

Binding of an RNA aptamer and a partial peptide of a prion protein: crucial importance of water entropy in molecular recognition

Tomohiko Hayashi, Hiraku Oshima, Tsukasa Mashima, Takashi Nagata, Masato Katahira and Masahiro Kinoshita*

Institute of Advanced Energy, Kyoto University, Uji, Kyoto 611-0011, Japan

Received February 10, 2014; Revised March 31, 2014; Accepted April 18, 2014

ABSTRACT

It is a central issue to elucidate the new type of molecular recognition accompanied by a global structural change of a molecule upon binding to its targets. Here we investigate the driving force for the binding of R12 (a ribonucleic acid aptamer) and P16 (a partial peptide of a prion protein) during which P16 exhibits the global structural change. We calculate changes in thermodynamic quantities upon the R12–P16 binding using a statistical-mechanical approach combined with molecular models for water which is currently best suited to studies on hydration of biomolecules. The binding is driven by a water-entropy gain originating primarily from an increase in the total volume available to the translational displacement of water molecules in the system. The energy decrease due to the gain of R12–P16 attractive (van der Waals and electrostatic) interactions is almost canceled out by the energy increase related to the loss of R12–water and P16–water attractive interactions. We can explain the general experimental result that stacking of flat moieties, hydrogen bonding and molecular-shape and electrostatic complementarities are frequently observed in the complexes. It is argued that the water-entropy gain is largely influenced by the geometric characteristics (overall shapes, sizes and detailed polyatomic structures) of the biomolecules.

INTRODUCTION

An intrinsically disordered protein is characterized by a lack of stable tertiary structure when it is present as an isolated polypeptide chain in aqueous solution under the physiological condition. Its folding is coupled with binding to its targets in the sense that it constructs a well-defined structure only after the binding is implemented (1). An aptamer, which is a folded single-stranded nucleic acid, pos-

sesses structural plasticity and binds to its targets by changing its structure in accordance with the target structures (2). The structure of the aptamer remains unchanged in some cases, and disordered, flexible portions of its targets become structured upon binding. This type of molecular recognition accompanied by a global structural change of a molecule upon binding to its targets appears to be substantially different from that described by the ‘lock and key’ (3) or ‘induced-fit’ (4) binding model. Elucidation of these molecular-recognition mechanisms is a central issue in biophysics, biochemistry and structural biology. In our opinion, all types of molecular recognitions share the mechanism based on the same physicochemical origin and can be explained within the same theoretical framework in a unified manner. The present study provides a remarkable first step toward achieving such elucidation by considering a ribonucleic acid (RNA) aptamer (5) as an important example.

Aptamers are RNAs and deoxyribonucleic acids originating from *in vitro* selection experiments which optimize the nucleic acids for high-affinity binding to their targets by starting from random sequence libraries (2,6,7). Up to now, many aptamers have been developed against a wide variety of targets such as small molecules (8), small interfering RNA (siRNAs) (9), peptides (10), sugars (11,12), proteins (13–17) and cells (18,19). Aptamers have many potential uses as effective therapeutics (20), drug delivery agents (21), purification agents for the therapeutic antibodies (16) and molecular probes (22). Elucidation of the molecular-recognition mechanism common in all of the aptamers presents much challenge not only from the scientific viewpoint but also for medical and technological applications. The RNA aptamer considered in the present study, r(GGAGGAGGAGGA) (R12), binds to and stabilizes a normal cellular form of a prion protein (PrP^C) (23), which is expected to prevent prion diseases (5). R12 takes a unique quadruplex structure forming a dimer (23). In this case, two portions (P1 and P16) of the N-terminal half of PrP^C are disordered and flexible, and they form well-defined structures in accordance with R12 upon binding.

*To whom correspondence should be addressed. Tel: +81 774 38 3503; Fax: +81 774 38 4695; Email: kinoshit@iae.kyoto-u.ac.jp

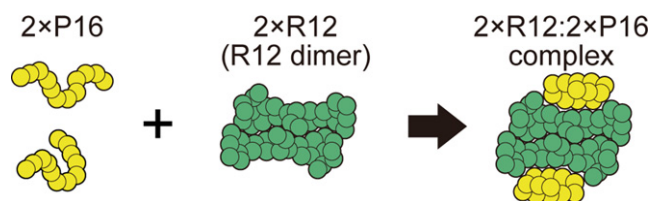


Figure 1. Binding of two P16s to an R12 dimer. The complex formed is denoted by '2xR12:2xP16'. The water molecules are omitted here.

Mashima *et al.* (5) performed experiments for binding of P16 (Gly-Gln-Trp-Asn-Lys-Pro-Ser-Lys-Pro-Lys-Thr-Asn) to R12, and we consider this fundamental binding process. In this process, two P16s bind to the R12 molecules forming the dimer, respectively, as illustrated in Figure 1. Since the upper and lower halves of the 2xR12:2xP16 complex share essentially the same structure, the theoretical investigation can be made for the constituent process in which a P16 binds to one of the R12 molecules, forming an R12:P16 complex.

The clue to the molecular-recognition mechanism is the driving force for the binding. The force has been investigated primarily from the viewpoint of structural biology on the basis of the three-dimensional (3D) structure of the aptamer–target complex. A prevailing view is that an aptamer binds to its targets via aptamer–target electrostatic attractive interactions due to the contact of groups with positive and negative charges, that is, the electrostatic complementarity. Actually, in several aptamer–target complexes, negatively charged phosphate backbones in the aptamer are in contact with positively charged moieties of the targets. For example, Toggle-25t (an RNA aptamer) binds to the positively charged arginine-rich surface of thrombin (15). In the R12:P16 complex, the contact of phosphate groups of R12 with lysine residues of P16 is observed as illustrated in Figure 2 (5) (Figures 2, 4 and 5 are drawn using the Visual Molecular Dynamics (VMD)(24)). However, there is an experimental result that raises a doubt with respect to this view: in the complex of Apt8 (an RNA aptamer) binding to the Fc fragment of Human IgG1 (hFc1), the aptamer-bound area of hFc1 consists of a less-positively charged surface (16,25). In the case of the R12–P16 binding, a π – π stacking interaction between a guanine nucleotide in R12 and an indole ring of tryptophan in P16 (see Figure 2) has been proposed as another important driving force (5). Taken together, the structural data have shown that precise stacking of flat moieties, specific hydrogen bonding and electrostatic complementarity frequently occur in the complexes (2).

The determination of the complex structure is an indispensable starting step toward the clarification of the driving force for the binding. However, it is not sufficient and to be followed by studies from the viewpoint of statistical thermodynamics of hydration. This is because the binding occurs in aqueous solution and water should play imperative roles. There are two principal points to be taken into account as illustrated in Figure 3. In the figure, the two solute molecules are rigid and their structures remain unchanged upon binding in (a), whereas one of them is structureless and becomes structured upon binding in (b). Firstly, the

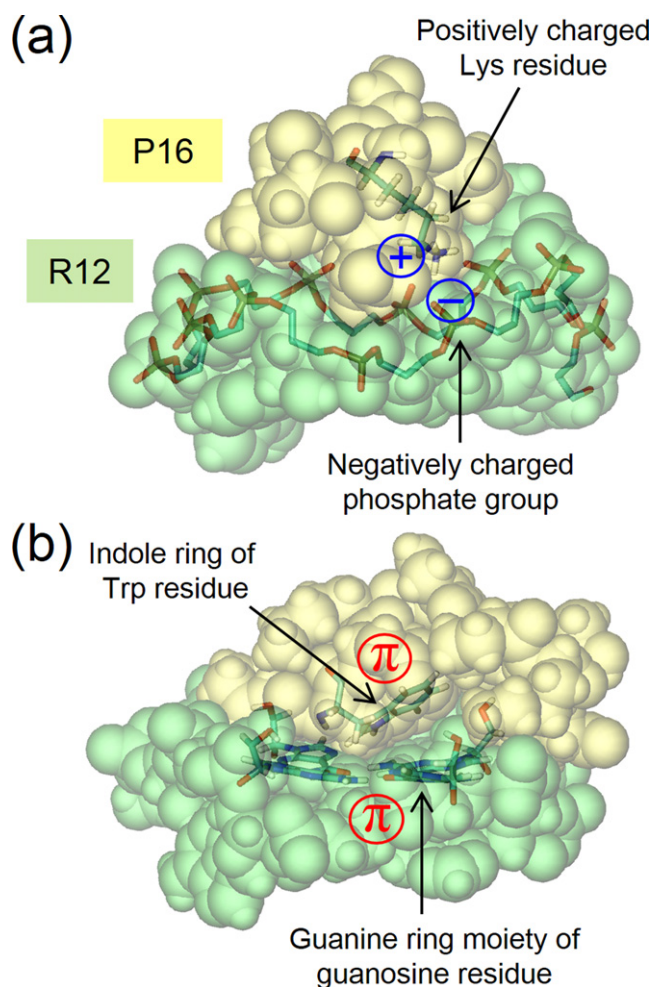


Figure 2. Structural images of the upper half of the R12:P16 complex. It is taken from an NMR structure (model 1) constructed as described in 'Structure Modeling for R12:P16 Complex'. R12 and P16 are drawn on the upper and lower sides, respectively, using the space-filled model. (a) Contact of a phosphate group in R12 with a lysine residue in P16. The phosphate backbone (C5'-C4'-C3'-PO⁴⁻) and Lys8 are depicted by the licorice model. (b) π – π stacking of a guanine nucleotide in R12 and an indole ring of tryptophan in P16. The quadruplex tetrad plane containing G:G:G:G (G denotes a guanine nucleotide) and Trp3 are depicted by the licorice model. The water molecules are omitted here. This figure is drawn using the VMD

two molecules are hydrated, and the regions that become unexposed to water upon binding undergo dehydrations. When a solute region is hydrated, unless the region is predominantly nonpolar, the water structure near the region is significantly perturbed due to solute–water electrostatic and van der Waals attractive interactions. Upon dehydration, the attractive interactions are lost, giving rise to an energy increase. At the same time, the water structure is reorganized and the water–water electrostatic and van der Waals attractive interactions are recovered, leading to an energy decrease. The energy increase is approximately two times larger than the energy decrease (26,27) and the net change in energy is positive and considerably large, which is referred to as 'energetic dehydration effect' hereafter. Since a hydrogen bond is usually expressed as an electrostatic at-

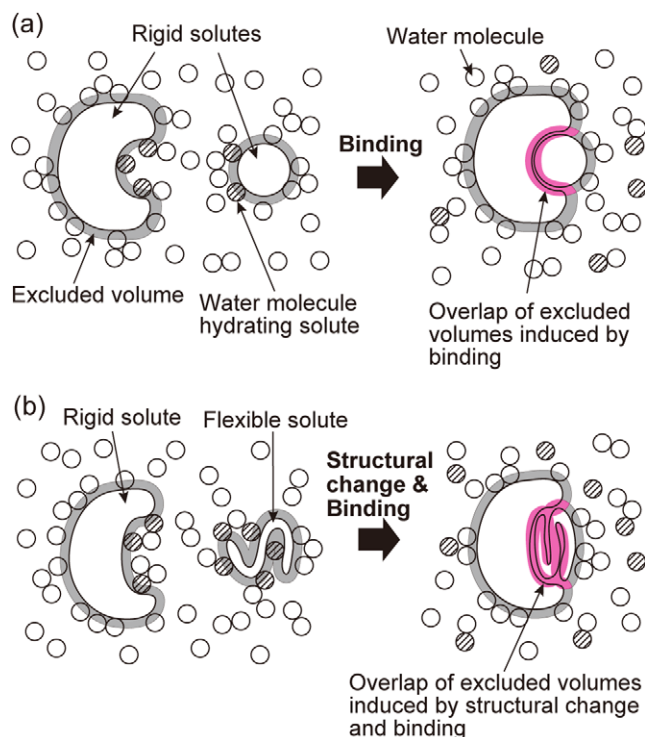


Figure 3. Cartoons illustrating the dehydration and the overlap of the excluded spaces accompanying the binding of two solute molecules. (a) Case where they are both rigid and their structures remain unchanged. (b) Case where one of them is flexible (i.e. structureless) and becomes structured after the binding.

tractive interaction between hydrogen and oxygen or nitrogen, we consider that solute–solute, solute–water and water–water hydrogen bonds can be regarded as electrostatic attractive interactions. Secondly, each molecule generates an excluded space where the centers of water molecules cannot enter. The volume of the excluded space is referred to as the excluded volume (EV). Upon binding, the two excluded spaces overlap and the total EV decreases by the overlapped volume as shown in Figure 3a, which is followed by a corresponding increase in the total volume available to the translational displacement of water molecules in the system. Hence, the binding leads to a large gain of the configurational entropy of water. When the structure of one of the molecules changes to a more compact one upon binding as in the case of Figure 3b, the water-entropy gain becomes considerably larger than in the case of Figure 3a, because such a structural change results in a larger decrease in the total EV. We have shown that the physical factors described above, which can be categorized as ‘entropic EV effect’, play essential roles in a variety of biological self-assembly processes (28–32).

In the present study, we revisit the R12–P16 binding from the viewpoint of statistical thermodynamics of hydration. Since P16 does not take a particular tertiary structure when it is isolated, the binding is accompanied by a global structural change of P16. We remark that obtaining only microscopic information using a computer (e.g. molecular dynamics) simulation does not yield the clarification of the

driving force. Calculating the changes of thermodynamic quantities upon binding is vitally important. We calculate the binding free energy using the molecular mechanics, hybrid method (30–37) in which the angle-dependent integral equation theory (ADIET) (38–42) applied to a multipolar water model (38,39) is combined with the morphometric approach (MA) (43,44) and 3D reference interaction site model (3D-RISM) theory (45–48). The ADIET and the 3D-RISM theory are statistical-mechanical theories for molecular liquids and hydration of solutes. The combination with the MA has recently been developed by us so that the ADIET can be applied to a complex biomolecule with drastic reduction in the computational load. The hybrid method and the 3D-RISM theory, which have been quite successful in analyzing a variety of self-assembly processes (27–32,49–52), are best suited to calculations of the hydration energy and entropy, respectively.

Our important findings are as follows. The binding is driven by a large gain of the water entropy; and the energy decrease due to the gain of R12–P16 attractive interactions is almost canceled out by the energy increase originating from the energetic dehydration effect upon binding. It should be noted, however, that the gain of R12–P16 attractive interactions is necessitated for compensating for the energetic dehydration effect. We point out that the general experimental observations manifesting precise stacking of flat moieties, electrostatic complementarity, including specific hydrogen bonding, and molecular shape complementarity can reasonably be interpreted. Further, we argue that the geometric characteristics (overall shapes, sizes and detailed polyatomic structures) of the solute molecules become the most important factors when the entropic EV effect dominates. This argument sheds new light on the elucidation of all types of molecular-recognition mechanisms within the same theoretical framework in a unified manner.

MODEL AND THEORY

Free-energy function

For a solute molecule immersed in water at infinite dilution, we define the free-energy function G expressed as

$$G = E_C - TS_C + \mu_H, \quad (1)$$

where E_C and S_C are the conformational (intramolecular) energy and entropy of the solute molecule, respectively, T is the absolute temperature and μ_H is the hydration free energy defined as the excess chemical potential of the solute. We note that μ_H is independent of the solute insertion condition, isobaric or isochoric (53), and we consider isochoric condition that is much more convenient in a theoretical treatment. Using the relation,

$$\mu_H = \epsilon_{VH} - TS_{VH}, \quad (2)$$

where the subscript ‘VH’ denotes the hydration under isochoric condition, ϵ_{VH} and S_{VH} are the hydration energy and entropy, respectively, Equation (1) is rewritten as

$$G = E_C - TS_C + \epsilon_{VH} - TS_{VH}. \quad (3)$$

Here, ϵ_{VH} represents the solute–water interaction energy generated plus the energy change due to the structural re-

organization of water upon solute insertion and S_{VH} represents the change in the water entropy upon solute insertion. G is also independent of the solute insertion condition.

E_C is calculated on the basis of a molecular mechanical potential and thereby decomposed into the bonded, non-electrostatic and electrostatic parts as

$$E_C = E_B + E_{\text{LJ}} + E_{\text{ES}}, \quad (4)$$

where E_B is the bond-type energy comprising the bond-stretching, angle-bending and torsion-angle energies, E_{LJ} is the Lennard–Jones (LJ) interaction energy and E_{ES} is the electrostatic interaction energy. We also decompose ϵ_{VH} as

$$\epsilon_{\text{VH}} = \epsilon_{\text{VH,LJ}} + \epsilon_{\text{VH,ES}}, \quad (5)$$

where $\epsilon_{\text{VH,LJ}}$ and $\epsilon_{\text{VH,ES}}$ are the nonelectrostatic and electrostatic contributions to ϵ_{VH} , respectively. The decomposition is made in the following manner. First, we calculate the hydration energy of a hypothetical solute molecule whose partial charges are all switched to zero, $\epsilon_{\text{VH,LJ}}$. Then, we obtain $\epsilon_{\text{VH,ES}}$ from $\epsilon_{\text{VH,ES}} = \epsilon_{\text{VH}} - \epsilon_{\text{VH,LJ}}$.

Substituting Equations (4) and (5) into Equation (3) yields

$$G = E_{\text{total}} - TS_C - TS_{\text{VH}}, \quad (6a)$$

$$E_{\text{total}} = E_B + (E_{\text{LJ}} + \epsilon_{\text{VH,LJ}}) + (E_{\text{ES}} + \epsilon_{\text{VH,ES}}). \quad (6b)$$

E_{total} , which is referred to as the total energy hereafter, consists of E_B and the LJ and electrostatic components denoted by $E_{\text{LJ}} + \epsilon_{\text{VH,LJ}}$ and $E_{\text{ES}} + \epsilon_{\text{VH,ES}}$, respectively.

Paths for calculating binding free energy

The free-energy function and its constituents defined above are applicable to R12, P16 and the R12:P16 complex. Figure 4 shows a schematic representation of the paths considered for calculating the binding free energy, the free-energy change upon the R12–P16 binding. In path I, R12 and P16 in the complex are simply separated with no structural changes, and the resulting structures are employed as the isolated molecules. That is, it is assumed that both of R12 and P16 isolated possess unique structures with no structural fluctuations. In the real system, isolated P16 takes an ensemble of rather extended structures and undergoes a global structural change upon binding though R12 exhibits almost no change in its structure (5,23). This binding process is mimicked by path II. Path III represents the structural change of P16 from the unstructured state (P16_{random coils}: it is assumed to be a set of random coils in the present study) to the compact structure taken from the complex (P16_{compact}).

Let $\Delta^M X$ ($M = \text{I, II, III}$) denote the change in X for path M . The binding free energies, which can be considered for paths I and II, are given by $\Delta^I G$ and $\Delta^{\text{II}} G$, respectively. They are formally expressed as

$$\begin{aligned} \Delta^I G &= G(\text{R12} : \text{P16}) - \{G(\text{R12}) + G(\text{P16}_{\text{compact}})\} \\ &= \Delta^I(E_{\text{LJ}} + \epsilon_{\text{VH,LJ}}) + \Delta^I(E_{\text{ES}} + \epsilon_{\text{VH,ES}}) - T\Delta^I S_{\text{VH}}, \quad (7) \end{aligned}$$

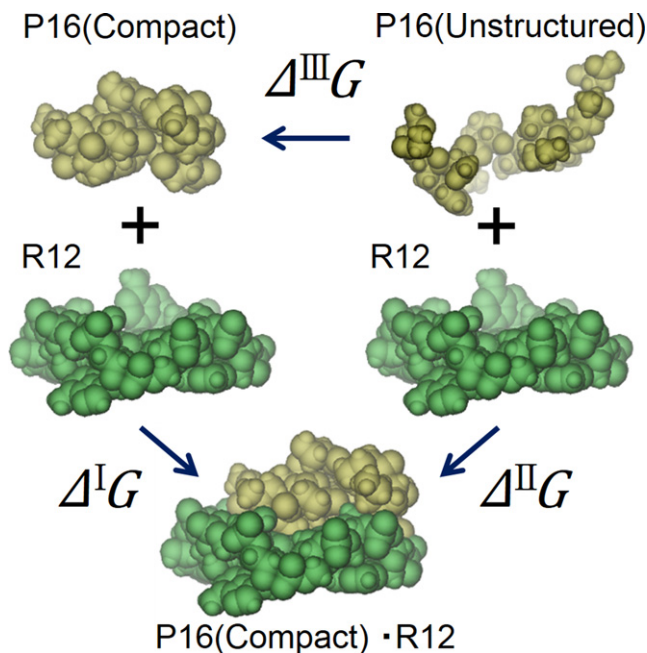


Figure 4. Schematic representation of the paths considered for calculating the binding free energy, the free-energy change upon the R12–P16 binding. The R12:P16 complex indicated by ‘P16(Compact)•R12’ is taken from an NMR structure (model 1) constructed as described in ‘Structure Modeling for R12:P16 Complex’. R12 and P16 are drawn on the upper and lower sides, respectively, using the space-filled model. $\Delta^I G$, $\Delta^{\text{II}} G$ and $\Delta^{\text{III}} G$ denote changes in the free-energy function along paths I, II and III, respectively. The water molecules are omitted here. This figure is drawn using the VMD.

$$\begin{aligned} \Delta^{\text{II}} G &= G(\text{R12} : \text{P16}) - \{G(\text{R12}) + G(\text{P16}_{\text{random coils}})\} \\ &= \Delta^{\text{II}} E_B + \Delta^{\text{II}}(E_{\text{LJ}} + \epsilon_{\text{VH,LJ}}) + \Delta^{\text{II}}(E_{\text{ES}} + \epsilon_{\text{VH,ES}}) \\ &\quad - T\Delta^{\text{II}} S_C - T\Delta^{\text{II}} S_{\text{VH}}. \quad (8) \end{aligned}$$

$G(\text{R12})$, for example, is the free-energy function for R12. $\Delta^I E_{\text{ES}}$, $\Delta^I \epsilon_{\text{VH,LJ}}$ and $\Delta^I \epsilon_{\text{VH,ES}}$, for example, originate from electrostatic attractive interactions between R12 and P16, from the energetic dehydration effect due to the loss of R12–water and P16–water van der Waals attractive interactions and from the energetic dehydration effect due to the loss of R12–water and P16–water electrostatic attractive interactions, respectively, and $\Delta^I S_{\text{VH}}$ represents the water-entropy gain in the binding process of path I. We remark that about half of the energy increase arising from the loss of R12–water and P16–water attractive interactions is canceled out by the energy decrease due to the structural reorganization of water upon binding (see the fourth paragraph of the Introduction section). In path I, $\Delta^I E_B = 0$ and $\Delta^I S_C = 0$. The free-energy change due to the global structural change of P16, $\Delta^{\text{III}} G$, is given by

$$\begin{aligned} \Delta^{\text{III}} G &= G(\text{P16}_{\text{compact}}) - G(\text{P16}_{\text{random coils}}). \\ &= \Delta^{\text{III}} E_B + \Delta^{\text{III}}(E_{\text{LJ}} + \epsilon_{\text{VH,LJ}}) + \Delta^{\text{III}}(E_{\text{ES}} + \epsilon_{\text{VH,ES}}) \\ &\quad - T\Delta^{\text{III}} S_C - T\Delta^{\text{III}} S_{\text{VH}}. \quad (9) \end{aligned}$$

It should be noted that $\Delta^{\text{II}} G = \Delta^I G + \Delta^{\text{III}} G$ and $\Delta^{\text{II}} S_C = \Delta^{\text{III}} S_C$.

Since P16_{random coils} is more stable than P16_{compact}, $\Delta^{\text{III}}G > 0$, $\Delta^{\text{II}}G < 0$ and it follows that $\Delta^{\text{I}}G < 0$ and $\Delta^{\text{I}}G > \Delta^{\text{II}}G$. $\Delta^{\text{II}}S_{\text{C}} < 0$ and $-T\Delta^{\text{II}}S_{\text{C}}$ ($= -T\Delta^{\text{III}}S_{\text{C}}$) cannot be a driving force for the binding. $\Delta^{\text{II}}E_{\text{B}}$ ($= \Delta^{\text{III}}E_{\text{B}}$) should have a relatively small contribution to the binding free energy. We present the theoretical results for $\Delta^{\text{II}}G$ and $\Delta^{\text{III}}G$ by omitting $-T\Delta^{\text{II}}S_{\text{C}}$ for the following reasons: the calculation of conformational entropy is a formidable task and modeling the unstructured state of isolated P16 is accompanied by significant uncertainty. However, the validity of our theoretical method can be checked in the following manner. The dissociation constant for the R12–P16 binding was determined by Mashima *et al.* (5) through a filter binding assay experiment by assuming that P16 bound to an R12 monomer. However, R12 is present as a dimer and two P16s bind to the R12 molecules forming the dimer, respectively (see Figure 1). In the present study, we recalculate the dissociation constant (the recalculated one is denoted by k_{D}) as described in Section A of Supplementary Appendices and obtain an approximate value of the binding free energy corresponding to $\Delta^{\text{II}}G$ as $-RT\ln(k_{\text{D}})$ that is denoted by $\Delta^{\text{II}}G_{\text{experimental}}$ (R is the gas constant). The theoretical value of $\Delta^{\text{II}}G$ in which $-T\Delta^{\text{II}}S_{\text{C}}$ is omitted, $\Delta^{\text{II}}G_{\text{theoretical}}$, must be sufficiently lower than $\Delta^{\text{II}}G_{\text{experimental}}$. This criterion is certainly satisfied as shown in the Results and Discussion section ($-T\Delta^{\text{II}}S_{\text{C}}$ is roughly estimated and a discussion is also made on the quantitative aspect of the theoretical results).

Our principal concern is to examine the signs and magnitudes of $\Delta^M(E_{\text{LJ}} + \epsilon_{\text{VH,LJ}})$, $\Delta^M(E_{\text{ES}} + \epsilon_{\text{VH,ES}})$ and $-T\Delta^M S_{\text{VH}}$ ($M = \text{I, II}$) which enable us to specify the driving force. Those for path I may be better suited to quantitative analyses because the binding free energy for path I given by Equation (7) does not include changes in the bond-type energy and the solute conformational entropy.

Calculation of hydration entropy

The ADIET (38–42) applied to a multipolar model for water (38,39) has been quite successful in studies on hydrophobic and hydrophilic hydrations. In the theory, the orientational dependency of the water–water and solute–water interaction potentials and correlations is explicitly taken into consideration. In most of our studies, a water molecule is modeled as a hard sphere with diameter $d_{\text{S}} = 0.28$ nm in which a point dipole and a point quadrupole of tetrahedral symmetry are embedded. The effect of the molecular polarizability is taken into account using the self-consistent mean field (SCMF) theory (38,39). 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 is about 1.42 times larger than the bare gas-phase dipole moment. The dielectric constant of bulk water, which is a good measure of the validity of a theory, is calculated to be ~ 83 that is in good agreement with the experimental value ~ 78 . Further, the theory is capable of reproducing the solubility of methane featuring the minimum at about 350 K (41). The hydration free energy of a spherical, nonpolar solute calculated is in perfect agreement with that obtained by computer simulations using more popular water models such as the trans-

ferable intermolecular potential four point (TIP4P) and the extended single point charge (SPC/E) (41). A more detailed description of the ADIET is given in Section B of Supplementary Appendices.

It is not straightforward to apply the ADIET to a large, complex solute with polyatomic structure due to the mathematical complexity. Fortunately, as far as the hydration entropy S_{VH} is concerned, it has been shown to be rather insensitive to the solute–water interaction potentials. For example, Imai *et al.* (47) examined S_{VH} for a total of eight peptides and proteins using the 3D-RISM theory combined with the all-atom potentials and found the following: even when the protein–water electrostatic interactions, which are quite strong, are completely shut off and only the LJ interactions are retained, S_{VH} changes merely by less than 5%. Thus, only the geometric characteristics are essential in the calculation of S_{VH} , and a solute molecule can be modeled as a set of fused, neutral hard spheres. The diameter of an atom in a solute molecule is set at the corresponding σ -value for the LJ potential. As a consequence, the MA (43,44) becomes a very powerful tool. We employ the hybrid method in which the ADIET is combined with the MA for calculating S_{VH} .

The idea of the MA is to express S_{VH} by the linear combination of only four geometric measures of a solute molecule:

$$S_{\text{VH}}/k_{\text{B}} = C_1 V_{\text{ex}} + C_2 A + C_3 X + C_4 Y. \quad (10)$$

Here, Equation (10) is referred to as the morphometric form, V_{ex} is the EV generated by the solute molecule, A is the water-accessible surface area, X and Y are the integrated mean and Gaussian curvatures of the water-accessible surface, respectively, and k_{B} is the Boltzmann constant. The water-accessible surface is the surface that is accessible to the centers of water molecules. The EV, as defined in the previous section, is the volume that is enclosed by this surface. We note that C_1 is completely independent of the solute–water interaction potential. Though S_{VH} is influenced by all the four terms, $C_1 V_{\text{ex}}$ is the principal term at normal temperature and pressure. This is the reason for the fair insensitivity of S_{VH} to the solute–water interaction potential. The contribution from the water molecules near the solute molecule is represented by the other three terms. In the MA, the solute shape enters S_{VH} 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_{VH} of hard-sphere solutes immersed in our model water. The ADIET is employed in the calculation. The four coefficients are determined by the least square fitting applied to the following equation (i.e. Equation 10 applied to hard-sphere solutes):

$$\begin{aligned} S_{\text{VH}}/k_{\text{B}} = \\ C_1(4\pi R^3/3) + C_2(4\pi R^2) + C_3(4\pi R) + C_4(4\pi), \quad (11) \\ R = (d_{\text{U}} + d_{\text{S}})/2. \end{aligned}$$

Here, d_{U} denotes the hard-sphere diameter and sufficiently many different values of d_{U} are considered. T is set at 298 K, and the number density of bulk water ρ_{S} is taken to be that of real water on the saturation curve, $\rho_{\text{S}} = 0.0333 \text{ \AA}^{-3}$. Once the four coefficients are determined, S_{VH} of the solute

molecule with a prescribed structure is obtained from Equation (10) only if the four geometric measures are calculated.

The high reliability of the hybrid method in calculating S_{VH} has been demonstrated in the following examples wherein the water-entropy effect is treated as the key factor: quantitative reproduction of the experimentally measured changes in thermodynamic quantities upon apoplastocyanin (apoPC) folding (30); elucidation of the microscopic mechanisms of pressure (33) and cold (34,35) denaturing of a protein; and proposal of a reliable measure of the thermodynamic stability of a protein (36,37).

Calculation of hydration energy

The hydration energy ϵ_{VH} , which is largely dependent on the solute–water interaction potential, cannot be calculated by the hybrid method described above. We employ the 3D-RISM theory (45–48) for the calculation of ϵ_{VH} . The LJ potential parameters and partial charges are assigned to the atoms constituting the solute molecules. The 3D-RISM theory has successfully been applied to important problems in biological systems such as the hydration properties of peptides and proteins (27,47), receptor–ligand binding processes (49,50) and association of protein molecules (51,52). However, the theory is not very good at elucidating the hydration of a hydrophobic molecule. For instance, the theory underestimates the absolute value of the water-entropy change upon protein folding though the hydration-energy change remains quantitatively reliable (27). Taken together, the 3D-RISM theory and the hybrid method are best suited to calculations of ϵ_{VH} and S_{VH} , respectively.

A description of the 3D-RISM theory is provided in Section C of Supplementary Appendices. First, the site–site correlation functions for bulk water are calculated using the dielectrically consistent RISM (DRISM) theory (54,55) coupled with the Kovalenko–Hirata (KH) closure equation (46). The site–site intermolecular potentials, water density ($\rho_{\text{S}} = 0.0333 \text{ \AA}^{-3}$) and absolute temperature ($T = 298 \text{ K}$) are served as the input data. The SPC/E model (56) is employed for water with a correction in terms of the LJ potential parameters for the hydrogen sites ($\sigma = 0.654 \text{ \AA}$, $\epsilon = 0.0155 \text{ kcal/mol}$). The water–solute correlation functions are then obtained for a solute molecule with a prescribed structure by solving the 3D-RISM/KH equations. The site–site correlation functions for bulk water and the water–solute interaction potentials form the input data. The LJ potential parameters and partial charges for the solute atoms are taken from the standard Amber99SB force field (57). The solution is performed on a 3D cubic grid. The grid spacing (Δx , Δy and Δz) is set at 0.5 \AA , 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 result obtained to be identical within convergence tolerance. The hydration free energy μ_{H} is calculated using the 3D extension of the Singer–Chandler formula adapted to the KH closure equation (46,58). Finally, ϵ_{VH} is obtained from $\epsilon_{\text{VH}} = \mu_{\text{H}} + TS_{\text{VH}}$, where S_{VH} is numerically evaluated as the temperature derivative of μ_{H} .

Structure modeling of R12:P16 complex

The ten 3D structures of the $2 \times \text{R12} : 2 \times \text{P16}$ complex were taken from the Protein Data Bank (PDB code: 2RSK). They were determined by Mashima *et al.* (5) using a simulated annealing protocol combined with the interatomic distance and dihedral angle constraints obtained from the nuclear magnetic resonance (NMR) experiments. In the simulated annealing protocol, the σ -values in the LJ potentials of the Amber99SB force field (57) were multiplied by 0.7 to widely explore the structural space for the complex.

We employ an advanced molecular mechanics refinement procedure using the Amber99SB force field with the full σ -values. This procedure, which utilizes the generalized Born and surface area (GB/SA) model solvent (59,60), is combined with the interatomic distance constraints obtained from the nuclear Overhauser effect spectroscopy (NOESY). NOESY intensities are taken into account as the limits of interatomic distances. These limits are about 1.4 ($= 1/0.7$) times larger than those employed by Mashima *et al.* in the previous study (5). The ten structures are refined with 100 steps of the steepest-descent energy minimization followed by the conjugated gradient minimization which is continued until the root mean square forces acting on the complex atoms become weaker than $1.0 \times 10^{-4} \text{ (kcal/mol)/\AA}$. All the calculations are carried out using the AMBER12 molecular dynamics package (61).

Although the effect of water is further taken into account in the present refinement procedure, the important structural features of the 3D structures determined by Mashima *et al.* (5) are conserved quite well. Example common features are the stacking between a guanine nucleotide in R12 and an indole ring of tryptophan in P16 and the contact of phosphate groups in R12 with lysine residues in P16 (see Figure 2).

Since the upper and lower halves of the $2 \times \text{R12} : 2 \times \text{P16}$ complex share almost the same 3D structure, the upper half is taken from each complex structure thus constructed and used as the structure of the R12:P16 complex. We calculate the free-energy function where the conformational entropy is set at zero for each model structure of the R12:P16 complex. The ten models obtained are ranked in ascending order with respect to this free-energy function, and the first to fifth models (models 1 through 5) are chosen for the present study. The coordinates of the ensemble of five structures thus obtained for the $2 \times \text{R12} : 2 \times \text{P16}$ complex have been deposited in the Protein Data Bank with the accession code 2RU7.

Structure modeling of P16 in unstructured state

The unstructured state of P16 is modeled as a set of random coils. First, P16 is taken from the R12:P16 complex and its structure is refined with 100 steps of the steepest-descent energy minimization followed by the conjugated gradient minimization with the GB/SA model solvent which is continued until the root mean square forces acting on the P16 atoms become weaker than $1.0 \times 10^{-4} \text{ (kcal/mol)/\AA}$. Then, starting from the structure thus obtained, we assign random numbers to the dihedral angles for the backbone chain, Φ and Ψ . The random numbers are limited to the range

from -180° to -30° for Φ and from -180° to -150° and from -90° to 180° for Ψ which correspond to the allowed regions in the Ramachandran map. For glycine, Φ is distributed from -180° to -30° and from 30° to 180° , and Ψ is from -180° to 180° . For proline, Φ is set at -65° and Ψ is set at 180° . The 200 models, which are obtained in this way, are refined in accordance with the minimization techniques described above. The models that give divergently high LJ potential energies due to the overlap of the constituent atoms are then excluded, and the remaining 103 models are employed as the random coils. All calculations are carried out using the AMBER12 molecular dynamics package (61). The average value of the free-energy function for these models is regarded as $G(\text{P16}_{\text{random coils}})$.

RESULTS AND DISCUSSION

Gain of water entropy upon binding: driving force

Table 1 gives the binding free energy and its two components, total-energy change and contribution from the water-entropy gain, upon the R12–P16 binding in path I (see Figure 4). The binding free energy, which is in the range from -47.05 to -30.31 kcal/mol, is governed by the contribution from the water-entropy gain for all the five models of the complex structure. The total-energy change is positive and against the binding for models 1, 3 and 4. It is negative and in favor of the binding for models 2 and 5, but the absolute value of the contribution from the water-entropy gain is roughly an order of magnitude larger than that of the total-energy change. We can conclude that the binding is driven by the water-entropy gain.

Table 2 gives the binding free energy and its two components upon the R12–P16 binding in path II. Unlike in path I, the structural change of P16 accompanying the binding is taken into account in path II. The binding free energy is in the range from -23.75 to -4.59 kcal/mol. The total-energy change takes a considerably large, positive value for all the five models of the complex structure, and it strongly opposes the binding. However, the water-entropy gain surpasses the increase in the total energy and drives the binding. Physicochemical interpretations of the results displayed in the two tables are provided in later sections. nullnull

Comparison between theoretical and experimental values of binding free energy

We compare the binding free energy calculated for path II, $\Delta^{\text{II}}G_{\text{theoretical}}$ ($\Delta^{\text{II}}G$ in Table 2), with the experimentally determined value, $\Delta^{\text{II}}G_{\text{experimental}}$. $\Delta^{\text{II}}G_{\text{experimental}}$ can be obtained as $-RT\ln(k_{\text{D}})$ where k_{D} is the dissociation constant for the R12–P16 binding: $\Delta^{\text{II}}G_{\text{experimental}} = -6.45$ kcal/mol (see Section A of Supplementary Appendices). In a strict sense, $\Delta^{\text{II}}G_{\text{theoretical}}$ cannot be in complete agreement with $\Delta^{\text{II}}G_{\text{experimental}}$. This is because for $\Delta^{\text{II}}G_{\text{experimental}} = -RT\ln(k_{\text{D}})$ the standard state is 1 mol/l and the activity coefficient is set at unity while in the theoretical calculation the standard state is the infinite dilution (62). Nevertheless, the two values should be close to each other. As stated above, $\Delta^{\text{II}}G_{\text{theoretical}}$ in which $-T\Delta^{\text{II}}S_{\text{C}}$ is omitted must be sufficiently lower than $\Delta^{\text{II}}G_{\text{experimental}}$. This requirement is certainly met except in the case of model 4.

Here, it is worthwhile to roughly estimate $-T\Delta^{\text{II}}S_{\text{C}}$. Since R12 undergoes essentially no structural change upon the R12–P16 binding, $\Delta^{\text{II}}S_{\text{C}}$ is almost equal to the conformational-entropy loss for P16 caused by the transition from the random-coil state to the compact structure, $\Delta^{\text{III}}S_{\text{C}}$. Fitter (63) studied the conformational-entropy loss upon protein folding using neutron spectroscopy. He estimated the temperature dependence of the radius parameter r representing a length scale within which each residue can freely move. The loss is given by $-3k_{\text{B}}N_{\text{r}}\ln(r_{\text{u}}/r_{\text{f}})$ where r_{u} and r_{f} are the radius parameters for the unfolded and folded states, respectively, and N_{r} is the number of residues. The estimation of the radius parameters was made at three temperatures: 303, 323 and 343 K. In order to obtain the loss at 298 K, we first performed the linear fitting of the temperature dependence of r_{u} and r_{f} , and then calculated their values at 298 K. The loss for P16 with $N_{\text{r}} = 12$ was thus obtained: $-T\Delta^{\text{II}}S_{\text{C}} = -T\Delta^{\text{III}}S_{\text{C}} \sim 8.83$ kcal/mol. The neutron scattering experiments cover only the picosecond time regime although the fluctuations in other time scales also affect the loss (63). For this reason, our method based on Fitter's experiments tends to give significant underestimation of the loss. However, P16 possesses two proline residues, making its random coils rather compact, with the result of an exceptionally small loss. For this reason, it may be justified to adopt the above result as a rough estimation.

If the contribution from the conformational-entropy loss, $-T\Delta^{\text{II}}S_{\text{C}} = 8.83$ kcal/mol, is incorporated in the theoretical calculation for the binding free energy, the result is in the range from -14.92 to 4.24 kcal/mol. The average value is -6.29 kcal/mol which is comparable with $\Delta^{\text{II}}G_{\text{experimental}} = -6.45$ kcal/mol. We can conclude that our theoretical results are reliable even in a quantitative sense. (Model 4 might be inappropriate because it gives the binding free energy a positive value when $-T\Delta^{\text{II}}S_{\text{C}}$ is incorporated.)

Decomposition of total-energy change into various components

The total-energy change upon the R12–P16 binding in path I is decomposed into $\Delta^{\text{I}}E_{\text{LJ}}$, $\Delta^{\text{I}}\epsilon_{\text{VH, LJ}}$, $\Delta^{\text{I}}E_{\text{ES}}$ and $\Delta^{\text{I}}\epsilon_{\text{VH, ES}}$ as displayed in Table 3. $\Delta^{\text{I}}E_{\text{LJ}}$ and $\Delta^{\text{I}}\epsilon_{\text{VH, LJ}}$ are both significantly large, but the former is negative while the latter is positive. $\Delta^{\text{I}}E_{\text{ES}}$ and $\Delta^{\text{I}}\epsilon_{\text{VH, ES}}$ are even much larger, but the former is negative while the latter is positive. $\Delta^{\text{I}}E_{\text{LJ}}$ and $\Delta^{\text{I}}\epsilon_{\text{VH, LJ}}$ or $\Delta^{\text{I}}E_{\text{ES}}$ and $\Delta^{\text{I}}\epsilon_{\text{VH, ES}}$ are compensating. Thus, the binding gives rise to an energy decrease arising from van der Waals attractive interactions between R12 and P16, but it is accompanied by an energy increase related to the loss of R12–water and P16–water van der Waals attractive interactions. The former is slightly larger and $\Delta^{\text{I}}(E_{\text{LJ}} + \epsilon_{\text{VH, LJ}})$ takes a rather small, negative value. The binding brings an energy decrease arising from electrostatic attractive interactions between R12 and P16, but it is accompanied by an energy increase related to the loss of R12–water and P16–water electrostatic attractive interactions. The latter is slightly larger and $\Delta^{\text{I}}(E_{\text{ES}} + \epsilon_{\text{VH, ES}})$ takes a rather small, positive value. As a matter of fact, the loss of R12–water and P16–water attractive interactions is quite large. However, about half of it is canceled out by the water reorganization energy (26,27) (i.e. the energy lowering due to the

Table 1. Binding free energy and its energetic and entropic components (in kcal/mol) upon the R12–P16 binding in path I

Model	$\Delta^I G$	$\Delta^I E_{\text{total}}$	$-T\Delta^I S_{\text{VH}}$
1	-36.16	5.47	-41.63
2	-39.80	-3.10	-36.71
3	-35.84	0.14	-35.98
4	-30.31	3.43	-33.73
5	-47.05	-6.58	-40.47

Five different models, which correspond to the upper halves of the $2 \times \text{R12}:2 \times \text{P16}$ complex structures in the Protein Data Bank with the accession code 2RU7, are considered for the structure of the R12:P16 complex.

Binding free energy, $\Delta^I G = G(\text{R12:P16}) - \{G(\text{R12}) + G(\text{P16}_{\text{compact}})\} = \Delta^I E_{\text{total}} - T\Delta^I S_{\text{VH}}$.

Total-energy change, $\Delta^I E_{\text{total}} = \Delta^I(E_{\text{LJ}} + \epsilon_{\text{VH, LJ}}) + \Delta^I(E_{\text{ES}} + \epsilon_{\text{VH, ES}})$.

Contribution from the water-entropy gain, $-T\Delta^I S_{\text{VH}}$.

Table 2. Binding free energy and its energetic and entropic components (in kcal/mol) upon the R12–P16 binding in path II

Model	$\Delta^{\text{II}} G$	$\Delta^{\text{II}} E_{\text{total}}$	$-T\Delta^{\text{II}} S_{\text{VH}}$
1	-21.67	36.26	-57.93
2	-14.04	40.32	-54.36
3	-23.75	27.43	-51.18
4	-4.59	45.10	-49.70
5	-11.53	48.84	-60.37

Five different models, which correspond to the upper halves of the $2 \times \text{R12}:2 \times \text{P16}$ complex structures in the Protein Data Bank with the accession code 2RU7, are considered for the structure of the R12:P16 complex.

Binding free energy, $\Delta^{\text{II}} G = G(\text{R12:P16}) - \{G(\text{R12}) + G(\text{P16}_{\text{random coils}})\}$, where $-T\Delta^{\text{II}} S_{\text{C}}$ is omitted: $\Delta^{\text{II}} G = \Delta^{\text{II}} E_{\text{total}} - T\Delta^{\text{II}} S_{\text{VH}}$.

Total-energy change, $\Delta^{\text{II}} E_{\text{total}} = \Delta^{\text{II}} E_{\text{B}} + \Delta^{\text{II}}(E_{\text{LJ}} + \epsilon_{\text{VH, LJ}}) + \Delta^{\text{II}}(E_{\text{ES}} + \epsilon_{\text{VH, ES}})$.

Contribution from the water-entropy gain, $-T\Delta^{\text{II}} S_{\text{VH}}$.

structural reorganization of water upon binding) with the result that the total loss becomes comparable with the gain of R12–P16 attractive interactions. $\Delta^I(E_{\text{LJ}} + \epsilon_{\text{VH, LJ}})$ and $\Delta^I(E_{\text{ES}} + \epsilon_{\text{VH, ES}})$ are also compensating, leading to a considerably small total-energy change (it is either positive or negative). nullnull

The energy decrease due to the gain of R12–P16 attractive interactions is almost canceled out by the energy increase caused by the energetic dehydration effect. The binding is driven by the entropic EV effect with the help of the gain of R12–P16 attractive interactions compensating the dehydrations. A loss of the conformational entropy of P16 opposes the binding, but this effect is much smaller than the driving force, i.e. the water-entropy gain.

Global structural change of P16 upon binding

Table 4 gives the free-energy change, total-energy change and contribution from the water-entropy gain arising from the global structural change of P16 referred to as path III (see Figure 4). The reason for the significantly large total-energy change can be summarized as follows. The energy decrease, which arises from intramolecular van der Waals attractive interactions brought by the structural change, is somewhat compensated by the energy increase related to the loss of P16–water van der Waals attractive interactions. As a consequence, $\Delta^{\text{III}}(E_{\text{LJ}} + \epsilon_{\text{VH, LJ}})$ takes a rather small, negative value. By contrast, the energy increase related to the loss of P16–water electrostatic attractive interactions is considerably larger than the energy decrease arising from intramolecular electrostatic attractive interactions brought by the structural change. As a consequence, $\Delta^{\text{III}}(E_{\text{ES}} + \epsilon_{\text{VH, ES}})$ takes a considerably large, positive value, giving rise to the significantly large total-energy change. The water-entropy

gain contributes to the free-energy change as a negative component, but this effect is smaller than the positive total-energy change. Thus, the free-energy change takes a significantly large, positive value. It increases further when the conformational-entropy loss, $-T\Delta^{\text{III}} S_{\text{C}} = 8.83$ kcal/mol, is incorporated. The positive free-energy change in path III is consistent with the experimental observation that isolated P16 is characterized by no well-defined structure. nullnull

The large electrostatic dehydration effect can be interpreted as follows. The amino-acid sequence of P16 is Gly-Gln-Trp-Asn-Lys-Pro-Ser-Lys-Pro-Lys-Thr-Asn. Except for Gly, P16 possesses 11 side chains. Three of them (three lysine residues) are positively charged, six of them (Gln, Trp, Asn, Ser, Thr and Asn) are polar and only two of them (two proline residues) are nonpolar. P16 is thus highly hydrophilic and substantially more stabilized when the water-accessible surface area is large, i.e. it is rather extended.

Physical origins of water-entropy gains upon R12–P16 binding

The principal term in the right hand of Equation (10) is the first one dependent on the EV at normal temperature and pressure. The water-entropy gain upon the R12–P16 binding can be discussed by decomposing it into the water-entropy gains accompanying paths III and I in Figure 4: (III) the global structural change of P16 from a random-coil state to the compact structure and (I) the contact of P16 possessing the compact structure with R12. In path III, many of the atoms in P16 contact one another, leading to a large decrease in the EV. In path I, the atoms of P16 and R12 at their interface are closely packed, leading to an even larger decrease in the EV. The EV decreases by

Table 3. Components of total-energy change (in kcal/mol) upon the R12–P16 binding in path I

Model	$\Delta^I E_{\text{total}}$	$\Delta^I(E_{\text{LJ}} + \epsilon_{\text{VH, LJ}})$	$\Delta^I(E_{\text{ES}} + \epsilon_{\text{VH, ES}})$	$\Delta^I E_{\text{LJ}}$	$\Delta^I \epsilon_{\text{VH, LJ}}$	$\Delta^I E_{\text{ES}}$	$\Delta^I \epsilon_{\text{VH, ES}}$
1	5.47	-3.94	9.41	-54.92	50.98	-1048.45	1057.85
2	-3.10	-6.33	3.23	-52.84	46.51	-1017.00	1020.23
3	0.14	-2.05	2.19	-50.55	48.50	-1070.42	1072.61
4	3.43	-5.04	8.47	-50.81	45.77	-1001.51	1009.98
5	-6.58	-9.55	2.97	-53.50	43.95	-1021.89	1024.85

Five different models, which correspond to the upper halves of the 2×R12:2×P16 complex structures in the Protein Data Bank with the accession code 2RU7, are considered for the structure of the R12:P16 complex.

Total-energy change, $\Delta^I E_{\text{total}} = \Delta^I(E_{\text{LJ}} + \epsilon_{\text{VH, LJ}}) + \Delta^I(E_{\text{ES}} + \epsilon_{\text{VH, ES}})$, comprises $\Delta^I E_{\text{LJ}}$, $\Delta^I \epsilon_{\text{VH, LJ}}$, $\Delta^I E_{\text{ES}}$ and $\Delta^I \epsilon_{\text{VH, ES}}$.

$\Delta^I(E_{\text{LJ}} + \epsilon_{\text{VH, LJ}})$ is the LJ component.

$\Delta^I(E_{\text{ES}} + \epsilon_{\text{VH, ES}})$ is the electrostatic component.

$\Delta^I E_{\text{LJ}}$,

$\Delta^I \epsilon_{\text{VH, LJ}}$,

$\Delta^I E_{\text{ES}}$, and

$\Delta^I \epsilon_{\text{VH, ES}}$ originate from van der Waals attractive interactions between R12 and P16, from electrostatic attractive interactions between R12 and P16, from the energetic dehydration effect related to the loss of R12–water and P16–water van der Waals attractive interactions and from the energetic dehydration effect related to the loss of R12–water and P16–water electrostatic attractive interactions, respectively. The contribution from the structural reorganization of water upon binding is included in $\Delta^I \epsilon_{\text{VH, LJ}}$ and $\Delta^I \epsilon_{\text{VH, ES}}$.

Table 4. Free-energy change and its energetic and entropic components (in kcal/mol) due to the global structural change of P16 upon the R12–P16 binding referred to as path III

Model	$\Delta^{\text{III}} G$	$\Delta^{\text{III}} E_{\text{total}}$	$-T\Delta^{\text{III}} S_{\text{VH}}$
1	14.49	30.79	-16.30
2	25.76	43.42	-17.65
3	12.09	27.29	-15.19
4	25.71	41.68	-15.96
5	35.52	55.42	-19.90

Five different models are considered for the structure of P16_{compact} (P16_{compact} is taken from the R12:P16 complex).

Free-energy change due to the global structural change of P16 upon the R12–P16 binding, $\Delta^{\text{III}} G = G(\text{P16}_{\text{compact}}) - G(\text{P16}_{\text{random coils}})$, where $-T\Delta^{\text{III}} S_{\text{C}}$ ($= -T\Delta^{\text{III}} S_{\text{C}}$) is omitted: $\Delta^{\text{III}} G = \Delta^{\text{III}} E_{\text{total}} - T\Delta^{\text{III}} S_{\text{VH}}$.

Total-energy change, $\Delta^{\text{III}} E_{\text{total}} = \Delta^{\text{III}} E_{\text{B}} + \Delta^{\text{III}}(E_{\text{LJ}} + \epsilon_{\text{VH, LJ}}) + \Delta^{\text{III}}(E_{\text{ES}} + \epsilon_{\text{VH, ES}})$.

Contribution from the water-entropy gain, $-T\Delta^{\text{III}} S_{\text{VH}}$.

$12.49d_{\text{S}}^3$ (d_{S} is the molecular diameter of water, 0.28 nm) in path III and by $33.33d_{\text{S}}^3$ in path I. These are the average values for the five different models of the R12:P16 complex. Since $\Delta^{\text{III}} V_{\text{ex}}/d_{\text{S}}^3 = -12.49$, $\Delta^I V_{\text{ex}}/d_{\text{S}}^3 = -33.33$ and $C_1 < 0$, the R12–P16 binding accompanied by the structural change of P16 leads to a large gain of the water entropy.

Changes in hydration entropy and enthalpy under isobaric condition

The theoretical calculations are made under isochoric condition while the experiments are performed under isobaric condition. The changes in hydration quantities under the two conditions are related to one another as (53,64)

$$\Delta H/(k_{\text{B}}T) = \Delta \epsilon_{\text{VH}}/(k_{\text{B}}T) + (\alpha^*/\kappa_{\text{T}}^*)\Delta V_{\text{PH}}/d_{\text{S}}^3, \quad (12a)$$

$$\Delta S_{\text{PH}}/k_{\text{B}} = \Delta S_{\text{VH}}/k_{\text{B}} + (\alpha^*/\kappa_{\text{T}}^*)\Delta V_{\text{PH}}/d_{\text{S}}^3, \quad (12b)$$

$$\begin{aligned} \Delta \mu_{\text{H}}/(k_{\text{B}}T) &= \Delta \epsilon_{\text{VH}}/(k_{\text{B}}T) - \Delta S_{\text{VH}}/k_{\text{B}} = \\ \Delta H/(k_{\text{B}}T) - \Delta S_{\text{PH}}/k_{\text{B}}. \end{aligned} \quad (13)$$

Here, the subscript ‘PH’ denotes the hydration under isobaric condition, H is the hydration enthalpy, ΔV_{PH} is the system-volume change and d_{S} is the molecular diameter of

water. α^* and κ_{T}^* , which depend only on the properties of pure water, are defined as

$$\alpha^* = \alpha T, \quad (14a)$$

$$\kappa_{\text{T}}^* = \kappa_{\text{T}} k_{\text{B}} T / d_{\text{S}}^3, \quad (14b)$$

where α is the isobaric thermal expansion coefficient and κ_{T} is the isothermal compressibility. ΔV_{PH} can be calculated by the 3D-RISM theory (45–48). The values obtained are $\Delta V_{\text{PH}}/d_{\text{S}}^3 = 5.51$, 5.47 and -0.04 in paths I, II and III, respectively. $k_{\text{B}}T(\alpha^*/\kappa_{\text{T}}^*)\Delta V_{\text{PH}}/d_{\text{S}}^3$ in path I, for example, is only 2.92 kcal/mol that is far smaller than $\Delta^I S_{\text{VH}}$, $\Delta^I \epsilon_{\text{VH, LJ}}$ and $\Delta^I \epsilon_{\text{VH, ES}}$. Thus, the changes in hydration quantities under isobaric condition are almost indistinguishable from those under isochoric condition.

It is worthwhile to comment on a general case. Upon binding of a hydrophobic receptor and a hydrophobic ligand under isobaric condition, for example, the compression of bulk water occurs, possibly leading to a water-entropy loss (31): due to considerably large, negative ΔV_{PH} , ΔS_{PH} can become negative despite positive ΔS_{VH} . Of course, the water entropy always increases upon the binding under isochoric condition. Thus, isochoric condition, which is free from the effects of compression or expansion of bulk water,

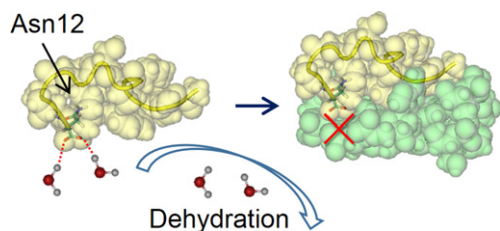


Figure 5. Image illustrating an example imperfection of the electrostatic complementarity in the R12:P16 complex. R12 and P16 are drawn on the upper and lower sides, respectively, using the space-filled model. An anionic C-terminal of P16 (Asn12) is buried within the contact surface but there is no counter cationic group in the vicinity of the buried C-terminal. Asn12 is depicted by the licorice model. The yellow ribbon represents the backbone of P16. The water molecules hydrating Asn12 are depicted by the ball-and-stick model. This figure is drawn using the VMD.

is more suited to the physical interpretation of a change in a thermodynamic quantity of hydration.

Significance of electrostatic complementarity

In the R12:P16 complex, the negatively charged phosphate groups of R12 are in contact with the positively charged ϵ -amino groups of lysine residues of P16, leading to R12–P16 energetic stabilization brought by electrostatic attractive interactions (see Figure 2) (5). This is a typical example of the electrostatic complementarity (2). Before the R12–P16 binding, the negatively and positively charged groups interact with water hydrogens with positive partial charges and with water oxygens with negative partial charges, respectively, maintaining R12–water and P16–water electrostatic attractive interactions. The binding is driven by the water-entropy gain, but it is unavoidably accompanied by an energy increase arising from the loss of R12–water and P16–water electrostatic attractive interactions. About half of the energy increase is canceled out by the energy decrease due to the structural reorganization of water upon binding (26,27), but the net effect acts as destabilization. This destabilization and the R12–P16 energetic stabilization are compensating. In other words, the latter is required for canceling the former. Similar arguments can be made for general binding processes for two solute molecules, and this is why the electrostatic complementarity is often observed in a complex. Looking at the structural data for a complex without accounting for the water roles can mislead to the conclusion that the R12–P16 binding, for instance, is driven by the R12–P16 electrostatic attractive interactions described above.

There is a case where the formation of the complete electrostatic complementarity is not favorable in terms of the water entropy. In the R12:P16 complex, for example, an anionic C-terminal of P16 is buried within the contact surface but there is no counter cationic group in the vicinity of the buried C-terminal, as observed in Figure 5. Such incomplete electrostatic complementarity, which makes a positive contribution to the total-energy change upon binding (the total energy is defined as the conformational energy plus the hydration energy), is accepted when the water-entropy gain dominates, or equivalently, when the priority should be given to the water-entropy gain. The incompleteness in

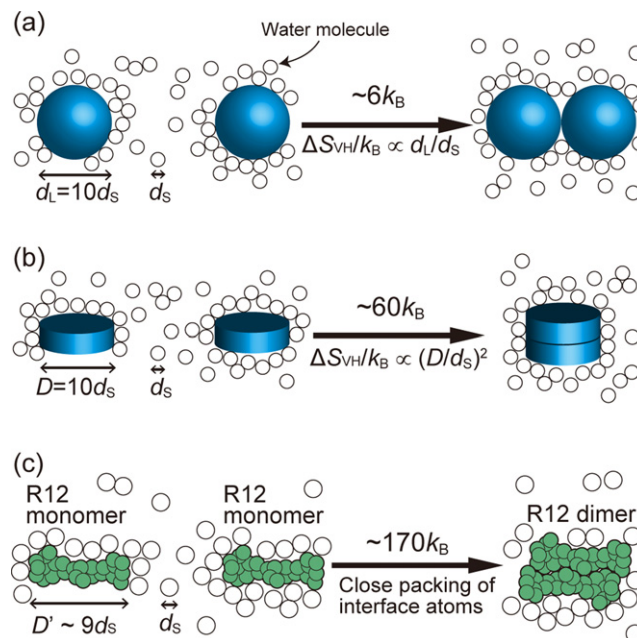


Figure 6. Water-entropy gain (k_B is the Boltzmann constant) upon the contact or stacking of two solutes immersed in water. (a) Contact of two spherical solutes with diameter $d_L = 10d_S$ (d_S denotes the water diameter). (b) Stacking of two disc-shaped solutes with surface diameter $D = 10d_S$. (c) Stacking of two R12 monomers each of which has disc-like shape with the span $D' \sim 9d_S$. The solutes in (a) and (b) possess no polyatomic structures whereas those in (c) possess them.

the electrostatic complementarity occurs in such a case, and the Apt8-hFc1 binding (16,25) mentioned in the Introduction section also provides a good example.

A hydrogen bond is usually expressed as an electrostatic attractive interaction between hydrogen and oxygen or nitrogen. The binding is accompanied by the break of solute–water hydrogen bonds compensated with the formation of solute–solute hydrogen bonds. Thus, the significance of the specific hydrogen bonding between the solute molecules upon binding can be discussed in a similar manner.

Crucial importance of geometric characteristics of solute molecules

The entropic EV effect becomes stronger as the molecular size of solvent decreases and/or the solvent number density increases. Thanks to the hydrogen bonding, water can exist in liquid state at ambient temperature and pressure despite its exceptionally small molecular size. The effect, which is relevant to the translational displacement of solvent molecules, becomes the largest when the solvent is water. The effect of the translational, configurational entropy of water plays imperative roles in sustaining life (65–67).

When two large spherical solutes with diameter d_L contact each other as shown in Figure 6a, the water-entropy gain is $\sim 6k_B$ and increases in proportion to d_L/d_S (d_S is the water diameter, 0.28 nm) (29). At $T = 298$ K, $k_B T$ corresponds to ~ -0.6 kcal/mol in terms of the contribution to the free-energy change. On the other hand, the water-entropy gain reaching $\sim 60k_B$ occurs upon the stacking of two large disc-shaped solutes with the surface diameter D ,

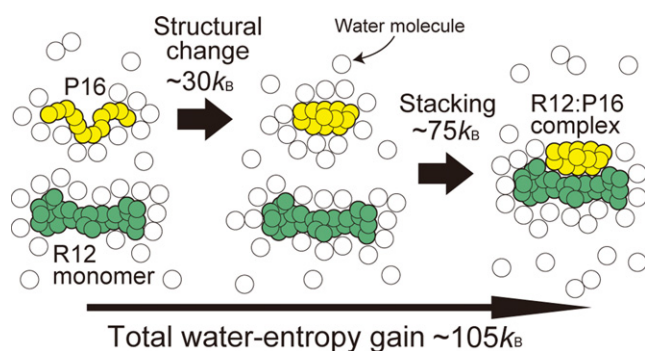


Figure 7. Water-entropy gain (k_B is the Boltzmann constant) upon each elemental process of the R12–P16 binding. The total water-entropy gain is also given.

and the gain increases in proportion to $(D/d_S)^2$ (see Figure 6b) (29). It is assumed that the spherical and disc-shaped solutes possess no polyatomic structures. Thus, the water-entropy gain originating from the entropic EV effect is remarkably influenced by the overall shapes and sizes of the solutes. It is also largely dependent on the details of the solute polyatomic structure. An R12 monomer has disc-like shape with the upper-surface diameter $D' \sim 9d_S$. The stacking of two R12 monomers to form a dimer illustrated in Figure 6c leads to the water-entropy gain of $\sim 170k_B$ that is far larger than the value expected from Figure 6b. This result arises from the close packing of the atoms constituting the interface regions. The decrease in the total EV brought by such packing is much larger than that brought by the contact of flat surfaces.

The water-entropy gain upon each elemental process of the R12–P16 binding is summarized in Figure 7. Each value given is the average taken over the five models of the complex structure. The gain upon the structural change of P16 from the random-coil state to the compact one is $\sim 30k_B$. Even for the structural change of such a small ($N_r = 12$) polypeptide, the water-entropy gain becomes this large. The gain arising from the binding of P16 to one of the surfaces of the R12 dimer is $\sim 75k_B$. This value is roughly a half of the gain in Figure 6c, partly because P16 is smaller than R12. The total gain is $\sim 105k_B$ ($105 = 30 + 75$). Thus, the binding in path II is accompanied by a very large gain of the water entropy.

Significance of stacking of flat moieties

In the R12:P16 complex, an indole ring of Trp 3 of P16 stacks on a guanidine ring of R12 (see Figure 2). This is why the π – π stacking interaction was regarded as a significant driving force for the binding (5). The π – π stacking interaction is often described by the van der Waals attractive interaction. When the stacking of flat moieties occurs, however, the van der Waals attractive interaction gained is accompanied by the loss of the moiety–water van der Waals attractive interactions, and these two factors are compensating. In other words, the stacking is required for canceling the loss. For this reason, in general, the stacking of flat moieties is often observed in a complex. The stacking is driven by a large gain of the water entropy as understood from Fig-

ure 6b or c. Looking at the structural data for a complex without accounting for the water roles can mislead to the conclusion that in the R12–P16 binding, for instance, the binding is driven by the π – π stacking interaction described above.

Significance of shape complementarity

The entropic EV effect is responsible for the shape complementarity often observed in a complex (2). For instance, it is obvious that the lock–key binding leads to a remarkable increase in the water entropy (see Figure 3a) and the molecular recognition is attributable to the lock–key shape complementarity which provides the largest decrease in the total EV. In the induced-fit binding, the lock structure exhibits a slight structural change so as to acquire the shape complementarity. In general binding processes, the shape complementarity between portions of the interface regions surely plays essential roles, and that at the atomic level also makes a remarkably large contribution to the water-entropy increase.

On the structural plasticity

The discussion on the significance of the electrostatic and shape complementarities made in the previous sections can also be applicable to the binding described by the lock–key or induced-fit model. We now discuss the molecular recognition accompanied by a global structural change of one of the solute molecules.

Close packing of the constituent atoms of a solute molecule leads to a large increase in the water entropy. However, such overall close packing cannot necessarily be achieved. When it is unachievable, the entropic EV effect is not large enough to surpass the energetic dehydration and conformational-entropy effects with the result that the solute molecule takes a rather extended, flexible structure. Even when the close packing can be accomplished, if the solute molecule is not sufficiently large and the energetic dehydration effect is substantial, it does not take a compact structure (68). P16 is a good example.

A protein is driven to take a structure in which the backbone and side chains are closely packed by the entropic EV effect. However, not all of the amino-acid sequences are amenable to such overall close packing. Even in cases where the overall close packing is not achievable, there are certainly the portions that can closely be packed. When preferential close packing of these portions is made, the other portions cannot participate in the packing and therefore become disordered and flexible. Nevertheless, the water-entropy gain brought by such preferential close packing is considerably larger than that resulting from an overall loose packing (69). An example case is found for PrP^C: its N-terminal half including P1 and P16 is disordered and flexible.

P16 and PrP^C can be referred to as ‘a soft molecule’ and ‘a molecule possessing a soft portion’, respectively. By contrast, ‘a rigid molecule’ possesses a structure in which overall close packing of the constituent atoms is successfully achieved, and even a slight change of its structure gives rise to an unacceptably large loss of the water entropy. R12 is

a good example of rigid molecules. It is possible even for a soft molecule or a soft portion of a molecule to construct a structure with overall close packing if the construction is made in concert with another molecule as a partner. In particular, when the partner is rigid, the soft molecule or the soft portion may bind to the partner by changing its structure in accordance with the partner structure, leading to the formation of a stable complex. Of course, there are not significantly many partners realizing such formation: only those which can realize it is successfully recognized.

CONCLUSIONS

We have investigated the binding of R12 and P16 as an important example of the new type of molecular recognition accompanied by a global structural change of a molecule upon binding to its targets. In this case, P16 exhibits the global change from an unstructured state to a compact structure while the structure of R12 remains almost unchanged (5,23). Changes in thermodynamic quantities upon binding are calculated using the molecular mechanics, hybrid method (30–37) in which the ADIET (38–42) is combined with the MA (43,44), and 3D-RISM theory (45–48). Molecular models are employed for water. The hybrid method and the 3D-RISM theory have been quite successful in studies on a variety of biological self-assembly processes (27–32,49–52) such as protein folding, association of protein molecules and receptor–ligand binding. In the present study, the hybrid method and the 3D-RISM theory are employed in calculating the hydration entropy and energy, respectively. This employment provides the most reliable results even in a quantitative sense.

The energy decrease due to the gain of R12–P16 attractive (i.e. van der Waals and electrostatic) interactions is almost canceled out by the energy increase originating from the energetic dehydration effect (i.e. by the energy increase caused by the loss of R12–water and P16–water attractive interactions plus the energy lowering arising from the structural reorganization of water). The binding is driven by the water-entropy gain that predominates over the conformational-entropy loss upon the global structural change of P16. The water-entropy gain originates primarily from the overlap of EVs of R12 and P16 followed by the increase in the total volume available to the translational displacement of water molecules in the system by the overlapped volume. Here, the EV is the volume of the space which the centers of water molecules cannot enter. The binding gives rise to an energy decrease arising from van der Waals attractive interactions between R12 and P16, but it is accompanied by an energy increase related to the loss of R12–water and P16–water van der Waals attractive interactions: the former is slightly larger and the net change (net change 1) is negative. The binding brings an energy decrease arising from electrostatic attractive interactions between R12 and P16, but it is accompanied by an energy increase related to the loss of R12–water and P16–water electrostatic attractive interactions; the latter is slightly larger and the net change (net change 2) is positive. The loss of R12–water and P16–water attractive interactions is quite large. However, about half of it is canceled out by the water reorganization energy (26,27), energy lowering due to the

structural reorganization of water upon binding. The resultant loss becomes comparable with the gain of R12–P16 attractive interactions. Net changes 1 and 2 are also compensating.

The picture of the binding can be summarized as follows. The binding is driven by the water-entropy gain which overcomes the dehydrations of R12 and P16 with the aid of a gain of R12–P16 van der Waals and electrostatic attractive interactions. A loss of the conformational entropy of P16, which opposes the binding, is surpassed by the water-entropy gain. It is very important to assure the gain of R12–P16 attractive interactions during the binding process though the gain itself is not a driving force. We believe that this picture is applicable to general binding processes. This is why stacking of flat moieties, specific hydrogen bonding and electrostatic complementarity (i.e. contact of oppositely charged groups) are frequently observed in the complexes (2). Stacking of flat moieties leads to a gain of intermolecular van der Waals attractive interactions, and specific hydrogen bonding and electrostatic complementarity lead to gains of intermolecular electrostatic attractive interactions. However, there is a case where it is not possible to accomplish both the water-entropy gain and the electrostatic complementarity. When the water-entropy gain dominates, or equivalently, when the priority should be given to the water-entropy gain, the electrostatic complementarity becomes incomplete. Such incompleteness has been found in the Apt8-hFcl binding (16,25).

The geometric characteristics (overall shapes, sizes and details of the polyatomic structures) of the solute molecules play crucially important roles in discussing the water-entropy gain. For instance, stacking of disc-shaped solutes brings a large decrease in the total EV followed by a correspondingly large gain of the water entropy. The gain becomes larger as the solute size increases (29). These effects are enlarged when the interface atoms are closely packed. The reason why stacking of flat moieties is frequently observed in the complexes (2) can also be understood from this viewpoint. Further, it is obvious that the water-entropy gain becomes quite large when the shape complementarity occurs within the interface region. The most striking example is found in the lock–key binding (see Figure 3a). The induced-fit binding, in which the lock structure exhibits a slight structural change so as to acquire the shape complementarity, is also driven by a large gain of the water entropy.

In what follows, we describe our proposition concerning the mechanism of molecular recognition accompanied by a global structural change of a molecule upon binding to its targets. When a solute molecule is not sufficiently large, the water-entropy gain is often incapable of surpassing the energetic dehydration and conformational-entropy effects with the result that the solute molecule takes a rather extended, flexible structure (68). Even for a large solute molecule like a protein, when its overall close packing is not achievable, only the portions that can closely be packed are preferentially packed: the other portions cannot participate in the close packing and therefore become disordered and flexible (69). (The water-entropy gain brought by such preferential close packing is considerably larger than that resulting from an overall loose packing.) These solute molecules can be referred to as ‘a soft molecule’ and ‘a molecule pos-

sessing a soft portion', respectively. By contrast, a solute molecule possessing a structure in which overall close packing of the constituent atoms is successfully achieved is 'rigid' because even a slight change of its structure gives rise to an unacceptably significant loss of the water entropy. It is possible even for a soft molecule or a soft portion of a molecule to construct a structure with overall close packing if the construction is made in concert with another molecule as a partner. In particular, when the partner is rigid, the soft molecule or the soft portion may bind to the partner by changing its structure in accordance with the partner structure, leading to the formation of a stable complex. Of course, there are not significantly many partners realizing such formation: only those which can realize it is successfully recognized.

Last, it is worthwhile to recapitulate the mechanism of the R12–P16 binding. For the binding, P16 changes its structure so that the R12–P16 interface atoms can closely be packed. At the same time, however, a decrease in R12–P16 intermolecular energy as well as that in P16 intramolecular energy (i.e. a gain of van der Waals and electrostatic attractive interactions between R12 and P16 and that within P16) is required to compensate for the energetic dehydration effect. As for the R12–P16 intermolecular energy, the formations of stacking structure and of electrostatic complementarity lead to decreases in van der Waals and electrostatic interaction energies, respectively. Since the stacking of flat moieties and the contact of unlike-charged groups bring decreases in the EVs, the formations are also favorable in terms of the water-entropy gain. The water-entropy gain brought by close R12–P16 interface packing, with the help of the decrease in the sum of P16 intramolecular and R12–P16 intermolecular energies, is large enough to surpass the energetic dehydration effect and the conformational-entropy loss of P16: the R12–P16 binding is achieved. If it was not large enough, the binding and the molecular recognition would not be realized. For the energetic component, the decrease in the sum of P16 intramolecular and R12–P16 energies, which is in favor of the binding, is considerably smaller than the opposing energetic dehydration effect. For the entropic component, by contrast, the water-entropy gain promoting the binding is far larger than the opposing conformational-entropy loss of P16.

To demonstrate our proposition mentioned in the last paragraph but one, we intend to investigate the binding of another RNA aptamer to its targets upon which the RNA aptamer exhibits a global structural change. The binding of an intrinsically disordered protein to its targets is also an interesting subject to be tackled. The ultimate goal is to elucidate all types of molecular recognitions including the 'lock and key' and 'induced-fit' binding models within the same theoretical framework in a unified manner.

ACCESSION NUMBER

The coordinates of the ensemble of five structures for the 2×R12:2×P16 complex have been deposited in the Protein Data Bank with the accession code 2RU7.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online, including [1–18].

ACKNOWLEDGMENTS

The computer program for the morphometric approach was developed with R. Roth and Y. Harano. We thank F. Nishikawa and S. Nishikawa for providing us with the data of the filter binding assay experiment.

FUNDING

Grant-in-Aid for Scientific Research, Ministry of Education, Culture, Sports, Science and Technology of Japan (25291035 to M. Kinoshita; 24121714, 25115507, 25291013 to M. Katahira; 23570146, 24113710 to T. Nagata); Grant-in-Aid for Japan Society for the Promotion of Science [to H. Oshima and T. Mashima]; CREST [to M. Katahira]; Sumitomo-Denko [to M. Katahira]; Iwatani Foundations [to M. Katahira].

Conflict of interest statement. None declared.

REFERENCES

- Uversky, V.N. and Dunker, A.K. (2010) Understanding protein non-folding. *Biochim. Biophys. Acta*, **1804**, 1231–1264.
- Hermann, T. and Patel, D.J. (2000) Adaptive recognition by nucleic acid aptamers. *Science*, **287**, 820–825.
- Fischer, E. (1894) Einfluss der configuration auf die wirkung der enzyme. *Ber. Dtsch. Chem. Ges.*, **27**, 2985–2993.
- Koshland, D.E. Jr (1958) Application of a theory of enzyme specificity to protein synthesis. *Proc. Natl. Acad. Sci. U.S.A.*, **44**, 98–104.
- Mashima, T., Nishikawa, F., Kamatari, Y.O., Fujiwara, H., Saimura, M., Nagata, T., Kodaki, T., Nishikawa, S., Kuwata, K. and Katahira, M. (2013) Anti-prion activity of an RNA aptamer and its structural basis. *Nucleic Acids Res.*, **41**, 1355–1362.
- Tuerk, C. and Gold, L. (1990) Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase. *Science*, **249**, 505–510.
- Ellington, A.D. and Szostak, J.W. (1990) *In vitro* selection of RNA molecules that bind specific ligands. *Nature*, **346**, 818–822.
- Burgstaller, P. and Famulok, M. (1994) Isolation of RNA aptamers for biological cofactors by *in vitro* selection. *Angew. Chem. Int. Ed. Engl.*, **33**, 1084–1087.
- Wang, J., Lu, Z., Wientjes, M.G. and Au, J.L.-S. (2010) Delivery of siRNA therapeutics: barriers and carriers. *AAPS J.*, **12**, 492–503.
- Nieuwlandt, D., Wecker, M. and Gold, L. (1995) *In vitro* selection of RNA ligands to substance P. *Biochemistry*, **34**, 5651–5659.
- Kawakami, J., Kawase, Y. and Sugimoto, N. (1998) *In vitro* selection of aptamers that recognize a monosaccharide. *Anal. Chimica Acta*, **365**, 95–100.
- Hermann, T. and Westhof, E. (1998) Aminoglycoside binding to the hammerhead ribozyme: a general model for the interaction of cationic antibiotics with RNA. *J. Mol. Biol.*, **276**, 903–912.
- Huang, D.-B., Vu, D., Cassidy, L.A., Zimmerman, J.M., Maher, L.J. III and Ghosh, G. (2003) Crystal structure of NF- κ B (p50)₂ complexed to a high-affinity RNA aptamer. *Proc. Natl. Acad. Sci. U.S.A.*, **100**, 9268–9273.
- Horn, W.T., Convery, M.A., Stonehouse, N.J., Adams, C.J., Liljas, L., Phillips, S.E.V. and Stockley, P.G. (2004) The crystal structure of a high affinity RNA stem-loop complexed with the bacteriophage MS2 capsid: further challenges in the modeling of ligand–RNA interactions. *RNA*, **10**, 1776–1782.
- Long, S.B., Long, M.B., White, R.R. and Sullenger, B.A. (2008) Crystal structure of an RNA aptamer bound to thrombin. *RNA*, **14**, 2504–2512.

16. Miyakawa, S., Nomura, Y., Sakamoto, T., Yamaguchi, Y., Kato, K., Yamazaki, S. and Nakamura, Y. (2008) Structural and molecular basis for hyperspecificity of RNA aptamer to human immunoglobulin G. *RNA*, **14**, 1154–1163.
17. Krauss, I.R., Merlino, A., Randazzo, A., Novellino, E., Mазzarella, L. and Sica, F. (2012) High-resolution structures of two complexes between thrombin and thrombin-binding aptamer shed light on the role of cations in the aptamer inhibitory activity. *Nucleic Acids Res.*, **40**, 8119–8128.
18. Daniels, D.A., Chen, H., Hicke, B.J., Swiderek, K.M. and Gold, L. (2003) A tenascin-C aptamer identified by tumor cell SELEX: systematic evolution of ligands by exponential enrichment. *Proc. Natl. Acad. Sci. U.S.A.*, **100**, 15416–15421.
19. Hernandez, L.I., Flenker, K.S., Hernandez, F.J., Klingelutz, A.J., McNamara, J.O. II and Giangrande, P.H. (2013) Methods for evaluating cell-specific, cell-internalizing RNA aptamers. *Pharmaceuticals*, **6**, 295–319.
20. Murakami, K., Nishikawa, F., Noda, K., Yokoyama, T. and Nishikawa, S. (2008) Anti-bovine prion protein RNA aptamer containing tandem GGA repeat interacts both with recombinant bovine prion protein and its β isoform with high affinity. *Prion*, **2**, 73–80.
21. McNamara, J.O. II, Andrechek, E.R., Wang, Y., Viles, K.D., Rempel, R.E., Gilboa, E., Sullenger, B.A. and Giangrande, P.H. (2006) Cell type-specific delivery of siRNAs with aptamer-siRNA chimeras. *Nat. Biotechnol.*, **24**, 1005–1015.
22. Xiang, Y. and Lu, Y. (2011) Using personal glucose meters and functional DNA sensors to quantify a variety of analytical targets. *Nat. Chem.*, **3**, 697–703.
23. Mashima, T., Matsugami, A., Nishikawa, F., Nishikawa, S. and Katahira, M. (2009) Unique quadruplex structure and interaction of an RNA aptamer against bovine prion protein. *Nucleic Acids Res.*, **37**, 6249–6258.
24. Humphrey, W., Dalke, A. and Schulten, K. (1996) VMD: visual molecular dynamics. *J. Mol. Graphics*, **14**, 33–38.
25. Nomura, Y., Sugiyama, S., Sakamoto, T., Miyakawa, S., Adachi, H., Takano, K., Murakami, S., Inoue, T., Mori, Y., Nakamura, Y. *et al.* (2010) Conformational plasticity of RNA for target recognition as revealed by the 2.15 Å crystal structure of a human IgG-aptamer complex. *Nucleic Acids Res.*, **38**, 7822–7829.
26. Sanchez, I.C., Truskett, T.M. and in't Veld, P.J. (1999) Configurational properties and corresponding states in simple fluids and water. *J. Phys. Chem. B*, **103**, 5106–5116.
27. Imai, T., Harano, Y., Kinoshita, M., Kovalenko, A. and Hirata, F. (2007) Theoretical analysis on changes in thermodynamic quantities upon protein folding: essential role of hydration. *J. Chem. Phys.*, **126**, 225102.
28. Kinoshita, M. (2002) Spatial distribution of a depletion potential between a big solute of arbitrary geometry and a big sphere immersed in small spheres. *J. Chem. Phys.*, **116**, 3493–3501.
29. Kinoshita, M. (2004) Ordered aggregation of big bodies with high asphericity in small spheres: a possible mechanism of the amyloid fibril formation. *Chem. Phys. Lett.*, **387**, 54–60.
30. Yoshidome, T., Kinoshita, M., Hirota, S., Baden, N. and Terazima, M. (2008) Thermodynamics of apolastocyanin folding: comparison between experimental and theoretical results. *J. Chem. Phys.*, **128**, 225104.
31. Oshima, H., Yasuda, S., Yoshidome, T., Ikeguchi, M. and Kinoshita, M. (2011) Crucial importance of the water-entropy effect in predicting hot spots in protein-protein complexes. *Phys. Chem. Chem. Phys.*, **13**, 16236–16246.
32. Chiba, S., Harano, Y., Roth, R., Kinoshita, M. and Sakurai, M. (2012) Evaluation of protein-ligand binding free energy focused on its entropic components. *J. Comput. Chem.*, **33**, 550–560.
33. Harano, Y., Yoshidome, T. and Kinoshita, M. (2008) Molecular mechanism of pressure denaturation of proteins. *J. Chem. Phys.*, **129**, 145103.
34. Yoshidome, T. and Kinoshita, M. (2009) Hydrophobicity at low temperatures and cold denaturation of a protein. *Phys. Rev. E*, **79**, 030905(R).
35. Oshima, H., Yoshidome, T., Amano, K. and Kinoshita, M. (2009) A theoretical analysis on characteristics of protein structures induced by cold denaturation. *J. Chem. Phys.*, **131**, 205102.
36. Amano, K., Yoshidome, T., Harano, Y., Oda, K. and Kinoshita, M. (2009) Theoretical analysis on thermal stability of a protein focused on the water entropy. *Chem. Phys. Lett.*, **474**, 190–194.
37. Oda, K., Kodama, R., Yoshidome, T., Yamanaka, M., Sambongi, Y. and Kinoshita, M. (2011) Effects of heme on the thermal stability of mesophilic and thermophilic cytochromes *c*: comparison between experimental and theoretical results. *J. Chem. Phys.*, **134**, 025101.
38. Kusalik, P.G. and Patey, G.N. (1988) On the molecular theory of aqueous electrolyte solutions. I. The solution of the RHNC approximation for models at finite concentration. *J. Chem. Phys.*, **88**, 7715–7738.
39. Kusalik, P.G. and Patey, G.N. (1988) The solution of the reference hypernetted-chain approximation for water-like models. *Mol. Phys.*, **65**, 1105–1119.
40. Kinoshita, M. and Bérard, D.R. (1996) Analysis of the bulk and surface-induced structure of electrolyte solutions using integral equation theories. *J. Comput. Phys.*, **124**, 230–241.
41. Kinoshita, M. (2008) Molecular origin of the hydrophobic effect: analysis using the angle-dependent integral equation theory. *J. Chem. Phys.*, **128**, 024507.
42. Kinoshita, M. and Yoshidome, T. (2009) Molecular origin of the negative heat capacity of hydrophilic hydration. *J. Chem. Phys.*, **130**, 144705.
43. König, P.-M., Roth, R. and Mecke, K.R. (2004) Morphological thermodynamics of fluids: shape dependence of free energies. *Phys. Rev. Lett.*, **93**, 160601.
44. Roth, R., Harano, Y. and Kinoshita, M. (2006) Morphometric approach to the solvation free energy of complex molecules. *Phys. Rev. Lett.*, **97**, 078101.
45. Beglov, D. and Roux, B. (1995) Numerical solution of the hypernetted chain equation for a solute of arbitrary geometry in three dimensions. *J. Chem. Phys.*, **103**, 360–364.
46. Kovalenko, A. and Hirata, F. (1999) Self-consistent description of a metal-water interface by the Kohn-Sham density functional theory and the three-dimensional reference interaction site model. *J. Chem. Phys.*, **110**, 10095–10112.
47. Imai, T., Harano, Y., Kinoshita, M., Kovalenko, A. and Hirata, F. (2006) A theoretical analysis on hydration thermodynamics of proteins. *J. Chem. Phys.*, **125**, 024911.
48. Luchko, T., Gusarov, S., Roe, D.R., Simmerling, C., Case, D.A., Tuszynski, J. and Kovalenko, A. (2010) Three-dimensional molecular theory of solvation coupled with molecular dynamics in Amber. *J. Chem. Theory Comput.*, **6**, 607–624.
49. Imai, T., Oda, K., Kovalenko, A., Hirata, F. and Kidera, A. (2009) Ligand mapping on protein surfaces by the 3D-RISM theory: toward computational fragment-based drug design. *J. Am. Chem. Soc.*, **131**, 12430–12440.
50. Imai, T., Miyashita, N., Sugita, Y., Kovalenko, A., Hirata, F. and Kidera, A. (2011) Functionality mapping on internal surfaces of multidrug transporter AcrB based on molecular theory of solvation: implications for drug efflux pathway. *J. Phys. Chem. B*, **115**, 8288–8295.
51. Chong, S.-H., Park, M. and Ham, S. (2012) Structural and thermodynamic characteristics that seed aggregation of amyloid- β protein in water. *J. Chem. Theory Comput.*, **8**, 724–734.
52. Chong, S.-H. and Ham, S. (2012) Impact of chemical heterogeneity on protein self-assembly in water. *Proc. Natl. Acad. Sci. U.S.A.*, **109**, 7636–7641.
53. Cann, N.M. and Patey, G.N. (1997) An investigation of the influence of solute size and insertion conditions on solvation thermodynamics. *J. Chem. Phys.*, **106**, 8165–8195.
54. Perkyns, J.S. and Pettitt, B.M. (1992) A dielectrically consistent interaction site theory for solvent-electrolyte mixtures. *Chem. Phys. Lett.*, **190**, 626–630.
55. Perkyns, J.S. and Pettitt, B.M. (1992) A site-site theory for finite concentration saline solutions. *J. Chem. Phys.*, **97**, 7656–7666.
56. Berendsen, H.J.C., Grigera, J.R. and Straatsma, T.P. (1987) The missing term in effective pair potentials. *J. Phys. Chem.*, **91**, 6269–6271.
57. Hornak, V., Abel, R., Okur, A., Strockbine, B., Roitberg, A. and Simmerling, C. (2006) Comparison of multiple Amber force fields and development of improved protein backbone parameters. *Proteins*, **65**, 712–725.
58. Singer, S.J. and Chandler, D. (1985) Free energy functions in the extended RISM approximation. *Mol. Phys.*, **55**, 621–625.

59. Hawkins,G.D., Cramer,C.J. and Truhlar,D.G. (1996) Parametrized models of aqueous free energies of solvation based on pairwise descreening of solute atomic charges from a dielectric medium. *J. Phys. Chem.*, **100**, 19824–19839.
60. Hawkins,G.D., Cramer,C.J. and Truhlar,D.G. (1995) Pairwise solute descreening of solute charges from a dielectric medium. *Chem. Phys. Lett.*, **246**, 122–129.
61. Case,D.A., Darden,T.A., Cheatham,T.E. III, Simmerling,C.L., Wang,J., Duke,R.E., Luo,R., Walker,R.C., Zhang,W., Merz,K.M. et al. (2012) *AMBER 12*. University of California, San Francisco, CA.
62. Gilson,M.K., Given,J.A., Bush,B.L. and McCammon,J.A. (1997) The statistical-thermodynamic basis for computation of binding affinities: a critical review. *Biophys. J.*, **72**, 1047–1069.
63. Fitter,J. (2003) A measure of conformational entropy change during thermal protein unfolding using neutron spectroscopy. *Biophys. J.*, **84**, 3924–3930.
64. Kinoshita,M., Harano,Y. and Akiyama,R. (2006) Changes in thermodynamic quantities upon contact of two solutes in solvent under isochoric and isobaric conditions. *J. Chem. Phys.*, **125**, 244504.
65. Kinoshita,M. (2009) Roles of translational motion of water molecules in sustaining life. *Front. Biosci.*, **14**, 3419–3454.
66. Kinoshita,M. (2009) Importance of translational entropy of water in biological self-assembly processes like protein folding. *Int. J. Mol. Sci.*, **10**, 1064–1080.
67. Kinoshita,M. (2013) A new theoretical approach to biological self-assembly. *Biophys. Rev.*, **5**, 283–293.
68. Harano,Y. and Kinoshita,M. (2005) Translational-entropy gain of solvent upon protein folding. *Biophys. J.*, **89**, 2701–2710.
69. Yoshidome,T., Ito,Y., Ikeguchi,M. and Kinoshita,M. (2011) Rotation mechanism of F₁-ATPase: crucial importance of the water entropy effect. *J. Am. Chem. Soc.*, **133**, 4030–4039.