JCTC Journal of Chemical Theory and Computation

Article

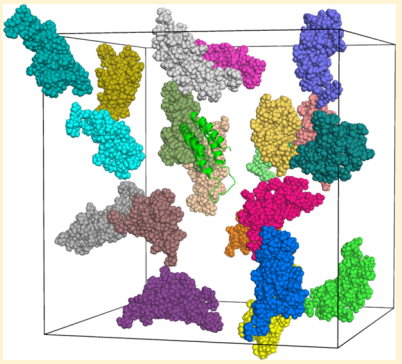
pubs.acs.org/JCTC

FFT-Based Method for Modeling Protein Folding and Binding under Crowding: Benchmarking on Ellipsoidal and All-Atom Crowders

Sanbo Qin and Huan-Xiang Zhou*

Department of Physics and Institute of Molecular Biophysics, Florida State University, Tallahassee, Florida, United States

ABSTRACT: It is now well recognized that macromolecular crowding can exert significant effects on protein folding and binding stability. In order to calculate such effects in direct simulations of proteins mixed with bystander macromolecules, the latter (referred to as crowders) are usually modeled as spheres and the proteins represented at a coarse-grained level. Our recently developed postprocessing approach allows the proteins to be represented at the all-atom level but, for computational efficiency, has only been implemented for spherical crowders. Modeling crowder molecules in cellular environments and *in vitro* experiments as spheres may distort their effects on protein stability. Here, we present a new method that is capable for treating aspherical crowders. The idea, borrowed from protein−protein docking, is to calculate the excess chemical potential of the proteins in crowded solution by fast Fourier transform (FFT). As the first application, we studied the effects of ellipsoidal crowders on the folding and binding free energies of all-atom proteins, and found, in agreement with previous direct simulations with coarse-grained protein models, that the aspherical crowders exert greater stabilization effects than spherical crowders of the same volume. Moreover, as demonstrated here, the FFT-based method has the important property that its computational cost does not increase strongly even when the level of details in representing the crowders is increased all the way to all-atom, thus significantly accelerating realistic modeling of protein folding and binding in cell-like environments.



I. INTRODUCTION

It is now widely recognized that the environments inside cells, crowded with bystander macromolecules, potentially can significantly affect the equilibria of protein folding and binding.[1,2] In many theoretical and simulation studies,[3−13] the effects of crowding have been calculated by assuming the bystander macromolecules, or "crowders", as spherical. This was done as a matter of convenience. Some recent simulation studies using simple bead representations of crowders have raised the prospect that the shape of crowders can strongly influence the magnitude of crowding effects.[14−17] Experimental studies have found that crowders of similar sizes can result in different effects on protein folding and binding stability;[18−21] variations in shapes could possibly account for some of the differences in crowding effects. Here, we present a general method that is capable of treating arbitrarily shaped crowders, including those represented at the all-atom level.

The effects of crowding on folding and binding stability can be calculated from direct simulations, in which the test protein is simulated along with the crowders.[3,4,9−12] To calculate how the crowders change the folding stability, one has to obtain the full free energy surface that covers both the folded and the unfolded states. To overcome the enormous computational cost of such calculations, the protein has to be represented at a coarse-grained level, and the crowders are often treated with an even simpler representation, such as a sphere or a few linked beads.

We have developed an alternative approach known as postprocessing,[5,6] and other have used the same idea,[22,23] in which the excess chemical potential arising from the interactions of the folded or unfolded protein with the crowders is calculated, and the difference between the chemical potentials for the two end states then yields the effect of crowding on the folding stability. This approach circumvents the need to simulate the protein in rarely sampled regions of the conformational space (e.g., around the transition state), since only the end states are used. More importantly, the expensive simulation of the protein in crowded solution is totally avoided. What is required is a method for calculating the excess chemical potential, via fictitiously inserting the protein into the crowded solution.

The postprocessing approach potentially holds tremendous advantages over the direct simulation approach.[24,25] In particular, the test protein can now be represented at the all-atom level. We have developed efficient algorithms to calculate the excess chemical potential for an atomistic protein interacting with spherical crowders, explicitly making use of the spherical shape of the crowders.[5,7] When applied to proteins represented at a coarse-grained level such that the direct simulation approach can be used, we have found that the postprocessing approach yields the same effects of crowding on the population ratio of different conformational states[6] and on the folding free energy surface.[25] In essence, from the conformational ensembles of the protein in two end states in dilute solution, the postprocessing approach can predict the change in the free energy difference between the two end states in crowded solution.

Received: June 18, 2013
Published: August 19, 2013

   dx.doi.org/10.1021/ct4005195 | *J. Chem. Theory Comput.* 2013, 9, 4633−4643

In principle, the postprocessing approach allows not only the test proteins but also the crowders to be modeled with arbitrary shapes and at any level of details. The present study aims to reach this full potential of the postprocessing approach, by developing a general method for calculating the excess chemical potential of test proteins in crowded solution. Our method was born out of the observation that the fictitious insertion of a test protein into a distribution of crowders is in spirit similar to the protein−protein or protein−ligand docking problem. For the latter problem, a powerful solution is based on fast Fourier transform (FFT).[26−30] We borrowed the FFT technique for calculating the excess chemical potential. The basic idea is to map the protein molecule and the crowders to a 3-dimensioanl grid, and express the protein-crowder interaction energy as a correlation function in the grid space. The correlation function is then evaluated via FFT.

As the first application of our FFT-based method, we calculated the effects of ellipsoidal crowders on the folding and binding free energies of all-atom proteins (Figure 1), and found

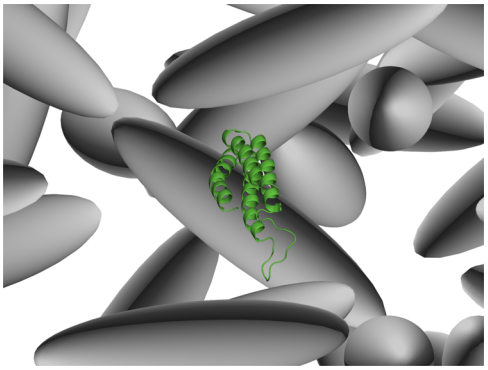


Figure 1. Illustration of native cytochrome b_{562} interacting with ellipsoidal crowders. The crowders have an axis ratio of 1:1:4 and an equivalent radius of 20 Å.

that the aspherical crowders exert greater stabilization effects than spherical crowders of the same volume. Because the protein and crowders are mapped to a grid, in principle, the computational cost of the FFT-based method remains the same when the level of details in representing the crowders is increased. We demonstrate the feasibility of the FFT-based method for dextran as crowders represented at the all-atom level.

II. THEORETICAL BASIS

In the postprocessing approach, the excess chemical potential, $\Delta\mu$, arising from the interactions of the test protein with a distribution of crowders is calculated as[24,25]

$$\exp(-\Delta\mu/k_BT) = \langle \exp[-U_{int}(\mathbf{X}, \Omega, \mathbf{R})/k_BT]\rangle_{0;crowd} \quad (1)$$

where $U_{int}(\mathbf{X},\mathbf{\Omega},\mathbf{R})$ is the protein−crowder interaction energy when the protein has conformation $\mathbf{X}$, orientation $\mathbf{\Omega}$, and position $\mathbf{R}$; k_B is Boltzmann's constant; T is the absolute temperature; $\langle\cdots\rangle_{0;crowd}$ means averaging over the conformation, orientation, and position of the test protein and over the configuration of the crowders; and the subscript "0" signifies that the protein conformations are those sampled in the absence of crowders (while the protein is restricted a given end state). For atomistic proteins and crowders, a direct calculation of $U_{int}(\mathbf{X},\mathbf{\Omega},\mathbf{R})$ requires exhaustive enumeration of all protein−crowder atom pairs, and the calculation needs to be repeated for many placement positions of the protein. It is this calculation that was replaced by FFT.

Many theoretical and simulation studies of crowding effects have modeled protein−crowder interactions as hard-core repulsion. In this case, $U_{int}(\mathbf{X},\mathbf{\Omega},\mathbf{R})$ has either a value of infinity when the protein clashes with any crowder or 0 otherwise; correspondingly, the Boltzmann factor $\exp[-U_{int}(\mathbf{X},\mathbf{\Omega},\mathbf{R})/k_BT]$ is either 0 or 1, and its average over $\mathbf{R}$ is the fraction, p, of placements of the protein that are free of clash with the crowders. The basic idea of the FFT-based method can be easily illustrated on this form of protein−crowder interactions (Figure 2). For each snapshot sampled

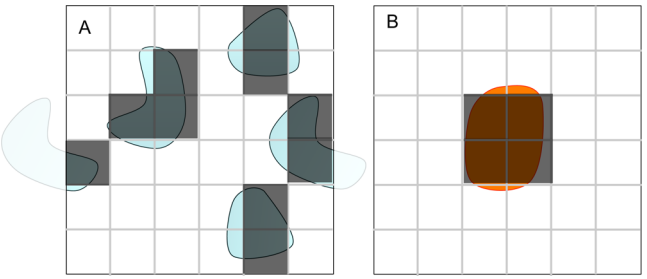


Figure 2. Illustration of the FFT-based method. The protein−crowder interactions are assumed to be hard-core repulsion. (A) In mapping the crowders to a grid, each voxel is assigned a value of 1 (dark shaded voxels) when its center (referred to as grid point) is inside a crowder; otherwise, a value of 0 is assigned (unshaded voxels). Note that the mapping works in the same way when crowders with different shapes and orientations are present. (B) A protein molecule, here located at the center of the grid, is mapped to the grid in a similar way.

from a simulation of the crowders with periodic boundary condition, we discretize the unit cell into a grid. We introduce a function $f(\mathbf{n})$, with value 1 when the grid point $\mathbf{n}$ is inside any crowder atom and 0 otherwise. Essentially, the crowders are mapped to the grid, where points with value 1 represent the region occupied by the crowders and points with value 0 represent the voids between crowders (Figure 2A). From another simulation of the protein alone, we take the protein molecule (with conformation $\mathbf{X}$ and orientation $\mathbf{\Omega}$) and placed it at the center of the grid, which for convenience is taken to define the grid point $\mathbf{n} = 0$ (i.e., the origin). We define another function $g(\mathbf{n})$, with value 1 when the grid point $\mathbf{n}$ is inside any protein atom and 0 otherwise. The protein is thus mapped to a set of grid points (identified by having $g(\mathbf{n}) = 1$; Figure 2B). Note that, when the protein is moved from the center of the grid to an arbitrary grid point $\mathbf{m}$, a grid point $\mathbf{n}$ in the body-fixed frame corresponds to grid point $\mathbf{n} + \mathbf{m}$ in the laboratory frame. The correlation function

$$c(m) = \sum_n f(n)g(n + m) \quad (2)$$

then represents the number of grid points where the crowders and the protein overlap. Any grid point where placement of the protein is free of clash with the crowders is distinguished by $c(\mathbf{m}) = 0$; grid points where placement of the protein results in clash with the crowders would have $c(\mathbf{m}) \geq 1$. The value of the Boltzmann factor $\exp[-U_{int}(\mathbf{X},\mathbf{\Omega},\mathbf{m})/k_BT]$ is related to $c(\mathbf{m})$ via

$$\exp[-U_{int}(\mathbf{X}, \Omega, \mathbf{m})/k_BT] = H[c(\mathbf{m})] \quad (3)$$

dx.doi.org/10.1021/ct4005195 | J. Chem. Theory Comput. 2013, 9, 4633−4643

where $H(l) = 1$ if $l = 0$ and 0 if $l \geq 1$.

Instead of directly evaluating the correlation function, we use FFT, taking advantage of the tremendous speed up of this technique. The Fourier transform of the correlation function can be expressed as

$$C(\mathbf{k}) = F(\mathbf{k})G^*(\mathbf{k}) \tag{4}$$

where $F(\mathbf{k})$ and $G(\mathbf{k})$ are the Fourier transforms of $f(\mathbf{n})$ and $g(\mathbf{n})$, respectively, and "*" denotes complex conjugate. So the FFT-based calculation of $c(\mathbf{m})$ involves first the forward Fourier transforms of $f(\mathbf{n})$ and $g(\mathbf{n})$ and then the inverse Fourier transform of $C(\mathbf{k})$. Note that this calculation yields at once the values of $c(\mathbf{m})$ and hence of $\exp[-U_{\text{int}}(\mathbf{X},\mathbf{\Omega},\mathbf{m})/k_BT]$ at all the grid points in the unit cell, allowing the average over placement position to be carried out to yield the clash-free fraction p.

The same idea can be extended to the case where the protein−crowder interactions are "soft". Intermolecular interactions are typically modeled as comprised of pairwise additive terms of the form

$$U_{\text{int}} = \sum_{i,j} \beta_i \beta_j u(r_{ij}) \tag{5}$$

where i and j are indices for the atoms of the crowders and the protein (in an atomistic representation), respectively, and r_{ij} are interatomic distances. For example, Coulomb interactions can be represented in this way, with β_i and β_j denoting partial charges and $u(r_{ij}) \propto 1/r_{ij}$. The repulsive and attractive components of van der Waals interactions can be likewise represented, with $u(r_{ij}) \propto 1/r_{ij}^{12}$ and $u(r_{ij}) \propto 1/r_{ij}^{6}$, respectively. We now use the terminology of Coulomb interactions to illustrate how this form of interaction energy can be expressed as a correlation function, thus suitable for evaluation via FFT.[27−30] The Coulomb interaction energy between two particles can be written as the product of the electrostatic potential of one particle and the charges on the other particle. We can thus identify the first function $f(\mathbf{n})$ as the electrostatic potential of the crowders, evaluated at grid point $\mathbf{n}$:

$$f(\mathbf{n}) = \sum_i \beta_i u(|\mathbf{r}_i - \mathbf{n}|) \tag{6}$$

where $\mathbf{r}_i$ denotes the position of crowder atom i. We then distribute the partial charges of the protein atoms to the grid points. Suppose that, when the protein is placed at the center of the grid, the partial charge accorded to grid point $\mathbf{n}$ is $g(\mathbf{n})$. The Coulomb interaction energy, when the protein is placed at grid point $\mathbf{m}$, is then given by the correlation function of eq 2.

As already explained, in the postprocessing approach, we calculate the excess chemical potentials for end states.[5,6,8,18] The FFT-based method allows the average of the Boltzmann factor $\exp[-U_{\text{int}}(\mathbf{X},\mathbf{\Omega},\mathbf{R})/k_BT]$ over the protein placement position to be carried out. From the simulation of each end state in dilute solution, we pick snapshots of the protein for the averaging over $\mathbf{X}$ and $\mathbf{\Omega}$. Averaging is also taken over configurations of the crowders, picked from a crowder-only simulation, finally yielding $\Delta\mu$. For studying protein folding stability under crowding, the end states are the native state ("N") and the denatured state ("D"); we denote the end-state excess chemical potentials as $\Delta\mu_N$ and $\Delta\mu_D$. The effect of crowding on the folding free energy is then given by

$$\Delta\Delta G_f = \Delta\mu_N - \Delta\mu_D \tag{7}$$

Similarly, for protein binding the end states are the bound state, in which the complex ("C") is the solute, and the unbound state, in which the two subunits ("A" and "B") are separately dissolved in solution. In this case, we carry out separate simulations of the complex and the two individual subunits and calculate the corresponding excess chemical potentials $\Delta\mu_C$, $\Delta\mu_A$, and $\Delta\mu_B$. The effect of crowding on the binding free energy is then

$$\Delta\Delta G_b = \Delta\mu_C - (\Delta\mu_A + \Delta\mu_B) \tag{8}$$

III. COMPUTATIONAL DETAILS

3.1. Background. For an all-atom protein interacting with spherical, repulsive crowders, we previously developed an efficient method for calculating p, the clash-free fraction of protein placements.[5] We mapped the covolume of a crowder and the protein onto a grid centered on the crowder. Because of the spherical crowder shape, the covolumes are the same for all the crowders and hence needs to be mapped only once. The set of grid points representing the covolume was then replicated to all the crowders in the unit cell. The grid points that are not part of any covolume are the ones where placement of the protein is clash-free; the ratio of the number of such grid points to the total number of grid points in the unit cell then equals p.

When a crowder has an aspherical shape, a mere change in its orientation will affect not only the orientation but also the shape and size of the covolume. Extending the way just outlined for calculating p to aspherical crowders will involve separately mapping the covolume to the grid for each crowder in the unit cell. This rigorous but slow calculation is not practical for routine use, but as further described below, we used it on a subset of cases to generate exact results for calibrating the FFT-based method.

For spherical crowders, we also developed a semianalytical method for calculating p, based on identifying the region in and around the protein that is inaccessible to any part of a crowder.[7] Again, because of the spherical crowder shape, this inaccessible region is the same for all the crowders in the unit cell.

3.2. Implementation of the FFT-Based Method. To benchmark the speed and accuracy of the FFT-based method, we studied the interactions of all-atom proteins with ellipsoidal, repulsive crowders (Figure 1). The unit cell of the crowder simulations (see below) was $10^3 \times 10^3 \times 10^3$ Å^3. We discretized the unit cell into a grid with 1-Å spacing, resulting in a total of 10^9 voxels (to be denoted as N_{tot}). The center of a voxel was taken as the grid point that represented that voxel. The function $f(\mathbf{n})$ for the crowders was assigned to 1 whenever $(n_x - x_i)^2/a^2 + (n_y - y_i)^2/a^2 + (n_z - z_i)^2/b^2 \leq 1$ for any crowder, where (n_x, n_y, n_z) are the projections of grid point $\mathbf{n}$ on the three principal axes of that crowder, (x_i, y_i, z_i) are the Cartesian coordinates of the center of the crowder in the principal-axes frame, and a and b are the transverse and longitudinal semiaxes of the crowder. Similarly, $g(\mathbf{n})$ was assigned 1 whenever $\mathbf{n}$ was within any atom of the protein.

We used the free library FFTW[31] for computing the discrete Fourier transforms. We ran the code in multithreading mode by compiling with the Linux pthread library, to take advantage of multiple-core processors. On an 8-core processor, maximum speed was achieved with 7 threads. The run time to calculate the clash-free fraction p for placing a single conformation of the

protein into a single configuration of the crowders was 46 s using 8 cores of an AMD Opteron 6174 processor.

Using the FFT-based method, the number of operations in a single calculation of p is of the order of $N_{tot}\ln_2 N_{tot}$. In comparison, a direct method where the crowder grid and protein grid are checked for clash would involve $N_{tot}N_{prot}$ operations, where N_{prot} is the number of grid points inside the protein molecule (corresponding to $g(\mathbf{n}) = 1$). With $N_{tot} = 10^9$ and N_{prot} of the order of 10^4, the speedup of the FFT-based is ~350-fold.

In calculating $\Delta\mu$ for each end state of a protein, we averaged over 10 configurations of the crowders and over up to 1000 protein conformations, similar to our previous work with spherical crowders.[5] Because we exhaustively sample all the positions inside a large unit cell of the crowders for possible placement of the test protein, the variations of $\Delta\mu$ among the configurations of the crowders are very small, usually less than $0.1k_BT$. The forward Fourier transform of the function $g(\mathbf{n})$ for a given protein conformation was calculated once and used for placement into different configurations of the crowders. Similarly, the forward Fourier transform of the function $f(\mathbf{n})$ for a given configuration of the crowders was calculated once and used for placing different protein conformations.

3.3. Generation of Crowder Configurations. Simulations of the ellipsoidal crowders were carried out by using PackLSD,[32] an event-driven simulation program. As already stated, the unit cell was a cube with side length of 10^3 Å, with periodic boundary condition applied. Each simulation started with the desired number of crowders with the desired axis ratio but a smaller size. The size of the crowders was then gradually grown to the desired value, followed by equilibration. The pair correlation function and pressure were checked to ensure full equilibration. Ten snapshots, separated by 10^6 collision events, were taken for $\Delta\mu$ calculations. For each crowder, the Cartesian coordinates of the center and the orientation in quaternions were saved.

Crowders with three axis ratios were studied: 4:4:1 (oblate); 1:1:4 (short prolate); and 1:1:8 (long prolate). For each shape, four sizes were studied, with volumes equal to those of spheres with radii of 15, 20, 30, 50 Å. We refer to these as the radii of the equivalent spheres, or simply the equivalent radii, and denote them with the symbol R_{eq}. The number of crowders in the unit cell was varied so the volume fraction of the crowders was 5%, 15%, 25%, or 35%. Corresponding conditions for spherical crowders were studied previously,[5] so a comparison with those results allowed us to identify the influences of crowder shape.

3.4. Generation of Protein Conformations. The protein conformations were identical to those in our previous study of spherical crowders.[5] We briefly describe these to provide essential information. The protein for studying folding stability was cytochrome b_{562}, a four-helix protein with 106 residues. One thousand conformations were taken from separate simulations at 300 and 500 K in explicit solvent, representing the native and denatured states, respectively.[33] The protein complex for studying binding stability was that formed by barnase and barstar. The bound state was represented by 548 conformations sampled from a 7.2-ns simulation of the complex in explicit solvent.[34] The subunits in these conformations were then separated to yield the conformations of unbound barnase and barstar. Protein atoms were assigned Bondi radii.[35]

3.5. Rigorous Calculation of p and Calibration of FFT-Based Method. We wanted to have a rigorous way of calculating the clash-free fraction so we could check the accuracy of the FFT-based method. For this purpose, we chose the covolume-based approach. That is, we first found the covolume for each crowder; the union of all the covolumes around all the crowders then constituted the region where placement of the protein would result in clash, and the rest of the unit cell constituted the clash-free region. This was the idea behind our previous method for spherical crowders.[5]

In this approach, we still had to sample the placement positions on a grid but avoided the more severe approximation of the FFT-based method in representing the crowders and the protein molecule on a grid. Our previous study found that, with a 1-Å spacing, the error caused by sampling the placement positions on a grid was negligible.[5] This was confirmed in the present study for ellipsoidal crowders.

The covolume was comprised of a set of grid points where placement of the protein would result in clash. We first included as part of the covolume all grid points inside the crowder. We then eliminated all grid points that are separated from the crowder surface by more than the distance between the most outlying atom and the center of the protein. Each remaining grid point was checked to see whether placement of the protein was allowed. To do this, the protein atoms, one at a time, were tested for possible clash with the crowder. This test consisted of two steps. The first crude but fast step was to check whether the center of the atom was inside the crowder after inflating all of its semiaxes by the radius of the atom. If the first step did not detect clash, the second rigorous but expensive step was followed. The latter consisted of calculating, from an analytical formula,[36] the distance of closest approach when the atom was moved toward the center of the crowder, and checking whether that distance was less than the radius of the atom. Whenever an atom was detected to clash with the crowder, the test stopped and the grid point was included as part of the covolume. If none of the protein atoms clashed with the crowder, the grid point was eliminated.

After the covolumes for all the crowders were found, the number, $\cup_{cov}$, of grid points in their union was counted. The clash-free fraction was the given by

$$p = 1 - \cup_{cov} / N_{tot} \tag{9}$$

As alluded to above, the main source of error for the FFT-based method came from representing the crowders and the protein molecule on a grid. This error could be reduced by decreasing the grid spacing, but that would significantly increase the computational cost. We sought a less expensive correction. The same error occurred in calculating the covolume for a single crowder. We noticed that the relative error in the number, N_{cov}, of grid points representing the covolume was largely invariant when the test protein was changed or the crowder shape was changed, depending mostly on the crowder size (as measured by the radius of the equivalent sphere). The relative errors were 3.65%, 3.06%, 2.38%, and 1.69% for R_{eq} = 15, 20, 30, and 50 Å, respectively, all in the direction of underestimation. We used this finding to correct the results from the FFT-based method. To this end we transformed eq 9 to

$$\ln p = \ln(1 - \cup_{cov} / N_{tot}) \approx - \cup_{cov} / N_{tot} \tag{10}$$

$\cup_{cov}$ and thus $\ln p$ should have errors similar to that of N_{cov}, which could be corrected through multiplying by a factor $1 + \alpha$, with α denoting the relative error stated above. We hence

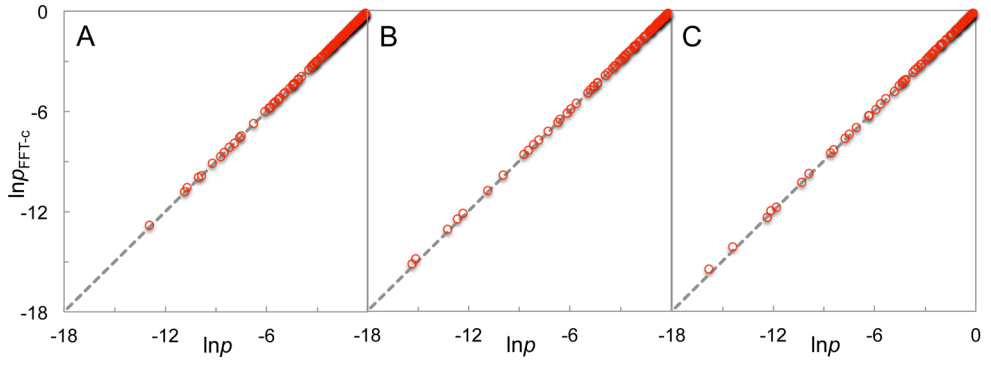


Figure 3. Comparison of the clash-free fractions calculated by the FFT-based method (after the correction of eq 11) and the rigorous, covolume-based method. Each data point displays as y and x coordinates the results of the two methods for a single protein conformation interacting with a single crowder configuration. A total of 16 crowder configurations representing different crowding conditions (4 crowder sizes combined with 4 crowder volume fractions) are considered. (A) Results for spherical crowders interacting with 8 different protein systems: native and denatured cytochrome b_{562}; barnase, barstar, and their complex; the ε and θ subunits of the DNA polymerase III holoenzyme, and the complex of the two subunits. The results of the covolume-based method are from our previous studies.[5,18] (B) Results for oblate ellipsoidal crowders (axis ratio at 4:4:1) interacting with the first 5 protein systems. (C) Corresponding results for prolate ellipsoidal crowders (axis ratio at 1:1:4).

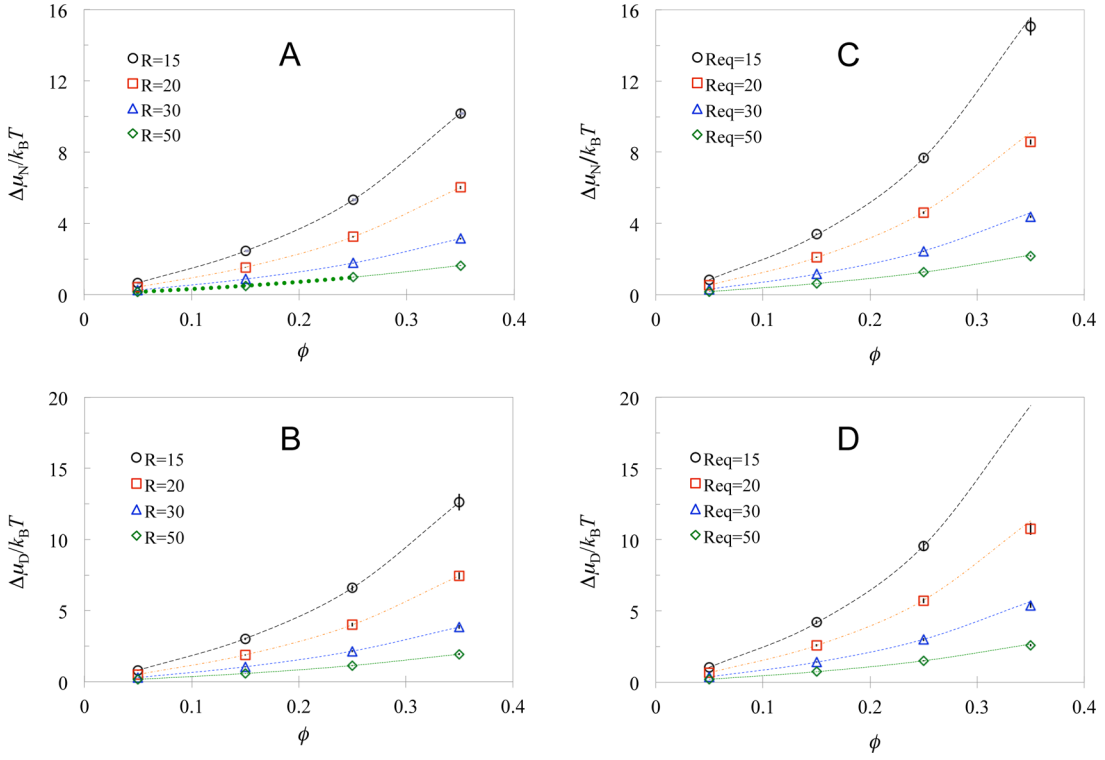


Figure 4. $\Delta\mu_N$ and $\Delta\mu_D$ for cytochrome b_{562} interacting with oblate ellipsoidal crowders and the equivalent spherical crowders. The results in (A) and (B) for the spherical crowders are from our previous study, with curves displaying fits to the scaled particle theory.[5] Symbols in (C) and (D) are results for the ellipsoidal crowders obtained from the FFT-based method; curves display fits to the fundamental measure theory.

applied the following correction to p_{FFT}, the clash-free fraction calculated by the FFT-based method:

$$\ln p_{FFT-c} = (1 + \alpha)\ln p_{FFT} \quad (11)$$

After this correction, results calculated by the FFT-based method for spherical crowders agreed to within 1.5% with those from our previous studies[5,18] (Figure 3A). The corrected FFT-based results also showed similar accuracy for prolate and oblate crowders, when benchmarked against the covolume-based rigorous calculations (Figure 3B, C). The $\Delta\mu$ results presented below for ellipsoidal crowders all had the correction of eq 11.

3.6. Fitting of $\Delta\mu$ Results to an Analytical Formula.

Our previous studies[5,18] showed that the dependence of $\Delta\mu$ on crowder volume fraction and on crowder size calculated for atomistic proteins can be fitted to the scaled particle theory, even though the latter was designed for convex test particles.[37] This motivated us to fit the results to the fundamental measure theory, which extends the scaled particle theory to convex crowders of arbitrary shapes, including ellipsoids.[38] The fundamental measure theory predicts the excess chemical potential as

dx.doi.org/10.1021/ct4005195 | J. Chem. Theory Comput. 2013, 9, 4633−4643

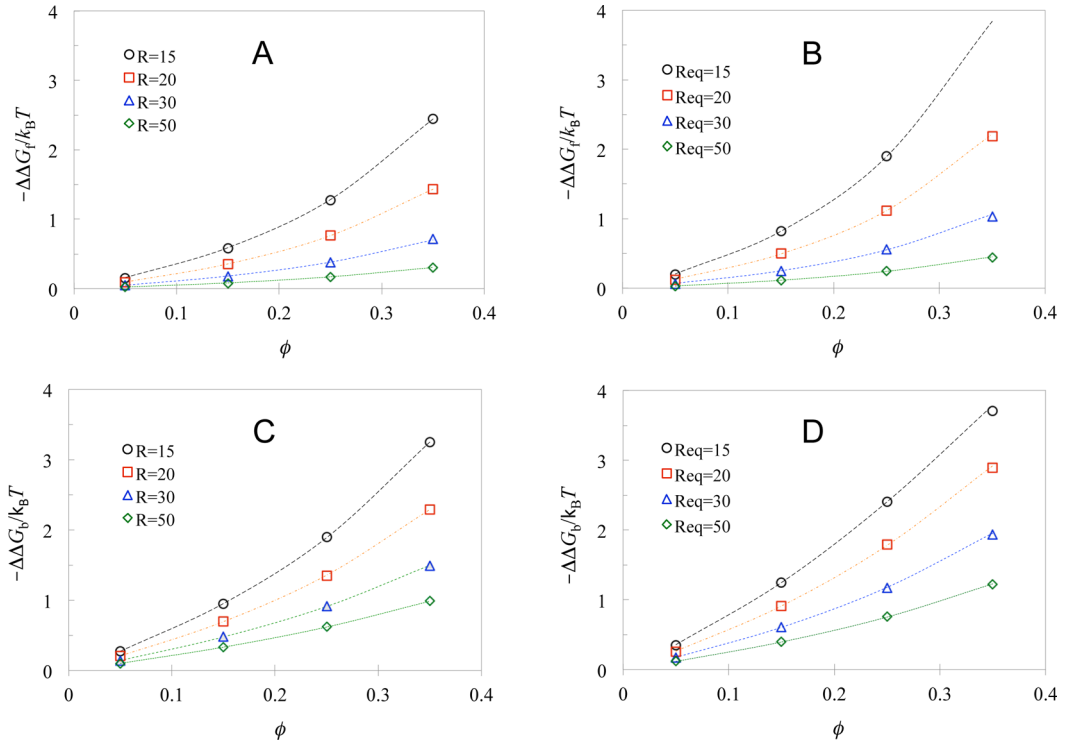


Figure 5. Shape-dependent effects of crowding on the folding stability of cytochrome b_{562} and on the binding stability of barnase and barstar. The results in (A) and (C) for the spherical crowders are from our previous study.[5] (B) and (D) display the results for the oblate ellipsoidal crowders obtained from the FFT-based method, and fits to the fundamental measure theory.

$$\frac{\Delta\mu}{k_BT} = -\ln(1-\phi) + A_1\phi/(1-\phi) + A_2[\phi/(1-\phi)]^2 + A_3[\phi/(1-\phi)]^3 \quad (12)$$

where ϕ is the crowder volume fraction and the coefficients are given by

$$A_1 v_c = s_c l_p + l_c s_p + v_p \quad (13)$$

$$A_2 v_c^2 = s_c^2 s_p/8\pi + l_c s_c v_p \quad (14)$$

$$A_3 v_c^3 = l_c^2 s_c^2 v_p/3 \quad (15)$$

In the last expressions, l_c, s_c, and v_c are the integration of the mean curvature over the surface, the surface area, and the volume of a crowder. These are $[b + a^2(|b^2 - a^2|)^{-1/2}\arccosh(b/a)]/2$, $2\pi a[a + b^2(|b^2 - a^2|)^{-1/2}\arccos(a/b)]$, and $4\pi a^2 b/3$, respectively, for a prolate ellipsoid; the results for an oblate ellipsoid are obtained when arccos and arccosh are exchanged. l_p, s_p, and v_p are the analogous quantities for the protein and were treated as floating parameters in the present work.

IV. RESULTS AND DISCUSSION

Biomacromolecules have a variety of shapes and sizes. There is a long tradition in modeling globular proteins and double-stranded DNA as ellipsoids in fitting experimental data, such as those from hydrodynamic and NMR measurements.[39−42] Disordered proteins and single-stranded DNA and RNA have extended conformations, which might be modeled as long ellipsoids. Dextran, a common crowding agent, has often been modeled as ellipsoids.[43−46]

While the method presented here is capable of treating arbitrarily shaped crowders, we use the application to ellipsoidal crowders to illustrate the basic idea and benchmark the accuracy and speed of its implementation. With ellipsoidal crowders, it is still possible to generate exact results by exhaustive enumeration for a limited number of cases. As shown in Figure 3, comparison to these exact results shows that the present FFT-based method is highly accurate. At the same time, ellipsoidal crowders are ideal for demonstrating the influences of crowder shape, as indicated by the results below.

4.1. Enhanced Effects of Crowding by Aspherical Shapes. A comparison of Figure 3A with Figure 3B,C already shows that the clash-free fractions for placing a protein into oblate or prolate ellipsoidal crowders are significantly lower than the counterpart for placing into the equivalent spherical crowders. Correspondingly, the excess chemical potentials for either the native or denatured cytochrome b_{562} interacting with the ellipsoidal crowders have significantly higher magnitudes than with the spherical crowders, as shown in Figure 4. Similarly, the magnitudes of the excess chemical potentials for barnase and barstar in either the bound or unbound state can be higher by several k_BTs while interacting with the ellipsoidal crowders than with the spherical crowders (not shown).

The observed shape effects are expected, because the equivalent spheres can be better packed than the ellipsoids, thus leaving more large voids for the placement of the test protein (see also ref 47).

4.2. Fitting of Excess Chemical Potential to Fundamental Measure Theory. Following our previous studies (Figure 4A, B),[5,18] we found here that the dependence of the excess chemical potential on crowder volume fraction and on crowder size can be fitted to the fundamental measure theory. The fit is illustrated in Figure 4C, D for the excess chemical

dx.doi.org/10.1021/ct4005195 | J. Chem. Theory Comput. 2013, 9, 4633−4643

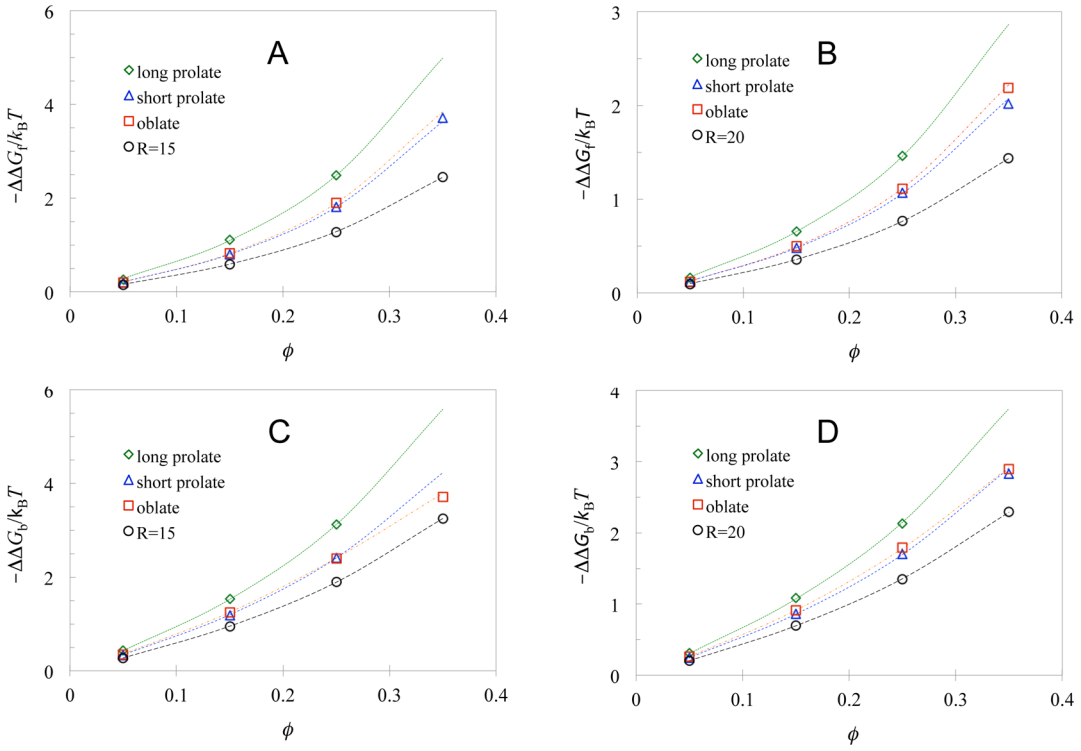


Figure 6. Additional shape-dependent results for $\Delta\Delta G_f$ and $\Delta\Delta G_b$. (A) $\Delta\Delta G_f$ for four different crowder shapes, all with an equivalent radius of 15 Å. (B) Corresponding results at an equivalent radius of 20 Å. (C) $\Delta\Delta G_b$ for four different crowder shapes, all with an equivalent radius of 15 Å. (D) Corresponding results at an equivalent radius of 20 Å.

potentials of the native and denatured cytochrome b_{562} interacting with the ellipsoidal crowders. In this fit (using data up to $\phi = 0.25$), for a given crowder shape, the integrated mean curvature, surface area, and volume of a test protein in an end state are treated as floating parameters. These "effective" parameters are not expected to be the same for all crowders. For native cytochrome b_{562} interacting with the oblate ellipsoidal crowders (axis ratio at 4:4:1), the values of l_p, s_p, and v_p are 18.8 Å, 5697 Å^2, and 10 439 Å^3, respectively. These are to be compared to the corresponding values of the protein interacting with the spherical crowders: 20.6 Å, 5165 Å^2, and 16 569 Å^3.[5] The l_p and v_p for the oblate ellipsoidal crowders are approximately 10% and 40% lower than for the spherical crowders, whereas s_p for the oblate ellipsoidal crowders is approximately 10% higher than for the spherical crowders. Still, the values of l_p, s_p, and v_p are in the range of magnitudes expected for the size, surface area, and volume of native cytochrome b_{562}. Similar crowder shape effects on the values of l_p, s_p, and v_p were observed for the denatured protein. These were 22.9 Å, 7362 Å^2, and 12 111 Å^3 for the oblate ellipsoidal crowders and 24.9 Å, 6873 Å^2, and 18 154 Å^3 for the spherical crowders.

An important application is that, with the fitted values for the geometric parameters of the test protein, the fundamental measure theory can be used to predict the excess chemical potential for crowding conditions where FFT-based calculations are not made. As an illustration, Figure 4C, D shows that the data for $\Delta\mu_N$ and $\Delta\mu_D$ produced by the FFT-based method at $\phi = 0.35$ agree well with the predictions of the parametrized fundamental measure theory, even though the parametrization excluded the data at $\phi = 0.35$. The value of the parametrized fundamental measure theory as a substitute of the FFT-based method can be especially important for calculating $\Delta\mu$ for denatured proteins at high ϕ. In these cases, the clash-free fractions have small magnitudes, and thus converged results would require fine grid spacings and many protein conformations, both of which increase the computational time of the FFT-based method.

For spherical crowders, the fit to the scaled particle theory further inspired us to develop a generalized fundamental measure theory.[7] In this semianalytical method for predicting the excess chemical potential, the protein geometric parameters l_p, s_p, and v_p are calculated *a priori*, based on a surface of the protein which encloses the region inaccessible to a spherical crowder with a given radius. The generalized fundamental measure theory does not require fitting to known $\Delta\mu$ results; that is, it makes *ab initio* predictions for $\Delta\mu$, not mere extrapolation to new conditions. We hope to develop a generalized fundamental measure theory for aspherical crowders in the future.

4.3. Shape-Dependent Effects of Crowding on Folding Free Energy. Figure 4 shows that, for a given crowding condition, the excess chemical potential for the denatured state is higher than the counterpart for the native state. This is because the conformations of the denatured protein are more open and hence result in lower clash-free fractions when placed into the distribution of crowders. Consequently, crowding produces a decrease in the folding free energy, corresponding to stabilization of the native state.

As can be anticipated from Figure 4 and shown in Figure 5A, B, the stabilization effects of the ellipsoidal crowders are stronger than those of the equivalent spherical crowders. At a volume fraction of 0.35, the oblate ellipsoidal crowders with an equivalent radius of 15 Å produced a stabilization effect of 3.8 k_BT. In comparison, the 15-Å spherical crowders produced a stabilization effect of 2.4 k_BT, which is one-third less.

dx.doi.org/10.1021/ct4005195 | J. Chem. Theory Comput. 2013, 9, 4633−4643

To further investigate the influences of crowder shape, we calculated $\Delta\Delta G_f$ for two prolate ellipsoidal crowders, with axis ratios at 1:1:4 and 1:1:8, respectively. As shown in Figure 6A, B, with equal crowder volume, the short prolate crowders exerted similar stabilization as the oblate crowders (both were stronger stabilizers than the spherical crowders). As the crowder shape became more elongated (as in the case of the long prolate crowders), the stabilization effect grew even stronger.

4.4. Shape-Dependent Effects of Crowding on Binding Free Energy. The shape-dependent effects on the binding free energy of barnase and barstar were similar to those on the folding free energy (Figure 5C, D). At ϕ = 0.35, the oblate ellipsoidal crowders with an equivalent radius of 15 Å produced a stabilization effect of 3.7 k_BT, which is to be compared with the stabilization effect of 3.2 k_BT produced by the equivalent spherical crowders. Figure 6C, D further shows that the oblate crowders and the short prolate crowders had similar effects but stabilization grew stronger with the long prolate crowders.

The forgoing results regarding crowder shape effects in protein folding and binding stability reinforce observations from previous direct simulations. Chritiansen et al.[15] simulated the folding free energy surface of a coarse-grained cytochrome c molecule (two beads per residue) in the presence of either a spherical crowder or a two-sphere dumbbell crowder with the same total volume. Consistent with the results in Figures 5A, B and 6A, B, they found that this "thin" dumbbell crowder exerted greater stabilization of cytochrome c. In comparison, a "fat" dumbbell crowder with twice the volume of the spherical crowder, when present at the same total volume fraction, exerted less stabilization than the spherical crowder. The latter observation can be understood because the fat dumbbell crowder is essentially a pair of spherical crowders but much better packed than most pairs found in the simulation with the spherical crowder. A similar study[14] a year earlier by the same groups on apoflavodoxin reached the opposite conclusion: the spherical crowder exerted more stabilization than the thin dumbbell crowder and less stabilization than the fat dumbbell crowder. This result seems difficult to rationalize.

Two other studies further highlight packing as the overriding factor in the shape effects of repulsive crowders. Kudley et al.[17] carried out simulations of a bead model of a homopolymer in the presence of a rod-like crowder (represented as a string of beads). This crowder favored a long, single helix for the homopolymer over a helix bundle, since with the former conformation the rod-like crowders could align along the polymer axis and avoid clashes between the crowders. In another study, O'Brien et al.[16] investigated the conformational preference of a peptide dimer (represented at the all-atom level) under crowding. Crowding was found to favor disordered dimeric aggregates over dimer β-sheets, because in this case the former was more compact, and the effects of rod-like crowders were greater than spherical crowders.

4.5. Extension to All-Atom Crowders. Up to now, we have used ellipsoidal crowders to demonstrate the accuracy and feasibility of the FFT-based method. Because we map the crowders to a grid (Figure 2A), the speed of our method is determined by the total number of grid points in the unit cell, not directly by the level of details in representing the crowders. The latter could influence the grid spacing needed to accurately map the crowders to the grid; that is, a more detailed representation of the crowders could require a finer grid spacing. Now we demonstrate that the FFT-based method is still very feasible for crowders represented at the all-atom level.

As crowders, we consider dextran, which is commonly used as a crowding agent in experimental studies. We generated an all-atom model for a dextran molecule with a molecular mass of 9923.8 Da (unpublished). We then randomly placed 20 copies of this molecule in a cubic unit cell with side length of 150 Å (while ensuring that the molecules did not clash), corresponding to a concentration of 98 g/L (Figure 7A). The random placement of the crowders was repeated, producing 10 independent crowder configurations. The test protein is native cytochrome b_{562}.

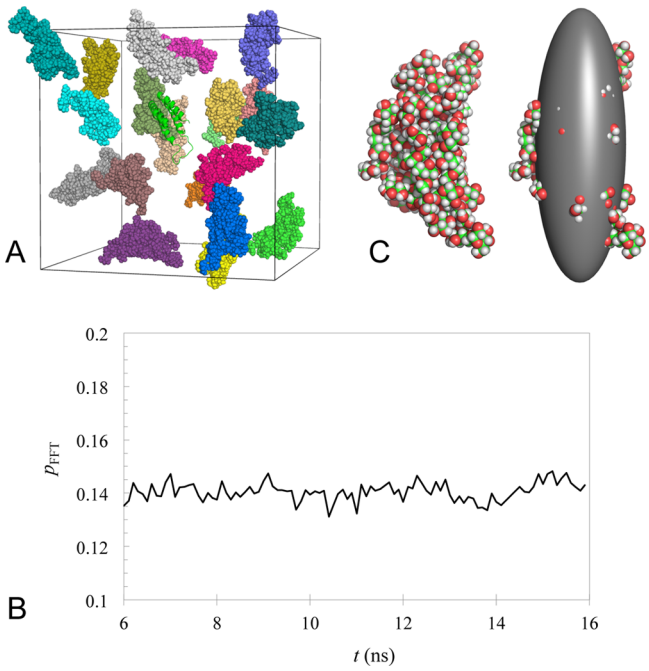


Figure 7. Application of the FFT-based method to native cytochrome b_{562} interacting with dextran, both represented at the all-atom level. (A) A cytochrome b_{562} molecule placed into a unit cell containing 20 copies of dextran with a molecular mass of 9923.8 Da. (B) Clash-free fractions calculated by the FFT-based method. For each of 100 snapshots along the trajectory of native cytochrome b_{562}, the clash-free fraction averaged over 10 crowder configurations is displayed. A grid spacing of 0.3 Å was used; for a single protein conformation interacting with a single crowder configuration, the computational time on a single core was 42.5 s. (C) The dextran molecule and the prolate ellipsoidal model, with semiaxes of 12, 12, and 36 Å.

We benchmarked the accuracy and computational time of the FFT-based method against two other methods, using a single protein conformation and a single crowder configuration. The first alternative method was to calculate the correlation function of eq 2 directly, instead of using FFT. The second was to exhaustively enumerate all pairs of atoms between the protein and the crowders. We refer to these as the direct-correlation method and atom-based method, respectively. The FFT-based method and direct-correlation method should produce identical results, which we confirmed. The major source of errors of these two methods was from mapping the crowders to the grid. This source of errors was absent in the atom-based method, but all the three methods had a minor source of errors, from placing the test protein on the grid.

With grid spacings of 0.75, 0.6, 0.5 Å for placing the test protein, the atom-based method yielded the same clash-free fraction of 0.141. In comparison, the FFT-based method

produced clash-free fractions in the narrow range from 0.129 to 0.141 with grid spacings of 0.75, 0.6, 0.5, 0.3, and 0.15 Å, suggesting that any of these grid spacings can be used for accurately calculating the clash-free fraction.

The computational times of the three methods were very different. With a 0.6-Å grid spacing, the timing on a single core was 2.7 s for the FFT-based method, 6274.6 s for the direct-correlation method, and 73295.2 s for the atom-based method.

In Figure 7B, we show the clash-free fractions averaged over the 10 crowder configurations for placing 100 conformations sampled from a trajectory of native cytochrome b_{562}. The clash-free fraction further averaged over the protein conformations was 0.141 ± 0.009, corresponding to a $\Delta\mu$ of 1.16 ± 0.04 kcal/mol.

We modeled the dextran molecule as a prolate ellipsoid, with semiaxes of 12, 12, and 36 Å (Figure 7C). Applying the fundamental measure theory parametrized for the short prolate crowders to this ellipsoidal model for dextran, we predicted an average clash-free fraction of 0.153. This value agrees well with result of the FFT-based method.

We also carried out calculations for Debye−Hückel and van der Waals types of protein−crowder interactions:

$$U_{\mathrm{DH}} = \sum_{i,j} q_i q_j \exp(-\kappa r_{ij})/\varepsilon r_{ij}) \tag{16}$$

$$U_{\mathrm{vdW}} = -\sum_{i,j} \xi B_{ij}/r_{ij}^6 \tag{17}$$

where κ is the Debye−Hückel screening parameter (calculated for a 1:1 salt at 50 mM); ε is the effective dielectric constant of the crowded solution (set at 39.3, i.e., half the value for water[48]); and ξ is a scaling constant (set at 0.2). We still imposed the hard-core repulsion (just as in the calculations of clash-free fractions); that is, the above "soft" interactions were included only for grid points that did not result in any protein−crowder clash. This explains the absence of the usual $1/r_{ij}^{12}$ term in eq 17. In this way, we obtained $\Delta\Delta\mu$, the additional excess chemical potential due to the inclusion of the soft interactions on top of the hard-core repulsion. The partial charges (q_i and q_j) and Lennerd-Jones parameters (B_{ij}) were taken from the Amber force field.[49] Since the purpose of the present study was to benchmark the FFT-based method, no attempt was made to tune any of the parameters. We used the combination rule $B_{ij} = (B_{ii}B_{jj})^{1/2}$ to allow for the calculation of U_{vdW} by FFT. The q_j and $B_{jj}^{1/2}$ values for each protein atom were distributed to the nearest eight grid points according to linear interpolation.[27,28]

In Figure 8, we show the $\Delta\Delta\mu$ results for the 100 conformations of cytochrome b_{562} due to either the Debye−Hückel or van der Waals interactions, averaged over the 10 configurations of the dextran molecules (grid spacing at 0.3 Å). Upon further averaging over the protein conformations, $\Delta\Delta\mu$ was −0.52 ± 0.03 kcal/mol for the Debye−Hückel interactions and −0.75 ± 0.04 kcal/mol for the van der Waals interactions. While these values depend on the parameters used, they do indicate that the soft interactions can counterbalance the effect of hard-core repulsion, in line with recent studies.[11,22]

With a 0.6-Å grid spacing, for a single protein conformation and a single crowder configuration, the time to calculate $\Delta\Delta\mu$ by the FFT-based method on a single core was 2.6 s for either the Debye−Hückel or van der Waals interactions, essentially unchanged from that for calculating the clash-free fraction. In comparison, by the atom-based method the time was 117 404.7 s for the Debye−Hückel interactions and 94 533.0 s for the van der Waals interactions. The speedup by the FFT-based method is ∼40 000-fold.

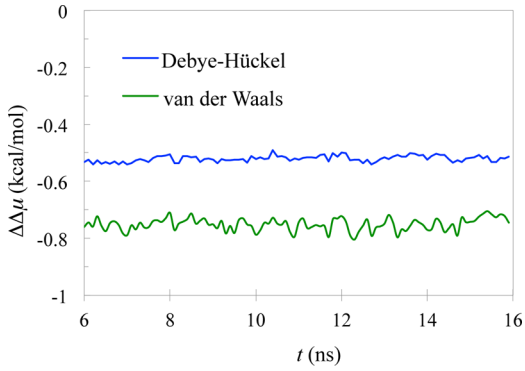


Figure 8. Additional excess chemical potential of native cytochrome b_{562}, due to electrostatic or van der Waals interactions with dextran molecules. Details of the FFT-based calculations are similar to those in Figure 7B.

McGuffee and Elcock[22] carried out Brownian dynamics simulations of 50 types of macromolecules to generate an atomically detailed model of the bacterial cytoplasm. They then used an approach conceptually similar to postprocessing but implemented by an atom-based method to calculate effects of the cytoplasm on folding and binding stability by considering electrostatic and van der Waals interactions between the test protein and crowders. They noted "very significant computational expense" associated with these calculations, similar to the timings of the atom-based method listed above. The 40 000-fold speed up illustrated here indicates that the FFT-based method can significantly accelerate realistic modeling of protein folding and binding in cell-like environments.

To apply the FFT-based method, one still needs to generate realistic configurations of the crowded solution. While that might be relatively easy to do when the crowders do not have any attractive interactions with each other (as was the case for the ellipsoidal crowders and dextran molecules considered here), it might not be easy to do when the crowders can attract each other. In such cases, one may need to do dynamic simulations in order to sample configurations where the crowders are associated with each other. The work of McGuffee and Elcock[22] represents the state of the art in simulations of cell-like crowded solution. We note that, once such simulations have been made, they can be used for studying many different kinds of test proteins, by, for example, the FFT-based method.

V. CONCLUSIONS

We have presented a general method for calculating the excess chemical potential of a test protein interacting with arbitrarily shaped crowders, including those represented at the all-atom level. The accuracy and feasibility of this method have been demonstrated on both ellipsoidal and all-atom crowders. It was found, in agreement with previous theoretical and simulation studies, that crowder shape has a significant influence on the magnitude of crowding effects on protein folding and binding stability. For the hard-core repulsive type of protein-crowder interactions, increasing crowder asphericity is associated with higher ability to break the free space into small pieces, leading to higher excess chemical potentials and stronger crowding effects. With soft interactions added, crowder shape effects can be expected to become much more complex; the results presented here provide a glimpse.

It is well-known that the shapes of biomacromolecules are a key factor in their functions. The present study suggests that shape not only plays a direct role in determining the functions of participating macromolecules but also an indirect role via influencing the magnitude of crowding effects when the macromolecules serve as crowders.

Cellular environments are composed of macromolecules with a variety of sizes and shapes. The FFT-based method will be able to treat such complex conditions without strong increase in computational cost, due to the fact that the crowder molecules are mapped to a grid before the most expensive calculations take place.

AUTHOR INFORMATION

Corresponding Author

*Phone: 850-645-1336. Fax: 850-644-7244. E-mail: hzhou4@fsu.edu.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This work was supported by National Institutes of Health Grant No. GM88187.

REFERENCES

(1) Zhou, H. X.; Rivas, G.; Minton, A. P. Macromolecular crowding and confinement: Biochemical, biophysical, and potential physiological consequences. *Annu. Rev. Biophys.* **2008**, *37*, 375−397.

(2) Zhou, H. X. Influence of crowded cellular environments on protein folding, binding, and oligomerization: Biological consequences and potentials of atomistic modeling. *FEBS Lett.* **2013**, *587*, 1053−1061.

(3) Cheung, M. S.; Klimov, D.; Thirumalai, D. Molecular crowding enhances native state stability and refolding rates of globular proteins. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 4753−4758.

(4) Minh, D. D.; Chang, C. E.; Trylska, J.; Tozzini, V.; McCammon, J. A. The influence of macromolecular crowding on HIV-1 protease internal dynamics. *J. Am. Chem. Soc.* **2006**, *128*, 6006−6007.

(5) Qin, S.; Zhou, H. X. Atomistic modeling of macromolecular crowding predicts modest increases in protein folding and binding stability. *Biophys. J.* **2009**, *97*, 12−19.

(6) Qin, S.; Minh, D. D. L.; McCammon, J. A.; Zhou, H.-X. Method to predict crowding effects by postprocessing molecular dynamics trajectories: application to the flap dynamics of HIV-1 protease. *J. Phys. Chem. Lett.* **2010**, *1*, 107−110.

(7) Qin, S.; Zhou, H.-X. Generalized fundamental measure theory for atomistic modeling of macromolecular crowding. *Phys. Rev. E* **2010**, *81*, 031919.

(8) Dong, H.; Qin, S.; Zhou, H. X. Effects of macromolecular crowding on protein conformational changes. *PLoS Comput. Biol.* **2010**, *6*, e1000833.

(9) Mittal, J.; Best, R. B. Dependence of protein folding stability and dynamics on the density and composition of macromolecular crowders. *Biophys. J.* **2010**, *98*, 315−320.

(10) Dhar, A.; Samiotakis, A.; Ebbinghaus, S.; Nienhaus, L.; Homouz, D.; Gruebele, M.; Cheung, M. S. Structure, function, and folding of phosphoglycerate kinase are strongly perturbed by macromolecular crowding. *Proc. Natl. Acad. Sci. U.S.A.* **2010**, *107*, 17586−17591.

(11) Rosen, J.; Kim, Y. C.; Mittal, J. Modest protein-crowder attractive interactions can counteract enhancement of protein association by intermolecular excluded volume interactions. *J. Phys. Chem. B* **2011**, *115*, 2683−2689.

(12) Kang, M.; Roberts, C.; Cheng, Y. H.; Chang, C. E. A. Gating and intermolecular interactions in ligand−protein association: Coarse-grained modeling of HIV-1 protease. *J. Chem. Theory Comput.* **2011**, *7*, 3438−3446.

(13) Nagarajan, S.; Amir, D.; Grupi, A.; Goldenberg, D.; Minton, A.; Haas, E. Modulation of functionally significant conformational equilibria in adenylate kinase by high concentrations of trimethylamine oxide attributed to volume exclusion. *Biophys. J.* **2011**, *100*, 2991−2999.

(14) Homouz, D.; Stagg, L.; Wittung-Stafshede, P.; Cheung, M. S. Macromolecular crowding modulates folding mechanism of α/β protein apoflavodoxin. *Biophys. J.* **2009**, *96*, 671−680.

(15) Christiansen, A.; Wang, Q.; Samiotakis, A.; Cheung, M. S.; Wittung-Stafshede, P. Factors defining effects of macromolecular crowding on protein stability: An in vitro/in silico case study using cytochrome c. *Biochemistry* **2010**, *49*, 6519−6530.

(16) O'Brien, E. P.; Straub, J. E.; Brooks, B. R.; Thirumalai, D. Influence of nanoparticle size and shape on oligomer formation of an amyloidogenic peptide. *J. Phys. Chem. Lett.* **2011**, *2*, 1171−1177.

(17) Kudlay, A.; Cheung, M. S.; Thirumalai, D. Influence of the shape of crowding particles on the structural transitions in a polymer. *J. Phys. Chem. B* **2012**, *116*, 8513−8522.

(18) Batra, J.; Xu, K.; Qin, S.; Zhou, H. X. Effect of macromolecular crowding on protein binding stability: Modest stabilization and significant biological consequences. *Biophys. J.* **2009**, *97*, 906−911.

(19) Batra, J.; Xu, K.; Zhou, H.-X. Nonaddtive effects of mixed crowding on protein stability. *Proteins* **2009**, *77*, 133−138.

(20) Jiao, M.; Li, H. T.; Chen, J.; Minton, A. P.; Liang, Y. Attractive protein−polymer interactions markedly alter the effect of macromolecular crowding on protein association equilibria. *Biophys. J.* **2010**, *99*, 914−923.

(21) Miklos, A. C.; Sarkar, M.; Wang, Y.; Pielak, G. J. Protein crowding tunes protein stability. *J. Am. Chem. Soc.* **2011**, *133*, 7116−7120.

(22) McGuffee, S. R.; Elcock, A. H. Diffusion, crowding, and protein stability in a dynamic molecular model of the bacterial cytoplasm. *PLoS Comput. Biol.* **2010**, *6*, e1000694.

(23) Johansen, D.; Jeffries, C. M. J.; Hammouda, B.; Trewhella, J.; Goldenberg, D. P. Effects of macromolecular crowding on an intrinsically disordered protein characterized by small-angle neutron scattering with contrast matching. *Biophys. J.* **2011**, *100*, 1120−1128.

(24) Zhou, H. X.; Qin, S. Simulation and modeling of crowding effects on the thermodynamic and kinetic properties of proteins with atomic details. *Biophys Rev.* **2013**, *5*, 207−215.

(25) Qin, S.; Mittal, J.; Zhou, H. X. Folding free energy surfaces of three small proteins under crowding: Validation of the postprocessing method by direct simulation. *Phys. Biol.* **2013**, *10*, 045001.

(26) Katchalski-Katzir, E.; Shariv, I.; Eisenstein, M.; Friesem, A. A.; Aflalo, C.; Vakser, I. A. Molecular surface recognition: determination of geometric fit between proteins and their ligands by correlation techniques. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 2195−2199.

(27) Harrison, R. W.; Kourinov, I. V.; Andrews, L. C. The Fourier−Green's function and the rapid evaluation of molecular potentials. *Protein Eng.* **1994**, *7*, 359−369.

(28) Bliznyuk, A. A.; Gready, J. E. Simple method for locating possible ligand binding sites on protein surfaces. *J. Comput. Chem.* **1999**, *20*, 983−988.

(29) Mandell, J. G.; Roberts, V. A.; Pique, M. E.; Kotlovyi, V.; Mitchell, J. C.; Nelson, E.; Tsigelny, I.; Ten Eyck, L. F. Protein docking using continuum electrostatics and geometric fit. *Protein Eng.* **2001**, *14*, 105−113.

(30) Kozakov, D.; Brenke, R.; Comeau, S. R.; Vajda, S. PIPER: An FFT-based protein docking program with pairwise potentials. *Proteins* **2006**, *65*, 392−406.

(31) Frigo, M.; Johnson, S. G. The design and implementation of FFTW3. *Proc. IEEE* **2005**, *93*, 216−231.

(32) Donev, A.; Torquato, S.; Stillinger, F. H. Neighbor list collision-driven molecular dynamics simulation for nonspherical hard particles. II. Applications to ellipses and ellipsoids. *J. Comput. Phys.* **2005**, *202*, 765−793.

(33) Tjong, H.; Zhou, H. X. The folding transition-state ensemble of a four-helix bundle protein: helix propensity as a determinant and macromolecular crowding as a probe. *Biophys. J.* **2010**, *98*, 2273−2280.

(34) Tjong, H.; Zhou, H. X. Accurate calculations of binding, folding, and transfer free energies by a scaled generalized Born method. *J. Chem. Theory Comput.* **2008**, *4*, 1733−1744.
(35) Bondi, A. van der Waals volumes and radii. *J. Phys. Chem.* **1964**, *68*, 441−451.
(36) Zheng, X.; Palffy-Muhoray, P. Distance of closest approach of two arbitrary hard ellipses in two dimensions. *Phys. Rev. E* **2007**, *75*, 061709.
(37) Lebowitz, J. L.; Helfand, E.; Praestgaard, E. Scaled particle theory of fluid mixtures. *J. Chem. Phys.* **1965**, *43*, 774−779.
(38) Oversteegen, S. M.; Roth, R. General methods for free-volume theory. *J. Chem. Phys.* **2005**, *122*, 214502.
(39) Cantor, C. R.; Schimmel, P. R., *Biophysical Chemistry*. W. H. Freeman: San Francisco, 1980.
(40) R Taylor, W.; M Thornton, J.; G Turnell, W. An ellipsoidal approximation of protein shape. *J. Mol. Graph.* **1983**, *1*, 30−38.
(41) Ryabov, Y. E.; Geraghty, C.; Varshney, A.; Fushman, D. An efficient computational method for predicting rotational diffusion tensors of globular proteins using an ellipsoid representation. *J. Am. Chem. Soc.* **2006**, *128*, 15432−15444.
(42) Schwieters, C. D.; Clore, G. M. A pseudopotential for improving the packing of ellipsoidal protein structures determined from NMR data. *J. Phys. Chem. B* **2008**, *112*, 6070−6073.
(43) Bohrer, M. P.; Deen, W. M.; Robertson, C. R.; Troy, J. L.; Brenner, B. M. Influence of molecular configuration on the passage of macromolecules across the glomerular capillary wall. *J. Gen. Physiol.* **1979**, *74*, 583−593.
(44) Nordmeier, E. Static and dynamic light-scattering solution behavior of pullulan and dextran in comparison. *J. Phys. Chem.* **1993**, *97*, 5770−5785.
(45) Wills, P. R.; Comper, W. D.; Winzor, D. J. Thermodynamic nonideality in macromolecular solutions: interpretation of virial coefficients. *Arch. Biochem. Biophys.* **1993**, *300*, 206−212.
(46) Liu, Z.; Weng, W.; Bookchin, R. M.; Lew, V. L.; Ferrone, F. A. Free energy of sickle hemoglobin polymerization: a scaled-particle treatment for use with dextran as a crowding agent. *Biophys. J.* **2008**, *94*, 3629−3634.
(47) Minton, A. P. Excluded volume as a determinant of macromolecular structure and reactivity. *Biopolymers* **1981**, *20*, 2093−2120.
(48) Tjong, H.; Zhou, H. X. Prediction of protein solubility from calculation of transfer free energy. *Biophys. J.* **2008**, *95*, 2601−2609.
(49) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. A second generation force field for the simulation of proteins, nucleic acids, and organic molecules. *J. Am. Chem. Soc.* **1995**, *117*, 5179−5197.