac{4\pi}{3}(R_\alpha + R_c)^3 = \int_{U_{\alpha,\text{rep}} = f_{\text{kT}}} r^2 dr d\Omega, \tag{5}$$

where the right-hand side represents the volume encompassed by the condition $U_{\alpha,\text{rep}}(r, \Omega) \geq f_{\text{kT}}$. Here, we use $f = 2$ that has been used successfully in previous studies [34].

Using thermodynamic perturbation theory approach, the attractive contribution, $\Delta F_{\alpha,\text{att}}$, can be expressed as (up to the first order),

$$\Delta F_{\alpha,\text{att}} \approx (U_{\alpha,\text{att}})_{\text{rep}} = \int \rho U_{\alpha,\text{att}}(r, \Omega) g_0(r) r^2 dr d\Omega, \tag{6}$$

where $\rho$ is the crowder number density related to $\phi$ via $\rho = \phi/(4\pi R_c^3)/3$, and $g_0(r)$ is the radial distribution function of the hard-sphere crowders between a protein and a crowder. Realizing that $g_0(r)$ has a maximum $g_0^{\text{max}}$ at contact and then decays rapidly to unity, we assume $g_0(r) = g_0^{\text{max}}$ for $r \in [r_\alpha^m, r_\alpha^m + \lambda]$ and $1$ for $r \in [r_\alpha^m + \lambda, \infty)$ with $\lambda = (2^{1/6} - 1)R_c \approx 0.12R_c$ [35]. We then approximate $\int_{\Omega}$ as,

$$\Delta F_{\alpha,\text{att}} \approx \rho \epsilon_\alpha S_\alpha \{\delta r + (g_0^{\text{max}} - 1)\lambda\}, \tag{7}$$

where $\epsilon_\alpha = \langle \epsilon_\alpha^m(\Omega) \rangle$ is the orientational average of $\epsilon_\alpha^m$, $S_\alpha = \int [r_\alpha^m(\Omega)]^2 d\Omega$ the surface area around the protein, and $\delta r$ the attraction range. Note that here we assume $\delta r \geq \lambda$.

To enhance the simplicity and practical value of our theory, we use the Carnahan-Starling (CS) equation of state for a hard sphere fluid to calculate $g_0^{\text{max}}$. The CS equation of state is known to reproduce the thermodynamic behavior of hard-sphere fluids from dilute gas to near the freezing transition. The CS expression for $g_0^{\text{max}}$ is given by

$$g_0^{\text{max}} = g_{\text{CS}}(\phi) = (1 - \phi/2)/(1 - \phi)^3, \tag{8}$$

and only depends on $\phi$. Note that the first term in Eq. (7) gives a linear order in $\phi$ while the term containing $g_0^{\text{max}}$ yields higher order terms. Combining together Eqs. (1), (3), (4), (7) and (8), one can easily obtain an estimate of crowding induced change in the binding free energy.

Next, we test this theory against REMC simulations of two protein complexes in a wide range of crowder sizes, packing fractions and interaction strengths.

**Model and simulation details.** A residue-based coarse-grained model is used to simulate protein-protein interactions [35]. This transferable protein-protein interaction model was shown to yield binding affinities and structures for moderate-to-weakly interacting protein complexes in accord with experiments [35] [37]. Crowding agents are represented by spheres interacting via a repulsive potential, $u_{\text{rep}}(r) = \epsilon_r \left( \frac{\sigma_r}{r - 2\sigma_r + \sigma_r} \right)^{12}$, where $\sigma_r$ is the interaction range set equal to 6Å. As our protein-protein interaction model only includes solvent (water) effects indirectly by accounting for it in the amino acid pair contact potentials, repulsive crowder-crowder interactions essentially mean that crowder-solvent interactions are assumed to be much stronger (to keep crowders dispersed in solution). See Supplemental Material (SM) at [URL will be inserted by publisher] for more details on models and simulation.

**Results and Discussion.** The spherical crowders interact with each other via the distance-dependent soft repulsive potential given by $u_{\text{rep}}(r)$ with a characteristic size $r_c$ and $\epsilon_r = 1.69k_BT$. To apply the SPT theory (4) for calculating the repulsive contribution of the binding free energy, it is necessary to obtain an effective hard-sphere radius for such crowders. We define the
FIG. 2: Plot of the overall interaction between the Ubq/UIM1 complex and a crowder, $U_{Ubq/UIM1}$, as a function of $r$ for different orientations $\Omega$. The solid curves are obtained from SM Eq. (1) with $\epsilon_c = \epsilon_{\text{min}}(\Omega)$ and $\sigma_c = r_0(\Omega)$.

effective hard-sphere radius, $R_c$, of crowders by the condition, $u_{\text{rep}}(2R_c) = f k_B T$ with the same $f$ as in Eq. (5). This yields $R_c = r_c + \gamma \sigma_r$ where $\gamma = \frac{1}{2} \left[ \left( \frac{1.69}{2} \right)^{1/2} - 1 \right]$. Note that although for $\epsilon_c = 1.69 k_B T$ one has $R_c \approx r_c$, in general, $R_c$ can be different from $r_c$. The effective packing fraction $\phi$ is then given by $\phi = \phi_0 (R_c/r_c)^3$.

Figure 2 presents the overall interaction between the complex Ubq/UIM1 and a crowder at five different orientations, illustrating a highly anisotropic and asymmetric nature of the interaction. It shows that the overall protein-crowder interaction follows the LJ shape of the residue-crowder interaction, SM Eq. 1, (see the solid curves), with a minimum $-\epsilon_m(\Omega)$ at $r = r_m(\Omega)$ for a given $\Omega$. However, the longer-distance tails are underestimated by the same formula as evident in the inset.

The effective radius, $R_\alpha$, for a protein $\alpha$, determined by Eq. (5) depends weakly on $r_\alpha$ and $\epsilon_c$ (see SM) as shown in Table I. For the repulsive protein-crowder interactions, such effective radii for proteins and complexes are sufficient enough to calculate the change in the binding free energy, $\Delta \Delta F_{\text{bind}}$, via Eq. (4). Figure 3 shows an excellent agreement between simulation results (black squares) and the theory (black solid curves) for the Ubq/UIM1 complex for different crowder sizes. As previously reported by us and others, the binding free energy decreases with increasing packing fraction $\phi$ and decreasing crowder size due to the excluded-volume effect.

Recent studies [28–30] have shown that attractive protein-crowder interactions can destabilize protein association. Figure 3 shows that indeed as the attraction strength, $\epsilon_c$, between a residue and a crowder increases the binding free energy also increases with the packing fraction $\phi$. For example, for a moderate strength $\epsilon_c = 0.6 k_B T$ the change in the binding free energy at $\phi = 0.3$ (close to the physiological condition) is up to about 4 $k_B T$ when the protein-crowder interaction switches from repulsive (black) to attractive (purple). For reference, hard sphere fluids undergo freezing transition at $\phi = 0.49$ and the random close packing is $\phi = 0.64$ [31]. After including the volume occupied by the proteins, it is clear that we are not simulating low crowder packing fractions for which linear expansion in $\phi$ can explain the observed trends. In order to apply our theory, Eqs. (1)–(5), to describe the simulation data for various $\epsilon_c$ and $r_c$, we calculate the average attraction strength, $\bar{\epsilon}_\alpha$, and the surface area, $S_\alpha$, for the individual proteins and the complex. Note that $\bar{\epsilon}_\alpha$ is proportional to $\epsilon_c$ while $S_\alpha$ is indepen-
FIG. 4: Binding free energy, $\Delta \Delta F_{\text{bind}}$, for the Cc/CcP complex as a function of $\phi$. Symbols and curves are same as in Fig. 3.

As previously suggested [29, 30], the enthalpic effects can be approximated to be proportional to protein’s surface area, $\alpha$ (close to $\sigma_r$) is used for all the crowder sizes and attraction strengths.

To check whether the theory can be transferable to other protein complexes, we calculate the binding free energies for the Cc/CcP complex (total 402 residues compared to 100 residues for the Ubq/UIM1 complex) as shown in Fig. 4. With the same $\delta r$, the theoretical predictions agree remarkably well with the simulation data.

The data in Figures 3 and 4 show the competition between entropic effects of the excluded volume and enthalpic effects by attractive crowder-protein interactions. As previously suggested [29, 30], the enthalpic effects can be approximated to be proportional to protein’s surface areas and our theory here provides its concrete foundation from microscopic nature of the protein-crowder interactions. At high attraction strengths, the enthalpic penalty for breaking the crowder-protein interactions (at the expense of protein-protein interactions) dominates, thus increasing the binding free energy. At some critical attraction $\epsilon_c^{\text{crit}}$, the two contributions are canceled out, and the binding energy in a crowded solution becomes equal to that in the absence of crowders (see green triangles and curve in Fig. 3b).

It was observed [30] that the critical attraction, $\epsilon_c^{\text{crit}}$, for which the effect of the excluded volume is canceled out exactly by that of the attractive contribution, (i.e., $\Delta \Delta F_{\text{bind}} = 0$), is approximately independent of the crowder packing fraction, $\phi$. This is owing to the fact that $\Delta \Delta F_{\text{bind}}$ is almost linear in $\phi$ for $\epsilon_c$ considered. To obtain $\epsilon_c^{\text{crit}}$ estimate, we combine Eqs. (4) and (7) and solve for $\epsilon_c$ that satisfies $\Delta \Delta F_{\text{bind}} = 0$ up to the linear order in $\phi$. One then obtains,

$$\epsilon_c^{\text{crit}} = \frac{\Delta Y}{\Delta W} + O(\phi),$$

where

$$\Delta Y = 3(y_A + y_B - y_{AB}) + 3(y_A^2 + y_B^2 - y_{AB}^2)$$
\[+ (y_A^3 + y_B^3 - y_{AB}^3) + 1, \quad (10)\]

$$\Delta W = 3(\epsilon_A S_A + \epsilon_B S_B - \epsilon_{AB} S_{AB}) \delta r/(4\pi R_c^2 \epsilon_c).$$

This yields $\epsilon_c^{\text{crit}}/k_B T \approx 0.19, 0.27, 0.36$ and 0.44 for $r_c = 8, 12, 16, 20\text{Å}$, for the Ubq/UIM1, and 0.28 and 0.35 for $r_c = 16$ and 20Å for the Cc/CcP, respectively, consistent with simulation data in Figs. 3 and 4. We can also plot $\epsilon_c^{\text{crit}}$ as it changes with crowder size $r_c$ as shown in Figure 5. For crowder-protein attraction values above this line, one will observe destabilization of protein association and stabilization below this line.

In summary, we have presented a quantitative theory for protein association equilibria in a crowded solution for both repulsive and attractive crowder-protein interactions. This work is important for providing a theoretical foundation of understanding the protein-protein interactions in a cellular environment in which proteins

---

**TABLE II**: Normalized average attraction strength, $\bar{\epsilon}_{\alpha}/\epsilon_c$, and the surface area, $S_{\alpha}$, (in Å$^2$) for Ubq, UIM1 and the Ubq/UIM1 complex

| $r_c$ | Ubq | UIM1 | Ubq/UIM1 |
|-------|-----|------|-----------|
|       | $\bar{\epsilon}_{\alpha}/\epsilon_c$ | $S_{\alpha}$ | $\bar{\epsilon}_{\alpha}/\epsilon_c$ | $S_{\alpha}$ | $\bar{\epsilon}_{\alpha}/\epsilon_c$ | $S_{\alpha}$ |
| 8     | 4.56 | 6543 | 4.01 | 4100 | 4.61 | 7522 |
| 12    | 4.71 | 9278 | 4.07 | 6418 | 4.75 | 10487 |
| 16    | 4.79 | 12409 | 4.07 | 9143 | 4.85 | 13830 |
| 20    | 4.85 | 15921 | 4.04 | 12273 | 4.91 | 17562 |

---

**FIG. 5**: Enthalpy-entropy compensation lines (i.e., $\Delta \Delta F_{\text{bind}} = 0$) in the parameter space ($\epsilon_c$, $r_c$) for Ubq/UIM1 and Cc/CcP complexes.
and crowding macromolecules exhibit non-specific interactions in addition to the excluded volume effects. The theory is based on the statistical mechanics and thermodynamics of a hard-sphere fluid. Even though proteins are highly anisometric, the repulsive contribution to the binding free energy is described well by the scaled particle theory of hard spheres. The expression for the attractive contribution is obtained by using thermodynamic perturbation theory and the radial distribution function of hard-sphere fluids. The theory is in excellent agreement with simulation results for the Ubq/UIM1 and Cc/CcP complexes over a wide range of the crowder sizes, packing fractions and attraction strengths.

We also observe crowding induced compensation for a critical protein-crowder interaction strength (independent of crowder packing fraction) leading to no change in the binding free energy with respect to bulk. Earlier Trout and co-workers had proposed a neutral-crowder hypothesis to explain the kinetic effect of small solution additives (crowders) that slow down the rate of protein association and dissociation without perturbing the equilibrium [35]. It will be interesting, in future, to explore the protein kinetics near this critical protein-crowder interaction strength to test if the neutral-crowder hypothesis is applicable in general. In future, we also plan to include attractions in the crowder-crowder interaction potential to study their interplay with protein-crowder and protein-protein interactions [39].

References

[1] C. Kleanthous, Protein-Protein Recognition (Oxford University Press, Oxford, UK, 2000).
[2] C. Tang, J. Iwahara, and G. M. Clore, Nature 444, 383 (2006).
[3] D. Boehr, P. Wright, et al., Science 320, 1429 (2008).
[4] A. Elcock, R. Gabdoulline, R. Wade, and J. McCammon, Journal of Molecular Biology 291, 149 (1999).
[5] M. Gilson and H.-X. Zhou, Annu. Rev. Biophys. Biomol. Struct. 36, 21 (2007).
[6] A. Fulton, Cell 30, 345 (1982).
[7] S. Zimmerman and A. Minton, Annual Review of Biophysics and Biomolecular Structure 22, 27 (1993).
[8] M. S. Cheung, D. Klimov, and D. Thirumalai, Proc. Natl. Acad. Sci. USA 102, 4753 (2005).
[9] H.-X. Zhou, G. Rivas, and A. P. Minton, Annu. Rev. Biophys. 37, 375 (2008).
[10] V. K. Shen, J. K. Cheung, J. R. Errington, and T. M. Truskett, J. Biomech. Eng. 131, 071002 (2009).
[11] J. S. Kim and A. Yethiraj, Biophys. J. 96, 1333 (2009).
[12] A. H. Elcock, Curr. Opin. Struct. Biol. 20, 196 (2010).
[13] J. Mittal and R. B. Best, Biophys. J. 98, 315 (2010).
[14] A. P. Minton and J. Wilf, Biochemistry 20, 4821 (1981).
[15] T. C. Jarvis, D. M. Ring, S. S. Daube, and P. H. von Hippel, J. Biol. Chem. 265, 15160 (1990).
[16] B. van den Berg, R. J. Ellis, and C. M. Dobson, EMBO J. 18, 6927 (1999).
[17] J. R. Wenner and V. A. Bloomfield, Biophysical journal 77, 3234 (1999).
[18] A. S. Morar, X. Wang, and G. J. Pielak, Biochemistry 40, 281 (2001).
[19] C. Patel, S. Noble, G. Weatherly, A. Tripathy, D. Winzor, and G. Pielak, Protein science 11, 997 (2002).
[20] N. Kozer and G. Schreiber, J. Mol. Biol. 336, 763 (2004).
[21] S. Zorrilla, G. Rivas, A. U. Acuña, and M. P. Lillo, Protein Sci. 13, 2960 (2004).
[22] Y. Phillip, E. Sherman, G. Haran, and G. Schreiber, Biophys. J. 97, 875 (2009).
[23] Q. Wang, A. Zhuravleva, and L. M. Gierasch, Biochemistry 50, 9225 (2011).
[24] A. Fodeke and A. Minton, J. Phys. Chem. B 115, 11261 (2011).
[25] A. P. Minton, Mol Cell. Biochem. 55, 119 (1983).
[26] H.-X. Zhou, J. Mol. Recognit. 17, 368 (2004).
[27] Y. C. Kim, R. B. Best, and J. Mittal, J. Chem. Phys. 133, 205101 (2010).
[28] J. F. Douglas, J. Dudowicz, and K. F. Fredd, Phys. Rev. Lett. 97, 875 (2009).
[29] M. Jiao, H.-T. Li, J. Chen, A. P. Minton, and Y. Liang, Biophys. J. 99, 914 (2010).
[30] J. Rosen, Y. C. Kim, and J. Mittal, J. Phys. Chem. B 115, 2683 (2011).
[31] S. Tarquato, T. M. Truskett, and P. G. Debenedetti, Phys. Rev. Lett. 84, 2064 (2000).
[32] Y. Phillip, V. Kiss, and G. Schreiber, Proc. Natl. Acad. Sci. USA 109, 1461 (2012).
[33] J. L. Lebowitz and J. S. Rowlinson, J. Chem. Phys. 41, 133 (1964).
[34] J. Mittal, J. Errington, and T. Truskett, Journal of Physical Chemistry B 111, 10054 (2007).
[35] S. Garde, A. E. García, L. R. Pratt, and G. Hummer, Biophys. Chem. 78, 21 (1999).
[36] Y. C. Kim and G. Hummer, J. Mol. Biol. 375, 1416 (2008).
[37] Y. C. Kim, C. Tang, G. M. Clore, and G. Hummer, Proc. Natl. Acad. Sci. USA 105, 12855 (2008).
[38] B. Baynes and B. Trout, Biophys. J. 87, 1631 (2004).
[39] J. Kim and A. Yethiraj, The Journal of Physical Chemistry B 115, 347 (2010).