

The Extent of Anisotropic Interactions Between Protein Molecules in Electrolyte Solutions

XUEYU SONG

Department of Chemistry, Iowa State University, Ames, IA 50011, USA

(Received October 2002; In final form October 2002)

In this report, we present a general formulation to calculate the electrostatic contribution to the effective interaction between two protein molecules in an electrolyte solution using boundary element method of solving linearized Poisson–Boltzmann (PB) equation. Our results for bovine pancreatic trypsin inhibitor (BPTI) at various relative orientations indicates that the anisotropy of the interaction is tens of $k_B T$. The implications of such strong anisotropy in protein crystallization is discussed.

Keywords: Bovine pancreatic trypsin inhibitor; Electrostatic contribution; Poisson–Boltzmann; Protein crystallization

INTRODUCTION

A crucial step in the determination of three-dimensional structures by X-ray crystallography is the production of suitable size crystals. This production is the bottleneck for most protein structure determination processes. For example, results from structural genomics show that the successful rate of proceeding from cloned proteins to structural determination is about 10%*. By July 31 of 2002, in a pilot human genome project documented at Brookhaven National Laboratory out of 124 proteins which are cloned, 63 were purified, 18 resulted in good crystals suitable for diffraction experiments and 14 structures have been determined*. Thus, even purified proteins are obtained the successful rate of producing a structural determination crystal is only 28%. Experiments clearly indicate that the success of protein crystallization depends sensitively on the physical conditions of the solution [1,2]. These conditions include temperature, salt concentration,

precipitant, pH and so on. The optimal crystallization condition often lies in a narrow window within a large set of possibilities. Traditional crystallization experiments are largely based on trial and error. It is therefore, useful to understand what kind of physical conditions might lead toward the optimal crystallization conditions and why.

Recently, we have shown that anisotropic interactions between proteins molecules play an important role in forming suitable crystals for the diffraction experiments [3]. As a first step towards a realistic model to access the anisotropy of protein–protein interactions, we will present calculations of the electrostatic contribution to the effective interaction between protein molecules in an electrolyte solution. We find that the anisotropy can be as large as tens of $k_B T$, therefore, could be used to stabilize the orientationally ordered crystal in contrast to orientationally disordered crystal which should be avoided in the search of optimal crystallization conditions.

In general, there are two types of protein–protein interactions in nature. One type of interaction is responsible for the protein–protein recognition to perform specific biological functions. In this case there are complementary regions on both proteins to recognize each other and hydrophobicity is the dominating factor [4]. On the other hand, the protein–protein interaction in protein crystallization does not necessarily involve complementary regions to establish protein–protein contacts. For example, we have recently analyzed the protein contacts for five lysozyme crystals from the protein data bank under five different crystallization conditions [5]. What we found is that the protein contacts can be formed from different parts of lysozyme surface

*See: <http://proteome.bnl.gov/progress.html>

residues depending upon solution conditions. Similar conclusions can be drawn from other studies as well. For example, Crosio *et al.* found that pancreatic ribonuclease uses nearly the entire protein surface residues to establish crystal-packing contacts under various crystallization conditions [6]. An extensive analysis on 78 protein crystals indicates that the amino acid composition involved in the protein contacts is indistinguishable from that of the protein surface accessible to the solvent [7]. These studies also suggest that crystal-packing contacts formed are sensitive to the solution conditions in contrast to the first type of protein–protein interactions where hydrophobic residues are favored [8,9]. Therefore, a universal model to capture the effective interaction between protein molecules may be developed for the second type of interactions.

For globular proteins there are two major universal anisotropic interactions, the geometric one due to uneven short-ranged repulsion and dispersion interactions of residues, and the electrostatic one due to the inhomogeneity of residue dipoles and surface charges. The former is determined by the residue distribution from protein structure and the latter can be tuned by the conditions of a protein solution. A full atomistic model is not practical under current technologies nor necessary since experimental modifications to the interaction are done at the residue level through mutagenesis. As a first step towards a realistic model of protein–protein interactions we will present a model at residue level and the electrostatic contribution to this model will be presented in this report.

AN EFFECTIVE INTERACTION MODEL BETWEEN PROTEIN MOLECULES AT RESIDUE LEVEL

In this model, each residue in a protein is represented by a sphere located at the geometric center of the residue determined by its native or approximate structure. The diameter of the sphere is determined by the molecular volume of the residue. The molecular surface of our model protein is defined as a Richard-Connolly surface spanned by the union of these residue spheres [10,11]. Each residue carries a permanent dipole moment located at the center of its sphere and the direction of the dipole is given by the amino acid type of protein's native structure. If a residue is charged the amount of charge is given by Henderson–Hasselbalch equation using the generic pK values of residues from Table II of Ref. [12], i.e. the local environmental effects on pK values are neglected. For each residue there is also a polarizable dipole at the center of the sphere, whose polarizability has been determined from our recent work [13]. There are three kinds of interactions: an

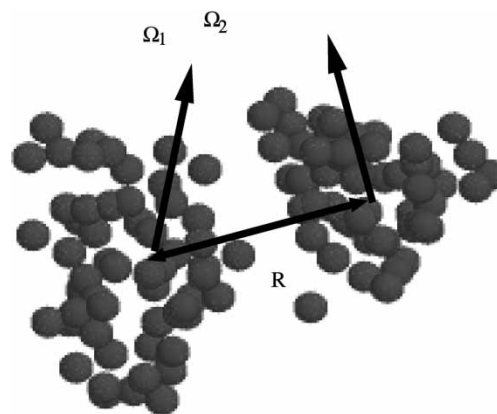


FIGURE 1 A schematic illustration of the effective interactions between two protein molecules rendered from the native structure of BPTI. Each sphere represents a residue and the electrostatic interaction energy due to residue dipoles and surface charges which can be obtained by solving the linearized PB equation. R is the center-to-center distance. Ω_1 and Ω_2 represent the orientations of both proteins, respectively.

electrostatic interaction, the Van der Waals attraction and a short range correction term to account for the short range interactions such as the desolvation energy, hydrophobic interaction and so on [14]. In this report, our calculations will be restricted to the situation where the anisotropy is due to the uneven charge distribution, which is expected to present the largest anisotropy (Fig. 1).

General Formulation

The electrostatic interaction is estimated from the Poisson–Boltzmann (PB) equation where the realistic shape of protein molecules are taken into consideration. The most popular method for solving the PB equation is the finite difference method (for example, see Refs. [15,16]). For the calculation of the effective interaction between two protein molecules, it is obviously difficult since a huge grid space is needed for the finite difference method. Another apparent drawback of this method is the treatment of the self-energy terms. An alternative method is the boundary element method [17–21]. Although the formulation for a single protein is well known a systematic presentation for two proteins case is not given yet [17]. Following Juffer *et al.* [18] we have generalized the integral equation of linearized PB equation in one domain to the two domains case (Fig. 2). Since the generalization is straightforward we only present the final results.

Consider the two molecular surfaces \sum_1 and \sum_2 spanned by the two protein molecules. There are N charges q_i at points \mathbf{r}_i inside a dielectric cavity enclosed by \sum_1 and there are N charges q_j at points \mathbf{r}_j inside a dielectric cavity enclosed by \sum_2 . Inside the dielectric cavities the dielectric constant is ϵ_1 and the dielectric constant of the solution is ϵ_2 . The ionic strength of the solution yields Debye screening

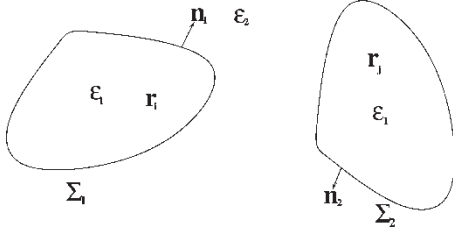


FIGURE 2 A schematic illustration of the electrostatic formulation of two protein molecules. The molecular surfaces of two proteins are Σ_1 and Σ_2 . The \mathbf{n}_1 and \mathbf{n}_2 are the outward unit normal to Σ_1 and Σ_2 , respectively. It should be noted that the sign of the gradient is changed due to this convention for the charges outside the cavity.

length κ . The potentials $[\varphi_1(\mathbf{r})$ and $\varphi_2(\mathbf{r})$] and their gradients $[(\partial\varphi_1/\partial n_1)(\mathbf{r})$ and $(\partial\varphi_2/\partial n_2)(\mathbf{r})$] over the molecular surfaces are given by the following integral equations

$$\begin{aligned} & \frac{1}{2} \left(1 + \frac{\epsilon_2}{\epsilon_1} \right) \varphi_1(\mathbf{r}_{01}) + \iint_{\Sigma_1} L_1(\mathbf{r}_1, \mathbf{r}_{01}) \varphi_1(\mathbf{r}_1) d\mathbf{r}_1 \\ & + \iint_{\Sigma_1} L_2(\mathbf{r}_1, \mathbf{r}_{01}) \frac{\partial \varphi_1}{\partial n_1}(\mathbf{r}_1) d\mathbf{r}_1 \\ & - \iint_{\Sigma_2} L_1(\mathbf{r}_2, \mathbf{r}_{01}) \varphi_2(\mathbf{r}_2) d\mathbf{r}_2 \\ & + \iint_{\Sigma_2} L_2(\mathbf{r}_2, \mathbf{r}_{01}) \frac{\partial \varphi_2}{\partial n_2}(\mathbf{r}_2) d\mathbf{r}_2 \\ & = \sum_{i=1}^{2N} q_i F(\mathbf{r}_i, \mathbf{r}_{01}) / \epsilon_1 \end{aligned} \quad (1)$$

$$\begin{aligned} & \frac{1}{2} \left(1 + \frac{\epsilon_1}{\epsilon_2} \right) \frac{\partial \varphi_1}{\partial n_1}(\mathbf{r}_{01}) + \iint_{\Sigma_1} L_3(\mathbf{r}_1, \mathbf{r}_{01}) \varphi_1(\mathbf{r}_1) d\mathbf{r}_1 \\ & + \iint_{\Sigma_1} L_4(\mathbf{r}_1, \mathbf{r}_{01}) \frac{\partial \varphi_1}{\partial n_1}(\mathbf{r}_1) d\mathbf{r}_1 \\ & - \iint_{\Sigma_2} L_3(\mathbf{r}_2, \mathbf{r}_{01}) \varphi_2(\mathbf{r}_2) d\mathbf{r}_2 \\ & + \iint_{\Sigma_2} L_4(\mathbf{r}_2, \mathbf{r}_{01}) \frac{\partial \varphi_2}{\partial n_2}(\mathbf{r}_2) d\mathbf{r}_2 \\ & = \sum_{i=1}^{2N} q_i \frac{\partial F(\mathbf{r}_i, \mathbf{r}_{01})}{\partial n_{01}} / \epsilon_1 \end{aligned} \quad (2)$$

$$\begin{aligned} & \frac{1}{2} \left(1 + \frac{\epsilon_2}{\epsilon_1} \right) \varphi_2(\mathbf{r}_{02}) - \iint_{\Sigma_1} L_1(\mathbf{r}_1, \mathbf{r}_{02}) \varphi_1(\mathbf{r}_1) d\mathbf{r}_1 \\ & + \iint_{\Sigma_1} L_2(\mathbf{r}_1, \mathbf{r}_{02}) \frac{\partial \varphi_1}{\partial n_1}(\mathbf{r}_1) d\mathbf{r}_1 \\ & + \iint_{\Sigma_2} L_1(\mathbf{r}_2, \mathbf{r}_{02}) \varphi_2(\mathbf{r}_2) d\mathbf{r}_2 \\ & + \iint_{\Sigma_2} L_2(\mathbf{r}_2, \mathbf{r}_{02}) \frac{\partial \varphi_2}{\partial n_2}(\mathbf{r}_2) d\mathbf{r}_2 \\ & = \sum_{i=1}^{2N} q_i F(\mathbf{r}_i, \mathbf{r}_{02}) / \epsilon_1 \end{aligned} \quad (3)$$

$$\begin{aligned} & \frac{1}{2} \left(1 + \frac{\epsilon_1}{\epsilon_2} \right) \frac{\partial \varphi_2}{\partial n_2}(\mathbf{r}_{02}) - \iint_{\Sigma_1} L_3(\mathbf{r}_1, \mathbf{r}_{02}) \varphi_1(\mathbf{r}_1) d\mathbf{r}_1 \\ & + \iint_{\Sigma_1} L_4(\mathbf{r}_1, \mathbf{r}_{02}) \frac{\partial \varphi_1}{\partial n_1}(\mathbf{r}_1) d\mathbf{r}_1 \\ & + \iint_{\Sigma_2} L_3(\mathbf{r}_2, \mathbf{r}_{02}) \varphi_2(\mathbf{r}_2) d\mathbf{r}_2 \\ & + \iint_{\Sigma_2} L_4(\mathbf{r}_2, \mathbf{r}_{02}) \frac{\partial \varphi_2}{\partial n_2}(\mathbf{r}_2) d\mathbf{r}_2 \\ & = \sum_{i=1}^{2N} q_i \frac{\partial F(\mathbf{r}_i, \mathbf{r}_{02})}{\partial n_{02}} / \epsilon_1, \end{aligned} \quad (4)$$

where

$$L_1(\mathbf{r}, \mathbf{r}_1) = \frac{\partial F}{\partial n}(\mathbf{r}, \mathbf{r}_1) - \frac{\partial P}{\partial n}(\mathbf{r}, \mathbf{r}_1) \epsilon_2 / \epsilon_1 \quad (5)$$

$$L_2(\mathbf{r}, \mathbf{r}_1) = P(\mathbf{r}, \mathbf{r}_1) - F(\mathbf{r}, \mathbf{r}_1) \quad (6)$$

$$L_3(\mathbf{r}, \mathbf{r}_1) = \frac{\partial^2 F}{\partial n_0 \partial n}(\mathbf{r}, \mathbf{r}_1) - \frac{\partial^2 P}{\partial n_0 \partial n}(\mathbf{r}, \mathbf{r}_1) \quad (7)$$

$$L_4(\mathbf{r}, \mathbf{r}_1) = -\frac{\partial F}{\partial n_0}(\mathbf{r}, \mathbf{r}_1) + \frac{\partial P}{\partial n_0}(\mathbf{r}, \mathbf{r}_1) \epsilon_1 / \epsilon_2, \quad (8)$$

and

$$F(\mathbf{r}, \mathbf{r}_1) = \frac{1}{4\pi|\mathbf{r} - \mathbf{r}_1|} \quad (9)$$

$$P(\mathbf{r}, \mathbf{r}_1) = \frac{e^{-\kappa|\mathbf{r} - \mathbf{r}_1|}}{4\pi|\mathbf{r} - \mathbf{r}_1|}. \quad (10)$$

Using collocation method by Atkinson and coworkers [22] the above integral equations can be solved and the potentials inside dielectric cavities are

$$\begin{aligned} \varphi_1(\mathbf{r}_{01}) &= - \iint_{\Sigma_1} L_1(\mathbf{r}_1, \mathbf{r}_{01}) \varphi_1(\mathbf{r}_1) d\mathbf{r}_1 \\ & - \iint_{\Sigma_1} L_2(\mathbf{r}_1, \mathbf{r}_{01}) \frac{\partial \varphi_1}{\partial n_1}(\mathbf{r}_1) d\mathbf{r}_1 \\ & + \sum_{i=1}^N \frac{q_i}{\epsilon_1} F(\mathbf{r}_i, \mathbf{r}_{01}) \end{aligned} \quad (11)$$

$$\begin{aligned} \varphi_2(\mathbf{r}_{02}) &= - \iint_{\Sigma_2} L_1(\mathbf{r}_2, \mathbf{r}_{02}) \varphi_2(\mathbf{r}_2) d\mathbf{r}_2 \\ & - \iint_{\Sigma_2} L_2(\mathbf{r}_2, \mathbf{r}_{02}) \frac{\partial \varphi_2}{\partial n_2}(\mathbf{r}_2) d\mathbf{r}_2 \\ & + \sum_{j=1}^N \frac{q_j}{\epsilon_1} F(\mathbf{r}_j, \mathbf{r}_{02}). \end{aligned} \quad (12)$$

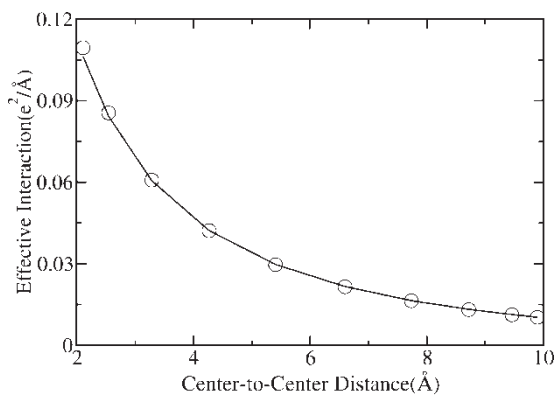


FIGURE 3 Effective interaction comparison between our calculation (circles) and the solution (solid line) based on Carnie and Chan analytical solution [24]. The radius of both spheres is 1.0 \AA and there is a unit charge at each center of the sphere. The dielectric constant inside spheres is 1.0 and outside sphere is 20.0. The Debye screening length is 0.1 \AA^{-1} .

With these potentials the solvation energy of the protein molecules at a center-to-center distance R and orientations Ω_1 and Ω_2 is

$$E(R, \Omega_1, \Omega_2) = \sum_{i=1}^N \frac{q_i}{\epsilon_1} \varphi_1(\mathbf{r}_i) + \sum_{j=1}^N \frac{q_j}{\epsilon_1} \varphi_2(\mathbf{r}_j).$$

The effective interaction between the two protein molecules is

$$\Delta E(R, \Omega_1, \Omega_2) = E(R, \Omega_1, \Omega_2) - E(R = \infty, \Omega_1, \Omega_2) - \sum_{i=1}^N \sum_{j=1}^N \frac{q_i q_j e^{-k|\mathbf{r}_i - \mathbf{r}_j|}}{\epsilon_1 |\mathbf{r}_i - \mathbf{r}_j|}.$$

The last term of the above equation is the interaction between the charges in proteins when the solution has the same dielectric constant ϵ_1 as the interior of the protein, but with Debye screening length k .

A Test Example

In order to test the above formulation we apply the above procedure to the effective interaction between two dielectric spheres with unit charges located at the center of respective spheres.

Figure 3 indicates that our formulation indeed gives correct effective interaction compared with the analytical solution of two spheres.

Effective Interactions Between Two Bovine Pancreatic Trypsin Inhibitor (BPTI) Molecules

Using the above formulation the effective interactions between two BPTI molecules are calculated at various distances and orientations. Figure 4 shows the results at four different relative orientations.

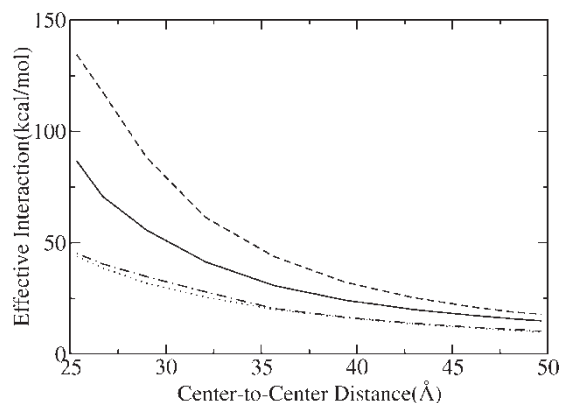


FIGURE 4 Electrostatic contribution to the effective interaction between two BPTI molecules at four different relative orientations. The molecular surface is generated by Sanner program [24]. The dielectric constant inside the proteins is 2.0 and outside the protein is 80.0. The Debye screening length is 0.1 \AA^{-1} and the temperature is 298 K. The error of the interaction is within 1 kcal/mol.

It clearly indicates that the effective interaction can vary tens of $k_B T$ among different relative orientations. Since the surface charge distribution can be tuned by changing the pH of a solution, hence, the orientation-dependent effective interaction between protein molecules. Therefore, we expect that the change of solution pH can have significant effect on the effective interaction anisotropy, thus, some particular pH may be optimal for the stabilization of orientationally ordered crystal.

Acknowledgements

The author is grateful for the financial support by a Petroleum Research Fund, administrated by American Chemical Society and the calculations in this work were performed in part on a IBM workstation cluster made possible by grants from IBM in the form of a Shared University Research grant, the United States Department of Energy, and the United States Air Force Office of Scientific Research.

References

- [1] Michel, H. ed. (1991) *Crystallization of Membrane Proteins*, 1st Ed. (CRC Press, Boca Raton, FL).
- [2] McPherson, A. (1999) *Crystallization of Biological Macromolecules*, 1st Ed. (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, USA).
- [3] Song, X. (2002) "The role of anisotropic interactions in protein crystallization", *Phys. Rev. E* **66**, 011906.
- [4] Elcock, A.H., Sept, D. and McCammon, J.A. (2001) "Computer simulation of protein-protein interactions", *J. Phys. Chem. B* **105**, 1504.
- [5] X. Song. The protein data bank entries are 6lyt,1lzt, 1lkr,0lzt,1lys and the crystallization conditions are from: www.bmcd.nist.gov:8080/bmcd, 2002. Unpublished results.
- [6] Janin, J., Crosio, M.P. and Jullien, M. (1992) "Crystal packing in six crystal forms of pancreatic ribonuclease", *J. Mol. Biol.* **228**, 243.

- [7] Carugo, O. and Argos, P. (1997) "Protein-protein crystal-packing contacts", *Protein Sci.* **6**, 2261.
- [8] Janin, J. and Rodier, F. (1995) "Protein-protein interactions at crystal contacts", *Proteins Struct. Funct. Genet.* **23**, 580.
- [9] Dasgupta, S., Iyer, G.H., Bryant, S.H., Lawrence, C.E. and Bell, J.A. (1997) "Extent and nature of contacts between protein molecules in crystal lattices and between subunits of protein oligomers", *Proteins Struct. Funct. Genet.* **28**, 494.
- [10] Connolly, M.L. (1983) "Solvent-accessible surfaces of proteins and nucleic acids", *Science* **221**, 709.
- [11] Lee, B. and Richards, F.M. (1971) "The interpretation of protein structures: estimation of static accessibility", *J. Mol. Biol.* **55**, 379.
- [12] Stryer, L. (1988) *Biochemistry*, 3rd Ed. (W.H. Freeman and Co. Press, New York, USA).
- [13] Song, X. (2002) "An inhomogeneous model of protein dielectric properties: intrinsic polarizabilities of amino acids", *J. Chem. Phys.* **106**, 9359.
- [14] Israelachvili, J. (1992) *Intermolecular and Surface Forces*, 2nd Ed. (Academic Press, London, UK).
- [15] Gilson, M.K., Sharp, K.A. and Honig, B.H. (1988) "Calculating the electrostatic potential of molecules in solution", *J. Comp. Chem.* **9**, 327.
- [16] Davis, M.E., Madura, J.D., Luty, B.A. and McCammon, J.A. (1991) "Electrostatics and diffusion of molecules in solution: simulations with the University of Houston Brownian Dynamics Program", *Comput. Phys. Commun.* **62**, 187.
- [17] Juffer, A.H., Botta, E.F.F., van Keulen, B.A.M., van der Ploeg, A. and Berendsen, H.J.C. (1991) "The electric potential of a macromolecule in a solvent: a fundamental approach", *J. Comp. Phys.* **97**, 144.
- [18] Zhou, H.X. (1993) "Boundary element solution of macromolecular electrostatics: interaction energy between two proteins", *Biophys. J.* **65**, 955.
- [19] Song, X. and Chandler, D. (1998) "Dielectric solvation dynamics of molecules of arbitrary shape and charge distribution", *J. Chem. Phys.* **108**, 2594.
- [20] Zauhar, R.J. and Morgan, R.S. (1985) "A new method for computing the macromolecular electric potential", *J. Mol. Biol.* **186**, 815.
- [21] Yoon, B.J. and Lenhoff, A.M. (1990) "A boundary element method for molecular electrostatics with electrolyte effects", *J. Comput. Chem.* **11**, 1080.
- [22] Atkinson, K. (1997) *The Numerical Solution of Fredholm Integral Equations of the Second kind*, 1st Ed. (Cambridge University Press, Cambridge, New York).
- [23] Carnie, S.L. and Chan, D.Y.C. (1993) "Interaction free energy between identical spherical colloidal particles: the linearized Poisson-Boltzmann theory", *J. Colloid Interface Sci.* **155**, 297.
- [24] Sanner, M.F. http://www.scripps.edu/pub/olson-web/people/sanner/html/msms_home.html