Potential of mean force between a large solute and a biomolecular complex: A model analysis on protein flux through chaperonin system

Ken-ich Amano, Hiraku Oshima, and Masahiro Kinoshita

Citation: J. Chem. Phys. **135**, 185101 (2011); doi: 10.1063/1.3657856
View online: http://dx.doi.org/10.1063/1.3657856
View Table of Contents: http://jcp.aip.org/resource/1/JCPSA6/v135/i18
Published by the American Institute of Physics.

Potential of mean force between a large solute and a biomolecular complex: A model analysis on protein flux through chaperonin system

Ken-ich Amano,[1] Hiraku Oshima,[2] and Masahiro Kinoshita[2,a)]
[1]*Graduate School of Energy Science, Kyoto University, Uji, Kyoto 611-0011, Japan*
[2]*Institute of Advanced Energy, Kyoto University, Uji, Kyoto 611-0011, Japan*

(Received 7 August 2011; accepted 13 October 2011; published online 9 November 2011)

Insertion of a large solute into an even larger vessel comprising biopolymers followed by release of the same solute from it is one of the important functions sustaining life. As a typical example, an unfolded protein is inserted into a chaperonin from bulk aqueous solution, a cochaperonin acting as a lid is attached to the chaperonin rim and the protein folds into its native structure within the closed cavity, the cochaperonin is detached after the folding is finished, and the folded protein is released back to the bulk solution. On the basis of the experimental observations manifesting that the basic aspects of the protein flux through the chaperonin system is independent of the chaperonin, cochaperonin, and protein species, we adopt a simple model system with which we can cover the whole cycle of the protein flux. We calculate the spatial distribution of the solvent-mediated potential of mean force (PMF) between a spherical solute and a cylindrical vessel or vessel/lid complex. The calculation is performed using the three-dimensional integral equation theory, and the PMF is decomposed into energetic and entropic components. We argue that an unfolded protein with a larger excluded volume (EV) and weak hydrophobicity is entropically inserted into the chaperonin cavity and constrained within a small space almost in its center. The switch from insertion to release is achieved by decreasing the EV and turning the protein surface hydrophilic in the folding process. For this release, in which the energetic component is a requisite, the feature that the chaperonin inner surface in the absence of the cochaperonin is not hydrophilic plays essential roles. On the other hand, the inner surface of the chaperonin/cochaperonin complex is hydrophilic, and the protein is energetically repelled from it: The protein remains constrained within the small space mentioned above without contacting the inner surface for correct folding. The structural and inner-surface properties of the chaperonin or complex are controlled by the adenosine triphosphate (ATP) binding to the chaperonin, hydrolysis of ATP into adenosine diphosphate (ADP) and Pi, and dissociation of ADP and Pi. The function of the chaperonin system is exhibited by synchronizing the chemical cycle of ATP hydrolysis with hydration properties of a protein in the water confined on the scale of a nanometer which are substantially different from those in the bulk water.
[doi:10.1063/1.3657856]

I. INTRODUCTION

A living organism is sustained by a variety of biological functions. Insertion of a large solute into an even larger vessel comprising biopolymers followed by release of the same solute from it is one of such functions. As a typical example, an unfolded protein is inserted into a chaperonin from bulk aqueous solution, protein folding occurs within it, and the folded protein is released back to the bulk solution.[1–7] Thus, chaperonin assists a protein to fold into its native structure. Otherwise, the unfolded proteins may subject to the aggregation preventing correct folding. In a chaperonin, the subunits form a cylindrical, ring-like structure enclosing a central cavity[1–8] within which protein folding occurs. There is a class of chaperonins functioning in conjunction with cochaperonins.[1–7] Chaperonin is an ATPase which utilizes the cycle comprising the binding of adenosine triphosphate (ATP) to it, hydrolysis of ATP into ADP (adenosine diphosphate) and Pi, and dissociation of ADP and Pi from it.[1–7]

A cochaperonin, which is attached to the chaperonin rim, acts as a lid and a protein folds into its native structure within the closed cavity of the chaperonin/cochaperonin complex. Two complexes, GroEL/GroES (Refs. 1 and 3–8) and Hsp60/Hsp10,[2,9] are well known and have been investigated rather extensively. GroEL works as a double-ring complex[1,3–7] in which the two rings act alternately to insert, encapsulate, and fold a variety of unfolded proteins. Protein folding occurs even within a single ring as shown in experimental studies, but the folded protein cannot be released in general because the dissociation of ADP and Pi followed by detachment of GroES is driven by a signal transmitted via the binding of ATP to the adjacent ring.[7] Unlike GroEL/GroES, Hsp60/Hsp10 fully exhibits its function without the formation of double rings.[2,9] Despite this difference, Hsp60/Hsp10 gives assistance to folding of essentially all proteins that are dependent on GroEL/GroES.[9] We are concerned not with the difference between GroEL/GroES and Hsp60/Hsp10, but

a) Author to whom correspondence should be addressed. Electronic mail: kinoshit@iae.kyoto-u.ac.jp.

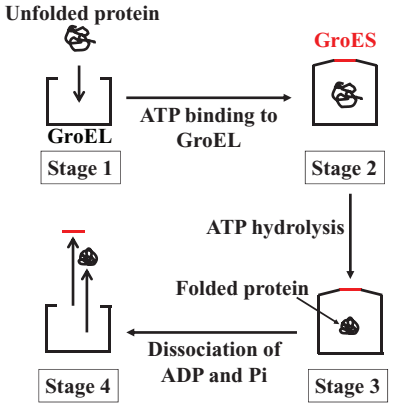

FIG. 1. One cycle during which a protein molecule is assisted by chaperonin/cochaperonin (GroEL/GroES) to fold into the correct native structure.

with the mechanism of the protein flux common in these two complexes.

A single ring is treated in the present study because it is best suited to the extraction of fundamental physics. The experimentally known properties of chaperonin that assist protein folding can be summarized as follows (see Fig. 1). Hereafter, GroEL/GroES is considered but the same explanation is valid for Hsp60/Hsp10 as well (simply replace GroEL and GroES by Hsp60 and Hsp10, respectively).

Stage 1: The GroEL vessel has a cavity and its inner surface is rather (weakly) hydrophobic in the open state.[5] An unfolded protein is inserted into the cavity from bulk aqueous solution and constrained within a small space almost in the center of the GroEL cavity.[10,11]

Stage 2: ATP binds to GroEL and this ATP binding causes a substantial structural change of GroEL so that it can be amenable to the attachment of GroES acting as a lid to the GroEL rim.[1,3–7] The inner surface of chaperonin turns hydrophilic and that of GroES is also hydrophilic.[5] The protein remains constrained within the small space almost in the center of the cavity.[11,12]

Stage 3: The protein, which is well separated from the inner surface of the GroEL/GroES complex, securely folds into its native structure within the hydrophilic chamber while ATP is hydrolyzed into ADP and Pi.[1,3–7]

Stage 4: After the folding is accomplished, ADP and Pi are dissociated from GroEL with the result that the detachment of GroES takes place.[1,3–7] The structure of GroEL returns to the one in the initial state whose cavity is open and inner surface is weakly hydrophobic. The folded protein is then released back to the bulk solution.[1,3–7]

Stages 1–4 form one cycle of the protein flux through the chaperonin system. Each cycle, which assists the folding of one protein molecule, accompanies the ATP binding, hydrolysis of ATP into ADP and Pi, and dissociation of ADP and Pi; in each cycle, the original configuration of chaperonin is recovered and the cycle is repeatable.

There is a prevailing view that an unfolded protein first binds to a hydrophobic patch on the interior rim of GroEL, and it is ejected into the cavity after the patch is removed by the GroEL structural change arising from the ATP binding followed by attachment of GroES.[1,3–7,11,13] In the course of this process, a partially unfolded (i.e., misfolded) protein, which is in a metastable state, could become completely unfolded for refolding to the correct native structure.[13] In the present study, however, we show that an unfolded protein can be inserted directly into the GroEL cavity and constrained within a small space almost in the center of the cavity, which is referred to as stage 1. A more detailed discussion on this point is given in Sec. III F.

To the best of our knowledge, no theoretical and computer simulation works have tackled the investigation of underlying mechanism of the *whole* cycle mentioned above. One might think that the mechanism can be elucidated only if details of the polyatomic structures of the protein and GroEL/GroES complex are taken into account by employing, for instance, a molecular dynamics (MD) simulation with all-atom potentials. With such an approach, however, only a single stage in the cycle can be explored with feasible computational effort. As typical examples, the characteristics of the folding process of a protein within the GroEL cavity have been studied using MD simulations.[14,15]

In order to cover the whole cycle for the first time in the present study, we adopt the simplest possible model that still captures essential physics of the protein flux through the chaperonin system. Specifically, we wish to answer the following questions: (i) How is the unfolded protein inserted into the cavity in stage 1?; (ii) Why is the folded protein released from the cavity in stage 4 despite that the unfolded one undergoes the insertion (how is the switch from insertion to release realized)?; (iii) Why should the inner surface be weakly hydrophobic in stages 1 and 4 (what is the significance of this weak hydrophobicity)?; (iv) Why should the inner surface be hydrophilic in stages 2 and 3?; (v) What is the role of GroES?; and (vi) How does the protein remain well separated from the inner surface for correct folding? In recent letters,[16,17] we have briefly provided physical insights into questions (i) and (ii), but in the present study we give the answers to all the six questions consistently on the basis of results of extensive parametric studies performed for the whole cycle of the protein flux.

Diverse proteins are inserted into the chaperonin cavity, assisted to fold into the correct native structures, and released back to the bulk solution. Moreover, despite the differences between GroEL/GroES and Hsp60/Hsp10 in terms of the details of their structures and apparent mechanisms, the latter gives assistance to folding of essentially all proteins that are dependent on the former.[9] It follows that the basic mechanism of the whole cycle, which is shared by these two complexes, must be independent of specific, chemical and structural characteristics of each protein. Only the properties of unfolded or folded molecules which are common in those proteins are crucially important. In our view, hydration properties of the protein are changed by the alteration of its conformation, and this plays crucially important roles in the switch from insertion to release (see stages 1 and 4). A protein becomes

much more compact upon folding. Further, the exposed surface of an unfolded protein comprises hydrophobic groups as well as hydrophilic groups, but the protein becomes dominantly hydrophilic after the folding is finished because hydrophobic groups are more preferentially buried. In our simple system, the unfolded and folded proteins are modeled as a larger, weakly solvophobic sphere and a smaller, solvophilic sphere, respectively. Likewise, we are concerned not with the specific differences between GroEL/GroES and Hsp60/Hsp10, but with the common properties shared by them. The chaperonin and cochaperonin are modeled as a vessel with cylindrical shape and a lid, respectively. We study the effects due to the magnitude of excluded volume (EV) generated by the spherical solute and the strength of solvophobicity or solvophilicity of the spherical solute and inner surfaces of the vessel and vessel/lid complex. Here, the EV is the volume of the space which the centers of solvent particles cannot enter, and a larger solute generates larger EV.

We show that fundamental physics can be understood through a simple model focused on solvation properties of a solute in the solvent confined on the scale of a nanometer which are substantially different from those in the bulk solvent. We analyze the potential of mean force (PMF) between the spherical solute and the vessel or vessel/lid complex, which are immersed in small spheres forming the solvent. The analysis is made using the three-dimensional (3D) integral equation theory (IET).[16–26] In one calculation, the 3D-IET gives the spatial distribution of the PMF, whereas a computer simulation (e.g., an MD simulation) gives only the value of the PMF on a single position. Further, in the theory, the PMF can readily be decomposed into entropic and energetic components which give physical insights into the PMF characteristics. It is, then, shown that the whole cycle comprising stages 1–4 can be explained *consistently* within the same theoretical framework.

II. MODEL AND THEORY

A. Model and parameter setting

In our model system, a large sphere with diameter d_B (solute 1) and a cylindrical vessel (solute 2) are immersed in the solvent at infinite dilution. A lid is either detached from or attached to the vessel (the vessel/lid complex is referred to as solute 3). The structural details of solutes 2 and 3 are illustrated in Figs. 2(a) and 2(b), respectively. The solvent is modeled as small spheres with diameter d_S and bulk density ρ_S interacting through the potential,

$$\begin{aligned} u_{SS}(r) &= \infty \quad \text{for} \quad r < d_S, \\ u_{SS}(r) &= -\varepsilon(d_S/r)^6 \quad \text{for} \quad r > d_S, \end{aligned} \tag{1}$$

where the subscript "S" denotes the solvent and r is the distance between the centers of two small spheres. In potential (1), the repulsive part of the Lennard-Jones potential is replaced by a hard-core interaction. $\rho_S d_S^3$ is set at the value for water under the normal condition, 0.7317 ($d_S = 0.28$ nm). $\varepsilon/(k_B T)$ (k_B is Boltzmann's constant) and the absolute temperature T are set at 1.0 and 298 K, respectively.

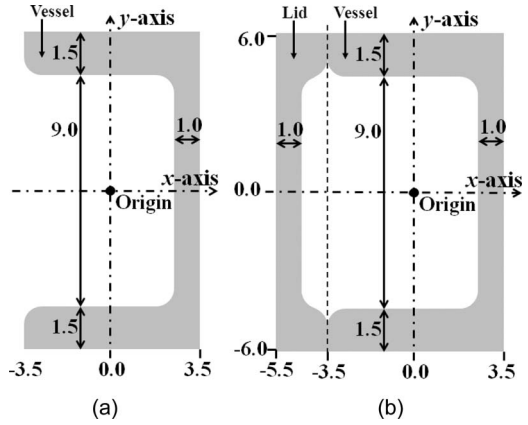

FIG. 2. Vessel model considered. The dimensions on the cross section of $z = 0$ are shown. The numbers given are scaled by d_S. When the large-sphere diameter d_B is set at $5d_S$, for example, the center of the large sphere in contact with the bottom wall is at the origin $(x, y) = (0.0, 0.0)$. (a) No lid is attached to the vessel. (b) A lid is attached to the vessel.

The solute I-solvent ($I = 1, 2, 3$) interaction potential is taken to be

$$\begin{aligned} u_{IS}(h) &= \infty \quad \text{for} \quad h < d_S/2, \\ u_{IS}(h) &= -(\xi_I/8)(d_S/h)^3 \exp[-\{h/(10d_S)\}^{10}] \\ &\qquad \text{for} \quad h > d_S/2, \; I = 1, 2, 3, \end{aligned} \tag{2}$$

where h is the distance between the center of a small sphere and the nearest surface of solute I. In potential (2), the repulsive part of the 9-3 type potential is simply replaced by a hard-core interaction and the potential at large separations is truncated.[27,28] The solute-solvent affinity (i.e., solvophobicity or solvophilicity of the surface of solute I) is controlled by the parameter $\xi_I > 0$. The affinity increases as ξ_I becomes larger.

The details of the 3D-IET are described in Sec. II B. Here, we explain how to specify potential (2) for $I = 2$ or 3. In this theory, the numerical values of the potential are calculated on 3D grid points. On a grid point, we determine the distance between the center of the small sphere placed on this grid point and the nearest surface of solute 2 or solute 3. The distance is then substituted into h for calculating potential (2). With this simple treatment, a small sphere feels significant, negative potential only in the close vicinity of the surface not only of the side but also of the base of the vessel (the potential becomes stronger as the small sphere approaches the surface), and the cylindrical dependence of the potential is fully taken into account. This property is all we need for the spatial distribution of the solute 2-solvent or solute 3-solvent potential. Further, we have verified that the results obtained are quite robust against how to truncate the potential at large separations (see Eq. (2)).

Thanks to hydrogen bonds, water exists as a dense liquid despite the exceptionally small molecular size. However, the hydrogen bonds themselves are not crucial in reproducing many of the interesting characteristics of hydration of solutes. For example, as shown in our earlier work,[29] the characteristics of *hydrophobicity* can reasonably be elucidated by modeling water as spherical particles interacting through strongly attractive potential like that expressed by Eq. (1), as long as

TABLE I. Systems considered. The solvent particles, solute, and vessel are all rigid bodies: No attractive interactions are incorporated in the solvent-solvent, solvent-solute, and solvent-vessel potentials. No lid is attached to the vessel. "(5)" and "(3)" in the system names represent that the values of d_B are set at $5d_S$ and $3d_S$, respectively. In the "Figure" column, it is mentioned that the result calculated for system 0(5), for example, is presented in Fig. 3(a).

System	d_B/d_S	$\varepsilon/(k_BT)$	$\xi_1/(k_BT)$	$\xi_2/(k_BT)$	Figure
0(5)	5	0.0	0.0	0.0	3(a)
0(3)	3	0.0	0.0	0.0	3(b)

the particle size and number density are set at the values for water, respectively. Therefore, the effects of hydrophobicity or hydrophilicity of solute I can well be investigated in terms of those of solvophobicity or solvophilicity using our simple model system.

We test a variety of systems for which the parameters, d_B and $\xi_I/(k_BT)$ ($I = 1, 2, 3$; $T = 298$ K), are set as given in Tables I–III, though the results of some of them are not explicitly discussed in the present article. The values of 0.0, 1.5, and 3.0 are adopted for $\xi_I/(k_BT)$ to represent high solvophobicity, weak solvophobicity, and solvophilicity, respectively. For instance, $\xi_1/(k_BT) = 0.0$, 1.5, and 3.0 mimic highly solvophobic, weakly solvophobic, and solvophilic spherical solutes, respectively. The diameter d_B of solute 1 is set at either $5d_S$ or $3d_S$. "(5)" and "(3)" in the system names represent that the values of d_B are set at $5d_S$ and $3d_S$, respectively. More specifically, the solute with $d_B = 5d_S$ and $\xi_1/(k_BT) = 1.5$ is a model of an unfolded protein possessing larger EV and almost random distribution of hydrophobic and hydrophilic residues on its surface. A folded protein, which is characterized by smaller EV and preferential exposure of hydrophilic residues, is modeled as the solute with $d_B = 3d_S$ and $\xi_1/(k_BT) = 3.0$. For GroEL in the open state, $\xi_2/(k_BT)$ is set at 1.5 because its inner surface is weakly hydrophobic. The setting of $\xi_3/(k_BT) = 3.0$ is employed for the GroEL/GroES complex whose inner surface is hydrophilic. It follows that systems 5(5), 10(5), 10(3), and 8(3) are prepared for stages 1, 2, 3, and 4, respectively.

TABLE II. Systems considered. The attractive interactions are incorporated in the solvent-solvent, solvent-solute, solvent-vessel, and solvent-vessel/lid complex potentials ($\varepsilon/(k_BT)$ is fixed at 1.0). "(5)" in the system names represents that d_B set at $5d_S$. $\xi_1/(k_BT) = 0.0$, 1.5, and 3.0 are adopted for highly solvophobic, weakly solvophobic, and solvophilic solutes, respectively. No lid is attached to the vessel except in systems 10(5) and 11(5). The value of $\xi_I/(k_BT)$ ($I = 2$ and 3 for the vessel and vessel/lid complex, respectively) is set at 0.0, 1.5, and 3.0 to mimic highly solvophobic, weakly solvophobic, and solvophilic inner surfaces, respectively. In the "Figure" column, it is mentioned that the result calculated for system 5(5), for example, is presented in Fig. 5. In the "correspondence" column, it is mentioned that systems 5(5) and 10(5) correspond to stages 1 and 2, respectively.

System	$\xi_1/(k_BT)$	$\xi_2/(k_BT)$	$\xi_3/(k_BT)$	Figure	Correspondence
1(5)	0.0	0.0	…	…	…
2(5)	0.0	1.5	…	…	…
3(5)	0.0	3.0	…	…	…
4(5)	1.5	0.0	…	10(a)	…
5(5)	1.5	1.5	…	5	Stage 1
6(5)	1.5	3.0	…	10(b)	…
7(5)	3.0	0.0	…	…	…
8(5)	3.0	1.5	…	9	…
9(5)	3.0	3.0	…	…	…
10(5)	1.5	…	3.0	14	Stage 2
11(5)	1.5	…	0.0	16(a)	…

TABLE III. Systems considered. The attractive interactions are incorporated in the solvent-solvent, solvent-solute, solvent-vessel, and solvent-vessel/lid complex potentials ($\varepsilon/(k_BT)$ is fixed at 1.0). "(3)" in the system names represents that d_B is set at $3d_S$. $\xi_1/(k_BT) = 0.0$, 1.5, and 3.0 are adopted for highly solvophobic, weakly solvophobic, and solvophilic solutes, respectively. No lid is attached to the vessel except in systems 10(3) and 11(3). The value of $\xi_I/(k_BT)$ ($I = 2$ and 3 for the vessel and vessel/lid complex, respectively) is set at 0.0, 1.5, and 3.0 to mimic highly solvophobic, weakly solvophobic, and solvophilic inner surfaces, respectively. In the "Figure" column, it is mentioned that the result calculated for system 7(3), for example, is presented in Figs. 11(a) and 13(a). In the "correspondence" column, it is mentioned that systems 8(3) and 10(3) correspond to stages 4 and 3, respectively.

System	$\xi_1/(k_BT)$	$\xi_2/(k_BT)$	$\xi_3/(k_BT)$	Figure	Correspondence
1(3)	0.0	0.0	…	12(a)	…
2(3)	0.0	1.5	…	7	…
3(3)	0.0	3.0	…	12(b)	…
4(3)	1.5	0.0	…	…	…
5(3)	1.5	1.5	…	6	…
6(3)	1.5	3.0	…	…	…
7(3)	3.0	0.0	…	11(a) and 13(a)	…
8(3)	3.0	1.5	…	8	Stage 4
9(3)	3.0	3.0	…	11(b) and 13(b)	…
10(3)	3.0	…	3.0	15	Stage 3
11(3)	3.0	…	0.0	16(b)	…

The parameter settings for $\varepsilon/(k_BT)$ and $\xi_I/(k_BT)$ ($I = 1$, 2, 3) described above are justified in the Appendix.

B. Three-dimensional integral equation theory

A solute of arbitrary geometry is immersed at infinite dilution in small spheres with diameter d_S forming the solvent. Solute I ($I = 2$, 3) described in Sec. II A is considered in the present study. The Ornstein-Zernike (OZ) equation in the Fourier space is expressed by[16–26]

$$W_{IS}(k_x, k_y, k_z) = \rho_S C_{IS}(k_x, k_y, k_z) H_{SS}(k) \tag{3}$$

and the hypernetted-chain (HNC) closure equation is written as[16–26]

$$c_{IS}(x, y, z) = \exp\{-u_{IS}(x, y, z)/(k_BT)\} \times \exp\{w_{IS}(x, y, z)\} - w_{IS}(x, y, z) - 1. \tag{4}$$

Here, $w = h - c$, c is the direct correlation function, h the total correlation function, and u the potential. The capital letters (C, H, and W) represent the Fourier transforms. $H_{SS}(k)$ ($k^2 = k_x^2 + k_y^2 + k_z^2$), which is calculated using the radial symmetric integral equation theory (RS-IET) for spherical particles with the HNC approximation,[30] is a part of the input data. We emphasize that the OZ equation is *exact*. Only the HNC closure neglecting the bridge function is approximate.

The numerical procedure can be summarized as follows:

(1) $u_{IS}(x, y, z)$ ($I = 2, 3$) is calculated at each 3D grid point.
(2) $w_{IS}(x, y, z)$ is initialized to zero.
(3) $c_{IS}(x, y, z)$ is calculated from Eq. (4), and $c_{IS}(x, y, z)$ is transformed to $C_{IS}(k_x, k_y, k_z)$ using the 3D fast Fourier transform (3D-FFT).
(4) $W_{IS}(k_x, k_y, k_z)$ is calculated from Eq. (3), and $W_{IS}(k_x, k_y, k_z)$ is inverted to $w_{IS}(x, y, z)$ using the 3D-FFT.
(5) Steps (3) and (4) are repeated until the input and output functions for $w_{IS}(x, y, z)$ become identical within convergence tolerance.

On grid points where a solvent particle and the solute overlap, $\exp\{-u_{IS}(x, y, z)/(k_BT)\}$ is zero. On those where a solvent particle is in contact with the solute, it is set at $\exp\{-u_{IS}^*/(k_BT)\}/2$ where u_{IS}^* represents $u_{IS}(h)$ at the limit, $h \to d_S/2 + 0$ (see Eq. (2)). The grid spacing ($\triangle x$, $\triangle y$, and $\triangle z$) is set at $0.1d_S$, and the grid resolution ($N_x \times N_y \times N_z$) is $512 \times 512 \times 512$. It has been verified that the spacing is sufficiently small and the box size ($N_x\triangle x$, $N_y\triangle y$, $N_z\triangle z$) is large enough for the correlation functions at the box surface to be essentially zero.

We consider solutes 1 and I ($I = 2, 3$). Solute 1 is a large sphere with diameter d_B. Solute 2 is a cylindrical vessel without a lid and solute 3 is the vessel/lid complex (see Fig. 2). First, the solute-1-solvent correlation functions (the Fourier transform of the total correlation function is denoted by $H_{1S}(k)$) are calculated using the RS-IET with the HNC approximation.[30] Second, the solute-I-solvent correlation functions ($I = 2, 3$; the Fourier transform of the direct correlation function is denoted by $C_{IS}(k_x, k_y, k_z)$) are calculated by following the procedure described above. The key quantity is the spatial distribution of the potential of mean force between solute I ($I = 2, 3$) and 1, $\Phi_{I1}(x, y, z)$. That is, the solvent-mediated interaction between solute I and 1 is described by the PMF by assuming that the solvent particles are always in equilibrium with each configuration of the two solutes. This assumption is justified because in the real system the hydration structure steadies in picoseconds, while a protein moves much slower.

$\Phi_{I1}(x, y, z)$ is obtained from

$$\Phi_{I1}(x, y, z)/(k_BT) = u_{I1}(x, y, z)/(k_BT) - w_{I1}(x, y, z), \quad I = 2, 3, \tag{5}$$

where $w_{I1}(x, y, z)$ is calculated by inverting $W_{I1}(k_x, k_y, k_z)$ given by

$$W_{I1}(k_x, k_y, k_z) = \rho_S C_{IS}(k_x, k_y, k_z) H_{1S}(k). \tag{6}$$

The physical meaning of $\Phi_{I1}(x, y, z)$ can be understood from

$$\Phi_{I1}(x, y, z) = F_{I1}(x, y, z) - F_{I1}(\infty, \infty, \infty) \tag{7}$$

and

$$g_{I1}(x, y, z) = \exp\{-\phi_{I1}(x, y, z)/(k_BT)\}, \quad g_{I1}(\infty, \infty, \infty) = 1. \tag{8}$$

Here, the origin of the coordinate system is chosen as illustrated in Fig. 2, $F_{I1}(x, y, z)$ is the free energy of the solvent in the case where solute 1 is at the position (x, y, z), and $g_{I1}(x, y, z)$ is the pair distribution function. A great advantage of the 3D integral equation theory is that the values of Φ_{I1} on all the grid points are obtained from only a single calculation, which is in marked contrast with the usual computer simulation.

The distribution of the solvent number density for a given configuration of solutes 1 and 2 or solutes 1 and 3 can be calculated by following steps (1)–(5) where solute I is treated as solutes 1 and 2 or solutes 1 and 3 in that configuration. The distribution is obtained as $g_{IS}(x,y,z) = h_{IS}(x,y,z) + 1$: $g_{IS}(x, y, z)$ has the physical meaning that the number of solvent molecules within the volume element $dxdydz$ is given by $\rho_S g_{IS}(x,y,z)dxdydz$.

Hereafter, the PMF between solutes 1 and 2 or solutes 1 and 3 is simply denoted by Φ. Under the isochoric condition, we decompose the PMF into entropic and energetic components (they are denoted by Φ_S and Φ_E, respectively) which give physical insights into the PMF characteristics. They are obtained from

$$\Phi_S = -(\partial\Phi/\partial T)_V = -\{\Phi(T + \delta T) - \Phi(T + \delta T)\}/(2\delta T), \quad \delta T = 5\,\mathrm{K}, \ \Phi_E = \Phi + T\Phi_S. \tag{9}$$

We are particularly interested in Φ, $-\Phi_S$, and Φ_E within the vessel cavity which are largely influenced by the solvent structure within it. The spatial distributions of the PMF and its entropic and energetic components are discussed using those of $\Phi/(k_BT)$, $-\Phi_S/k_B$, and $\Phi_E/(k_BT)$, respectively, on the cross section of $z = 0$.

The IET with the HNC approximation was shown to give quite satisfactory results, even in a quantitative sense for the following examples: the PMF between large spherical solutes immersed in pure solvent calculated by the RS-IET,[30] the PMF between large spherical and nonspherical solutes immersed in pure solvent calculated by the 3D-IET,[21] and the PMF between large spherical solutes immersed in a multicomponent solvent calculated by the RS-IET.[31] The reliability test was performed by comparing the IET results obtained by the HNC closures with those either from the density functional theory (DFT),[21,31] which was shown to give the results indistinguishable from those by a MD simulation for rigid-body systems, or from the IET with the exact bridge functions.[30]

III. RESULTS AND DISCUSSION

A. Entropic insertion of a sufficiently large solute into vessel cavity

Hereafter, the spherical solute (solute 1) is simply referred to as "solute." We first consider systems 0(5) and 0(3) (see Table I) where the solute, vessel, and solvent particles are all rigid bodies. No attractive interactions are incorporated in the potentials. In these systems, all of the allowed system configurations share the same energy and the system behavior is purely entropic in origin. The PMF possesses no energetic component: $\Phi/(k_BT) = -\Phi_S/k_B$.

The spatial distributions of the PMFs calculated for systems 0(5) and 0(3) are shown in Figs. 3(a) and 3(b), respectively. In Fig. 3(a), a domain within which the solute is highly stabilized appears around the position $(x/d_S, y/d_S) = (-1.0, 0.0)$. The value of the PMF at this position is $-7.8k_BT$. It

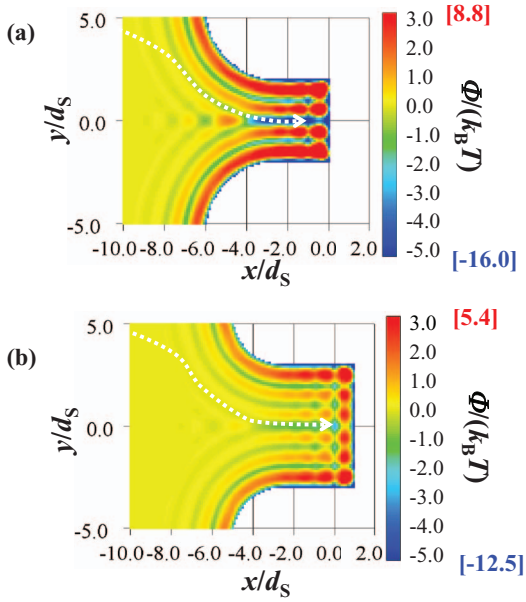

FIG. 3. Potential of mean force $\Phi/(k_BT)$ on the cross section of $z = 0$. As the color approaches thick blue, the value of $\Phi/(k_BT)$ becomes lower, and as the color approaches thick red, it becomes higher. The numbers written in red and in blue represent the maximum and minimum values of $\Phi/(k_BT)$, respectively. The center of the large sphere cannot enter the domain drawn in white. (a) System 0(5). (b) System 0(3). The parameter settings in these systems are given in Table I.

is difficult for the solute to overcome a free-energy barrier well exceeding k_BT. The barrier along the straight route from $(-1.0, 0.0)$ to $(0.0, 0.0)$ is $4.1k_BT$. Therefore, the solute can hardly overcome the barrier to contact the bottom surface of the vessel. (Obviously, the solute cannot contact the side surface, either.) The solute is most likely to be inserted into the vessel cavity through the route indicated by the white dotted arrow (there is no significant barrier) and constrained within the small space around the position $(-1.0, 0.0)$. The route that is symmetrical about the x-axis is also equally probable.

When the solute diameter d_B is made smaller (i.e., it is changed from $5d_S$ to $3d_S$), the amplitudes of the PMF are reduced as observed in Fig. 3(b). The solute tends to be inserted into the vessel cavity through the route indicated by the white dotted arrow and constrained within the small space around the position $(0.0, 0.0)$. The value of the PMF at this position is $-2.8k_BT$ and the barrier along the straight route from $(0.0, 0.0)$ to $(1.0, 0.0)$ is $3.2k_BT$. The powers of entropic insertion and constraint become weaker as the EV generated by the solute decreases.

We then explain how the stripe pattern of the PMF is formed along the y-axis as observed in Fig. 3. When the separation between the nearest solute and vessel surfaces, which is denoted by η here, is not sufficiently close to nd_S ($n = 0, 1, 2, \ldots$), spaces unavailable to the translational displacement of solvent particles appear as indicated in Fig. 4(a). By contrast, in cases of $\eta \sim nd_S$, such unfavorable spaces do not appear and the solvent particles can efficiently be packed within the domain confined between two surfaces as illustrated in Fig. 4(b). The configurations in Figs. 4(a) and 4(b) are entropically unfavorable and favorable, respectively, leading to

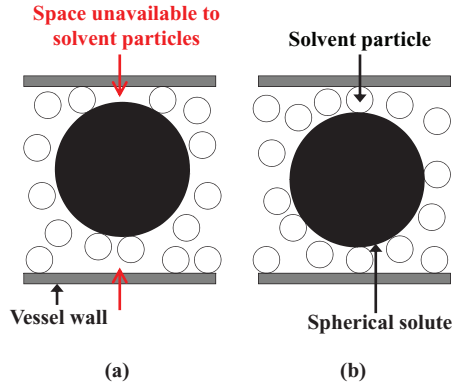

FIG. 4. Cartoons illustrating how the stripe pattern of the potential of mean force in Fig. 3 is formed along the y-axis. (a) Separation between the nearest solute and vessel surfaces, which is denoted by η, is not sufficiently close to nd_S ($n = 0, 1, 2, \ldots$). (b) In the case of $\eta \sim nd_S$.

the stripe pattern formed along the y-axis. A similar pattern is formed along the x-axis near the bottom surface of the vessel.

B. Insertion of a sufficiently large, weakly solvophobic solute into vessel cavity: Model of stage 1

Incorporating attractive interactions in the solvent-solvent, solvent-solute, and solvent-vessel potentials, we calculate the PMF and its energetic and entropic components. In systems 5(5) and 5(3) (see Tables II and III) considered in this section, the parameter setting features a solute and inner surface of the vessel with weak solvophobicity. The spatial distributions of $\Phi/(k_BT)$, $\Phi_E/(k_BT)$, and $-\Phi_S/k_B$ for the two systems are shown in Figs. 5 and 6, respectively.

The pattern of $\Phi/(k_BT)$ in Fig. 3(a) and that of $-\Phi_S/k_B$ in Fig. 5(c) share qualitatively the same characteristics. Also, the pattern of $\Phi/(k_BT)$ in Fig. 3(b) is qualitatively the same as that of $-\Phi_S/k_B$ in Fig. 6(c). For a given value of d_B, the behavior of $-\Phi_S/k_B$ is similar to that of $\Phi/(k_BT)$ for the rigid-body model considered in Sec. III A. By comparing Fig. 5(a) with Fig. 5(c), and Fig. 6(a) with Fig. 6(c), we notice that the basic pattern of the PMF is determined primarily by its entropic component. The energetic components shown in Figs. 5(b) and 6(b) exhibit qualitatively the same characteristics.

As observed in Fig. 5(a), the solute is most likely to be inserted into the vessel cavity through the route indicated by the white dotted arrow (there is no significant barrier) and constrained within the small space around the position $(-1.0, 0.0)$. The value of the PMF at this position is $-5.9k_BT$. The barrier along the straight route from $(-1.0, 0.0)$ to $(0.0, 0.0)$ is $4.1k_BT$. This result is similar to that obtained in system 0(5). In Fig. 6(a), the solute tends to be inserted into the vessel cavity through the route indicated by the white dotted arrow and constrained within the small space around the position $(0.0, 0.0)$. However, the PMF-value at $(0.0, 0.0)$ is only $-1.1k_BT$ that is not sufficiently deep unlike in Fig. 5(a). This difference is ascribed to the smaller EV generated by the solute.

As explained in Sec. II A and proved in this section, the solute insertion in system 5(5) illustrated in Fig. 5 is a model of stage 1 (see Fig. 1). According to the result from this model, an unfolded protein is entropically inserted into the

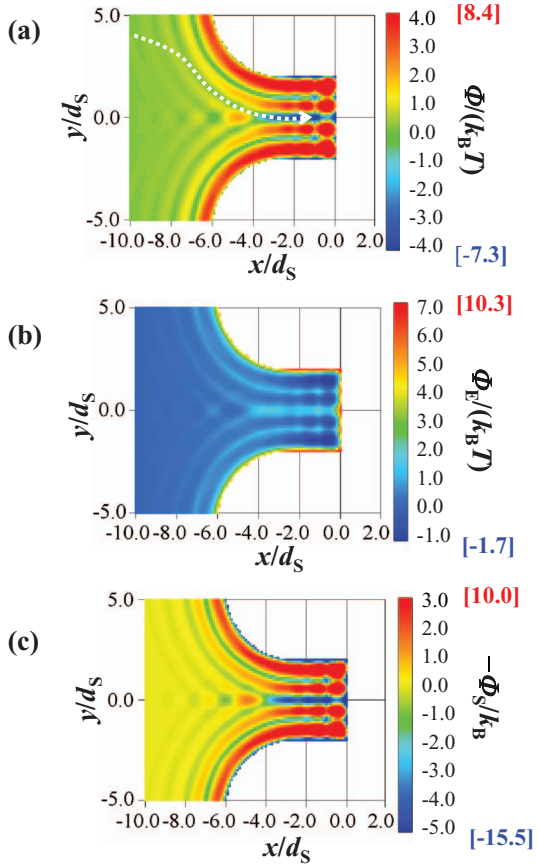

FIG. 5. Potential of mean force $\Phi/(k_BT)$ (a), energetic component $\Phi_E/(k_BT)$ (b), and entropic component $-\Phi_S/k_B$ (c) on the cross section of $z = 0$ calculated for system 5(5). $\Phi = \Phi_E - T\Phi_S$: $\Phi/(k_BT) = \Phi_E/(k_BT) - \Phi_S/k_B$. As the color approaches thick blue, the value of $\Phi/(k_BT)$, $\Phi_E/(k_BT)$, or $-\Phi_S/k_B$ becomes lower, and as the color approaches thick red, it becomes higher. The numbers written in red and in blue represent the maximum and minimum values of $\Phi/(k_BT)$, $\Phi_E/(k_BT)$, or $-\Phi_S/k_B$, respectively. The positions where $\Phi_E/(k_BT)$ takes the maximum and minimum values are different from those where $-\Phi_S/k_B$ takes the maximum and minimum values, respectively. Hence, the maximum value of $\Phi/(k_BT)$, for instance, is not equal to that of $\Phi_E/(k_BT)$ plus that of $-\Phi_S/k_B$. The parameter setting in this system is given in Table II.

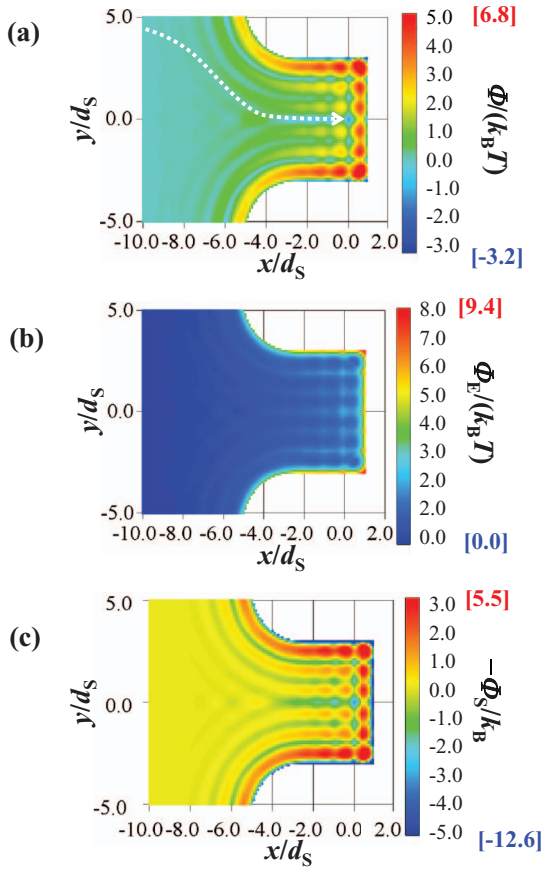

FIG. 6. Potential of mean force $\Phi/(k_BT)$ (a), energetic component $\Phi_E/(k_BT)$ (b), and entropic component $-\Phi_S/k_B$ (c) on the cross section of $z = 0$ calculated for system 5(3). The parameter setting in this system is given in Table III.

GroEL cavity. After the insertion, the protein is constrained almost in the center of the cavity and ready to fold, which is consistent with the experimental observation.[10,11] We note that the constraint as well as the insertion is achieved by the entropic component.

C. Energetic release of a sufficiently small, solvophilic solute to bulk solvent: Model of stage 4

The results presented in Figs. 3, 5, and 6 indicate that the entropic component, which always drives the insertion of the solute, predominates over the energetic component. The powers of entropic insertion and constraint are weaker in Fig. 6 than in Fig. 5, but they are still dominant. The release of the solute may be realized, if the power is sufficiently reduced and the energetic component effectively acts for the release. In this section, we investigate the effects of the solute solvophobicity or solvophilicity represented by $\xi_1/(k_BT)$ on the energetic component and show the following: When the solute turns solvophilic, the energetic component strongly promotes the release; it can surpass the powers of entropic insertion and constraint when the solute becomes more compact with smaller EV; and the switch from insertion to release is realized.

We consider systems 2(3) and 8(3) given in Table III. The inner surface of the vessel remains weakly solvophobic. The values of $\xi_1/(k_BT)$ are set at 0.0 and 3.0 to mimic highly solvophobic and solvophilic solutes in systems 2(3) and 8(3), respectively. $\Phi/(k_BT)$, $\Phi_E/(k_BT)$, and $-\Phi_S/k_B$ for the two systems are shown in Figs. 7 and 8, respectively. It is observed in Figs. 7(c) and 8(c) that the entropic component is rather insensitive to $\xi_1/(k_BT)$. Our major concern is the effects of $\xi_1/(k_BT)$ on the energetic component: Figs. 7(b) and 8(b) should also be compared with Fig. 6(b) for which $\xi_1/(k_BT)$ is set at 1.5. When the solute is highly solvophobic, as shown in Fig. 7(b), the solute is energetically inserted into the vessel cavity and driven to contact the inner surface of the vessel. The insertion originates from the feature that a highly solvophobic solute is strongly excluded from the bulk solvent, and the contact is attributed to the solvophobic interaction. As a consequence, the PMF acts for nothing but strong insertion as observed in Fig. 7(a). By contrast, a solvophilic solute, which is preferentially solvated in the bulk solvent, is strongly released from the vessel cavity by the energetic component as shown in Fig. 8(b).

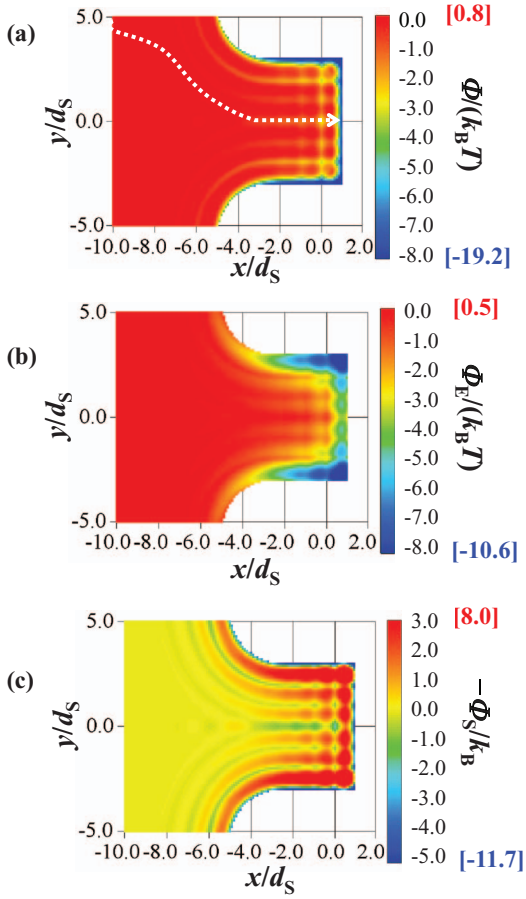

FIG. 7. Potential of mean force $\Phi/(k_BT)$ (a), energetic component $\Phi_E/(k_BT)$ (b), and entropic component $-\Phi_S/k_B$ (c) on the cross section of $z = 0$ calculated for system 2(3). The parameter setting in this system is given in Table III.

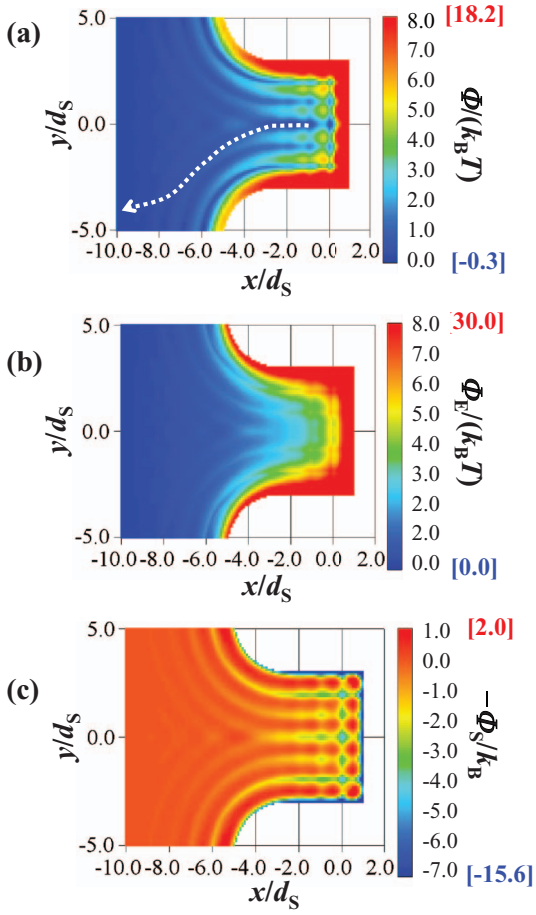

FIG. 8. Potential of mean force $\Phi/(k_BT)$ (a), energetic component $\Phi_E/(k_BT)$ (b), and entropic component $-\Phi_S/k_B$ (c) on the cross section of $z = 0$ calculated for system 8(3). The parameter setting in this system is given in Table III.

When the solute turns sufficiently solvophilic, the energetic component acts for the release. At the same time, the solute needs to become sufficiently more compact. The EV generated by the solute then decreases, leading to reduced powers of entropic insertion and constraint. It is clear from the PMF for system 8(5) (the parameter setting is the same as that in system 8(3) except that $d_B = 5d_S$) shown in Fig. 9 that the release is not achieved unless the solute becomes sufficiently more compact. As explained in Sec. II A and proved in this section, the solute release in system 8(3) illustrated in Fig. 8 is a model of stage 4 (see Fig. 1).

D. Significance of weak solvophobicity of vessel inner surface in stage 1

The solute insertion in system 5(5) illustrated in Fig. 5 is a model of stage 1. In stage 1, the inner surface of the vessel is weakly solvophobic: $\xi_2/(k_BT) = 1.5$. A question then arises: What will happen if the inner surface is highly solvophobic or solvophilic? To answer this question, we consider systems 4(5) and 6(5) with $\xi_2/(k_BT) = 0.0$ and 3.0, respectively (see Table II). The inner surfaces in systems 4(5), 5(5), and 6(5) are highly solvophobic, weakly solvophobic, and solvophilic, respectively.

Figure 10(a) shows the PMF calculated for system 4(5). This figure should be compared with Fig. 5(a) drawn for system 5(5). In Fig. 10(a), the solute tends to be inserted into the vessel cavity, for example, through the route indicated by the white dotted arrow. The PMF-value at $(-1.0, 0.0)$ is $-4.5k_BT$. The barrier along the straight route from $(-1.0, 0.0)$ to $(0.0, 0.0)$ is $1.9k_BT$ that is much lower than $4.1k_BT$ in the case of system 5(5). When the inner surface of the vessel is highly solvophobic, the possibility that the solute comes in contact with the bottom surface becomes much higher.

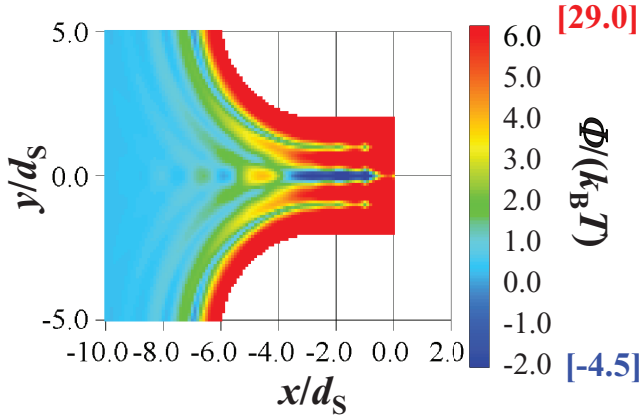

FIG. 9. Potential of mean force $\Phi/(k_BT)$ on the cross section of $z = 0$ calculated for system 8(5). The parameter setting in this system is given in Table II.

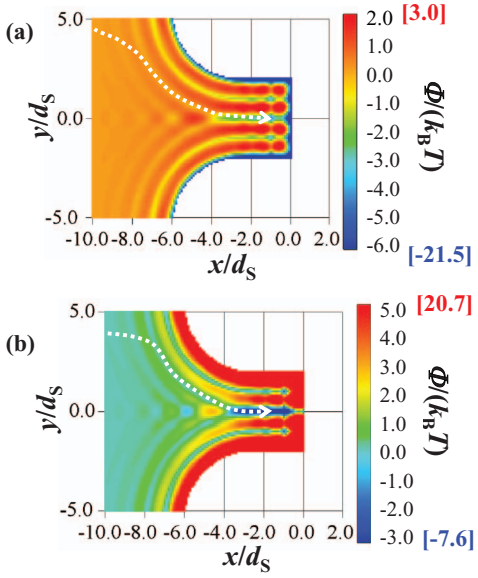

FIG. 10. Potential of mean force $\Phi/(k_BT)$ on the cross section of $z = 0$. (a) System 4(5). (b) System 6(5). The parameter settings in these systems are given in Table II.

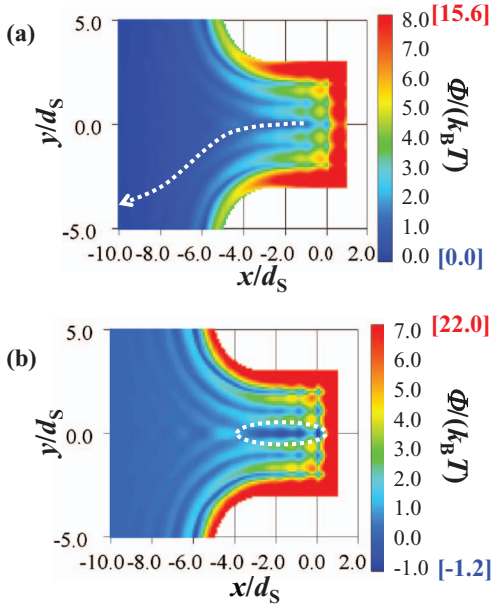

FIG. 11. Potential of mean force $\Phi/(k_BT)$ on the cross section of $z = 0$. (a) System 7(3). (b) System 9(3). The parameter settings in these systems are given in Table III.

Figure 10(b) shows the PMF calculated for system 6(5). The solute is most likely to be inserted into the vessel cavity through the route indicated by the white dotted arrow. There are two positions at which the PMF takes a large, negative value along the x-axis: $(-1.8, 0.0)$ with $\Phi = -6.9k_BT$ and $(-1.0, 0.0)$ with $\Phi = -7.6k_BT$. The barrier along the straight route from $(-1.8, 0.0)$ to $(-1.0, 0.0)$ is only $1.4k_BT$, whereas that from $(-1.0, 0.0)$ to $(0.0, 0.0)$ reaches $8.5k_BT$. The solute-bottom surface contact is very difficult.

In stage 1, $\xi_2/(k_BT)$ has significant effects primarily on the barrier for the solute to overcome for contacting the bottom surface of the vessel. Since the contact is to be avoided, a highly solvophobic inner surface of the vessel is somewhat problematic. A solvophilic inner surface is the most favorable, but a weakly solvophobic one is quite acceptable.

E. Significance of weak solvophobicity of vessel inner surface in stage 4

The solute release in system 8(3) illustrated in Fig. 8 is a model of stage 4. In stage 4, the inner surface of the vessel is weakly solvophobic as in stage 1: $\xi_2/(k_BT) = 1.5$. To explore the significance of this weak solvophobicity, we consider systems 7(3) and 9(3) with $\xi_2/(k_BT) = 0.0$ and 3.0, respectively (see Table III). The inner surfaces in systems 7(3), 8(3), and 9(3) are highly solvophobic, weakly solvophobic, and solvophilic, respectively.

Figure 11(a) shows the PMF calculated for system 7(3). This figure should be compared with Fig. 8(a) drawn for system 8(3). The PMFs in the two figures share qualitatively the same characteristics. If the solute is located at the position $(-1.0, 0.0)$ in Fig. 11(a), the solute is expected to be released to the outside, for example, through the route indicated by the white dotted arrow.

Figure 11(b) shows the PMF calculated for system 9(3). It is found that the solute is weakly constrained within the domain indicated by the white dotted ellipse (the PMF-value at $(-1.0, 0.0)$ is $-1.0k_BT$). If the inner surface of the vessel is solvophilic, there are more solvent particles within the vessel cavity than in the case where the inner surface is solvophobic. As a consequence, the solvent environment for a solvophilic solute in the center of the vessel cavity is energetically as good as that in the bulk, and the trend that the solute is preferentially solved in the bulk is significantly reduced. (More details are discussed below using Fig. 13(b).) This reduction leads to the slight dominance of the entropic constraint.

In stage 4, $\xi_2/(k_BT)$ has significant effects primarily on the power of the solute release. The inner surface of the vessel should not be solvophilic: High or weak solvophobicity is required for the inner surface. In stage 1, as described in Sec. III D, a highly solvophobic inner surface needs to be avoided, a solvophilic one is the most favorable, and a weakly solvophobic one is quite acceptable. The best compromise meeting both of the requirements in stages 1 and 4 is the weak solvophobicity of the inner surface.

F. Effects of affinities of solute and vessel inner surface with solvent on energetic component of potential of mean force

The energetic component $\Phi_E/(k_BT)$ of the PMF is much more sensitive to the parameters, $\xi_1/(k_BT)$ and $\xi_2/(k_BT)$, than the entropic component. Here, it is worthwhile to investigate $\Phi_E/(k_BT)$ in more detail. The energetic components calculated for systems 1(3) and 3(3) are shown in Figs. 12(a) and 12(b), respectively, and those for systems 7(3) and 9(3) are shown in Figs. 13(a) and 13(b), respectively. There are four combinations: (a) both of the solute and the inner surface are highly solvophobic, system 1(3) (Fig. 12(a)); (b) the solute is highly solvophobic but the inner surface is solvophilic, system 3(3) (Fig. 12(b)); (c) the solute is solvophilic but the inner surface

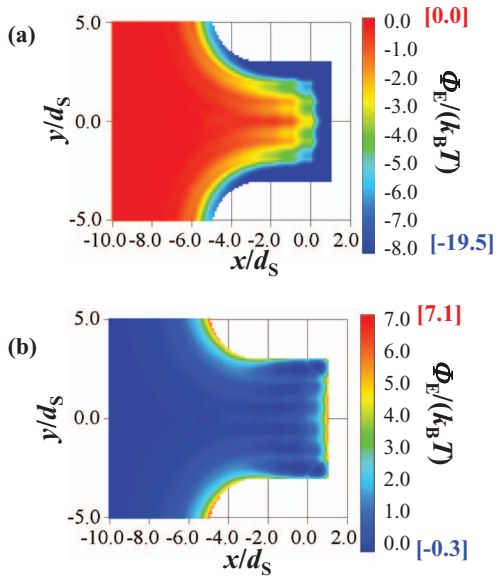

FIG. 12. Energetic component of the potential of mean force $\Phi_E/(k_BT)$ on the cross section of $z = 0$. (a) System 1(3). (b) System 3(3). The parameter settings in these systems are given in Table III.

is highly solvophobic, system 7(3) (Fig. 13(a)); and (d) both of them are solvophilic, system 9(3) (Fig. 13(b)).

When the inner surface is solvophilic, there are more solvent particles within the vessel cavity than in the case where the inner surface is solvophobic. Especially, the number density of solvent particles in the vicinity of the inner surface becomes higher than in the bulk solvent. However, the opposite is true in the case where the inner surface is solvophobic. $\Phi_E/(k_BT)$ behaves differently for the four combinations:

(a) $\Phi_E/(k_BT)$ is negative inside the cavity. The solute is excluded from the bulk solvent and strongly inserted into the cavity (the solvent environment within the cavity is more favorable for the solute). $\Phi_E/(k_BT)$ takes large, negative values near the inner surface, and the solute-inner surface contact is strongly promoted.

(b) $\Phi_E/(k_BT)$ is essentially zero except in the immediate vicinity of the inner surface where it is positive. The solute is almost equally stabilized inside the cavity and in the bulk solvent but the solute-inner surface contact is destabilized.

(c) $\Phi_E/(k_BT)$ is positive inside the cavity. The solute is preferentially solvated in the bulk solvent. $\Phi_E/(k_BT)$ takes large, positive values near the inner surface, and the solute cannot come in contact with the inner surface.

(d) The solute is roughly equally stabilized inside the cavity and in the bulk solvent except near the inner surface where $\Phi_E/(k_BT)$ is largely positive. The solute-inner surface contact is highly destabilized.

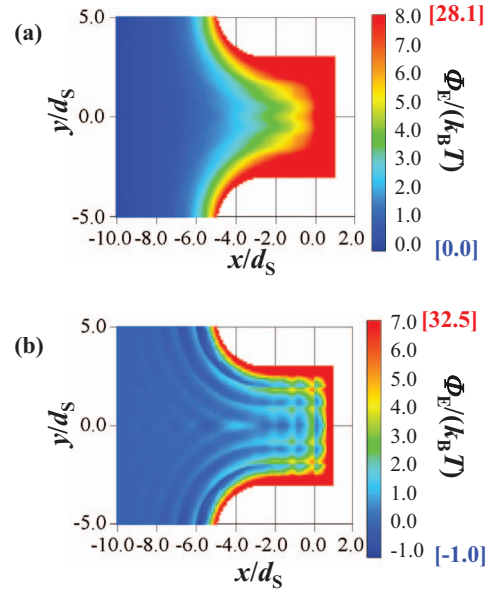

FIG. 13. Energetic component of the potential of mean force $\Phi_E/(k_BT)$ on the cross section of $z = 0$. (a) System 7(3). (b) System 9(3). The parameter settings in these systems are given in Table III.

The values of $\Phi_E/(k_BT)$ at (0.0, 0.0) in cases (a), (b), (c), and (d) are −2.6, 1.4, 8.3, and 1.0, respectively. As important information, when at least one of the solute and inner surface is solvophilic, the solute-inner surface contact is destabilized due to the difficulty in desolvating the solvophilic solute or inner surface.

According to a prevailing view, an unfolded protein first binds to a hydrophobic patch on the interior rim of GroEL and it is ejected into the cavity after the patch is removed by the GroEL structural change,[1,3–7,11,13] which is not quite coincident with stage 1. The binding can be understood in terms of the result from case (a). The ejection process is not straightforwardly treated in the present study, but it is definite that the solvent plays crucially important roles. For instance, when the GroEL inner surface in contact with the unfolded protein turns hydrophilic, the contact becomes destabilized as suggested by the result from case (b). Further, the insertion mechanism argued for systems 5(5) and 6(5) should be utilized before the protein starts to fold in the closed cavity of the GroEL/GroES complex. We also believe that it is relevant to the insertion of an antibiotic molecule into a cell-membrane protein such as ATP-binding cassette (ABC) transporter[32,33] and TolC.[34,35]

G. Roles of a lid: Models of stages 2 and 3

We consider system 10(5) to simulate stage 2 (see Table II). What we calculate is the PMF between the solute and the vessel/lid complex. The inner surface of the complex is solvophilic, while the solute is weakly solvophobic: $\xi_3/(k_BT)$ and $\xi_1/(k_BT)$ are set at 3.0 and 1.5, respectively. Figure 14 shows the PMF and its energetic and entropic components for system 10(5). As observed in Fig. 14(a), the solute remains constrained within the small space around the position (−1.0, 0.0). The solute is energetically repelled from the inner surface and entropically stabilized at (−1.0, 0.0) (see Figs. 14(b) and 14(c)). The barrier along the straight route from (−1.0, 0.0) to (0.0, 0.0) in Fig. 14(a) reaches $9.4k_BT$. The result, which suggests that the protein is constrained almost in the center of the complex cavity, is consistent with the experimental observation.[11,12]

We then consider the system 10(3) to simulate stage 3 (see Table III). The inner surfaces of the complex and the

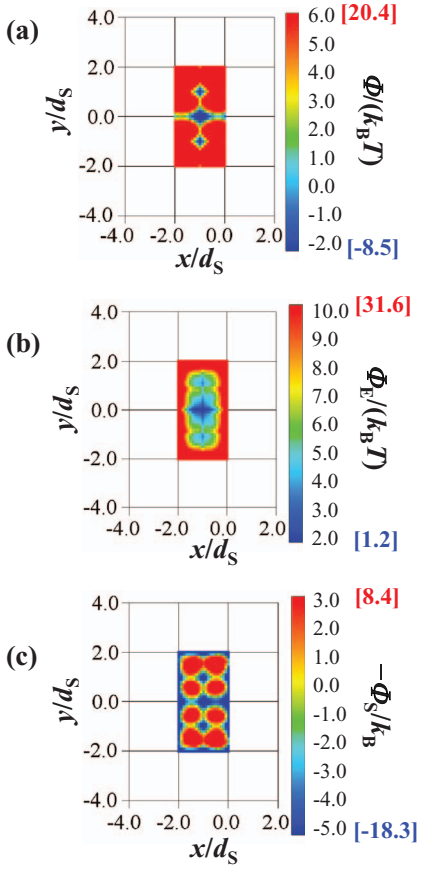

FIG. 14. Potential of mean force $\Phi/(k_BT)$ (a), energetic component $\Phi_E/(k_BT)$ (b), and entropic component $-\Phi_S/k_B$ (c) on the cross section of $z = 0$ calculated for system 10(5). The parameter setting in this system is given in Table II.

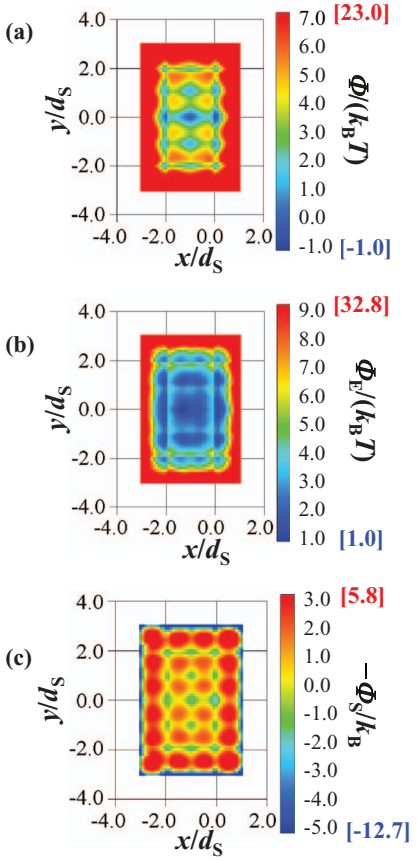

FIG. 15. Potential of mean force $\Phi/(k_BT)$ (a), energetic component $\Phi_E/(k_BT)$ (b), and entropic component $-\Phi_S/k_B$ (c) on the cross section of $z = 0$ calculated for system 10(3). The parameter setting in this system is given in Table III.

solute are solvophilic: Both $\xi_3/(k_BT)$ and $\xi_1/(k_BT)$ are set at 3.0. Figure 15 shows the PMF and its energetic and entropic components calculated for system 10(3). The barrier along the straight route from (−1.0, 0.0) to (1.0, 0.0) in Fig. 15(a) reaches $10.5k_BT$. The qualitative aspects of the PMF and its components are not significantly different from those in stage 2. The folded protein is also stabilized almost in the center of the complex cavity and remains well separated from the complex inner surface.

H. Significance of solvophilicity of complex inner surface in stages 2 and 3

To examine the significance of the solvophilicity of complex inner surface in stages 2 and 3, we consider systems 11(5) and 11(3) for which $\xi_3/(k_BT)$ is set at 0.0 to mimic a highly solvophobic inner surface (see Tables II and III). The PMFs calculated for the two systems are shown in Fig. 16. In Fig. 16(a) where the solute is weakly solvophobic and the inner surface is highly solvophobic, the contact of the protein with the inner surface is highly stabilized due to the energetic component as well as the entropic component of the PMF. This is readily understood from the result in Sec. III F. When the solute is located at (−1.0, 0.0), it has to overcome the barrier of $3.5k_BT$ along the straight route from (−1.0, 0.0) to (0.0, 0.0): The solute can still be constrained within a small space around (−1.0, 0.0), but the barrier is only about one-third of that in Fig. 14(a). The constraint is still retained in stage 3 as well (see Fig. 16(b)), but the barrier along the straight route from (−1.0, 0.0) to (1.0, 0.0) is $5.5k_BT$ that is about half of that in Fig. 15(a). Considering that the solute is much more strongly repelled from the inner surface due to the energetic component and constrained within the small space in systems 10(5) and 10(3), we conclude that the solvophilicity of the inner surface is definitely the most favorable in stages 2 and 3.

The solvophilicity of complex inner surface plays essential roles in another way. England *et al.*[36] have found the following: The hydrophilic inner surface imports more water molecules into the cavity of the GroEL/GroES complex than the hydrophobic one, which is expected to play essential roles in facilitation of protein folding. In summary, the affinity of the GroEL inner surface with water is judiciously adjusted (i.e., the inner surface is weakly hydrophobic in stages 1 and 4, while it is made hydrophilic in stages 2 and 3) to assure the optimum environment for the insertion of an unfolded protein into the cavity, protein folding, and release of the folded one back to the bulk aqueous solution.

IV. MECHANISM OF PROTEIN FLUX THROUGH CHAPERONIN SYSTEM

On the basis of our theoretical results and experimentally available information, we can uncover the mechanism of the

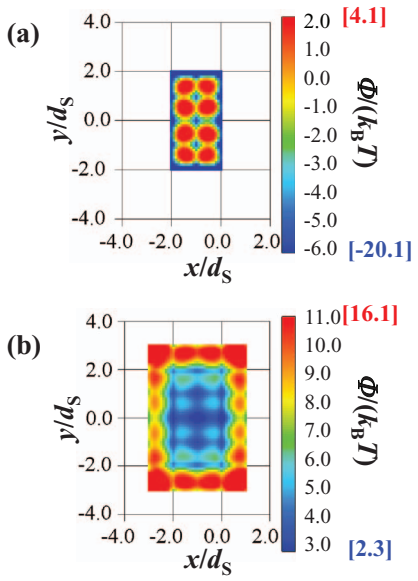

FIG. 16. Potential of mean force $\Phi/(k_BT)$ on the cross section of $z = 0$. (a) System 11(5). (b) System 11(3). The parameter settings in these systems are given in Tables II and III, respectively.

whole cycle comprising stages 1 through 4. First, some of the important results concerning the entropic and energetic components of the PMF are recapitulated here.

Entropic component of the PMF:

(S-1) $-\Phi_S/k_B$ always drives a solute to be inserted into the vessel cavity and confined within a small space almost in its center. The insertion and constraint become more powerful as the EV generated by the solute increases.

(S-2) $-\Phi_S/k_B$ always stabilizes the solute-inner surface contact, but there is a barrier for the solute to overcome for coming in contact with the inner surface.

(S-3) The basic pattern of $-\Phi_S/k_B$ is much less sensitive to the solvophobicity or solvophilicity of the solute and vessel inner surface than that of the energetic component.

(S-4) As the EV generated by the solute decreases, the amplitudes of $-\Phi_S/k_B$ reduce.

Energetic component of the PMF:

(E-1) $\Phi_E/(k_BT)$ is considerably dependent on the solvophobicity or solvophilicity of the solute and vessel inner surface, exhibiting large variation.

(E-2) When both of the solute and the vessel inner surface are solvophobic, $\Phi_E/(k_BT)$ excludes the solute from the bulk solvent and inserts it into the vessel cavity. The solute-inner surface contact is readily reached and stabilized.

(E-3) When the solute is solvophilic but the inner surface is solvophobic, $\Phi_E/(k_BT)$ makes the solute be preferentially solvated in the bulk solvent. The solute cannot come in contact with the inner surface.

(E-4) In cases where either the solute or the inner surface is solvophilic, the solute-inner surface contact is destabilized. In particular, both of the solute and the inner surface possess solvophilicity, the solute is strongly repelled from the inner surface.

The mechanism of the protein flux through the chaperonin system can be described as follows (GroEL/GroES is considered but the same explanation is valid for Hsp60/Hsp10 as well):

Stage 1. For an unfolded protein, the EV generated is significantly large and the surface is weakly hydrophobic. In this stage, the GroEL inner surface is also weakly hydrophobic.[5] As a consequence, the PMF is governed by its entropic component. The protein is inserted into the GroEL cavity and constrained within a small space almost in the center of the cavity. The insertion and constraint are primarily entropic in origin.

Stage 2. The ATP binding followed by attachment of GroES to the GroEL rim as a lid leads to the change that the GroEL/GroES inner surface turns hydrophilic. The protein is energetically repelled from the inner surface and remains constrained within the small space mentioned in stage 1.

Stage 3. The protein securely folds into its native structure without contacting the inner surface, during which ATP is hydrolyzed into ADP and Pi. For the folded protein, the EV generated is much smaller and the surface is now hydrophilic.

Stage 4. The dissociation of ADP and Pi occurs with the result of the detachment of GroES. The structural and inner-surface properties of GroEL return to those in stage 1. Due to the much smaller EV of the folded protein, the powers of insertion and constraint by the entropic component of the PMF are suppressed. The protein is then energetically released back to the bulk solvent due to the hydrophilicity of the folded protein. The release arises from factor (E-3). It is important to note that if the GroEL inner surface remained hydrophilic, it would become rather difficult even for the folded protein to be released (see Fig. 11(b)). It is apparent that GroES acting as a lid is useful for safely preventing the energetic release before the folding is finished.

An unfolded protein with larger EV and weak hydrophobicity is entropically inserted into the chaperonin cavity. The switch from insertion to release is achieved by decreasing the EV and turning the protein surface hydrophilic in the folding process. For this release, in which the energetic component is a requisite, the feature that the vessel inner surface is not hydrophilic plays an essential role. On the other hand, the inner surface of the GroEL/GroES complex is hydrophilic,[5] and the protein is energetically repelled from it: the protein remains constrained within a small space well separated from them, which is relevant to correct folding.

For a slow folder such as mitochondrial rhodanese, the folding is not finished in one cycle of the ATP binding, hydrolysis, and dissociation of ADP and Pi.[37] That is, it may not be released from chaperonin even when it opens by the detachment of cochaperonin. It has been observed for rhodanese *in vitro* that ~25% of non-native proteins are released and ~75% of them remain within chaperonin (in the absence

of crowding agents as assumed in the present study).[38,39] This could be interpreted as follows. A protein whose structure differs significantly from the native structure is not released even when chaperonin enters the open state: The protein remains within chaperonin until its structure becomes sufficiently close to the native structure. This interpretation does not contradict the results from the present study. On the other hand, it has been suggested that the protein is inserted and released in non-native structures multiple times.[37,40] If this is true, the mechanism proposed in the present study is not always applicable to all cases, or some other factors which we do not account for also come into play.

V. CONCLUDING REMARKS

We have investigated the behavior of a large solute which is inserted into or released from an even larger cylindrical vessel by analyzing the solvent-mediated, solute-vessel potential of mean force. A lid is either detached from or attached to the vessel. It provides a model of the protein flux through the chaperonin system which inserts an unfolded protein from bulk aqueous solution, facilitates protein folding, and releases the folded protein back to the bulk solution. The three-dimensional integral equation theory is employed in the analyses. The theory is characterized by the following advantages: The spatial distribution of the PMF is obtained only in one calculation; and the PMF can readily be decomposed into entropic and energetic components which give physical insights into the PMF characteristics.

On the basis of the experimental observations manifesting that the basic aspects of the protein flux is independent of the chaperonin, cochaperonin, and protein species,[2,5,9] we adopt a simple model system which is free from specific chemical and structural characteristics of a protein as well as the chaperonin/cochaperonin complex. We have succeeded in consistently elucidating the underlying mechanism of the whole cycle comprising stages 1–4 mentioned in the "Introduction" and illustrated in Fig. 1. The mechanism uncovered as well as the important results concerning the entropic and energetic components of the PMF is summarized in Sec. IV. The function of chaperonin/cochaperonin, assistance of protein folding, is exhibited by the three key factors: the hydration properties of a protein in the water confined on the scale of a nanometer which are substantially different from those in the bulk water; changes of the conformation and surface properties of the protein upon folding; and adjustment of the structure and properties of the chaperonin system by the chemical cycle of ATP hydrolysis (i.e., the ATP binding, hydrolysis of ATP into ADP and Pi, and dissociation of ADP and Pi). Some of the results obtained should be applicable to the transportation of an antibiotic molecule across a cell-membrane protein such as ABC transporter[32,33] and TolC.[34,35]

A variety of self-assembling and ordering processes in biological systems, which occur at molecular levels, are sustaining life. Examples of such processes are protein folding leading to a unique native structure, molecular recognition, association of protein molecules forming an ordered and often symmetrical complex, and lipid-membrane formation. Water plays imperative roles in these processes.[41,42] As another aspect, there are proteins or protein complexes called "molecular machines" which exhibit a variety of high functions.[1–13,32,33,43–49] These molecular machines share the feature that they function by utilizing the chemical cycle of ATP hydrolysis. There is a general trend that only the roles of ATP are emphasized. We remark, however, that the exhibition of their functions of all those molecular machines is realized by the cooperation of water and ATP. From this viewpoint, we have recently elucidated the mechanisms of the unidirectional movement of a linear-motor protein along a filament[26] and the directed rotation of a subunit within a rotatory-motor protein complex.[50] In the present study, the mechanism of the protein flux through the chaperonin system is successfully elucidated. The water-mediated potential field felt by a protein is strongly dependent on the structure and properties of the chaperonin system as well as those of the protein; the chemical cycle of ATP hydrolysis adjusts the former, thus controlling the potential field.

ACKNOWLEDGEMENTS

This work was supported by the Grant-in-Aid for Scientific Research (B) (No. 22300100) and that on Innovative Areas (No. 20118004) from the Ministry of Education, Culture, Sports, Science and Technology of Japan, by the Grand Challenges in Next-Generation Integrated Nanoscience, MEXT, Japan, by the Grant-in-Aid for JSPS (Japan Society for the Promotion of Science) fellows, and by the Kyoto University Global Center of Excellence of Energy Science.

APPENDIX: JUSTIFICATION OF THE PARAMETER SETTINGS FOR $\varepsilon/(k_B T)$ and $\xi_l/(k_B T)$ ($l = 1, 2, 3$)

In potential (1), ε is the lowest value of potential energy between two solvent molecules. On the other hand, the strength of a hydrogen bond between two water molecules is larger than $5k_BT$ at $T = 298$ K. One might think that the setting of $\varepsilon/(k_BT) > 5$ should be employed. However, this setting is pathological for the following reason. In the simple-fluid model, increasing $\varepsilon/(k_BT)$ eventually leads to the divergence of the correlation length (and that of the isothermal compressibility); It occurs at $\varepsilon/(k_BT) \sim 2$ beyond which a single phase cannot exist even as a metastable state. $\varepsilon/(k_BT) \sim 2$ is the spinodal point for the liquid-gas phase transition. For a dipolar fluid, when the dipole moment is progressively increased, the spinodal point for the liquid-nematic phase transition is eventually encountered; the divergence of the orientational correlation length occurs at the value that is much smaller than the dipole moment of water.[51] Thus, the equalization of the lowest value of potential energy between two solvent molecules for different models is physically meaningless. We have already shown that most of the characteristics of *hydrophobicity* can well be reproduced by the simple-fluid model with $\varepsilon/(k_BT) = 1.5$.[29,52] In the present study, $\varepsilon/(k_BT)$ is set at 1.0, but this is still large enough to represent solvophobicity or solvophilicity of a solute as described below.

One of the good criteria which allows us to determine if a solute is solvophobic or solvophilic (and its degree) is the reduced number-density profile of solvent near the solute. In

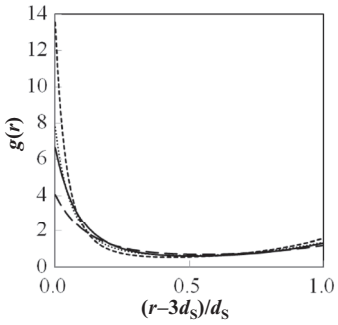

FIG. 17. Reduced number-density profiles of solvent molecules near a solute with diameter $d_B = 5d_S$ (d_S is the diameter of solvent molecules and $3d_S = (d_B + d_S)/2$). Solid curve: system A($\varepsilon/(k_BT) = 0.0$ and $\xi_1/(k_BT) = 0.0$); long-dashed curve: system B ($\varepsilon/(k_BT) = 1.0$ and $\xi_1/(k_BT) = 0.0$); dotted curve: system C ($\varepsilon/(k_BT) = 1.0$ and $\xi_1/(k_BT) = 1.5$); short-dashed curve: system D ($\varepsilon/(k_BT) = 1.0$ and $\xi_1/(k_BT) = 3.0$).

Fig. 17, the profiles are shown for the four systems: $\varepsilon/(k_BT) = 0.0$ and $\xi_1/(k_BT) = 0.0$ (system A); $\varepsilon/(k_BT) = 1.0$ and $\xi_1/(k_BT) = 0.0$ (system B); $\varepsilon/(k_BT) = 1.0$ and $\xi_1/(k_BT) = 1.5$ (system C); and $\varepsilon/(k_BT) = 1.0$ and $\xi_1/(k_BT) = 3.0$ (system D). The solute is spherical and its diameter d_B is set at $5d_S$ in all the systems. Even in system A, a layer within which the solvent enrichment occurs is formed near the solute due to the purely entropic effect.[41] The enrichment is significantly reduced in system B because the solute is unfavorable for the solvent, almost the same in system C, and significantly enhanced in system D because the solute is favorable for the solvent. Therefore, the solutes in systems B and D are solvophobic and solvophilic, respectively. It is rather difficult to characterize the solute in system C. However, the solvation free energy of the solute, another good criterion, takes a large, positive value for system C (due to the large solute size), and we consider the solute weakly solvophobic. The solute in system B is then referred to as "a highly solvophobic solute." The solute in system C could be considered neutral or even weakly solvophilic. However, what is the most important is that the solute in system C (an unfolded protein) must be significantly less solvophilic than the solute in system D (the folded protein). This requirement is certainly met by the settings, $\xi_1/(k_BT) = 1.5$ and $\xi_1/(k_BT) = 3.0$, respectively.

We have verified that the qualitative aspects of the results described in the previous paragraph are also true for different values of d_B (e.g., $d_B = 3d_S$). Further, they are not altered even when a curved or flat surface is treated; the surfaces with $\xi_I/(k_BT) = 0.0$, $\xi_I/(k_BT) = 1.5$, and $\xi_I/(k_BT) = 3.0$ ($I = 2, 3$) can be considered highly solvophobic, weakly solvophobic, and solvophilic, respectively. All we need is that the vessel inner surfaces in stages 1 and 4 are significantly less solvophilic than those in stages 2 and 3. This requirement is certainly met by the settings, $\xi_I/(k_BT) = 1.5$ and $\xi_I/(k_BT) = 3.0$, respectively.

[1] J. S. Weissman, C. M. Hohl, O. Kovalenko, Y. Kashi, S. Chen, K. Braig, H. R. Saibil, W. A. Fenton, and A. L. Horwich, Cell **83**, 577 (1995).
[2] K. L. Nielsen and N. J. Cowan, Mol. Cell **2**, 93 (1998).
[3] H. S. Rye, A. M. Roseman, S. Chen, K. Furtak, W. A. Fenton, H. R. Saibil, and A. L. Horwich, Cell **97**, 325 (1999).
[4] A. Brinker, G. Pfeifer, M. J. Kerner, D. J. Naylor, F. U. Hartl, and M. Hayer-Hartl, Cell **107**, 223 (2001).
[5] A. L. Horwich, W. A. Fenton, E. Chapman, and G. W. Farr, Annu. Rev. Cell Dev. Biol. **23**, 115 (2007).
[6] T. K. Chaudhuri, V. K. Verma, and A. Maheshwari, Prog. Biophys. Mol. Biol. **99**, 42 (2009).
[7] S. Tanaka, Y. Kawata, G. Otting, N. E. Dixon, K. Matsuzaki, and M. Hoshino, Biochim. Biophys. Acta **1804**, 866 (2010).
[8] Z. Xu, A. L. Horwich, and P. B. Sigler, Nature (London) **388**, 741 (1997).
[9] K. L. Nielsen, N. McLennan, M. Masters, and N. J. Cowan, J. Bacteriol. **181**, 5871 (1999).
[10] H. Saibil and S. Wood, Curr. Opin. Struct. Biol. **3**, 207 (1993).
[11] D. K. Clare, P. J. Bakkes, H. van Heerikhuizen, S. M. van der Vies, and H. R. Saibil, Nature (London) **457**, 107 (2009).
[12] R. Kanno, A. Koike-Takeshita, K. Yokoyama, H. Taguchi, and K. Mitsuoka, Structure (London) **17**, 287 (2009).
[13] M. Shtilerman, G. H. Lorimer, and S. W. Englander, Science **284**, 822 (1999).
[14] F. Takagi, N. Koga, and S. Takada, Proc. Natl. Acad. Sci. U.S.A. **100**, 11367 (2003).
[15] W. Xu, J. Wang, and W. Wang, Proteins **61**, 777 (2005).
[16] K. Amano and M. Kinoshita, Chem. Phys. Lett. **488**, 1 (2010).
[17] K. Amano and M. Kinoshita, Chem. Phys. Lett. **504**, 221 (2011).
[18] D. Beglov and B. Roux, J. Chem. Phys. **103**, 360 (1995).
[19] M. Ikeguchi and J. Doi, J. Chem. Phys. **103**, 5011 (1995).
[20] M. Kinoshita and T. Oguni, Chem. Phys. Lett. **351**, 79 (2002).
[21] M. Kinoshita, J. Chem. Phys. **116**, 3493 (2002).
[22] M. Kinoshita, Chem. Phys. Lett. **387**, 54 (2004).
[23] Y. Harano and M. Kinoshita, Chem. Phys. Lett. **399**, 342 (2004).
[24] Y. Harano and M. Kinoshita, Biophys. J. **89**, 2701 (2005).
[25] M. Kinoshita, Chem. Eng. Sci. **61**, 2150 (2006).
[26] K. Amano, T. Yoshidome, M. Iwaki, M. Suzuki, and M. Kinoshita, J. Chem. Phys. **133**, 045103 (2010).
[27] M. Kinoshita, Mol. Phys. **94**, 485 (1998).
[28] M. Kinoshita, Mol. Phys. **96**, 71 (1999).
[29] M. Kinoshita, J. Chem. Phys. **128**, 024507 (2008).
[30] M. Kinoshita, S. Iba, K. Kuwamoto, and M. Harada, J. Chem. Phys. **105**, 7177 (1996).
[31] R. Roth and M. Kinoshita, J. Chem. Phys. **125**, 084910 (2006).
[32] K. Hollenstein, R. J. P. Dawson, and K. P. Locher, Curr. Opin. Struct. Biol. **17**, 412 (2007).
[33] A. Ward, C. L. Reyes, J. Yu, C. B. Roth, and G. Chang, Proc. Natl. Acad. Sci. U.S.A. **104**, 19005 (2007).
[34] V. Koronakis, A. Sharff, E. Koronakis, B. Luisi, and C. Hughes, Nature (London) **405**, 914 (2000).
[35] L. Federici, F. Walas, and B. Luisi, Curr. Sci. **87**, 190 (2004).
[36] J. L. England, D. Lucent, and V. S. Pande, J. Am. Chem. Soc. **130**, 11838 (2008).
[37] A. I. Jewett and J.-E. Shea, Cell. Mol. Life Sci. **67**, 255 (2010).
[38] J. Martin and F.-U. Hartl, Proc. Natl. Acad. Sci. U.S.A. **94**, 1107 (1997).
[39] S. G. Burston, J. S. Weissman, G. W. Farr, W. A. Fenton, and A. L. Horwich, Nature (London) **383**, 96 (1996).
[40] J. Martin, T. Langer, R. Boteva, A. Schramel, A. L. Horwich, and F.-U. Hartl, Nature (London) **352**, 36 (1991).
[41] M. Kinoshita, Front. Biosci. **14**, 3419 (2009).
[42] M. Kinoshita, Int. J. Mol. Sci. **10**, 1064 (2009).
[43] E. Hirakawa, H. Higuchi, and Y. Y. Toyoshima, Proc. Natl. Acad. Sci. U.S.A. **99**, 2533 (2000).
[44] S. M. Block, Biophys. J. **92**, 2986 (2007).
[45] J. L. Ross, M. Y. Ali, and D. M. Warshaw, Curr. Opin. Cell Biol. **20**, 41 (2008).
[46] H. Noji, R. Yasuda, M. Yoshida, and K. Kinosita, Jr., Nature (London) **386**, 299 (1997).
[47] R. Yasuda, H. Noji, K. Kinoshita, Jr., and M. Yoshida, Cell **93**, 1117 (1998).
[48] T. Masaike, F. Koyama-Horibe, K. Oiwa, M. Yoshida, and T. Nishizaka, Nat. Struct. Mol. Biol. **15**, 1326 (2008).
[49] S. Furuike, M. D. Hossain, Y. Maki, K. Adachi, T. Suzuki, A. Kohori, H. Itoh, M. Yoshida, and K. Kinoshita, Jr., Science **319**, 955 (2008).
[50] T. Yoshidome, Y. Ito, M. Ikeguchi, and M. Kinoshita, J. Am. Chem. Soc. **133**, 4030 (2011).
[51] M. Kinoshita and M. Harada, Mol. Phys. **74**, 443 (1991).
[52] M. Kinoshita, Y. Harano, and R. Akiyama, J. Chem. Phys. **125**, 244504 (2006).