doi:10.1016/j.jmb.2005.08.009 J. Mol. Biol. (2005) 353, 196–210

JMB

Available online at www.sciencedirect.com

SCIENCE DIRECT®

ELSEVIER

Conformational Equilibria and Free Energy Profiles for the Allosteric Transition of the Ribose-binding Protein

Krishna Pratap Ravindranathan, Emilio Gallicchio and Ronald M. Levy*

Department of Chemistry and Chemical Biology and BioMaPS Institute for Quantitative Biology Rutgers, The State University of New Jersey Piscataway, NJ 08854, USA

*Corresponding author

The ribose-binding protein (RBP) is a sugar-binding bacterial periplasmic protein whose function is associated with a large allosteric conformational change from an open to a closed conformation upon binding to ribose. The crystal structures of RBP in open and closed conformations have been solved. It has been hypothesized that the open and closed conformations exist in a dynamic equilibrium in solution, and that sugar binding shifts the population from open conformations to closed conformations. Here, we study by computer simulations the thermodynamic changes that accompany this conformational change, and model the structural changes that accompany the allosteric transition, using umbrella sampling molecular dynamics and the weighted histogram analysis method. The open state is comprised of a diverse ensemble of conformations; the open ribose-free X-ray crystal conformations being representative of this ensemble. The unligated open form of RBP is stabilized by conformational entropy. The simulations predict detectable populations of closed ribose-free conformations in solution. Additional interdomain hydrogen bonds stabilize this state. The predicted shift in equilibrium from the open to the closed state on binding to ribose is in agreement with experiments. This is driven by the energetic stabilization of the closed conformation due to ribose–protein interactions. We also observe a significant population of a hitherto unobserved ribose-bound partially open state. We believe that this state is the one that has been suggested to play a role in the transfer of ribose to the membrane-bound permease complex.

Keywords: ribose-binding protein; ligand-induced conformational change; protein allostery; umbrella sampling molecular dynamics; free energy profiles

Introduction

Ligand-induced conformational changes are essential for the functioning of biological systems. They are important in a variety of protein functions such as transport, signaling and enzymatic activity.[1] Large conformational changes in multi-domain proteins often involve motion of one domain as a relatively rigid subunit with respect to another.[2,3] Various simulation techniques have been used to study these macromolecular motions, including molecular mechanics,[4] targeted and steered molecular dynamics,[5–7] and elastic network theory.[8,9] Modeling techniques using low-resolution data from FRET[10] and cryo-EM[11,12] measurements have also been developed.

Here, we model by computer simulations the allosteric conformational change of the ribose-binding protein (RBP). RBP is a sugar-binding protein found in the periplasmic space of some bacteria. It serves as the initial receptor for transport of ribose across the cytoplasmic membrane into the cell, and it is the first component in the chemosensory pathway of bacterial chemotaxis.[13] On binding to ribose, the protein undergoes a large conformational change (see Figure 1), which allows recognition by membrane components that are responsible for transport and chemotaxis.[13,14] The structures of the closed-ribose bound conformation,[15] and three open apo conformations of the protein,[16] have been determined by X-ray crystallography. RBP is a 271 residue protein, consisting of

Abbreviations used: RBP, ribose-binding protein; WHAM, weighted histogram analysis method.

E-mail address of the corresponding author: ronlevy@lutece.rutgers.edu

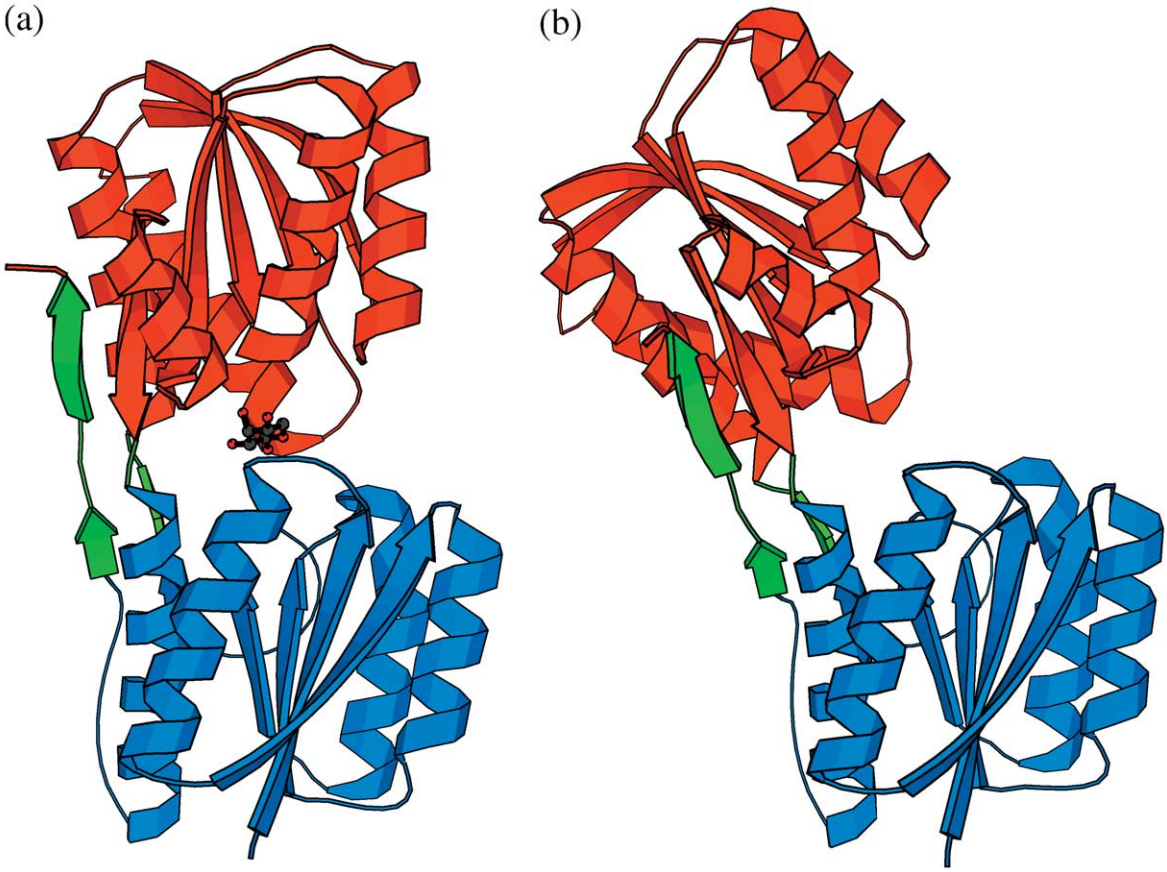

Figure 1. (a) The closed ribose-bound crystal structure of RBP (PDB id 2dri) and (b) the open ribose-free crystal structure of RBP (PDB id 1urp). Ribose (D-ribopyranose) is shown in ball-and-stick representation, carbon atoms are colored dark grey, oxygen atoms in red. The N-terminal domain of RBP is shown in blue, the C-terminal domain is shown in red and the hinge region in green. The Figures were generated using the program MOLSCRIPT.[47]

two structurally homologous domains linked by a three-stranded hinge (see Figure 1). The three strands that connect the two domains serve as a flexible hinge that allows opening and closing of the cleft between the two domains, which also forms the ribose-binding site. In the ribose-bound closed conformation, ribose is engulfed completely by the two domains (see Figure 1(a)). Comparison of the closed (see Figure 1(a)) and the open (see Figure 1(b)) crystal structures[16] revealed that the conformational change can be described as a rigid rotation of the two domains with respect to each other, whereas the internal structure of the two domains remains largely intact. Small-angle X-ray scattering studies[17] indicate a reduction of the radius of gyration of the protein by roughly 1.5 Å in the presence of ribose, confirming the transition in solution from an extended open conformation to a compact closed conformation upon ligand binding. NMR cleft angle studies[18] and disulfide-trapping[19] studies on the related glucose binding protein[20] further confirm these conclusions.

In the present study, our aim is to correlate structural changes with thermodynamic changes and to identify possible paths for the allosteric transition of RBP, using simulations based on detailed effective atomic potentials. The protein is assumed to exist in solution in an ensemble of conformations that can be described by statistical populations that follow statistical mechanical laws,[21,22] and the distribution of populations is expected to shift in favor of closed conformations on binding to ribose.

Standard molecular dynamics simulations using physics-based atomic force-fields are generally unsuitable for the study of large conformational changes in macromolecules, as the timescales involved are beyond what is currently achievable. Advanced sampling techniques are therefore necessary to study the conformational change of RBP. In this work, we perform molecular dynamics simulations using the angle of opening of RBP (the angle between the centers of mass of the two domains and the center of mass of the hinge segments, see Figure 2(a)) as the principal order parameter to describe the conformational change. This angle varies between 109° and 130° from the closed to the open X-ray conformations; the opening angles of the two known open conformations of a mutant of RBP are 134° and 145°.[16] A biasing potential based on the opening angle is defined and the free energy profile and population distribution

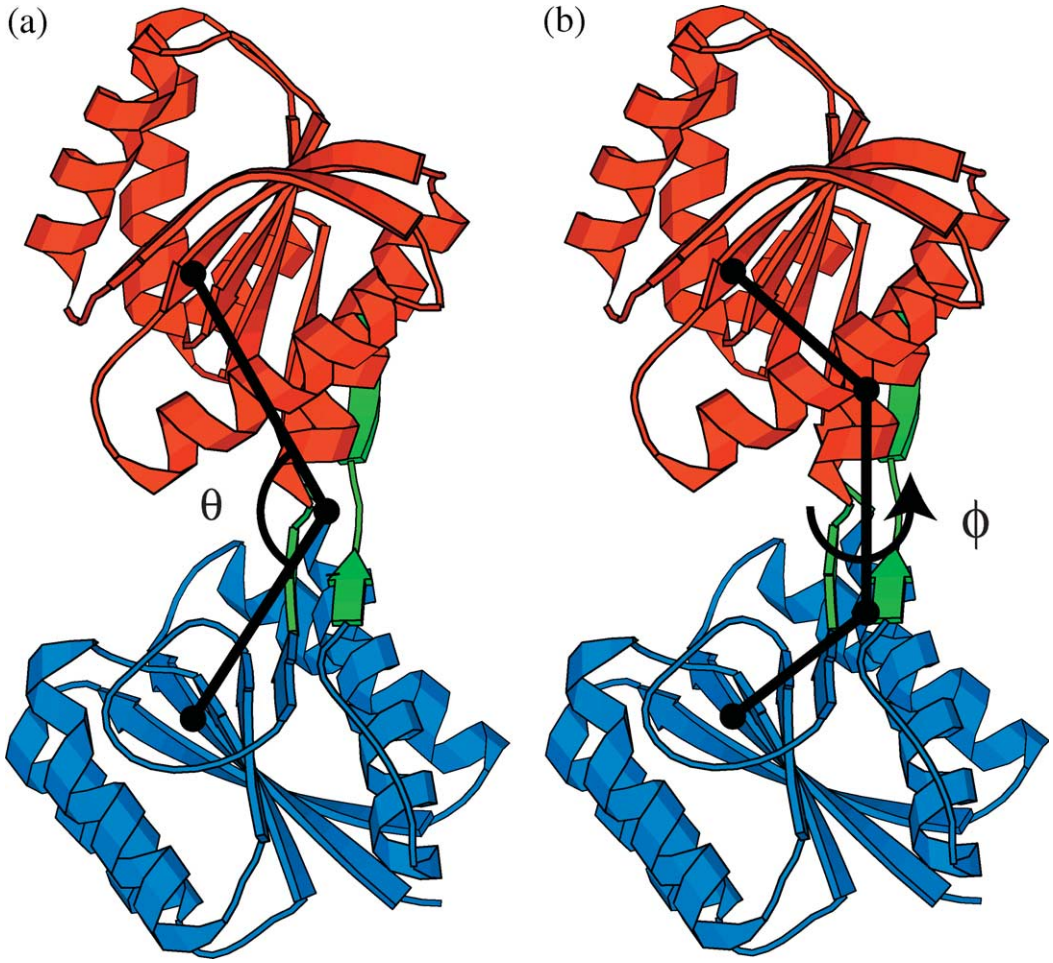

Figure 2. (a) Definition of the hinge angle θ, and (b) definition of the twist angle ϕ. The N-terminal domain of RBP is shown in blue, the C-terminal domain is shown in red and the hinge region in green. In (a) the segments joining the centers of mass of the N-terminal domain (bottom) and of the C-terminal (top) to the center of mass of the hinge region are shown in black. The angle between the segments is defined as the hinge angle. In (b) the centers of mass of the N-terminal domain (bottom), the base of the N-terminal domain (the base of a domain is defined by the three residues lining the three hinge strands), the C-terminal domain (top) and the base of the C-terminal domain define three segments, which are shown in black. The dihedral angle formed by these three segments is defined as the twist angle.

of protein conformations along this order parameter are then obtained using umbrella sampling[23] and the weighted histogram analysis method (WHAM)[24,25] unbiasing methodology. The free energy profile is calculated in the presence and in the absence of ribose to study the effect of ribose on the relative populations of the closed and open states. The free energy is further decomposed into energetic and entropic components. The thermodynamic forces that drive the transition from open to closed conformations on binding to ribose and the concomitant structural changes are investigated.

Results and Discussion

Conformational equilibrium and allostery of RBP

The calculated free energy profile of ribose-free RBP as a function of the hinge angle θ (which measures the degree of opening; see Figure 2(a)) is shown in Figure 3(a). The closed state corresponds to the free energy minimum at θ = 108°. The open state corresponds to a shallow basin that covers a wide span of hinge angles from 125° to 155°. The global minimum of the free energy profile is at θ = 129°, in correspondence with the known crystal structure of the wild-type apo form of RBP (θ = 130°).[16] The free energy minimum corresponding to the open state is lower than that corresponding to the closed state by roughly 2 kcal/mol. A free energy barrier at θ = 115, 2 kcal/mol higher than the closed state, separates the closed state from the open state. The calculated population distribution, shown in Figure 3(b), shows that, in the absence of ribose, the protein is predicted to exist mainly in the open state, in agreement with experimental evidence. The ratio of the open state to closed state populations is roughly 95:5. Thus, some population

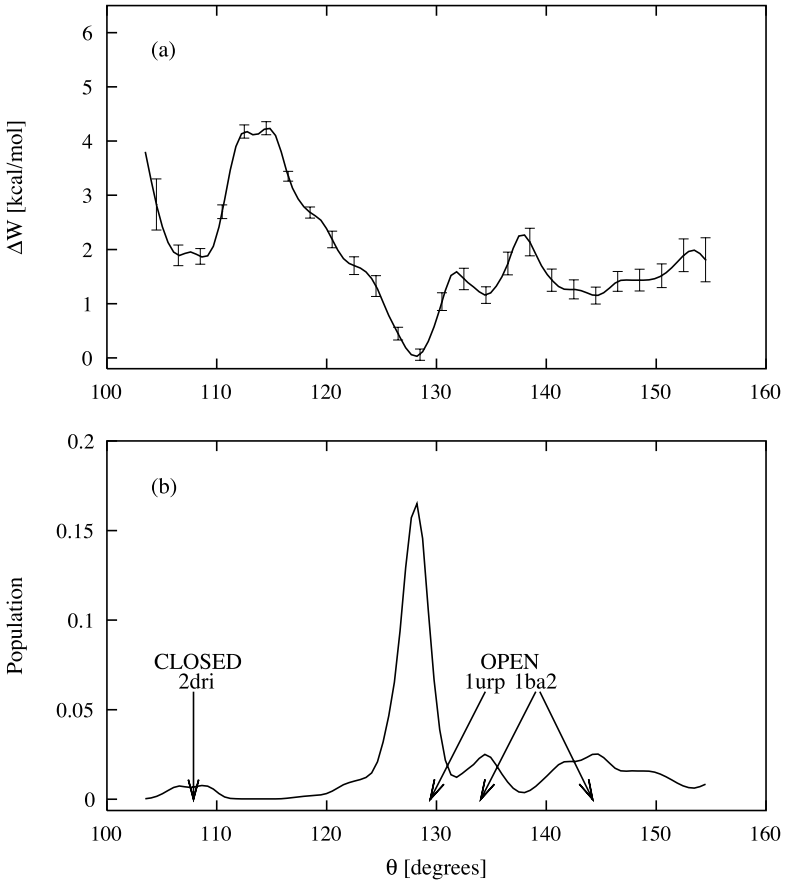

Figure 3. (a) The potential of mean force and (b) the population distribution of ribose-free RBP as a function of the hinge angle θ. The arrows indicate the values of the hinge angle θ of the crystal structures of the closed ribose-bound (PDB id 2dri) conformation, the ribose-free open (PDB id 1urp) conformation and of the two conformations of the RBP mutant (PDB id 1ba2).

of the closed state exists even in the absence of ribose. Significant population of the open state, which spans hinge angles between θ=125° and 155°, is found. The population peak centered at 129°, covering hinge angles between 125° and 131°, corresponds to conformations similar to that of the wild-type RBP open crystal structure (PDB id 1urp); 61% of the population of RBP is predicted to lie in this peak. In all, 35% of the population of RBP corresponds to conformations with larger hinge angles between 131° and 155° outside this main peak. The population at larger values of θ corresponds to conformations similar to the two X-ray crystal conformations of a mutant of RBP,[16] which have hinge angles θ=135° and θ=145°. Our calculations support the hypothesis that in solution, ribose-free RBP exists in a wide range of conformations; the wild-type and the two mutant crystal conformations being representative solution conformations stabilized by the crystal environment.[16] Our calculations confirm that the crystal structure of wild-type RBP is representative of the conformations of highest concentration in solution.[16]

The calculated free energy profile of ribose-bound RBP is shown in Figure 4(a). In the presence of ribose, the hinge angle of the closed conformation of RBP increases, on average, by 3° due to the extra interdomain space required to accommodate the ligand. Correspondingly, the free energy barrier separating the closed and the open ribose-bound states is shifted to θ=118°, from θ=115° in the absence of ribose. The height of the barrier is larger than in the absence of ribose by 1 kcal/mol. As seen in Figure 4(a), the free energy minimum of the closed state of ribose-bound RBP is lower than that of the open state by roughly 1 kcal/mol. The free energy profile indicates that in solution, the ribose-bound protein exists in two distinct states; a predominant closed state (θ=109°–112°) with 70% population, representative of the ribose-bound X-ray crystal structure (see Figure 4(b)),[15] and a previously unobserved, partially open ribose-bound state with 25% population centered at θ=122° and ranging from θ=119°–130°.

The predicted shift in population from the open to the closed state on binding to ribose is in agreement with experiments.[16,17] A crystal structure of the bound closed[15] form has been determined, whereas only open structures for the apo form of RBP have been identified.[16]

Small-angle X-ray scattering studies have revealed that the radius of gyration of RBP in solution is reduced by 1.4(±0.2) Å when ribose is introduced.[17] This further indicates a shift in population from the more extended open conformations to the more compact closed conformations upon binding to ribose. We have measured the average radius of gyration from the simulations

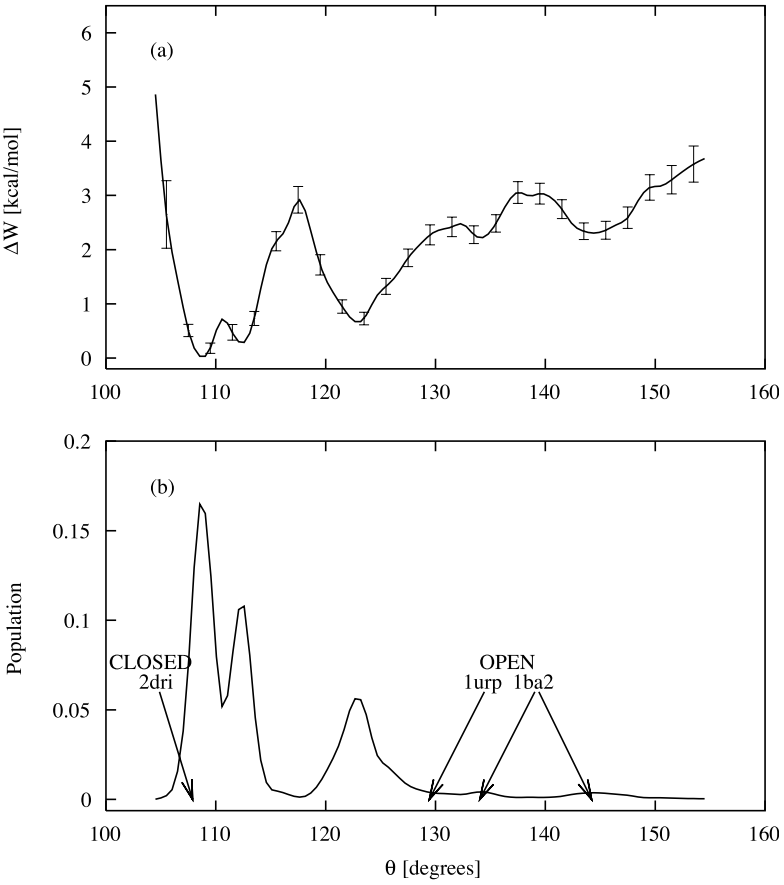

Figure 4. (a) The potential of mean force and (b) the population distribution of ribose-bound RBP as a function of the hinge angle θ. The arrows indicate the values of the hinge angle θ of the crystal structures of the closed ribose-bound (PDB id 2dri) conformation, the ribose-free open (PDB id 1urp) conformation and of the two conformations of an RBP mutant (PDB id 1ba2).

with and without ribose. The average radius of gyration of the ribose-bound and ribose-free forms of the protein is obtained by integrating the product of the radius of gyration and the corresponding population distribution as a function of the hinge angle. We predict a change in the radius of gyration upon binding to ribose of 0.8 Å, in reasonable agreement with the experiment (1.4(±0.2) Å). The difference between the measured and predicted radius of gyration change may be due partly to hydration shell effects, which were not considered in our analysis. The underprediction of the radius of gyration change may also be an indication that our model overpredicts the population of the ribose-bound open state and/or underpredicts the population of conformations with large hinge angles in the absence of ribose.

Thermodynamic analysis of the conformational equilibrium

The simulations indicate that in the absence of ribose the open state is more stable than the closed state, despite the fact that the average effective potential energy (see Table 1) favors the closed state. The open state of RBP has larger conformational entropy than the closed state and, indeed, inspection of the trajectories reveals that in the open state the two domains of RBP can rotate with respect to each other more freely than in the closed state. Thus, due to the higher entropy associated with the open state, which is sufficient to overcome the energetic stabilization of the closed state in the absence of ribose, the open state is the major species in solution.

In the presence of ribose, however, the entropic stabilization of the open state is not sufficient to overcome the increased energetic stabilization of the closed state. Whereas the average effective energy of the apo closed conformation is 14 kcal/mol smaller than the average effective energy of the apo open conformations (see Table 1), the ribose-bound closed conformation is

Table 1. Free energy difference ΔG, effective energy difference ΔE, and entropy difference $-T\Delta S$, between the open and closed states of ribose-free and ribose-bound RBP

	ΔG (kcal/mol)[a]	ΔE (kcal/mol)[a]	$-T\Delta S$ (kcal/mol)[a,b]
Ribose-free	−1.9	14.0	−15.9
Ribose-bound	0.6	30.2	−29.6

[a] Open form is defined to span hinge angles between 117° and 155°. Closed form is defined to span hinge angles between 104° and 116°.
[b] Entropies obtained using the equation $-T\Delta S = \Delta G - \Delta E$, T = 300 K.

stabilized energetically with respect to the open state by roughly 30 kcal/mol (see Table 1). The increase of the relative population of the closed state is thus a consequence of the energetic stabilization of the closed state due to the presence of ribose.

In Table 1, the higher conformational entropy loss of ribose-bound RBP in going from the open state to the closed state compared to ribose-free RBP should be noted (third column of Table 1). This is a consequence of the lower conformational entropy of the closed state of ribose-bound RBP relative to the closed state of ribose-free RBP. This is caused by the specific interactions of ribose with RBP (see below), which restrict interdomain motion.

The agreement between our calculations and the experiments[15–17] on the allosteric equilibrium of RBP would have been difficult to obtain with a low-resolution representation of the RBP–ribose complex. The difference in the energy between the open and closed states is small relative to the energy of each state, and the shift in free energy on ribose binding is even smaller. Moreover, the allosteric change is caused by the interaction of RBP with a small molecule for which an atomic-resolution model is well suited. The OPLS-AA/AGBNP all-atom force-field coupled with the morphing and umbrella sampling/WHAM methodology provides estimates of thermodynamic changes associated with structural changes predicted by the model that are consistent with the structural changes observed experimentally by X-ray crystallography and small-angle X-ray scattering.

Mechanism of opening

In addition to the hinge angle θ, the ensemble of conformations generated by the simulations is represented by another order parameter; the twist angle ϕ between the two domains. The twist angle is defined as the dihedral angle formed by the centers of mass of the N-terminal and C-terminal domains and the centers of mass of the regions at the base of each domain (see Figure 2(b)). The base of each domain is defined by the three residues lining the three hinge strands. Figure 5(a) and (b) show the calculated population distribution p^0 (θ,ϕ) as a function of the hinge angle θ and twist angle ϕ for ribose-free and ribose-bound RBP, respectively, obtained by the WHAM analysis described in Materials and Methods. It is evident from Figure 5(a) and (b) that, although the θ and ϕ order parameters are generally correlated (the larger the hinge angle the larger the twist angle), for any given value of the hinge angle the protein can assume a variety of conformations that span a wide range of twist angles.

In terms of the (θ,ϕ) order parameters, the ribose-bound closed state corresponds to the population peak spanning θ = 105°–115° and ϕ = 30°–70° (Figure 5(b)). The crystal structure of ribose-bound RBP (indicated by a cross in the lower left of Figure 5(b)) is located at (θ = 110°, ϕ = 50°), well within the high population region of the closed ribose-bound RBP predicted by the calculations. In the absence of ribose (Figure 5(a)), the closed state is much less populated and the corresponding peak is shifted to slightly lower values of the hinge and twist angles. The ribose-free open state (Figure 5(a)) is characterized by a broad region corresponding to conformations with hinge angles spanning 120–155° (the upper limit explored in this study) and twist angles between 60° and 160°. In this region, two high population regions can be identified, one centered at (θ = 129°, ϕ = 95°) in correspondence with the wild-type ribose-free crystal structure located at (θ = 130°, ϕ = 90°) and another, composed of more open conformations with hinge angles θ between 140° and 155°, in proximity to the crystal

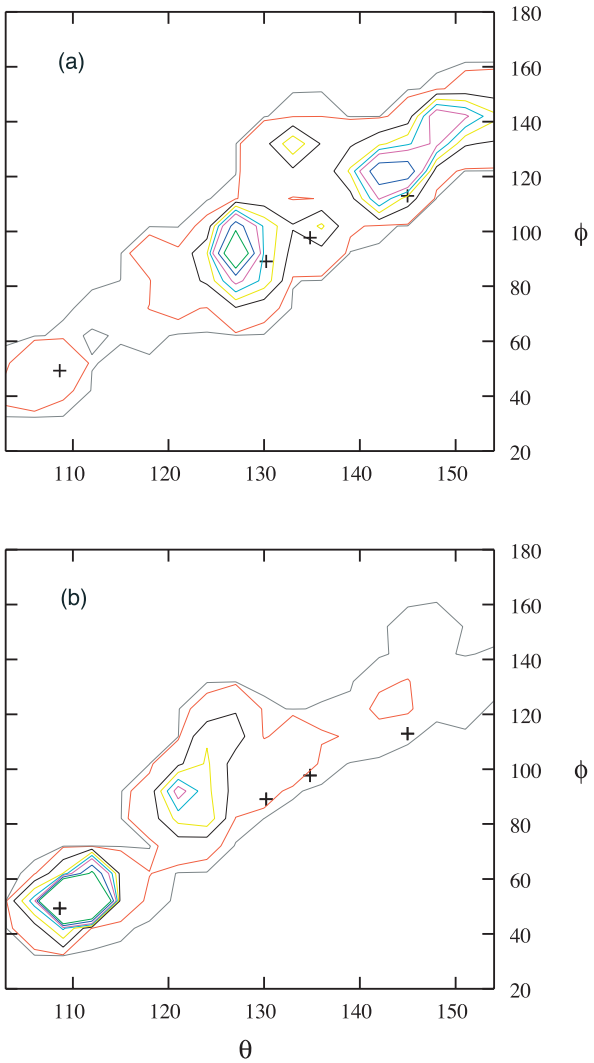

Figure 5. (a) Population distributions of ribose-free RBP and (b) of the ribose-bound RBP as a function of hinge angle θ and twist angle ϕ. Contours are drawn at 0.08, 0.06, 0.04, 0.03, 0.02, 0.01, 0.001, and 0.0001 relative populations. The crosses indicate the (θ,ϕ) coordinates of the crystal structures of, with increasing hinge angle θ, the closed ribose-bound, open ribose-free and the open ribose-free conformations of the RBP mutant.

structure of a mutant of RBP at ($\theta = 145°$, $\phi = 113°$). An additional populated region between the two major open populations (see Figure 5(a)) can be seen, which corresponds to the crystal structure of another conformation of the RBP mutant located at ($\theta = 135°$, $\phi = 100°$).

All of the known crystal structures of RBP (including the mutant), indicated by the crosses in Figure 5(a) and (b), lie within regions predicted to be populated in solution. This correspondence is unlikely to be an artifact due to the choice of starting conformations, because the simulations were started both from the crystal structure conformations and from conformations distributed uniformly along the morphing paths connecting one crystal structure conformation to another (see Materials and Methods). Moreover, the ribose-bound simulations were started from similar initial conditions but resulted in very different distribution of populations (compare Figure 5(a) and (b)). This result supports the hypothesis that the range of RBP conformations observed in different crystal forms reflects the distribution of conformations that exists in solution.[15,16]

In the presence of ribose, the open state ($\theta > 117°$, see Figure 5(b)) is significantly less populated than the open state of ribose-free RBP (see Figure 5(a)). Moreover, the population of this state is concentrated in a region centered at ($\theta = 122°$, $\phi = 95°$) corresponding to less open conformations than the most populated state of ribose-free RBP centered at ($\theta = 129°$, $\phi = 100°$). It has been hypothesized that, in order for ribose to be transferred to the membrane-bound permease complex, closed ribose-bound RBP first interacts with the permease complex and then opens partially, allowing ribose to be transferred to the complex.[15] This partially open form is still capable of binding the permease complex, and it is thought to be stabilized by interactions between ribose and specific binding site residues. Once ribose is transferred to the complex, the partially open conformation is destabilized in favor of the open conformation. Because it interacts weakly with the permease complex, once the open conformation is formed the protein leaves the complex and returns to solution. The open state at ($\theta = 122°$, $\phi = 95°$) of ribose-bound RBP predicted by our simulations could correspond to this hypothesized intermediate involved in the mechanism of ribose release. Later, we present a detailed structural analysis of this conformational state, for which a crystal structure is not available.

By comparing the closed and open crystal structures of RBP, Björkman & Mowbray have determined the axis of rotation of one domain with respect to the other domain and, observing that similar rotation axes are obtained when comparing the closed conformation to each of the three open conformations, have concluded that the mechanism of opening of RBP can likely be described by a continuous rotation of one domain with respect to the other around an approximately fixed axis of rotation.[16] The crystal structures of RBP (indicated by the crosses in Figure 5(a) and (b)) trace a linear path in the (θ,ϕ) space. Similarly, the high population regions of ribose-free RBP (Figure 5(a)) can be connected by a straight line contained within the populated region of the (θ,ϕ) plane, thus supporting the mechanism of opening proposed by Björkman & Mowbray.[16] In the presence of ribose (Figure 5(b)), however, a straight line connecting the closed state with the partially open state would inevitably intersect the region of very low population at ($\theta = 115°$, $\phi = 75°$). The distribution of populations of ribose-bound RBP suggests that the mechanism of opening in the presence of ribose is more complex than in the absence of ribose. On the basis of the shape of the population distribution and the analysis of the molecular dynamics trajectories, we formulate a mechanism of opening whereby, starting at the closed state, ribose-bound RBP first opens primarily by hinge bending until it reaches an opening hinge angle θ of approximately 118°, at which point a sharp 10° change of the twist angle occurs (from 65° to 75°). Then the protein continues to open with a coordinated change of both hinge angle and twist angle, as for ribose-free RBP. The initial phase of opening of ribose-bound RBP is therefore best described as two successive rotations around perpendicular axes, one perpendicular to and the other parallel with the hinge strands, whose overall effect is an apparent rotation around an axis with intermediate orientation similar to that identified by Björkman & Mowbray.[16]

Structural analysis

We now present an analysis of the key interactions between the two domains of RBP and

Table 2. Interdomain hydrogen bonds in the closed ribose-bound conformations and in the closed ribose-free conformations

Ribose-bound conformations		Ribose-free conformations	
Domain 1[a]	Domain 2[b]	Domain 1[a]	Domain 2[b]
N Ser68	O Gly134	N^{δ} Asn12	$O^{\delta 2}$ Asp165
O^{γ} Ser68	N Ser136	N^{δ} Asn13	O Phe164
N Gln91	O^{γ} Ser136	N^{δ} Asn13	$O^{\varepsilon 2}$ Glu192

A hydrogen bond D-H···A-R is said to exist if the distance between H and the acceptor atom (A) is less than 2.5 Å, the D-H···A angle (where D is the donor atom) is at least 120° and the H···A-R angle is at least 90°.

[a] Represents the N-terminal domain.

[b] Represents the C-terminal domain.

between ribose and RBP, placing particular emphasis on the ribose-free closed and ribose-bound open conformational states that have not been characterized experimentally.

Closed ribose-free conformations

Six interdomain hydrogen bonds are present in the closed ribose-free conformations as compared to the three interdomain hydrogen bonds that are found in the closed ribose-bound conformations. The three interdomain hydrogen bonds (see Table 2) found in the closed ribose-bound crystal structure,[15] are retained in the closed ribose-free conformations. The residues forming these three hydrogen bonds are on one side of the ribose-binding site (the right-hand side of Figure 6). Three new interdomain hydrogen bonds (see Table 2) are found in the closed ribose-free conformations on the opposite side of the ribose-binding site (left-hand side of Figure 6). Formation of these additional interdomain bonds is aided by the reduction of the hinge angle by approximately 3° in the absence of ribose, which brings the two domains into closer proximity.

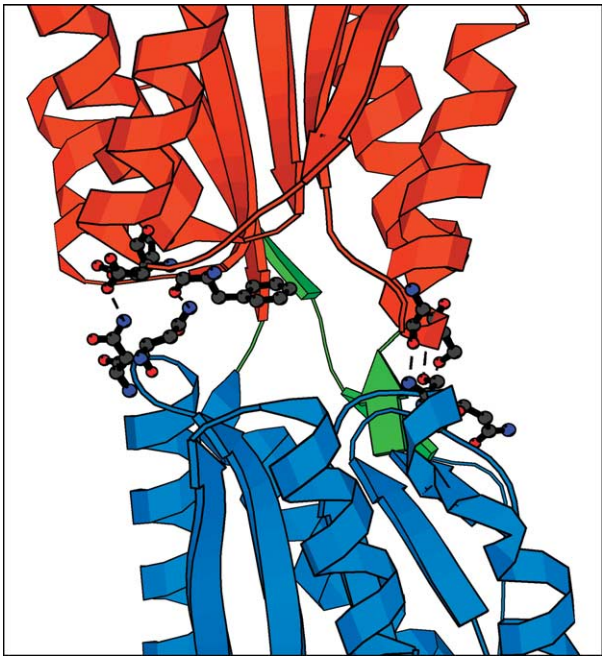

Figure 6. Representative structure of a ribose-free closed conformation of RBP. The N-terminal domain of RBP is shown in blue, the C-terminal domain is shown in red and the hinge region in green. Residues involved in interdomain hydrogen bonds are shown in ball-and-stick representation. The interdomain hydrogen bonds are shown as broken lines. The interdomain hydrogen bonds that are found in the closed ribose-bound crystal structure are those that are on the right side of the picture. The additional interdomain hydrogen bonds that are formed in the ribose-free closed conformation are those that are on the left side of the picture.

Partially open ribose-bound conformations

Eleven hydrogen bonds (see Figure 7(a) and Table 3) are formed between ribose and the binding site residues in the closed ribose-bound conformation. Five of these hydrogen bonds are formed with the N-terminal domain residues, five with the C-terminal domain residues, and one with a hinge residue. Three hydrophobic residues, Phe215, Phe15 and Phe16, make hydrophobic contacts with ribose in the closed conformation of RBP. In the partially open ribose-bound conformations (see Figure 7(b)), corresponding to the population peak at ($\theta = 122°$, $\phi = 95°$) of Figure 5(b), most of these ribose-RBP interactions are broken, ribose is rotated and it forms three new hydrogen bonds with RBP (see Figure 7(b) and Table 3). Gln235, which is a hinge residue that binds to O2 of ribose in the closed ribose-bound conformations (see Figure 7(a)), in the open conformations binds to O4 of ribose instead of O2 (see Figure 7(b)). This residue was hypothesized to play a role in stabilizing the ribose-bound open conformation,[16] and our model indicates that it is the only binding site residue in the closed state that continues to remain a binding site residue in the open conformations. Furthermore, the hydrophobic contact of ribose with Phe215, belonging to the N-terminal domain, is absent from the partially open conformations (the hydrophobic contacts with Phe15 and Phe16 are retained). In summary, ribose is bound more weakly to partially open conformations than to closed conformations, due to fewer hydrogen bonds and fewer hydrophobic contacts between ribose and RBP in partially open conformations. This presumably facilitates the release of ribose into the membrane bound permease.

In more open conformations, ribose interacts even more weakly with RBP, more often only with binding site residues of the C-terminal domain. The substantial loss of ligand–protein contacts in going from the closed to the open conformations supports the hypothesis that it is the strong interaction between ribose and RBP that stabilizes the closed conformation. Another point to be noted is that ribose intramolecular hydrogen bonds frequently replace ribose-RBP hydrogen bonds in open conformations. Formation of intramolecular hydrogen bonds is perhaps involved in the mechanism of ribose release.

An important feature of the ribose-bound closed conformation is the clustering of residues (see Figure 8(a)) that are known to be important in transport and chemotaxis.[26] This cluster of residues colored yellow in Figure 8(a) and (b) form a contiguous cluster distributed on the surface of RBP facing the ribose-binding pocket. The contiguity of this cluster of residues is conserved in the partially open conformations (Figure 8(b)) despite the large rotation of the N-terminal domain (shown in red in Figure 8(a) and (b)) with respect to the C-terminal domain (shown in blue in Figure 8(a) and (b)). The contiguity of residues involved in binding to the membrane-bound permease is

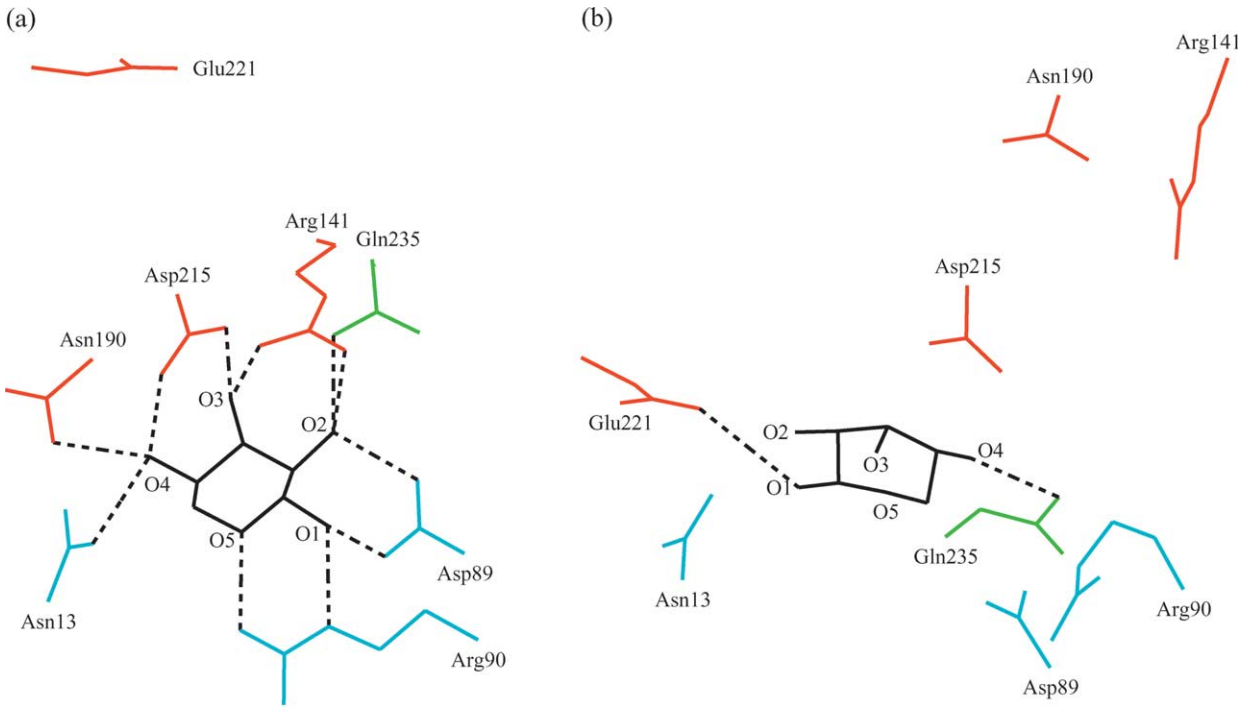

Figure 7. (a) Pattern of hydrogen bonds between ribose and RBP in a representative closed-ribose-bound conformation from the simulation, and (b) pattern of hydrogen bonds in a representative conformation of the ribose-bound open conformation. Hydrogen bonds between ribose and RBP are shown as broken lines. Residues belonging to the C-terminal domain are shown in blue. Residues belonging to the N-terminal domain are shown in red. Residues belonging to the hinge region are shown in green. Ribose is shown in black. It can be seen that in the open conformation the position of the residues belonging to the N-terminal domain (in red) changes significantly with respect to the closed ribose-bound crystal conformation in (a) due to the twisting of the N-terminal domain with respect to the C-terminal domain. Ribose forms hydrogen bonds to two new residues in the open conformation. One residue, Glu221, belongs to the N-terminal domain. The other residue, Gln235, is a hinge residue that forms a hydrogen bond with ribose in the closed ribose-bound conformation, although with a different oxygen atom of the ribose. Note that 11 hydrogen bonds between ribose and the protein are formed in the closed conformation in (a), whereas only two hydrogen bonds are present in the open conformation in (b).

important if the partially open conformations have to retain the capacity to bind to the permease. The cluster of residues in the partially open ribose-bound conformations (Figure 8(b)) is more symmetrical with respect to the position of ribose than the closed ribose-bound conformation (Figure 8(a)), and forms a channel that presumably aids the transfer of ribose to the membrane-bound permease. The observed shift of the position of ribose (Figure 8(b)) towards the face of RBP that is expected to bind to the permease, is also consistent with the proposal that the partially open ribose-bound state of RBP observed in this study corresponds to a relevant intermediate for the release of ribose into the permease.

Table 3. Hydrogen bonds between RBP and ribose in closed conformations and in partially open conformations

Closed conformations		Partially open conformations	
RBP	Ribose	RBP	Ribose
N^{δ} Asn13	O4	N^{δ} Gln235	O4
O Asp89	O1	O Gly216	O4
O Asp89	O2	O^{δ} Glu221	O1
N^{ε} Arg90	O1		
N^{η} Arg90	O5		
$N^{\eta 1}$ Arg141	O2		
$N^{\eta 2}$ Arg141	O3		
N^{δ} Asn190	O4		
O Asp215	O3		
O Asp215	O4		
N^{ε} Gln235	O2		

Conclusions

We have used advanced simulation techniques and high-resolution modeling to study the allosteric equilibrium of the ribose-binding protein. Experiments show that ribose binding induces a conformational change from an open conformation to a closed conformation that allows the protein to bind to a membrane-bound permease complex. Similar ligand-induced conformational changes for other sugar-binding proteins are used as signaling events in bacterial chemotaxis and sugar transport.

We employ the OPLS-AA/AGBNP all-atom force-field to model RBP and the associated effector

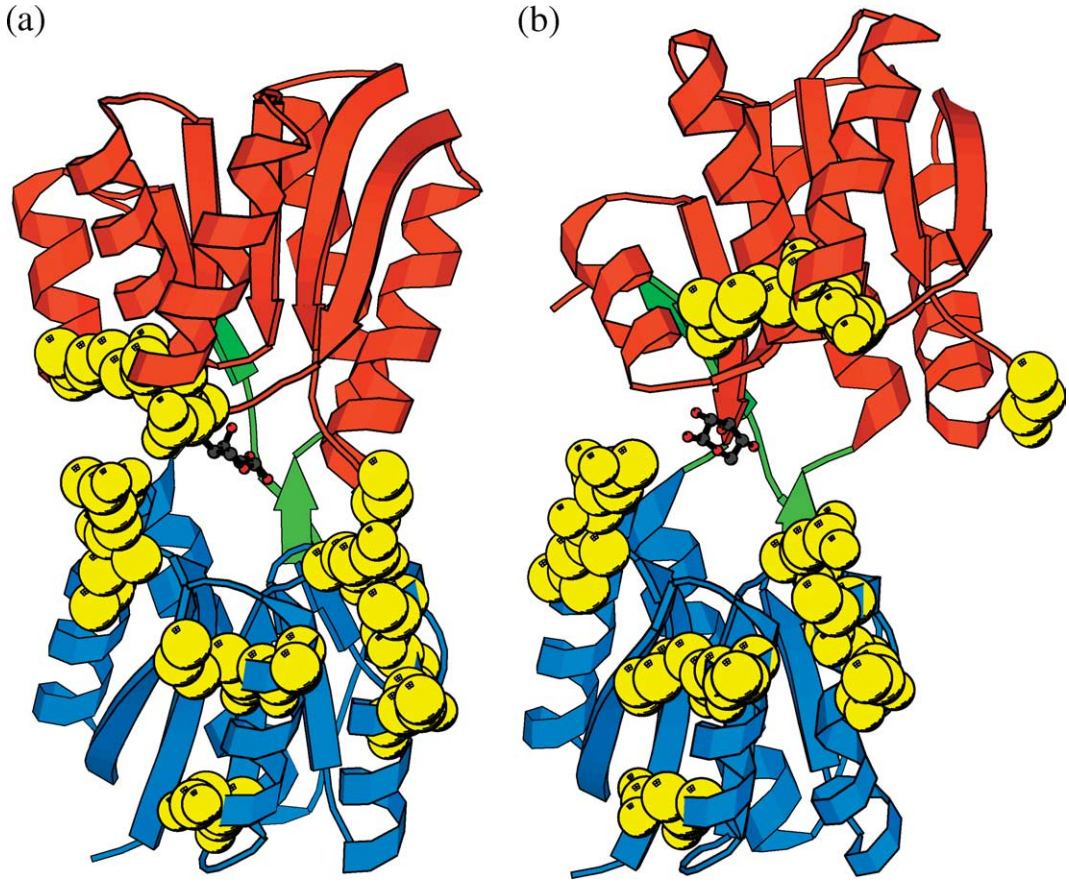

Figure 8. (a) Representative closed ribose-bound and (b) ribose-bound partially open conformations of RBP. Ribose is shown in ball-and-stick representation. Ribose carbon atoms are colored in dark grey, oxygen atoms in red. Residues known to be important in chemotaxis and transport are colored yellow in a space-filling representation.[26] The N-terminal domain of RBP is shown in blue, the C-terminal domain is shown in red and the hinge region in green. In the partially open conformation, the contiguity of residues (shown in yellow) seen in the closed bound conformation is retained despite the rotation of the N-terminal domain with respect to the C-terminal domain. In this conformation, ribose is shifted towards the face of RBP that is involved in binding to the permease complex and in the mechanism of ribose release.

molecule at atomic resolution. The free energy as a function of the interdomain hinge angle is computed using the umbrella sampling methodology and the WHAM unbiasing scheme. We show that the free energy ranking of the closed and the open states with and without ligand bound is in agreement with experiments. The free energy of open ribose-free RBP is lower than the free energy of closed ribose-free RBP. Conversely, with ribose bound, the free energy of the open state is higher than in the closed state. Thus, binding of ribose shifts the conformational equilibrium to the closed state. We observe that the open state of ribose-free RBP is stabilized by conformational entropy. Binding to ribose stabilizes the closed state energetically and shifts the equilibrium in favor of the closed state.

The free energy surface of the open state of ribose-free RBP calculated from the simulation is shallow and broad, and therefore ribose-free RBP is predicted to exist in a wide range of open conformations in solution. In the open state, the protein is highly mobile and can bend around the hinge region, giving rise to a wide array of conformations. The population distribution of protein conformations peaks at a hinge angle of 129°, which is representative of the wild-type open crystal conformation. A total of 61% of the open state population lies in this conformational macrostate, while the rest of the population of open conformations is distributed at large values of the hinge angle, up to $\theta = 155°$, encompassing both mutant RBP crystal conformations. The protein shows considerable variability even with respect to the twist angle, which describes the rotation of one domain with respect to the other around an axis parallel with the hinge strands.

We predict that 5% of the population of the ribose-free RBP conformations exists in the closed state. Although no X-ray crystal structure of closed ribose-free RBP is known, a closed ligand-free crystal structure of the related glucose-binding protein has been solved.[27] Furthermore, disulfide trapping experiments on the glucose-binding

protein indicate the existence of a closed ligand-free state in solution.[19] The existence of a minor population of closed RBP in solution in the absence of ribose indicates that the closing of the protein on binding to ribose is better viewed as a shift in equilibrium rather than as a triggered change.[14] The functional context of RBP (of transport and chemotaxis) does not demand a high-fidelity trigger (where no significant population of closed ribose-free conformations or of open ribose-bound conformations exists). Moreover, the need for the protein to partially open in order to release ribose is incompatible with a high-fidelity trigger mechanism.

The analysis of the ensemble of conformations of ribose-free RBP suggests that the mechanism of opening of the ribose-free protein is likely well represented by a smooth rotation around a fixed axis of rotation centered on the hinge domain, as proposed previously.[16] However, the analysis of the ensemble of conformations of ribose-bound RBP suggests that the mechanism of opening in the presence of ribose is better represented by a two-step process whereby the protein first opens primarily by hinge bending and subsequently undergoes a sharp transition involving mainly a twisting motion.

We were able to characterize two important conformational states of RBP that have not been studied by experimental means; the closed conformations in the absence of ribose and the partially open conformations in the presence of ribose. The ribose-free closed state of the protein has three new interdomain hydrogen bonds that are clustered together on the opposite side of the binding pocket with respect to the cluster of the three interdomain hydrogen bonds already observed in the ligand bound crystal structure.[15] In open ribose-bound conformations, all of the hydrogen bonds that are present between ribose and RBP in the closed conformations are absent and three new hydrogen bonds are formed, two with the C-terminal domain and one with the hinge region. In these conformations, ribose is rotated by almost 180° with respect to the closed conformation. This partially open state could correspond to that hypothesized to be involved in the mechanism of transfer of ribose to the permease complex.[16] The transition from the closed conformations to the open conformations is predicted to be an activated process that goes through an intermediate state of unfavorable free energy, where the interactions present in the closed state are broken and have not yet fully been replaced by new ones. The free energy barrier separating the closed and open conformations is higher in the presence of ribose due to the more extensive rearrangement of ribose–RBP interactions.

Allostery is a phenomenon of fundamental importance in biology. The analysis of static conformations as obtained by X-ray crystallography is often not sufficient to determine the driving forces that regulate the allosteric equilibria. The concept of the free energy landscape in the context of protein function constitutes a framework to interpret allosteric mechanisms taking into account conformational flexibility.[21,22] In this work, we have employed a high-resolution effective potential and advanced sampling techniques to characterize the free energy landscape of a multi-domain allosteric protein, obtaining both thermodynamic and structural information at an atomic level for the relevant protein conformational transitions.

Materials and Methods

Choice of coordinates

The ribose binding protein is a two-domain protein (see Figure 1). The three strands (three residues in each strand) connecting the two domains serve as a hinge for the motion from the closed to the open conformations, which can be described as a rigid rotation of one domain with respect to the other (Figure 2(a)).

Björkman and Mowbray have described the opening motion of RBP in terms of the axis of rotation and angle of rotation necessary to align two corresponding domains of the closed and open conformations, oriented in such a way that the two other domains are superimposed.[16] They observed that similar axes of rotation are obtained by comparing the closed X-ray conformation with each of the open X-ray conformations. The rotation angle obtained by comparing the closed X-ray conformation[15] and the open X-ray conformation of wild-type RBP is 43°.[16] The corresponding rotation angles for the two X-ray conformations of a mutant of RBP are 50° and 63°.

The coordinates employed by Bjorkman & Mowbray[16] are not convenient as order parameters for the umbrella sampling molecular dynamics calculations described below. We define the hinge opening angle θ, formed between the segments connecting the centers of mass of the N-terminal (colored red in Figure 2(a)) and C-terminal (colored blue) domains with the center of mass of the hinge domain (colored grey). Residues 1–100 and 236–259 constitute the C-teminal domain, residues 108–231 and 269–271 constitute the N-terminal domain, and residues 101–107, 232–235 and 260–268 constitute the hinge domain. In addition to the hinge angle θ, another order parameter, the twist angle ϕ between the two domains, is useful to describe the conformational change from the closed to the open conformations. The twist angle is defined as the dihedral angle formed by the centers of mass of the N-terminal and C-terminal domains, and the centers of mass of the regions at the base of each domain (see Figure 2(b)). The base of each domain is defined by the six residues lining the three hinge strands; residues 99–100, 236–237 and 258–259 on the C-terminal side, and residues 108–109, 230–231 and 269–270 for the N-terminal domain. In terms of our definition of the hinge opening angle and the twist angle, the coordinates (θ,ϕ) of the X-ray conformations are (109°,50°) for the closed ribose-bound X-ray crystal conformation (PDB id 2dri), (130°,90°) for the open wild-type ribose-free X-ray conformation (PDB id 1urp), and (135°,100°) and (145°,113°) for the X-ray conformations of the RBP mutant (PDB id 1ba2).

We wish to calculate the population distribution along the hinge angle coordinate and the twist angle coordinate. Exhaustive sampling of the hinge angle θ and twist angle

ϕ is not possible in the time-scale of ordinary molecular dynamics simulations. An advanced sampling methodology is thus employed to generate the population distributions (or equivalently the potential of mean force) along these angles. We first generate an adiabatic morphing path connecting the known closed and open conformations, which is used as a starting point for the umbrella sampling molecular dynamics simulations. From these simulations, we collect biased distributions of the angles θ and ϕ, which are then unbiased using WHAM. This methodology is elaborated below.

Morphing path

An adiabatic morphing path[28,29] is generated using as the two end conformations the crystal structures of the closed conformation (PDB id 2dri) and of the open conformation (PDB id 1urp). The two endpoint conformations are energy minimized. The N-terminal domains of the two structures are then superimposed. The energy-minimized closed conformation is represented by a $3N$-dimensional (where N is the number of atoms) position vector $\mathbf{X}_1$and the energy-minimized open conformation by $\mathbf{X}_M$. All the atoms of the minimized closed conformation are then displaced towards their positions in the minimized open conformation by an amount $(\mathbf{X}_1-\mathbf{Y}_M)/(M-1)$, where M is the desired number of conformations along the interpolation path (here, $M=100$). The resulting structure is energy-minimized to give a new conformation represented by $\mathbf{X}_2$. This interpolation procedure is then repeated iteratively to give M conformations $\mathbf{X}_i$ along the morphing path.

The morphing procedure allows generation of low-energy conformations connecting the two end conformations. An identical procedure is used to generate low-energy conformations between the open crystal conformation at $\theta=130°$ and the open crystal conformation of the RBP mutant at $\theta=135°$, and from this structure to the crystal conformation of the RBP mutant at $\theta=145°$, after replacing the mutated residue with the corresponding wild-type residue. Energy-minimized conformations along the morphing pathway are selected at equally spaced values of θ from 109° to 145° and used as starting points for the umbrella sampling molecular dynamics simulations. Two sets of morphing paths were generated, one with ribose and one without it.

The purpose of starting the molecular dynamics simulations near the morphing path is to ensure uniform sampling along the path that connects the open and closed states. In principle, a morphing path is not necessary; it should be possible to start the umbrella sampling simulations at the closed state, say, and progressively vary the umbrella potential to force opening of the protein towards the open state. In practice, however, we observed that, unless the rate of change of the umbrella potential is made very small, this process does not produce an open conformation in agreement with the observed X-ray open structure. In particular, we observed that the amount of twist (see below) of one domain with respect to the other of the open conformation generated in this fashion was significantly smaller than in the X-ray open conformation.

Umbrella sampling

We choose the umbrella sampling coordinate to be the hinge angle θ. A harmonic biasing potential given by:[24]

$$V_i = \frac{1}{2}k_F(\theta-\theta_i)^2 \tag{1}$$

is added to the Hamiltonian of the system, where k_F is the force constant, θ is the conformation dependent hinge angle, θ_i is the reference value of the hinge angle set for each simulation, which is varied at equal intervals from 109° to 154°. The biasing potential forces the system to sample conformations with hinge angles close to the set reference angle θ_i. The forces that result from the biasing potential on each atom of the protein necessary for molecular dynamics (MD) sampling[29] are derived from equation (1) and implemented into the IMPACT molecular simulation software package.[30,31]

MD simulations are carried out with various values of the reference angle θ_i such that, over all simulations, conformational space is sampled nearly uniformly. The biased distributions $p_i(\theta)$ of the hinge angle θ generated from each simulation are then unbiased using WHAM to give the unbiased distribution function $p^0(\theta)$.[24,32] The free energy profile (potential of mean force) $W(\theta)$ is then calculated from the distribution function $p^0(\theta)$ as:

$$\Delta W(\theta) = W(\theta)-W(\theta_0) = -KT\ln\frac{p^0(\theta)}{p^0(\theta_0)} \tag{2}$$

where θ_0 is a reference value of the hinge angle. A similar procedure, described below, yields the joint probability distribution $p^0(\theta,\phi)$ of the hinge and twist angles, and estimation of the average effective potential energy.

Force-field and implicit solvent model

Simulations were carried out using the OPLS-AA all-atom force-field.[33] In the present study, we have adopted an implicit solvent model to mimic the water environment. Implicit solvent models have the advantage over explicit solvent models of being less computationally intensive, because of the very much smaller number of atomic interactions that need to be calculated at every step and, perhaps more importantly, because of the need of long simulations to properly average the fluctuating effects of the explicit solvent reaction field on the biomolecular solute. The simulations reported in this work would have required two orders of magnitude longer computer time if they had been performed with an explicit solvent representation of the solvent. Depending on the underlying model and implementation, however, implicit solvent models may not be as accurate as explicit solvent models, particularly in describing specific short-ranged solute–water interactions. In recent years, accurate and computationally efficient implicit solvent models have been developed that have been applied successfully to the study of biomolecules.[34,35] We have adopted the AGBNP implicit solvent model,[36] which is based on a novel implementation of the pairwise descreening scheme of the generalized Born model[37,38] for the electrostatic component, and a novel non-polar hydration free energy estimator. The non-polar term consists of an estimator for the solute solvent van der Waals dispersion energy designed to mimic explicit solvent solute–solvent van der Waals interaction energies, in addition to a surface area term corresponding to the work of cavity formation.[39–41] AGBNP makes use of a new parameter-free algorithm to calculate the scaling coefficients used in the pairwise descreening scheme to take into account atomic overlaps. The same algorithm is used to calculate the solute surface area. The parameter-free approach used in AGBNP and the sensitivity of the

AGBNP model with respect to large and small conformational changes make it suitable for high-resolution modeling. It has been used to model peptide folding,[42,43] protein loop modeling and protein ligand binding.[36] The model is fully analytical with first derivatives.

Simulation details

The crystal structures[15,16] of RBP are taken from the RCSB Protein Data Bank (PDB id 2dri, 1urp, and 1ba2).[44] All calculations were carried out using the IMPACT molecular simulations software package.[30,31]

A total of 32 umbrella sampling simulations (16 of RBP complexed with ribose and 16 of apo RBP) were carried out. For each simulation, the system is heated to 300 K over 3 ps. The system is then equilibrated for 225 ps at 300 K. After equilibration, data are gathered for 800 ps. Velocity rescaling was used with a temperature relaxation time of 0.4 ps.[45] The time-step of the simulations is 1 fs. A value of 0.3 kcal/mol per deg^2. was used for the force constant k_F in equation (1). All atoms were treated explicitly. Hydrogen atoms on the protein were generated by IMPACT. A residue-based, non-bonded neighbor list with a 12 Å cutoff radius is used.

WHAM

We use WHAM to merge the data from the umbrella sampling simulations and to unbias the histograms from each simulation to obtain thermodynamic data independent of the biasing potential.[32,46] WHAM finds the unbiased probability distribution $\mathbf{p}^0$ that maximizes the likelihood of the observed biased histogram counts n, assuming that the observed histogram counts are distributed according to a multinomial distribution:

$$P(\mathbf{n}|\mathbf{p}^0) \propto \prod_{i=1}^{S}\prod_{j=1}^{M}(f_i c_{ij} p_j^0)^{n_{ij}} \quad (3)$$

where S is the number of simulations, M is the number of bins, n_{ij} is the observed count in bin j from simulation i, p_j^0 is the unknown unbiased probability of observing a sample in bin j, $c_{ij} = \exp(-w_{ij}/k_bT)$, where w_{ij} is the value of the biasing potential in bin j (evaluated at the mid point of the bin) for simulation i, and:

$$f_i = \left[\sum_{j=1}^{M} c_{ij} p_j^0\right]^{-1} \quad (4)$$

is a normalization factor related to the free energy of the system in simulation i.[32] Maximization of $P(\mathbf{n}|\mathbf{p}^0)$ with respect to $\mathbf{p}^0$ leads to a system of non-linear equations, which are solved iteratively.

To obtain the potential of mean force with respect to the hinge angle θ the range of hinge angles from 100° to 136° is divided into 36 equally spaced bins and the number of times that in simulation i a protein conformation is found within each bin, n_{ij}, is recorded. These counts are then reduced by the factor $g = (1 + 2\tau/\Delta t)^{-1}$,[25] where $\tau/\Delta t$ is the correlation time of the samples in units of the time interval Δt between samples, to take into account statistical correlations in the data. The reduced counts are then inserted in the WHAM equation to obtain the unbiased probability distribution of the hinge angle $p^0(\theta)$. To obtain unbiased probability distributions of quantities other than the hinge angle θ (such as the average effective potential energy E and the twist angle φ) joint histograms of the hinge angle and the property of interest are considered. For example, to obtain the average energy we collect joint histograms of the hinge angle θ and the effective potential energy E by creating a 2-D grid of bins, each corresponding to a particular value θ_k of the hinge angle and a particular value E_l of the effective potential energy. The histogram counts in each bin are collected from each simulation and processed using WHAM,[32] noticing that the umbrella potential factors $w_{i,kl}$ depend only on index k (the one corresponding to the angle θ). The result of the WHAM processing is the joint discretized probability distribution p_{kl}^0 of θ and E. The unbiased average energy:

$$\bar{E}(\theta_k) = \sum_l E_l p_{kl}^0$$

for each hinge angle is then obtained from the joint discretized probability distribution. The unbiased population distribution $p^0(\theta,\phi)$ with respect to the opening angle θ and twist angle φ is computed in a similar way. Statistical uncertainties have been obtained by a numerical method that we have developed recently.[32]

Acknowledgements

We thank David Talaga for stimulating our interest in the structural biology of sugar-binding proteins. This work has been supported by a grant from the NIH (GM 30580).

References

1. Alberts, B., Bray, D., Lewis, J., Raff, M., Roberts, K. & Watson, J. D. (2002). *Molecular Biology of the Cell*, Garland Science, New York.
2. Gerstein, M., Lesk, A. & Chothia, C. (1994). Structural mechanisms for domain movements in proteins. *Biochemistry*, **33**, 6739–6749.
3. Gerstein, M. & Krebs, W. A. (1998). Database of molecular motions. *Nucl. Acids Res.* **26**, 4280–4290.
4. Mao, B. & McCammon, J. A. (1983). Theoretical study of hinge-binding protein: internal energy and free energy changes. *J. Biol. Chem.* **258**, 12543–12547.
5. Ma, J., Sigler, P. B., Xu, Z. & Karplus, M. (2000). A dynamic model for the allosteric mechanism of GroEl. *J. Mol. Biol.* **302**, 303–313.
6. Young, M. A., Gonfloni, S., Superti-Furga, G., Roux, B. & Kuriyan, J. (2001). Dynamic coupling between the SH2 and SH3 Domains of c-Src and Hck underlies their inactivation by C-terminal tyrosine phosphorylation. *Cell*, **105**, 115–126.
7. Villa, E., Balaeff, A. & Schulten, K. (2005). Structural dynamics of the lac repressor-DNA complex revealed by a multiscale simulation. *Proc. Natl Acad. Sci. USA*, **102**, 6683–6788.
8. Tirion, M. M. (1996). Large amplitude elastic motions in proteins from a single-parameter, atomic analysis. *Phys. Rev. Letters*, **77**, 1905–1908.
9. Wang, Y., Rader, A. J., Bahar, I. & Jernigan, R. (2004). Global ribosome motions revealed by elastic network model. *J. Struct. Biol.* **147**, 302–314.
10. Mekler, V., Kortkhonjia, E., Mukhopadhyay, J., Knight, J., Revyakin, A., Kapanidis, A. N. *et al.*

(2002). Structural organisation of bacterial RNA polymerase holoenzyme and the RNA polymerase-promoter open complex. *Cell*, **108**, 599–614.
11. Darst, S. A., Opalka, N., Chacon, P., Polyakov, P., Richter, C., Zhang, G. & Wriggers, W. (2002). Conformational flexibility of bacterial RNA polymerase. *Proc. Natl Acad. Sci. USA*, **99**, 4296–4301.
12. Tama, F., Wriggers, W. & Brooks, C. L., III (2002). Exploring global distortions of biological macromolecules and assembles from low-resolution structural information and elastic network theory. *J. Mol. Biol.* **321**, 297–305.
13. Falke, J. F., Bass, R. B., Butler, S. L., Chervitz, S. A. & Danielson, M. A. (1997). The two component signalling pathway of bacterial chemotaxis: a molecular view of signal transduction by receptors, kinases, and adaptation enzymes. *Annu. Rev. Cell. Dev. Biol.* **13**, 457–512.
14. Stock, A. M. & Mowbray, S. L. (1995). Bacterial chemotaxis: a field in motion. *Curr. Opin. Struct. Biol.* **5**, 744–751.
15. Mowbray, S. & Cole, L. (1992). The 1.7 Å X-ray structure of the periplasmic ribose receptor of *Escherichia coli*. *J. Mol. Biol.* **225**, 155–175.
16. Björkman, A. & Mowbray, S. (1998). Multiple open forms of ribose binding protein trace the path of its conformational change. *J. Mol. Biol.* **279**, 651–664.
17. Shilton, B. H., Flocco, M. M., Nilsson, M. & Mowbray, S. L. (1996). Conformational changes of three periplasmic receptors for bacterial chemotaxis and transport: the maltose-,glucose/galactose- and ribose-binding proteins. *J. Mol. Biol.* **264**, 350–363.
18. Luck, L. A. & Falke, J. J. (1991). Open conformation of substrate binding cleft: 19F NMR studies of cleft angle in the D-galactose chemosensory receptor. *Biochemistry*, **30**, 6484–6490.
19. Careaga, C. L., Sutherland, J., Sabeti, J. & Falke, J. J. (1995). Large amplitude twisting motions of an interdomain hinge: a disulfide trapping study of the glucose-galactose binding protein. *Biochemistry*, **34**, 3048–3055.
20. Vyas, N. K., Vyas, M. N. & Quiocho, F. A. (1988). Sugar and signal-transducer binding sites of the *Escherichia coli* galactose chemoreceptor protein. *Science*, **242**, 1290–1295.
21. Kumar, S., Ma, B., Tsai, C. J., Sinha, N. & Nussinov, R. (2000). Folding and binding cascades: dynamic landscapes and population shifts. *Protein Sci.* **2000**, 10–19.
22. Gunasekaran, K., Ma, B. & Nussinov, R. (2004). Is allostery an intrinsic property of all dynamic proteins? *Proteins: Struct. Funct. Bioinf.* **57**, 433–443.
23. Berneche, S. & Roux, B. (2001). Energetics of ion conduction through the K^+ channel. *Nature*, **414**, 73–77.
24. Roux, B. (1995). The calculation of the potential of mean force using computer simulations. *Comput. Phys. Commun.* **91**, 275–282.
25. Kumar, S., Bouzida, D., Swendsen, R. H., Kollman, P. A. & Rosenberg, J. M. (1992). The weighted histogram analysis method for free energy calculations of biomolecules. I. The method. *J. Comput. Chem.* **13**, 1011–1021.
26. Binnie, R. A., Zhang, H., Mowbray, S. & Hermodson, M. A. (1992). Functional mapping of the surface of *Escherichia coli* ribose-binding protein: mutations that affect chemotaxis and transport. *Protein Sci.* **1**, 1642–1651.
27. Flocco, M. M. & Mowbray, S. L. (1994). The 1.9 Å X-ray structure of a closed unliganded form of the periplasmic glucose/galactose receptor from *Salmonella typhimurium*. *J. Biol. Chem.* **269**, 8931–8936.
28. Krebs, W. G. & Gerstein, M. (2000). The Morph server: a standardised system for analyzing and visualizing macromolecular motions in a database framework. *Nucl. Acids Res.* **28**, 1665–1675.
29. McCammon, J. A. & Harvey, S. C. (1987). *Dynamics of Proteins and Nucleic Acids*, Cambridge University Press, Cambridge.
30. Kitchen, D. B., Hirata, D. A., Kofke, J. D., Westbrook, M., Yarmush, M. & Levy, R. M. (1990). Conserving energy during molecular dynamics simulations of water, proteins and proteins in water. *J. Comput. Chem.* **11**, 1169–1180.
31. Banks, J. L., Beard, S. B., Cao, Y., Cho, A. E., Damm, W., Farid, R. *et al*. (2005). Integrated modelling program, applied chemical theory (IMPACT). *J. Comput. Chem*. In the press.
32. Gallicchio, E., Andrec, M., Felts, A. K. & Levy, R. M. (2005). T-WHAM, replica exchange, and transition paths. *J. Phys. Chem. B*, **109**, 6722–6731.
33. Jorgensen, W. L., Maxwell, D. S. & Tirado-Rives, J. (1996). Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids. *J. Am. Chem. Soc.* **118**, 11225–11236.
34. Trylska, J., McCammon, J. A. & Brooks, C. L., III (2005) . Exploring assembly energetics of the 30S ribosomal subunit using an implicit solvent approach. *J. Am. Chem. Soc.* **127**, 11125–11133.
35. Vendruscolo, M., Paci, E., Karplus, M. & Dobson, C. M. (2003). Structures and relative free energies of partially folded states of proteins. *Proc. Natl Acad. Sci. USA*, **100**, 14817–14821.
36. Gallicchio, E. & Levy, R. M. (2004). AGBNP: an analytic implicit solvent model suitable for molecular dynamics simulations and high-resolution modeling. *J. Comput. Chem.* **25**, 479–499.
37. Qui, D., Shenkin, P. S., Hollinger, F. P. & Still, C. W. (1997). The GB/SA continuum model for solvation. A fast analytical method for the calculation of approximate Born radii. *J. Phys. Chem. A*, **101**, 3005–3014.
38. Bashford, D. & Case, D. A. (2000). Generalized Born models of macromolecular solvation effects. *Annu. Rev. Phys. Chem.* **51**, 129–152.
39. Gallicchio, E., Zhang, L. Y. & Levy, R. M. (2002). The SGB/NP hydration free energy model based on the surface generalized Born solvent reaction field and novel non-polar hydration free energy estimators. *J. Comput. Chem.* **23**, 517–529.
40. Levy, R. M., Zhang, L. Y., Gallicchio, E. & Felts, A. K. (2003). On the non-polar hydration free energy of proteins: surface area and continuum solvent models for the solute–solvent interaction energy. *J. Am. Chem. Soc.* **125**, 9523–9530.
41. Su, Y. & Gallicchio, E. (2004). The non-polar solvent potential of mean force for the dimerization of alanine dipeptide: the role of solute-solvent van der Waals interactions. *Biophys. Chem.* **109**, 251–260.
42. Felts, A. K., Harano, Y., Gallicchio, E. & Levy, R. M. (2004). Free energy surfaces of beta-hairpin and alpha-helical peptides generated by replica exchange molecular dynamics with the AGBNP implicit solvent model. *Proteins: Struct. Funct. Bioinf.* **56**, 310–321.
43. Andrec, M., Felts, A. K., Gallicchio, E. & Levy, R. M. (2005). Protein folding pathways from replica exchange simulations and a kinetic network model. *Proc. Natl Acad. Sci. USA*, **102**, 6801–6806.

44. Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H. *et al*. (2000). The Protein Data Bank. *Nucl. Acids Res.* **28**, 235–242.
45. Berendsen, H. C., Postma, J. P. M., van Gunsternen, W. F., DiNola, A. & Haak, J. R. (1984). Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* **81**, 3684–3690.
46. Bartels, C. & Karplus, M. (1997). Multidimensional adaptive umbrella sampling: applications to main chain and side chain peptide conformations. *J. Comput. Chem.* **18**, 1450–1462.
47. Kraulis, P. J. (1991). MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallog.* **24**, 946–950.

Edited by M. Levitt

(Received 4 June 2005; received in revised form 8 August 2005; accepted 9 August 2005)
Available online 24 August 2005