

**Studies Based on Statistical Mechanics for
Mechanism of Multidrug Efflux of AcrA/AcrB/TolC**

Hirokazu Mishima

*Department of Fundamental Energy Science
Graduate School of Energy Science
Kyoto University*

February 2015

Table of Contents

Chapter 1. General Introduction

1.1 Scope and significance of this study	1
1.2. Essential roles of solvent entropy in biological self-assembly processes	
1.2.1. Self-assembly processes at molecular levels	3
1.2.2. Entropic excluded-volume effect	3
1.2.3. Driving force of protein folding	3
1.2.4. Integral equation theory	5
1.2.5. Hybrid of angle dependent integral equation theory and morphometric approach	5
1.2.6. Free energy function for a protein	6
1.2.7. Characterization of native-structure models of a protein	6
1.2.8. Justifiability of using hard-sphere solvent model	7
1.3. Essential roles of solvent entropy in biological orderly processes	
1.3.1. Orderly processes at molecular levels	7
1.3.2. Potential of mean force (PMF)	9
1.4. Studies on solute insertion and release	
1.4.1. Entropic and energetic components of PMF	9
1.4.2. Entropic insertion of a solute	10
1.4.3. Protein flux through chaperonin system	10
1.4.4. Problems unresolved	11
1.5. New developments in this study	
1.5.1. Entropic component of PMF and its relevance to multidrug efflux	12
1.5.2. Elucidation of functional-rotation mechanism and construction of energetics for AcrB	12
1.6. Synopsis of this thesis	13
References	15

Chapter 2. Entropic release of a big sphere from a cylindrical vessel

2.1. Introduction	18
2.2. Model and Theory	20
2.3. Results and Discussion	
2.3.1. Entropic potential for vessel geometry (a) in Fig. 2.1	23
2.3.2. Entropic potentials for vessel geometries (b), (c), and (d) in Fig.2.1	25
2.3.3. Entropic potentials for vessel geometries (d), (e), and (f) in Fig. 2.1	27
2.4. Concluding Remarks	28
References	30

Chapter 3. On the physics of multidrug efflux through a biomolecular complex	
3.1. Introduction	31
3.2. Model and Theory	
3.2.1. Three-dimensional integral equation theory applied to rigid-body models	34
3.2.2. Variation of vessel geometry	37
3.3. Results and Discussion	
3.3.1. Entropic potential for vessel geometry (a) in Fig. 3.2	38
3.3.2. Entropic potentials for vessel geometries (b), (c), and (d) in Fig. 3.2	42
3.3.3. Entropic potentials for vessel geometries (d), (e), and (f) in Fig. 3.2	47
3.3.4. Case where solute size differs from integral multiplication of d_s	49
3.3.5. Another method of vessel-geometry variation	54
3.3.6. Applicability of results obtained to AcrB and ABC transporter	58
3.4. Conclusions	58
References	61

Chapter 4. Statistical Thermodynamics for Functionally Rotating Mechanism of the Multidrug Efflux Transporter AcrB

4.1. Introduction	63
4.2. Functional Rotation	
4.2.1. Effects of Drug on Structural Properties of AcrB	66
4.2.2. Basic Physical Picture of Functional Rotation	67
4.2.3. Roles of Proton Action during One Cycle	68
4.2.4. Free-Energy Changes during One Cycle	70
4.2.5. Crucial Importance of Solvent-Entropy Effect	71
4.3. Model and Theoretical Method	
4.3.1. Crystal Structure of AcrB	72
4.3.2. Solvent Model	72
4.3.3. Integral Equation Theories	73
4.3.4. Solvation Entropy	73
4.3.5. Protein Model	74
4.3.6. Morphometric Approach	74
4.4. Solvation Entropy Closely Related to Packing Efficiency	
4.4.1. Packing Efficiency of Protomer	75
4.4.2. Packing Efficiency of Interface between two Protomers	75
4.5. Results and Discussion	
4.5.1. Ununiform Packing Structure of AcrB	76
4.5.2. Structural Change of Transmembrane Domain Accompanying Proton Binding	77
4.5.3. Conformational Reorganization Induced by Proton Binding or Dissociation	79
4.5.4. Functional Rotation Induced by Solvent-Entropy Effect	80

4.5.5. Structural Perturbation Caused by Proton Binding	82
4.5.6. Structural Perturbation Caused by Proton Dissociation	82
4.5.7. Significance of Trimer Formation	82
4.5.8. Comparison between AcrB and F ₁ -ATPase	83
4.5.9. Effects of Energetic Component	83
4.6. Conclusion	83
References	86
Supplement	89
Chapter 5. General Conclusion	103
Appendix	107
List of Publications	135
Copyright and Publisher's Link	136
List of Presentations	137
Acknowledgement	139

Chapter 1. General Introduction

1.1 Scope and significance of this study

A multidrug resistance of *Escherichia coli* (*E. coli*) is a serious medical problem. AcrA/AcrB/TolC tripartite complex,¹⁻⁸ which is a multidrug efflux transporter, is the strongest system for the multidrug resistance of *E. coli*. It extends from the cell inner membrane to the outer membrane and drives drugs with diverse properties from the periplasm space of the cell. The multidrug efflux transporter can handle a variety of drugs independent of their properties: sizes and affinities with water (i.e., hydrophobic or hydrophilic). This function requires a proton gradient (Fig. 1.1). AcrB trimer and TolC trimer anchor inner and outer membranes of the cell, respectively. AcrA hexamer connects the edges of these two trimers. Drugs are inserted into the pocket of AcrB protomer and released through the tunnel of TolC to the outside of the cell.

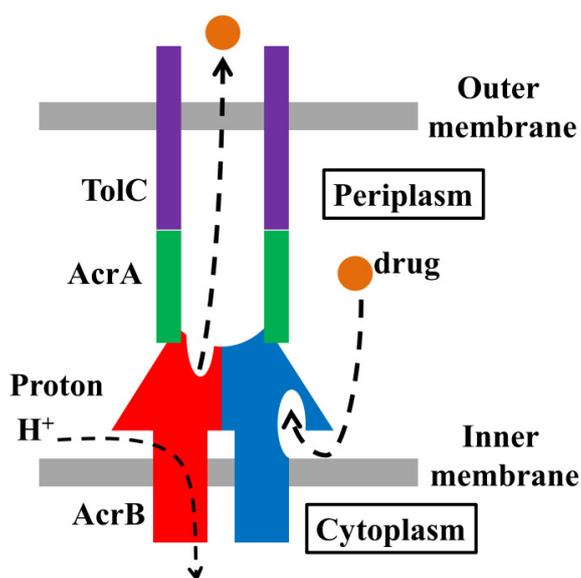


Fig. 1.1. Cartoon illustrating an efflux transporter, AcrA/AcrB/TolC.

Experimental studies based on the X-ray crystallography have shown that AcrB protomers take three different conformations which are in access (A), binding (B) and extrusion (E) states along the drug transport cycle.³⁻⁵ The functionally rotating transport picture has then been proposed for the drug efflux mechanism. In the trimer, each protomer exhibits a sequential conformational change represented as (A, B, E)→(B, E, A)→(E, A, B)→(A, B, E) by utilizing proton binding and dissociation. This picture has received much attention. TolC trimer possesses a simple cylindrical structure and works as a tunnel for

drugs.^{1,6} AcrA also works as a tunnel and is considered to transmit the conformational change of AcrB to TolC.⁷ A recent research suggests that the additional inner membrane protein, AcrZ, contributes to the multidrug ability of the transporter.⁸ A mutant strain of *E. coli* lacking AcrZ is sensitive to some, but not all, antibiotics that are exported by the tripartite complex.

There are the three principal problems to solve: (i) How can the apparently opposite events, insertion and release, successively occur for the same solute in the same system (i.e., AcrB protomer or TolC)?; (ii) what is the mechanism of the multidrug efflux?; and (iii) what are the mechanism and significance of the functional rotation of AcrB? The first problem has received very little attention so far. It is generally thought that the two events occur through simple diffusion processes. However, solutes must be transferred rapidly, unidirectionally, and successively, and this requirement can hardly be met by the simple diffusion processes. It has been suggested for the second problem that there is a multifunctional binding site within a AcrB protomer.³ A variety of solutes can bind to the site, which is in marked contrast to the well-known receptor-ligand binding. However, once a solute binds to the site, it may be difficult for the solute to be released. Further, we emphasize that the multidrug efflux of AcrA/AcrB/TolC can be realized only when TolC as well as AcrB exhibits this feature. TolC possesses an uncomplicated structure of a cylindrical shape with an open space inside it and works as a tunnel for drugs. The existence of the multifunctional binding site can hardly be applied to TolC, implying that the multidrug efflux is implemented through a different mechanism. As for the third problem, it has never been tackled so far.

In this study, we intend to solve the three problems described above using the statistical-mechanical theories developed at Kinoshita's laboratory.⁹⁻³⁵ The emphasis is placed on the entropic effect originating from the translational displacement of water molecules that coexist with the proteins and protein complexes in the system. The crucial importance of the entropic effect was pointed out by Kinoshita *et al.* for the biological self-assembly processes.⁹⁻³⁵ We show that the entropic effect plays, in different manners, imperative roles in the solute insertion and release (i.e., the switch from insertion to release), multidrug efflux, and functionally rotating mechanism of AcrB. The achievements for the solute insertion and release and for the multidrug efflux are applicable to ABC transporter as well as to TolC, and those for the functionally rotating mechanism of AcrB expand the understanding of the rotational mechanism of F₁-ATPase. (ABC transporter and F₁-ATPase, which can be categorized as ATP-driven proteins, are described in a later section.) In the course of this study, we elucidate the roles of the so-called proton-motive force (i.e., proton binding and dissociation) and the ATP hydrolysis cycle. This study thus provides a significant progress toward understanding the mechanism of functional expression by the proton-motivated and ATP-driven proteins in a unified manner.

1.2. Essential roles of water entropy in biological self-assembly processes

1.2.1. Self-assembly processes at molecular levels

A variety of self-assembly processes at molecular levels in biological systems are sustaining life. Typical examples of such processes are as follows. A long protein molecule spontaneously folds into the native structure in aqueous solution under the physiological condition. An enzyme combines only with a particular ligand and discriminates the ligand from the others, and it possesses very high ability of molecular recognition. When proteins aggregate, ordered and often symmetrical structure is formed. The membrane is completed by the penetration of proteins into the lipid bilayer. Each protein species knows where to penetrate.

1.2.2. Entropic excluded-volume effect

Kinoshita *et al.* have shown that the self-assembly processes in biological systems can be explained on the basis of the entropic effect originating from the translational displacement of water molecules.^{18,23,24,27} A solute generates a volume which the centers of water molecules cannot enter. The volume is referred to as the excluded volume (EV). When two solutes contact each other, the overlap of the EVs occurs and the total EV decreases by the overlapped volume. This is followed by increases in the total volume available to the translational displacement of water molecules and in the number of the accessible configurations of water, leading to a gain of the translational entropy of water. This factor, which we refer to as “entropic EV effect”, promotes the solute contact. It also promotes a variety of self-assembly processes. An example is argued in the next section.

The structures formed by the biological self-assembly are collapsed by the application of high pressures. Typical examples are denaturation of a protein, dissociation of filamentous actin (F-actin) into actin monomers, and destruction of amyloid fibril. At low temperatures, the power of forming the structures becomes considerably weaker. For example, a protein is unfolded and the binding of myosin to F-actin is weakened. This weakening is also relevant to the following: Upon temperature lowering, the solubility of methane increases, the critical micelle concentration becomes higher, and the average size of micelles for nonionic amphiphilic molecules becomes smaller. These phenomena are suggestive that there are common features of the biological self-assembly and a certain physical factor universally plays a dominant role as the driving force. Kinoshita *et al.* have shown that this physical factor is a gain in the translational entropy of water brought by the reduction in water crowding in the system, i.e., the water-entropy effect arising from the biomolecule-water many-body correlations.

1.2.3. Driving force of protein folding

The self-assembly processes are accompanied by a loss of entropy of biomolecules. For

example, a loss of the protein intramolecular conformational entropy occurs upon protein folding. Upon protein folding, an energetic decrease occurs due to the formation of intramolecular hydrogen bonds and van der Waals attractive interactions. However, the folding accompanies serious dehydration, that is, the break of hydrogen bonds, loss of van der Waals attractive interactions *with water molecules*, and structural reorganization of water molecules near the protein surface. Terazima *et al.* have measured the enthalpic change upon folding of apoplastocyanin (apoPC) by their novel experimental technique. The enthalpic change takes a large, positive value reaching 870 kJ/mol.¹⁸ This result implies that the dehydration penalty predominates over the decrease in the intramolecular energy. Therefore, a large gain of water entropy is required as the driving force.

When the backbone forms α -helix and β -sheet, the overlap of the EVs occurs and the total EV decreases by the overlapped volume, leading to a water-entropy gain (Fig. 1.2(a), (b)). Close packing of side chains, which reduces the total EV, also leads to a water-entropy gain (Fig. 1.2(c)). Yoshidome *et al.* at Kinoshita's laboratory have shown that the water-entropy gain upon protein folding is large enough to surpass the loss of the intramolecular conformational entropy plus the enthalpic increase.¹⁸

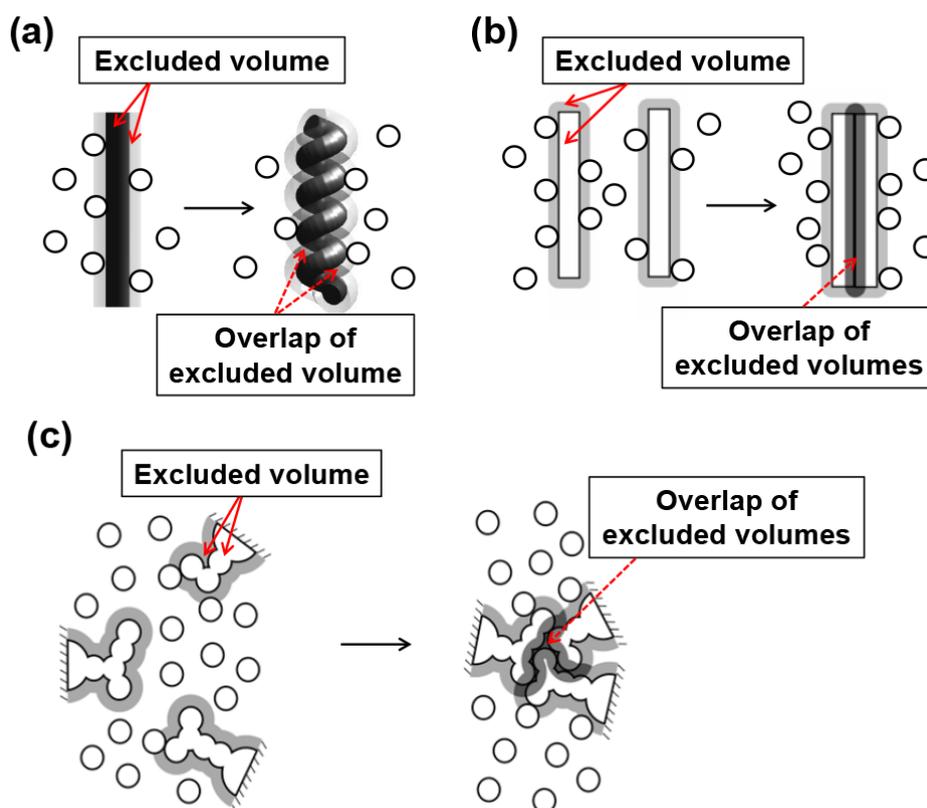


Fig. 1.2. (a) Formation of α -helix by a portion of the backbone. (b) Formation of β -sheet by portions of the backbone. (c) Close packing of side chains.

1.2.4. Integral equation theory

The integral equation theory (IET) is based on classical statistical mechanics.^{9,17,36-39} In this theory, from the system partition function, various correlation functions are defined, and the basic equations satisfied by them are derived. The many-body correlations are also approximately taken into account. As far as the equilibrium properties are concerned, the results obtained are as detailed as those from a computer simulation. In the case of bulk solvent of one component, for example, the temperature, number density, and interaction potential form the input data. By numerically solving the basic equations to obtain the correlation functions, we can calculate the microscopic structure and thermodynamic quantities. The IET is also applicable to analyses on solvation properties of a spherical solute. Thermodynamic quantities of solvation can readily be calculated from the solute-solvent correlation functions obtained via the hypernetted-chain (HNC) closure. A thermodynamic quantity of solvation is the change in the thermodynamic quantity upon solute insertion into the solvent.

When the solvent is water, the water-water and solute-water potentials and correlations are dependent not only on the distance between centers of water molecules but also on the orientation of each water molecule represented by the three Euler angles. Therefore, we use the angle-dependent IET (ADIET).^{9,17,37-39} In the ADIET the effect of the molecular polarizability is taken into account using the self-consistent mean field (SCMF) theory.^{37,38} At the SCMF level the many-body induced interactions are reduced to pairwise additive potentials involving an effective dipole moment. The effective dipole moment thus determined at 298 K and 1 atm is about 1.42 times larger than the *bare* gas-phase dipole moment.

1.2.5. Hybrid of angle dependent integral equation theory and morphometric approach

The calculation of the hydration entropy of the solute, which represents a water-entropy loss upon the protein insertion with a given structure under the isochoric condition, has been in the hub of the studies at Kinoshita's laboratory. However, the ADIET cannot be acceptable directly to the calculation of the hydration entropy of a complicated solute such as a protein because of the mathematical difficulty on dealing with the water molecular model. This problem was solved by combining the ADIET and the morphometric approach (MA).⁴⁰ The idea of the MA is to express a hydration quantity such as the hydration entropy S by the linear combination of only four geometric measures of a solute molecule:

$$S/k_B = C_1 V_{\text{ex}} + C_2 A + C_3 X + C_4 Y$$

Here, V_{ex} is the excluded volume, A is the water-accessible surface area, and X and Y are the integrated mean and Gaussian curvatures of the accessible surface, respectively. The

water-accessible surface is the surface that is accessible to the centers of water molecules. The volume that is enclosed by this surface is the excluded volume. In the approach, the solute shape enters S only via the four geometric measures.

The calculation of the hydration entropy can be finished quite rapidly by the hybrid method. The high reliability of the hybrid method has been demonstrated in the following examples: quantitative reproduction of the experimentally measured changes in thermodynamic quantities upon apoPC folding¹⁸; elucidation of the molecular mechanism of pressure²⁰ and cold²¹ denaturation of proteins; and proposal of a reliable measure of the thermal stability of proteins²².

1.2.6. Free energy function for a protein

Kinoshita *et al.* have developed a free energy function expressed by

$$F = (A - TS)/(k_B T_0), T_0 = 298\text{K},$$

where T is set at T_0 in the present study.^{16,19,30} S is the hydration entropy and A is the total dehydration penalty. Both $-S$ and A are positive quantities. A is equal to the sum of the protein intramolecular energy and the hydration energy. This energetic term is calculated as follows. Compared to the fully extended structure with $A = 0$, in a more compact structure some donors and acceptors of hydrogen bonds are buried in the interior after the break of hydrogen bonds with water molecules. There is no problem, if the intramolecular hydrogen bonds are formed. However, such hydrogen bonds are not always formed, giving rise to an energetic penalty and positive A . Our procedure of calculating A can be summarized as follows. When a donor and an acceptor are buried in the interior after the break of hydrogen bonds with water molecules, if they form an intermolecular hydrogen bond, we impose no penalty. On the other hand, when a donor or an acceptor is buried with no intermolecular hydrogen bond formed, we impose the penalty.

Yoshidome *et al.* at Kinoshita's laboratory have examined the performance of the free-energy function in discriminating the native fold from a number of misfolded decoys.^{16,19,30} The examination is carried out for a total of 133 proteins in 8 decoy sets. The free energy function is shown to be far superior to any of the previously reported functions in terms of the performance of discriminating the native fold from misfolded decoys.

1.2.7. Characterization of native-structure models of a protein

We have developed a method of characterizing the native-structure models of a protein using the free-energy function described above.³⁵ The native-structure (NS) models are constructed by a structure calculation upon which the structural information experimentally obtained as a set of constraints is imposed. Therefore, the NS models are substantially

influenced by the experimental technique and the structure calculation employed. It is not rare that there are many NS models for the same protein. Nevertheless, the characteristics of all those NS models are not always unambiguous. It is strongly desired that a reliable method, which is based on energetics of a protein accounting for its hydration thermodynamics, be developed for characterizing the NS models. Our characterization method can select the best model and point out the drawbacks of the other models. The characterization method is capable of clarifying benefits and drawbacks of the models and selecting the best one (see Appendix).

1.2.8. Justifiability of using hard-sphere solvent model

The insertion of a solute reduces the translational and rotational freedoms of water molecules whereas the self-assembly increases them. The degree of the translational-freedom change is substantially larger than that of the rotational-freedom change.^{23,24} When the discussion is limited to the water-entropy change at the normal temperature and pressure without the consideration of its temperature and pressure dependences, water can be modeled as a hard-sphere solvent, as long as the diameter and number density are set at those of water.

1.3. Essential roles of water entropy in biological orderly processes

1.3.1. Orderly processes at molecular levels

Adenosine triphosphate (ATP) is regarded as a universal currency of energy in living organisms. Motor proteins, which play imperative roles in sustaining life, function through their significantly large conformational changes induced by such chemical processes as ATP binding, ATP hydrolysis, and release of products, adenosine diphosphate (ADP) and inorganic phosphate (Pi). Myosin, a linear motor protein, moves only in one direction along filamentous actin (F-actin) similar in function to a railway track (Fig. 1.3(a)).⁴¹⁻⁴³ Myosin utilizing the ATP hydrolysis cycle does not make a random walk. ATP-binding cassette (ABC) transporter is a typical multidrug efflux transporter.⁴⁴⁻⁴⁵ It takes two conformations, inward-facing and outward-facing conformations. Harmful solutes such as drugs are inserted into ABC transporter in the inward-facing conformation from the inside of the cell. The transporter changes its conformation to the outward-facing one by the ATP binding, and the solute is then released to the outside of the cell. The transporter conformation returns to the inward-facing one after the hydrolysis and dissociation of the products, ADP and Pi (Fig. 1.3(b)). F₁-ATPase is a rotary motor comprising three α subunits, three β subunits, and one γ subunit.⁴⁶⁻⁵⁰ In the $\alpha_3\beta_3\gamma$ complex of F₁-ATPase, three α subunits and three β subunits are hexagonally arranged around the γ subunit. The rotation of the central stalk (i.e., γ subunit) of F₁-ATPase has directly been observed in single-molecule experiments and the correspondence between the chemical processes and the rotational angles of the central stalk

has been made rather clear (Fig. 1.3(c)). Each β subunit utilizes the ATP hydrolysis cycle. A chaperonin is a cylindrical vessel, and a cochaperonin acts as a lid of the vessel.⁵¹⁻⁵³ An unfolded protein is inserted into the vessel cavity from a bulk aqueous solution, the lid is attached to the vessel and the protein folds into its native structure within the closed cavity, the lid is detached after the folding is finished, and the folded protein is released back to the bulk solution (Fig. 1.3(d)). As for the drug efflux process controlled by AcrA/AcrB/TolC, drugs are unidirectionally moved from the cell. AcrB trimer, which pumps out drugs actively, rotates functionally in one direction utilizing the binding and dissociation of a proton which transfers from the high-concentration side to the low-concentration one.

The processes described in this section can be categorized as “orderly processes”. We believe that the water entropy plays the pivotal role in these orderly processes as well as in the self-assembly processes described in Sec. 1.2, but in different ways.

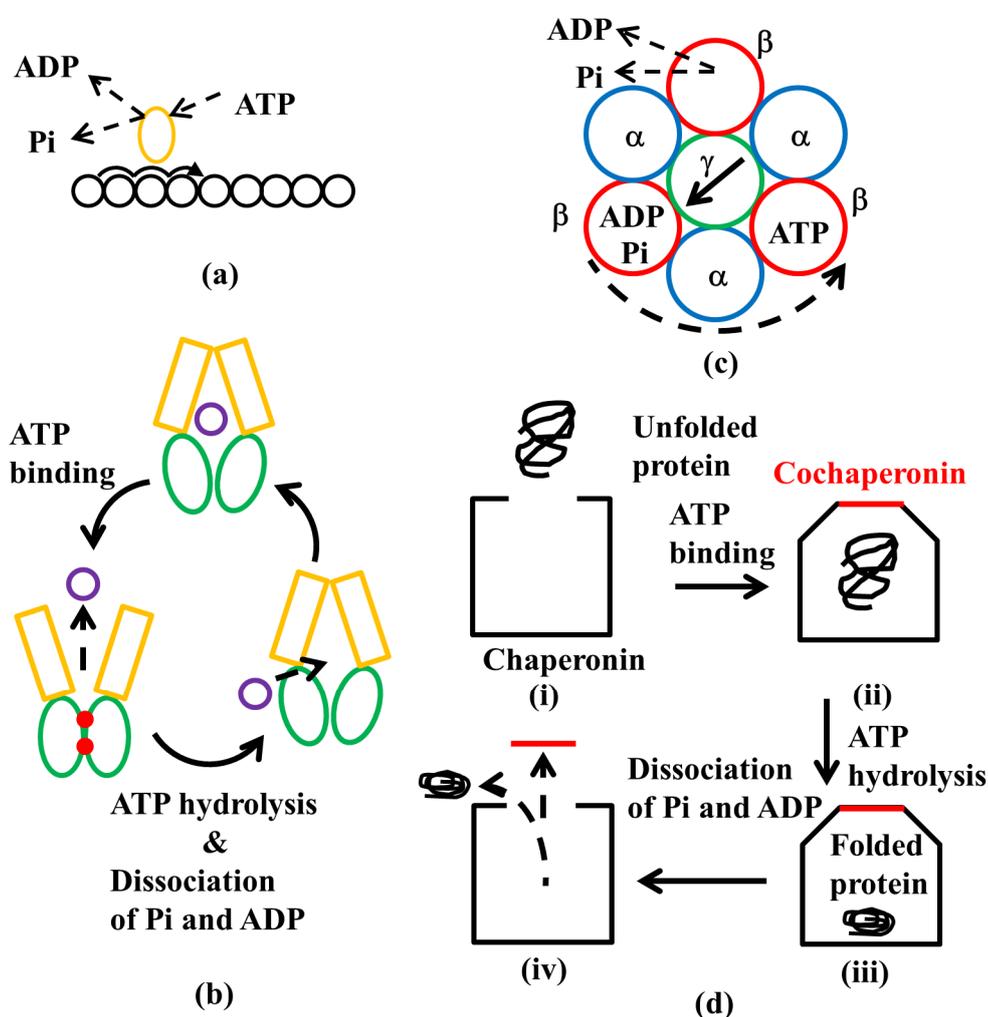


Fig. 1.3. Cartoon illustrating four orderly processes utilizing ATP hydrolysis cycle. (a) Unidirectional movement of a liner-motor protein (myosin) along F-actin. (b) Drug efflux by ABC transporter. (c) Rotation of F₁-ATPase. (d) One cycle during which a protein is assisted by chaperonin/cochaperonin to fold into the correct native structure.

1.3.2. Potential of mean force (PMF)

Amano *et al.* at Kinoshita's laboratory have shown that in the unidirectional movement of myosin head S1 along F-actin and in the assistance in protein folding performed by chaperonin, the potential of mean force (PMF) formed by water between F-actin and S1 or a chaperonin and a protein plays essential roles.^{26,28,29} The PMF is defined as “free energy of water for a fixed configuration of the solute pair” minus “that for the configuration where the two solutes are infinitely separated”. Here, “solutes” are S1 and F-actin or the chaperonin and the protein. The location of F-actin or the chaperonin is considered to be fixed, and the PMF is treated as a function of the position of S1 or the protein. We can take the view that S1 or the protein feels a potential field represented by the PMF near F-actin or the chaperonin.

We concentrate on the chaperonin-protein pair hereafter and chaperonin is referred to as “vessel”. In the vessel, the PMF becomes largely positive and largely negative with the periodicity of the molecular diameter of water, 0.28 nm. This is because the spatial distribution of the PMF cannot be described by the overlap of the excluded volumes generated by the vessel and the solute alone. It is substantially influenced by the structuring of water molecules within the domain confined by the vessel and the solute. The PMF is calculated on the assumption that water molecules are always in equilibrium with each configuration of the vessel-protein pair. This assumption is justified because the time scale of the water-structure steadying is orders of magnitude shorter than that of the protein movement. The three-dimensional integral equation theory is employed to calculate the spatial distribution of the PMF.^{10-14,25,26,28,29,54,55} It is an elaborate statistical-mechanical theory and gives the spatial distribution of the PMF in a single calculation, whereas an MD simulation gives only the value of the PMF on a single position.

1.4. Studies on solute insertion and release

1.4.1. Entropic and energetic components of PMF

Insertion of a solute into a vessel comprising biopolymers followed by release of the same solute from it is a fundamental function in biological systems. We consider the PMF formed between a solute (e.g., protein) and a vessel (e.g., chaperonin). The PMF can be regarded as the potential field which varies depending on the position of the solute when the vessel location is fixed. The PMF comprises the energetic and entropic components.

The energetic component acts for either insertion or release depending on whether the vessel inner-surface and the solute are hydrophobic or hydrophilic. When the inner surface of the vessel is hydrophilic, the number density of water molecules within the vessel cavity (especially in the immediate vicinity of the vessel inner surface) is higher than in the bulk. A hydrophilic solute is preferentially hydrated within the vessel or is favorably inserted into the vessel cavity, whereas a hydrophobic solute prefers to be hydrated in the bulk and is

favorably released from the vessel cavity. When the vessel-inner surface is hydrophobic, the number density of water molecules within the vessel cavity is lower than in the bulk. A hydrophobic solute is preferentially hydrated within the vessel and is favorably inserted into the vessel cavity (if there was no entropic effect described below, the solute could readily contact the vessel inner surface), whereas a hydrophilic solute prefers to be hydrated in the bulk and is favorably released from the vessel cavity. These energetic effects of insertion and release are induced by the solvent-solvent, solvent-vessel, and solvent-solute interactions.

On the other hand, the entropic component of the PMF is rather insensitive to affinities of the vessel inner-surface and the solute with water. It has been shown by Amano *et al.* at Kinoshita's laboratory that the entropic component always drives the solute to be inserted and the power of insertion becomes stronger as the EV generated by the solute increases.

1.4.2. Entropic insertion of a solute

Amano *et al.* at Kinoshita's laboratory have investigated the behavior of the solute coexisting with the vessel by calculating the spatial distribution of the entropic potential (i.e., the entropic component of the PMF) formed between a big sphere and an even bigger cylindrical vessel, which are immersed in small spheres forming the solvent (i.e., water).²⁵ In the calculation, the hard-body models are employed for the solute, vessel, and small particles for the following reason. In biological systems, the solvent is water characterized by hydrogen bonds. However, in the entropic gain upon the solute contact and related processes, the translational entropy predominates over the rotational entropy. The entropic effect originating from the translational displacement of water molecules can be described by modeling water as hard spheres except when its temperature or pressure dependence plays essential roles, as long as the diameter and number density of the hard-sphere solvent are set at those of water. In the hard-body models, all the accessible system configurations share the same energy, and the system behavior is purely entropic in origin. In other words, the PMF possesses only the entropic component, and we can focus on the entropic potential.

In the bulk water, no entropic potential field acting on the big sphere is formed. However, in the water confined on the scale of a nanometer, there is a spatial distribution in which the entropic potential becomes largely positive and largely negative with the periodicity of the water diameter. At the positions where the sphere is in contact with the vessel inner surface, the entropic potential becomes largely negative. However, this contact cannot be reached due to high entropic potential barrier formed. For a fixed value of the inner diameter, a big sphere whose diameter lies in a specific range is spontaneously inserted into the vessel and strongly confined within a small space almost at the center of the vessel cavity.

1.4.3. Protein flux through chaperonin system

Amano *et al.* at Kinoshita's laboratory have explained the function of a chaperonin and a

cochaperonin by calculating the PMF with a simple model.^{28,29} The chaperonin is a cylindrical vessel and the cochaperonin acts as a lid of the vessel. An unfolded protein is inserted into the vessel cavity from the bulk aqueous solution, the lid is attached to the vessel and the protein folds into its native structure within the closed cavity, the lid is detached after the folding is finished, and the folded protein is released back to the bulk solution. This function is exhibited by utilizing the ATP hydrolysis cycle (Fig. 1.3d). The inner surface of the open vessel (Fig. 1.3d(i), (iv)) is hydrophobic and that of the closed vessel (Fig. 1.3d(ii), (iii)) is hydrophilic. An unfolded protein has a hydrophobic surface while a folded protein has a hydrophilic surface. The size of the unfolded protein is larger than that of the folded protein. The chaperonin is modeled as a cylindrical vessel with the hydrophobic inner surface and the closed chaperonin is modeled as that with the hydrophilic surface. The model of the unfolded protein is a large, spherical solute with hydrophobic surface. The model of the folded protein is a spherical solute with hydrophilic surface. The EV generated by the folded protein is smaller than that of the unfolded protein. Although these models are simple, they still capture essential physics of the protein flux through the chaperonin system. The spatial distribution of the PMF between the solute and the vessel is calculated using the three-dimensional integral equation theory. Both of the energetic and entropic components of the PMF drive the unfolded protein to be inserted into the vessel cavity. The protein is confined within a small space by the entropic component. After protein folding, the energetic component of the PMF drives the protein to be released and the insertion power by the entropic component is weakened because of the decreased EV of the protein. As a result, the net effect of the PMF acts for release. Thus, the switch from insertion to release is achieved by the change in the structural properties and affinity with water of the protein arising from the folding.

1.4.4. Problems unresolved

We expect that the multidrug efflux mechanism can be also explained by the PMF. If the insertion and release processes in the multidrug efflux transporter were controlled by the energetic component of the PMF like the chaperonin, it would be impossible to handle both of hydrophobic and hydrophilic solutes because the affinities of drugs with water and their sizes remain unchanged unlike proteins. Therefore, the switch from insertion to release in the multidrug efflux transporter should occur through a different mechanism. To explain the multidrug efflux process, we must answer the two questions: (i) How can the apparently opposite events, insertion and release, successively occur for the same solute in the same system?; and (ii) what is the mechanism of the multidrug efflux? The second problem can potentially be explained when the entropic component of the PMF dominates, because the entropic component of the PMF is rather insensitive to the affinity of the solute with water. However, Amano *et al.* at Kinoshita's laboratory have shown that the entropic component of the PMF always drives a solute to be inserted and confined within the vessel cavity.²⁵ It

follows that we have to show how to release the inserted solute *entropically*. For example, X-ray crystallography has shown that ABC transporter takes two conformations, inward-facing and outward facing. It is probable that the transporter changes its conformation from the inward-facing one to the outward-facing one upon the ATP binding to it. This conformational change is a potentially important key to the entropic release because the spatial distribution of the entropic potential depends on the vessel geometry.

As explained above, it is desired for a protein or a complex of proteins that the backbones and side chains be closely (efficiently) packed, like a three-dimensional jigsaw puzzle. However, this is not always possible, depending on the amino-acid sequence. Even in cases where the overall close packing is not achievable, there are certainly the portions that can closely be packed. It is important to pack such portions preferentially. Thus, the packing efficiency of a protein or a protein complex is often ununiform. Yoshidome *et al.* at Kinoshita's laboratory have developed a physical picture of the rotation mechanism of the F₁-ATPase on the basis of the water-entropy effect.^{32,33} They find packing ununiformity in F₁-ATPase by the calculation of the solvation entropies of portions of F₁-ATPase using the hybrid of the ADIET for a molecular model for water and the morphometric approach. It is interesting that uniform packing results in a looser, overall packing and lower water entropy, which is avoided. The ununiformity plays the critical role for the rotation. The functional rotation of AcrB trimer resembles the rotation of the F₁-ATPase. However, it is unclear how the fundamental idea of the mechanism of the latter can be applied to that of the former.

1.5. New developments in this study

1.5.1. Entropic component of PMF and its relevance to multidrug efflux

The mechanism of the switch from insertion to release in the chaperonin/cochaperonin system cannot be applied to the multidrug efflux transporter. This is because the affinities with water and sizes of drugs remain unchanged unlike in cases of proteins. If the energetic component of the PMF controls the insertion and release of the solute, the switch from insertion to release is made possible by the change in the affinity of the vessel inner surface with water. However, this energetic mechanism cannot handle both of the hydrophobic and hydrophilic solutes in the same manner. On the other hand, the entropic component of the PMF is suitable to the multidrug efflux because it almost equally acts on a solute independent of its affinity with water. When the vessel inner surface is neither hydrophobic nor hydrophilic, the water within the vessel cavity and that in the bulk offer almost the same environment to any solute with hydrophobicity or hydrophilicity. As a consequence, the energetic component becomes much smaller than the entropic component. The entropic insertion of the solute has already been demonstrated as mentioned above. The multidrug efflux process is realized by the mechanism which releases the inserted solute by the entropic

component of the PMF. The entropic release should be achieved by the manner which is acceptable to a simple conformational vessel such as TolC.

In this study, for the first time, we show that the solute release can be achieved entropically. The entropic component of the PMF always acts for insertion when the vessel geometry is fixed as Amano *et al.* at Kinoshita's laboratory showed, and it was also confirmed by them that the entropic release cannot be realized by any stepwise conformational change of the vessel. The entropic component of the PMF can act for release only by a continuous variation of the vessel geometry. The result is applicable to ABC transporter as well because the conformational change of ABC transporter should occur not stepwise but continuously. We also elucidate the multidrug efflux mechanism.

1.5.2. Elucidation of functional-rotation mechanism and construction of energetics for AcrB

Up to now, molecular dynamics simulations have been performed for AcrB rather extensively. However, any of them is focused on a particular aspect or an elementary process in the drug transport cycle. By contrast, we present a theoretical analysis and construct *energetics* on the basis of statistical thermodynamics for the *whole* functional rotation cycle to unravel its mechanism. The solvent entropy is regarded as the key quantity which is correlated with the conformational change of AcrB protomer because it is most sensitive to the geometric feature of the protein. The solvation entropy is calculated by the combination of the integral equation theories and the morphometric approach. Significant ununiformity is found in the packing efficiency of AcrB and this ununiformity is shown to play the pivotal role.

At the same time, we show a new interpretation of the so-called proton motive force and discuss the similarity between the AcrB functional rotation mechanism and the F₁-ATPase rotation mechanism from a thermodynamic viewpoint. The result provides a significant progress toward understanding the mechanism of functional expression by the so-called proton-motivated and ATP-driven proteins in a unified manner.

1.6. Synopsis of this thesis

In Chapter 2, we explore the possibility of releasing the solute by the entropic component of the PMF alone. The entropic component of the PMF can be expected to realize the multidrug efflux when it predominates over the energetic component. In this study, we consider the PMF possessing only the entropic component. The system model is composed of a big sphere and a cylindrical vessel with open ends immersed in small spheres, and they are represented by rigid bodies. The vessel location is fixed and the PMF is regarded as the entropic potential field. Unlike in the chaperonin/cochaperonin system, the vessel geometry is

made variable after the solute insertion. We argue that a continuous variation of the vessel geometry is the key factor of release.

In Chapter 3, we show that solutes with a wide range of sizes are entropically released using the same vessel-geometry variation. When the vessel inner surface is neither hydrophobic nor hydrophilic, the water within the vessel cavity and the bulk water offer almost the same environment to any solute with hydrophobicity and hydrophilicity, and the energetic component becomes much less powerful than the entropic component. The entropic component of the PMF is considerably dependent on the solute size. Therefore, to realize the multidrug efflux by the water-entropy effect, it is necessary to show that the solutes with a variety of sizes can be released in the same manner. The model calculation of the entropic potential field is similar to that in Chapter 2 except that the diameter of the sphere is made variable for handling drugs with a variety of sizes. We treat a model of TolC, a cylindrical vessel possessing an entrance at one end and an exit at the other end for the solute.

In Chapter 4, we investigate the conformational property of AcrB on the basis of the solvent-entropy effect. The solvent-entropy effect in this chapter originates from the translational displacement of water molecules and CH₂, CH₃, and CH groups constituting nonpolar chains of lipid molecules. Although the proton binding and dissociation is utilized in the functional rotation of AcrB, its significance is still unknown. We outline the whole cycle to understand the mechanism and energetics of the functional rotation, using our theoretical method consisting of the integral equation theories combined with the morphometric approach. We calculate the solvation entropies of portions of AcrB and evaluate the packing efficiency of each protomers and interfaces. It is found that the packing efficiency is not uniform in AcrB trimer, and this ununiformity often plays crucially important roles in its functioning. We propose the new interpretation of the roles of the so-called proton motive force. The similarities to the rotation mechanism of F₁-ATPase are also discussed in detail and the roles of the ATP hydrolysis cycle are newly interpreted.

Finally, some of the important conclusions are recapitulated and the subjects to be pursued in future studies are briefly described in Chapter 5.

References

- ¹V. Koronakis, A. Sharff, E. Koronakis, B. Luisi, and C. Hughes, *Nature* **405**, 914 (2000).
- ²S. Murakami, R. Nakashima, E. Yamashita, and A. Yamaguchi, *Nature* **419**, 587 (2002).
- ³S. Murakami, R. Nakashima, E. Yamashita, T. Matsumoto, and A. Yamaguchi, *Nature* **443**, 173 (2006).
- ⁴M. A. Seeger, A. Schiefner, T. Eicher, F. Verrey, K. Diederichs, and K. M. Pos, *Science* **313**, 1295 (2006).
- ⁵G. Sennhauser, P. Amstutz, C. Briand, O. Storchenegger, and M. G. Grütter, *PLoS Biol.* **5**, e7 (2007).
- ⁶L. Vaccaro, K. A. Scott, and M. S. P. Sansom, *Biophys. J.* **95**, 5681 (2008).
- ⁷Y. Xu, M. Lee, A. Moeller, S. Song, B.-Y. Yoon, H.-M. Kim, S. Y. Jun, K. Lee, and N.-C. Ha, *J. Biol. Chem.* **286**, 17910 (2011).
- ⁸D. Du, Z. Wang, N. R. James, J. E. Voss, E. Klimont, T. Ohene-Agyei, H. Venter, W. Chiu, and B. F. Luisi, *Nature* **509**, 512 (2014).
- ⁹M. Kinoshita, D. R. Bérard, *J. Comput. Phys.* **124**, 230 (1996).
- ¹⁰M. Kinoshita and T. Oguni, *Chem. Phys. Lett.* **351**, 79 (2002).
- ¹¹M. Kinoshita, *J. Chem. Phys.* **116**, 3493 (2002).
- ¹²M. Kinoshita, *Chem. Phys. Lett.* **387**, 47 (2004).
- ¹³Y. Harano and M. Kinoshita, *Biophys. J.* **89**, 2701 (2005).
- ¹⁴M. Kinoshita, *Chem. Eng. Sci.* **61**, 2150 (2006).
- ¹⁵R. Roth, Y. Harano, and M. Kinoshita, *Phys. Rev. Lett.* **97**, 078101 (2006).
- ¹⁶Y. Harano, R. Roth, Y. Sugita, M. Ikeguchi, and M. kinoshita, *Chem. Phys. Lett.* **437**, 112 (2007).
- ¹⁷M. Kinoshita, *J. Chem. Phys.* **128**, 024507 (2008).
- ¹⁸T. Yoshidome, M. Kinoshita, S. Hirota, N. Baden, and M. Terazima, *J. Chem. Phys.* **128**, 225104 (2008).
- ¹⁹T. Yoshidome, K. Oda, Y. Harano, R. Roth, Y. Sugita, M. Ikeguchi, and M. Kinoshita, **77**, 950 (2009).
- ²⁰T. Yoshidome, Y. Harano, and M. Kinoshita, *Phys. Rev. E* **79**, 011912 (2009).
- ²¹T. Yoshidome and M. Kinoshita, *Phys. Rev. E* **79**, 030905(R) (2009).
- ²²K. Amano, T. Yoshidome, K. Oda, Y. Harano, M. Kinoshita, *Chem. Phys. Lett.* **474**, 190 (2009).
- ²³M. Kinoshita, *Front. Biosci.* **14**, 3419 (2009).
- ²⁴M. Kinoshita, *Int. J. Mol. Sci.* **10**, 1064 (2009).
- ²⁵K. Amano and M. Kinoshita, *Chem. Phys. Lett.* **488**, 1 (2010).(insertion)
- ²⁶K.-i. Amano, T. Yoshidome, M. Iwaki, M. Suzuki, and M. Kinoshita, *J. Chem. Phys.* **133**, 045103 (2010).

- ²⁷S. Yasuda, T. Yoshidome, H. Oshima, R. Kodama, Y. Harano, and M. Kinoshita, *J. Chem. Phys.* **132**, 065105 (2010).
- ²⁸K. Amano and M. Kinoshita, *Chem. Phys. Lett.* **504**, 221 (2011).
- ²⁹K. Amano, H. Oshima, and M. Kinoshita, *J. Chem. Phys.* **135**, 185101 (2011).
- ³⁰S. Yasuda, T. Yoshidome, Y. Harano, R. Roth, H. Oshima, K. Oda, Y. Sugita, M. Ikeguchi, M. Kinoshita, *Proteins*, **79**, 2161, (2011).
- ³¹H. Oshima, S. Yasuda, T. Yoshidome, M. Ikeguchi, and M. Kinoshita, *Phys. Chem. Chem. Phys.* **13**, 16236 (2011).
- ³²T. Yoshidome, Y. Ito, M. Ikeguchi, and M. Kinoshita, *J. Am. Chem. Soc.* **133**, 4030 (2011).
- ³³T. Yoshidome, Y. Ito, N. Matubayasi, M. Ikeguchi, and M. Kinoshita, *J. Chem. Phys.* **137**, 035102 (2012).
- ³⁴S. Yasuda, H. Oshima, and M. Kinoshita, *J. Chem. Phys.* **137**, 135103 (2012).
- ³⁵H. Mishima, S. Yasuda, T. Yoshidome, H. Oshima, Y. Harano, M. Ikeguchi, and M. Kinoshita, *J. Phy. Chem. B* **116**, 7776 (2012).
- ³⁶J.-P. Hansen and I. R. McDonald, *Theory of Simple Liquids*, 3rd ed. Academic Press: London, (2006).
- ³⁷P. G. Kusalik and G. N. Patey, *J. Chem. Phys.* **88**, 7715 (1988).
- ³⁸P. G. Kusalik and G. N. Patey, *Mol. Phys.* **65**, 1105 (1988).
- ³⁹N. M. Cann and G. N. Patey, *J. Chem. Phys.* **106**, 8165 (1997).
- ⁴⁰P.-M. König, R. Roth, and K. R. Mecke, *Phys. Rev. Lett.* **93**, 160601 (2004).
- ⁴¹K. Kitamura, M. Tokunaga, A. H. Iwane, and T. Yanagida, *Nature* **397**, 129 (1999).
- ⁴²K. Kitamura, M. Tokunaga, S. Esaki, A. H. Iwane, and T. Yanagida, *Biophysics* **1**, 1 (2005).
- ⁴³T. Yanagida, M. Iwaki, and Y. Ishii, *Philos. Trans. R. Soc. London, Ser. B* **363**, 2123 (2008).
- ⁴⁴K. Hollenstein, R. J. P. Dawson, and K. P. Locher, *Curr. Opin. Struct. Biol.* **17**, 412 (2007).
- ⁴⁵A. Ward, C. L. Reyes, J. Yu, C. B. Roth, and G. Chang, *Proc. Natl. Acad. Sci. U.S.A.* **104**, 19005 (2007).
- ⁴⁶J. P. Abrahams, A. G. Leslie, R. Lutter, and J. E. Walker, *Nature* **370**, 621 (1994).
- ⁴⁷H. Noji, R. Yasuda, M. Yoshida, and K. Kinoshita, Jr. *Nature* **386**, 299 (1997).
- ⁴⁸S. Furuike, M. D. Hossain, Y. Maki, K. Adachi, T. Suzuki, A. Kohori, H. Itoh, M. Yoshida, and K. Kinoshita, Jr. *Science* **319**, 955 (2008).
- ⁴⁹S. Toyabe, T. Okamoto, T. Watanabe-Nakayama, H. Taketani, S. Kudo, and E. Muneyuki, *Phys. Rev. Lett.* **104**, 198103 (2010).
- ⁵⁰T. Uchihashi, R. Iino, T. Ando, and H. Noji, *Science* **333**, 755 (2011).
- ⁵¹A. L. Horwich, W. A. Fenton, E. Chapman, and G. W. Farr, *Annu. Rev. Cell Dev. Biol.* **23**, 115 (2007).
- ⁵²T. K. Chaudhuri, V. K. Verma, and A. Maheshwari, *Prog. Biophys. Mol. Biol.* **99**, 42

(2009).

⁵³S. Tanaka, Y. Kawata, G. Otting, N. E. Dixon, K. Matsuzaki, and M. Hoshino, *Biochim. Biophys. Acta* **1804**, 866 (2010).

⁵⁴D. Beglov and B. Roux, *J. Chem. Phys.* **103**, 360 (1995).

⁵⁵M. Ikeguchi and J. Doi, *J. Chem. Phys.* **103**, 5011 (1995).

Chapter 2. Entropic release of a big sphere from a cylindrical vessel

2.1. Introduction

Insertion of a solute into a vessel comprising biopolymers followed by release of the same solute from it is a fundamental function in biological systems. It is quite interesting that the two apparently *opposite* events, insertion and release, successively occur in a system. A typical example of such an insertion/release process is found for TolC: an important component of the tripartite efflux system, AcrA/AcrB/TolC¹⁻³. AcrB is an efflux pump protein, TolC is a cylindrical vessel possessing an entrance at one end and an exit at the other end for the solute, and AcrA mediates the contact between AcrB and TolC. AcrB interacting with TolC at the entrance sends the solute to the central position within the vessel cavity of TolC, and then the solute is moved to the exit.

In earlier works⁴⁻⁶, we showed that the solute-vessel potential of mean force (PMF denoted by Φ) formed by the solvent plays imperative roles for insertion. The PMF represents “the free energy of the solvent for a fixed configuration of the solute-vessel pair” minus “that for the configuration where the solute is infinitely far from the vessel”. It is physically insightful to decompose the PMF scaled by $k_B T$ (k_B is Boltzmann’s constant and T is the absolute temperature), $\Phi/(k_B T)$, into its energetic and entropic components denoted by $\Phi_E/(k_B T)$ and $-\Phi_S/k_B$, respectively: $\Phi/(k_B T) = \Phi_E/(k_B T) - \Phi_S/k_B$.

A target of our earlier works was the protein flux through a chaperonin system⁵⁻⁹: An unfolded protein is inserted into the chaperonin from bulk aqueous solution, protein folding occurs within the chaperonin cavity, and the folded protein is released back to the bulk. The conclusions drawn were as follows^{5,6}.

(1) $-\Phi_S/k_B$ represents the solvent-entropy effect originating from the translational displacement of solvent molecules. It is closely related to the excluded volume (EV) generated by the solute. (The solute generates a space which the centers of solvent molecules cannot enter, and the volume of this space is the EV.) $-\Phi_S/k_B$ is rather insensitive to the solute-solvent and vessel inner surface-solvent affinities, namely, to whether the solute or the vessel inner surface is solvophobic or solvophilic (on condition that they are neither too solvophobic nor too solvophilic). $-\Phi_S/k_B$ always drives the solute to be inserted into the vessel cavity and constrained within a small space almost in the center. The power of insertion and constraint becomes stronger as the EV increases.

(2) $\Phi_E/(k_B T)$ is strongly dependent on the solute-solvent and vessel inner surface-solvent affinities. The roles of $\Phi_E/(k_B T)$ can be summarized as follows. When the inner surface of the vessel is solvophilic, the solvent number density within the vessel cavity is higher than that in the bulk aqueous solution. As a result, a solvophilic solute is preferentially solvated within the cavity, whereas a solvophobic solute is more stabilized in the bulk solution. When the inner surface is solvophobic, the solvent number density within the cavity is lower than that

in the bulk. Consequently, a solvophobic solute is more stabilized within the cavity, whereas a solvophilic solute is preferentially solvated in the bulk.

(3) As for the protein flux through a chaperonin system, when a protein is inserted or released, the inner surface of chaperonin is weakly hydrophobic⁵⁻⁹. An unfolded protein with large EV and weak hydrophobicity is driven to be weakly inserted by $\Phi_E/(k_B T)$ and strongly inserted by $-\Phi_S/k_B$, and the net action is insertion. The switch from insertion to release is achieved by altering the protein properties as follows. The folding reduces the EV and makes the protein more hydrophilic. $-\Phi_S/k_B$ and the power of insertion then become weaker, with the result that $\Phi_E/(k_B T)$ dominates. The folded protein is released back to the bulk aqueous solution to become preferentially hydrated in the bulk. The switch from insertion to release is thus realized, and $-\Phi_S/k_B$ drives insertion while $\Phi_E/(k_B T)$ is requisite in release.

In TolC, by contrast, the solute properties remain unchanged for insertion and release, and the switch from insertion to release should occur through a different mechanism. AcrA/AcrB/TolC extrudes a variety of drug molecules (i.e., nonpolar, polar, and charged solutes), and this variability is known as “multidrug efflux”¹⁻³. However, its microscopic mechanism remains rather mysterious and has not been elucidated yet. We emphasize that the multidrug efflux of AcrA/AcrB/TolC can be performed only when TolC as well as AcrB exhibits this feature. In this Chapter, we are concerned with the multidrug efflux of TolC. The multidrug efflux is indicative that its mechanism cannot be chemically specific but rather has to be based on a physical factor. We show that this factor is a time-dependent entropic force acting on a solute, which originates from the solvent-mediated interaction between the solute and TolC whose geometry is time dependent. The model system employed is very simple but takes this important factor into consideration.

On the basis of the results summarized in (2), if $\Phi_E/(k_B T)$ dominates, we can conclude the following. In cases where the inner surface of the vessel is hydrophilic, only hydrophilic solutes are inserted, and they are released when the inner surface turns hydrophobic. In cases where the inner surface is hydrophobic, only hydrophobic solutes are inserted, and they are released when the inner surface turns hydrophilic. It is thus difficult to treat both of hydrophilic and hydrophobic solutes: The multidrug efflux cannot be exhibited if $\Phi_E/(k_B T)$ dominates. To realize the multidrug efflux, the PMF must be governed by $-\Phi_S/k_B$ with the insensitivity mentioned in (1). An important point is that when the vessel inner surface is neither hydrophilic nor hydrophobic and the solute within the cavity and that in the bulk aqueous solution share almost the same stability, $\Phi_E/(k_B T)$ is much less powerful than $-\Phi_S/k_B$ (i.e., $-\Phi_S/k_B$ dominates). This can be relevant to TolC whose inner surface possesses a mixture of predominantly nonpolar and isolated electronegative patches¹. However, $-\Phi_S/k_B$ *always* acts for insertion. A question then arises: Can a solute be released without adjusting $\Phi_E/(k_B T)$ unlike in the chaperonin case?

Here we explore the possibility of releasing the solute by $-\Phi_S/k_B$ alone. We calculate the spatial distribution of the PMF between a big sphere and a cylindrical vessel with open ends immersed in small spheres using the three-dimensional (3D) integral equation theory^{4-6,10-15} combined with rigid-body models. With the rigid-body models, all of the allowed system

configurations share the same energy and the system behavior is purely entropic in origin¹⁵. Namely, $\Phi/(k_B T)$ possesses only the entropic component: $\Phi/(k_B T) = -\Phi_S/k_B$. Unlike in our earlier works, the vessel geometry is made variable after the solute insertion. It is demonstrated that release can also be achieved by $-\Phi_S/k_B$: The solute is entropically moved from the entrance to the exit by a continuous variation of the vessel geometry. Since the spatial distribution of the PMF becomes largely positive and largely negative with the periodicity of the molecular diameter of the solvent, $d_S=0.28$ nm, even a variation of the vessel geometry within the scale of d_S leads to a drastic change in the spatial distribution, thus enabling the vessel to control the solute motion.

In biological systems, the solvent is water characterized by hydrogen bonds. However, in hydration thermodynamics of a solute, the translational entropy predominates over the rotational entropy¹⁶. The entropic effect originating from the translational displacement of water molecules can be described by modeling water as hard spheres except when its temperature dependence plays essential roles (e.g., in the elucidation of cold denaturation of a protein^{17,18}), as long as the diameter and number density of the hard-sphere solvent are set at those of water¹⁹⁻²¹. We note that the hydrogen-bonding allows water to exist as a dense liquid despite its quite a small molecular size, leading to an exceptionally large entropic effect^{20,21}.

In AcrA/AcrB/TolC, the proton motive force causes structural changes of AcrB which are transmitted to TolC through AcrA¹⁻³. This action may lead to a continuous vessel-geometry variation of TolC. A molecular dynamics (MD) simulation study² has also suggested that TolC can vary its geometric characteristics (e.g., they have observed a peristaltic motion of the periplasmic domain). However, the details of the variation during the insertion/release process are still unknown. What we emphasize in this Chapter is that release as well as insertion can be accomplished by $-\Phi_S/k_B$ which is rather insensitive to the solute-solvent and vessel inner surface-solvent affinities and that a continuous vessel-geometry variation is a key factor. This result may provide us with a clue to the microscopic mechanism of the multidrug efflux of TolC, pending further studies.

2.2. Model and Theory

$-T\Phi_S$ considered in this Chapter is essentially the so-called depletion potential whose physical essence was first discussed by the Asakura-Oosawa theory^{22,23}. The depletion potential between big bodies with simple geometries (e.g., big spheres, a big sphere and a convex, planer, or concave surface, etc.)²⁴⁻²⁸ or between big, nonspherical bodies^{13,29,30} has been analyzed using more advanced theories. A feature of the present study is that one of the big bodies possesses a complex geometry^{4-6,12-15}.

We consider a big vessel and a big sphere immersed at infinite dilution in small spheres with diameter d_S forming the solvent. The big sphere, whose diameter is denoted by d_B , corresponds to a solute. The initial geometry of the vessel is illustrated in Fig. 2.1(a). After the solute insertion, the vessel geometry is varied for the solute release as explained in Fig. 2.1(b)-(f). The Cartesian coordinate system is chosen as illustrated in Fig. 2.1(a). The cross section of $z=0$ is shown for each geometry. More details are described in a later paragraph.

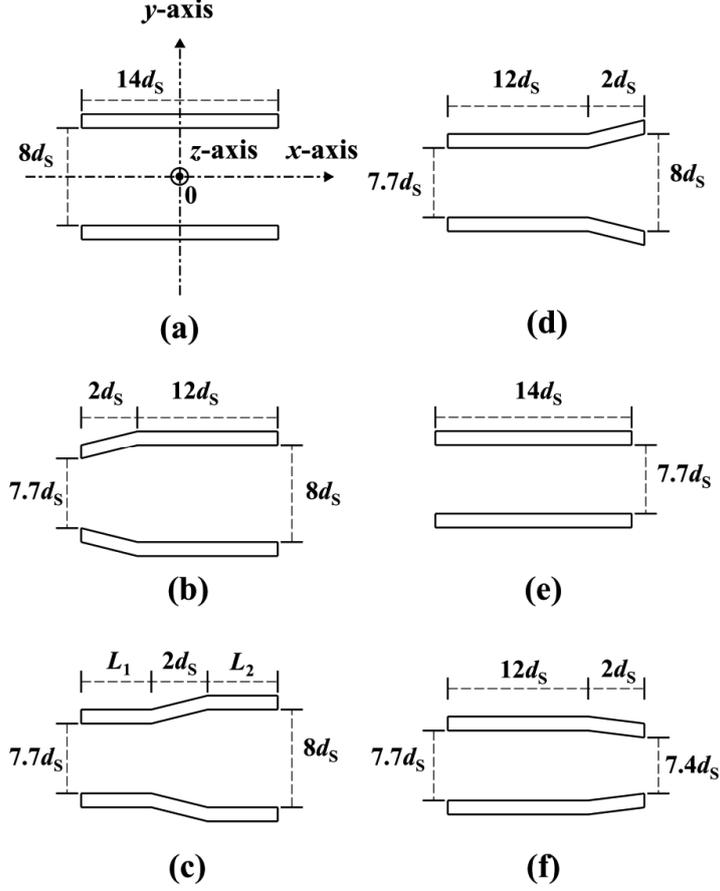


Fig. 2.1. Variation of vessel geometry: (a)→(b)→(c)→(d)→(e)→(f). The geometry in (a) is a cylinder with inner diameter $8d_s$, length $14d_s$, and thickness d_s . The coordinate system is chosen as illustrated here. The cross section of $z=0$ is shown for each geometry. Though the geometry variation is illustrated in a stepwise manner in this figure, it is made *continuously*.

The vessel with a prescribed geometry is considered. First, the vessel-solvent correlation functions are calculated using the 3D integral equation theory^{4,6,10-15} described below. Second, the solute-solvent correlation functions are calculated using the radial-symmetric HNC theory²⁵ for spherical particles. The PMF between the vessel and the solute is then calculated by assuming that the solvent particles are always in equilibrium with each configuration of the vessel-solute pair^{6,14}.

The Ornstein-Zernike (OZ) equation in the Fourier space is expressed by

$$W_{1S}(k_x, k_y, k_z) = \rho_S C_{1S}(k_x, k_y, k_z) H_{SS}(k) \quad (2.1)$$

and the hypernetted-chain (HNC) closure equation is written as

$$c_{1S}(x, y, z) = \exp\{-u_{1S}(x, y, z)/(k_B T)\} \exp\{w_{1S}(x, y, z)\} - w_{1S}(x, y, z) - 1. \quad (2.2)$$

Here, the subscripts “1” and “S” denote the vessel and the solvent, respectively, $w=h-c$, c is the direct correlation function, h the total correlation function, u the potential, and ρ_S the bulk density. The molecular diameter d_S is set at the value of water, 0.28 nm. The reduced number density $\rho_S d_S^3$ is taken to be the value of water at 298 K and 1 atm, 0.7317. The capital letters (C , H , and W) represent the Fourier transforms. $H_{SS}(k)$ ($k^2=k_x^2+k_y^2+k_z^2$) calculated using the radial-symmetric HNC theory for spherical particles is part of the input data. We emphasize that the OZ equation is *exact*. On the other hand, the bridge function is neglected in the HNC closure equation. However, it has been verified that the 3D-OZ-HNC theory gives quantitatively reliable results¹².

The numerical procedure is briefly summarized as follows: (1) $u_{1S}(x, y, z)$ is calculated at each 3D grid point, (2) $w_{1S}(x, y, z)$ is initialized to zero, (3) $c_{1S}(x, y, z)$ is calculated from Eq. (2.2), and $c_{1S}(x, y, z)$ is transformed to $C_{1S}(k_x, k_y, k_z)$ using the 3D fast Fourier transform (3D-FFT), (4) $W_{1S}(k_x, k_y, k_z)$ is calculated from Eq. (2.1), and $W_{1S}(k_x, k_y, k_z)$ is inverted to $w_{1S}(x, y, z)$ using the 3D-FFT, and (5) steps (3) and (4) are repeated until the input and output functions for $w_{1S}(x, y, z)$ become identical within convergence tolerance. On grid points where a solvent particle and the solute overlap, $\exp\{-u_{1S}(x, y, z)/(k_B T)\}$ is zero. On those where a solvent particle is in contact with the solute, it is set at 0.5, and otherwise it is unity. The grid spacing (Δx , Δy , and Δz) is set at $0.1d_S$, and the grid resolution ($N_x \times N_y \times N_z$) is $256 \times 256 \times 256$. It has been verified that the spacing is sufficiently small and the box size ($N_x \Delta x$, $N_y \Delta y$, $N_z \Delta z$) is large enough for the correlation functions at the box surfaces to be essentially zero.

First, the Fourier transform of the vessel-solvent direct correlation function denoted by $C_{1S}(k_x, k_y, k_z)$ is calculated by following the procedure described above. Second, the Fourier transforms of the solute-solvent total correlation function denoted by $H_{2S}(k)$ (the subscript “2” denotes the solute) is calculated using the radial symmetric HNC theory for spherical particles. The PMF between the vessel and the solute, $\Phi_{12}(x, y, z)$, is then obtained from

$$\Phi_{12}(x, y, z)/(k_B T) = u_{12}(x, y, z)/(k_B T) - w_{12}(x, y, z), \quad (2.3)$$

where $w_{12}(x, y, z)$ is calculated by inverting $W_{12}(k_x, k_y, k_z)$ given by

$$W_{12}(k_x, k_y, k_z) = \rho_S C_{1S}(k_x, k_y, k_z) H_{2S}(k). \quad (2.4)$$

The physical meaning of $\Phi_{12}(x, y, z)$ ^{6,14} can be understood from

$$\Phi_{12}(x, y, z) = F(x, y, z) - F(\infty, \infty, \infty) \quad (2.5)$$

and

$$g_{12}(x, y, z) = \exp\{-\Phi_{12}(x, y, z)/(k_B T)\}, \quad g_{12}(\infty, \infty, \infty) = 1. \quad (2.6)$$

$F(x, y, z)$ is the free energy of small spheres in the case where the big-sphere center is at the position (x, y, z) , and $g_{12}(x, y, z)$ the pair distribution function. For rigid-body models, the

behavior of $\Phi_{12}(x, y, z)$ is purely entropic in origin. Due to the microscopic structure of small spheres formed within the domain confined by the big sphere and the vessel, $\Phi_{12}(x, y, z)$ exhibits a complex spatial distribution. A great advantage of the 3D integral equation theory is that the values of Φ_{12} on all the grid points are obtained from only a single calculation, which is in marked contrast with the usual computer simulation.

We emphasize that the geometry variation is made *continuously* though it is presented in a stepwise manner in Fig. 2.1. Fig. 2.1(a) shows the initial geometry of the vessel, a cylinder with inner diameter $8d_S$, length $14d_S$, and thickness d_S . These dimensions roughly mimic those of TolC except that the length is set at a much shorter value (the length of TolC is $\sim 35d_S$). This is because all we need in this Chapter is a length which is sufficiently longer than the diameter of the solute d_B set at $4d_S$. The left and right ends of the cylindrical vessel are the entrance and exit, respectively. The inner diameter at the entrance is first reduced to $7.7d_S$ (Fig. 2.1(b)). L_1 and L_2 in Fig. 2.1(c) (lengths of the portions with inner diameter $7.7d_S$ and $8d_S$, respectively) are gradually increased and decreased, respectively. The length of the tapering portion is set at $2d_S$ ($L_1+L_2=12d_S$). The PMF for $L_1=L_2=6d_S$ is presented in this Chapter. After the geometry shown in Fig. 2.1(d), the inner diameter uniformly becomes $7.7d_S$ (Fig. 2.1(e)). As the final geometry, the inner diameter at the exit is further reduced to $7.4d_S$ (Fig. 2.1(f)). Hereafter, the vessel geometries shown in Fig. 2.1(a), (b), (c), (d), (e), and (f) are referred to as vessel geometries a, b, c, d, e, and f, respectively.

2.3. Results and Discussion

2.3.1. Entropic potential for vessel geometry (a) in Fig. 2.1

Hereafter, the PMF scaled by $k_B T$ between the vessel and the solute are denoted simply by $-\Phi_S/k_B$. We refer to $-\Phi_S$ as “entropic potential”. The distribution of $-\Phi_S/k_B$ on the cross section of $z=0$ for the initial vessel geometry (see Fig. 2.1(a)) is shown in Fig. 2.2(a). $-\Phi_S/k_B$ becomes lower as the color approaches dark blue, and it becomes higher as the color approaches dark red. The center of the big sphere cannot enter the domain drawn in white. The profile of $-\Phi_S/k_B$ along the x -axis ($y=z=0$) is shown in Fig. 2.2(b). There are three narrow domains within which $-\Phi_S/k_B$ is negative as indicated in Fig. 2.2(a). In general, it is difficult for the solute to overcome a free-energy barrier well exceeding $k_B T^4$ (an important exception is described in the second paragraph of Sec. 2.3.3). The solute can spontaneously be inserted into domain 2 or domain 3 due to the essentially zero barrier. The solute can also enter domain 1 without difficulty: The trajectory indicated by the black arrow in Fig. 2.2(a) possesses the lowest value of the barrier, $\sim 0.8k_B T$. However, we postulate that the solute is inserted into domain 3, the narrow domain around the central axis of the vessel (i.e., the x -axis). In the real system, AcrB interacting with TolC at its entrance sends the solute to the central position within the vessel cavity, and our postulation can thus be justified. Once the solute enters domain 3, it cannot move in the radial direction. At $x/d_S \sim -5.5$ (the position of the lowest value of $-\Phi_S$), the barrier for the solute to overcome for moving from domain 3 to domain 2 is $\sim 8.5k_B T$ and that for moving from domain 2 to domain 1 is $\sim 6.7k_B T$. These barriers remain quite high for $-6 < x/d_S < 6$. If the vessel geometry is not changed, the solute will be confined within the region of $-6 < x/d_S < 6$ (see Fig. 2.2(b)).

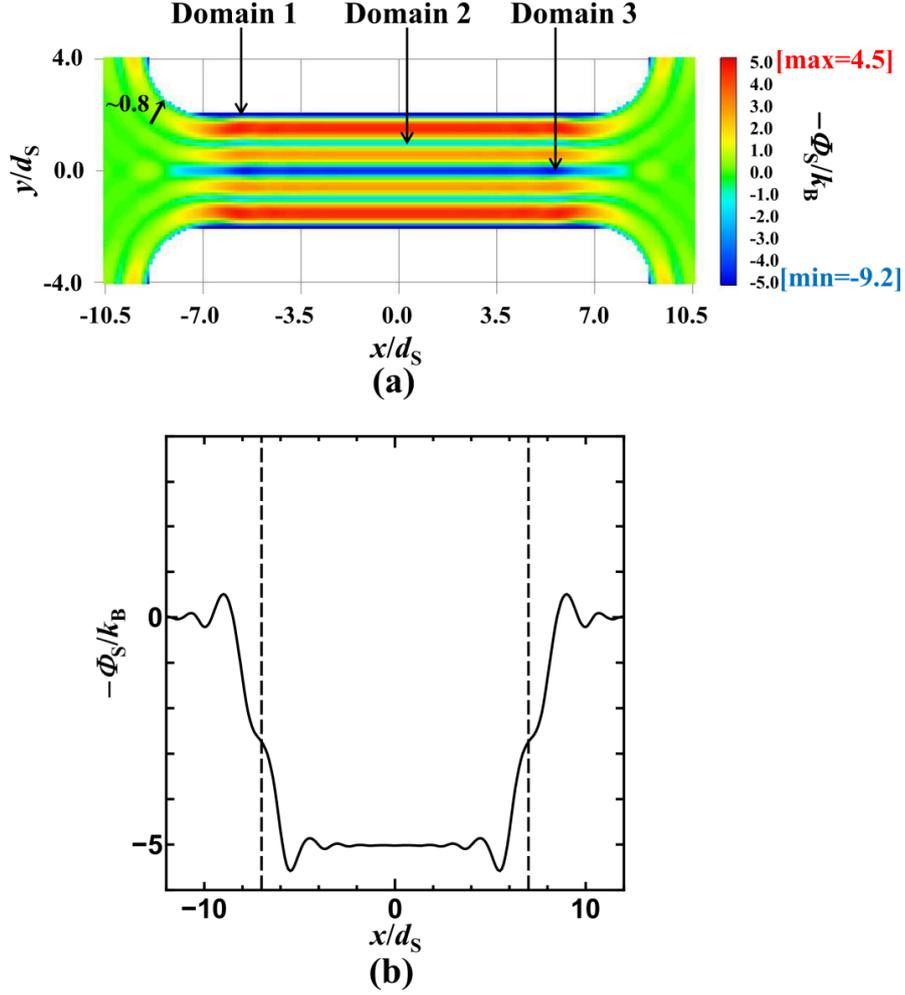


Fig. 2.2. (a) Distribution of $-\Phi_S/k_B$ on the cross section of $z=0$ for the initial vessel geometry shown in Fig. 2.1(a). $-\Phi_S/k_B$ becomes lower as the color approaches dark blue, and it becomes higher as the color approaches dark red (“max” and “min” represent the maximum and minimum values, respectively). The center of the large sphere cannot enter the domain drawn in white. “ ~ 0.8 ” represents that the free-energy barrier along the black arrow is $\sim 0.8k_B T$. (b) Profile of $-\Phi_S/k_B$ along the x -axis ($y=z=0$) for the initial vessel geometry shown in Fig. 2.1(a). The two broken lines represent positions of the vessel ends, $x/d_S=\pm 7$.

Here, we explain how the stripe pattern of the entropic potential is formed along the y -axis as observed in Fig. 2.2(a). When the separation between the nearest solute and vessel inner surfaces, which is denoted by η , is not sufficiently close to nd_S ($n=0, 1, 2, \dots$), spaces unavailable to the translational displacement of solvent molecules appear as indicated in Fig. 2.3(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. 2.3(b). The configuration in Fig. 2.3(a) is entropically unfavorable, while that in Fig. 2.3(b) is entropically favorable, leading to the stripe pattern formed along the y -axis.

Spaces unavailable to the translational displacement of solvent molecules

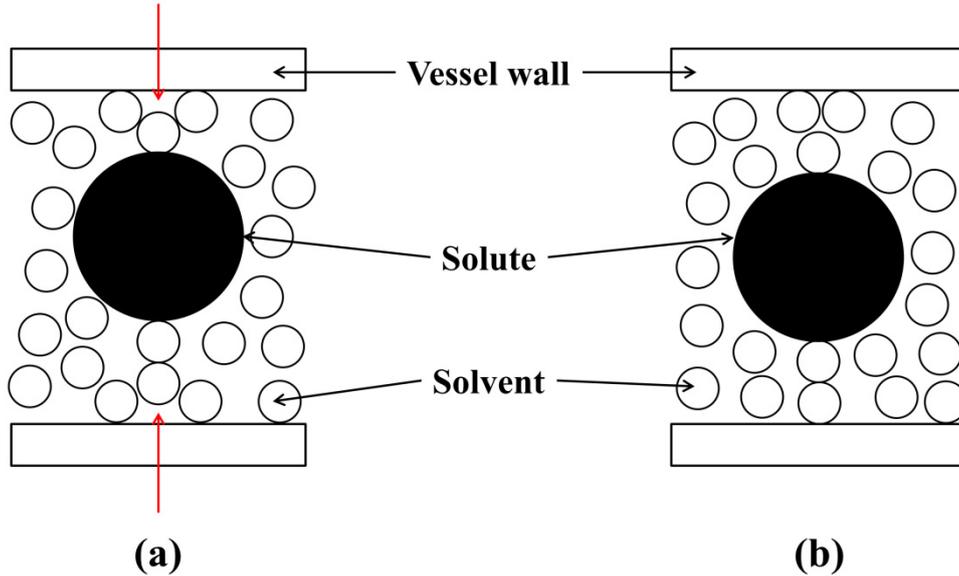


Fig. 2.3. Cartoons illustrating how the stripe pattern of the entropic potential shown in Fig. 2.2(a) is formed along the y -axis. (a) Separation between the nearest solute and vessel inner surfaces, which is denoted by η , is not sufficiently close to nd_S ($n=0, 1, 2, \dots$). (b) In the case of $\eta \sim nd_S$.

2.3.2. Entropic potentials for vessel geometries (b), (c), and (d) in Fig. 2.1

The distributions of $-\Phi_S/k_B$ on the cross section of $z=0$ for vessel geometries b, c, and d (see Fig. 2.1(b)-(d)) are shown in Fig. 2.4(a), (b), and (c), respectively. The profile of $-\Phi_S/k_B$ along the x -axis ($y=z=0$) for each entropic potential is shown in Fig. 2.5: Fig. 2.5(a), (b), and (c) correspond to Fig. 2.4(a), (b), and (c), respectively. As shown in Fig. 2.2(a) and explained in Fig. 2.3, $-\Phi_S/k_B$ becomes largely positive and largely negative with the periodicity of the molecular diameter of the solvent, $d_S=0.28$ nm. Therefore, even a variation of the vessel geometry within the scale of d_S leads to a drastic change in $-\Phi_S/k_B$. In the case of the variation illustrated in Fig. 2.1(b)-(d), $-\Phi_S/k_B$ in the narrow domain around the central axis of the vessel turns largely positive for the region with inner diameter $7.7d_S$. Since this region is gradually lengthened, the solute is driven to move in the right direction along the x -axis (see Figs. 2.4 and 2.5).

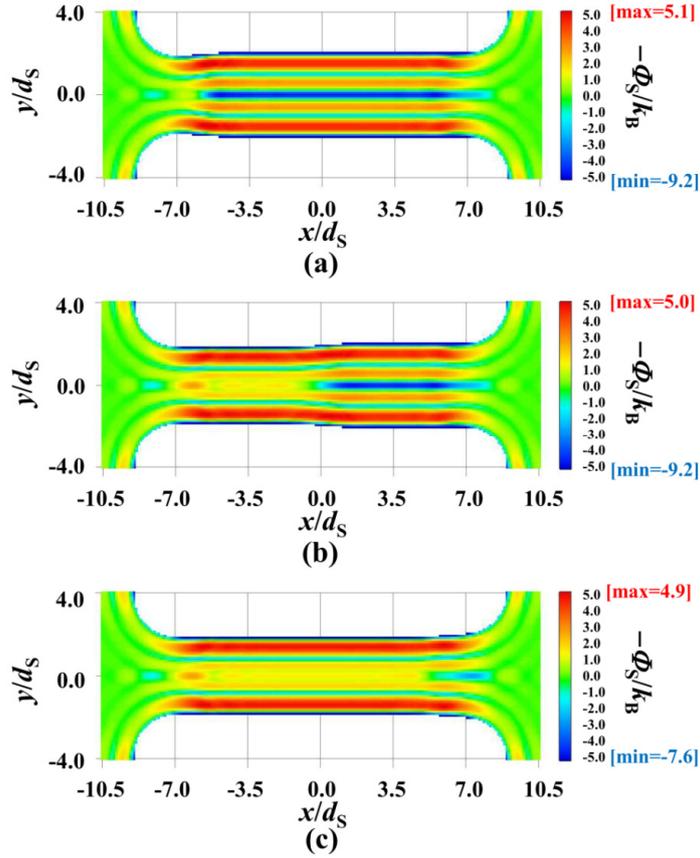


Fig. 2.4. (a) Distribution of $-\Phi_S/k_B$ on the cross section of $z=0$ for the vessel geometry shown in Fig. 2.1(b). (b) That for the vessel geometry shown in Fig. 2.1(c) (both of L_1 and L_2 are set at $6d_S$). (c) That for the vessel geometry shown in Fig. 2.1(d). $-\Phi_S/k_B$ becomes lower as the color approaches dark blue, and it becomes higher as the color approaches dark red (“max” and “min” represent the maximum and minimum values, respectively). The center of the large sphere cannot enter the domain drawn in white.

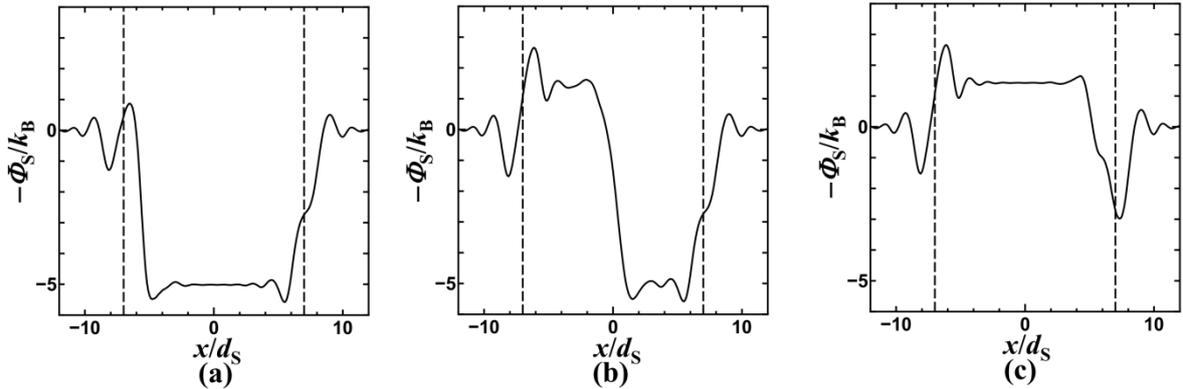


Fig. 2.5. (a) Profile of $-\Phi_S/k_B$ along the x -axis ($y=z=0$) for the entropic potential shown in Fig. 2.4(a). (b) That shown in Fig. 2.4(b). (c) That shown in Fig. 2.4(c). The two broken lines in each plot represent positions of the vessel ends, $x/d_S=\pm 7$.

An important point is that the entropic potential does not remain constant: It is time dependent. There are three different time scales: time scales of the solvent motion, variation in the vessel geometry, and solute motion. The time scale of the solvent motion is doubtlessly the fastest. The solvent is practically in equilibrium with the solute-vessel configuration all the time. Here, we assume that the variation of the vessel geometry (that is, the variation of the entropic potential) is relatively faster than the solute motion. The faster variation of the vessel geometry could be realized by structural changes of AcrB caused by the proton motive force. An entropic force, which is given as $-\{\partial(-T\Phi_S)/\partial x\}$, *continuously* acts on the solute in the right direction along the x -axis and accelerates its motion during the variation of the vessel geometry. When vessel geometry d is reached, the solute possesses a considerably high velocity. As a consequence, the solute is capable of overcoming even a barrier significantly exceeding $k_B T$.

2.3.3. Entropic potentials for vessel geometries (d), (e), and (f) in Fig. 2.1

The profile of $-\Phi_S/k_B$ along the x -axis ($y=z=0$) for vessel geometry d is redrawn for $4 \leq x/d_S \leq 12$ in Fig. 2.6(a). The minimum of $-\Phi_S/k_B$ is located at $(x/d_S, y/d_S, z/d_S)=(7.3, 0, 0)$. The coordinate, $(x/d_S, y/d_S, z/d_S)=(7, 0, 0)$, corresponds to the position where the right half of the solute is outside the vessel. It is observed in Fig. 2.6(a) that the solute must overcome the barrier $\sim 3.6k_B T$ to be released to the bulk in the right direction along the x -axis. It is not definite if the solute velocity is high enough to overcome this barrier. However, we find the following: As the inner diameter at the exit D decreases as (d)→(e)→(f) in Fig. 2.1(d)-(f) (the values of D are $8d_S$, $7.7d_S$, and $7.4d_S$ in vessel geometries d , e , and f , respectively), the solute is driven to move further in the right direction along the x -axis and the barrier becomes progressively lower. The profiles of $-\Phi_S/k_B$ along the x -axis ($y=z=0$) for vessel geometries e and f are shown in Fig. 2.6(b) and (c), respectively. The barriers in vessel geometries e and f are $\sim 2.0k_B T$ and $\sim 1.2k_B T$, respectively. The barrier can further be reduced by decreasing D .

The entropic potential is an effective one, which means that it is subject to fluctuation. When the solute has reached the potential minimum in Fig. 2.6(c), it must overcome the barrier $\sim 1.2k_B T$. However, due to the potential fluctuation, the actual barrier is higher than this value at some times and lower at other times. For this reason, the probability for the solute to overcome such an effective barrier is higher than one might expect from the barrier height. Further, we consider rigid-body interactions. Potential barriers in such a model system are higher than in the real system. If the vessel wall and the solute are soft, the oscillation of the entropic potential should be weaker than that calculated in this Chapter. It is probable that the barrier for the solute to overcome is smaller than in the rigid-body system.

From the above argument, we conclude that the solute can definitely be released to the bulk. The key idea is a continuous variation of the vessel geometry accompanying that of the spatial distribution of the entropic potential. We remark that the solute velocity in the final vessel geometry becomes higher as the vessel length increases. In other words, a longer vessel enables the solute to overcome a higher barrier for being released to the bulk. Even a barrier well exceeding $k_B T$ could be overcome. The very large value of the length of TolC, $\sim 35d_S$, may play essential roles in the solute release.

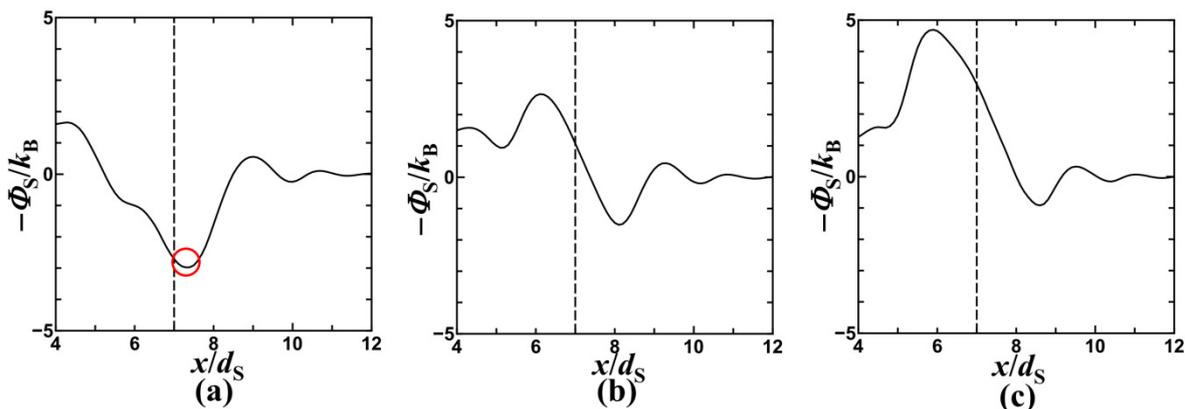


Fig. 2.6. (a) Profile of $-\Phi_S/k_B$ along the x -axis ($y=z=0$) for the vessel geometry shown in Fig. 2.1(d). (b) That for the vessel geometry shown in Fig. 2.1(e). (c) That for the vessel geometry shown in Fig. 2.1(f). The broken line in each plot represents position of the right end of the vessel, $x/d_S=7$.

2.4. Concluding Remarks

We have investigated insertion and release of a solute into and from a cylindrical vessel possessing an entrance at one end and an exit at the other end for the solute. The entropic potential formed by the solvent is calculated using the 3D integral equation theory and rigid-body models. With the rigid-body models, all of the allowed system configurations share the same energy, and the system behavior is purely entropic in origin¹⁵: The insertion/release process is described by the solvent-entropy effect arising from the translational displacement of solvent molecules. It has been demonstrated that the two opposite events, insertion of a solute into the vessel and release of the solute from the vessel, can be driven by the same mechanism. It is quite interesting and important that such rich behavior is observed in a very simple model system.

The results obtained are recapitulated as follows. As long as the vessel geometry is fixed, the entropic component always acts for insertion. We show, however, that a solute which has been inserted can also be released by the solvent-entropy effect using a continuous variation of the vessel geometry. An example of the variation is illustrated in Fig. 2.1. In the real system, the insertion/release process is to be described in terms of the potential of mean force (PMF) consisting of the energetic and entropic components. The energetic component is strongly dependent on the solute-solvent and vessel inner surface-solvent affinities, whereas the entropic component is rather insensitive to them⁶. The entropic component dominates when the inner surface of the vessel is neither hydrophobic nor hydrophilic. The present study, which shows that both of insertion and release can be achieved solely by the entropic component being uninfluenced by the solute-solvent affinity, is an important first step toward elucidating the microscopic mechanism of the multidrug efflux of TolC¹⁻³.

To complete a physical picture of the multidrug efflux, we have to find a manner of the

entropic release (i.e., a reliable manner of varying the geometric features of the vessel) which can handle a fairly wide range of the solute size. To know how high a barrier can be overcome in the final stage of the solute release (see Sec. 2.3.3), we need to know dynamic aspects of the entropic release by simulating a biased Brownian motion of the solute in a time-dependent force field. In any case, a continuous variation of the vessel geometry is the key as pointed out in this Chapter. The details of the polyatomic structures of the vessel and/or the solute are also important factors. Further, the solvent-entropy effect should play crucially important roles even in the functioning of AcrB, which is to be investigated. Works in these directions are in progress.

References

- ¹V. Koronakis, A. Sharff, E. Koronakis, B. Luisi, and C. Hughes, *Nature* **405**, 914 (2000).
- ²L. Vaccaro, K. A. Scott, and M. S. P. Sansom, *Biophys. J.* **95**, 5681 (2008).
- ³K. M. Pos, *Biochim. Biophys. Acta* **1794**, 782 (2009).
- ⁴K. Amano and M. Kinoshita, *Chem. Phys. Lett.* **488**, 1 (2010).
- ⁵K. Amano and M. Kinoshita, *Chem. Phys. Lett.* **504**, 221 (2011).
- ⁶K. Amano and H. Oshima, M. Kinoshita, *J. Chem. Phys.* **135**, 185101 (2011).
- ⁷A. L. Horwich, W. A. Fenton, E. Chapman, and G. W. Farr, *Annu. Rev. Cell. Dev. Biol.* **23**, 115 (2007).
- ⁸T. K. Chaudhuri, V. K. Verma, and A. Maheshwari, *Prog. Biophys. Mol. Biol.* **99**, 42 (2009).
- ⁹S. Tanaka, Y. Kawata, G. Otting, N. E. Dixon, K. Matsuzaki, and M. Hoshino, *Biochim. Biophys. Acta* **1804**, 866 (2010).
- ¹⁰D. Beglov and B. Roux, *J. Chem. Phys.* **103**, 360 (1995).
- ¹¹M. Ikeguchi and J. Doi, *J. Chem. Phys.* **103**, 5011 (1995).
- ¹²M. Kinoshita, *J. Chem. Phys.* **116**, 3493 (2002).
- ¹³M. Kinoshita, *Chem. Phys. Lett.* **387**, 47 (2004).
- ¹⁴K. Amano, T. Yoshidome, M. Iwaki, M. Suzuki, and M. Kinoshita, *J. Chem. Phys.* **133**, 045103 (2010).
- ¹⁵M. Kinoshita, *Chem. Eng. Sci.* **61**, 2150 (2006).
- ¹⁶T. Yoshidome, M. Kinoshita, S. Hirota, N. Baden, and M. Terazima, *J. Chem. Phys.* **128**, 225104 (2008).
- ¹⁷T. Yoshidome, and M. Kinoshita, *Phys. Rev. E* **79**, 090305(R) (2009).
- ¹⁸H. Oshima, T. Yoshidome, K. Amano, and M. Kinoshita, *J. Chem. Phys.* **131**, 205102 (2009).
- ¹⁹Y. Harano and M. Kinoshita, *Biophys. J.* **89**, 2701 (2005).
- ²⁰M. Kinoshita, *Front. Biosci.* **14**, 3419 (2009).
- ²¹M. Kinoshita, *Int. J. Mol. Sci.* **10**, 1064 (2009).
- ²²S. Asakura and F. Oosawa, *J. Chem. Phys.* **22**, 1255 (1954).
- ²³S. Asakura and F. Oosawa, *J. Polym. Sci.* **33**, 183 (1958).
- ²⁴P. Attard and G. N. Patey, *J. Chem. Phys.* **92**, 4970 (1990).
- ²⁵M. Kinoshita, S. Iba, K. Kuwamoto, and M. Harada, *J. Chem. Phys.* **105**, 7177 (1996).
- ²⁶R. Roth, B. Götzelmann, and S. Dietrich, *Phys. Rev. Lett.* **83**, 448 (1999).
- ²⁷M. Kinoshita, *Chem. Phys. Lett.* **353**, 259 (2002).
- ²⁸R. Roth and M. Kinoshita, *J. Chem. Phys.* **125**, 084910 (2006).
- ²⁹P.-M. König, R. Roth, and S. Dietrich, *Phys. Rev. E* **74**, 041404 (2006).
- ³⁰P.-M. König, R. Roth, and S. Dietrich, *Europhys. Lett.* **84**, 68006 (2008).

Chapter 3. On the physics of multidrug efflux through a biomolecular complex

3.1. Introduction

Drug extrusion via efflux through a tripartite complex constructing the Resistance-Nodulation-Cell Division (RND) system is a widely used mechanism in Gram-negative bacteria. The complex comprises AcrA, AcrB, and TolC as illustrated in Fig. 3.1.¹⁻⁸ AcrB, which is referred to as an efflux pump protein, comprises three protomers. It has been suggested that each protomer undergoes a sequential structural change among three states in which a drug insertion, binding, and release take place, respectively.^{5,6,8} The outer membrane protein, TolC, is a cylindrical vessel possessing an entrance at one end and an exit at the other end for the solute. The periplasmic linker protein, AcrA, mediates the contact between AcrB and TolC.^{5,6,8} AcrB interacting with TolC in this way sends the solute to the central position within the vessel cavity of TolC at the entrance, and then the solute is moved to the exit. Thus, the transporter enables a passage of the drug from the periplasm to the external medium. A feature of the transporter is that it is capable of handling drugs with diverse properties (i.e., solvophobic and solvophilic solutes with a wide range of sizes).⁶⁻⁸ This feature, which is known as “multidrug efflux”, is in marked contrast with the high selectivity in the receptor-ligand binding. In our earlier works, it was shown that the high selectivity is ascribed to geometric characteristics of the receptor pocket and ligand, which are made substantial by the entropic effect originating from the translational displacement of solvent molecules.⁹⁻¹¹ Here we show that the solvent-entropy effect plays crucially important roles in the multidrug efflux as well.

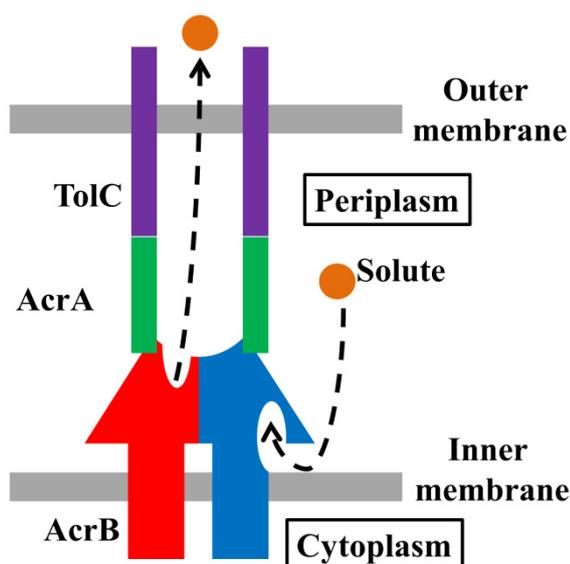


Fig. 3.1. Cartoon illustrating an efflux transporter, AcrA/AcrB/TolC.

A prevailing view is the following: The multidrug efflux stems from multifunctional ligand-binding sites^{4,5,12} of AcrB which recognize various types of functional groups; once the solute enters TolC, it goes to the external medium merely by “diffusion”¹² in the presence of no particular potential field; and only AcrB is responsible for the multidrug efflux (in general, AcrB has caught more attention than TolC for analyses on the solute motion^{12,13}). However, there is an important point to be noticed: Once AcrB sends the solute to the central position within the vessel cavity of TolC (the length of TolC reaches ~14 nm that is far larger than the solute size) at the entrance, the solute must be ejected to the external medium through the exit before AcrB sends the next solute. The ejection of every solute needs to be finished with sufficient rapidness, but it is uncertain whether such ejection is achievable by the mere diffusion mentioned above or not. The achievement may be assured if the solute motion is directed only toward the exit (i.e., the solute motion is unidirectional). Also, we emphasize that the multidrug efflux of AcrA/AcrB/TolC can be realized only when TolC as well as AcrB exhibits this feature. TolC possesses an uncomplicated structure of a cylindrical shape with an open space inside it and works as a tunnel for drugs. The existence of the multifunctional ligand-binding sites can hardly be applied to TolC, implying that the multidrug efflux is implemented through a different mechanism. Further, it is not definite that such sites actually exist in AcrB. Even if they exist, a solute once bound to one of the sites is not likely to be released from it since the solute properties remain unchanged. The binding must be strong enough for the solute to remain captured, whereas it must be sufficiently weak for the solute dissociation in the next step. It is not straightforward to meet both of these requirements. In the present study, we revisit the mechanism of the multidrug efflux by considering TolC, from the standpoint that the mechanism cannot be chemically specific but rather has to be based on a physical factor. It is demonstrated that the multidrug efflux can be realized without assuming the existence of the multifunctional ligand-binding sites.

A principal concern is insertion of a solute into a vessel comprising biopolymers followed by release of the same solute from it. It is quite interesting that the two apparently *opposite* events, insertion and release, successively occur in a system. In earlier works,¹⁴⁻¹⁷ we showed that the spatial distribution of the solute-vessel potential of mean force (PMF denoted by Φ) formed by the solvent plays imperative roles for the insertion/release process. The PMF represents “the free energy of the solvent for a fixed configuration of the solute-vessel pair” minus “that for the configuration where the solute is infinitely far from the vessel”. The spatial distribution of the PMF becomes largely positive and largely negative with the periodicity of the molecular diameter of the solvent, $d_S=0.28$ nm. The motion of a solute immersed in a solvent confined on the scale of a nanometer is influenced by the distribution.

It is physically insightful to decompose the PMF scaled by $k_B T$ (k_B is Boltzmann’s constant and T is the absolute temperature), $\Phi/(k_B T)$, into its energetic and entropic components denoted by $\Phi_E/(k_B T)$ and $-\Phi_S/k_B$, respectively: $\Phi/(k_B T)=\Phi_E/(k_B T)-\Phi_S/k_B$. In what follows, the conclusions drawn from our earlier works^{15,16} are recapitulated. $\Phi_E/(k_B T)$ is strongly dependent on the solute-solvent and vessel inner surface-solvent affinities. When the inner surface of the vessel is solvophilic, the average solvent number density within the

vessel cavity is higher than that in the bulk solvent. As a result, the solvent environment becomes more favorable within the cavity than in the bulk for a solvophilic solute, while the opposite is true for a solvophobic solute: The solvophilic solute is preferentially solvated within a small space almost in the center of the cavity but the solvophobic solute is more stabilized in the bulk. A solvophobic inner surface gives rise to a lower solvent number density on an average within the cavity, leading to preferential solvation of a solvophilic solute in the bulk and more stabilization of a solvophobic solute within the cavity. $-\Phi_S/k_B$, which originates from the translational displacement of solvent molecules, is rather insensitive to the solute-solvent and vessel inner surface-solvent affinities, namely, to whether the solute or the vessel inner surface is solvophobic or solvophilic (on condition that they are neither too solvophobic nor too solvophilic). It is closely related to the excluded volume (EV) generated by the solute. (The solute generates a space which the centers of solvent molecules cannot enter, and the volume of this space is the EV.) $-\Phi_S/k_B$ always drives the solute to be inserted into the vessel cavity and constrained within a small space almost in the center. The power of insertion and constraint becomes stronger as the EV increases. The release can be performed only by $\Phi_E/(k_B T)$.

In the protein flux through a chaperonin system,^{15,16,18-20} where an unfolded protein is inserted into the chaperonin from the bulk solvent and the folded protein is released back to the bulk, the vessel properties (i.e., geometry and vessel inner surface-solvent affinity) are the same for insertion and release. The inner surface is weakly solvophobic. However, the solute properties are different for the two events. Since the unfolded protein possesses large EV and weak solvophobicity, it is driven to be strongly inserted by $-\Phi_S/k_B$ and weakly inserted by $\Phi_E/(k_B T)$, and the net action is insertion. The folded protein is characterized by small EV and solvophilicity with the result that the power of insertion by $-\Phi_S/k_B$ becomes weaker and $\Phi_E/(k_B T)$ dominates. The folded protein is released back to the bulk solvent for preferential solvation. The switch from insertion to release is thus realized. An important point is that $\Phi_E/(k_B T)$ is requisite in release. In TolC, by contrast, the solute properties remain unchanged for insertion and release. To accomplish the release, the vessel inner surface needs to be solvophobic for a solvophilic solute, whereas it needs to be solvophilic for a solvophobic solute. It turns out that releasing both of the solvophilic and solvophobic solutes is not achievable.

Our basic idea is as follows: Only the entropic component of the PMF possesses the insensitivity to the solute-solvent and vessel inner surface-solvent affinities; and the multidrug efflux can be realized if the entropic component predominates over the energetic one. When the vessel inner surface is neither solvophobic nor solvophilic, the solvents within the vessel cavity and in the bulk offer almost the same environment to any solute with solvophobicity or solvophilicity, and the energetic component becomes much less powerful than the entropic component (i.e., the latter dominates). This can be relevant to TolC whose inner surface possesses a mixture of predominantly nonpolar and isolated electronegative patches.¹ The remaining issue is to explore how to release a solute *entropically*.

In the present study, we calculate the spatial distribution of the PMF between a big sphere and a cylindrical vessel with two open ends immersed in small spheres using the three-dimensional (3D) integral equation theory^{9,10,14-17,21-26} combined with rigid-body models

in which the constituents interact only through hard-body potentials. The big sphere and small spheres correspond to a solute and solvent, respectively, and the vessel is a model of TolC. To focus our analysis on the entropic component of the PMF, we employ the rigid-body models mentioned above where all of the allowed system configurations share the same energy and the system behavior is purely entropic in origin. Namely, we look at $\Phi/(k_B T)$ possessing only the entropic component: $\Phi/(k_B T) = -\Phi_S/k_B$. Unlike in our earlier works,¹⁴⁻¹⁶ the vessel geometry is made variable after the solute insertion. (For a fixed geometry of the vessel, the solute remains confined within the vessel cavity once it is inserted.) It is demonstrated that release can also be realized by the entropic component of the PMF: The solute is entropically moved from the entrance to the exit by a continuous variation of the vessel geometry. Due to the PMF periodicity of $d_S=0.28$ nm explained above, even a vessel-geometry variation of this scale leads to a drastic change in the spatial distribution of the PMF, thus enabling the vessel to control the solute motion. The key factor is a time-dependent entropic force acting on a solute, which originates from the solvent-mediated interaction between the solute and the vessel whose geometry is continuously varied and time dependent. Solutes with a wide range of sizes are entropically released using the same vessel-geometry variation. Such rich behavior is observed even in our simplified model neglecting the potentials other than hard-body potentials and polyatomic structures of the vessel and the solute.

A brief report has already been published by us in a Letter,¹⁷ but much more extensive parametric studies are carried out with detailed arguments in the present article. In particular, we show for the first time that solutes with a wide range of sizes can be handled, thus developing a sounder physical basis of the multidrug efflux. The conclusions drawn are fairly general and also applicable to other proteins (or protein complexes) exhibiting the multidrug efflux such as AcrB and ATP-binding cassette (ABC) transporter.^{27,28}

3.2. Model and Theory

3.2.1. Three-dimensional integral equation theory applied to rigid-body models

In biological systems, the solvent is water characterized by hydrogen bonds. Solute insertion into water, for instance, causes restrictions of translational and rotational freedoms of water molecules. However, the contribution from the translational restriction is much larger: In hydration thermodynamics of a solute, the translational entropy predominates over the rotational entropy.^{29,30} In many cases, the translational-entropy effect can be described by modeling water as hard spheres as long as the diameter and number density of the hard-sphere solvent are set at those of water.^{31,32} (An exception is found in the elucidation of cold denaturation of a protein³³⁻³⁵ where the weakening of the entropic effect at low temperatures plays essential roles and a suitable molecular model is necessitated for water.) We note that the hydrogen bonding allows water to exist as a dense liquid despite its quite a small molecular size, leading to an exceptionally large entropic effect.

We consider rigid-body models in which the constituents interact only through hard-body potentials: a big hard vessel and a big hard sphere immersed at infinite dilution in

small hard spheres with diameter d_s forming the solvent. The big sphere corresponds to a solute and its diameter is denoted by d_B . The initial geometry of the vessel is illustrated in Fig. 3.2(a). After the solute insertion, the vessel geometry is varied for the solute release as explained in Figs. 3.2(b)-(f). The Cartesian coordinate system is chosen as illustrated in Fig. 3.2(a). The cross section of $z=0$ is shown for each geometry. More details are described in 3.2.2.

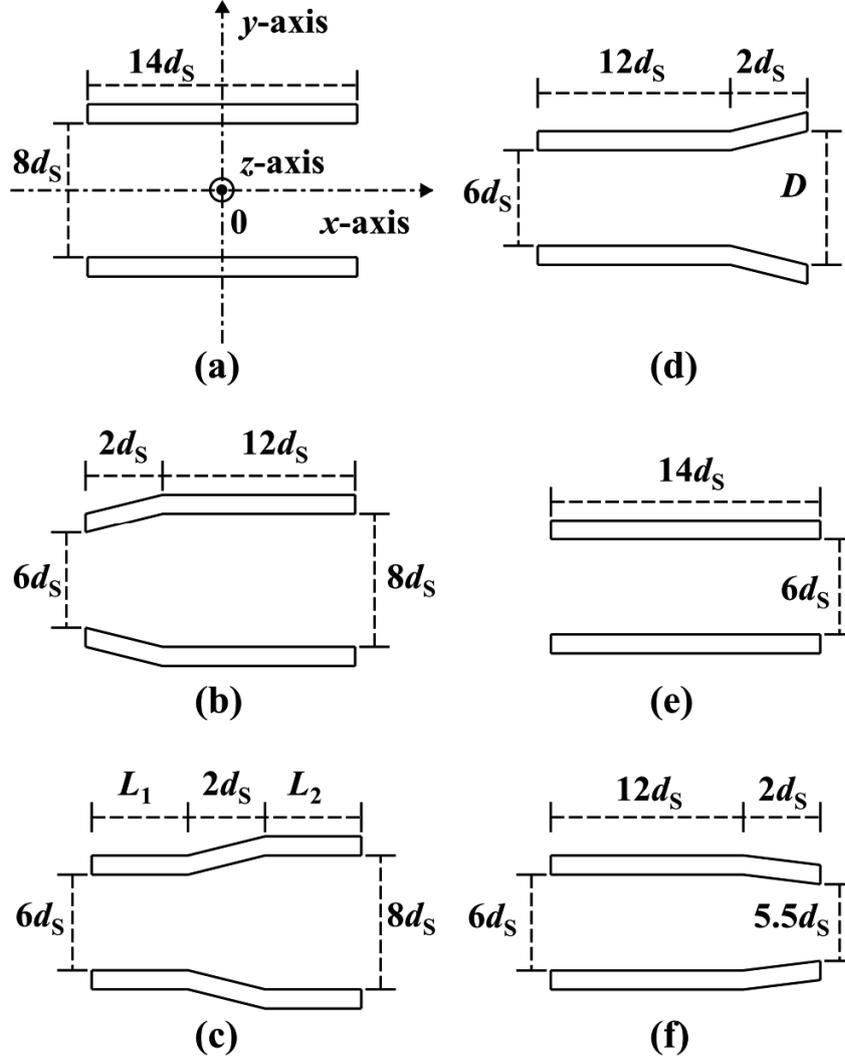


Fig. 3.2. Variation of vessel geometry: (a)→(b)→(c)→(d)→(e)→(f). The geometry in (a) is a cylinder with inner diameter $8d_s$, length $14d_s$, and thickness d_s (d_s is the molecular diameter of the solvent). The coordinate system is chosen as illustrated here. In the geometry in (d), $D=8d_s$. D is reduced as (d)→(e)→(f). The cross section of $z=0$ is shown for each geometry. Though the geometry variation is illustrated in a stepwise manner, it is made *continuously*.

The vessel with a prescribed geometry is considered. First, the solvent-solvent correlation functions are calculated using the radial-symmetric (RS) hypernetted-chain (HNC) theory for spherical particles.³⁶ Second, the vessel-solvent correlation functions are calculated by means of the 3D integral equation theory described below. This

theory^{9,10,14-17,21-26} and the density functional theory³⁷⁻⁴⁰ have been applied to a variety of important problems in biophysics and colloidal science with the emphasis on the entropic force^{41,42} induced between big, spherical and nonspherical bodies immersed in small spheres.^{9,10,14-17,23,25,26,37-40} Third, the solute-solvent correlation functions are calculated using the RS-HNC theory for spherical particles.³⁶ The PMF between the vessel and the solute is then calculated by assuming that the solvent particles are always in equilibrium with each configuration of the vessel-solute pair.

The Ornstein-Zernike (OZ) equation in the Fourier space is expressed by^{9,10,14-17,21-26}

$$W_{1S}(k_x, k_y, k_z) = \rho_S C_{1S}(k_x, k_y, k_z) H_{SS}(k). \quad (3.1)$$

This is coupled with the HNC closure equation written as^{9,10,14-17,21-26}

$$c_{1S}(x, y, z) = \exp\{-u_{1S}(x, y, z)/(k_B T)\} \exp\{w_{1S}(x, y, z)\} - w_{1S}(x, y, z) - 1. \quad (3.2)$$

In these equations, the subscripts “1” and “S” denote the vessel and the solvent, respectively, $w = h - c$, c is the direct correlation function, h the total correlation function, u the potential, and ρ_S the bulk density. The molecular diameter d_S is set at the value of water, 0.28 nm. The reduced number density $\rho_S d_S^3$ is taken to be the value of water at 298 K and 1 atm, 0.7317. C , H , and W represent the Fourier transforms of c , h , and w , respectively. $H_{SS}(k)$ ($k^2 = k_x^2 + k_y^2 + k_z^2$) calculated using the RS-HNC theory for spherical particles³⁶ is part of the input data. We emphasize that the OZ equation is *exact*.⁴³ It has been stated that the OZ equation includes serious approximations in a publication,³⁸ but this statement is incorrect. On the other hand, the bridge function is neglected in the HNC closure equation. However, it has been corroborated that the 3D-OZ-HNC theory gives quantitatively reliable results.¹⁰

The numerical procedure is briefly summarized as follows: (1) $u_{1S}(x, y, z)$ is calculated at each 3D grid point, (2) $w_{1S}(x, y, z)$ is initialized to zero, (3) $c_{1S}(x, y, z)$ is calculated from Eq. (3.2), and $c_{1S}(x, y, z)$ is transformed to $C_{1S}(k_x, k_y, k_z)$ using the 3D fast Fourier transform (3D-FFT), (4) $W_{1S}(k_x, k_y, k_z)$ is calculated from Eq. (3.1), and $W_{1S}(k_x, k_y, k_z)$ is inverted to $w_{1S}(x, y, z)$ using the 3D-FFT, and (5) steps (3) and (4) are repeated until the input and output functions for $w_{1S}(x, y, z)$ become identical within convergence tolerance. On grid points where a solvent particle and the solute overlap, $\exp\{-u_{1S}(x, y, z)/(k_B T)\}$ is zero. On those where a solvent particle is in contact with the solute, it is set at 0.5, and otherwise it is unity. The grid spacing (Δx , Δy , and Δz) is set at $0.1d_S$, and the grid resolution ($N_x \times N_y \times N_z$) is $256 \times 256 \times 256$. It has been confirmed that the spacing is sufficiently small and the box size ($N_x \Delta x$, $N_y \Delta y$, $N_z \Delta z$) is large enough for the correlation functions at the box surfaces to be essentially zero.

The Fourier transform of the vessel-solvent direct correlation function denoted by $C_{1S}(k_x, k_y, k_z)$ is calculated in accordance with the procedure described above. The Fourier transforms of the solute-solvent total correlation function denoted by $H_{2S}(k)$ (the subscript “2” denotes the solute) is calculated using the radial symmetric HNC theory for spherical particles where $H_{SS}(k)$ is part of the input data. The PMF between the vessel and the solute, $\Phi_{12}(x, y, z)$, is then obtained from

$$\Phi_{12}(x, y, z)/(k_B T) = u_{12}(x, y, z)/(k_B T) - w_{12}(x, y, z), \quad (3.3)$$

where $w_{12}(x, y, z)$ is calculated by inverting $W_{12}(k_x, k_y, k_z)$ given by

$$W_{12}(k_x, k_y, k_z) = \rho_S C_{1S}(k_x, k_y, k_z) H_{2S}(k). \quad (3.4)$$

The two equations,^{16,26}

$$\Phi_{12}(x, y, z) = F(x, y, z) - F_\infty \quad (3.5)$$

and

$$g_{12}(x, y, z) = \exp\{-\Phi_{12}(x, y, z)/(k_B T)\}, \quad g_{12,\infty} = 1, \quad (3.6)$$

allow us to understand the physical meaning of $\Phi_{12}(x, y, z)$. Here, $F(x, y, z)$ is the free energy of small spheres in the case where the solute center is at the position (x, y, z) , and $g_{12}(x, y, z)$ is the pair distribution function. The subscript “ ∞ ” denotes the value in the case where the solute is infinitely separated from the vessel. For rigid-body models, the behavior of $\Phi_{12}(x, y, z)$ is purely entropic in origin. Due to the microscopic structure of small spheres formed within the domain confined by the solute and the vessel, $\Phi_{12}(x, y, z)$ exhibits a complex spatial distribution. A great advantage of the 3D integral equation theory is that the values of Φ_{12} on all the grid points are obtained from only a single calculation, which is not inherent in the usual computer simulation (e.g., a molecular dynamics (MD) simulation).

3.2.2. Variation of vessel geometry

We emphasize that the vessel-geometry variation is made *continuously* though it is represented in a stepwise manner in Fig. 3.2. Fig. 3.2(a) shows the initial geometry of the vessel, a cylinder with inner diameter $8d_S$, length $14d_S$, and thickness d_S . These dimensions roughly mimic those of TolC except that the length is set at a much shorter value (the length of TolC in the real system is $\sim 50d_S$).⁴⁴ This is because all we need in the present study is a length which is sufficiently longer than the diameter of the solute d_B set at $2d_S - 5d_S$. The left and right ends of the cylindrical vessel are the entrance and the exit, respectively. The inner diameter at the entrance is first reduced to $6d_S$ (Fig. 3.2(b)). L_1 and L_2 , lengths of the portions with inner diameter $6d_S$ and $8d_S$, respectively, are gradually increased and decreased, respectively (Fig. 3.2(c)). Since the length of the tapering portion is set at $2d_S$, $L_1 + L_2 = 12d_S$. The inner diameter at the exit D is gradually reduced after the geometry with $D = 8d_S$ shown in Fig. 3.2(d) is reached. In the case of $D = 6d_S$, the inner diameter uniformly becomes $6d_S$ (Fig. 3.2(e)). As the final geometry, D is made smaller than $6d_S$ ($D = 5.5d_S$ in Fig. 3.2(f)). Hereafter, the vessel geometries shown in Figs. 3.2(a), (b), (c), (d), (e), and (f) are referred to as vessel geometries (a), (b), (c), (d), (e), and (f), respectively.

In AcrA/AcrB/TolC, the proton motive force causes structural changes of AcrB, which are transmitted to TolC through AcrA.^{5,6,8} This action may lead to a continuous

vessel-geometry variation of TolC. On the other hand, a recent MD simulation study⁴⁴ has suggested that TolC can vary its geometric characteristics by itself: They have observed a peristaltic motion of the periplasmic domain and closing and opening for both of the periplasmic and extracellular mouths. This observation is consistent with the geometry variation illustrated in Fig. 3.2. We start with the geometry of Fig. 3.2(a) where the left end corresponding to the periplasmic mouth is opened. In the geometry of Fig. 3.2 (f), the right end corresponding to the extracellular mouth begins to change its geometry toward closing. Of course, details of the geometry variation in the real system during the insertion/release process are not known. However, our conclusions are independent of the details as discussed in Sec. 3.5. What we emphasize in the present study is that release as well as insertion of solutes with a wide range of sizes can be achieved entropically and that a continuous vessel-geometry variation plays essential roles.

3.3. Results and Discussion

3.3.1. Entropic potential for vessel geometry (a) in Fig. 3.2

Hereafter, the PMF scaled by $k_B T$ between the vessel and the solute are denoted simply by $-\Phi_S/k_B$ because the equation, $\Phi/(k_B T) = -\Phi_S/k_B$, holds for our model system. We refer to $-\Phi_S$ as “entropic potential”. Unless otherwise mentioned, the solute size d_B is set at md_S ($m=2, 3, 4, 5$).

Fig. 3.3 shows the distributions of $-\Phi_S/k_B$ on the cross section of $z=0$ for vessel geometry (a) (see Fig. 3.2(a)) for the four values of d_B . As the color approaches dark blue, $-\Phi_S/k_B$ and the free energy of the solvent become lower and the solute is more stabilized there. As the color approaches dark red, they become higher and the solute is more destabilized there. The center of the big sphere cannot enter the domain drawn in white. The stripe pattern of the entropic potential formed along the y -axis can physically be interpreted as follows. When the separation between the nearest solute and vessel inner surfaces, which is denoted by η , is not sufficiently close to nd_S ($n=0, 1, 2, \dots$), spaces unavailable to the translational displacement of solvent molecules appear as indicated in Fig. 3.4(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. 3.4(b). The configuration in Fig. 3.4(a) is entropically unfavorable, while that in Fig. 3.4(b) is entropically favorable, leading to the stripe pattern formed along the y -axis. The inner diameter of vessel geometry (a) is $8d_S$. The entropic potential in the central region around $y=0$ is negative when $(8d_S - d_B)/d_S$ is even (i.e., $d_B=2d_S$ or $4d_S$), whereas it is positive when $(8d_S - d_B)/d_S$ is odd (i.e., $d_B=3d_S$ or $5d_S$), as observed in Fig. 3.3. The amplitude of the stripe pattern formed along the y -axis becomes progressively larger as the solute size increases.

It was experimentally shown for the entropic interaction between large spheres immersed in small spheres that a free-energy barrier well exceeding $k_B T$ cannot readily be overcome and that with the barrier of $\sim 2k_B T$ the large spheres come together only rarely.⁴⁵ It is true, however, that even a significantly high barrier will eventually be overcome. To estimate the time required for overcoming the barrier, we have to analyze the dynamic

behavior of the solute by solving the Fokker-Planck equation. Nevertheless, pending such a detailed analysis, we discuss barriers in reference to the experimental observation mentioned above. (As argued below, the time required for overcoming the barrier in the final stage of the solute release at the exit can be an important quantity: It is better to make this time as short as possible by keeping the barrier as low as possible. See the fourth and fifth paragraphs in “Conclusions” for a more detailed discussion.)

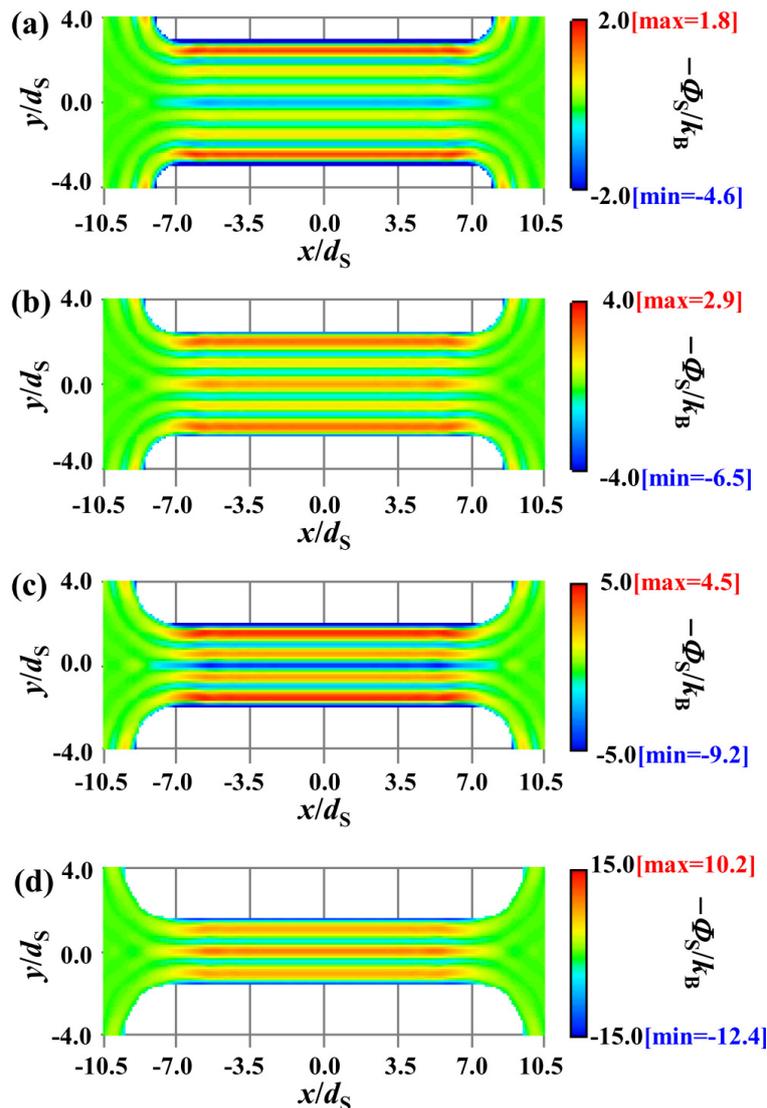


Fig. 3.3. Distributions of $-\Phi_S/k_B$ on the cross section of $z=0$ for the initial vessel geometry illustrated in Fig. 3.2(a) for the solutes with $d_B=2d_S$ (a), $d_B=3d_S$ (b), $d_B=4d_S$ (c), and $d_B=5d_S$ (d). $-\Phi_S/k_B$ becomes lower as the color approaches dark blue, and it becomes higher as the color approaches dark red (“max” and “min” represent the maximum and minimum values, respectively). The center of the big sphere cannot enter the domain drawn in white.

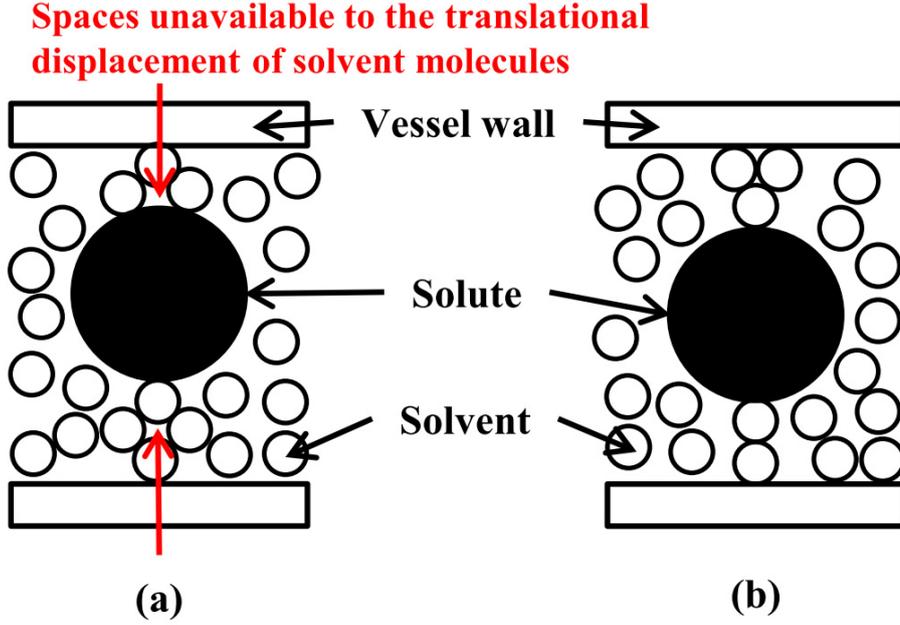


Fig. 3.4. Cartoons illustrating how the stripe pattern of the entropic potential shown in Fig. 3.2(a) is formed along the y -axis. (a) Separation between the nearest solute and vessel inner surfaces, which is denoted by η , is not sufficiently close to nd_S ($n=0, 1, 2, \dots$). (b) In the case of $\eta \sim nd_S$.

As observed in Fig. 3.3, the solute is most likely to be inserted into the central region around $y=0$ for $d_B=2d_S$ or $4d_S$ and into the region next to the central one around $y/d_S=\pm 0.5$ for $d_B=3d_S$ or $5d_S$ (these regions are referred to as the most stable regions). However, for all the solute sizes tested there is a possibility that the solute is inserted into a different region with negative $-\Phi_S/k_B$ due to the presence of a trajectory possessing a sufficiently low barrier. As shown in our earlier works,¹⁴⁻¹⁶ a modification of the vessel geometry (e.g., truncation of the edge at the entrance) increases the barriers in a trajectory in the region $x/d_S < -7$ along which the solute can touch the vessel and may enable the solute to be inserted into the most stable region with higher probability. In any case, we postulate that the solute is inserted into the most stable region for each solute size. This postulation is justified because in the real system AcrB interacting with TolC at its entrance sends the solute to the central position within the vessel cavity.

The profiles of $-\Phi_S/k_B$ along “ $y=0$ and $z=0$ ” (i.e., the x -axis) are shown in Fig. 3.5(a) for $d_B=2d_S$ and $4d_S$. Those along “ $y/d_S=\pm 0.5$ and $z=0$ ” are presented in Fig. 3.5(b) for $d_B=3d_S$ and $5d_S$. The dotted line indicates the position near the entrance where $-\Phi_S/k_B$ takes the lowest value. The position is located at $x/d_S=-5.5$ in Fig. 3.5(a) and at $x/d_S=-5.7$ in Fig. 3.5(b). Fig. 3.6 shows the profiles of $-\Phi_S/k_B$ in the radial direction along “ $x/d_S=-5.5$ and $z=0$ ” for $d_B=2d_S$ and $4d_S$ together with those along “ $x/d_S=-5.7$ and $z=0$ ” for $d_B=3d_S$ and $5d_S$. The open circle indicates the position where the solute is stabilized with the highest probability right after the solute insertion. If the vessel geometry underwent no change, the solutes with $d_B=4d_S$ and $5d_S$ would remain confined within the region of $-6 < x/d_S < 6$ (see Fig. 3.5). A problem is that the solutes with $d_B=2d_S$ and $3d_S$ may go in the left direction by overcoming the barriers of $\sim k_B T$.

Moreover, the solute with $d_B=2d_S$ may move in the radial direction without difficulty (see Fig. 3.6). These unfavorable motions of the smaller solutes can be prevented by the variation in the vessel geometry which is made as soon as the solute is inserted into the most stable region for each solute size. More details are discussed in 3.3.2.

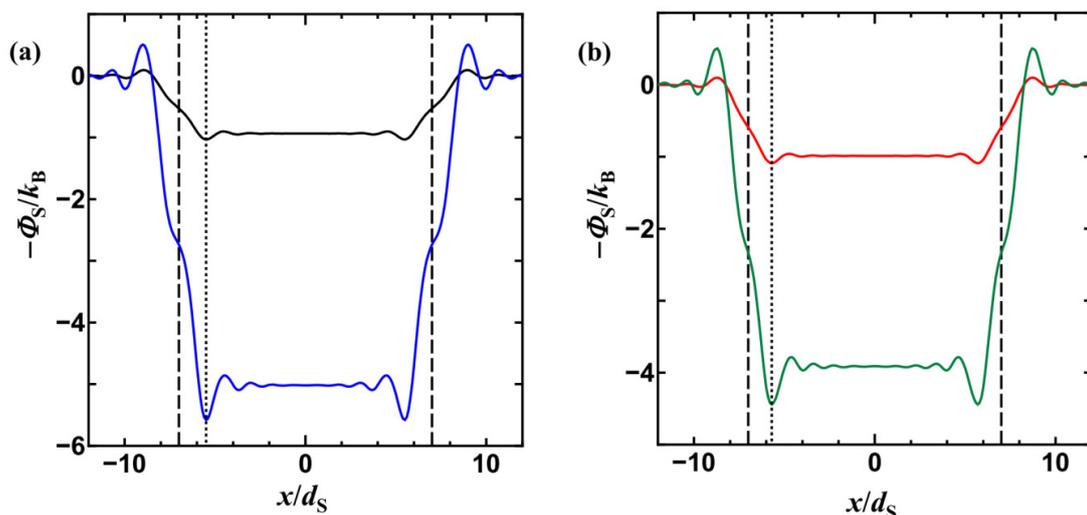


Fig. 3.5. (a) Profiles of $-\Phi_S/k_B$ along “ $y=0$ and $z=0$ ” for $d_B=2d_S$ (black) and $4d_S$ (blue). (b) Those along “ $y/d_S=\pm 0.5$ and $z=0$ ” for $d_B=3d_S$ (red) and $5d_S$ (green). The two broken lines in each plot represent positions of the vessel ends, $x/d_S=\pm 7$. The dotted line indicates the position near the entrance where $-\Phi_S/k_B$ takes the lowest value.

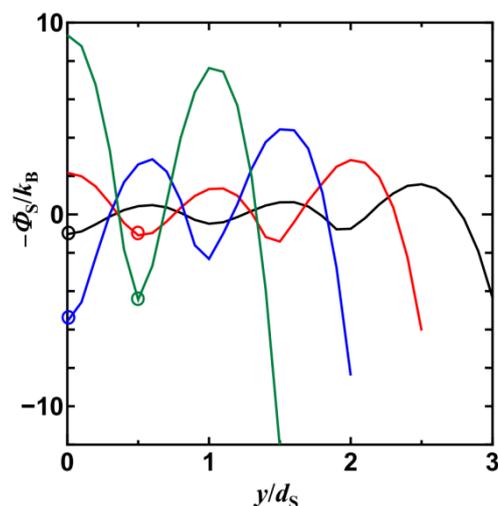


Fig. 3.6. Profiles of $-\Phi_S/k_B$ in the radial direction along “ $x/d_S=-5.5$ and $z=0$ ” for $d_B=2d_S$ (black) and $4d_S$ (blue), and those along “ $x/d_S=-5.7$ and $z=0$ ” for $d_B=3d_S$ (red) and $5d_S$ (green). The open circle indicates the position where the solute is stabilized with the highest probability right after the solute insertion.

3.3.2. Entropic potentials for vessel geometries (b), (c), and (d) in Fig. 3.2

For the solute with $d_B=2d_S$, the distributions of $-\Phi_S/k_B$ on the cross section of $z=0$ for vessel geometries (b), (c), and (d) (see Figs. 3.2(b)-(d)) are shown in Figs. 3.7(a), (b), and (c), respectively. Those for the solutes with $d_B=3d_S$, $4d_S$, and $5d_S$ are presented in Figs. 3.8, 3.9, and 3.10, respectively. In vessel geometry (c), $L_1=L_2=6d_S$. The open circle indicates the position where the entropic potential is locally minimum. For the portion of length L_1 as well as for that of length L_2 (see Fig. 3.2(c)), the entropic potential in the central region around $y=0$ is negative when $(8d_S-d_B)/d_S$ is even (i.e., $d_B=2d_S$ or $4d_S$), whereas it is positive when $(8d_S-d_B)/d_S$ is odd (i.e., $d_B=3d_S$ or $5d_S$). Within the intermediate portion (i.e., portion of length $2d_S$) around the x -axis, a small domain with positive $-\Phi_S/k_B$ is formed for $d_B=2d_S$ or $4d_S$, while that with negative $-\Phi_S/k_B$ is formed for $d_B=3d_S$ or $5d_S$. Next to the solute on the left side, the potential becomes positive regardless of the solute size. Since L_1 is continuously lengthened as illustrated in Fig. 3.2, the solutes with all the sizes tested are driven to move in the right direction along the x -axis toward the exit (see Figs. 3.7–3.10).

A notable point is that the entropic potential does not remain constant: It is time dependent. There are three different time scales: those of the solvent motion, variation in the vessel geometry, and solute motion. The time scale of the solvent motion is doubtlessly the fastest. The solvent is practically in equilibrium with the solute-vessel configuration all the time. Here, we assume that the variation of the vessel geometry (that is, the variation of the entropic potential) is relatively faster than the solute motion. The faster variation of the vessel geometry could be realized by structural changes of AcrB caused by the proton motive force. It follows that an entropic force, which is given as $F_S = -\{\partial(-T\Phi_S)/\partial x\}$, *continuously* acts on the solute in the right direction along the x -axis and continues to accelerate its motion during the variation of the vessel geometry. When vessel geometry (d) is reached, the solute should possess a considerably high velocity. As a consequence, even a barrier well exceeding $k_B T$ can readily be overcome by the solute.

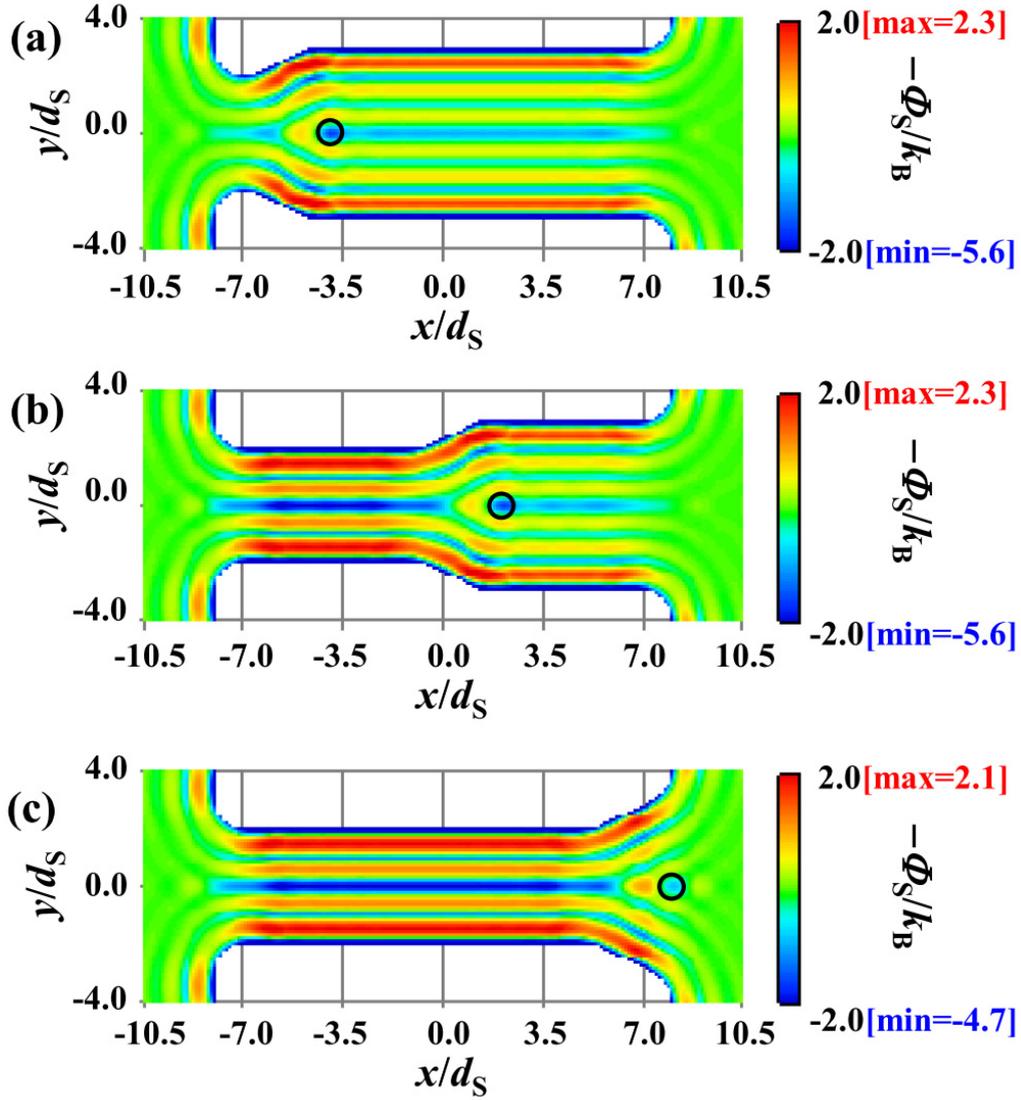


Fig. 3.7. Distributions of $-\Phi_S/k_B$ on the cross section of $z=0$ for vessel geometries (b), (c), and (d) illustrated in Figs. 3.2(b), (c) and (d), respectively. The solute size d_B is $2d_S$. $-\Phi_S/k_B$ becomes lower as the color approaches dark blue, and it becomes higher as the color approaches dark red (“max” and “min” represent the maximum and minimum values, respectively). The open circle indicates the position where the entropic potential is locally minimum.

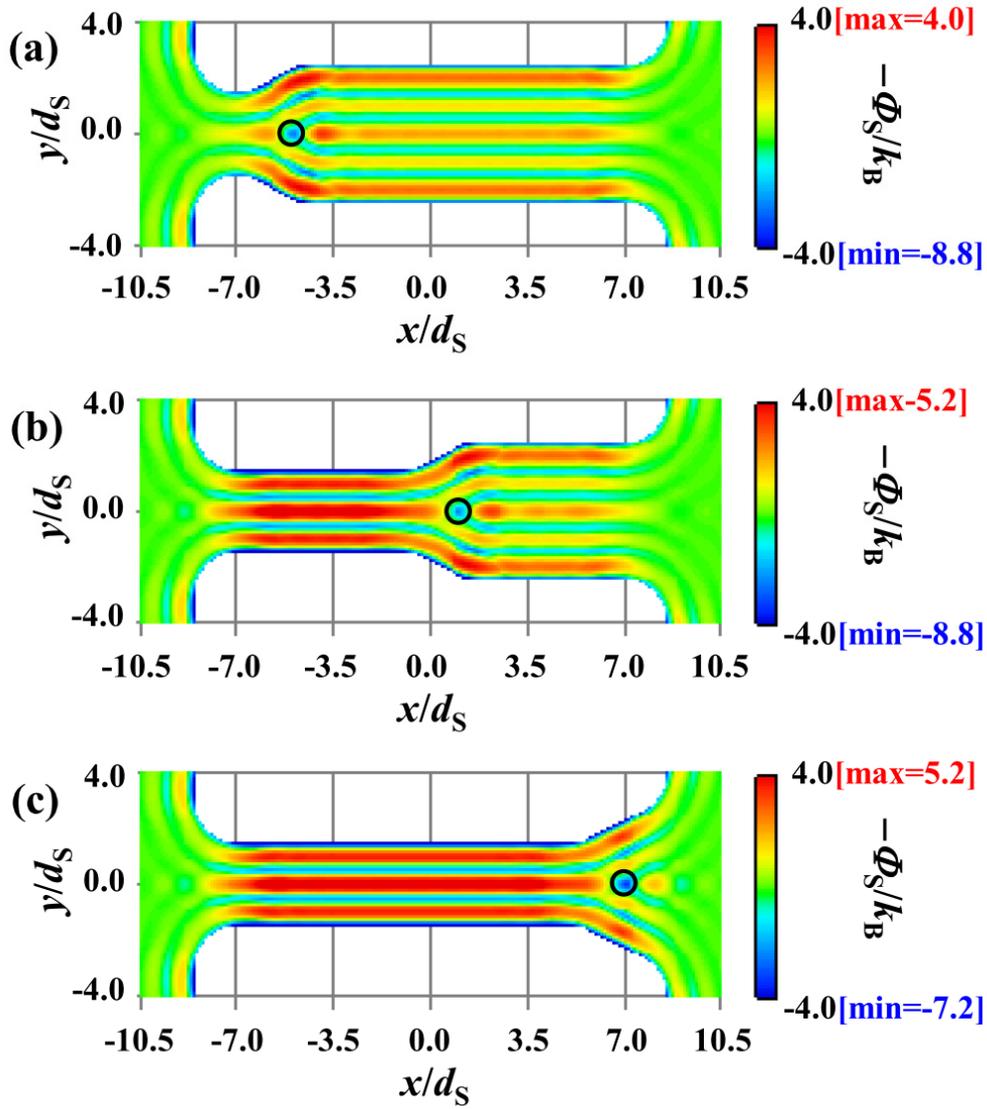


Fig. 3.8. Distributions of $-\Phi_S/k_B$ on the cross section of $z=0$ for vessel geometries (b), (c), and (d) illustrated in Figs. 3.2(b), (c) and (d), respectively. The solute size d_B is $3d_S$. $-\Phi_S/k_B$ becomes lower as the color approaches dark blue, and it becomes higher as the color approaches dark red (“max” and “min” represent the maximum and minimum values, respectively). The open circle indicates the position where the entropic potential is locally minimum.

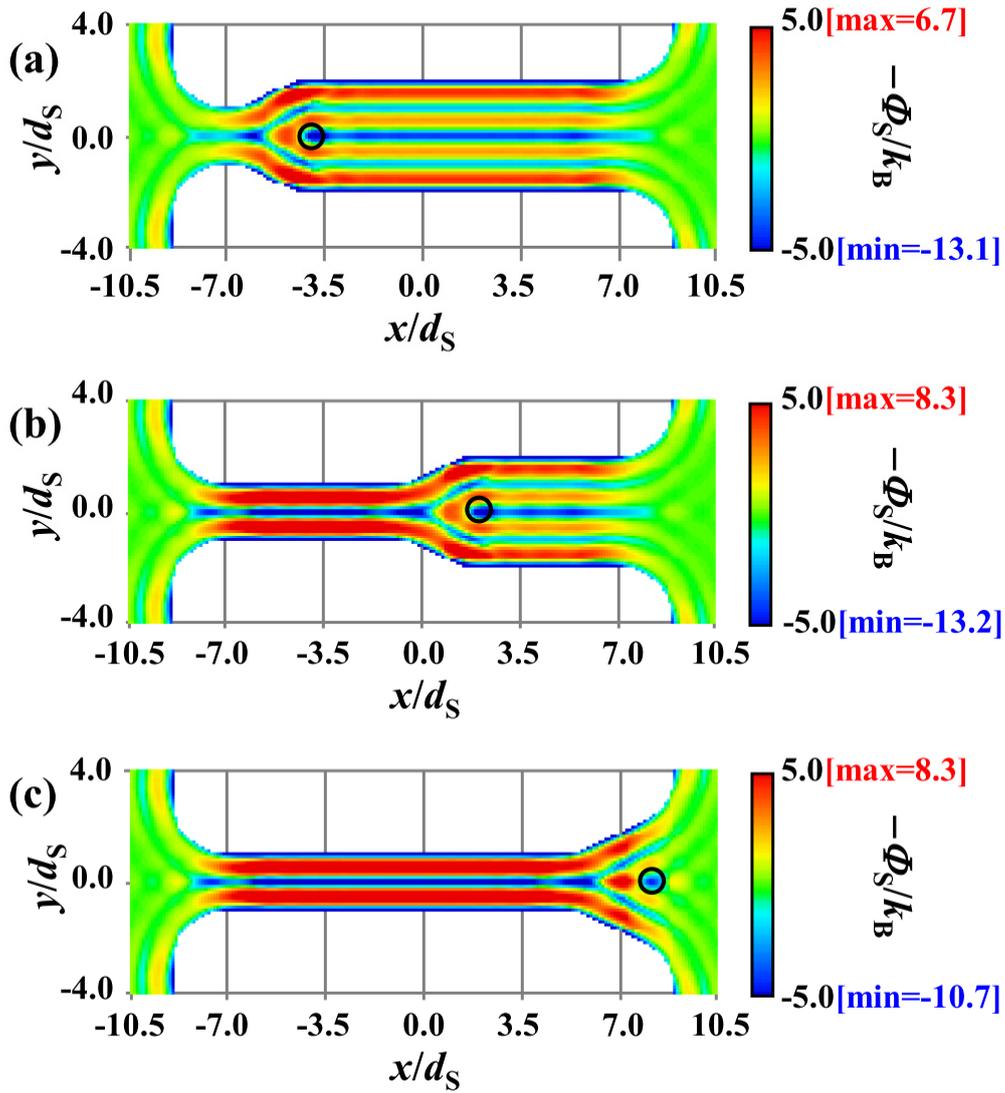


Fig. 3.9. Distributions of $-\Phi_S/k_B$ on the cross section of $z=0$ for vessel geometries (b), (c), and (d) illustrated in Figs. 3.2(b), (c) and (d), respectively. The solute size d_B is $4d_S$. $-\Phi_S/k_B$ becomes lower as the color approaches dark blue, and it becomes higher as the color approaches dark red (“max” and “min” represent the maximum and minimum values, respectively). The open circle indicates the position where the entropic potential is locally minimum.

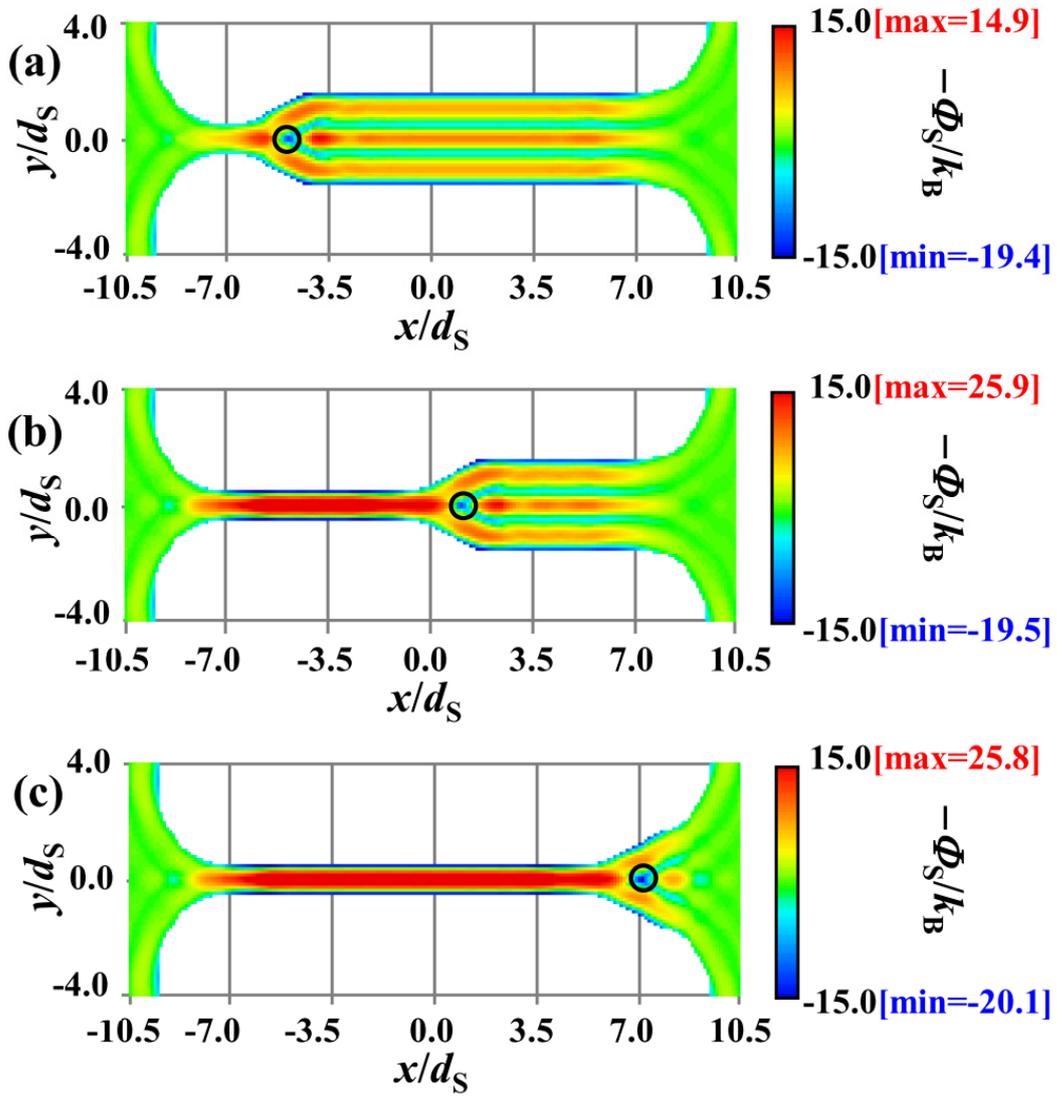


Fig. 3.10. Distributions of $-\Phi_S/k_B$ on the cross section of $z=0$ for vessel geometries (b), (c), and (d) illustrated in Figs. 3.2(b), (c) and (d), respectively. The solute size d_B is $5d_S$. $-\Phi_S/k_B$ becomes lower as the color approaches dark blue, and it becomes higher as the color approaches dark red (“max” and “min” represent the maximum and minimum values, respectively). The open circle indicates the position where the entropic potential is locally minimum.

Fig. 3.11(a) shows the entropic forces acting on the solute with $d_B=2d_S$ in vessel geometry (c). They are for $(L_1, L_2)=(6d_S, 6d_S)$ and for $(L_1, L_2)=(6.1d_S, 5.9d_S)$, respectively. A similar plot is made in Fig. 3.11(b) for the solute with $d_B=5d_S$. When the solute is at the location indicated by the dash-dot line representing the potential minimum for $(L_1, L_2)=(6d_S, 6d_S)$, the force looking like that for $(L_1, L_2)=(6.1d_S, 5.9d_S)$ is applied to the solute. This is because the solute is always on the left side of (and close to) the location with the potential minimum when the variation of the vessel geometry is relatively faster than the solute motion. The force continuously accelerates the solute motion in the right direction along the x -axis. The maximum value of the force increases remarkably as the solute size becomes larger: For instance, it is approximately 25 times stronger for $d_B=2d_S$ than for $d_B=5d_S$.

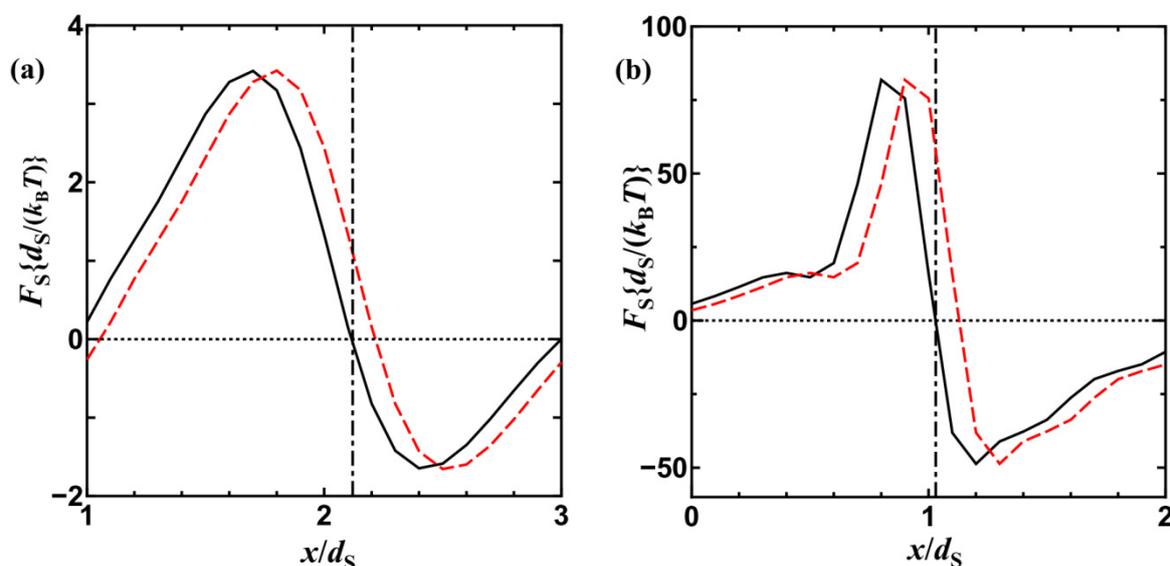


Fig. 3.11. Entropic forces acting on the solute in vessel geometry (c) illustrated in Fig. 3.2(c) for $(L_1, L_2)=(6d_S, 6d_S)$ (black and solid) and for $(L_1, L_2)=(6.1d_S, 5.9d_S)$ (red and broken). The solute size d_B is $2d_S$ in (a) and $5d_S$ in (b). The position of the potential minimum for $(L_1, L_2)=(6d_S, 6d_S)$ is indicated by the dash-dot line.

3.3.3. Entropic potentials for vessel geometries (d), (e), and (f) in Fig. 3.2

To complete the release of the solute, the inner diameter at the exit D is gradually reduced as illustrated in Figs. 3.2(d), (e), and (f). Fig. 3.12(a) shows the profiles of $-\Phi_S/k_B$ along the x -axis for vessel geometries (d) ($D=8d_S$), (e) ($D=7d_S$), (f) ($D=6d_S$), and (g) ($D=5.5d_S$) for the solute with $d_B=2d_S$. Those for the solute with $d_B=5d_S$ are displayed in Fig. 3.12(b). The open circle indicates the position where the entropic potential is locally minimum. The coordinate, $(x/d_S, y/d_S, z/d_S)=(7, 0, 0)$, corresponds to the position where the right half of the solute is outside the vessel. As D decreases, the solute is driven to move further in the right direction along the x -axis and the barrier becomes progressively lower. It should be quite easy for the solute with $d_B=2d_S$ to overcome the barrier for being released to the bulk, even in vessel geometry (d) ($D=8d_S$). However, the barriers become higher as the

solute size increases. For the solute with $d_B=5d_S$, though the velocity of the solute in the right direction along the x -axis is expected to be much higher than for the solute with $d_B=2d_S$, it is not definite if the barriers in vessel geometries (d) ($D=8d_S$), (d) with $D=7d_S$, and (e) ($D=6d_S$) can be overcome. Fortunately, the barrier decreases to $\sim 3.5k_B T$ in vessel geometry (f) ($D=5.5d_S$). This value should be sufficiently low in the light of the very strong entropic force discussed in 3.3.2.

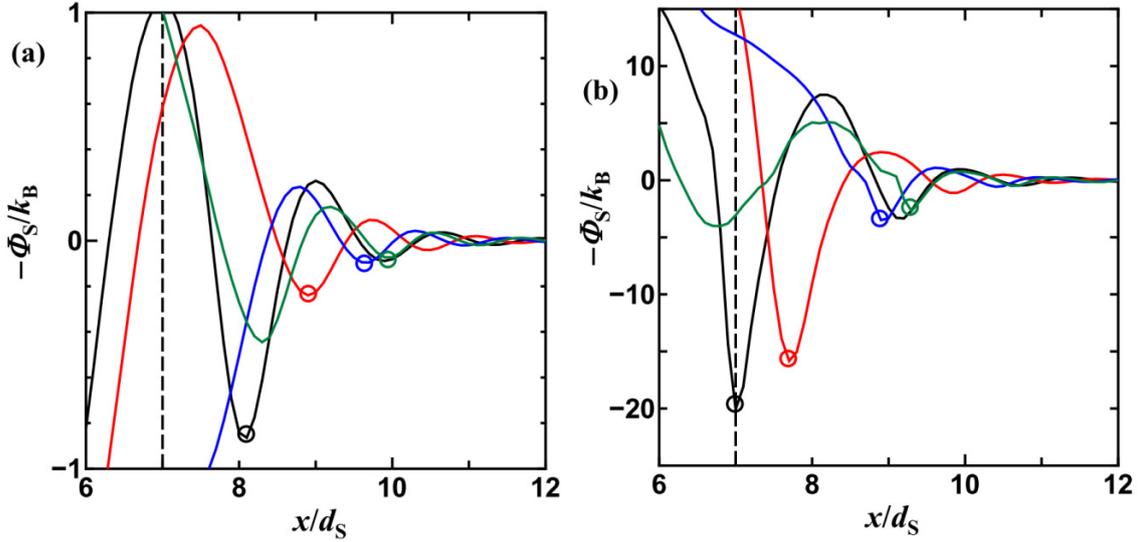


Fig. 3.12. Profiles of $-\Phi_S/k_B$ along the x -axis for vessel geometries (d) with $D=8d_S$ (black), (d) with $D=7d_S$ (red), (e) (blue), and (f) (green). The solute size d_B is $2d_S$ in (a) and $5d_S$ in (b). The open circle indicates the position where the entropic potential is locally minimum. The coordinate, $(x/d_S, y/d_S, z/d_S)=(7, 0, 0)$, corresponds to the position where the right half of the solute is outside the vessel. The broken line in each plot represents position of the right end of the vessel, $x/d_S=7$.

From the above argument, we conclude that solutes with a wide range of sizes can be released to the bulk. The key idea is a continuous variation of the vessel geometry accompanying that of the spatial distribution of the entropic potential. We remark that the solute velocity in the final vessel geometry considered becomes higher as the vessel length increases. In other words, a longer vessel enables the solute to overcome a higher barrier for being released to the bulk. Even a barrier that is considerably higher than $k_B T$ could readily be overcome. The very large value of the length of TolC, $\sim 50d_S$, may play crucially important roles in the solute release.

Within the cylindrical vessel with two open ends (i.e., entrance and exit for the solute), the solute motion is thus directed only toward the right end. Such unidirectional motion cannot be performed by the mere diffusion: The solute release from the right end is achievable only with the probability of 0.5 at most; and the actual probability is even lower than 0.5 since the initial position of the solute is in the vicinity of the left end. Even when the

left end is closed (it is confined by AcrB in the real system), the probability that the solute remains near the left end is considerably high.

3.3.4. Case where solute size differs from integral multiplication of d_S

We consider the solute sizes that differ from the integral multiplication of d_S : Solutes with the eleven sizes, $d_B=3d_S+0.1md_S$ ($m=0-10$), are compared. We find that the stripe pattern observed in Fig. 3.3 persists even for the solute sizes differing from the integral multiplication of d_S . Fig. 3.13 shows the profiles of $-\Phi_S/k_B$ in the radial direction along “ $x/d_S=-p$ and $z=0$ ” for $m=0-10$ ($p=5.7, 5.6, 5.8, 6.1, 5.9, 6.0, 6.0, 5.9, 5.7, 5.6$, and 5.5 , respectively). This figure corresponds exactly to Fig. 3.6, and the open circle indicates the position where the solute is stabilized with the highest probability right after the solute insertion. It is observed that the barrier to overcome for the solute to reach $y/d_S=1$ is sufficiently high for all the solute sizes tested. At $y=0$, $-\Phi_S/k_B$ takes the maximum value for $m=0-2$ (category 1) whereas it takes the minimum value for $m=7-10$ (category 2). $-\Phi_S/k_B$ for $m=3$ behaves like that in category 1 in the sense that the solute rarely comes to the center of $y=0$. The profiles of $-\Phi_S/k_B$ for $m=5$ and 6 are qualitatively similar to those in category 2. $-\Phi_S/k_B$ for $m=4$ ($d_B=3.4d_S$) exhibits intermediate behavior.

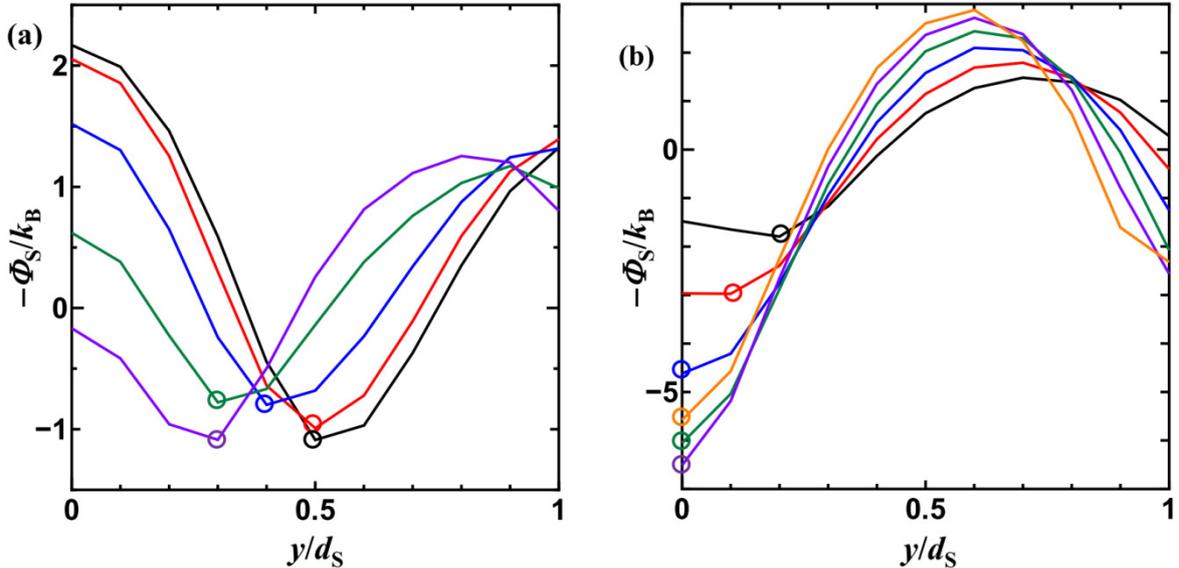


Fig. 3.13. Profiles of $-\Phi_S/k_B$ in the radial direction for solutes with the eleven different sizes, $d_B=3d_S+0.1md_S$ ($m=0-10$). (a) $m=0$ (black), $m=1$ (red), $m=2$ (blue), $m=3$ (green), and $m=4$ (purple). (b) $m=5$ (black), $m=6$ (red), $m=7$ (blue), $m=8$ (green), $m=9$ (purple), and $m=10$ (orange). This figure corresponds exactly to Fig. 3.6. The open circle indicates the position where the solute is stabilized with the highest probability right after the solute insertion.

The distributions of $-\Phi_S/k_B$ on the cross section of $z=0$ for vessel geometries (b), (c), and (d) (see Figs. 3.2(b)-(d)) are shown for $m=3$ in Fig. 3.14. Those for $m=4$ and 5 are

presented in Figs. 3.15 and 3.16, respectively. We find that the basic patterns of the distributions for $m=1-3$ is qualitatively similar to those for $m=0$ ($d_B=3d_S$). This sentence is valid when “ $m=1-3$ ” and “ $m=0$ ($d_B=3d_S$)” are replaced by “ $m=5-9$ ” and “ $m=10$ ($d_B=4d_S$)”, respectively. Even for $m=4$, the potential becomes positive next to the solute on the left side and an entropic force continuously acts on the solute in the right direction along the x -axis and continues to accelerate its motion during the variation of the vessel geometry. The figure corresponding to Fig. 3.12 is presented for $m=4$ as Fig. 3.17. The solute release to the bulk is completed by means of the vessel-geometry variation, (d)→(e)→(f), illustrated in Fig. 3.2. Similar arguments are possible for the solutes with $d_B=2d_S+0.1md_S$ and $d_B=4d_S+0.1md_S$ ($m=0-10$). Thus, solutes with a wide range of sizes can be released *using the same vessel-geometry variation*.

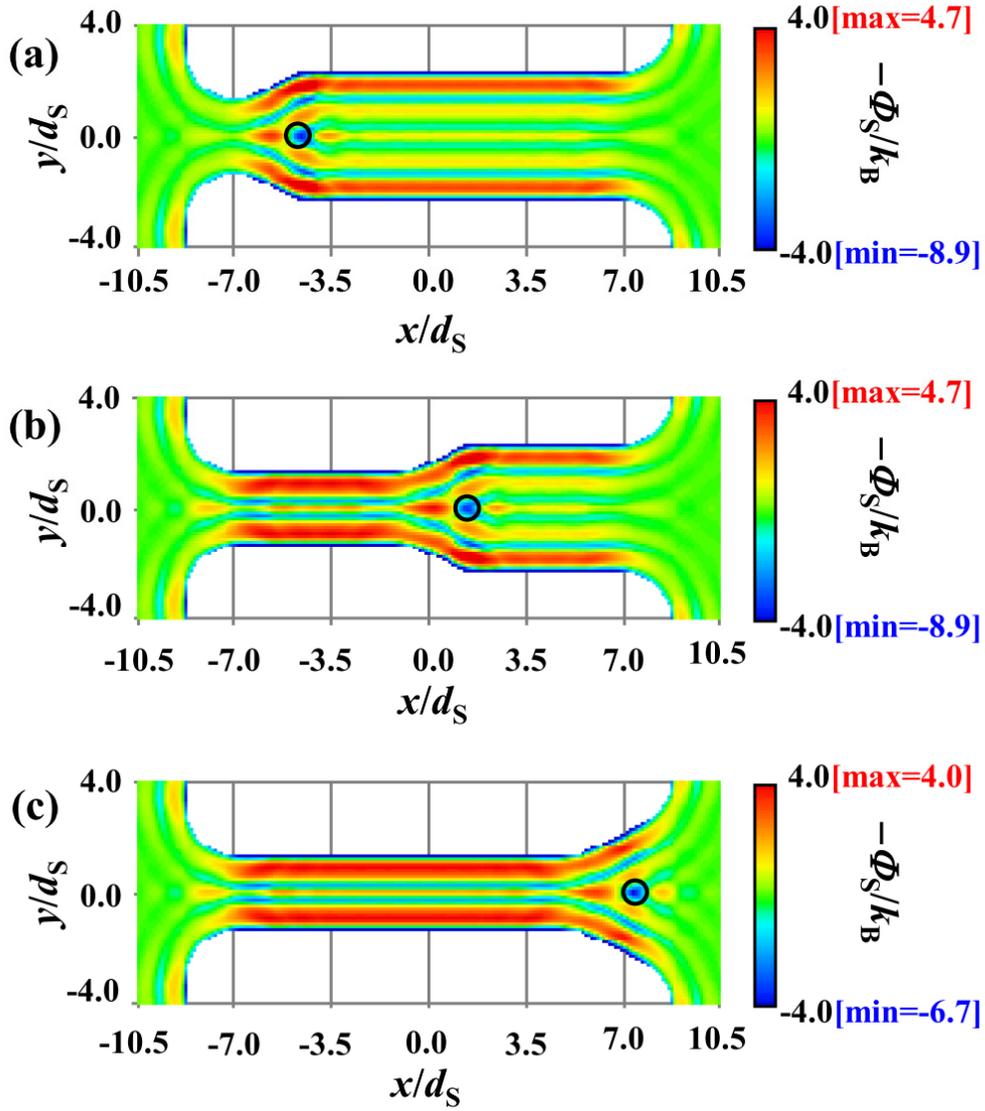


Fig. 3.14. Distributions of $-\Phi_S/k_B$ on the cross section of $z=0$ for vessel geometries (b), (c), and (d) (see Figs. 2(b)-(d)) for the solute with $d_B=3.3d_S$. $-\Phi_S/k_B$ becomes lower as the color approaches dark blue, and it becomes higher as the color approaches dark red (“max” and “min” represent the maximum and minimum values, respectively). The open circle indicates the position where the entropic potential is locally minimum. This figure is to be compared with Figs. 3.8 and 3.9.

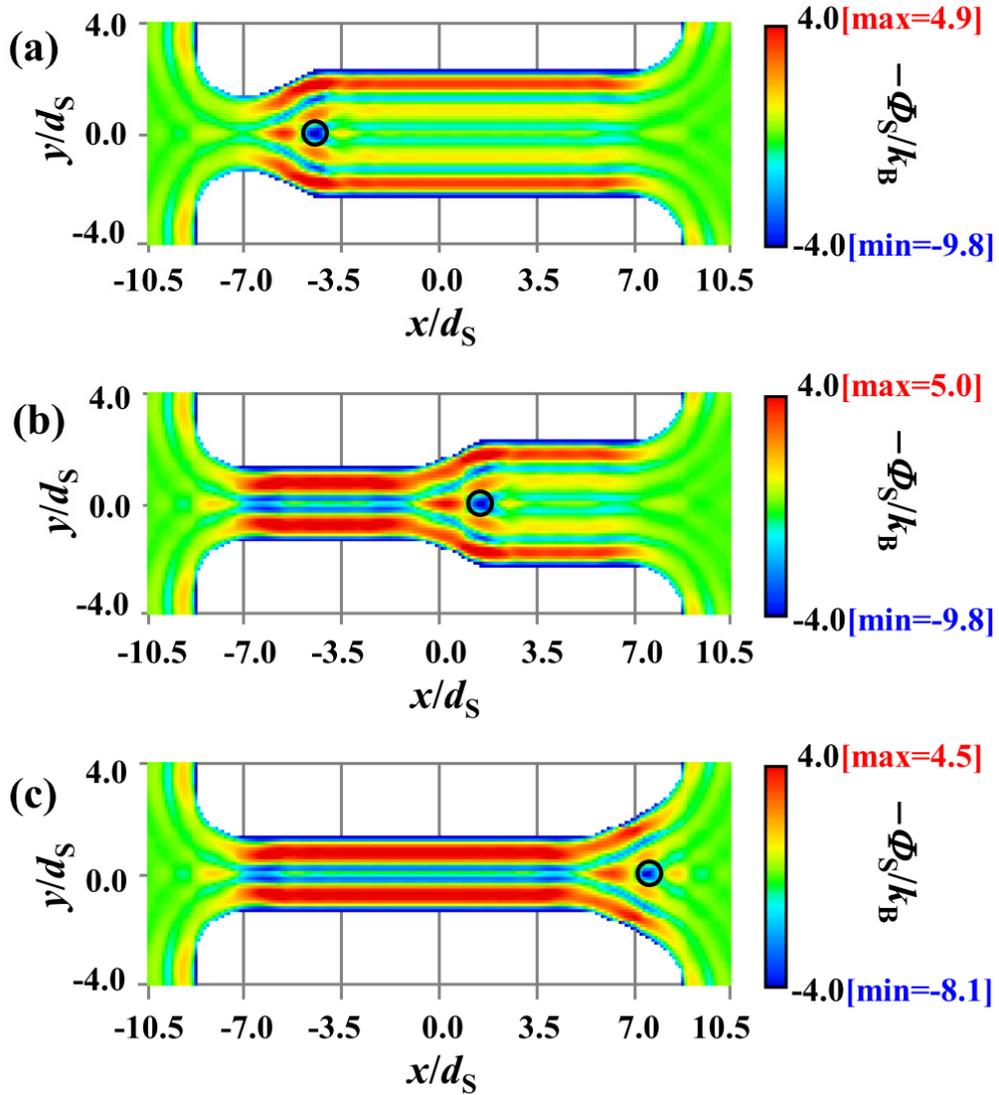


Fig. 3.15. Distributions of $-\Phi_S/k_B$ on the cross section of $z=0$ for vessel geometries (b), (c), and (d) (see Figs. 2(b)-(d)) for the solute with $d_B=3.4d_S$. $-\Phi_S/k_B$ becomes lower as the color approaches dark blue, and it becomes higher as the color approaches dark red (“max” and “min” represent the maximum and minimum values, respectively). The open circle indicates the position where the entropic potential is locally minimum. This figure is to be compared with Figs. 3.8 and 3.9.

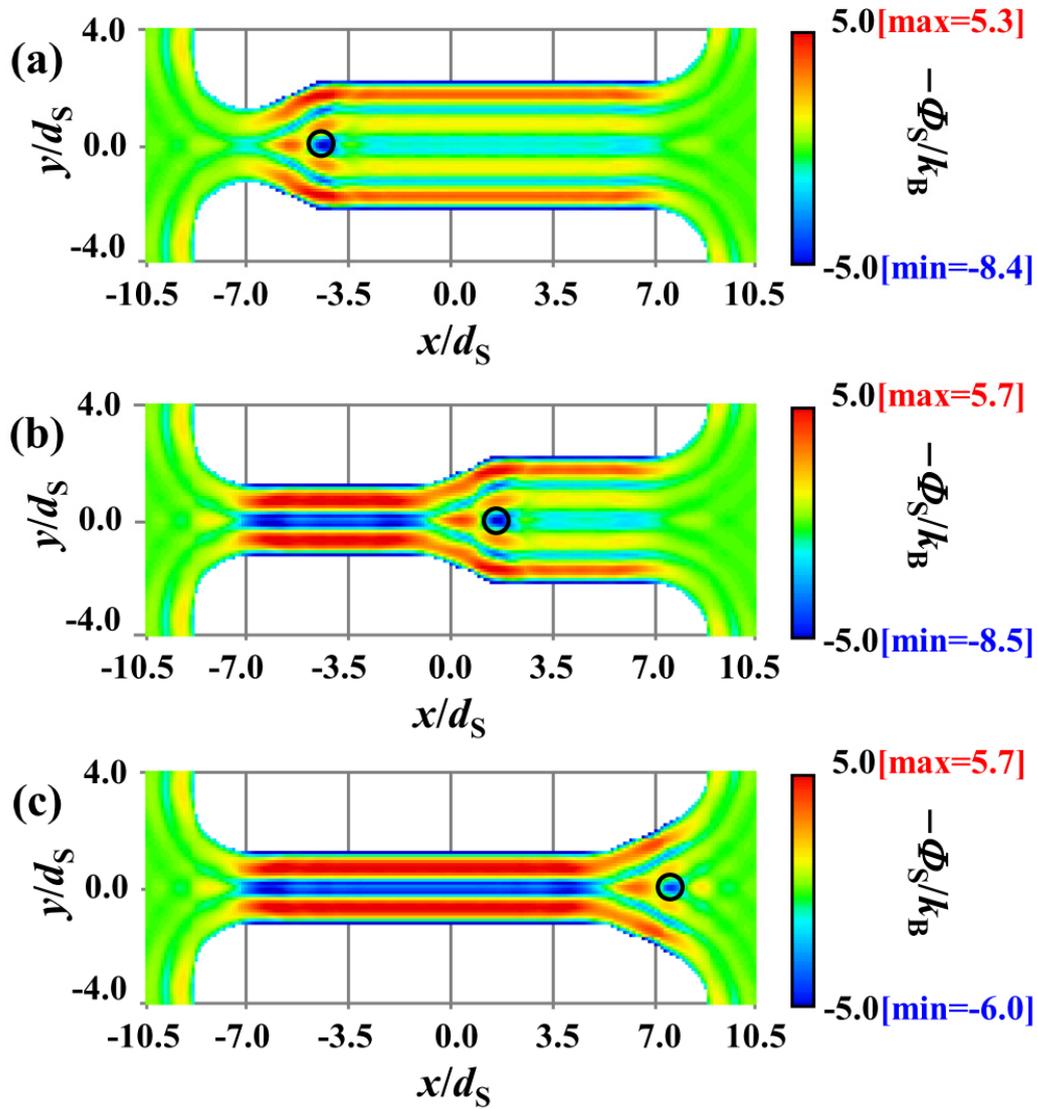


Fig. 3.16. Distributions of $-\Phi_S/k_B$ on the cross section of $z=0$ for vessel geometries (b), (c), and (d) (see Figs. 2(b)-(d)) for the solute with $d_B=3.5d_S$. $-\Phi_S/k_B$ becomes lower as the color approaches dark blue, and it becomes higher as the color approaches dark red (“max” and “min” represent the maximum and minimum values, respectively). The open circle indicates the position where the entropic potential is locally minimum. This figure is to be compared with Figs. 3.8 and 3.9.

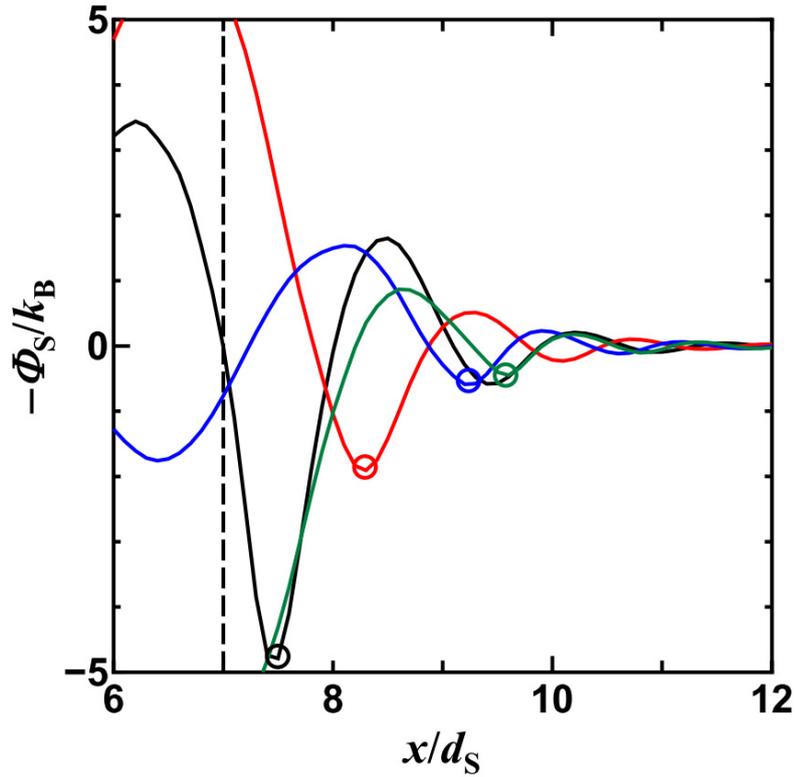


Fig. 3.17. Plot for the solute with $d_B=3.4d_S$ corresponding to Fig. 3.12. The open circle indicates the position where the entropic potential is locally minimum. The coordinate, $(x/d_S, y/d_S, z/d_S)=(7, 0, 0)$, corresponds to the position where the right half of the solute is outside the vessel, $x/d_S=7$.

3.3.5. Another method of vessel-geometry variation

The method of the vessel-geometry variation achieving the entropic release for solutes with a wide range of sizes is not unique. To demonstrate this, we consider another type of variation illustrated in Fig. 3.18. Tapering in the diameter from $8d_S$ to $6d_S$ is applied only to a small portion of the vessel, and the portion is continuously moved in the right direction toward the exit. Fig. 3.19 shows the distributions of $-\Phi_S/k_B$ on the cross section of $z=0$ for vessel geometries (b), (c) with $L_1=L_2=5d_S$, (d), and (e) (see Fig. 3.18) for the solute with $d_B=3d_S$. The open circle indicates the position where the entropic potential is locally minimum. Those for the solute with $d_B=4d_S$ are displayed in Fig. 3.20. An entropic force continuously acts on the solute in the right direction along the x -axis and continues to accelerate its motion during the variation of the vessel geometry. The solute is released to the bulk without difficulty.

We have thus obtained qualitatively the same results for the vessel-geometry variations illustrated in Figs. 3.2 and 3.18. A vessel in the real system is expected to exhibit a more complex variation. However, the geometric properties play essential roles and should appropriately be designed just for the portion around the position where the entropic force acts on the solute. The geometric variation for the other portions can be different from that assumed in the present study.

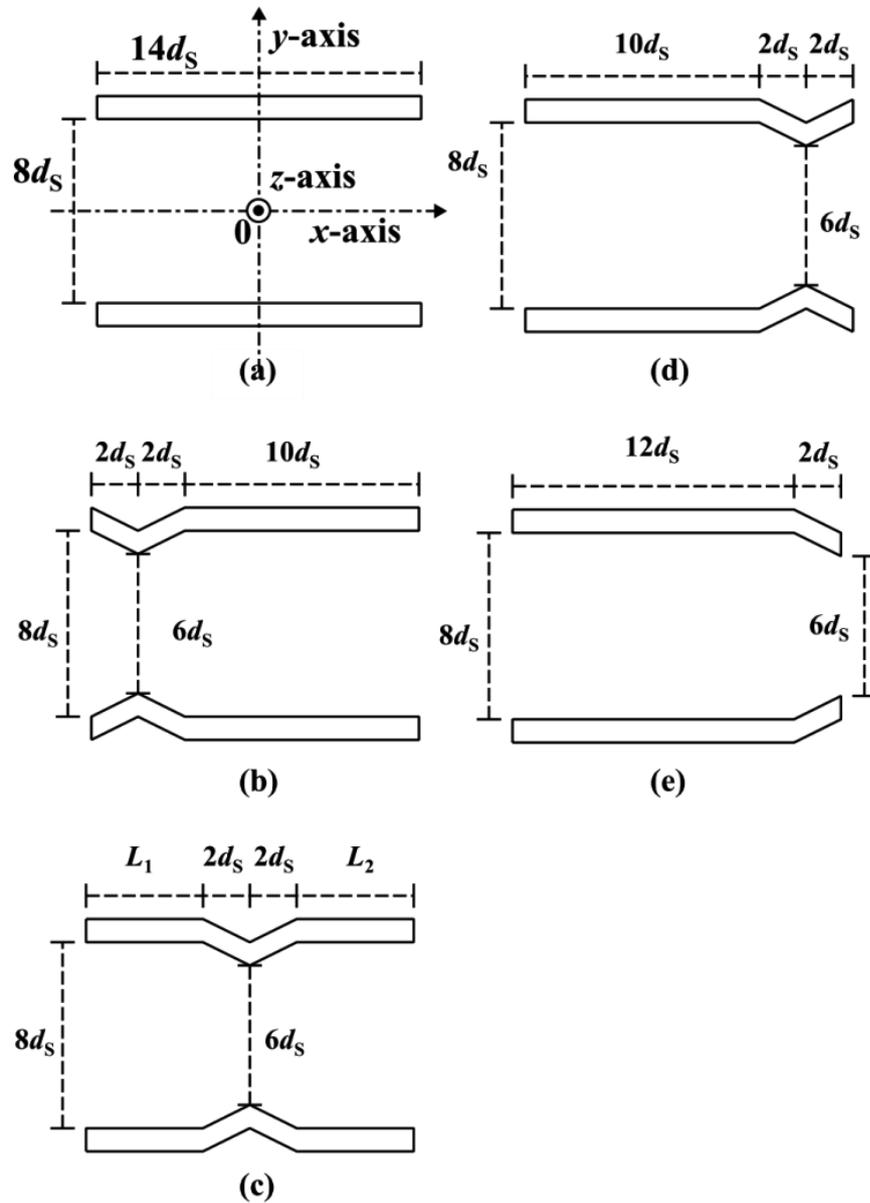


Fig. 3.18. Variation of vessel geometry: (a)→(b)→(c)→(d)→(e). The geometry in (a) is exactly the same as that in Fig. 3.2. The coordinate system is chosen as illustrated here. L_1 and L_2 are gradually increased and decreased, respectively. The cross section of $z=0$ is shown for each geometry. Though the geometry variation is illustrated in a stepwise manner, it is made *continuously*.

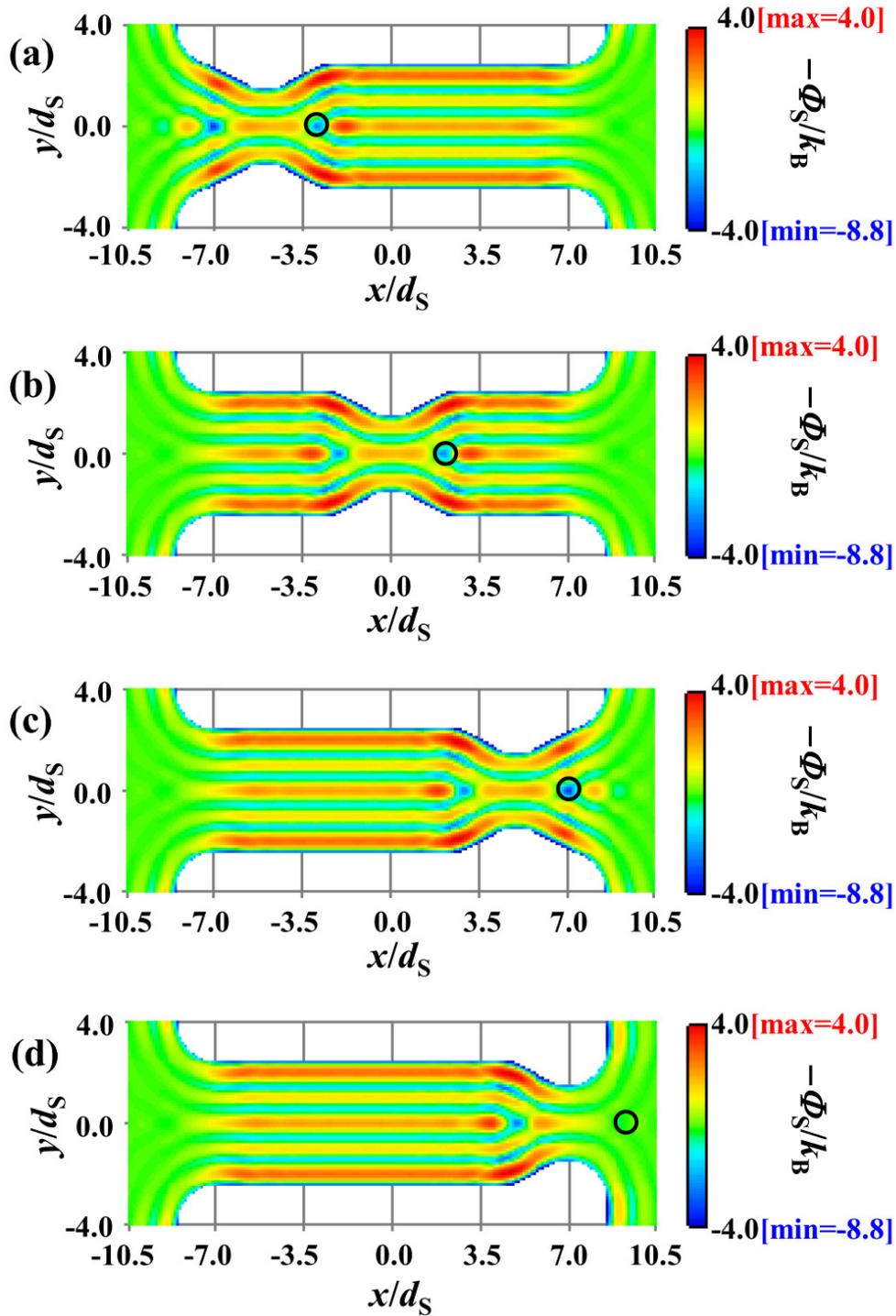


Fig. 3.19. Distributions of $-\Phi_S/k_B$ on the cross section of $z=0$ for vessel geometries (b), (c) with $L_1=L_2=5d_S$, (d), and (e) illustrated in Fig. 3.18. The solute size d_B is $3d_S$. $-\Phi_S/k_B$ becomes lower as the color approaches dark blue, and it becomes higher as the color approaches dark red (“max” and “min” represent the maximum and minimum values, respectively). The open circle indicates the position where the entropic potential is locally minimum.

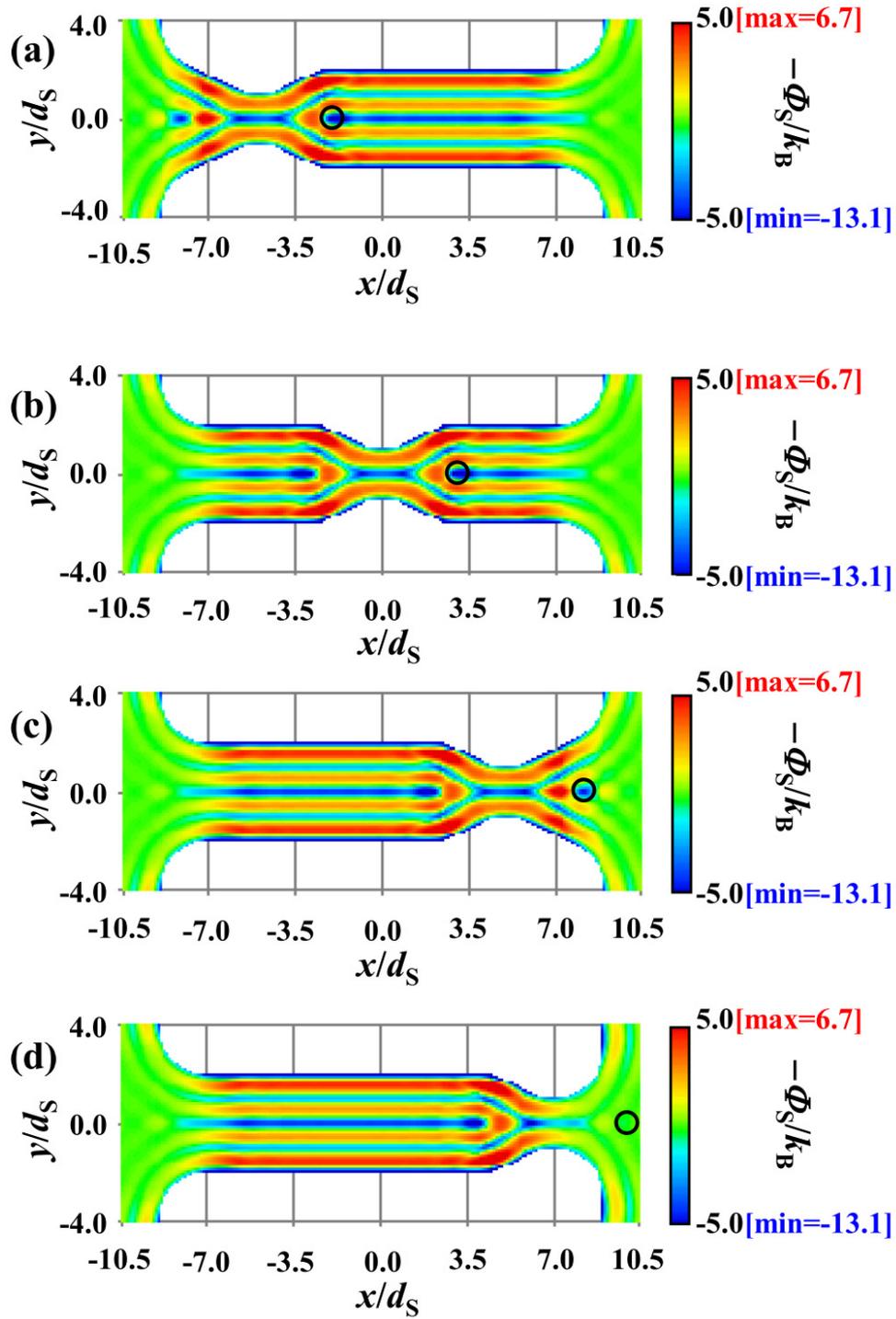


Fig. 3.20. Distributions of $-\Phi_S/k_B$ on the cross section of $z=0$ for vessel geometries (b), (c) with $L_1=L_2=5d_S$, (d), and (e) illustrated in Fig. 3.18. The solute size d_B is $4d_S$. $-\Phi_S/k_B$ becomes lower as the color approaches dark blue, and it becomes higher as the color approaches dark red (“max” and “min” represent the maximum and minimum values, respectively). The open circle indicates the position where the entropic potential is locally minimum.

3.3.6. Applicability of results obtained to AcrB and ABC transporter

The principal results obtained are as follows: The entropic release of the solute is not feasible as long as the vessel geometry is fixed; it can be accomplished by a continuous variation of the vessel geometry; and solutes with a wide range of sizes can be handled. Namely, a time-dependent entropic force continuing to accelerate the solute motion in the axial direction toward the exit is the key to the multidrug efflux. These results, which are obtained using simple model calculations, seem to be fairly general and applicable to other types of efflux transporters. In the ABC transporter^{27,28} with the inward-facing structure and in AcrB¹⁻⁸ with the structure drawn in blue in Fig. 3.1, a solute can spontaneously be inserted into a cavity within the transporter or AcrB. In the transporter, after the solute insertion, the inward-facing structure is gradually changed to the outward-facing one, performing the switch from insertion to release. Similarly, in AcrB the solute release is promoted by a gradual variation from the structure drawn in blue to that drawn in red (see Fig. 3.1). Our view thus suggested concerning the insertion/release process for AcrB is substantially different from the one relying on the mere diffusion which is unable to perform the unidirectional solute motion.

In AcrB the variation of the structure (i.e., the geometry) is induced by the proton motive force,¹⁻⁸ while in ABC transporter it is controlled by the cycle comprising the binding of adenosine triphosphate (ATP), hydrolysis of ATP, and dissociation of Pi and adenosine diphosphate (ADP).^{27,28} Despite this difference, these two protein complexes share the feature that the structure is continuously changed to carry out the switch from insertion to release for the solute.

3.4. Conclusions

We have investigated insertion and release of a solute into and from a cylindrical vessel possessing an entrance at one end and an exit at the other end for the solute. This model vessel mimics TolC, an important component of the multidrug efflux transporter, AcrA/AcrB/TolC.⁴⁴ The spatial distribution of the solute-vessel potential of mean force (PMF) formed by the solvent is calculated using the three-dimensional integral equation theory^{9,10,14-17,21-26} combined with rigid-body models in which the constituents interact only through hard-body potentials. With such models, all of the allowed system configurations share the same energy, and the system behavior is purely entropic in origin. It has been demonstrated that the two opposite events, insertion and release, can be driven solely by the entropic effect arising not from the electrostatic and van der Waals interactions for solute-vessel, solute-solvent, solvent-vessel, and solvent-solvent pairs but from the translational displacement of solvent molecules.

A principal aim of the present study is to develop a novel physical picture of the multidrug efflux⁶⁻⁸ implying that solutes such as drug molecules with diverse properties (i.e., solvophobic and solvophilic solutes with a wide range of sizes) can be handled. The insertion/release process can be described in terms of the PMF consisting of the energetic and entropic components. The energetic component is strongly dependent on the solute-solvent

and vessel inner surface-solvent affinities, whereas the entropic component is rather insensitive to them. The basis of our physical picture is that the multidrug efflux can be assured under a condition where the entropic component dominates. This condition is satisfied when the inner surface of the vessel is neither solvophobic nor solvophilic. As long as the vessel geometry is fixed, however, the entropic component does not work for release. A finding is that a solute which has been inserted can also be released entropically using a continuous variation of the vessel geometry. Two typical examples of the variation are illustrated in Figs. 3.2 and 3.18. The variation is never unrealistic for the following reasons: The proton motive force causes structural changes of AcrB, they are transmitted to TolC through AcrA,^{5,6,8} possibly leading to a continuous vessel-geometry variation of TolC; and a recent molecular dynamics simulation study⁴⁴ has suggested that TolC can vary its geometric characteristics even by itself (an observed variation is a peristaltic motion of the periplasmic domain).

The variation of the vessel geometry is initiated as soon as the solute is inserted into the vessel at its entrance. In the real system, the solute insertion is made with the aid of AcrB: AcrB interacting with TolC sends the solute to the central position within the vessel cavity of TolC. The manner of the vessel-geometry variation proposed in the present study generates a time-dependent entropic force. There are three different time scales: those of the solvent motion, variation in the vessel geometry, and solute motion. The time scale of the solvent motion is the fastest, and the solvent is practically in equilibrium with the solute-vessel configuration all the time. We assume that the variation of the vessel geometry (that is, the variation of the entropic potential) is relatively faster than the solute motion. The faster variation of the vessel geometry could be realized by structural changes of AcrB caused by the proton motive force. The time-dependent entropic force then continues to accelerate the solute motion in the axial direction toward the exit. A larger solute is subjected to a stronger force. The solute velocity at the exit becomes higher as the solute size or the vessel length increases. Solutes with a wide range of sizes can be released using the same manner of the variation. It is quite interesting and important that such rich behavior is observed in the very simple model system adopted in the present study.

Once a solute is sent from AcrB to the central position within the TolC cavity, the solute must be ejected to the external medium through the exit before the next solute is sent. Namely, the ejection of every solute needs to be finished with sufficient rapidness. This requirement can be met with much more certainty when the solute motion is directed only toward the exit. This unidirectional motion is assured by the time-dependent entropic force arising from the vessel-geometry variation proposed in the present study. Our physical picture of the multidrug efflux is clearly distinguished from the previously reported one^{4,5,12} assuming the existence of multifunctional ligand-binding sites which recognize various types of solutes and relying on the mere diffusion (i.e., the diffusion in the presence of no particular potential field) by which the unidirectional solute motion is not realized. In earlier works,⁹⁻¹¹ we showed that the solvent entropy plays crucially important roles in the high selectivity pertaining to the receptor-ligand binding. Interestingly, it is also true in the multidrug efflux which is in marked contrast with the high selectivity.

To complement the physical picture of the multidrug efflux, we need to analyze dynamic

aspects of the entropic release by solving the Fokker-Planck equation in the presence of a time-dependent force field calculated in the present study. The time required for overcoming the barrier in the final stage of the solute release at the exit is an important quantity. However, it is to be argued in the light of the interval between the two times when a solute and the next solute are sent from AcrB to the central position within the TolC cavity at the entrance. Even if the former time is shorter than the interval, the vessel-geometry variation should still be required. This is because the solute cannot not always arrive at the exit in time without its unidirectional motion toward the exit: The long length of TolC () is a noticeable factor. When the vessel-geometry variation is applied, the long length may inversely play crucial roles in the solute release as discussed in 3.3.3. (The interval mentioned above is unknown and to be measured in future experimental and theoretical works.)

The details of the polyatomic structures of the vessel and/or the solute are also essential factors whose effects are to be examined. We remark that our physical picture is fairly general and also applicable to AcrB and ATP-binding cassette (ABC) transporter^{27,28} as discussed in Sec. 3.6, pending detailed analyses. It is challenging to investigate the sequential structural change exhibited by each protomer of AcrB among three states in which a drug insertion, binding, and release take place.^{5,6,8} We have recently developed a novel physical picture⁴⁶⁻⁴⁸ for F₁-ATPase in which the $\alpha_3\beta_3$ complex always tries to form three regions which are tightly packed, moderately packed, and loosely packed, respectively, and these regions are cyclically exchanged. The inhomogeneous packing structure and the cyclic exchange are attributed to the maintenance of the maximal solvent entropy. This picture, which is consistent with the experimental observations for F₁-ATPase without the γ -subunit,⁴⁹ may also be relevant to the sequential structural change in AcrB. The most important matter is treating the solvent as an ensemble of particles with finite sizes to account for the imperative solvent-entropy effect (this has often been neglected as in a recent simulation study⁵⁰). Works along these lines are in progress in our group.

References

- ¹V. Koronakis, A. Sharff, E. Koronakis, B. Luisi, and C. Hughes, *Nature* **405**, 914 (2000).
- ²S. K. Buchanan, *Trends Biochem. Sci.* **26**, 3 (2001).
- ³V. Koronakis, J. Eswaran, and C. Hughes, *Annu. Rev. Biochem.* **73**, 467 (2004).
- ⁴S. Murakami, R. Nakashima, E. Yamashita, T. Matsumoto, and A. Yamaguchi, *Nature* **443**, 173 (2006).
- ⁵S. Murakami, *Curr. Opin. Struct. Biol.* **18**, 459 (2008).
- ⁶M. A. Seeger, K. Diederichs, T. Eicher, L. Brandstätter, A. Schiefner, F. Verrey, and K. M. Pos, *Curr. Drug Targets* **9**, 729 (2008).
- ⁷X.-Z. Li and H. Nikaido, *Drugs* **69**, 1555 (2009).
- ⁸K. M. Pos, *Biochim. Biophys. Acta* **1794**, 782 (2009).
- ⁹M. Kinoshita and T. Oguni, *Chem. Phys. Lett.* **351**, 79 (2002).
- ¹⁰M. Kinoshita, *J. Chem. Phys.* **116**, 3493 (2002).
- ¹¹H. Oshima, S. Yasuda, T. Yoshidome, M. Ikeguchi, and M. Kinoshita, *Phys. Chem. Chem. Phys.* **13**, 16236 (2011).
- ¹²T. Imai, N. Miyashita, Y. Sugita, A. Kovalenko, F. Hirata, and A. Kidera, *J. Phys. Chem. B* **115**, 8288 (2011).
- ¹³R. Schulz, A. V. Vargiu, P. Ruggerone, and U. Kleinekathöfer, *J. Phys. Chem. B* **115**, 8278 (2011).
- ¹⁴K. Amano and M. Kinoshita, *Chem. Phys. Lett.* **488**, 1 (2010).
- ¹⁵K. Amano and M. Kinoshita, *Chem. Phys. Lett.* **504**, 221 (2011).
- ¹⁶K. Amano, H. Oshima, and M. Kinoshita, *J. Chem. Phys.* **135**, 185101 (2011).
- ¹⁷H. Mishima, H. Oshima, S. Yasuda, K. Amano, and M. Kinoshita, *Chem. Phys. Lett.* **561-562**, 159 (2013).
- ¹⁸A. L. Horwich, W. A. Fenton, E. Chapman, and G. W. Farr, *Annu. Rev. Cell Dev. Biol.* **23**, 115 (2007).
- ¹⁹T. K. Chaudhuri, V. K. Verma, and A. Maheshwari, *Prog. Biophys. Mol. Biol.* **99**, 42 (2009).
- ²⁰S. Tanaka, Y. Kawata, G. Otting, N. E. Dixon, K. Matsuzaki, and M. Hoshino, *Biochim. Biophys. Acta* **1804**, 866 (2010).
- ²¹D. Beglov and B. Roux, *J. Chem. Phys.* **103**, 360 (1995).
- ²²M. Ikeguchi and J. Doi, *J. Chem. Phys.* **103**, 5011 (1995).
- ²³M. Kinoshita, *Chem. Phys. Lett.* **387**, 47 (2004).
- ²⁴Y. Harano and M. Kinoshita, *Biophys. J.* **89**, 2701 (2005).
- ²⁵M. Kinoshita, *Chem. Eng. Sci.* **61**, 2150 (2006).
- ²⁶K. Amano, T. Yoshidome, M. Iwaki, M. Suzuki, and M. Kinoshita, *J. Chem. Phys.* **133**, 045103 (2010).
- ²⁷K. Hollenstein, R. J. P. Dawson, and K. P. Locher, *Curr. Opin. Struct. Biol.* **17**, 412 (2007).
- ²⁸A. Ward, C. L. Reyes, J. Yu, C. B. Roth, and G. Chang, *Proc. Natl. Acad. Sci. U.S.A.* **104**, 19005 (2007).
- ²⁹M. Kinoshita, *J. Chem. Phys.* **128**, 024507 (2008).
- ³⁰T. Yoshidome, M. Kinoshita, S. Hirota, N. Baden, and M. Terazima, *J. Chem. Phys.* **128**,

225104 (2008).

- ³¹M. Kinoshita, *Front. Biosci.* **14**, 3419 (2009).
- ³²H. Oshima and M. Kinoshita, *J. Chem. Phys.* **138**, 245101 (2013).
- ³³T. Yoshidome and M. Kinoshita, *Phys. Rev. E* **79**, 090305(R) (2009).
- ³⁴H. Oshima, T. Yoshidome, K. Amano, and M. Kinoshita, *J. Chem. Phys.* **131**, 205102 (2009).
- ³⁵T. Yoshidome and M. Kinoshita, *Phys. Chem. Chem. Phys.* **14**, 14554 (2012).
- ³⁶J.-P. Hansen and L. R. McDonald, *Theory of Simple Liquids*, 3rd ed. (Academic, London, 2006).
- ³⁷R. Roth, B. Götzelmann, and S. Dietrich, *Phys. Rev. Lett.* **83**, 448 (1999).
- ³⁸P. Bryk, R. Roth, M. Schoen, and S. Dietrich, *Europhys. Lett.* **63**, 233 (2003).
- ³⁹P.-M. König, R. Roth, and S. Dietrich, *Phys. Rev. E* **74**, 041404 (2006).
- ⁴⁰P.-M. König, R. Roth, and S. Dietrich, *Europhys. Lett.* **84**, 68006 (2008).
- ⁴¹S. Asakura and F. Oosawa, *J. Chem. Phys.* **22**, 1255 (1954).
- ⁴²S. Asakura and F. Oosawa, *J. Polym. Sci.* **33**, 183 (1958).
- ⁴³P. Attard and G. N. Patey, *J. Chem. Phys.* **92**, 4970 (1990).
- ⁴⁴L. Vaccaro, K. A. Scott, and M. S. P. Sansom, *Biophys. J.* **95**, 5681 (2008).
- ⁴⁵J. C. Crocker, J. A. Matteo, A. D. Dinsmore, and A. G. Yodh, *Phys. Rev. Lett.* **82**, 4352 (1999).
- ⁴⁶T. Yoshidome, Y. Ito, M. Ikeguchi, and M. Kinoshita, *J. Am. Chem. Soc.* **133**, 4030 (2011).
- ⁴⁷T. Yoshidome, Y. Ito, N. Matubayasi, M. Ikeguchi, and M. Kinoshita, *J. Chem. Phys.* **137**, 035102 (2012).
- ⁴⁸Y. Ito, T. Yoshidome, N. Matubayasi, M. Kinoshita, and M. Ikeguchi, *J. Phys. Chem. B* **117**, 3298 (2013).
- ⁴⁹T. Uchihashi, R. Iino, T. Ando, and H. Noji, *Science* **333**, 755 (2011).
- ⁵⁰X.-Q. Yao, H. Kenzaki, S. Murakami, and S. Takada, *Nat. Commun.* **1**, 117 (2010).

Chapter 4. Statistical Thermodynamics for Functionally Rotating Mechanism of the Multidrug Efflux Transporter AcrB

4.1. Introduction

The *Escherichia coli* AcrA–AcrB–TolC tripartite complex¹⁻⁸ comprises the polytopic inner membrane protein AcrB,²⁻⁵ periplasmic adaptor protein AcrA,⁷ and outer membrane channel TolC.^{1,6} It has recently been found that AcrZ associates with the complex in an AcrB-dependent manner.⁸ The complex provides a good model for multidrug efflux pumps which are capable of extruding a variety of drugs. This capability causes a serious problem in medical practice. AcrB is a member of the resistance-nodulation-division (RND) superfamily⁹ found in all domains of life. AcrB, which cooperates with AcrA, TolC, and AcrZ, is in charge of the principal part of pumping drugs out of the cell from the inner membrane or periplasm through the TolC channel as illustrated in Fig. 4.1(a) (AcrZ may only enhance the ability of the complex to export certain classes of substrates).⁷⁻⁹

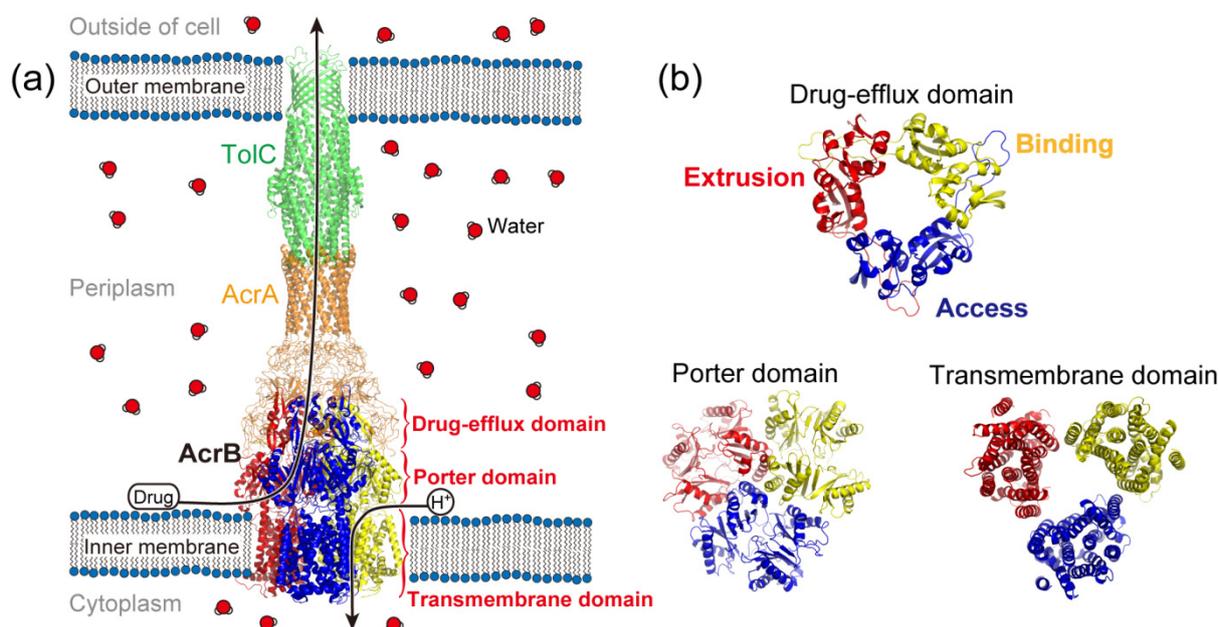


Fig. 4.1. (a) Schematic diagram of AcrA–AcrB–TolC tripartite complex. (b) Ribbon representation of the three domains of the protomers viewed from the top. “Access”, “binding”, and “extrusion” represent the protomers in three different states. This figure is drawn using PyMOL.

AcrB is a homotrimer with a triangular-prism shape, and each protomer comprises three domains: the transmembrane (TM), porter, and drug-efflux domains.²⁻⁵ Another important feature of AcrB is that its protomers take three different structures which are in access (A), binding (B), and extrusion (E) states along the drug transport cycle³⁻⁵ (see Fig. 4.1(b)). A “functionally rotating” picture has then been proposed for the transport of drugs by AcrB.³⁻⁵ In this picture, each protomer exhibits a sequential conformational change represented as (A, B, E)→(B, E, A)→(E, A, B)→(A, B, E) by utilizing proton binding and dissociation (i.e., the so-called proton-motive force). The result from a recent molecular dynamics (MD) simulation¹⁰ suggests that a single proton binds to Asp408 in the TM domain (the proton-binding site) in one protomer and the proton translocation stoichiometry is a single proton per step (e.g., (A, B, E)→(B, E, A)) in the trimer.

One might associate F₁-ATPase,¹¹⁻¹⁴ the $\alpha_3\beta_3\gamma$ complex, with AcrB.^{3,4} In F₁-ATPase, the γ subunit achieves a unidirectional rotation utilizing the ATP hydrolysis cycle (i.e., ATP binding, ATP hydrolysis, and dissociation of ADP and Pi). According to a recent experimental study,¹⁵ even the three β subunits in the $\alpha_3\beta_3$ complex *without* the γ subunit undergo cyclic conformational changes. It follows that this conformational change induces the rotation of the γ subunit. In recent years, the rotational mechanism of AcrB as well as that of F₁-ATPase has received much attention as a main topic in modern chemistry.

To date, MD simulations have been performed for AcrB rather extensively.^{10,16-22} However, each of them is focused on a particular aspect or an elementary process in the drug transport cycle such as drug extrusion upon induction of conformational changes,¹⁶ effect of a mutation on drug extrusion process,¹⁷ water roles during the drug extrusion,¹⁸ water distribution within each protomer relevant to possible routes of the proton transfer,¹⁹ unidirectional peristaltic movement of a drug toward the extrusion funnel of TolC,²⁰ multidrug binding properties of the pocket,²¹ proton translocation stoichiometry,¹⁰ and opening and closing motions of the porter domains.²² An analysis has also been performed for multidrug binding properties of the pocket using a statistical-mechanical theory for molecular fluids.²³ Although a significant amount of useful information has been made available by these studies, the mechanism of the functional rotation and energetics of the whole cycle (in the present article, one cycle corresponds to (A, B, E)→(B, E, A)) remain unresolved. What is strongly desired in the next step is an analysis for outlining the whole cycle using statistical thermodynamics rather than a detailed computer simulation of a particular process.

Let us consider the entropic excluded volume (EV) effect on the structure formation of a protein.²⁴⁻²⁸ The backbone and side chains of a protein generate EVs which the centers of solvent molecules cannot enter. The formation of α -helix or β -sheet and close packing of side chains is accompanied by overlaps of the EVs (see Fig. 4.2).²⁸ Consequently, the total EV decreases, leading to a corresponding increase in the total volume available to the translational displacement of solvent molecules in the system followed by a gain of solvent

entropy. As the backbone and side chains are more efficiently packed, the solvent entropy becomes higher: It is desired for a protein or a complex of proteins that the backbones and side chains be closely packed, like a three-dimensional jigsaw puzzle. However, this is not always possible. Even in cases where the overall close packing is not achievable, there are certainly the portions that can closely be packed. It is important to pack such portions preferentially. Thus, in general the packing efficiency is not uniform in a protein or a complex of proteins, and this ununiformity often plays crucially important roles in its functioning.^{29,30} Refer to Sec. S1 of the Supplement for more details.

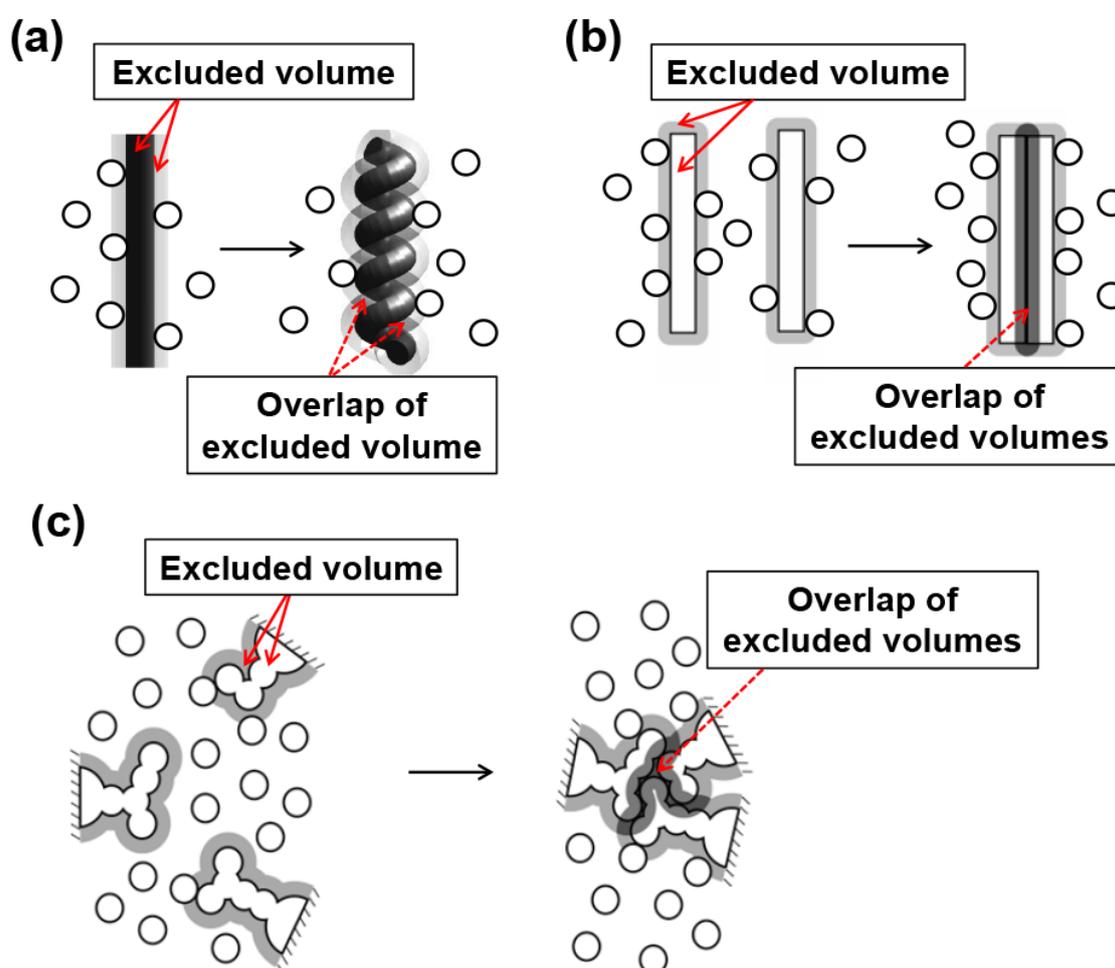


Fig. 4.2. (a) Formation of α -helix by a portion of the backbone. (b) Lateral contact of (formation of β -sheet by) portions of the backbone. (c) Close packing of side chains. Total excluded volume decreases by the overlapped volume marked in dark gray, leading to a corresponding increase in the total volume available to the translational displacement of solvent molecules in the system.

The porter and drug-efflux domains are immersed in water and the TM domain is within nonpolar chains of lipid molecules. The solvent-entropy effect emphasized in the present article originates from the translational displacement of water molecules²⁴⁻²⁸ and CH₂, CH₃, and CH groups³¹ constituting the nonpolar chains. We report the following developments: (1) The mechanism of the functional rotation is elucidated using our theoretical method based on statistical thermodynamics applied to the whole cycle; (2) significant ununiformity is found in the packing efficiency of AcrB; (3) this ununiformity is shown to play critical roles primarily through the solvent-entropy effect; (4) the significance of proton binding and dissociation is argued in a unique manner; (5) the reason why the AcrB proteins have to form a trimer is clarified; and (6) the similarities to the rotation mechanism of F₁-ATPase are discussed from a thermodynamic point of view. The theoretical method, which consists of the integral equation theories for simple and molecular fluids³²⁻³⁷ combined with the morphometric approach,^{38,39} allows us to calculate the solvation entropies of the three pairs of adjacent protomers as well as each protomer by accounting for the polyatomic structures in atomic details. These developments confer a firm physical basis upon the functionally rotating picture previously proposed for AcrB. They also shed new light on the understanding of the mechanism of functional expression by the so-called proton-motivated and ATP-driven proteins in a unified manner.

4.2. Functional Rotation

4.2.1. Effects of Drug on Structural Properties of AcrB

As a feature of AcrA/AcrB/TolC, which is known as “multidrug efflux”, it is capable of handling drugs with diverse properties. The multidrug efflux of AcrA/AcrB/TolC can be realized only when TolC as well as AcrB exhibits this feature. Taking TolC as an example, we have recently shown that the multidrug efflux can be reproduced by a theoretical approach which is focused on the entropic potential field originating from the translational displacement of solvent molecules.⁴⁰⁻⁴² In the bulk solvent, there is no potential field formed which acts on a drug. However, in the solvent confined on the scale of a nanometer, there is a spatial distribution in which the entropic potential becomes largely positive and largely negative with the periodicity of the solvent diameter. The drug motion is substantially influenced by the potential field. The result obtained from the theoretical approach is applicable to AcrB as well. The existence of the so-called multifunctional ligand-binding sites³ need not be assumed. When a drug is inserted into one of the protomers of AcrB, it can be stabilized in a cavity without contacting the protomer inner surface.⁴⁰ The insertion and stabilization are ascribed to the entropic potential field acting on the drug. The switch from insertion to release of the drug is realized by the change in the potential field caused by a structural change of the protomer.^{41,42} It follows that the presence of a drug has only few

effects on the structural properties of the trimer as well as of each protomer. In our view, the functional rotation is controlled by proton binding and dissociation, and drug insertion and release occur in accordance with the structural change of each protomer during the rotation.

4.2.2. Basic Physical Picture of Functional Rotation

The protomers in three different structures are referred to as protomers A (access), B (binding), and E (extrusion). A drug is ready to be accommodated in protomer A, a drug is accommodated in protomer B, and a drug has been extruded from protomer E.³⁻⁵ Here, we discuss one cycle represented by the 120° rotation, (A, B, E)→(B, E, A), in more detail. The aqueous solutions where the proton concentrations are higher and lower are referred to as higher-concentration and lower-concentration sides (i.e., periplasm and cytoplasm), respectively. A single proton is transferred from the higher-concentration side to the lower-concentration one per cycle.¹⁰ When Asp408 with no proton bound is exposed to the higher-concentration side, proton binding occurs because it leads to a decrease in the system free energy. On the other hand, when Asp408 with a proton bound is exposed to the lower-concentration side, proton dissociation occurs because it also leads to a decrease in the system free energy. More details are described in Sec. S2 of the Supplement.

In the initial conformation in the cycle (see Fig. 4.3(a)), a proton is bound only to Asp408 of protomer E.¹⁰ The protomers next to protomer E in the clockwise and counterclockwise directions are referred to as protomers B and A, respectively. In protomer E there are no open routes of proton transfer for both of the two sides,¹⁹ which has been suggested by MD simulation results. Proton binding to Asp408 of protomer E persists, and the next event occurring is proton binding to Asp408 of either of protomers B and A from the higher-concentration side. For the rotation to continue only in one direction, the binding must necessarily occur in the protomer next to protomer E in the same direction (clockwise or counterclockwise) during every cycle. The binding occurs to protomer B as shown in Fig. 4.3(b) (Asp408 of protomer B is exposed to the higher-concentration side). The binding induces structural changes of all three protomers: The resultant protomers are referred to as protomers A', B', and E' (i.e., protomers A, B, and E change to protomers A', B', and E', respectively). Due to the structural changes, the drug is extruded from protomer B', the route of proton transfer in protomer E' becomes open to the lower-concentration side, and the proton dissociates from Asp408 of protomer E'. The dissociation induces structural changes of all three protomers, leading to the conformation shown in Fig. 4.3(c).

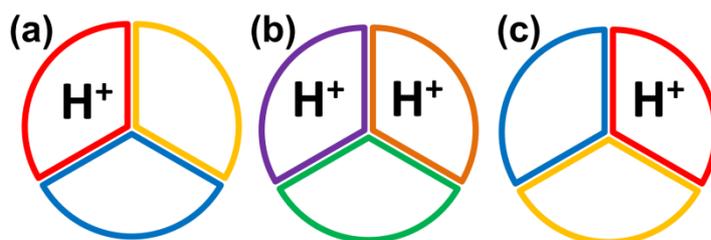


Fig. 4.3. Functional rotation of AcrB per cycle. Protomers A, B, E, A', B', and E' are drawn in blue, yellow, red, green, orange, and purple, respectively. A different color represents a different structure. (a) Initial conformation. (b) Conformation after proton binding to protomer B. (c) Conformation after proton dissociation from protomer E'. The conformations in (a) and in (c) are the same.

4.2.3. Roles of Proton Action during One Cycle

We refer to proton binding from the higher-concentration side and proton dissociation to the lower-concentration one as “proton action”. Our basic concept of energetics of the whole cycle and roles of the proton action are described here.

We first consider the case where a proton is bound to none of the three protomers. In this case (conformation 1 shown in Fig. 4.4(a)), the three protomers should share the same structure. Without the proton action, once the trimer reaches the most stable conformation, this conformation remains unchanged. In the presence of the proton action, when the proton binding sites are exposed to the higher-concentration side, proton binds to one of the protomers (conformation 2 shown in Fig. 4.4(b)) because the binding leads to a decrease in the system free energy. The trimer conformation is then perturbed, being changed toward a different, the most stable one (conformation 3 shown in Fig. 4.4(c) which is the same as Fig. 4.3(a)), that is, reorganization of the trimer conformation takes place. However, further proton binding occurs as described in “Basic Physical Picture” because it leads to a decrease in the system free energy (conformation 4 shown in Fig. 4.4(d)). The trimer conformation is then perturbed, being changed toward a different, the most stable one (conformation 5 shown in Fig. 4.4(e) which is the same as Fig. 4.3(b)). Afterward, proton dissociation occurs as described in “Basic Physical Picture” because it leads to a decrease in the system free energy (conformation 6 shown in Fig. 4.4(f)). The trimer conformation is then perturbed, being changed toward a different, the most stable one: It is conformation 3 shown in Fig. 4.4(g), which is the same as Fig. 4.3(c). One cycle is thus completed. The most stable conformation of the trimer is successively modified by the proton action.

In Fig. 4.4(c), the structures of the two protomers without proton binding are different. The reason for this difference is as follows. Let us consider the two edges of a protomer in

Fig. 4.4(a), each of which forms the interface with an edge of one of the other two protomers. Since the two edges possess different structures, their structural changes caused by proton binding influence the other two protomers in different ways.

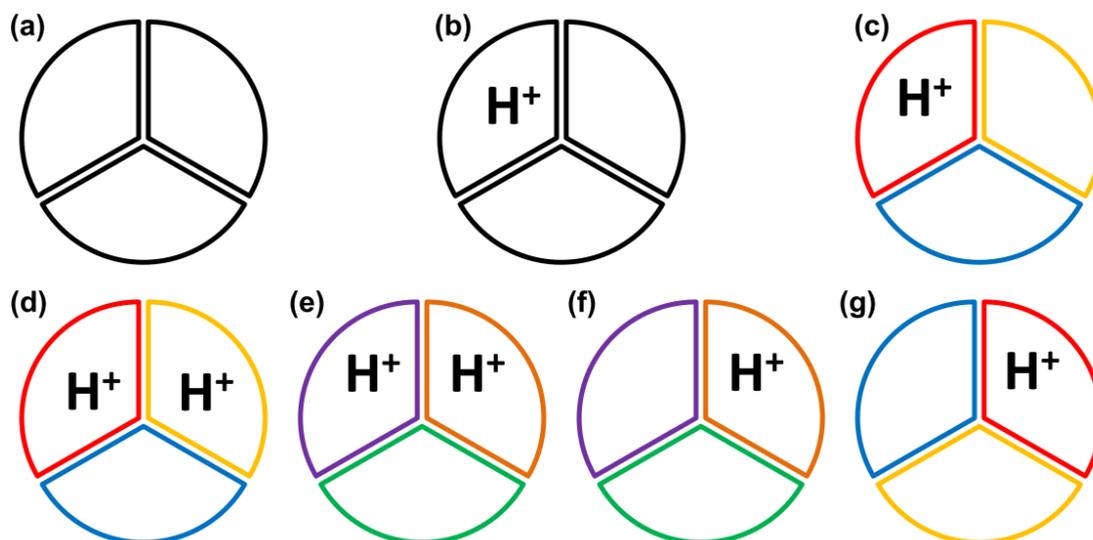


Fig. 4.4. Conformations of AcrB. Protomers A, B, E, A', B', and E' are drawn in blue, yellow, red, green, orange, and purple, respectively. A different color represents a different structure. (a) Conformation 1: No proton is bound to AcrB. (b) Conformation 2: Right after a proton binds to one of the protomers. (c) Conformation 3: After reorganization of AcrB conformation. (d) Conformation 4: Right after a proton binds to protomer B. (e) Conformation 5: After reorganization of AcrB conformation. (f) Conformation 6: Right after a proton dissociates from protomer E'. (g) Conformation 7: After reorganization of AcrB conformation. Conformations 3 and 7 are the same. (c)→(d)→(e)→(f)→(g) forms one cycle. The trimer conformation changes as (c)→(d)→(e)→(f)→(g)=(c)→(d)→(e)→(f)→(g)=(c)→
•••.

4.2.4. Free-Energy Changes during One Cycle

One cycle comprises proton binding, conformational reorganization of the trimer, proton dissociation, and conformational reorganization of the trimer. The system free energy continues to decrease as illustrated in Fig. 4.5. The net decrease in the system free energy in each cycle is only $\sim -8k_B T$ (k_B is the Boltzmann constant, T is the absolute temperature, and $T = 298$ K). $-8k_B T$ results from the increase of $\sim 1.7k_B$ in the system entropy and the decrease of $\sim -6.5k_B T$ in the system energy. The entropic increase arises from the transfer of one proton from the higher-concentration side to the lower-concentration one. The energetic decrease is ascribed to the transfer of a positive charge along the electrostatic-potential gradient. More details are described in Sec. S3 of the Supplement.

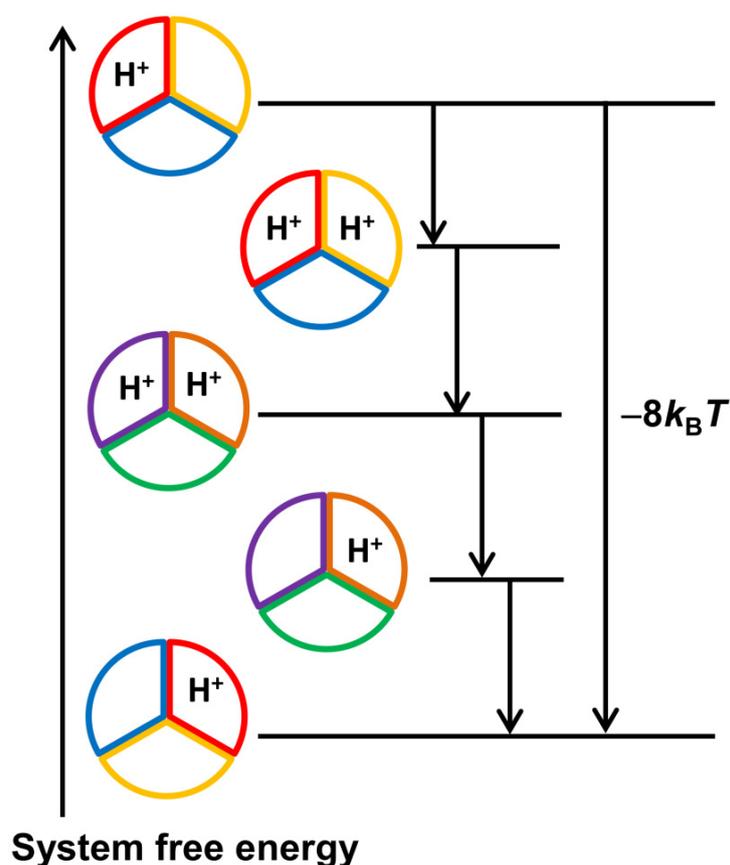


Fig. 4.5. Change in the system free energy accompanying the conformational change of the trimer (c)→(d)→(e)→(f)→(g)=(c). Protomers A, B, E, A', B', and E' are drawn in blue, yellow, red, green, orange, and purple, respectively. A different color represents a different structure. The system free energy changes by $-8k_B T$ (k_B is the Boltzmann constant, T is the absolute temperature, and $T = 298$ K) after this conformational change.

4.2.5. Crucial Importance of Solvent-Entropy Effect

The solvation entropy plus the conformational (protein intramolecular) entropy and the solvation energy plus the conformational energy of the trimer can be considered to remain essentially constant during each cycle irrespective of the number of protons bound (one or two). This is because these thermodynamic quantities are orders of magnitude larger than the values, $1.7k_B$ and $-6.5k_B T$, respectively. How about these thermodynamic quantities of each protomer? As described in “Crystal Structure of AcrB”, there are only slight structural differences among protomers A, B, and E. Let us consider the solvation entropy, conformational entropy, and energetic component defined as the sum of solvation and conformational energies. First, the solvation entropy is quite sensitive to a structural change of a protein.²⁸⁻³¹ Even when the differences in the protein structure are slight, those in the solvation entropy are significant. For example, the root mean square deviation (RMSD) for C_α atoms between protomers A and B, which is as small as 0.211 nm, causes the difference in the solvation entropy reaching $300k_B$ (see later sections). Though the solvation entropy of the trimer remains essentially constant, the solvation entropies of the three protomers are considerably different from one another.

By contrast, the three protomers with very compact structures should share roughly the same value of the conformational entropy. Further, we assume that the energetic component (i.e., the sum of solvation and conformational energies) is significantly less sensitive to a structural change of a protomer than the solvation entropy. The reason for this assumption is as follows. In water, a gain of protein intramolecular electrostatic and van der Waals attractive interactions upon a structural change is almost canceled out by the loss of protein-water electrostatic and van der Waals attractive interactions, and a loss of the former upon a structural change is almost canceled out by the gain of the latter.⁴³ The electrostatic attractive interactions include hydrogen bonds. Thus, the changes in the solvation and conformational energies are rather compensating. The energy decrease due to proton binding to Asp408 in aqueous solution is over an order of magnitude smaller than in vacuum (this is why a proton can readily be dissociated),⁴⁴ also justifying the assumption mentioned above. In nonpolar chains of lipid molecules, a gain of protein intramolecular van der Waals attractive interactions is almost canceled out by the loss of protein-chain van der Waals attractive interactions, and a loss of the former is almost canceled out by the gain of the latter. For electrostatic attractive interactions, however, the cancellation does not occur unlike in water. Nevertheless, the differences in the energetic component among the three protomers should be smaller than those in the solvation entropy. This is because only about 40% (in terms of the number of constituent atoms) of each protomer is within the TM domain.

Hereafter, we argue the functionally rotating mechanism and energetics focused on the solvent-entropy effect. Even if the energetic component makes a significant contribution to energetics, the conclusions drawn are not likely to be altered by the incorporation of this

component: Changes are to be made only for the quantitative descriptions. A more detailed discussion is given in “Effects of Energetic Component”.

4.3. Model and Theoretical Method

4.3.1. Crystal Structure of AcrB

As the structure data of AcrB where the three protomers are in different structures, we adopt the model whose PDB code is 2gif.⁴ It was obtained by X-ray crystallography and free from DARPin often used for crystallization (the inclusion of DARPin might affect the AcrB structure). Further, no drug is bound to AcrB.

In this model, the number of residues is 1032 for protomers A and E, but it is 1044 for protomer B. The 12 residues which are present only in protomer B are removed: Since they are outside the TM domain, the removal is not likely to alter the result of our theoretical calculations. The coordinates of hydrogen atoms are not obtainable by X-ray diffraction. We give hydrogen atoms to the model using the CHARMM biomolecular simulation program⁴⁵ through the Multiscale Modeling Tools in the Structural Biology (MMTSB) program.⁴⁶

The root mean square deviation (RMSD) for C_{α} atoms is 0.211 nm between the structures of protomers A and B, 0.309 nm between those of protomers A and E, and 0.310 nm between those of protomers B and E.⁴ These values are far smaller than, for example, the overall height of AcrB (~ 12 nm).² It follows that there are only slight structural differences among the three protomers (only the structure of protomer E is significantly different, but the difference is still quite small).

We remark that we test the different structure data of AcrB (PDB code: 4dx5)⁴⁷ and verified that the results obtained change only quantitatively and the conclusions drawn are not altered (see “Conclusion”).

4.3.2. Solvent Model

A water molecule is modeled as a hard sphere with diameter $d_s=0.28$ nm in which a point dipole and a point quadrupole of tetrahedral symmetry are embedded.^{33,34} The number density is taken to be that of real water at 298 K and 1 atm: $\rho_s d_s^3=0.7317$. The porter and drug-efflux domains of each protomer are immersed in this model water.

For the TM domain, the entropic effect is attributable primarily to the translational displacement of the CH_2 , CH_3 , and CH groups (the CH_2 groups form the major component) which constitute nonpolar chains of lipid molecules.³¹ The ensemble of groups is regarded as “solvent”. The solvent is formed by neutral hard spheres, and the particle diameter and packing fraction of the solvent are set at those of water. In general, the solvent-entropy effect becomes larger as the solvent diameter decreases or the packing fraction increases. The diameters of the CH_2 , CH_3 , and CH groups are larger than the molecular diameter of water

but their packing fraction is higher than the water value. (Water can be characterized by its relatively low packing fraction due to hydrogen bonds.) These two properties are rather compensating, and the solvent was shown to provide a good model of the membrane environment in our earlier work.³¹ The TM domain of each protomer is immersed in this model solvent.

Though water molecules are also present within the TM domain of each protomer, they are modeled as neutral hard spheres in our treatment mentioned above. In our earlier works, however, we showed the following: In an analysis on the entropic effect arising from the translational displacement of water molecules at normal temperature and pressure, it becomes justifiable to model water as neutral hard spheres as long as the hard-sphere diameter and packing fraction are set at the values pertinent to water.²⁶ When the solvent is water, the solvation entropy comprises the translational and rotational components. As shown in our earlier works,²⁵⁻²⁷ however, the translational component predominates over the rotational one, which is the reason for the justification.

4.3.3. Integral Equation Theories

In the present study, the integral equation theory (IET) based on classical statistical mechanics (see Sec. S4 of the Supplement) is employed for calculating the solvation entropy for a spherical solute. The application to a solute with polyatomic structure is made through the morphometric approach (MA). We use the radial symmetric IET (RSIET)³² for the hard-sphere solvent. For water, the water-water and solute-water potentials and correlations are dependent not only on the distance between centers of water molecules but also on the orientation of each water molecule represented by the three Euler angles. Therefore, we use the angle-dependent IET (ADIET).³³⁻³⁷

We calculated the hydration free energy of a hard-sphere solute with a diameter of 0.28 nm using our water model and the ADIET with the hypernetted-chain (HNC) closure: The value obtained is 3.56 kcal/mol at 300 K, which is in excellent agreement with the values from Monte Carlo simulations for more popular water models: 3.56 kcal/mol at 300 K for TIP4P and 3.65 kcal/mol at 298 K for SPC/E.³⁷

4.3.4. Solvation Entropy

It should be noted that the solvation free energy μ is independent of the protein insertion condition, isobaric or isochoric, but the solvation energy E and the solvation entropy S are not.^{36,48} We consider isochoric condition for the following reasons: (i) It is free from the effects of compression or expansion of the bulk solvent and more suited to physical interpretation of a change in a thermodynamic quantity of solvation; (ii) the structural transition of a protein occurs with the system pressure and volume almost unchanged^{25,49} (the EV of a more compact structure is smaller but the partial molar volume is almost independent

of the compactness); and (iii) it is much more convenient in a theoretical treatment. The solvation energy and energetic component under isochoric condition are almost equal to the solvation enthalpy and enthalpic component under isobaric condition, respectively.

4.3.5. Protein Model

S is fairly insensitive to the solute-water interaction potential as proved in our earlier work.⁵⁰ For example, the three quantities, μ , S , and E , are calculated for a spherical solute with diameter 0.28 nm at $T = 298$ K using the ADIET. For the hard-sphere solute with zero charge, the calculated values are $\mu = 5.95k_B T$, $S = -9.22k_B$, and $E = -3.27k_B T$. When the point charge $-0.5e$ (e is the elementary electric charge) is embedded at its center, the calculated values are $\mu = -32.32k_B T$, $S = -10.11k_B$, and $E = -42.43k_B T$. Therefore, a protein can be modeled as a set of fused hard spheres for calculating its S .

4.3.6. Morphometric Approach

The calculation of S for a large, complex solute molecule like a protein (e.g., a protomer of AcrB) is performed by combining the RSIET or ADIET with the MA.^{38,39} This combination allows us to finish the calculation quite rapidly. In the MA, S is expressed by the linear combination of only four geometric measures of a solute molecule, V_{ex} , A , X , and Y ^{38,39}

$$S/k_B = C_1 V_{\text{ex}} + C_2 A + C_3 X + C_4 Y. \quad (4.1)$$

Here, V_{ex} is the EV, A is the solvent-accessible surface area, and X and Y are the integrated mean and Gaussian curvatures of the accessible surface, respectively. The solvent-accessible surface is the surface that is accessible to the centers of solvent molecules. The volume that is enclosed by this surface is the EV. C_1 is independent of the solute-solvent interaction potential. Though S is influenced by all four terms, $C_1 V_{\text{ex}}$ is the principal term at normal temperature and pressure. In the MA, the solute shape enters S only via the four geometric measures. The four coefficients (C_1 – C_4), which are dependent only on the solvent species and its thermodynamic state, can be determined for simple solute geometries: hard-sphere solutes with various diameters immersed in the solvent. The RSIET or ADIET is employed in calculating the values of S for the hard-sphere solutes. The procedure of calculating S of a protein with a prescribed structure is described in Sec. S5 of the Supplement.

We emphasize that the four geometric measures, V_{ex} , A , X , and Y , are largely dependent on the protein structure. When the solvent is a simple fluid (i.e., the solvent particles interact through a radial-symmetric potential like the Lennard-Jones one), the solvation entropy of a protein with a prescribed structure can be calculated via two routes. One of them is the three-dimensional integral equation theory^{51,52} which is capable of accounting for the structure in the atomic details. The other is the hybrid of the radial symmetric integral

equation theory and the MA. We calculated the solvation entropies for protein G (the number of residues is 56) with a number of different structures via the two routes. The error in the hybrid is only within $\pm 2\%$. Thus, the MA is highly accurate in accounting for the polyatomic structure.

4.4. Solvation Entropy Closely Related to Packing Efficiency

4.4.1. Packing Efficiency of Protomer

We discuss the packing efficiency, which is intimately related to the solvation entropy, by decomposing each protomer into the three portions: those within the TM, porter, and drug-efflux domains, respectively. The portions within the three domains are referred to as portions T, P, and D, respectively (“T”, “P”, and “D” denote the TM, porter, and drug-efflux domains, respectively). Let S_I^M ($I=A, B, E$; $M=T, P, D$) denote the solvation entropy of portion M of protomer I . Smaller $|S_I^M|$ implies that the backbone and side chains of the portion are more efficiently (closely) packed.

4.4.2. Packing Efficiency of Interface between two Protomers

We also analyze the packing efficiency of the interface (i.e., tightness of the interface packing) between two protomers. The decomposition of each protomer into the three portions is again adopted. ΔS_{IJ}^M is defined as

$$\Delta S_{IJ}^M = S_{IJ}^M - (S_I^M + S_J^M), (I, J) = (A, B), (A, E), (B, E), M = T, P, D \quad (4.2)$$

where S_{IJ}^M is the solvation entropy of the pair of “portion M of protomer I and portion M of protomer J ” taken from the trimer, and S_I^M and S_J^M , respectively, are the solvation entropies of “portion M of protomer I ” and “portion M of protomer J ” obtained simply by separating them with their structures kept unchanged. ΔS_{IJ}^M is positive because the overlaps of EVs occur upon the contact of the two portions. There are interfaces between portions within drug-efflux and porter domains, in which case we define ΔS_{IJ}^{PD} as

$$\Delta S_{IJ}^{PD} = S_{IJ}^{PD} - (S_I^P + S_J^D), (I, J) = (A, B), (B, A), (E, B) \quad (4.3)$$

where S_{IJ}^{PD} is the solvation entropy of the pair of “portion P of protomer I and portion D of protomer J ” taken from the trimer.

The very high correlation between ΔS_{IJ}^M or ΔS_{IJ}^{PD} and tightness of the interface packing was demonstrated with the aid of an MD simulation⁵³ in our earlier work²⁹ for F₁-ATPase. Larger ΔS_{IJ}^M or ΔS_{IJ}^{PD} implies that the interface between the two portions is more efficiently packed. Smaller $|S_I^M|$, larger ΔS_{IJ}^M , and larger ΔS_{IJ}^{PD} are more favorable with respect to the

solvent entropy. We emphasize the following:

“For the trimer to maintain its high stability, ΔS_{IJ}^M and $\Delta S_{IJ}^{\text{PD}}$ must be sufficiently large for all pairs of (I, J) and all M .”

4.5. Results and Discussion

4.5.1. Ununiform Packing Structure of AcrB

Each protomer is considered by decomposing it into three portions, those within the TM, porter, and drug-efflux domains. Fig. 4.6(a)–(c) shows the values of $|S_I^M|$ for the three portions of each protomer (the numbers within parentheses) and those *relative to* the smallest value. Within the drug-efflux domain, the differences among the three protomers in the packing efficiency are relatively small. Within the porter domain, protomer B is overall less closely packed than the other two protomers probably due to the presence of an accommodation space for a drug. Within the TM domain, protomer E is overall more closely packed, which is ascribed to proton binding (see “Structural Change of Transmembrane Domain Accompanying Proton Binding”).

Here, we comment on the reliability of the difference between $|S_I^M|$ and $|S_J^M|$ ($I \neq J$). The solvation entropy of a portion is determined primarily by its EV, but most of the EV is taken by the molecular volume of the portion *itself* which has nothing to do with the solvent: The EV is only slightly larger than the molecular volume. Since the molecular volumes of portion M of protomer I and portion M of protomer J share the same value, the solvation entropies of these two portions, which are quite large, are not very different from each other. The difference between the solvation entropies calculated becomes much smaller, but this is because the large contributions to the solvation entropies from the molecular volumes are canceled out. No serious cancellation of significant digits actually occurs. The EV minus the molecular volume and its contribution to the solvation entropy can accurately be calculated for a given structure of a portion. Thus, the difference between $|S_I^M|$ and $|S_J^M|$ is much more reliable than one might expect. A similar argument can be made for any of the solvation-entropy changes presented in the later sections.

Fig. 4.6(d)–(g) shows the values of ΔS_{IJ}^M and $\Delta S_{IJ}^{\text{PD}}$ for the interfaces of two portions in contact. Overall, the interfaces between portions within the drug-efflux domain are tightly packed. This is because part of each portion is penetrating into the adjacent portion (see Sec. S6 of the Supplement). The tight packing through the penetration within the drug-efflux domain makes a substantial contribution to the stability of the trimer. Within each domain, the packing efficiencies of the three interfaces are not significantly different from one another. This is suggestive that impartial, the closest possible packing of all the three interfaces is crucial for maintenance of the high stability of the trimer.

In what follows, we are concerned not with accurate values of $|S_I^M|$, ΔS_{IJ}^M , and ΔS_{IJ}^{PD} but with their approximate values capturing the physical essence. It is therefore assumed for simplicity that any of the portions within the drug-efflux domain preserves the same structure during each cycle due to the tight packing.

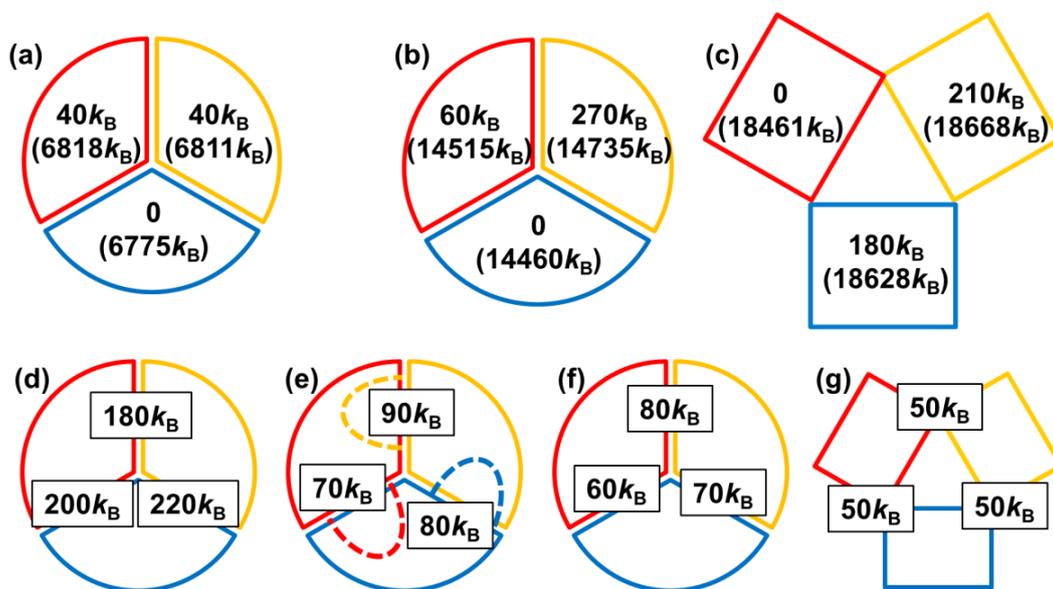


Fig. 4.6. (a)–(c) Values of $|S_I^M|$ for the three portions of each protomer (numbers within parentheses) and those *relative to* the smallest value. The Boltzmann constant is denoted by k_B . Smaller $|S_I^M|$ implies that the backbone and side chains of the portion are more efficiently (closely) packed. Conformation 3 shown in Fig. 4.4(c) is considered. Protomers A, B, and E are drawn in blue, yellow, and red, respectively. A different color represents a different structure. (a) Drug-efflux domain. (b) Porter domain. (c) Transmembrane domain. (d)–(g) Values of ΔS_{IJ}^M and ΔS_{IJ}^{PD} for the interfaces of two portions in contact. They are given within rectangles. Larger ΔS_{IJ}^M or ΔS_{IJ}^{PD} implies that the interface between the two portions is more efficiently packed. Conformation 3 shown in Fig. 4.4(c) is considered. (d) Interfaces between protomers within drug-efflux domain. (e) Interfaces between protomers within drug-efflux and porter domains, respectively (see Fig. S1 in the Supplement). (f) Interfaces between protomers within porter domain. (g) Interfaces between protomers within the transmembrane domain.

4.5.2. Structural Change of Transmembrane Domain Accompanying Proton Binding

A TM domain comprises the 12 α -helices, TM1–TM12. A proton binds to Asp408 in TM4.¹⁰ For protomers B and E, we calculate the efficiency of the interface packing between each pair of α -helices through the solvent-entropy effect. Slight structural differences are observed between any of the 12 α -helices in protomer B and the same helix in protomer E, but they are taken into account in the calculation. The numerical values given below come

primarily from changes in the interface-packing efficiencies, and the contributions from the structural differences mentioned above are much smaller. Since proton binding is present in protomer E while it is absent in protomer B,¹⁰ the change in the efficiency induced by binding can be estimated. The result is summarized in Fig. 4.7 where $-26k_B$, for example, represents that the solvent-entropy gain originating from the interface packing between TM10 and TM4 becomes lower (i.e., the interface packing becomes less efficient) by $26k_B$ upon the binding. Without the binding, Asp408 in TM4 and Lys940 in TM10 are attracting each other through screened, weak electrostatic interaction. Upon binding, the negative charge of Asp408 vanishes in essence and the attractive interaction is lost, with the result that the interface between TM4 and TM10 becomes less closely packed. Instead, the interfaces of TM10-TM5, TM10-TM12, TM4-TM2, and TM4-TM3 pairs undergo closer packing properties: The overall packing efficiency becomes higher. Inversely, upon proton dissociation, the opposite changes occur, leading to lower overall packing efficiency. The structural change of the TM domain accompanying proton binding or dissociation becomes a trigger of conformational reorganization of the trimer.

In the trimer, any two of the three TM domains are in contact with each other only through the packing of TM8 and TM1. It has been suggested that the structural modification of TM8 may play significant roles for drug transport.^{3,4,10} In protomer E, the TM8 helix is extended to the periplasmic side, and the drug entrance is closed. In protomers A and B, by contrast, the periplasmic extension of TM8 is unwound, and the drug entrance is in the vicinity of the unwound loop and opened. We note that this structural modification of TM8 is taken into account in all of our solvation-entropy calculations through the crystal structure⁴ employed. However, we do not know why and how the structural modification occurs. This subject is to be pursued in a future study.

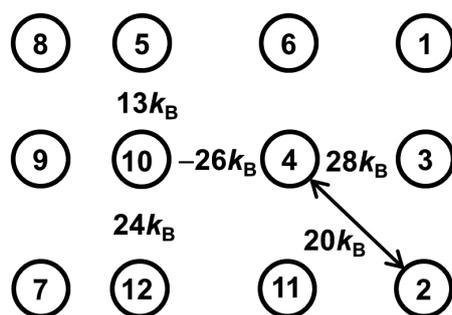


Fig. 4.7. Changes in solvent-entropy gains originating from the interface packing between α -helices within the transmembrane domain. The Boltzmann constant is denoted by k_B . The changes are induced by proton binding. “1–12” represent “TM1–TM12”, respectively.

4.5.3. Conformational Reorganization Induced by Proton Binding or Dissociation

The information on the conformational reorganizations caused by proton binding and proton dissociation is included in the conformational transitions (a)→(c) and (c)→(a), respectively, of Fig. 4.4. It is illustrated in Fig. 4.8. In (a), the three protomers share the same structure. The solvation entropy of the trimer in (a) is approximately equal to that in (c). $|S_I^M|$ and ΔS_{IJ}^M in (c) are calculated by assuming that the closest possible packing properties are achieved at the three interfaces within the porter and TM domains, and they therefore remain unchanged.

The following propositions can then be made:

- I. When a proton binds to a protomer, for the portions within the TM domain (see Fig. 4.8(a)), (i) $|S_I^T|$ of this protomer decreases by $130k_B$; (ii) $|S_I^T|$ of the protomer next to it in the clockwise direction increases by $80k_B$; and (iii) $|S_I^T|$ of the protomer next to it in the counterclockwise direction increases by $50k_B$. The change of (i) is consistent with the increased overall packing efficiency for this protomer within the TM domain described in the last section. For the portions within the porter domain (see Fig. 4.8(b)), (iv) $|S_I^P|$ of this protomer decreases by $50k_B$; (v) $|S_I^P|$ of the protomer next to it in the clockwise direction increases by $160k_B$; and (vi) $|S_I^P|$ of the protomer next to it in the counterclockwise direction decreases by $110k_B$. The change of (i) also induces an increase in overall packing efficiency for this protomer within the porter domain, giving rise to the change of (iv). The change of (v), which is consistent with generation of accommodation space for a drug, induces the decrease described in (vi).
- II. When a proton dissociates from a protomer, for the portions within the TM domain (see Fig. 4.8(a)), (vii) $|S_I^T|$ of this protomer increases by $130k_B$; (viii) $|S_I^T|$ of the protomer next to it in the clockwise direction decreases by $80k_B$; (ix) and $|S_I^T|$ of the protomer next to it in the counterclockwise direction decreases by $50k_B$. The change of (vii) is consistent with the decreased overall packing efficiency for this protomer within the TM domain described in the last section. For the portions within the porter domain (see Fig. 4.8(b)), (vii) $|S_I^P|$ of this protomer increases by $50k_B$; (viii) $|S_I^P|$ of the protomer next to it in the clockwise direction decreases by $160k_B$; (ix) and $|S_I^P|$ of the protomer next to it in the counterclockwise direction increases by $110k_B$. The change of (viii) is consistent with the disappearance of accommodation space for a drug.

The physical meaning of these propositions is discussed in the succeeding three sections.

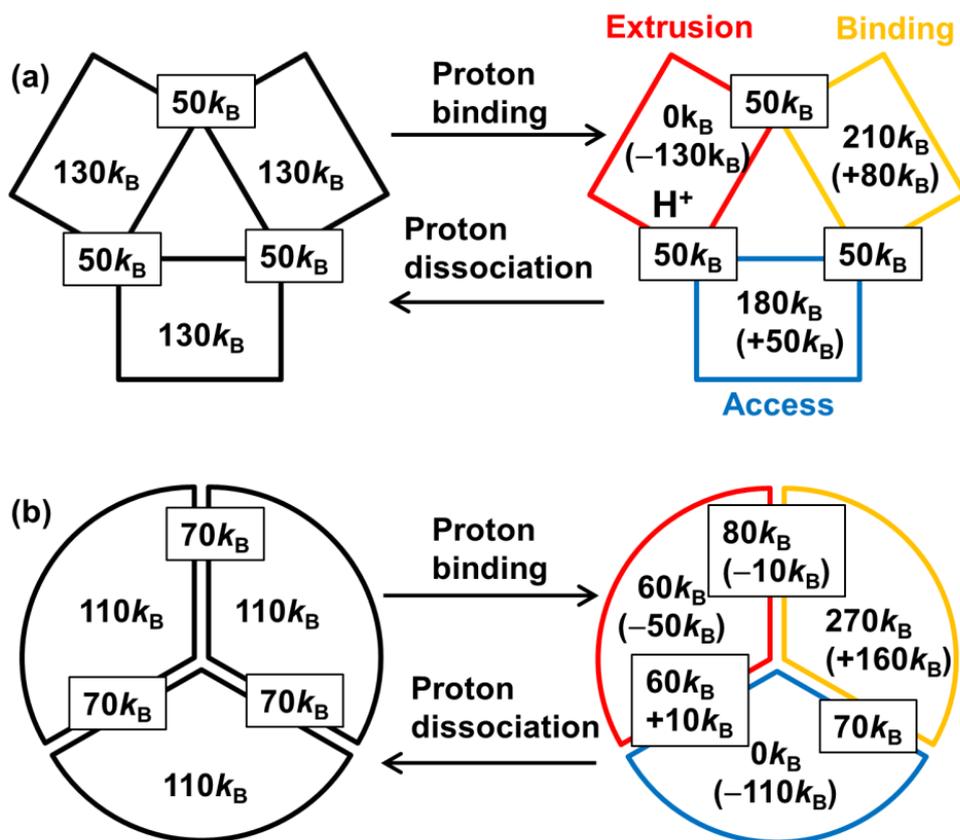


Fig. 4.8. Information on the structural reorganizations caused by proton binding and proton dissociation. The Boltzmann constant is denoted by k_B . Proton binding and proton dissociation correspond to conformational transitions $1 \rightarrow 3$ and $3 \rightarrow 1$, respectively (see Fig. 4.4(a), and (c)). Relative values of $|S_I^M|$ and values of ΔS_{IJ}^M (within rectangles) are given. Numbers within parentheses represent the changes caused by proton binding. Protomers A, B, and E are drawn in blue, yellow, and red, respectively. A different color represents a different structure. (a) Transmembrane domain. (b) Porter domain.

4.5.4. Functional Rotation Induced by Solvent-Entropy Effect

We assume that the proposition made in the last section is always applicable to the conformational reorganization accompanying proton binding or dissociation. The functional rotation can then be interpreted as illustrated in Fig. 4.9. The left, middle, and right conformations in this figure correspond to those of (c), (e), and (g) in Fig. 4.4, respectively. The upper three are within the porter domain, and the lower three are within the TM domain. The numbers denote the values of $|S_I^M|$ for the three portions of each protomer and those within parentheses denote the changes caused by proton binding (middle) or proton

dissociation (right)). An essential point in the functionally rotating mechanism is the following:

“When a proton binds to or dissociates from a protomer, the packing properties of this protomer and its two interfaces are perturbed on the whole in the direction that the solvent entropy is lowered. Hence, the packing properties of the other two protomers are reorganized with the recovery of closely packed interfaces so that the solvent-entropy loss can be compensated and the solvent entropy can be kept almost constant.”

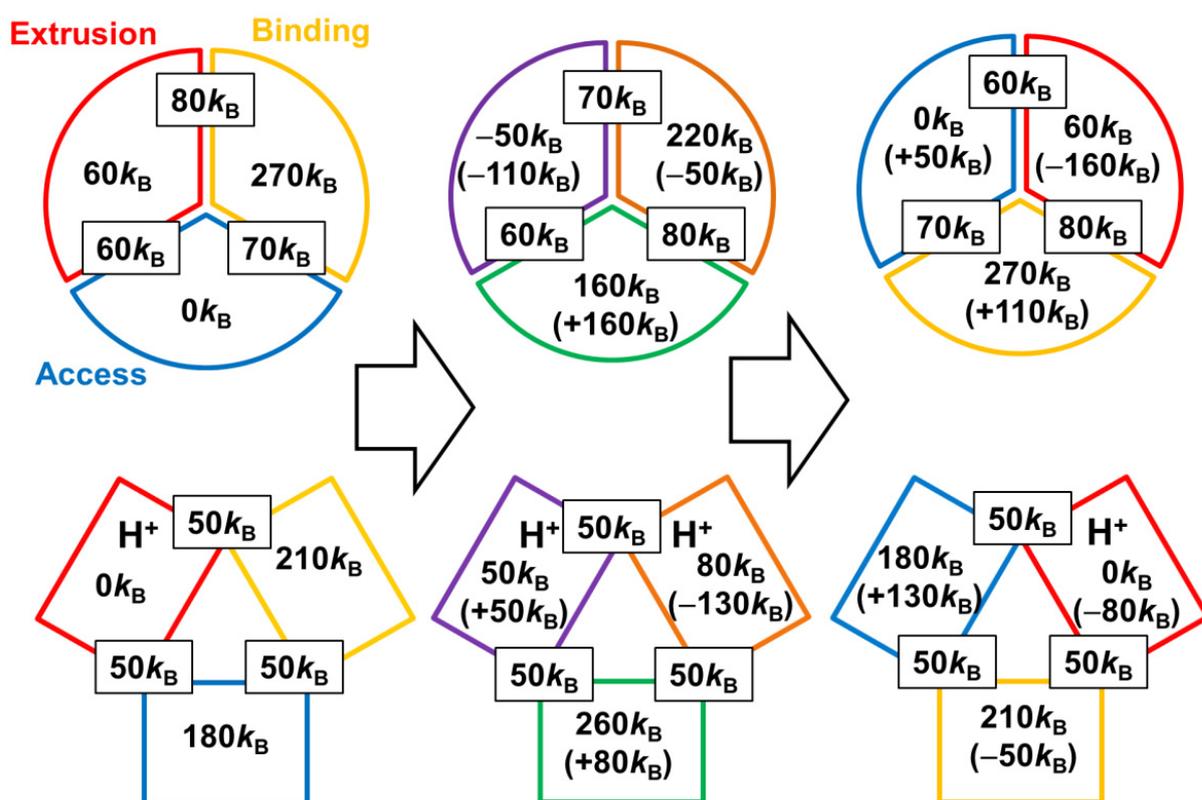


Fig. 4.9. Interpretation of the functional rotation in terms of relative values of $|S_I^M|$ and values of ΔS_{IJ}^M (within rectangles). The Boltzmann constant is denoted by k_B . Smaller $|S_I^M|$ implies that the backbone and side chains of the portion are more efficiently (closely) packed. Larger ΔS_{IJ}^M implies that the interface between the two portions is more efficiently packed. Numbers within parentheses represent the changes arising from each conformational reorganization of the trimer. Left, middle, and right conformations correspond to conformations 3, 5, and 7 (see Fig. 4.4(c), (e), and (g)), respectively. Protomers A, B, E, A', B', and E' are drawn in blue, yellow, red, green, orange, and purple, respectively. A different color represents a different structure. (Top) Porter domain. (Bottom) Transmembrane domain.

4.5.5. Structural Perturbation Caused by Proton Binding

When a proton binds to protomer B as (left)→(middle) in Fig. 4.9, its structure is perturbed in the direction that the solvent entropy is lowered. (The resultant protomer is referred to as protomer B'.) This initially sounds contradictory because $|S_{B'}^T|$ and $|S_{B'}^P|$ are lower than $|S_B^T|$ and $|S_B^P|$, respectively (the backbone and side chains of this protomer become more closely packed). It should be noted, however, that the closer packing followed by reduction of the EV gives rise to looser packing of the B-E and B-A interfaces. Within the TM domain, proton binding tries to induce not only the solvent-entropy gain of $130k_B$ arising from the closer packing of protomer B but also the solvent-entropy loss of “ $50k_B + 50k_B = 100k_B$ ” at most caused by the looser packing of the interfaces (see Fig. 4.6(g)). Within the porter domain, proton binding tries to induce the solvent-entropy loss of “ $90k_B + 80k_B + 80k_B + 70k_B = 320k_B$ ” at most originating from the looser packing of the interfaces (see Fig. 4.6(e), and (f)) as well as the solvent-entropy gain of $50k_B$ brought by the closer packing of protomer B. The net gains and losses are $180k_B$ and $420k_B$, respectively; $420k_B$ is the maximum value, but it is probable that the net loss is larger because even slightly looser interface packing causes a significantly large decrease in solvent entropy. Taken together, proton binding to protomer B tries to induce a significantly large loss of solvent entropy. This loss is compensated with the structural reorganization of the other two protomers primarily to recover the close packing of the interfaces.

4.5.6. Structural Perturbation Caused by Proton Dissociation

When a proton dissociates from protomer E' shown in the middle of Fig. 4.9, the structure of this protomer is perturbed in the direction that the solvent entropy is lowered. This can readily be understood because $|S_{A'}^T|$ and $|S_{A'}^P|$ are higher than $|S_E^T|$ and $|S_E^P|$, respectively (i.e., the backbone and side chains of this protomer become less closely packed). In this case, the two interfaces are already packed with sufficient tightness and not significantly affected by the proton dissociation. The solvent-entropy loss is compensated with the structural reorganization of the other two protomers, which still retains the close packing of the interfaces.

4.5.7. Significance of Trimer Formation

The solvation entropies of protomers A, B, and E are $-36700k_B$, $-37000k_B$, and $-36400k_B$, respectively. An isolated protomer cannot realize the functional structural change “A→B→E→A” by itself. This is because “A→B” and “E→A” would give rise to the solvent-entropy loss of $300k_B$. By forming a trimer, as argued above, the solvent-entropy loss caused by a protomer is always canceled out by the solvent-entropy gain brought by the other two protomers. It is surprising that as a consequence each protomer accomplishes the functional structural change “A→B→E→A” using the free-energy decrease arising from the

transfer of only a single proton, which is as small as $-8k_B T$, with no free-energy barriers. The polyatomic structure of a protomer should possess the feature that the closest overall packing is achievable when three (e.g., neither two nor four) protomers aggregate.

4.5.8. Comparison between AcrB and F₁-ATPase

We refer to the binding of ATP, hydrolysis of ATP into ADP and Pi, and dissociation of ADP and Pi as “ATP action”. Without the ATP action, once the $\alpha_3\beta_3\gamma$ complex (i.e., F₁-ATPase) reaches the most stable conformation, this conformation does not change any more. In our earlier studies,^{29,30} the complex was decomposed into three portions. It was found that in the most stable conformation the packing efficiencies of the three portions are substantially different from one another. Due to the ATP action, the most stable conformation is perturbed and the conformational reorganization of the complex occurs to reach a different, the most stable conformation. More specifically, when the structure of one of the three portions is perturbed in the direction that a solvent-entropy loss is caused, the structures of the other two portions are reorganized to compensate for the loss. In other words, when one of the three portions is less closely packed, the other two portions are more closely packed. The conformational reorganization induces the rotation of the γ subunit (120° rotation per ATP hydrolysis cycle). The rotation is accomplished by the free-energy decrease arising from the hydrolysis of only a single ATP molecule, which is as small as $-15k_B T$.³⁰

4.5.9. Effects of Energetic Component

The basic framework of the functionally rotating mechanism is not altered by incorporation of the energetic component. An additional factor which comes into play is the maintenance of intramolecular hydrogen bonds in the interfaces between two adjacent protomers especially within the TM domain. The total number of bonds is calculated to be three: The corresponding hydrogen-bonding energy is $\sim -40k_B T$. There can be the additional requirement that the total number of intramolecular hydrogen bonds in the trimer be kept almost constant during each cycle. In any case, the essential point can be rewritten as

“When a proton binds to or dissociates from a protomer, the structural properties of this protomer and its two interfaces are perturbed on the whole in the direction that the free energy becomes higher. Hence, the structural properties of the other two protomers are reorganized so that the free-energy increase can be compensated and the free energy can be kept almost constant.”

4.6. Conclusion

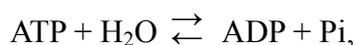
We elucidated the functionally rotating mechanism of AcrB with emphasis on the

solvent-entropy effect. “Solvent entropy” in the present article represents the translational entropy of water molecules and CH₂, CH₃, and CH groups constituting nonpolar chains of lipid molecules. Energetics of the whole cycle has also been developed. The theoretical tools employed are the integral equation theories³²⁻³⁷ and our recently developed morphometric approach.^{38,39} The structure of AcrB is considered in atomic detail. Some of the important results are recapitulated here.

- (1) The functional structural change “A→B→E→A” is essentially infeasible by an isolated protomer. This is because “A→B” and “E→A” would cause the solvent-entropy loss of $300k_B$. By forming a trimer, any solvent-entropy loss caused by a protomer is always canceled out by the solvent-entropy gain brought by the other two protomers. In the trimer each protomer is allowed to accomplish the functional structural change.
- (2) When a proton binds to or dissociates from a protomer, the packing properties of this protomer and its two interfaces are perturbed on the whole in the direction that the solvent entropy is lowered. The packing properties of the other two protomers are then reorganized with the recovery or maintenance of closely packed interfaces, so that the solvent-entropy loss can be compensated and the solvent entropy can be kept almost constant.
- (3) Only slight differences are found among protomers A, B, and E in geometric characteristics of the structures. However, the solvation entropy is quite sensitive to the details of the structure, with the result that the solvent-entropy effect is powerful enough to induce the functional rotation.
- (4) In each cycle, the trimer conformation at the beginning is the same as that at the end. Nevertheless, the transition from the former to the latter occurs, because the system free energy changes by $-8k_B T$. The roles of the proton action (i.e., the so-called proton-motive force) can thus be understood.
- (5) Even when the energetic component is taken into account, the essence of our physical picture is not deteriorated at all (see “Effects of Energetic Component”). The argument is made simply by analyzing not only the solvation entropy but also the energetic component. By forming a trimer, any free-energy increase caused by a protomer is almost canceled out by the free-energy decrease brought by the other two protomers. These free-energy changes originate primarily from the solvent-entropy changes.
- (6) By forming a trimer, each protomer accomplishes the functional structural change using the free-energy decrease arising from the transfer of only a single proton, which is as small as $-8k_B T$, with no free-energy barriers.
- (7) The mechanism of the functional rotation is similar to that of F₁-ATPase in many respects as discussed in “Comparison between AcrB and F₁-ATPase”.

The results obtained are rather robust against adoption of different structural data of AcrB. Even when “PDB code 2gif”⁴ is replaced by “PDB code 4dx5”,⁴⁷ for example, only the numerical values given change (their signs do not change) and the interpretation of the results and the discussion made remain completely the same. The data of 4dx5 was obtained by X-ray crystallography in which designed ankyrin repeat protein (DARPin) was used for crystallization. The DARPin molecules bind to protomers A and B. Further, a drug molecule (minocycline) binds to protomer B. These molecules are removed in the calculations. Our proposition mentioned in “Effects of Drug on Structural Properties of AcrB”, the presence of a drug has only minor effects on structural properties of the trimer as well as of each protomer, has thus been supported. Detailed description is presented in Sec. S7 of the Supplement.

Protein folding, for example, can readily be understood as an irreversible process accompanying a decrease in the system free energy, because the protein conformations before and after folding are different. By contrast, the conformations of F₁-ATPase or those of AcrB before and after rotation are the same. The rotation becomes an irreversible process thanks to the action of ATP or proton. In the case of F₁-ATPase, the chemical reaction,



is not in equilibrium in aqueous solution: The concentrations of ATP and ADP are too high and too low, respectively, and the reaction proceeds in the right direction. It proceeds, however, only with the catalyst, β subunit. ATP binds to the β subunit from the aqueous solution, hydrolysis occurs, and ADP and P_i are released to the aqueous solution. This hydrolysis cycle is accompanied by the conformational reorganization of the $\alpha_3\beta_3\gamma$ complex, which induces the rotation of the γ subunit. Before and after the rotation, the conformations of the $\alpha_3\beta_3\gamma$ complex are the same. Nevertheless, the rotation proceeds because it is an irreversible process accompanying a decrease in the system free energy. In the case of AcrB, the two aqueous solutions (higher-concentration and lower-concentration sides) are not in equilibrium with each other in terms of the proton concentration. Hence, the proton transfer occurs, but only through the binding to AcrB from the higher-concentration side followed by dissociation from AcrB to the lower-concentration side. This transfer induces the functional rotation. Before and after rotation, the conformations of AcrB are the same. Nevertheless, the rotation proceeds because it is an irreversible process accompanying a decrease in the system free energy. Thus, the change in the most stable conformation of F₁-ATPase or AcrB is made possible by the concert with a nonequilibrium process (i.e., the chemical reaction described above or the proton transfer). Moreover, the change is repeatable because the nonequilibrium process occurs successively.

References

- ¹V. Koronakis, A. Sharff, E. Koronakis, B. Luisi, and C. Hughes, *Nature* **405**, 914 (2000).
- ²S. Murakami, R. Nakashima, E. Yamashita, and A. Yamaguchi, *Nature* **419**, 587 (2002).
- ³S. Murakami, R. Nakashima, E. Yamashita, T. Matsumoto, and A. Yamaguchi, *Nature* **443**, 173 (2006).
- ⁴M. A. Seeger, A. Schiefner, T. Eicher, F. Verrey, K. Diederichs, and K. M. Pos, *Science* **313**, 1295 (2006).
- ⁵G. Sennhauser, P. Amstutz, C. Briand, O. Storchenegger, and M. G. Grütter, *PLoS Biol.* **5**, e7 (2007).
- ⁶L. Vaccaro, K. A. Scott, and M. S. P. Sansom, *Biophys. J.* **95**, 5681 (2008).
- ⁷Y. Xu, M. Lee, A. Moeller, S. Song, B.-Y. Yoon, H.-M. Kim, S. Y. Jun, K. Lee, and N.-C. Ha, *J. Biol. Chem.* **286**, 17910 (2011).
- ⁸D. Du, Z. Wang, N. R. James, J. E. Voss, E. Klimont, T. Ohene-Agyei, H. Venter, W. Chiu, and B. F. Luisi, *Nature* **509**, 512 (2014).
- ⁹E.-H. Kim, D. H. Nies, M. M. McEvoy, and C. Rensing, *J. Bacteriol.* **193**, 2381 (2011).
- ¹⁰T. Yamane, S. Murakami, and M. Ikeguchi, *Biochemistry* **52**, 7648 (2013).
- ¹¹J. P. Abrahams, A. G. Leslie, R. Lutter, and J. E. Walker, *Nature* **370**, 621 (1994).
- ¹²H. Noji, R. Yasuda, M. Yoshida, and K. Kinosita, Jr. *Nature* **386**, 299 (1997).
- ¹³S. Furuike, M. D. Hossain, Y. Maki, K. Adachi, T. Suzuki, A. Kohori, H. Itoh, M. Yoshida, and K. Kinosita, Jr. *Science* **319**, 955 (2008).
- ¹⁴S. Toyabe, T. Okamoto, T. Watanabe-Nakayama, H. Taketani, S. Kudo, and E. Muneyuki, *Phys. Rev. Lett.* **104**, 198103 (2010).
- ¹⁵T. Uchihashi, R. Iino, T. Ando, and H. Noji, *Science* **333**, 755 (2011).
- ¹⁶R. Schulz, A. V. Vargiu, F. Collu, U. Kleinekathöfer, and P. Ruggerone, *PLoS Comput. Biol.* **6**, e1000806 (2010).
- ¹⁷A. V. Vargiu, F. Collu, R. Schulz, K. M. Pos, M. Zacharias, U. Kleinekathöfer, and P. Ruggerone, *J. Am. Chem. Soc.* **133**, 10704 (2011).
- ¹⁸R. Schulz, A. V. Vargiu, P. Ruggerone, and U. Kleinekathöfer, *J. Phys. Chem. B* **115**, 8278 (2011).
- ¹⁹N. Fischer, and C. Kandt, *Proteins* **79**, 2871 (2011).
- ²⁰Z. Feng, T. Hou, and Y. Li, *Mol. Biosyst.* **8**, 2699 (2012).
- ²¹A. V. Vargiu and H. Nikaido, *Proc. Natl. Acad. Sci. U. S. A.* **109**, 20637 (2012).
- ²²N. Fischer and C. Kandt, *Biochim. Biophys. Acta* **1828**, 632 (2013).
- ²³T. Imai, N. Miyashita, Y. Sugita, A. Kovalenko, F. Hirata, and A. Kidera, *J. Phys. Chem. B* **115**, 8288 (2011).
- ²⁴Y. Harano and M. Kinoshita, *Biophys. J.* **89**, 2701 (2005).
- ²⁵T. Yoshidome, M. Kinoshita, S. Hirota, N. Baden, and M. Terazima, *J. Chem. Phys.* **128**,

- 225104 (2008).
- ²⁶M. Kinoshita, *Front. Biosci.* **14**, 3419 (2009).
- ²⁷M. Kinoshita, *Int. J. Mol. Sci.* **10**, 1064 (2009).
- ²⁸S. Yasuda, T. Yoshidome, H. Oshima, R. Kodama, Y. Harano, and M. Kinoshita, *J. Chem. Phys.* **132**, 065105 (2010).
- ²⁹T. Yoshidome, Y. Ito, M. Ikeguchi, and M. Kinoshita, *J. Am. Chem. Soc.* **133**, 4030 (2011).
- ³⁰T. Yoshidome, Y. Ito, N. Matubayasi, M. Ikeguchi, and M. Kinoshita, *J. Chem. Phys.* **137**, 035102 (2012).
- ³¹S. Yasuda, H. Oshima, and M. Kinoshita, *J. Chem. Phys.* **137**, 135103 (2012).
- ³²J.-P. Hansen, I. and R. McDonald, *Theory of Simple Liquids*, 3rd ed. Academic Press: London, (2006).
- ³³P. G. Kusalik and G. N. Patey, *J. Chem. Phys.* **88**, 7715 (1988).
- ³⁴P. G. Kusalik and G. N. Patey, *Mol. Phys.* **65**, 1105 (1988).
- ³⁵M. Kinoshita and D. R. Bérard, *J. Comput. Phys.* **124**, 230 (1996).
- ³⁶N. M. Cann and G. N. Patey, *J. Chem. Phys.* **106**, 8165 (1997).
- ³⁷M. Kinoshita, *J. Chem. Phys.* **128**, 024507 (2008).
- ³⁸P.-M. König, R. Roth, and K. R. Mecke, *Phys. Rev. Lett.* **93**, 160601 (2004).
- ³⁹R. Roth, Y. Harano and M. Kinoshita, *Phys. Rev. Lett.* **97**, 078101 (2006).
- ⁴⁰K. Amano and M. Kinoshita, *Chem. Phys. Lett.* **488**, 1 (2010).
- ⁴¹H. Mishima, H. Oshima, S. Yasuda, K.-i. Amano, and M. Kinoshita, *Chem. Phys. Lett.* **561-562**, 159 (2013).
- ⁴²H. Mishima, H. Oshima, S. Yasuda, K. Amano, and M. Kinoshita, *J. Chem. Phys.* **139**, 205102 (2013).
- ⁴³T. Imai, Y. Harano, M. Kinoshita, A. Kovalenko, and F. Hirata, *J. Chem. Phys.* **126**, 225102 (2007).
- ⁴⁴M. Kinoshita and Y. Harano, *Bull. Chem. Soc. Jpn.* **78**, 1431 (2005).
- ⁴⁵B. R. Brooks, R. E. Brucoleri, B. D. Olafson, D. J. States, S. Swaminathan, and M. Karplus, *J. Comput. Chem.* **4**, 187 (1983).
- ⁴⁶M. Feig, J. Karanicolas, and C. L. Brooks, III *J. Mol. Graphics Modell.* **22**, 377 (2004).
- ⁴⁷T. Eicher, H.-J. Cha, M. A. Seeger, L. Brandstätter, J. El-Delik, J. A. Bohnert, W. V. Kern, F. Verrey, M. G. Grütter, K. Diederichs, and K. M. Pos, *Proc. Natl. Acad. Sci. U.S.A.* **109**, 5687 (2012)
- ⁴⁸M. Kinoshita, Y. Harano, and R. Akiyama, *J. Chem. Phys.* **125**, 244504 (2006)
- ⁴⁹S.-H. Chong and S. Ham, *Chem. Phys. Lett.* **535**, 152 (2012)
- ⁵⁰T. Imai, Y. Harano, M. Kinoshita, A. Kovalenko, and F. Hirata, *J. Chem. Phys.* **125**, 024911 (2006)
- ⁵¹M. Ikeguchi and J. Doi, *J. Chem. Phys.* **103**, 5011 (1995).
- ⁵²M. Kinoshita, *J. Chem. Phys.* **116**, 3493 (2002).

⁵³Y. Ito and M. Ikeguchi, *J. Comput. Chem.* **31**, 2175 (2010).

Supplement

S1. Ununiform Packing of a Protein or Protein Complex

As explained in Fig. 4.2 in the main article, it is desired for a protein or a complex of proteins that the backbones and side chains be closely (efficiently) packed, like a three-dimensional jigsaw puzzle.^{1,2} However, this is not always possible, depending on the amino-acid sequence. Even in cases where the overall close packing is not achievable, there are certainly the portions that can closely be packed. It is important to pack such portions preferentially: The other portions cannot participate in the close packing and often become disordered and flexible. For example, the native structure of yeast frataxin³ has a large valley and a tail. Nevertheless, $|S|$ (S is the solvation entropy) of the native structure is almost minimized⁴ because the other portions are closely packed. If an impartial packing was undertaken, the valley and/or the tail could be removed, but the resultant overall packing would become rather loose, causing a larger value of $|S|$.

The preferential packing described above occurs in the case of the AcrB trimer as well. Upon proton binding to one of the three protomers, the structure of this protomer undergoes a significant change with the result that the packing efficiencies of the three protomers become different from one another. Namely, the packing efficiency of the trimer is ununiform. Uniform packing would give rise to lower solvent entropy.

S2. Proton Binding and Dissociation

Let us consider the binding of proton (P) to a protomer (Q) of AcrB, $P+Q \rightarrow PQ$. Denoting the chemical potential of component J ($J=P, Q, PQ$) by μ_J and setting the activity coefficients at unity yields

$$\mu_P = \mu_{P0} + RT \ln(C_P/C_{P0}), \mu_Q = \mu_{Q0} + RT \ln(C_Q/C_{Q0}), \mu_{PQ} = \mu_{PQ0} + RT \ln(C_{PQ}/C_{PQ0}) \quad (S1)$$

where R is the gas constant, T the absolute temperature, C_J (mol/l) the concentration of component J , and the subscript “0” represents the standard state. Setting C_{P0} , C_{Q0} , and C_{PQ0} at 1 mol/l gives

$$\mu_P = \mu_{P0} + RT \ln(C_P), \mu_Q = \mu_{Q0} + RT \ln(C_Q), \mu_{PQ} = \mu_{PQ0} + RT \ln(C_{PQ}). \quad (S2)$$

C_P , C_Q , and C_{PQ} are now dimensionless concentrations. The free-energy change upon the binding ΔG is expressed as

$$\Delta G = \mu_{PQ} - \mu_P - \mu_Q = \mu_{PQ0} - \mu_{P0} - \mu_{Q0} + RT \ln \{C_{PQ}/(C_P C_Q)\}. \quad (\text{S3})$$

Denoting $\mu_{PQ0} - \mu_{P0} - \mu_{Q0}$ by ΔG_0 gives

$$\Delta G = \Delta G_0 + RT \ln \{C_{PQ}/(C_P C_Q)\}. \quad (\text{S4})$$

ΔG_0 represents the free-energy change upon the production of 1 mol of PQ by the binding of 1 mol of P to 1 mol of Q. ΔG_0 is a negative quantity. When the proton concentration C_P is sufficiently low, $RT \ln \{C_{PQ}/(C_P C_Q)\}$ is positive and large enough to make ΔG positive: Proton dissociation, $PQ \rightarrow P+Q$, occurs. Otherwise, ΔG is negative and proton binding, $P+Q \rightarrow PQ$, occurs.

Thus, when the proton binding site is exposed to the lower-concentration side, proton dissociation occurs because it also leads to a decrease in the system free energy.

S3. Free-energy Decrease upon Proton Transfer

The free-energy decrease upon the transfer of a single proton from the higher-concentration side to the lower-concentration one can be estimated for the mitochondrial membrane of the liver⁵ as follows. It comprises the entropic and energetic components.

We first consider the entropic component. The chemical potentials of proton in the higher-concentration and lower-concentration sides, μ_H and μ_L , are expressed as

$$\mu_H = \mu_{H0} + RT \ln(C_H/C_{H0}), \mu_L = \mu_{L0} + RT \ln(C_L/C_{L0}) \quad (\text{S5})$$

where C_H (mol/l) and C_L (mol/l), respectively, denote proton concentrations in the higher-concentration and lower-concentration sides, and the activity coefficients are set at unity. Setting C_{H0} and C_{L0} at 1 mol/l gives

$$\mu_H = \mu_{H0} + RT \ln(C_H), \mu_L = \mu_{L0} + RT \ln(C_L). \quad (\text{S6})$$

C_H and C_L are now dimensionless concentrations. The free-energy change upon the proton transfer from the higher-concentration side to the lower-concentration one, ΔG , is given by

$$\Delta G = \mu_L - \mu_H = \mu_{L0} - \mu_{H0} + RT \ln(C_L/C_H). \quad (\text{S7})$$

The use of $\mu_{L,0} = \mu_{H,0}$ yields

$$\Delta G = RT \ln(C_L/C_H). \quad (\text{S8})$$

The pH-value in the higher-concentration side is higher than that in the lower-concentration one by 0.75:⁵ $\log(C_L/C_H) = -0.75$ and $C_L/C_H = 10^{-0.75}$. Therefore, $\Delta G \sim -1.7RT$. The free-energy decrease upon the transfer of a single proton from the higher-concentration side to the lower-concentration one is $\sim -1.7k_B T$. This comes from the entropic component originating from the entropy of mixing, $\sim 1.7k_B$.

The energetic component can be estimated as follows. The difference between the higher-concentration and lower-concentration sides in the electrostatic potential is 0.168 V.⁵ Since the thickness of the membrane is approximately 8 nm,⁵ there is the electrostatic-potential gradient (i.e., electric field) of 210,000 V/cm. The energy decrease brought by the transfer of a single proton along the electrostatic-potential gradient is calculated to be $\sim -6.5k_B T$.

The free-energy decrease upon the transfer of a single proton from the higher-concentration side to the lower-concentration one is $\sim -8k_B T$. The free-energy decrease is dependent on the organism, experimental condition, and experimental technique employed. However, the value for the *Escherichia coli* cell is not significantly different: The reported values are around $-6k_B T$.⁶ In the main article, we adopt $-8k_B T$ whose entropic and energetic components ($-1.7k_B T$ and $-6.5k_B T$, respectively) are known, and our conclusions are not affected by this adoption at all.

S4. Integral Equation Theory

The integral equation theory (IET) is based on classical statistical mechanics.⁷ In this theory, from the system partition function, various correlation functions are defined, and the basic equations satisfied by them are derived. The many-body correlations are also approximately taken into account. As far as the equilibrium properties are concerned, the results obtained are as detailed as those from a computer simulation. In the case of bulk solvent of one component, for example, the temperature, number density, and interaction potential form the input data. By numerically solving the basic equations to obtain the correlation functions, we can calculate the microscopic structure and thermodynamic quantities.

The IET is applicable to analyses on solvation properties of a spherical solute. Thermodynamic quantities of solvation can readily be calculated from the solute-solvent correlation functions obtained via the hypernetted-chain (HNC) closure. A thermodynamic

quantity of solvation is the change in the thermodynamic quantity upon solute insertion into the solvent.

When the solvent is water, the water-water and solute-water potentials and correlations are dependent not only on the distance between centers of water molecules but also on the orientation of each water molecule represented by the three Euler angles. Therefore, we use the angle-dependent IET (ADIET).⁸⁻¹² In the ADIET the effect of the molecular polarizability is taken into account using the self-consistent mean field (SCMF) theory.^{8,9} At the SCMF level the many-body induced interactions are reduced to pairwise additive potentials involving an effective dipole moment. The effective dipole moment thus determined at 298 K and 1 atm is about 1.42 times larger than the *bare* gas-phase dipole moment.

S5. Morphometric Approach

Morphometry is an academic field which treats the shape of a complex object quantitatively. The calculation of S for a large, complex solute molecule like a protein (e.g., a protomer of AcrB) is performed by combining the radial symmetric IET (RSIET) for simple fluids⁴ or the ADIET for molecular fluids⁸⁻¹² with the morphometric approach (MA).^{13,14} The procedure of calculating S of a protein with a prescribed structure comprises the following four steps.

- (1) S of a hard-sphere solute with diameter d_U is calculated using the RSIET or ADIET with the HNC closure. The values of S are prepared for sufficiently many different values of d_U ($0.6 \leq d_U/d_S \leq 10$).
- (2) The four coefficients are determined by the least square fitting applied to the following equation for hard-sphere solutes (i.e., Eq. (4.1) in the main article applied to hard-sphere solutes):

$$S/k_B = C_1(4\pi R^3/3) + C_2(4\pi R^2) + C_3(4\pi R) + C_4(4\pi), R = (d_U + d_S)/2. \quad (\text{S9})$$

- (3) The four geometric measures of a protein (V_{ex} , A , X , and Y) with a prescribed structure are calculated by means of an extension¹⁴ of Connolly's algorithm.^{15,16}
- (4) S of a protein with a prescribed structure is obtained from Eq. (4.1) in the main article where the four coefficients determined in step (2) are used. Smaller $-S$ implies a closer, more efficient packing of the backbone and side chains.

The high reliability of the ADIET-MA hybrid method has been demonstrated for such subjects as the quantitative reproduction of the experimentally measured changes in

thermodynamic quantities upon apoplastocyanin folding,¹⁷ elucidation of the mechanisms of cold^{4,18} and pressure¹⁹ denaturing of a protein, and proposal of a reliable measure of the thermal stability of a protein.²⁰

S6. Tight Packing of Drug-efflux Domains

Overall, the interfaces between portions within the drug-efflux domain are tightly packed. This is because part of each portion is penetrating into the adjacent portion as illustrated in Fig. S1.

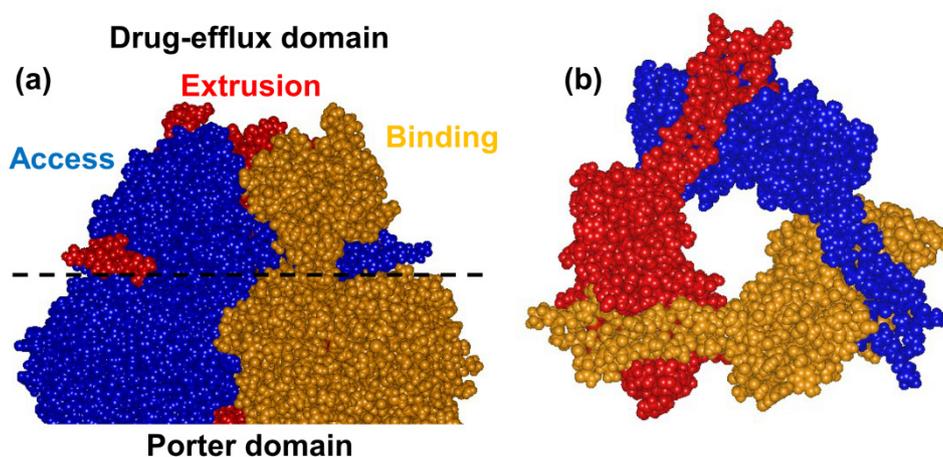


Fig. S1. (a) Space-filled representation of drug-efflux (above the broken line) and porter (below the broken line) domains of protomers A (access) and B (binding). (b) Space-filled representation of drug-efflux domain viewed from porter-domain side. This figure is drawn using the DS visualizer 2.5.

S7. Calculations Using a Different Structural Data of AcrB

The structural data in “PDB Code: 4dx5”²¹ was obtained by X-ray crystallography in which designed ankyrin repeat protein (DARPin) was used for crystallization. The DARPin molecules bind to protomers A and B. Further, a drug molecule (minocycline) binds to protomer B. These molecules are removed from AcrB in the calculations. In this structural model, the number of residues is 1033 for protomers B and E and it is 1044 for protomer A. The 11 residues which are present only in protomer A are removed: Since they are outside the TM domain, the removal is not likely to alter the result of our theoretical calculations. The coordinates of hydrogen atoms cannot be obtained by the X-ray diffraction. We give

hydrogen atoms to the model using the CHARMM biomolecular simulation program²² through the Multi-scale Modeling Tools in Structural Biology (MMTSB) program.²³ The LJ potential energy for AcrB model is positive and large due to unrealistic overlaps of protein atoms. Such overlaps are removed by the minimization of the energy function using the CHARMM and MMTSB programs. The root mean square deviation (RMSD) for C_α atoms is 0.241 nm between the structures of protomers A and B, 0.324 nm between those of protomers A and E, 0.319 nm between those of protomers B and E.

In what follows, the seven sections in “Results and Discussion” are copied and pasted with the modifications of the numerical values underlined.

Ununiform Packing Structure of AcrB

Each protomer is considered by decomposing it into three portions, those within the TM, porter, and drug-efflux domains, respectively. Fig. S2(a)–(c) shows the values of $|S_I^M|$ for the three portions of each protomer (the numbers within parentheses) and those *relative to* the smallest value. Within the drug-efflux domain, the differences among the three protomers in the packing efficiency are relatively small. Within the porter domain, protomer B is overall less closely packed than the other two protomers probably due to the presence of an accommodation space for a drug. Within the TM domain, protomer E is overall more closely packed, which is ascribed to proton binding (see “Structural Change of Transmembrane Domain Accompanying Proton Binding”).

Here, we comment on the reliability of the difference between $|S_I^M|$ and $|S_J^M|$ ($I \neq J$). The solvation entropy of a portion is determined primarily from its EV, but most of the EV is taken by the molecular volume of the portion *itself* which has nothing to do with the solvent: The EV is only slightly larger than the molecular volume. Since the molecular volumes of portion M of protomer I and portion M of protomer J share the same value, the solvation entropies of these two portions are not very different from each other. The difference between the solvation entropies calculated becomes much smaller, but this is because the very large values arising from the molecular volumes are canceled out. No serious cancellation of significant digits actually occurs. The EV minus the molecular volume and its contribution to the solvation entropy can accurately be calculated for a given structure of a portion. Thus, the difference between $|S_I^M|$ and $|S_J^M|$ is much more reliable than one might expect. A similar argument can be made for any of the solvation-entropy changes presented in the later sections.

Fig. S2(d)–(g) shows the values of ΔS_{IJ}^M and ΔS_{IJ}^{PD} for the interfaces of two portions in contact. Overall, the interfaces between portions within the drug-efflux domain are tightly packed. This is because part of each portion is penetrating into the adjacent portion (see Sec. S6 of the Supplement). The tight packing through the penetration within the drug-efflux

domain makes a substantial contribution to the stability of the trimer. Within each domain, the packing efficiencies of the three interfaces are not significantly different from one another. This is suggestive that impartial, the closest possible packing of all the three interfaces is crucial for the maintenance of high stability of the trimer.

In what follows, we are concerned not with accurate values of $|S_I^M|$, ΔS_{IJ}^M , and ΔS_{IJ}^{PD} but with their approximate values capturing the physical essence. It is therefore assumed for simplicity that any of the portions within the drug-efflux domain preserves the same structure during each cycle due to the tight packing.

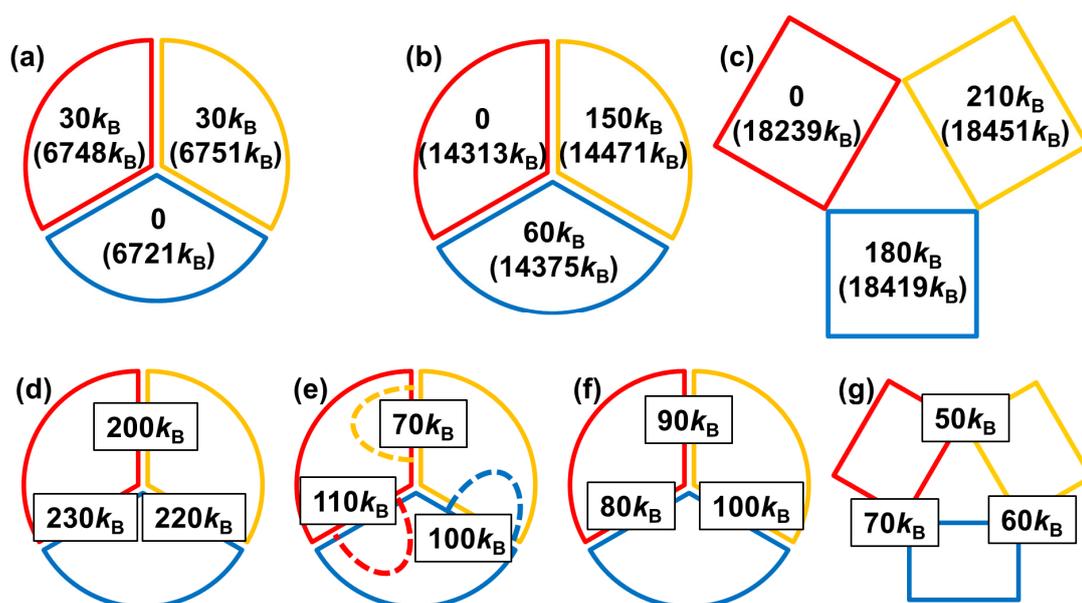


Fig. S2. (a)–(c) Values of $|S_I^M|$ for the three portions of each protomer (the numbers within parentheses) and those *relative to* the smallest value. The Boltzmann constant is denoted by k_B . Smaller $|S_I^M|$ implies that the backbone and side chains of the portion are more efficiently (closely) packed. Conformation 3 shown in Fig. 4.4(c) is considered. Protomers A, B, and E are drawn in blue, yellow, and red, respectively. A different color represents a different structure. (a) Drug-efflux domain. (b) Porter domain. (c) Transmembrane domain. (d)–(g) Values of ΔS_{IJ}^M and ΔS_{IJ}^{PD} for the interfaces of two portions in contact. They are given within rectangles. Larger ΔS_{IJ}^M or ΔS_{IJ}^{PD} implies that the interface between the two portions is more efficiently packed. Conformation 3 shown in Fig. 4.4(c) is considered. (d) Interfaces between protomers within drug-efflux domain. (e) Interfaces between protomers within drug-efflux and porter domains, respectively (see Fig. S1 in the Supplement). (f) Interfaces between protomers within porter domain. (g) Interfaces between protomers within transmembrane domain.

Structural Change of Transmembrane Domain Accompanying Proton Binding

A TM domain comprises the 12 α -helices, TM1–TM12. A proton binds to Asp408 in TM4.²⁴ For protomers B and E, we calculate the efficiency of the interface packing between each pair of α -helices through the solvent-entropy effect. Slight structural differences are observed between any of the 12 α -helices in protomer B and the same helix in protomer E, but they are taken into account in the calculation. The numerical values given below come primarily from changes in the interface-packing efficiencies, and the contributions from the structural differences mentioned above are much smaller. Since proton binding is present in protomer E while it is absent in protomer B,²⁴ the change in the efficiency induced by the binding can be estimated. The result is summarized in Fig. S3 where $-27k_B$, for example, represents that the solvent-entropy gain originating from the interface packing between TM10 and TM4 becomes lower (i.e., the interface packing becomes less efficient) by $27k_B$ upon the binding. Without the binding, Asp408 in TM4 and Lys940 in TM10 are attracting each other through screened, weak electrostatic interaction. Upon the binding, the negative charge of Asp408 vanishes in essence and the attractive interaction is lost, with the result that the interface between TM4 and TM10 becomes less closely packed. Instead, the interfaces of TM10-TM5, TM10-TM12, TM4-TM2, and TM4-TM3 pairs undergo closer packing properties: The overall packing efficiency becomes higher. Inversely, upon proton dissociation, the opposite changes occur, leading to lower overall packing efficiency. The structural change of the TM domain accompanying proton binding or dissociation becomes a trigger of conformational reorganization of the trimer.

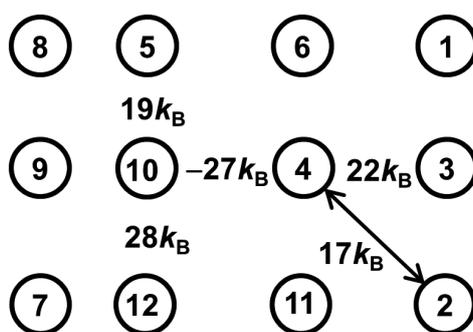


Fig. S3. Changes in solvent-entropy gains originating from the interface packing between α -helices within transmembrane domain. The Boltzmann constant is denoted by k_B . The changes are induced by proton binding. “1–12” represent “TM1–TM12”, respectively.

Conformational Reorganization Induced by Proton Binding or Dissociation.

The information on the conformational reorganizations caused by proton binding and proton dissociation is included in the conformational transitions (a)→(c) and (c)→(a),

respectively, of Fig. 4.4. It is illustrated in Fig. S4. In (a), the three protomers share the same structure. The solvation entropy of the trimer in (a) is approximately equal to that in (c). $|S_I^M|$ and ΔS_{IJ}^M in (c) are calculated by assuming that the closest possible packing properties are achieved at the three interfaces within the porter and TM domains and they therefore remain unchanged.

The following propositions can then be made:

- I. When a proton binds to a protomer, for the portions within the TM domain (see Fig. S4(a)), (i) $|S_I^T|$ of this protomer decreases by $130k_B$; (ii) $|S_I^T|$ of the protomer next to it in the clockwise direction increases by $80k_B$; and (iii) $|S_I^T|$ of the protomer next to it in the counterclockwise direction increases by $50k_B$. The change of (i) is consistent with the increased overall packing efficiency for this protomer within the TM domain described in the last section. For the portions within the porter domain (see Fig. S4(b)), (iv) $|S_I^P|$ of this protomer decreases by $70k_B$; (v) $|S_I^P|$ of the protomer next to it in the clockwise direction increases by $80k_B$; and (vi) $|S_I^P|$ of the protomer next to it in the counterclockwise direction decreases by $10k_B$. The change of (i) also induces an increase in overall packing efficiency for this protomer within the porter domain, giving rise to the change of (iv). The change of (v), which is consistent with the generation of accommodation space for a drug, induces the decrease described in (vi).

- II. When a proton dissociates from a protomer, for the portions within the TM domain (see Fig. S4(a)), (vii) $|S_I^T|$ of this protomer increases by $130k_B$; (viii) $|S_I^T|$ of the protomer next to it in the clockwise direction decreases by $80k_B$; (ix) and $|S_I^T|$ of the protomer next to it in the counterclockwise direction decreases by $50k_B$. The change of (vii) is consistent with the decreased overall packing efficiency for this protomer within the TM domain described in the last section. For the portions within the porter domain (see Fig. S4(b)), (vii) $|S_I^P|$ of this protomer increases by $70k_B$; (viii) $|S_I^P|$ of the protomer next to it in the clockwise direction decreases by $80k_B$; (ix) and $|S_I^P|$ of the protomer next to it in the counterclockwise direction increases by $10k_B$. The change of (viii) is consistent with the disappearance of accommodation space for a drug.

The physical meaning of these propositions is discussed in the succeeding three sections.

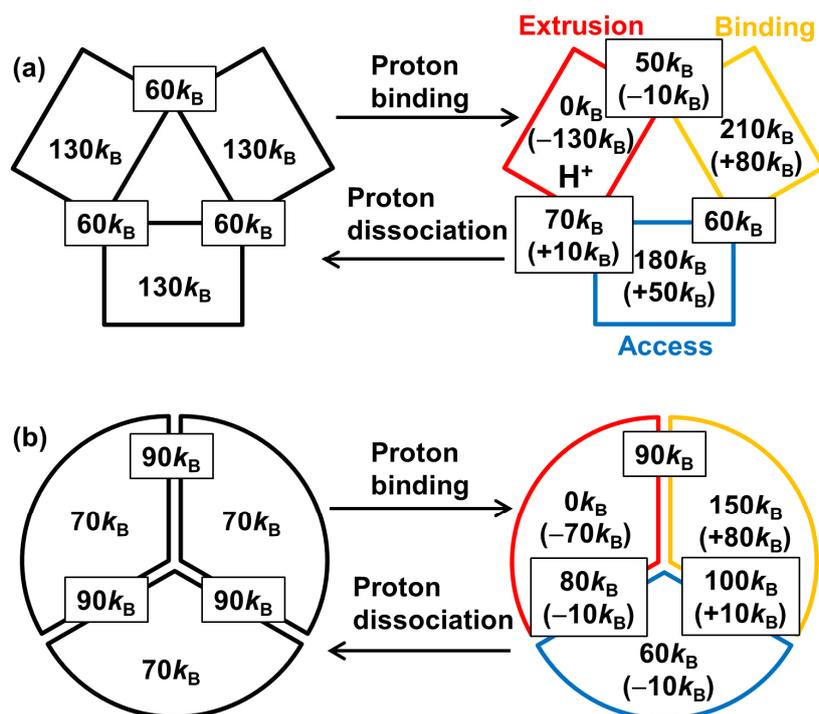


Fig. S4. Information on the structural reorganizations caused by proton binding and proton dissociation. The Boltzmann constant is denoted by k_B . Proton binding and proton dissociation correspond to conformational transitions $1 \rightarrow 3$ and $3 \rightarrow 1$, respectively (see Fig. 4.4(a), (c)). Relative values of $|S_I^M|$ and values of ΔS_{IJ}^M (within rectangles) are given. The numbers within parentheses represent the changes caused by proton binding. Protomers A, B, and E are drawn in blue, yellow, and red, respectively. A different color represents a different structure. (a) Transmembrane domain. (b) Porter domain.

Functional Rotation Induced by Solvent-Entropy Effect

We assume that the proposition made in the last section is always applicable to the conformational reorganization accompanying proton binding or dissociation. The functional rotation can then be interpreted as illustrated in Fig. S5. The left, middle, and right conformations in this figure correspond to those of (c), (e), and (g) in Fig. 4.4, respectively. The upper three are within the porter domain and the lower three are within the TM domain. The numbers denote the values of $|S_I^M|$ for the three portions of each protomer and those within parentheses denote the changes caused by proton binding (middle) or proton dissociation (right)). An essential point in the functionally rotating mechanism is the following:

“When a proton binds to or dissociates from a protomer, the packing properties of this

protomer and its two interfaces are perturbed on the whole in the direction that the solvent entropy is lowered. Hence, the packing properties of the other two protomers are reorganized with the recovery of closely packed interfaces so that the solvent-entropy loss can be compensated and the solvent entropy can be kept almost constant.”

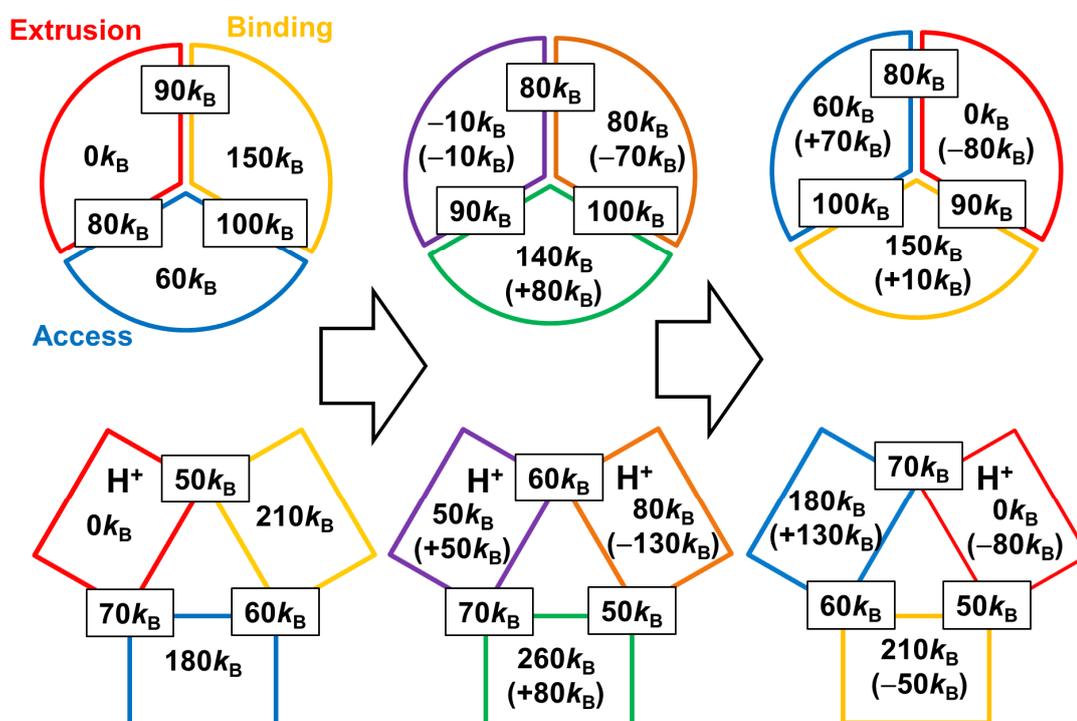


Fig. S5. Interpretation of the functional rotation in terms of relative values of $|S_I^M|$ and values of ΔS_I^M (within rectangles). The Boltzmann constant is denoted by k_B . Smaller $|S_I^M|$ implies that the backbone and side chains of the portion are more efficiently (closely) packed. Larger ΔS_I^M implies that the interface between the two portions is more efficiently packed. The numbers within parentheses represent the changes arising from each conformational reorganization of the trimer. The left, middle, and right conformations (they are referred to as (left), (middle), and (right), respectively) correspond to conformations 3, 5, and 7 (see Fig. 4.4(c), (e), (g)), respectively. Protomers A, B, E, A', B', and E' are drawn in blue, yellow, red, green, orange, and purple, respectively. A different color represents a different structure. The top is for porter domain. The bottom is for transmembrane domain.

Structural Perturbation Caused by Proton Binding

When a proton binds to protomer B as (left)→(middle) in Fig. S5, its structure is perturbed in the direction that the solvent entropy is lowered. (The resultant protomer is

referred to as protomer B'.) This initially sounds contradictory because $|S_{B^T}|$ and $|S_{B^P}|$ are lower than $|S_B^T|$ and $|S_B^P|$, respectively (the backbone and side chains of this protomer become more closely packed). It should be noted, however, that the closer packing followed by reduction of the EV gives rise to looser packing of the B-E and B-A interfaces. Within the TM domain, proton binding tries to induce not only the solvent-entropy gain of $130k_B$ arising from the closer packing of protomer B but also the solvent-entropy loss of " $50k_B + 60k_B = 110k_B$ " at most caused by the looser packing of the interfaces (see Fig. S2(g)). Within the porter domain, proton binding tries to induce the solvent-entropy loss of " $70k_B + 100k_B + 90k_B + 100k_B = 360k_B$ " at most originating from the looser packing of the interfaces (see Fig. S2(e), (f)) as well as the solvent-entropy gain of $70k_B$ brought by the closer packing of protomer B. The net gains and losses are $200k_B$ and $470k_B$, respectively. $470k_B$ is the maximum value, but it is probable that the net loss is larger because even slightly looser interface packing causes a significantly large decrease in solvent entropy. Taken together, proton binding to protomer B tries to induce a significantly large loss of solvent entropy. This loss is compensated with the structural reorganization of the other two protomers primarily to recover the close packing of the interfaces.

Structural Perturbation Caused by Proton Dissociation

When a proton dissociates from protomer E' shown in the middle of Fig. S5, the structure of this protomer is perturbed in the direction that the solvent entropy is lowered. This can readily be understood because $|S_{A^T}|$ and $|S_{A^P}|$ are higher than $|S_{E^T}|$ and $|S_{E^P}|$, respectively (i.e., the backbone and side chains of this protomer become less closely packed). In this case, the two interfaces are already packed with sufficient tightness and not significantly affected by the proton dissociation. The solvent-entropy loss is compensated with the structural reorganization of the other two protomers which still retains the close packing of the interfaces.

Significance of Trimer Formation.

The solvation entropies of protomers A, B, and E are $-36100k_B$, $-36300k_B$, and $-35700k_B$, respectively. An isolated protomer cannot realize the functional structural change " $A \rightarrow B \rightarrow E \rightarrow A$ " by itself. This is because " $A \rightarrow B$ " and " $E \rightarrow A$ " would give rise to the solvent-entropy losses of $200k_B$ and $400k_B$, respectively. By forming a trimer, as argued above, the solvent-entropy loss caused by a protomer is always canceled out by the solvent-entropy gain brought by the other two protomers. It is surprising that as a consequence each protomer accomplishes the functional structural change " $A \rightarrow B \rightarrow E \rightarrow A$ " using the free-energy decrease arising from the transfer of only a single proton, which is as small as $-8k_B T$, with no free-energy barriers. The polyatomic structure of a protomer should possess the feature that the closest overall packing is achievable when three (e.g., neither two

nor four) protomers aggregate.

The first recapitulation in “CONCLUSION” is copied and pasted with the modifications of the numerical values underlined.

- (1) The functional structural change “ $A \rightarrow B \rightarrow E \rightarrow A$ ” is essentially infeasible by an isolated protomer. This is because “ $A \rightarrow B$ ” and “ $E \rightarrow A$ ” would cause the solvent-entropy losses of $200k_B$ and $400k_B$, respectively. By forming a trimer, any solvent-entropy loss caused by a protomer is always canceled out by the solvent-entropy gain brought by the other two protomers. In the trimer each protomer is allowed to accomplish the functional structural change.

References

- ¹M. Kinoshita, Chem. Eng. Sci. **61**, 2150 (2006).
- ²S. Yasuda, T. Yoshidome, H. Oshima, R. Kodama, Y. Harano, and M. Kinoshita, J. Chem. Phys. **132**, 065105 (2010).
- ³S. Adinolfi, M. Nair, A. Politou, E. Bayer, S. Martin, P. Temussi, and A. A. Pastore, Biochemistry **43**, 6511 (2004).
- ⁴H. Oshima, T. Yoshidome, K. Amano, and M. Kinoshita, J. Chem. Phys. **131**, 205102 (2009).
- ⁵D. Voet and J. Voet, G. *Biochemistry*, 3rd ed. John Wiley & Sons: New York, (2004).
- ⁶E. R. Kashket, Ann. Rev. Microbiol. **39**, 219 (1985).
- ⁷J.-P. Hansen and I. R. McDonald, *Theory of Simple Liquids*, 3rd ed. Academic Press: London, (2006).
- ⁸P. G. Kusalik and G. N. Patey, J. Chem. Phys. **88**, 7715 (1988).
- ⁹P. G. Kusalik and G. N. Patey, Mol. Phys. **65**, 1105 (1988).
- ¹⁰M. Kinoshita, D. R. Bérard, J. Comput. Phys. **124**, 230 (1996).
- ¹¹N. M. Cann and G. N. Patey, J. Chem. Phys. **106**, 8165 (1997).
- ¹²M. Kinoshita, J. Chem. Phys. **128**, 024507 (2008).
- ¹³P.-M. König, R. Roth, and K. R. Mecke, Phys. Rev. Lett. **93**, 160601 (2004).
- ¹⁴R. Roth, Y. Harano, and M. Kinoshita, Phys. Rev. Lett. **97**, 078101 (2006).
- ¹⁵M. L. Connolly, J. Appl. Crystallogr. **16**, 548 (1983).
- ¹⁶M. L. Connolly, J. Am. Chem. Soc. **107**, 1118 (1985).
- ¹⁷T. Yoshidome, M. Kinoshita, S. Hirota, N. Baden, and M. Terazima, J. Chem. Phys. **128**, 225104 (2008).
- ¹⁸T. Yoshidome and M. Kinoshita, Phys. Chem. Chem. Phys. **14**, 14554 (2012).
- ¹⁹Y. Harano, T. Yoshidome, and M. Kinoshita, J. Chem. Phys. **129**, 145103 (2008).
- ²⁰K. Oda, R. Kodama, T. Yoshidome, M. Yamanaka, Y. Sambongi, and M. Kinoshita, J. Chem. Phys. **134**, 025101 (2011).

Chapter 5. General conclusion

This study is concerned primarily with the mechanism of the functional expression of AcrA/AcrB/TolC tripartite complex. The entropic effect originating from the translational displacement of water molecules in the system is treated as the pivotal factor. The multidrug efflux transporter, which presents much challenge because it strongly contributes to the multidrug resistance of *E.coli*, inserts a variety of drugs into its cavity and releases them to the outside of the cell. AcrB trimer pumps drugs actively by its functional rotation mechanism in which the protomers take three conformations in turn along the drug efflux process. There are three objectives of this study. The first objective is to show how the apparently opposite events, insertion and release of a solute, successively occur for the same solute in the same system, or equivalently, how the switch from insertion to release is realized. The second one is to show how a variety of drugs independent of their properties (i.e., sizes and affinities with water) can be handled for insertion and release. The third one is to elucidate the mechanism and significance of the functional rotation of AcrB.

Using the original statistical-thermodynamic analyses, Kinoshita *et al.* have shown that the entropic effect arising from the translational displacement of water molecules coexisting with biomolecules play essential roles in the self-assembly and orderly processes in biological systems. We intend to achieve the three objectives mentioned above on the basis of these statistical-thermodynamic analyses and the concept of entropically driven self-assembly and orderly processes. As the key factor, we focus on the potential of mean force (PMF), especially, its entropic component. The entropic component of the PMF is expected to control the insertion and release of a variety of drugs because it is rather insensitive to the affinity of the solute with water. However, it has been suggested at Kinoshita's laboratory that the entropic component of the PMF always inserts the solute into the vessel. Hence, a remaining, important subject is to show that this result is not necessarily true: The solute release can also be accomplished *entropically*. On the other hand, Yoshidome *et al.* at Kinoshita's laboratory have shown that the solvation entropy plays imperative roles in the rotation mechanism of F₁-ATPase. The next subject to be tackled is to demonstrate that the fundamental idea of the rotation mechanism of F₁-ATPase is applicable to the AcrB functional rotation.

Protein folding, for example, can readily be understood as an irreversible process accompanying a decrease in the system free energy, because the protein conformations before and after the folding are different. By contrast, the conformations of F₁-ATPase or those of AcrB before and after the rotation are the same: These are the most stable conformations. Therefore, it is generally believed that a special concept is necessitated to understand the rotation. AcrB and F₁-ATPase are categorized as a proton-motivated protein and an ATP-driven protein, respectively. The prevailing view is the following: The free energy,

which arises from the transfer of protons from the higher concentration side to the lower concentration one or the ATP hydrolysis, is converted to *work*; the work is required for performing the rotation. However, we disagree with this view. In this study, we present a different view and discuss the roles of the so-called proton-motive force (i.e., proton binding and dissociation) and the ATP hydrolysis cycle.

The principal conclusions drawn in each chapter are recapitulated as follows. In Chapter 2, we have investigated insertion and release of a solute into and from a cylindrical vessel possessing an entrance at one end and an exit at the other end for the solute. A water-mediated potential (the potential of mean force, PMF) field acts on the solute near the vessel (the vessel location is fixed). The system constituents are represented by rigid-body models and the potential field is purely entropic in origin (i.e., the PMF possesses only the entropic component). The solute is spontaneously inserted into the vessel due to the spatial distribution of the entropic potential. The inserted solute is confined within a small space in the vessel cavity as long as the vessel geometry is not changed. We show that a continuous variation of the vessel geometry, which changes the spatial distribution of the entropic potential, generates a time-dependent entropic force which acts on the inserted solute *unidirectionally* and strongly. The solute is then released from the exit to the bulk with significantly high speed. This study is an important first step toward elucidating the multidrug efflux mechanism. The result can be applied to such transporters as TolC, AcrB protomer, and ABC transporter.

Since the entropic potential is rather insensitive to the affinities of the solutes with water, the result from Chapter 2 suggests that both hydrophilic and hydrophobic solutes can be handled. The remaining subject is to show that solutes with a variety of sizes can be handled. In Chapter 3, we have shown that solutes with a variety of sizes can be inserted and released using the manner of the continuous variation of the vessel described in Chapter 2. When the inner surface of the vessel is neither strongly hydrophobic nor strongly hydrophilic, the behavior of the solute is determined by the entropic component of the PMF because the entropic component of the PMF predominates over the energetic component. Therefore, we have succeeded in reproducing the multidrug efflux with only weakly hydrophilic or weakly hydrophobic inner surface of the vessel. A possible variation of the vessel geometry is the following: The exit is opened and the diameter of a portion in the entrance side is made smaller by $2d_s$ (d_s is the molecular diameter of water), and the portion is continuously lengthened. There are also other possible variation methods. The variation of the vessel geometry is realistic because the structural geometry of ABC transporter is changed upon the ATP binding. Even for TolC, a peristaltic motion of the TolC tunnel has been observed in a recent molecular dynamics simulation. The theoretical tool employed in Chapters 2 and 3 is the three-dimensional integral equation theory.

In Chapter 4, we have elucidated the functionally rotating mechanism of AcrB with the

emphasis on the solvent-entropy effect. In this case, the solvent is water or hydrocarbon groups of nonpolar chains of lipid molecules. Each is referred to as “solvent molecules”. The conformational properties of AcrB are analyzed in terms of the solvation entropy originating from the translational displacement of solvent molecules. The solvation entropy is calculated using the combination of the integral equation theories and the morphometric approach developed by Kinoshita *et al.* We have found that the packing efficiency of AcrB is significantly ununiform. Although the conformational differences between three protomers are small, the conformational change of an isolated protomer is essentially infeasible because the large solvent-entropy loss (it reaches $+300k_B T$ when converted to the free energy) is unavoidably accompanied. By forming a trimer, any solvent entropy loss caused by a protomer is always canceled out by the solvent-entropy gain brought by the other two protomers. The proton binding to and dissociation from a protomer perturbs the packing properties of the protomer and its two interfaces in the direction that the solvent entropy is lowered. The packing properties of the other two protomers are then reorganized with the recovery or maintenance of closely packed interfaces so that the solvent-entropy loss can be compensated and the solvent entropy can be kept almost constant. Each protomer accomplishes the functional structural change using the free-energy decrease arising from the transfer of only a single proton, which is as small as $-8k_B T$, with no free-energy barriers. The above picture keeps its essence even when “solvent-entropy loss” is replaced with “free-energy increase”. The rotation is repeatable because the irreversible process, proton transfer, occurs successively. The similarity between the functional rotation of AcrB and the unidirectional rotation mechanism of F_1 -ATPase are also discussed in detail. The achievement in Chapter 3 also deepens the understanding of the rotation mechanism of F_1 -ATPase suggested by Yoshidome *et al.* at Kinoshita’s laboratory. Although the conformations of F_1 -ATPase or AcrB trimer before and after the rotation are the same, the functional rotation occurs because the rotation is coupled with the free energy decrease caused by the irreversible process, the ATP hydrolysis or the proton transfer. In other words, the change in the most stable conformation of F_1 -ATPase or AcrB is made possible by the concert with the irreversible process. In chapter 4, the roles of the so-called proton motive force and the ATP hydrolysis are interpreted in a new manner.

Though we have succeeded in reproducing the multidrug efflux, our result is indicative that even the useful solutes such as sugars and amino acids may be extruded from the cell. We believe that this problem can be solved by introducing the energetic component of the PMF as well. In the multidrug efflux, it is postulated that the vessel inner surface is either weakly hydrophobic or weakly hydrophilic is essential, and we should adhere to this postulation. In general, the useful solutes are hydrophilic whereas toxic solutes are hydrophobic. When the vessel inner surface is kept weakly hydrophobic, the cavity within the vessel presents rather unfavorable environment for water with the result that the average

number density of water within the vessel cavity becomes lower than that in the bulk. It follows that a hydrophilic solute prefers to be hydrated in the bulk water, and the insertion of the solute is prevented. This type of interesting behavior has been unveiled by our recent theoretical analysis. We intend to carry out an extensive, systematic analysis based on this concept. We may be able to obtain valuable information on the design of drugs which are not excluded, leading to the solution of the serious medical problem.

Taken together, the physical basis of the multidrug efflux has been uncovered and the mechanism of the functional rotation has been elucidated. These accomplishments are based on the concept of the solvent-entropy effect originating from the translational displacement of solvent molecules. It is worth noting that the hydration properties of solutes in the water confined on the scale of a nanometer are shown to be substantially different from those in the bulk water. This difference plays the pivotal role in the functioning of the proton-motivated and ATP-driven proteins. This thesis provides a significant progress toward understanding the mechanism of functional expression by the proton-motivated and ATP-driven proteins in a unified manner. As an eminent example, AcrB is shown to achieve a high function in the homotrimer by utilizing only remarkably low energy, which features a highly efficient energy conversion process in biological systems.

Appendix. Characterization of Experimentally Determined Native-Structure Models of a Protein Using Energetic and Entropic Components of Free-Energy Function

A1. Introduction

The X-ray crystallography and nuclear magnetic resonance (NMR) are popular and important means of investigating structural properties of biomolecules such as proteins. A protein folds into its unique native structure (NS) in aqueous solution under the physiological condition. Understanding the NS at the atomic level is essential because it is closely related to the protein function. The NS models are represented by the coordinates of the constituent atoms and registered in Protein Data Bank (PDB). The X-ray crystallography,¹⁻³ which has been employed for many years, can be performed in an almost established manner once the solvent condition leading to protein crystallization is found, while the NMR is still in a developing stage. A variety of solution-NMR approaches⁴⁻⁶ have been devised to make the NMR applicable even to a large protein by improving its relaxation properties. Solid-state NMR spectroscopy^{7,8} has also been proposed as a useful alternative to solution NMR. The NS models are constructed by a structure calculation upon which the structural information experimentally obtained as a set of constraints is imposed.⁴⁻⁸ Typical constraints are the nuclear Overhauser effect (NOE), residual dipolar coupling (RDC), hydrogen bonding, and dihedral angle restraints. The structure calculation usually yields a number of candidate models which can be significantly different from one another. Thus, the NS models are substantially influenced by the experimental technique and the structure calculation employed.

A great advantage of solution NMR is that it provides insightful information on dynamical properties associated with backbone and side-chain mobilities which are crucially important in discussing the protein function. To elucidate such properties, the native state of a protein needs to be represented not as a single structure but as an ensemble of structures.⁹⁻¹³ Such a structural ensemble is usually constructed by employing the ensemble refinement protocol^{9-11,13} started from an NS model determined by the X-ray crystallography or solution NMR. The refinement is performed in a variety of manners using a computer simulation with all-atom potentials (e.g., an ensemble molecular dynamics (MD) simulation) or a method based on geometric restrictions, which is restrained by NOEs, RDCs, and order parameters S^2 obtained in NMR experiments. The order parameters contain atomic-detailed information about the amplitude of molecular motion. The restraints are enforced not on a single structure

but on the average calculated over an ensemble of structures. The NS models thus obtained as a structural ensemble represent accessible structures of a protein in aqueous solution rather than candidate models for a fixed NS. The ensembles are diversified by the details of the ensemble refinement protocol and the restraints effectively taken into consideration.

It is not rare that there are many NS models for the same protein: candidate models for a fixed NS and ensembles which represent fluctuating structures of the native state in aqueous solution. Nevertheless, the characteristics of all those NS models are often ambiguous. It is strongly desired that a reliable method, which is based on *energetics* of a protein accounting for its hydration thermodynamics, be developed for characterizing the NS models. In the present study, we show how to characterize the NS models using our recently developed free-energy function¹⁴⁻¹⁶ F and its energetic and entropic components, A and S . F has been demonstrated to be far superior to any of the previously reported functions in terms of the performance of discriminating the native fold from misfolded decoys. Ubiquitin with 76 residues is chosen as an example protein. It is known to be stable over a wide range of pH and temperature.¹ It has frequently been adopted as a benchmark protein for which a new NMR approach is illustrated.⁴⁻⁸ Thus, a comprehensive range of high-quality X-ray and NMR data of ubiquitin are available. For these reasons, ubiquitin is best suited to the present study.

We consider three X-ray models (PDB codes: 1ubq,¹ 1ubi,² and 3n32³) and a number of models in three solution NMR data (PDB codes: 1d3z,⁴ 1g6j,⁵ and 2klg⁶) and in two solid-state NMR data (PDB codes: 2jzz⁷ and 2l3z⁸) which were obtained as candidate models for a fixed NS. The model sets in 1d3z, 1g6j, 2klg, 2jzz, and 2l3z are categorized as “model sets of type 1”. Five ensembles which were constructed for representing the structural variability of the native state (PDB codes: 1xqq,⁹ 2nr2,¹⁰ 2k39,¹¹ 2kn5,¹² and 2kox¹³) are also considered. The model sets in 1xqq, 2nr2, 2k39, 2kn5, and 2kox are categorized as “model sets of type 2”. The models belonging to the two types of model sets are referred to as “NMR models” because they were constructed such that the constraints obtained in NMR experiments were satisfied. It should be emphasized that only the models for isolated monomers are selected in the present study.

We propose a set of figures where F , A , S , and additional parameters judiciously defined are plotted. The X-ray models resemble one another with respect to the structural features, F , A , and S , whereas the NMR models exhibit diverse properties. In general, the X-ray models give lower values of F than the NMR models. There is an overall trend that as a model deviates more from the reference model (one of the three X-ray models), its F becomes higher. Model sets of type 1 and those of type 2, respectively, exhibit two apparently different properties in terms of the correlation between the deviation and F . On the whole, the average value of F tends to become lower as the structures in a model set are better converged. Concerning the correlation between the structural convergence and F_{av} (the subscript “av” denotes the average value in each model set), two apparently different properties are

exhibited by model sets of type 1 and those of type 2, respectively. There are model sets which depart from the correlation observed, but the physical reasoning for the departure can be made, thus uncovering their characteristics. It is argued that F becomes lower with the increase in the total amount of constraints effectively incorporated in the structure calculation or the ensemble refinement protocol. Further, by examining A and S , we can clarify structural properties of each model set with respect to intramolecular hydrogen bonding and packing efficiency of the backbone and side chains.

We are successful in characterizing the NS models of ubiquitin, finding the features of the X-ray and NMR models, and clarifying the overall properties of the structures in a particular model set. It should be emphasized that the results described above are achievable only by a free-energy function like ours capturing essential physics of the structural stability of a protein in aqueous solution. The characterization method thus developed is expected to be useful for the following applications: the evaluation of a set of NS models of ubiquitin or any other protein determined via a new NMR approach by comparing them to the models which are already available; and refinement of an NMR model by rectifying its weak points found. Further, our free-energy function is well suited not only to the selection of the best model from among many candidate NMR models but also to the original construction of the best candidate model or a good structural ensemble for the native state, on the basis of the experimentally obtained constraints such as NOEs and RDCs.

A2 Free-Energy Function

A2.1. Definition.

We treat a number of different structures of a protein. Our free-energy function F is expressed for the protein with a prescribed structure by¹⁴⁻¹⁶

$$F=(E_I+\mu)/(k_B T_0), T_0=298\text{ K}, \quad (\text{A1})$$

where E_I is the protein intramolecular energy, μ is the hydration free energy (i.e., excess chemical potential) that is the most important thermodynamic quantity of protein hydration, and k_B is the Boltzmann constant. We note that μ is the same irrespective of the protein insertion condition: isobaric or isochoric,¹⁷ and we consider the isochoric condition that is much more convenient in a theoretical treatment. Using the relation,

$$\mu=E-TS, \quad (\text{A2})$$

where E is the hydration energy, S is the hydration entropy, and T is the absolute temperature

and defining A by

$$A=E_1+E, \quad (\text{A3})$$

we obtain

$$F=(A-TS)/(k_B T_0), \quad T_0=298 \text{ K} \quad (\text{A4})$$

where T is set at T_0 in the present study. A is referred to as the total dehydration penalty and S represents a water-entropy loss upon the protein insertion. Both A and $-S$ are positive quantities. A , S , and F are substantially dependent on the protein structure. The procedures of calculating the entropic component S and the energetic component A are briefly described below (more details are described in our earlier publications¹⁴⁻¹⁶).

A2.2. Entropic Component: Hydration Entropy.

Unlike the previously reported functions where water was regarded as a dielectric continuum, a molecular model is employed in the calculation of S , which would require a heavy burden. However, the calculation is finished quite rapidly by combining the angle-dependent integral equation theory¹⁷⁻²³ (ADIET), a statistical-mechanical theory for molecular liquids, and the morphometric approach (MA).²⁴⁻²⁶

We employ a multipolar model for water.¹⁹ A water is modeled as a hard sphere with diameter $d_s=0.28$ nm in which a point dipole and a point quadrupole of tetrahedral symmetry are embedded. In the ADIET the effect of the molecular polarizability is taken into account using the self-consistent mean field (SCMF) theory.^{18,19} At the SCMF level the many-body induced interactions are reduced to pairwise additive potentials involving an effective dipole moment. As proved in our earlier work,²³ the ADIET predominates over the reference interaction site model (RISM) and related theories.²⁷⁻²⁹ The hydration entropy under the isochoric condition is fairly insensitive to the solute-water interaction potential as proved in our earlier works.^{30,31} For example, the three quantities, μ , S , and E , under the isochoric condition are calculated for a spherical solute with diameter 0.28 nm using the ADIET¹⁷⁻²³ combined with the multipolar water model.¹⁹ For the hard-sphere solute with zero charge, the calculated values are $\mu=5.95k_B T$, $S=-9.22k_B$, and $E=-3.27k_B T$. When the point charge $-0.5e$ (e is the electronic charge) is embedded at its center, the calculated values are $\mu=-32.32k_B T$, $S=-10.11k_B$, and $E=-42.43k_B T$. Thus, S is fairly insensitive to the solute-water interaction potential while μ and E are largely influenced by it. Therefore, a protein can be modeled as a set of fused hard spheres just for calculating its hydration entropy. (The hydration energy, which is influenced by the protein-water interaction potential, is separately treated in the total dehydration penalty.)

The idea of the morphometric approach is to express a hydration quantity such as S by the linear combination of only four geometric measures of a solute molecule.^{25,26}

$$S/k_B = C_1 V_{\text{ex}} + C_2 A + C_3 X + C_4 Y. \quad (\text{A5})$$

Here, V_{ex} is the excluded volume, A is the water-accessible surface area, and X and Y are the integrated mean and Gaussian curvatures of the accessible surface, respectively. The water-accessible surface is the surface that is accessible to the centers of water molecules. The volume that is enclosed by this surface is the excluded volume. Though $C_1 V_{\text{ex}}$ is the principal term, the other three terms also influence S . In the approach, the solute shape enters S only via the four geometric measures. Therefore, the four coefficients (C_1 – C_4) can be determined in simple geometries. They are calculated from the values of S for hard-sphere solutes with various diameters immersed in our model water. The ADIET¹⁷⁻²³ is employed in the calculation.

The procedure of calculating S of a protein with a prescribed structure comprises the following four steps.

- (1) S of a hard-sphere solute with diameter d_U is calculated using the ADIET. The values of S are prepared for sufficiently many different values of d_U ($0 \leq d_U \leq 10d_S$; changing $10d_S$ to $30d_S$, for example, leads to no changes in the result from step (2).)
- (2) The four coefficients are determined by the least square fitting applied to the following equation for hard-sphere solutes (i.e., Eq. (A5) applied to hard-sphere solutes):

$$S/k_B = C_1(4\pi R^3/3) + C_2(4\pi R^2) + C_3(4\pi R) + C_4(4\pi), \quad R = (d_U + d_S)/2. \quad (\text{A6})$$

- (3) The four geometric measures of a protein (V_{ex} , A , X , and Y) with a prescribed structure are calculated by means of an extension²⁵ of Connolly's algorithm.^{32,33} The x - y - z coordinates of the protein atoms used as part of the input data to account for the polyatomic structure at the atomic level. The diameter of each atom is set at the sigma value of the Lennard-Jones (LJ) potential parameters which are taken from the CHARMM22.³⁴
- (4) S of a protein with a prescribed structure is obtained from Eq. (A5) in which the four coefficients determined in step (3) are used. It should be emphasized that the computation time required for step (4) is only ~ 0.1 sec on our workstation.

The high reliability of the hybrid of ADIET and MA in calculating S has been demonstrated in the following examples: quantitative reproduction of the experimentally

measured changes in thermodynamic quantities upon apoplastocyanin folding;³⁵ elucidation of the molecular mechanisms of pressure³⁶ and cold^{37,38} denaturing of proteins; and proposal of a reliable measure of the thermal stability of proteins.^{39,40}

A2.3. Energetic Component: Total Dehydration Penalty.

The energetic component is calculated in accordance with a simple manner which still accounts for the physically important factors (the time required for this calculation is only ~ 0.1 sec per protein structure on our workstation). A defined by Eq. (A3) is calculated by choosing a fully extended structure as the standard one. The fully extended structure possesses the maximum number of hydrogen bonds with water molecules but no intramolecular hydrogen bonds. “ E_I+E ” corresponds to the total dehydration penalty A occurring upon the transition to a more compact structure. Compared to the fully extended structure with $A=0$, in a more compact structure some donors and acceptors (e.g., N and O, respectively) are buried in the interior after the break of hydrogen bonds with water molecules (CO \cdots W, NH \cdots W, etc.). There is no problem if the intramolecular hydrogen bonds (CO \cdots HN, etc.) are formed. However, such hydrogen bonds are not always formed, giving rise to an energetic penalty.

Our procedure of calculating A can be summarized as follows. When a donor and an acceptor are buried in the interior after the break of hydrogen bonds with water molecules, if they form an intramolecular hydrogen bond, we impose no penalty. On the other hand, when a donor or an acceptor is buried with no intramolecular hydrogen bond formed, we impose the penalty of $7k_B T_0$ ($T_0=298$ K). The value, $7k_B T_0$, is based on the energy-decrease of $-14k_B T_0$ arising from hydrogen-bond formation between two formamide molecules in a nonpolar liquid.⁴¹ We have to determine whether or not each of the donors and acceptors is buried. The water-accessible surface area is calculated for each of them by means of Connolly’s algorithm^{32,33}. If it is smaller than A_0 , the donor or acceptor is considered buried. A_0 is set at 0.001 \AA^2 . We examine all the donors and acceptors for backbone-backbone, backbone-side chain, and side chain-side chain intramolecular hydrogen bonds and determine if each of the donors and acceptors is buried or not.

The energetic component is not considered for nonpolar groups. This is justifiable because the break of hydrogen bonds with water molecules, when they are not compensated by the intramolecular hydrogen bonds, should be more serious and form a principal component of the total dehydration penalty. The torsion energy is not considered, either. The structures to be treated share the property that the torsion energy is reasonably low (i.e., only the structures with sufficiently low torsion energies are chosen), and the difference between two structures in the torsion energy makes no essential contribution to the difference in the energetic component.

A2.4. Performance of Discriminating Native Fold from Misfolded Decoys.

We have examined the performance of our free-energy function F in discriminating the native fold from a number of misfolded decoys.^{15,16} The examination is carried out for a total of 133 proteins in 8 decoy sets. F is shown to be far superior to any of the previously reported functions. When the NS model is determined by the X-ray crystallography, the discrimination is always successful. In the case of NMR models, as long as a sufficiently good NS model is included in the candidate models, the discrimination is accomplished with 100% accuracy. Thus, the approximations employed in calculating F can be justified by this success. F captures the features of the NS of a protein such that it is optimized in terms of the sum of the hydration entropy and the total dehydration penalty.

A3. Native-Structure Models Considered

We consider three X-ray models (PDB codes: 1ubq,¹ 1ubi,² and 3n32³) and a number of models in three solution NMR data (PDB codes: 1d3z,⁴ 1g6j,⁵ and 2klg⁶) and two solid-state NMR data (PDB codes: 2jzz⁷ and 2l3z⁸) which were obtained as candidate models for a fixed NS. The numbers of the models registered in 1d3z, 1g6j, 2klg, 2jzz, and 2l3z are 10, 32, 20, 20, and 20, respectively. We also consider five sets of models constructed for representing the native state which comprises an ensemble of structures (PDB codes: 1xqq,⁹ 2nr2,¹⁰ 2k39,¹¹ 2kn5,¹² and 2kox¹³). A total of 128, 144, 116, 50, and 640 models are registered in 1xqq, 2nr2, 2k39, 2kn5, and 2kox, respectively. The model sets in 1d3z, 1g6j, 2klg, 2jzz, and 2l3z are categorized as “model sets of type 1”. Those in 1xqq, 2nr2, 2k39, 2kn5, and 2kox are categorized as “model sets of type 2”.

A3.1. Three Models Obtained from X-ray Crystallographic Experiments.

1ubq is a model of natural human ubiquitin (hUb) refined at 1.8 Å resolution using a restrained least-squares procedure,¹ while 1ubi is a model of synthetic ubiquitin refined with the coordinates of 1ubq as the starting point of the refinement using the same procedure against the synthetic X-ray data.² 3n32 is a model of ubiquitin to which platinum ions bind (Pt₃-hUb)³ (Pt is omitted in our calculations).

A3.2. Five Sets of Candidate Models Obtained from NMR Experiments.

The model sets of 1d3z, 1g6j, 2klg, 2jzz, and 2l3z give candidate models for a fixed NS. 1d3z was obtained from a structure calculation on the basis of 2727 NOEs, 98 dihedral angle constraints, and 372 RDC restraints.⁴ 1g6j was obtained from a structure calculation based on 1291 NOEs, 63 dihedral angle constraints, and 23 hydrogen bonding restraints.⁵ A distinctive aspect of 1g6j is the encapsulation of the protein within a reverse micelle. The encapsulation

was undertaken for the purpose of improving the relaxation properties even for a large protein. 2klg was obtained via an NMR experiment in which inert paramagnetic molecules were added so that a rather small number of NOEs could be complemented with the restraints from paramagnetic relaxation enhancements (PREs).⁶ 2jzz and 2l3z were the products from solid-state NMR experiments intended for obtaining long-range distance information, which was suitable to a large, insoluble protein.^{7,8} It should be noted that the environment is considerably different from that in solution NMR. The structure calculation is usually performed using standard methods for generating protein structures satisfying geometric restrictions determined by solution or solid-state NMR. The most popular method is the CYANA program⁴² of an MD simulation with the idea of simulated annealing which uses torsion angles instead of Cartesian coordinates as the degrees of freedom.

A3.3. Five Ensembles of Structures Constructed for Representing Structural Variability of Native State.

The models in 1xqq, 2nr2, and 2kox were constructed using the ensemble refinement protocol started from the model in 1ubq.^{9,10,13} The ensemble refinement protocol was also employed in 2k39, but the initial structures were generated by amending random coils using the CONCOORD method⁴³ based on the geometric restrictions of NOEs.¹¹ The structural ensemble was constructed on the basis of such constraints as NOEs, RDCs, and order parameters S^2 obtained in previously reported NMR experiments.^{9-11,13} The construction of the ensemble was made to represent structural fluctuations rather than candidate structures for a fixed NS with the emphasis on protein dynamics in aqueous solution. The ensemble was obtained by employing an all-atom MD simulation with explicit water in 1xqq, 2k39, or 2kox while a hybrid-type approach of MD simulations with explicit water and in vacuum was adopted in 2nr2. The approach in 2kn5 first used unrestrained structural sampling with a Monte Carlo (MC) protocol and the Backrub motional model⁴⁴ to generate a large set of structures, starting from the model in 1ubq.¹² It then selected an ensemble optimizing the agreement with RDCs. Due to the particular method employed in 2kn5, the root mean square deviation (RMSD) for C_α atoms from the model in 1ubq (and from that in 1ubi) becomes intrinsically small.

A3.4. Slight Modification of Native-Structure Models.

The coordinates of hydrogen atoms cannot be obtained by the X-ray diffraction. We give hydrogen atoms to each model using the CHARMM biomolecular simulation program⁴⁵ through the Multi-scale Modeling Tools in Structural Biology (MMTSB) program.⁴⁶

The LJ potential energy for many of the NS models is positive and quite large due to unrealistic overlaps of protein atoms. Such overlaps are removed by the minimization of the energy function using the CHARMM and MMTSB programs. The minimization is performed

so that the original structures can be retained as much as possible. We employ the CHARMM22³⁴ with the CMAP correction⁴⁷ as the force-field parameters. Electrostatic and non-bonded interactions are all evaluated without any cut-off. The Generalized-Born (GBMV/SA) approximation⁴⁸⁻⁵⁰ is employed for the electrostatic part of the hydration free energy. After the minimization, there are no unrealistic overlaps of protein atoms. Moreover, it is verified for each structure that the RMSD for C_α atoms from the structure before the minimization is quite small: 0.5-0.7 Å in 2kn5, 0.1-0.3 Å in 1g6j, 2jzz, and 2l3z, and 0.03-0.06 Å in the others. Each structure is then switched to a set of fused hard spheres in calculating the hydration entropy.

A4. Results and Discussion

A4.1. Free-Energy Function Plotted against RMSD or TM-score.

Fig. A1 shows ribbon representation of the X-ray model in 1ubq and NMR models with the highest and lowest values of F in 1d3z, 2klg, 2k39, 2kox, and 2l3z. The three X-ray models, which are almost the same, possess three and one-half turns of α -helix, two short pieces of 3_{10} -helix, a mixed β -sheet containing five strands, and seven reverse turns.¹⁻³ It is apparent that the secondary structures of the models in 2klg are incomplete. Even the models looking almost indistinguishable can be considerably different from one another in terms of F , S , and A (i.e., the details of the secondary structures, backbone and side-chain packing, and intramolecular hydrogen bonding).

The values of our free-energy function F for the three X-ray models follow the order, 1ubi~3n32<1ubq though the differences are quite small. In general, the X-ray models give lower F than the NMR models: Only 14 models in 2kox and 1 model in 1xqq give lower F than the X-ray model in 1ubq. The model with the lowest F is in 2kox. The X-ray model in 1ubi, whose F is the second lowest, is regarded as *the reference model*. We consider the Template Modeling score (TM-score)⁵¹ as well as the RMSD for C_α atoms to look at the similarity between two protein structures. The TM-score indicates the structural difference by a score in the range (0, 1]. The score 1 implies a perfect match of the structures. It can be assumed that a score higher than 0.5 implies significantly high similarity.

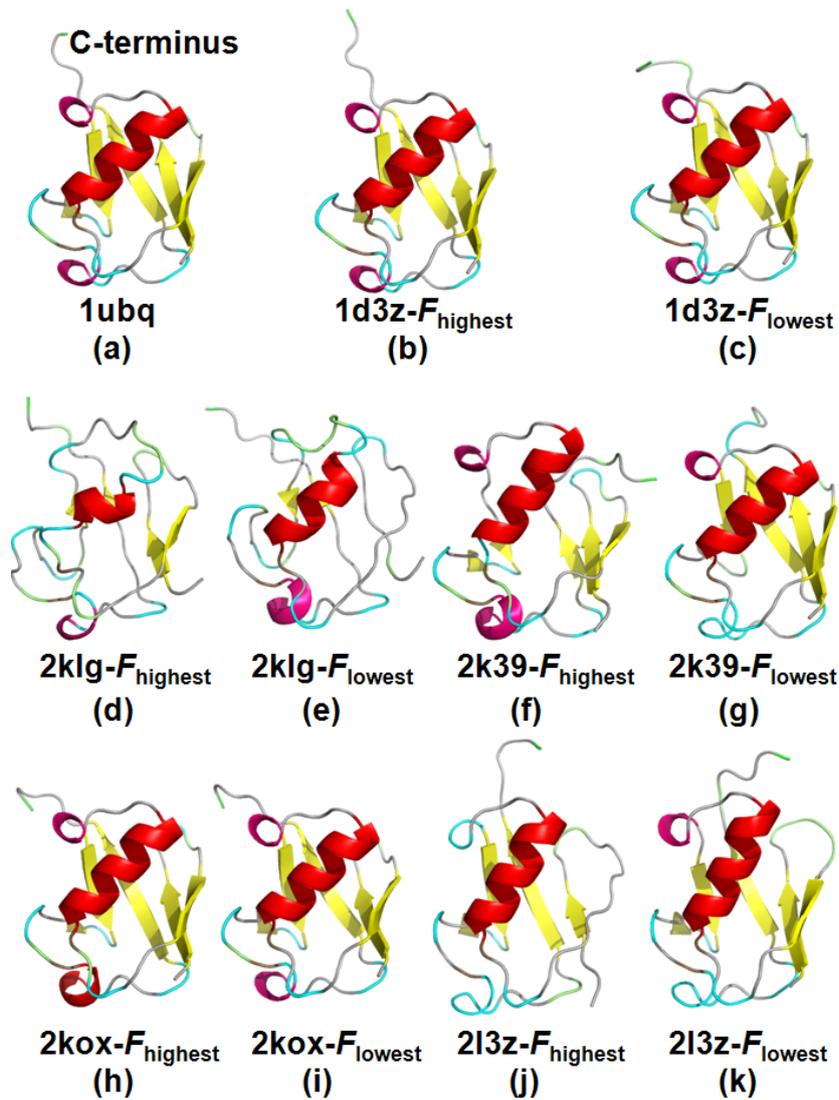


Fig. A1. Ribbon representation of an X-ray model and NMR models with the highest and lowest values of F in representative model sets. The models with the highest and lowest values of F in 1d3z, for instance, are denoted by 1d3z- F_{highest} and 1d3z- F_{lowest} , respectively. (a) 1ubq, (b) 1d3z- F_{highest} , (c) 1d3z- F_{lowest} , (d) 2klg- F_{highest} , (e) 2klg- F_{lowest} , (f) 2k39- F_{highest} , (g) 2k39- F_{lowest} , (h) 2kox- F_{highest} , (i) 2kox- F_{lowest} , (j) 2l3z- F_{highest} , and (k) 2l3z- F_{lowest} . Residues 71-76 on the C-terminus side are significantly flexible. This figure was drawn by PyMol 1.3.

Figs. A2 and A3 show the relation between $(F-F_{1ubi})_{av}$ (the subscripts “av” and “1ubi” denote the average value in each model set and the value for the reference model, respectively) and the RMSD or TM-score calculated with the reference model as the standard structure. The standard deviation of $F-F_{1ubi}$, $(F-F_{1ubi})_{sd}$ (the subscript “sd” denotes the standard deviation), and that of the RMSD or TM-score are indicated as error bars. In the calculation of the TM-score and RMSD, all the residues (1-76) are considered in Fig. A2 while only the core region comprising residues 1-70 are considered in Fig. A3. It is known that residues 71-76 which we do not include in the core region are significantly flexible:¹¹ They are disordered as observed in Fig. A1. This is because they do not participate in the packing. Since the correlation between $(F-F_{1ubi})_{av}$ and the RMSD is influenced by these flexible residues, it is not appropriate to discuss the correlation by looking at Fig. A2(a). As observed in Figs. A2(b) and A3, there is a general trend that $(F-F_{1ubi})_{av}$ becomes higher as the TM-score decreases or the RMSD increases and the model deviates more from the reference model in terms of the backbone structure. The correlation between $(F-F_{1ubi})_{av}$ and the TM-score or RMSD can be represented by a curve (not drawn) for 1d3z, 1g6j, 2klg, 2k39, and 2l3z. The results for 2kox, 2nr2, 1xqq, and 2jzz deviate from the curve in a downward direction, while the result for 2kn5 deviates from it in an upward direction.

We first discuss model sets of type 1. On an average, the models in 1d3z are the closest to the reference model in term of the TM-score, RMSD, and F , and those in 1g6j are the second closest. In 2klg, the models exhibit considerably high F and the largest deviation from the reference model in terms of the structural properties. Almost the same characteristics are observed in 2l3z. As for model sets of type 2, on an average, the models in 2kox, 2nr2, and 1xqq possess relatively lower F . Compared with the models in 1d3z, those in 2kox, 2nr2, and 1xqq are more different from the reference model with respect to the structural properties. However, the values of F for 2kox, 2nr2, and 1xqq are as low as those for 1d3z. We find that the models in 2kn5 give higher F than all the models in 2kox, 2nr2, and 1xqq. Among model sets of type 2, 2kn5 gives the highest value of F_{av} .

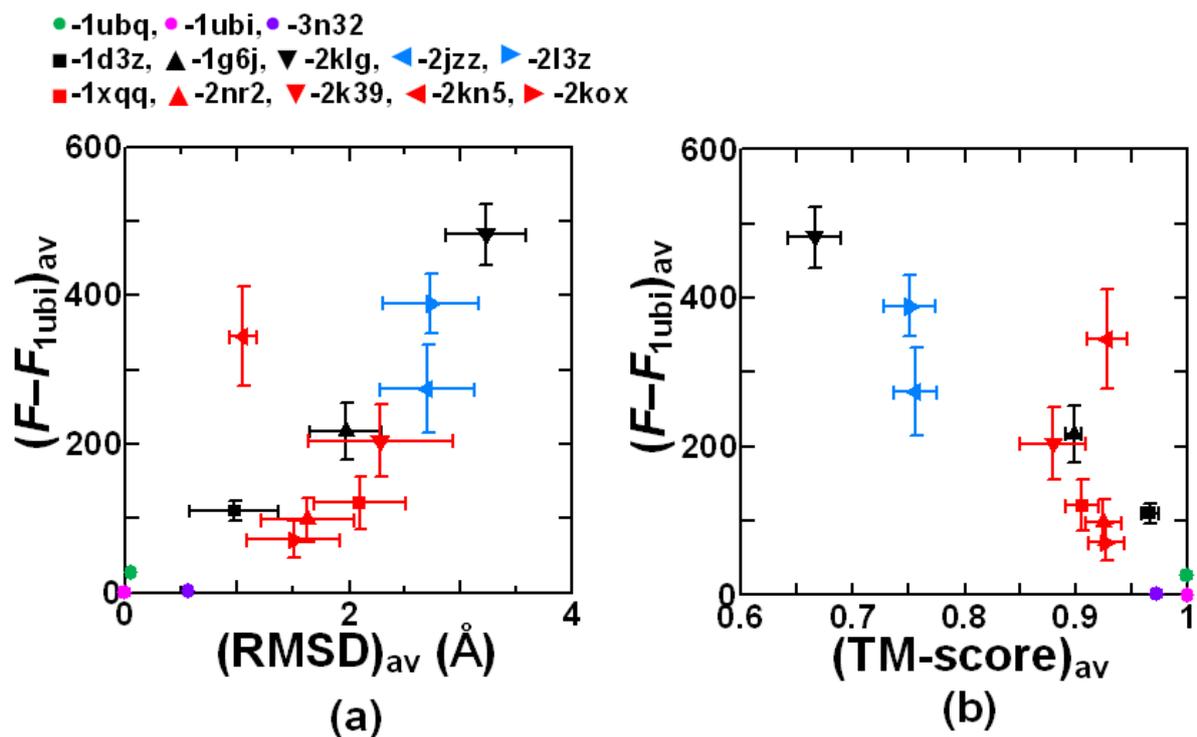


Fig. A2. Relation between $(F-F_{1ubi})_{av}$ (the subscripts “av” and “1ubi” denote the average value in each model set and the value for the X-ray model in 1ubi, respectively) and RMSD or TM-score for C_{α} atoms calculated with the model in 1ubi as the standard structure. The standard deviation of $F-F_{1ubi}$, $(F-F_{1ubi})_{sd}$ (the subscripts “sd” denote the standard deviation), and that of the RMSD or TM-score are indicated as error bars. All the residues (1-76) are considered in the calculation of the TM-score and RMSD. (a): RMSD. (b): TM-score. Black: solution NMR in model sets of type 1, blue: solid-state NMR in model sets of type 1, and red: ensembles in model sets of type 2.

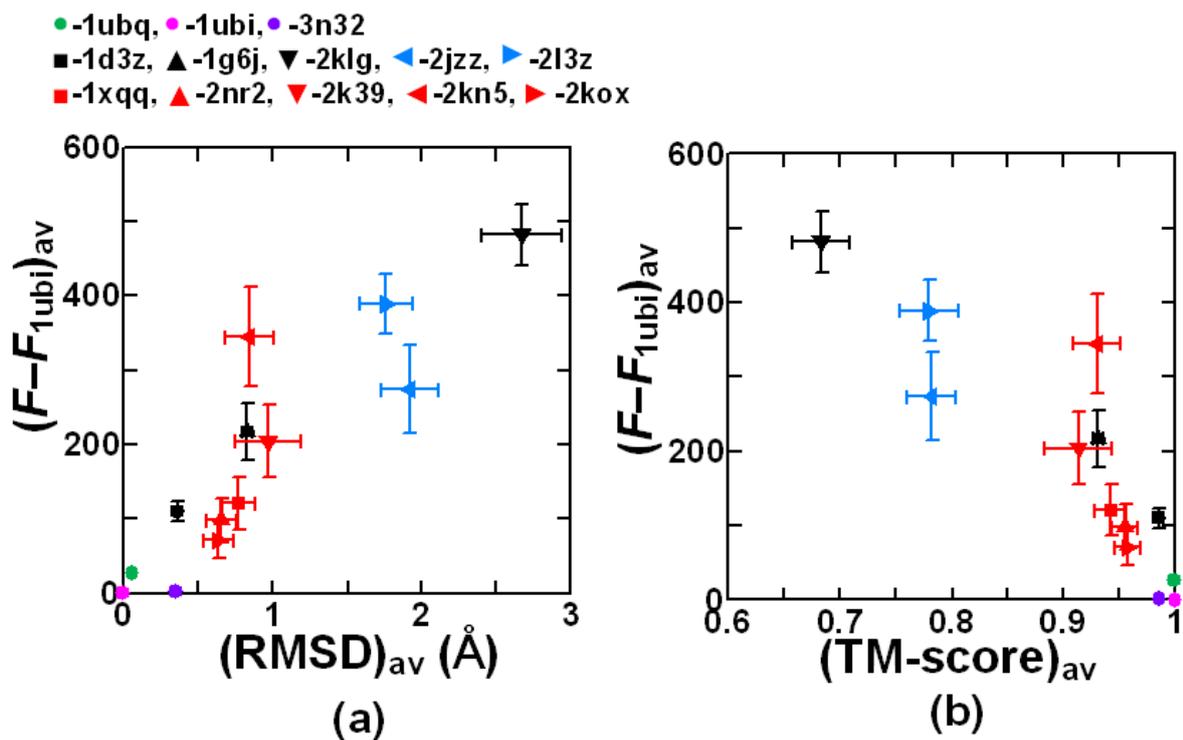


Fig. A3. Relation between $(F-F_{1ubi})_{av}$ (the subscripts “av” and “1ubi” denote the average value in each model set and the value for the X-ray model in 1ubi, respectively) and RMSD or TM-score for C_α atoms calculated with the model in 1ubi as the standard structure. The standard deviation of $F-F_{1ubi}$, $(F-F_{1ubi})_{sd}$ (the subscripts “sd” denote the standard deviation), and that of the RMSD or TM-score are indicated as error bars. The core region comprising residues 1-70 are considered in the calculation of the TM-score and RMSD. (a): RMSD. (b): TM-score. Black: solution NMR in model sets of type 1, blue: solid-state NMR in model sets of type 1, and red: ensembles in model sets of type 2.

A4.2. Relation between Properties of Free-Energy Function and Structural Convergence in Model Set.

In order to examine the relation between the properties of F and the convergence of model structures, we propose the following procedure for each model set: First, calculate the averaged structure; second, calculate RMSDs for heavy atoms in the backbone and side chains *from the averaged structure*; calculate the average value $(\text{RMSD})_{\text{av}}$ and standard deviation $(\text{RMSD})_{\text{sd}}$; and plot the relation between $(\text{RMSD})_{\text{av}}$ and F_{av} as a figure. In the figure, $(\text{RMSD})_{\text{sd}}$ and F_{sd} are also indicated as error bars. Residues 71-76 are excluded from the calculation of RMSDs.

Fig. A4 shows the plots explained above. A curve (curve 1; not drawn) correlating F_{av} with $(\text{RMSD})_{\text{av}}$ is found for the three X-ray models and the models in 1d3z, 1g6j, 2klg, 2l3z, and 2kn5. The correlation between F_{av} and $(\text{RMSD})_{\text{av}}$ for 2kox, 2nr2, 1xqq, and 2k39 is expressed by another curve (curve 2; not drawn). The result for the models in 2jzz seems to deviate from both of curves 1 and 2: it is in-between. In each curve, there is a strong tendency that $(\text{RMSD})_{\text{sd}}$ as well as F_{av} decreases as $(\text{RMSD})_{\text{av}}$ becomes smaller. Smaller $(\text{RMSD})_{\text{av}}$ and $(\text{RMSD})_{\text{sd}}$ imply better convergence of the model structures. Curve 2 is shifted in a downward direction in comparison with curve 1. Namely, for model sets of type 2 excluding 2kn5, F_{av} is relatively lower for a given value of $(\text{RMSD})_{\text{av}}$. When the model structures are better converged, not only F_{av} but also F_{sd} tends to become lower, though the models in 2kn5 possess an exceptionally high value of F_{sd} . Better convergence of F or the model structures is indicative of a larger total amount of constraints effectively taken into account in the structure calculation or the ensemble refinement protocol. It is important to know how effectively the experimentally obtained constraints have been incorporated in the models constructed. However, there are a variety of constraints (e.g., NOEs, RDC, hydrogen bonding, and dihedral angle restraints) and different constraints are used in different amounts. The details of the way employed for imposing the constraints on the model structures are also variable. Nevertheless, impartial comparison among the structures in model sets can be made through the plots like Fig. A4 using our free-energy function with regard to the total amount of constraints effectively taken into account.

Among model sets of type 1, 1d3z gives the smallest values of $(\text{RMSD})_{\text{av}}$ and F_{av} . The model structures in 1d3z are the best converged and on the whole they are the closest to the reference model, the X-ray model in 1ubi. Those in 1g6j are the second best converged and the second closest to the reference model. $(\text{RMSD})_{\text{av}}$ and F_{av} take the largest values for 2klg. As for model sets of type 2, 2kox, 2nr2, and 1xqq are successful in the construction of fluctuating structures in aqueous solution. This is particularly true for the structural ensemble in 2kox. The success is also reflected on the deviation from the curve in a downward direction observed in Fig. A2(b) and A3. Last, we remark the following: If there is a data set for which $(\text{RMSD})_{\text{av}}$ and $(\text{RMSD})_{\text{sd}}$ are small though F_{av} and F_{sd} are high, its structures are

well converged to a different model, an NS model somewhat vitiated by unfavorable solvent condition (such a case is not found in Fig. A4).

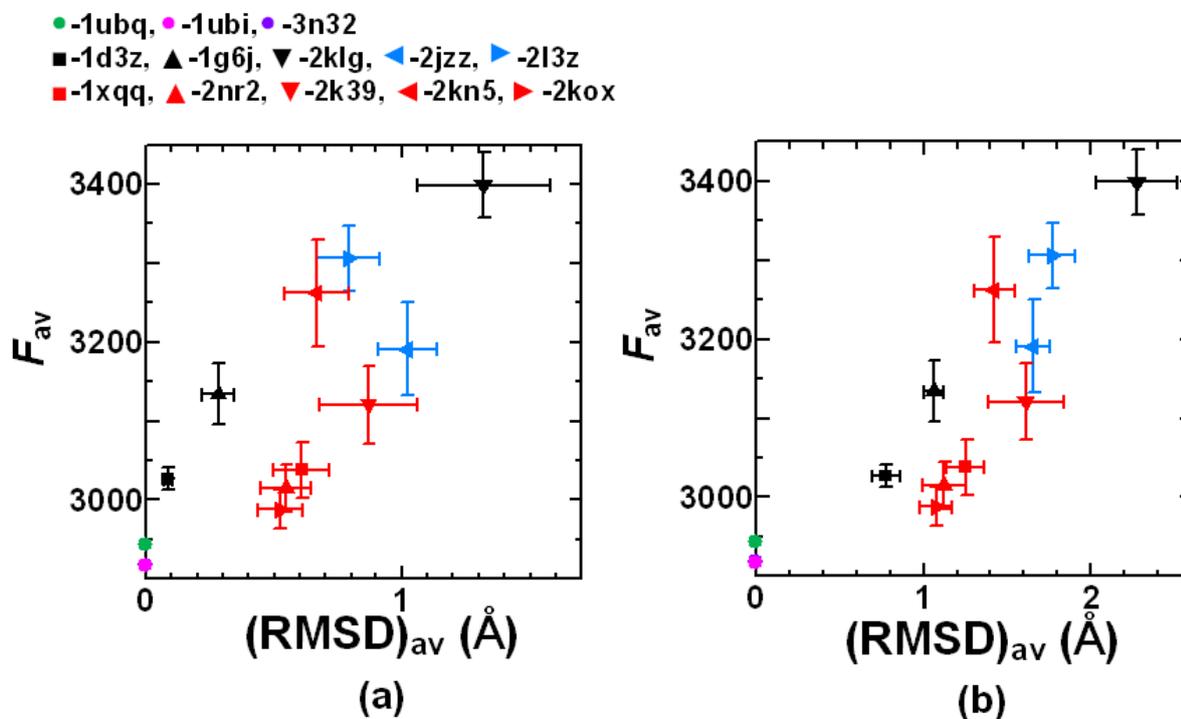


Fig. A4. F_{av} plotted against $(\text{RMSD})_{av}$ (the subscript “av” denotes the average value). The core region comprising residues 1-70 are considered. The RMSD, which is for heavy atoms, represents the deviation from the averaged structure calculated for the structures in each model set. Thus, the definition of the RMSD is different from that in Fig. A2 or A3. F_{sd} and $(\text{RMSD})_{sd}$ (the subscript “sd” denotes the standard deviation) are indicated as the error bars. (a) Plot for the backbone. (b) Plot for the side chains. Black: solution NMR in model sets of type 1, blue: solid-state NMR in model sets of type 1, and red: ensembles in model sets of type 2. The points for 1ubi and 3n32 are almost indistinguishable.

A4.3. Entropic and Energetic Components.

We define X and Y as

$$X = \{A/(k_B T)\} - \{A/(k_B T)\}_{1ubi}, \quad (\text{A7})$$

and

$$Y = (-S/k_B) - (-S/k_B)_{1ubi}, \quad (\text{A8})$$

respectively. Λ or X is a measure of the assurance of intramolecular hydrogen bonds for decreasing the total dehydration penalty.¹⁴⁻¹⁶ A higher value of X implies less intramolecular hydrogen bonds formed. Close packing of the backbone and side chains leads to the reduction of the excluded volume generated for water molecules by a protein followed by the relaxation of water crowding.^{31,52-54} $-S$ or Y represents the efficiency of backbone and side-chain packing for making the hydration entropy as small as possible.^{14-16,31} Low efficiency results in a higher value of Y .

The plot of Y_{av} against X_{av} is shown for the NMR models in Fig. A5: Y_{sd} and X_{sd} are indicated as error bars. The three X-ray models, which share almost the same characteristics with respect to Λ and S , are also included in the plot. There is a general trend that as X_{av} becomes higher, Y_{av} also increases. The points of each data set tend to gather in its own way, reflecting the specifications of the experimental technique and the structure calculation or the ensemble refinement protocol employed. For the models in 2l3z and 2kn5, Y_{av} is relatively higher, indicating less efficient packing of the backbone and/or side chains. By contrast, those in 1g6j and 2jzz suffer relatively higher values of X_{av} and less intramolecular hydrogen bonds formed.

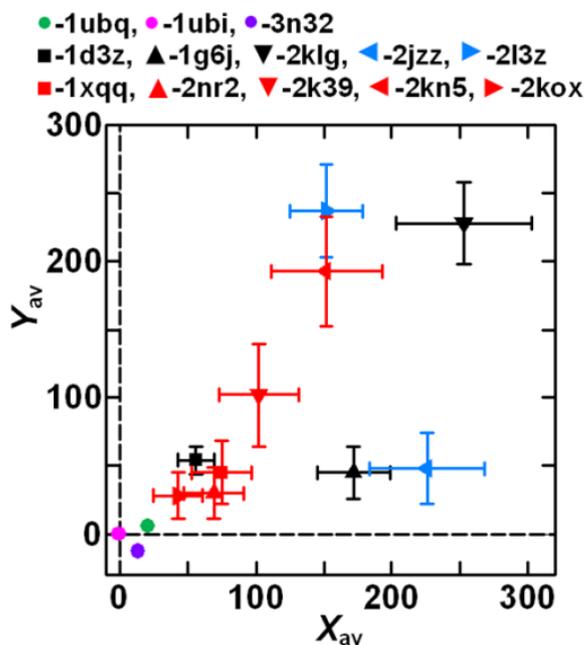


Fig. A5. Y_{av} plotted against X_{av} (the subscript “av” denotes the average value). X and Y are defined by Eq. (A7) and Eq. (A8), respectively. Y_{sd} and X_{sd} (the subscript “sd” denotes the standard deviation) are indicated as error bars. Black: solution NMR in model sets of type 1, blue: solid-state NMR in model sets of type 1, and red: ensembles in model sets of type 2.

A4.4. Definition of Contributions to Energetic and Entropic Components from Backbone and Side Chains.

It is physically insightful to separate the effect of side chains from that of the backbone. The contributions from the backbone and side chains to A , which are denoted by A_b and A_{sc} , respectively, can readily be obtained. X_b and X_{sc} are defined by Eq. A2 where A_b and A_{sc} are substituted for A , respectively. To perform the separation for the entropic component, we replace all residues in each structure by Gly using the CHARMM and MMTSB programs.^{45,46} The replacement is carried out after the slight modification of the structure described in “Slight Modification of Native-Structure Models”. The structure thus made has essentially no side chains (hereafter, these are referred to as “the structures without side chains”). $-S$ represents the loss of the water entropy upon the insertion of a protein with a prescribed structure. The information on the effect of side chains is contained in “ $-S$ of a structure with side chains” – “ $-S$ of the corresponding structure without side chains (i.e., with the backbone alone)”: The latter is denoted by $-S_b$ and $-S = -S_b + (-S_{sc})$. $-S_b$ and $-S_{sc}$ denote the contributions from the backbone and side chains to $-S$, respectively. Y_b and Y_{sc} are defined by Eq. A3 where $-S_b$ and $-S_{sc}$ are substituted for $-S$, respectively.

A4.5. Physical Origin of Superiority or Inferiority of Energetic Component for NMR Models.

Fig. A6 shows the average values of X_b and X_{sc} (they are denoted by $X_{b,av}$ and $X_{sc,av}$, respectively) for the NMR models. It is observed that X_b is much larger than X_{sc} . In comparison with the X-ray models, the NMR models undergo larger total dehydration penalty in the backbone than in side chains. This is particularly true for 2klg and 2jzz. Among model sets of type 1, $X_{b,av}$ is the smallest for 1d3z and the largest for 2klg. Among those of type 2, $X_{b,av}$ is the smallest for 2kox and the largest for 2kn5.

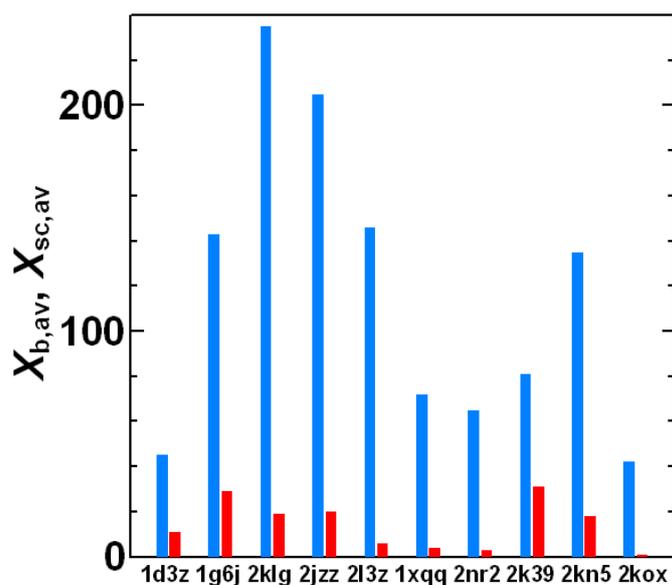


Fig. A6. Average values of X_b and X_{sc} (they are denoted by $X_{b,av}$ and $X_{sc,av}$, respectively) for the NMR models. X is defined by eq A7. X_b and X_{sc} , respectively, are the contributions from the backbone and side chains to X . X_b and X_{sc} are marked in blue and red, respectively.

The contribution from each residue to $\Lambda_{b,av}$ is shown in Fig. A7 where the results for 1ubi, 1d3z, and 2klg or those from 1ubi, 2kox, and 2kn5 are compared. At the top of the figure, the portions of α -helix, β -sheet, and 3_{10} -helix for 1ubi are indicated in three different colors. These secondary structures are identified using the DSSP program.⁵⁵ The three X-ray models share the characteristic that the loop portions as well as the secondary structures in the backbone exhibit only small dehydration penalty and sufficiently many intramolecular hydrogen bonds are formed. On an average, the models in 1d3z and 2kox undergo larger dehydration penalty primarily in the loop portions. However, we find that significantly many models in 2kox possess as many intramolecular hydrogen bonds as the three X-ray models in the secondary structures and the loop portions. As for the models in 2klg and 2kn5, intramolecular hydrogen bonds are not always formed even in the secondary structures, causing considerably large dehydration penalty.

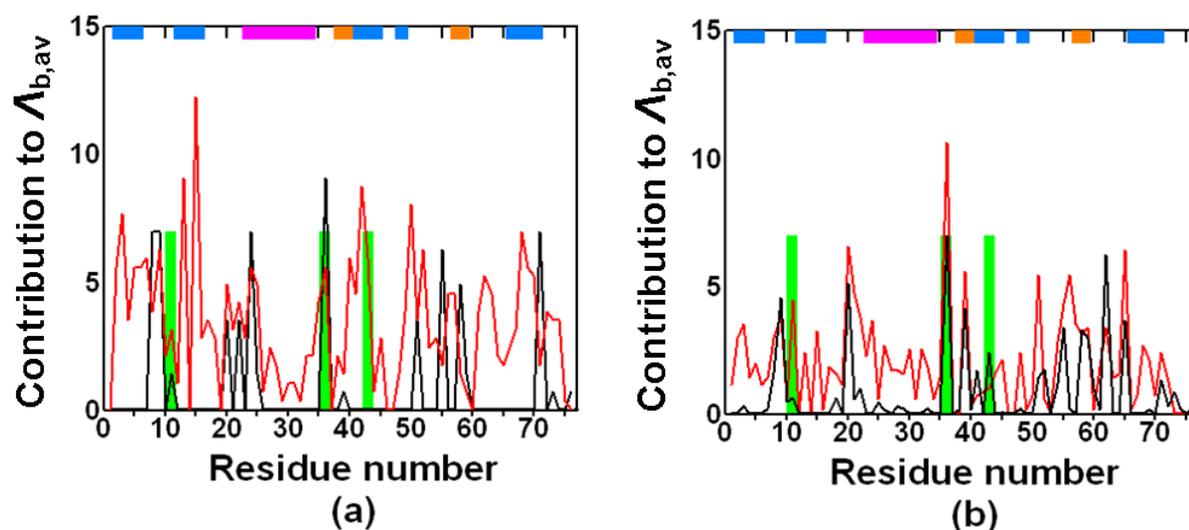


Fig. A7. Contribution from each residue to the average value of Λ_b (it is denoted by $\Lambda_{b,av}$). Λ is the total dehydration penalty in Eq. A3. The results for 1ubi (green), 1d3z (black), and 2klg (red) are compared in (a), and those for 1ubi (green), 2kox (black), and 2kn5 (red) are compared in (b). Only residues 11, 36, and 43 undergo the dehydration penalty in 1ubi, and the contribution is shown by a vertical bar. At the top of the figure, the portions of α -helix, β -sheet, and 3_{10} -helix for 1ubi are indicated in pink, blue, orange, respectively: For example, residues 23-34 form α -helix. These secondary structures are not always complete in 2klg and 2kn5.

A4.6. Physical Origin of Superiority or Inferiority of Entropic Component for NMR Models.

Fig. A8 shows the average values of Y_b and Y_{sc} (they are denoted by $Y_{b,av}$ and $Y_{sc,av}$, respectively) for the NMR models. Except for the models in 2jzz, Y_{sc} is larger than Y_b . In the NMR models, the deterioration of the packing efficiency is more significant in side chains than in the backbone, which is in contrast to the case of the total dehydration penalty. The side-chain packing is relatively more inefficient for 2l3z, 2kn5, and 2klg. Among model sets of type 2, 2kn5 features appreciably inefficient packing of side chains.

As illustrated in Fig. A9, there is an overall trend that as the side-chain packing becomes less efficient, the efficiency of the backbone packing is also lowered. On an average, the packing of the backbone and side chains for the models in 2kox and 2nr2 is as close as that for the three X-ray models.

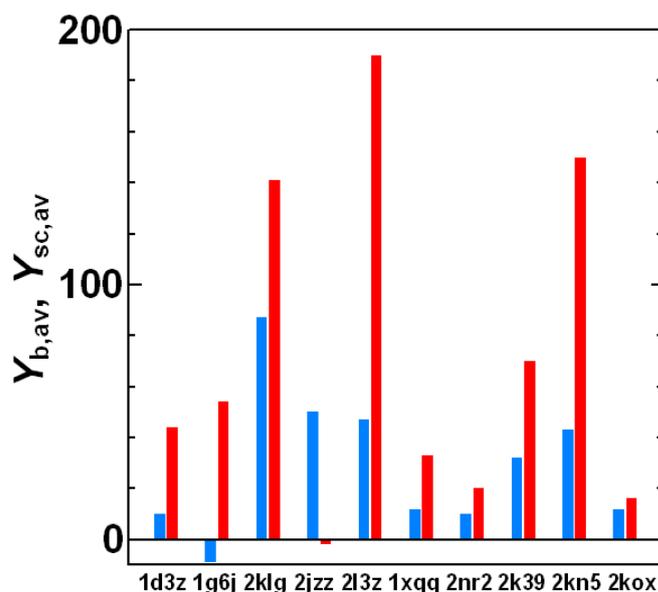


Fig. A8. Average values of Y_b and Y_{sc} (they are denoted by $Y_{b,av}$ and $Y_{sc,av}$, respectively) for the NMR models. Y is defined by Eq. A8. Y_b and Y_{sc} , respectively, are the contributions from the backbone and side chains to Y . Y_b and Y_{sc} are marked in blue and red, respectively.

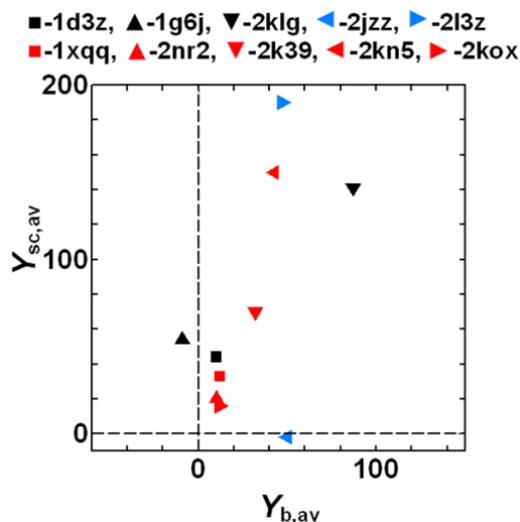


Fig. A9. Relation between $Y_{b,av}$ and $Y_{sc,av}$ denoting the average values of Y_b and Y_{sc} , respectively, for the NMR models. Y is defined by Eq. A3. Y_b and Y_{sc} , respectively, are the contributions from the backbone and side chains to Y . Black: solution NMR in model sets of type 1, blue: solid-state NMR in model sets of type 1, and red: ensembles in model sets of type 2.

A4.7. Relevance to Specificity of NMR Structures in Model Set.

Since any NMR approach has its own advantages, we can never grade the previously reported approaches. Yet, the characteristics of the NMR models can be summarized as follows.

We first discuss model sets of type 1. 1d3z merits the largest total amount of constraints effectively taken into account and the best converged model structures, leading to the lowest values of F and A and the second lowest value of $-S$ on the whole. The structural characteristics as well as the values of F , A , and $-S$ are close to those of the three X-ray models. From this result, we can conclude that the NS of ubiquitin is influenced neither by crystallization in the X-ray crystallography nor by solvent environment adjusted in the NMR experiment for 1d3z. The characteristics of the models in 1g6j are fairly close to those in 1d3z except the relatively larger total dehydration penalty in the backbone. The structure of ubiquitin remains almost undisturbed upon encapsulation within a reverse micelle. As for the models in 2klg, the total amount of constraints effectively taken into account is not sufficiently large and the models overall suffer the lack in firmly formed intramolecular hydrogen bonds in the backbone. They feature incomplete secondary structures, which can be appreciated even in visualized ribbon representation of the model structures (see Fig. A1). It is interesting that the side-chain packing for 2l3z is the least efficient while that for 2jzz is the most efficient despite that both 2l3z and 2jzz were obtained through solid-state NMR. The

differences are ascribed probably to the details of the structure calculations employed. The effects of such details on the NS models generated are usually ambiguous, and this is one of the reasons why the characterization of the models becomes useful. The most efficient packing of side chains for 2jzz is the principal reason for the deviations from the correlations among 1d3z, 1g6j, 2klg, and 2l3z observed in Figs. A2-4.

We then discuss model sets of type 2. The structural variability of the native state in aqueous solution is well represented by the models in 2kox, 2nr2, and 1xqq. In particular, 2kox provides successful representation of the structures fluctuating around the X-ray structure in 1ubq for the following reasons: The value of F for 14 models in 2kox is lower than that for the X-ray structure; as mentioned above, the structure with the lowest value of F is included in 2kox; the RMSD and TM-score of this structure calculated with the reference model as the standard structure (the core region comprising residues 1-70 are considered) are 0.66 Å and 0.96, respectively; and it is observed in Fig. A4 that 2kox merits the largest total amount of constraints effectively taken into account. For 2k39, 2kn5, and 2kox, in Fig. A10 we show the relation between $F-F_{1ubi}$ of all the models and the RMSD or TM-score for C_α atoms calculated with the reference model as the standard structure. The F -value in 2k39 varies considerably from model to model, and the maximum difference in F observed is as large as ~ 300 . 2k39 includes significantly many structures with high values of F . On the whole, the models in 2kn5 suffer the largest deviations from the X-ray structures and the highest values of F primarily due to the rather unfavorable side-chain packing. This is why 2kn5 departs from the correlations among 2kox, 2nr2, 1xqq, and 2k39 as observed in Figs. A2-4.

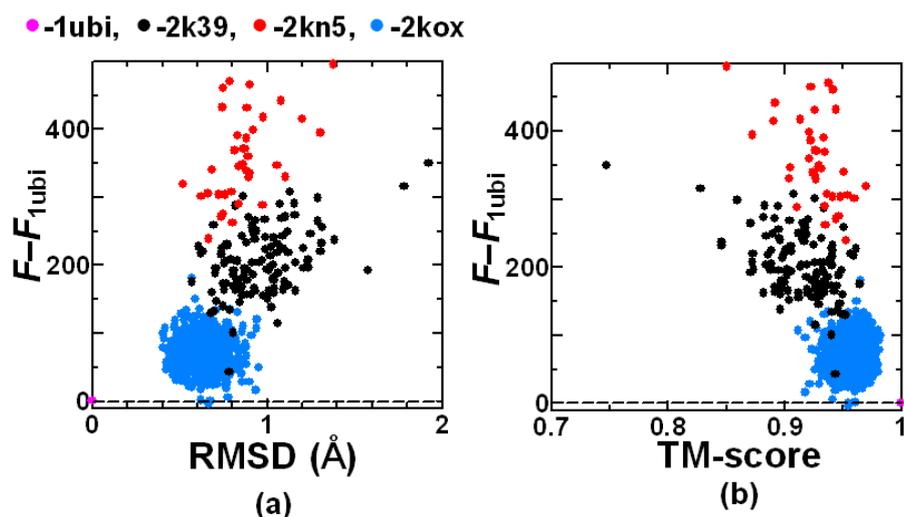


Fig. A10. Relation between $F - F_{1ubi}$ and RMSD or TM-score for C_α atoms calculated with the model in 1ubi as the standard structure. The core region comprising residues 1-70 are considered in the calculation of the RMSD or TM-score. All the models in 2k39, 2kn5, and 2kox are individually treated. (a): RMSD. (b): TM-score.

A5. Conclusion

We have developed a reliable method of characterizing the NS models of a protein determined through a variety of routes. It is illustrated for ubiquitin for which a comprehensive range of high-quality X-ray and NMR data are available. The characterization is based on not mere geometric analyses but energetics of a protein accounting for its hydration thermodynamics wherein a molecular model is employed for water. The NS models can be classified into X-ray models¹⁻³ and two types of NMR model sets. A model set of type 1 comprises candidate models for a fixed NS determined by solution or solid-state NMR with the aid of a structure calculation upon which the constraints experimentally obtained⁴⁻⁸ (e.g., NOEs, RDCs, hydrogen bonding and dihedral angle restraints) are imposed. That of type 2 forms an ensemble of structures representing the structural variability of the native state in aqueous solution.⁹⁻¹³ It is usually constructed by employing the ensemble refinement protocol^{9-11,13} started from an NS model determined by the X-ray crystallography or solution NMR. The refinement is performed in a variety of manners using a computer simulation with all-atom potentials or a method based on geometric restrictions which is restrained by NOEs, RDCs, and order parameters S^2 obtained in NMR experiments. Our recently developed free-energy function F and its energetic and entropic components¹⁴⁻¹⁶ (A and S , respectively) are shown to be very useful to the characterization.

In general, the X-ray models give lower values of F than the NMR models. The X-ray models resemble one another with respect to the structural features, F , A , and S , whereas the NMR models exhibit diverse properties. These results never deny the usefulness of NMR. In the X-ray crystallography, it may be difficult to find out the solvent condition leading to the crystallization of a protein which is not far from the physiological condition. Moreover, only a smaller amount of information on dynamical properties of a protein structure is provided by the X-ray crystallography.

The model giving the lowest value to F among the X-ray models is chosen as the reference model. If there are only NMR models, the model with the lowest value of F among them is the best one for the reference one. The plot of $(F-F_{\text{ref}})_{\text{av}}$ (the subscripts “av” and “ref” denote the average value and the value of the reference model, respectively) against the TM-score⁵¹ for C_{α} atoms calculated with the reference model as the standard structure allows us to evaluate the quality of each model set. $(F-F_{\text{ref}})_{\text{sd}}$ (the subscript “sd” denote the standard deviation) and the standard deviation of the TM-score are indicated as error bars. There is a general trend that $(F-F_{\text{ref}})_{\text{av}}$ becomes higher as the TM-score decreases and the model deviates more from the reference model in terms of the backbone structure (this trend is required in any characterization method). Namely, $(F-F_{\text{ref}})_{\text{av}}$ is correlated with the TM-score. This correlation can be described by a curve, but some model sets exhibit deviations from the curve. In particular, model sets of type 2, when they successfully represent fluctuating

structures of the native state in aqueous solution, deviate from the curve in a downward direction. A good structural ensemble for the native state should include significantly many structures for which the value of F is lower than that for the reference model. When the TM-score is replaced by the RMSD, the correlation between $(F-F_{\text{ref}})_{\text{av}}$ and the RMSD is observed only if the residues with high flexibility are removed (compare Figs. A2(a) and A3(a)). The flexible residues, which exhibit large fluctuations, are visualized by superposing the set of structures. Such residues can perfectly be identified by the plots like Figs. A2(a) and A3(a) together with the superposition. In general, when the RMSD is smaller than 1 Å or the TM-score is larger than 0.9 between two structures, they are considered almost the same. Our free-energy function, which possesses high resolution power, is capable of distinguishing even those structures.

In each model set, the relation between the structural convergence and properties of the free-energy function can be explored in the following manner: First, calculate the averaged structure; second, calculate RMSDs for heavy atoms in the backbone and those in the side chains from the averaged structure; calculate the average value $(\text{RMSD})_{\text{av}}$ and standard deviation $(\text{RMSD})_{\text{sd}}$; and plot the relation between $(\text{RMSD})_{\text{av}}$ and F_{av} as a figure. In the figure, $(\text{RMSD})_{\text{sd}}$ and F_{sd} are also indicated as error bars. The flexible residues which can be identified as described above (see Fig. A4) are excluded from the calculation of RMSDs. Model sets of type 1 and those of type 2 can be described by different correlation curves. In each curve, there is a strong tendency that as $(\text{RMSD})_{\text{av}}$ becomes smaller, $(\text{RMSD})_{\text{sd}}$ and F_{av} decrease. However, F_{av} is lower for a given value of $(\text{RMSD})_{\text{av}}$ in model sets of type 2 than in those of type 1. Smaller $(\text{RMSD})_{\text{av}}$ and $(\text{RMSD})_{\text{sd}}$ imply better convergence of the model structures. When the model structures are better converged, not only F_{av} but also F_{sd} tends to become lower (this property is required in any characterization method). Even when there is a model set which deviates from the correlation curve, the physical reasoning for the deviation can be made, thus uncovering its characteristics. If we find a data set for which $(\text{RMSD})_{\text{av}}$ and $(\text{RMSD})_{\text{sd}}$ are small though F_{av} and F_{sd} are high, the structures are well converged to an incorrect NS model, a model somewhat vitiated by unfavorable solvent condition. We remark that impartial comparison among the structures in model sets can be made through the plots like Fig. A4 using our free-energy function with regard to the total amount of constraints effectively taken into account.

The plot of X_{av} against Y_{av} ($X = \{\Lambda/(k_{\text{B}}T)\} - \{\Lambda/(k_{\text{B}}T)\}_{\text{ref}}$ and $Y = (-S/k_{\text{B}}) - (-S/k_{\text{B}})_{\text{ref}}$) with X_{sd} and Y_{sd} indicated as error bars gives useful information on the assurance of intramolecular hydrogen bonds, efficiency of backbone and side-chain packing, and balance of these two principal factors¹⁴⁻¹⁶ (see Fig. A5). By analyzing the contributions from the backbone and side chains to X_{av} and Y_{av} and drawing the figures such as Figs. A6 through A9, we can clarify detailed weak points of the models in an NMR model set in terms of intramolecular hydrogen bonding and packing efficiency in the backbone and side chains.

We emphasize that the results described above are achievable only by a free-energy function capturing essential physics of the structural stability of a protein in aqueous solution. It has been shown that F is certainly this type of function. The characterization method thus developed should be useful for the following applications: the evaluation of a set of NS models of a protein determined via a new NMR approach by comparing them to the models which are already available; and refinement of an NMR model by rectifying its weak points found. Further, our free-energy function is well suited not only to the selection of the best model from among many candidate NMR models but also to the original construction of the best candidate model or a good structural ensemble for the native state, on the basis of the experimentally obtained constraints such as NOEs and RDCs. Works in these directions are in progress.

References

- ¹S. Vijay-Kumar, C. E. Bugg, and W. J. Cook, *J. Mol. Biol.* **194**, 531 (1987).
- ²D. Alexeev, S. M. Bury, M. A. Turner, O. M. Ogunjobi, T.W. Muir, R. Ramage, and L. Sawyer, *Biochem. J.* **299**, 159 (1994).
- ³F. Arnesano, B. D. Belviso, R. Caliendo, G. Falini, S. Fermani, G. Natile, and D. Siliqi, *Chem. Eur. J.* **17**, 1569 (2011).
- ⁴G. Cornilescu, J. L. Marquardt, M. Ottiger, and A. Bax, *J. Am. Chem. Soc.* **120**, 6836 (1998).
- ⁵C. R. Babu, P. F. Flynn, and A. J. Wand, *J. Am. Chem. Soc.* **123**, 2691 (2001).
- ⁶T. Madl, W. Bermel, and K. Zangger, *Angew. Chem. Int. Ed.* **48**, 8259 (2009).
- ⁷T. Manolikas, T. Herrmann, and B. H. Meier, *J. Am. Chem. Soc.* **130**, 3959 (2008).
- ⁸M. Huber, S. Hiller, P. Schanda, M. Ernst, A. Böckmann, R. Verel, and B. H. Meier, *ChemPhysChem* **12**, 915 (2011).
- ⁹K. Lindorff-Larsen, R. B. Best, M. A. DePristo, C. M. Dobson, and M. Vendruscolo, *Nature* **433**, 128 (2005).
- ¹⁰B. Richter, J. Gsponer, P. Várnai, X. Salvatella, and M. Vendruscolo, *J. Biomol. NMR* **37**, 117 (2007).
- ¹¹O. F. Lange, N. A. Lakomek, C. Farès, G. F. Schröder, K. F. A. Walter, S. Becker, J. Meiler, H. Grubmüller, C. Griesinger, and B. L. de Groot, *Science* **320**, 1471 (2008).
- ¹²G. D. Friedland, N. A. Lakomek, C. Griesinger, J. Meiler, and T. Kortemme, *PLoS Comput. Biol.* **5**, e1000393 (2009).
- ¹³R. B. Fenwick, S. Esteban-Martín, B. Richter, D. Lee, K. F. A. Walter, D. Milovanovic, S. Becker, N. A. Lakomek, C. Griesinger, and X. Salvatella, *J. Am. Chem. Soc.* **2011**, 133, 10336.
- ¹⁴Y. Harano, R. Roth, Y. Sugita, M. Ikeguchi, and M. Kinoshita, *Chem. Phys. Lett.* **437**, 112 (2007).
- ¹⁵S. Yasuda, T. Yoshidome, Y. Harano, R. Roth, H. Oshima, K. Oda, Y. Sugita, M. Ikeguchi, and M. Kinoshita, *Proteins* **79**, 2161 (2011).
- ¹⁶S. Yasuda, T. Yoshidome, Y. Harano, R. Roth, H. Oshima, K. Oda, Y. Sugita, M. Ikeguchi, and M. Kinoshita, *Proteins* **79**, 2161 (2011).
- ¹⁷N. M. Cann and G. N. Patey, *J. Chem. Phys.* **106**, 8165 (1997).
- ¹⁸P. G. Kusalik and G. N. Patey, *J. Chem. Phys.* **88**, 7715 (1988).
- ¹⁹P. G. Kusalik and G. N. Patey, *Mol. Phys.* **65**, 1105 (1988).
- ²⁰M. Kinoshita and M. Harada, *Mol. Phys.* **81**, 1473 (1994).
- ²¹M. Kinoshita, S. Iba, and M. Harada, *J. Chem. Phys.* **105**, 2487 (1996).
- ²²M. Kinoshita and D. R. Bérard, *J. Comput. Phys.* **124**, 230 (1996).
- ²³M. Kinoshita, *J. Chem. Phys.* **128**, 024507 (2008).
- ²⁴P. M. König, R. Roth, and K. R. Mecke, *Phys. Rev. Lett.* **93**, 160601 (2004).

- ²⁵R. Roth, Y. Harano, and M. Kinoshita, *Phys. Rev. Lett.* **97**, 078101 (2006).
- ²⁶R. Kodama, R. Roth, Y. Harano, and M. Kinoshita, *J. Chem. Phys.* **135**, 045103 (2011).
- ²⁷F. Hirata and P. J. Rossky, *Chem. Phys. Lett.* **83**, 329 (1981).
- ²⁸J. S. Perkyins and B. M. Pettitt, *Chem. Phys. Lett.* **190**, 626 (1992).
- ²⁹J. S. Perkyins and B. M. Pettitt, *J. Chem. Phys.* **97**, 7656 (1992).
- ³⁰T. Imai, Y. Harano, M. Kinoshita, A. Kovalenko, and F. Hirata, *J. Chem. Phys.* **125**, 024911 (2006).
- ³¹S. Yasuda, T. Yoshidome, H. Oshima, R. Kodama, Y. Harano, and M. Kinoshita, *J. Chem. Phys.* **132**, 065105 (2010).
- ³²M. L. Connolly and *J. Appl. Crystallogr.* **16**, 548 (1983).
- ³³M. L. Connolly and *J. Am. Chem. Soc.* **107**, 1118 (1985).
- ³⁴A. D. MacKerell, Jr. D. Bashford, M. Bellott, R. L. Dunbrack, Jr. J. D. Evanseck, M. J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, D. Joseph-McCarthy, L. Kuchnir, K. Kuczera, F. T. K. Lau, C. Mattos, S. Michnick, T. Ngo, D. T. Nguyen, B. Prodhom, W. E. Reiher, III B. Roux, M. Schlenkrich, J. C. Smith, R. Stote, J. Straub, M. Watanabe, J. Wiorcikiewicz-Kuczera, D. Yin, and M. Karplus, *J. Phys. Chem. B* **102**, 3586 (1998).
- ³⁵T. Yoshidome, M. Kinoshita, S. Hirota, N. Baden, and M. Terazima, *J. Chem. Phys.* **128**, 225104 (2008).
- ³⁶Y. Harano, T. Yoshidome, and M. Kinoshita, *J. Chem. Phys.* **129**, 145103 (2008).
- ³⁷T. Yoshidome and M. Kinoshita, *Phys. Rev. E* **79**, 030905(R) (2009).
- ³⁸H. Oshima, T. Yoshidome, K. Amano, and M. Kinoshita, *J. Chem. Phys.* **131**, 205102 (2009).
- ³⁹K. Amano, T. Yoshidome, K. Oda, Y. Harano, and M. Kinoshita, *Chem. Phys. Lett.* **474**, 190 (2009).
- ⁴⁰K. Oda, R. Kodama, T. Yoshidome, M. Yamanaka, Y. Sambongi, and M. Kinoshita, *J. Chem. Phys.* **134**, 025101 (2011).
- ⁴¹S. F. Sneddon, D. J. Tobias, and C. L. Brooks III. *J. Mol. Biol.* **209**, 817 (1989).
- ⁴²P. Güntert, C. Mumenthaler, and K. Wüthrich, *J. Mol. Biol.* **273**, 283 (1997).
- ⁴³B. L. de Groot, D. M. F. van Aalten, R. M. Scheek, A. Amadei, G. Vriend, and H. J. C. Berendsen, *Proteins* **29**, 240 (1997).
- ⁴⁴I. W. Davis, W. B. Arendall 3rd D. C. Richardson, and J. S. Richardson, *Structure* **14**, 265 (2006).
- ⁴⁵B. R. Brooks, R. E. Bruccoleri, B. D. Olafson, D. J. States, S. Swaminathan, and M. Karplus, *J. Comput. Chem.* **4**, 187 (1983).
- ⁴⁶M. Feig, J. Karanicolas, and C. L. Brooks III, *J. Mol. Graphics* **22**, 377-395 (2004).
- ⁴⁷A. D. MacKerell Jr. M. Feig, and C. L. Brooks III. *J. Comput. Chem.* **25**, 1400 (2004).
- ⁴⁸M. S. Lee, F. R. Salsbury Jr., and C. L. Brooks III. *J. Chem. Phys.* **116**, 10606 (2002).
- ⁴⁹M. S. Lee, M. Feig, F. R. Salsbury, Jr., and C. L. Brooks III. *J. Comput. Chem.* **24**, 1348

(2003).

⁵⁰J. Chocholousšová and M. Feig, *J. Comput. Chem.* **27**, 719 (2006).

⁵¹Y. Zhang and J. Skolnick, *Proteins* **57**, 702 (2004).

⁵²Y. Harano and M. Kinoshita, *Biophys. J.* **89**, 2701 (2005).

⁵³M. Kinoshita, *Front. Biosci.* **14**, 3419 (2009).

⁵⁴M. Kinoshita, *Int. J. Mol. Sci.* **10**, 1064 (2009).

⁵⁵W. Kabsch and C. Sander, *Biopolymers* **22**, 2577 (1983).

List of Publications

Chapter 1, Appendix

“Characterization of Experimentally Determined Native-Structure Models of a Protein Using Energetic and Entropic Components of Free-Energy Function”,

Hirokazu Mishima, Satoshi Yasuda, Takashi Yoshidome, Hiraku Oshima, Yuichi Harano, Mitsunori Ikeguchi, and Masahiro Kinoshita,

The Journal of Physical Chemistry B, **116**, 7776-7786 (2012), ACS (American Chemical Society) Publications.

Chapter 2

“Entropic release of a big sphere from a cylindrical vessel”,

Hirikazu Mishima, Hiraku Oshima, Satoshi Yasuda, Ken-ichi Amano, and Masahiro Kinoshita,

Chemical Physics Letters, **561-62**, 159-165 (2013), Elsevier.

Chapter 3

“On the physics of multidrug efflux through a biomolecular complex”,

Hirikazu Mishima, Hiraku Oshima, Satoshi Yasuda, Ken-ichi Amano, and Masahiro Kinoshita,

The Journal of Chemical Physics, **139**, 205102(1-13) (2013), AIP (American Institute of Physics) Publishing.

Chapter 4

“Statistical Thermodynamics for Functionally Rotating Mechanism of the Multidrug Efflux Transporter AcrB”,

Hirikazu Mishima, Hiraku Oshima, Satoshi Yasuda, and Masahiro Kinoshita,

The Journal of Physical Chemistry B, **119**, 3423-3433 (2015), ACS (American Chemical Society) Publications.

Copyright and Publisher's Link

Chapter 1, Appendix

Adapted with permission from The Journal of Physical Chemistry B, **116**, 7776-7786 (2012), ACS Publications. Copyright 2015 American Chemical Society.

<http://pubs.acs.org/doi/abs/10.1021/jp301541z>

Chapter 2

This is the author's version of a work that was accepted for publication in Chemical Physics Letters. The work was published by Elsevier in Chemical Physics Letters, **561-62**, 159-165 (2013), <http://dx.doi.org/10.1016/j.cplett.2013.01.045>

Chapter 3

This is the author's version of a work that was accepted for publication in The Journal of Chemical Physics. The work was published by the American Institute of Physics in The Journal of Chemical Physics, **139**, 205102(1-13) (2013), <http://dx.doi.org/10.1063/1.4832896>

Chapter 4

Adapted with permission from The Journal of Physical Chemistry B, **119**, 3423-3433 (2015), ACS Publications. Copyright 2015 American Chemical Society.

<http://pubs.acs.org/doi/abs/10.1021/jp5120724>

List of Presentations

International Conference

Poster Presentation

Hirokazu Mishima, Hiraku Oshima, Satoshi Yasuda, Ken-ichi Amano and Masahiro Kinoshita, Entropic release of a big sphere from a cylindrical vessel, Symposium on Hydration and ATP energy, March 6-8, 2013, Akiu, Sendai, Japan.

Hirokazu Mishima, Hiraku Oshima, Satoshi Yasuda, Ken-ichi Amano and Masahiro Kinoshita, PHYSICS OF MULTIDRUG EFFLUX THROUGH A BIOMOLECULAR COMPLEX, Biophysical Society 58th Annual Meeting, February 15-19, 2014, Mosecone Center, San Francisco, California.

Domestic Conference

Oral Presentation

三嶋浩和、尾嶋拓、安田賢司、木下正弘、AcrB による薬剤の取込・排出機構に対する物理描像の構築、日本物理学会 2014 年秋季大会, 2014 年 9 月 7 日-9 月 10 日, 中部大学春日井キャンパス.

Poster Presentation

三嶋浩和、安田賢司、吉留崇、池口満徳、木下正弘、実験で得られた蛋白質天然構造モデルのキャラクタリゼーション, 第 34 回溶液化学シンポジウム, 2011 年 11 月 15 日-17 日, 名古屋大学 ES 総合館 ES ホール.

三嶋浩和、安田賢司、吉留崇、尾嶋拓、原野雄一、池口満徳、木下正弘、実験で得られた蛋白質天然構造モデルのキャラクタリゼーション, ナノ分野グランドチャレンジ研究開発 第 6 回公開シンポジウム, 2012 年 3 月 5 日-6 日, ニチイ学館 神戸ポートアイランドセンター.

Hirokazu Mishima, Hiraku Oshima, Satoshi Yasuda, Ken-ichi Amano and Masahiro Kinoshita, Entropic release of a big sphere from a cylindrical vessel, 日本生物物理学会第 50 回年会, 2012 年 9 月 22 日-24 日, 名古屋大学東山キャンパス.

三嶋浩和、尾嶋拓、安田賢司、天野健一、木下正弘、エントロピー力による円筒状容器からの大球の放出, 2012 GCOE Annual Report Meeting, 2013 年 1 月 16 日, 京都大学百周年時計台記念館国際交流ホール.

Hirokazu Mishima, Hiraku Oshima, Satoshi Yasuda, Ken-ichi Amano and Masahiro Kinoshita, On the Physics of Multidrug Efflux through a Biomolecular Complex, 日本生物物理学会第 51 回年会, 2013 年 10 月 28 日-30 日, 国立京都国際会館.

Acknowledgement

This thesis has been accomplished with support of many people. I would like to express my gratitude to them.

I would like to express my cordial gratitude to Prof. Masahiro Kinoshita at Institute of Advanced Energy, Kyoto University for his earnest guidance, proper suggestions, constructive discussions, and constant encouragement throughout my academic life. Since the enrollment to Graduate School of Energy Science, he has given me the basic knowledge of biophysics and stirred up my curiosity about life phenomena.

I am deeply grateful to Associate Prof. Takashi Yoshidome at Graduate School of Engineering, Tohoku University for his proper guidance and tremendous supports. His professional attitude on the study has motivated me.

I am sincerely grateful to Dr. Hiraku Oshima at Kinoshita's laboratory for his useful advice, constructive suggestions, and constant encouragement. He has been keeping our academic research environment in the best condition.

I would like to thank Associate Prof. Ken-ichi Amano at Graduate School of Engineering, Kyoto University for his encouragement and useful comments. He has made a great contribution to studies described in Chapters 2 and 3.

I am very thankful to Dr. Satoshi Yasuda at Kinoshita's laboratory for his helpful comments and continuing supports. Especially, he has made a significant contribution to the study described in Chapter 4.

I would like to thank Associate Prof. Mitstumori Ikeguchi at Graduate School of Nanobioscience, Yokohama City University for useful suggestions and technical comments. His broad and deep knowledge has been most helpful to my research.

I am very thankful to Prof. Takashi Morii and Prof. Masato Katahira at Institute of Advanced Energy, Kyoto University for the critical reading of my doctoral thesis and useful comments. Their experimental knowledge and perspective has benefited to the improvement of the thesis.

I appreciate Dr. Tomohiko Hayashi, Mr. Ryota Kodama, Mr. Shota Murakami, and Mr. Yuta Kajiwara at Kinoshita's laboratory for their encouragements and helpful advices.

I have received education from Associate Prof. Takashi Yoshidome, Dr. Hiraku Oshima and Dr. Satoshi Yasuda. They strengthened my computational skills, information-gathering ability, and physics knowledge.

This research was supported by Grants-in-Aid for Scientific Research on Innovative Areas (No. 20118004) and that on (B) (No. 22300100 and No. 25291035) from the Ministry of Education, Culture, Sports, Science and Technology of Japan, by Kyoto University Global Center of Excellence of Energy Science, and by Grants-in-Aid for JSPS (Japan Society for

the Promotion of Science) fellows.

I have received big help and support from many people. Although I cannot list up all the people here, I convey thankfulness to them. Finally, I thank my family for supporting my mind.

February, 2015
Hirokazu Mishima