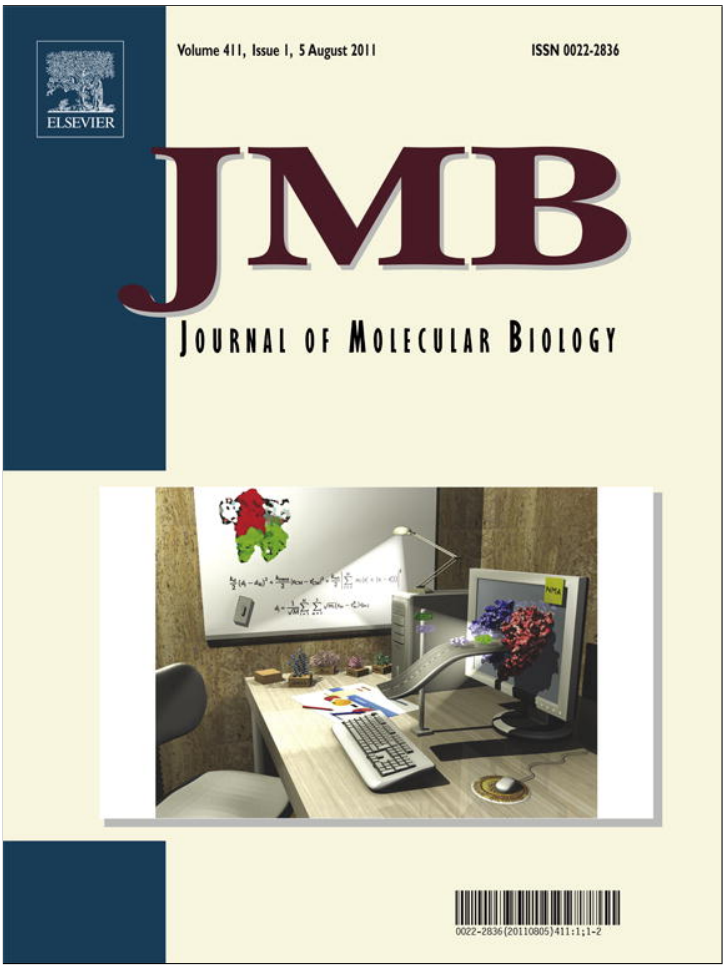

doi:10.1016/j.jmb.2011.05.039 J. Mol. Biol. (2011) 411, 36–52

Contents lists available at www.sciencedirect.com

Journal of Molecular Biology

journal homepage: http://ees.elsevier.com.jmb

Arg149 Is Involved in Switching the Low Affinity, Open State of the Binding Protein *Af*ProX into Its High Affinity, Closed State

Britta Tschapek[1]†, Marco Pittelkow[2]†, Linda Sohn-Bösser[2], Gudrun Holtmann[2], Sander H. J. Smits[1], Holger Gohlke[3], Erhard Bremer[2] and Lutz Schmitt[1]*

[1]*Institute of Biochemistry, Heinrich Heine University Duesseldorf, Universitaetsstrasse 1, D-40225 Duesseldorf, Germany*
[2]*Laboratory for Microbiology, Department of Biology, Philipps University Marburg, Karl-von-Frisch Strasse 8, D-35032 Marburg, Germany*
[3]*Institute of Pharmaceutical and Medicinal Chemistry, Heinrich Heine University Duesseldorf, Universitaetsstrasse 1, D-40225 Duesseldorf, Germany*

Received 8 April 2011; received in revised form 21 May 2011; accepted 25 May 2011
Available online 2 June 2011

Edited by R. Huber

Keywords:
ABC transporter; substrate binding protein; glycine betaine; substrate recognition; ligand binding site

The substrate binding protein *Af*ProX from the *Archaeoglobus fulgidus* ProU ATP binding cassette transporter is highly selective for the compatible solutes glycine betaine (GB) and proline betaine, which confer thermoprotection to this hyperthermophilic archaeon. A detailed mutational analysis of the substrate binding site revealed the contribution of individual amino acids for ligand binding. Replacement of Arg149 by an Ala residue displayed the largest impact on substrate binding. The structure of a mutant *Af*ProX protein (substitution of Tyr111 with Ala) in complex with GB was solved in the open liganded conformation to gain further insight into ligand binding. In this crystal structure, GB is bound differently compared to the GB closed liganded structure of the wild-type *Af*ProX protein. We found that a network of amino acid side chains communicates the presence of GB toward Arg149, which increases ligand affinity and induces domain closure of *Af*ProX. These results were corroborated by molecular dynamics studies and support the view that Arg149 finalizes the high-affinity state of the *Af*ProX substrate binding protein.

Introduction

Microorganisms import a variety of chemical compounds from environmental sources through high-affinity transport systems, many of which belong to the family of ATP binding cassette (ABC) transporters.[1] These systems depend strictly on a so-called substrate binding protein (SBP) that captures the ligand with high affinity and subsequently delivers it to its cognate transporter.[2] The SBP thereby determines the directionality of the overall transport reaction, and its interaction with the transmembrane domain of the transporter regulates the ATPase activity of the transport complex.[2,3]

Structural studies of SBPs revealed a common fold with a bilobal organization connected via a linker region.[4,5] In the ligand-free, open conformation, the

*Corresponding author. E-mail address: lutz.schmitt@hhu.de.

† B.T. and M.P. contributed equally to this work.

Abbreviations used: ABC, ATP binding cassette; MD, molecular dynamics; SBP, substrate binding protein; MBP, maltose binding protein; GB, glycine betaine; PB, proline betaine; PDB, Protein Data Bank.

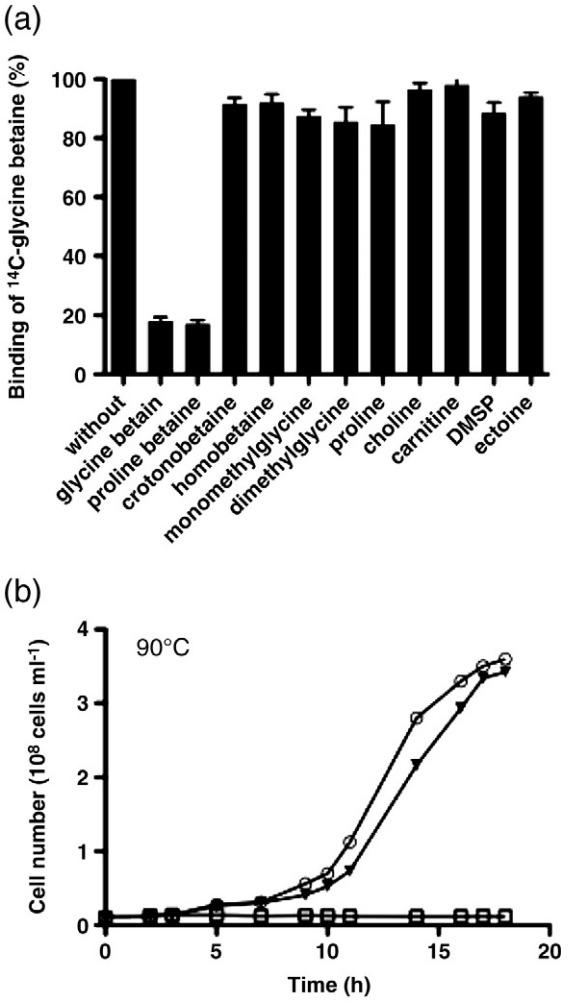

Fig. 1. *Af*ProX-mediated thermoprotection of *A. fulgidus*. (a) Substrate specificity of *Af*ProX. The substrate specificity of ProX was investigated in a competitive radioactive binding assay. We incubated 5 μM *Af*ProX in 100 μl of 10 mM Tris–HCl (pH 7.5) at 85 °C with 5 μM [1-^{14}C]GB and 40-fold excess of different unlabeled compatible solutes. As control, the same experiment was performed with no competitive substrate, and that value was set to 100%. (b) Influence of GB and PB on the growth of *A. fulgidus* at elevated temperature. Pre-warmed media lacking or containing 2 mM GB or 2 mM PB were inoculated with 5% exponentially growing *A. fulgidus* cells and incubated on a rotary shaker at 90 °C. Squares: *A. fulgidus* cells grown in the absence of either GB or PB; diamonds: *A. fulgidus* cells grown in the presence of 2 mM GB; triangles: *A. fulgidus* cells grown in the presence of 2 mM PB.

two domains are separated from each other, thereby forming a deep solvent-exposed cleft, which harbors the substrate binding site. Upon ligand binding, both domains move toward each other through a hinge-bending motion or rotation, which results in the so-called closed liganded conformation. As a consequence of this movement, residues originating from both domains generate the ligand binding site and trap the ligand within a deep cleft formed by the two lobes of the SBP.[4] In the absence of a ligand, open and closed unliganded states of the SBP are in equilibrium, and the ligand solely shifts this equilibrium toward the closed liganded state. This sequence of events has been coined the "Venus flytrap mechanism",[6–8] which is supported by crystal structures in the absence and presence of a ligand[9,10] and other biophysical techniques.[4]

For the maltose binding protein (MBP),[2] it has been shown that both domains are dynamically fluctuating in the absence of the ligand.[11] NMR spectroscopy revealed that the ligand-free form of MBP consists of a predominantly open species (95%) and a minor species (5%) that corresponds to a partially closed state.[12] The open form of MBP observed by NMR is similar to the crystal structure of the open unliganded conformation.[13] However, the partially closed species detected by NMR[12] does not correspond to the ligand-bound, fully closed crystal form. Instead, it represents an intermediate, partially closed conformation,[14] suggesting that the substrate is required to reach the final, closed liganded conformation.

Millet *et al*. have shown that every degree of domain closure in the unliganded MBP requires conformational deformation energy of approximately 212 cal/mol. Since MBP closes by about 20°, this would lead to an energetic cost of ~4 kcal/mol when forming the closed unliganded conformation of MBP.[15] This closed unliganded state, first observed by X-ray crystallography, is therefore highly unfavorable and resembles a loaded spring.[15] Changing the stiffness of this region by introducing amino acids with bulky side chains changes the degree of closure of MBP, which directly correlates with maltose binding affinity.[15,16] As a consequence, mutations in the hinge region likely alter the equilibrium between the open and partially closed states of MBP.[12] Upon maltose binding, a further shift in this equilibrium occurs toward the fully closed liganded conformation, which becomes accessible because maltose forms favorable interactions with the protein.

Generally, the hinge region provides MBP with a large degree of flexibility. This allows MBP to bind a variety of malto-oligosaccharides, some of which lead to domain closure.[13,17] Although this highlights a distinct role of the substrate in domain closure, the exact nature of the substrate-induced closing mechanism of MBPs and SBPs in general is unknown. Furthermore, SBPs bind their substrates with high affinity,[4] but it is not fully understood how the substrate is released to its cognate ABC importer. One scenario could be that interactions between the ABC transporter and the SBP during the transport cycle lead to a modulation of the SBP's affinity to its substrate.[3]

Table 1. Mutational analysis of the binding site of *Af*ProX

Mutant	K_d GB (μM)	K_d PB (μM)
Wild type	0.06 ± 0.01	0.05 ± 0.01
K13A	107 ± 20	101 ± 12
Y63A	149 ± 17	288 ± 26
T66A	1.8 ± 0.2	18 ± 1.4
Y111A	76 ± 4	148 ± 28
E145A	2.7 ± 0.5	23.2 ± 4
F146A	6.6 ± 0.6	4.0 ± 0.5
Y190A	67 ± 9	19 ± 5
R149A	320 ± 60	n.b.
Y214A	3.5 ± 0.7	3.5 ± 0.7
Double tyrosine mutations	n.b.	n.b.

The apparent affinities of wild-type *Af*ProX and *Af*ProX mutants are summarized for GB and PB. No binding for any of the double mutants was observed; hence, these mutants are not listed in the table. n.b. indicates that no binding was observed under this experimental setup.

In the thermophilic archaeon *Archaeoglobus fulgidus*, the ABC importer ProU has been annotated as a specific uptake system for the compatible solutes glycine betaine (GB) and proline betaine (PB).[18] The corresponding SBP (*Af*ProX) has been crystallized in different conformations: a closed liganded conformation in complex with either GB or PB and in an open unliganded conformation.[19]

Here, we present the structure of *Af*ProX in the open liganded conformation together with a systematic mutational analysis of the binding site of the *Af*ProX protein and molecular dynamics (MD) simulations of liganded *Af*ProX, using starting structures of different conformational states. Based on our results, residue Arg149 is crucial to switch *Af*ProX from a low-affine open structure to a high-affine closed state. Our data imply that the "Venus flytrap" mechanism is composed of distinct molecular events to ensure the specific biological function of SBPs because the presence of the substrate is communicated through a network of amino acids located in both domains, triggering domain closure.

Results

Thermoprotection of the hyperthermophilic archaeon *A. fulgidus* by GB and PB

*Af*ProX possesses striking substrate specificity: from all of the osmoprotectants tested, only GB and PB are substrates (Fig. 1a). Both GB and PB typically function as osmostress protectants in microorganisms. However, we observed that these compounds had no osmoprotective effect for *A. fulgidus* (Supplementary Fig. 1). Therefore, the uptake of GB or PB must serve another physiological role for this hyperthermophilic archaeon. Since GB has been shown previously to serve as a thermoprotectant in different bacteria,[20] we explored a possible thermoprotective effect of GB and PB for *A. fulgidus*. Neither GB nor PB had any effect on the growth of *A. fulgidus* when it was cultivated at its optimal growth temperature of 83 °C (Supplementary Fig. 1). In contrast, both compounds exerted a strong thermoprotective effect for cells cultivated at the elevated growth temperature of 90 °C. Without the addition of GB or PB to the growth medium, *A. fulgidus* was unable to grow at this temperature (Fig. 1b). Such a striking thermoprotection of a hyperthermophilic archaeon has never been observed before.

Mutational analysis of the ligand binding site in *Af*ProX

*Af*ProX has previously been crystallized with substrate at a resolution of 2.1 Å and 1.9 Å, respectively.[19] The positively charged trimethylammonium head group of GB and the dimethylammonium head group of PB are wedged into an aromatic cage that is formed by the main-chain carbonyl of Asp109 and by four Tyr residues (Tyr63, Tyr111, Tyr119, and Tyr214). Here, cation–π interactions are key determinants for coordinating the head group. The carboxylic tails of GB and PB protrude from this cage and form two salt bridges and one hydrogen bond with Lys13, Thr66, and Arg149, respectively.[19] Binding of GB and PB to purified *Af*ProX influenced the spectroscopic properties of *Af*ProX. Substrate binding of these compatible solutes leads to an increase of the intrinsic tryptophan fluorescence. We exploited this feature to determine the apparent affinity constants (K_d) for both substrates. *Af*ProX binds GB and PB with high apparent affinity, 60 ± 10 nM and 50 ± 10 nM, respectively (Table 1). These data represent high, but not unusual, apparent affinities of SBPs[5,21] of ABC transporters. To further support our findings on the affinity of *Af*ProX for GB and PB, we also determined a similar binding constant of GB by isothermal calorimetry, revealing a K_d of 100 ± 30 nM (Supplementary Fig. 2). It is important to note that, with both techniques, we only observe the binding of the substrate, but we cannot differentiate between an initial binding of the substrate and subsequent domain closure.

According to the structure of *Af*ProX,[19] it is evident that the substrate binding pocket is built up by four tyrosines (Tyr63, Tyr111, Tyr190, and Tyr214) and Lys13, Thr66, and Arg149. To further understand the residues' roles in substrate binding, we mutated them to alanine, and we measured the substrate apparent affinity of the corresponding mutants (Table 1). Mutation of any of the four tyrosines reduced the apparent affinity, ranging from 3.5 ± 0.7 μM for Tyr214Ala to 149 ± 17 μM for Tyr63Ala, respectively (Table 1). Any one of the double Tyr-to-Ala mutants was not able to bind GB at all.

Table 2. Crystallographic parameters

Space group	P_1
a, *b*, *c* (Å)	33.6, 36.9, 57.8
α, β, γ (°)	83.8, 80.5, 95.9
Data collection and processing	
Wavelength (Å)	0.8726
Resolution (Å)	20–2.0
Mean redundancy	4.2 (4.3)
Unique reflections	17,040
Completeness (%)	92.6 (93.6)
I/sigma	47.7 (29.0)
R_{sym}[a]	2.1 (4.3)
Refinement	
R_F[b] (%)	17.8
R_{free}[c] (%)	24.4
Overall *B*-factor from Wilson scaling (Å^2)	20.9
RMSD from ideal	
Bond lengths (Å)	0.023
Bond angles (°)	0.956
Average *B*-factors (Å^2)	7.6
Ramachandran plot (%)	
Most favored	93.4
Allowed	6.2
Disallowed	0.4
Model content	
Monomers per asymmetric unit	1
Protein residues	6–275
Ligands	
GB	1
Hepes	1
H_2O	300

Crystal parameters and data collection statistics are derived from XDS.[24] Refinement statistics were obtained from REFMAC5.[25] Ramachandran analysis was performed using PROCHECK. Numbers in parentheses correspond to the highest-resolution shell (2.1–2.0 Å).

[a] R_{sym} is defined as $R_{sym} = \sum_{hkl} \sum_i |I_i(hkl) - \langle I(hkl) \rangle| / \sum_{hkl} \sum_I I_i(hkl)$.

[b] R_F is defined as $R_F = \sum_{hkl} |F_{obs}| - |F_{calc}| / \sum_{hkl} |F_{obs}|$.

[c] R_{free} is calculated as R_F but for 5% randomly chosen reflections that were omitted from all refinement steps.

The Lys13Ala and Thr66Ala mutants also showed a lower apparent affinity, 1.8 ± 0.2 μM and 107 ± 20 μM, respectively (Table 1). The largest decrease in apparent affinity in the case of a single mutation was observed for the Arg149Ala mutation. Here, the apparent K_d dropped to 320 ± 60 μM for GB, and no binding could be observed for PB (Table 1). This suggests a specific role of Arg149 in the substrate binding to *Af*ProX.

Crystal structure of Tyr111Ala *Af*ProX

The Tyr mutations of the aromatic box had an unexpectedly large effect on the apparent affinity of *Af*ProX for GB. Therefore, we crystallized one of these mutants (Tyr111Ala) and solved its structure in the presence of GB at 2.0 Å. Tyr111Ala has an apparent affinity of 76 ± 4 μM toward GB (Table 1). Crystals were grown as described in Material and Methods. A data set of the *Af*ProX Tyr111Ala in complex with GB was collected at beamline ID23 (European Synchrotron Radiation Facility, Grenoble, France) and processed using XDS.[22] Initial phases were obtained by molecular replacement using the program Phaser,[23] with the open unliganded *Af*ProX structure as template [Protein Data Bank (PDB) entry: 1SW5].[19] The final structure was refined to a resolution of 2.0 Å. Data, refinement statistics, and model content are summarized in Table 2.

The structure of Tyr111Ala *Af*ProX consists of two domains, the ligand GB and a Hepes molecule captured from the crystallization solution. Due to the binding of Hepes in close proximity, Asn16 adopts an unusual conformation (as highlighted in the Ramachandran plot). A comparison with the open liganded structure and a closed liganded structure revealed that the structure of Tyr111Ala *Af*ProX adopts an open liganded conformation, with the hinge region positioned between residues 109–111 and residues 213–215 (Fig. 2). The RMSD between the two, open and closed liganded structures of *Af*ProX, was analyzed by the DynDom server.[26] The RMSD after the superimposition of all C^α atoms of the Tyr111Ala structure with respect to the closed liganded structure of *Af*ProX (PDB code: 1SW2) is 4.75 Å, whereas only minimal structural differences were found with respect to the open unliganded structure (PDB code: 1SW5) (RMSD < 0.2 Å). In order to determine how much domain II moves with respect to domain I when going from the open to the closed structures, we superimposed residues 1–105 and 207–270 of domain I; subsequently, the RMSD was calculated for C^α atoms of domain II only, which yields 13.3 Å. When reporting the MD simulation results below, the latter RMSD definition is used.

A comparison between the open liganded structure and the open unliganded structure revealed that only one loop (amino acids 142–153) differs (RMSD, 1.6 Å). This loop contains Arg149 that is part of the substrate binding site in the closed liganded structure and will be referred to as "Arg loop".

The substrate binding site of the *Af*ProX Tyr111Ala variant

In the *Af*ProX Tyr111Ala mutant, the trimethylammonium moiety of GB is located at the same position as in the closed liganded structure (Fig. 3). Although GB lacks interactions with Tyr111, it is bound by Tyr63, Tyr190, and Tyr214 of the aromatic cage and further stabilized by interactions with Thr66, Asp109, and the piperazinyl ring of the bound Hepes. Notably, when compared to the closed structure of the wild-type protein, the carboxyl tail of GB is rotated by almost 180° in the *Af*ProX Tyr111Ala mutant. Here, GB interacts through a water molecule with Thr66 and through two water molecules with the side chain of Tyr190. As the carboxyl group of GB

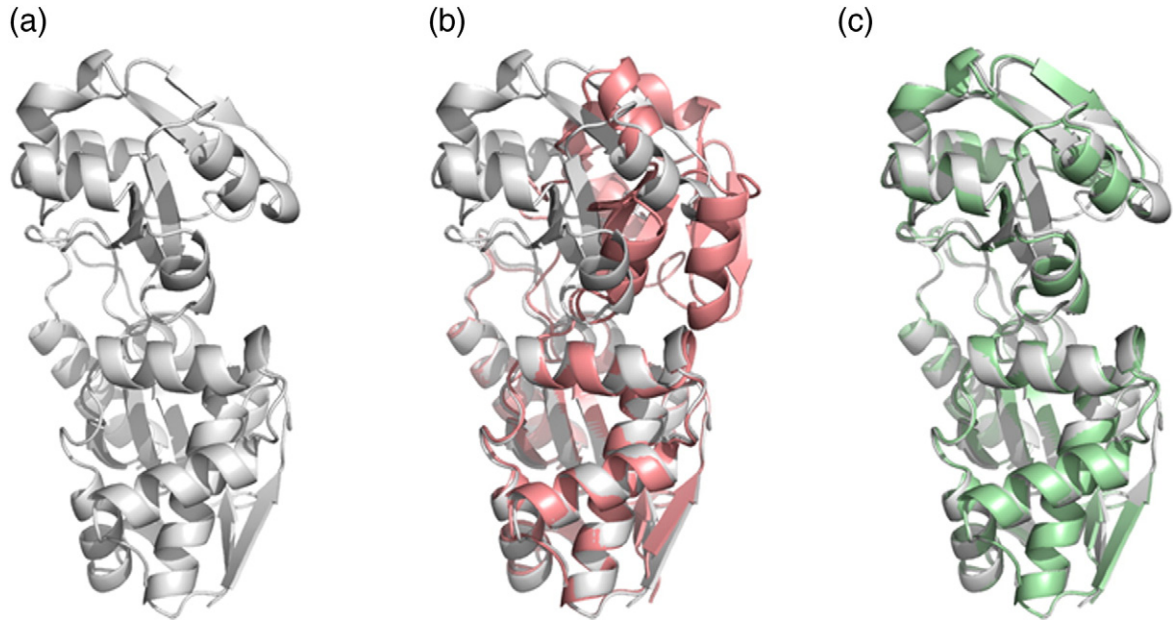

Fig. 2. Overlay of Tyr111Ala *Af*ProX with the closed liganded and the open unliganded structures of *Af*ProX. (a) Overall structure of *Af*ProX highlighted as cartoons. (b) Overlay of the Tyr111Ala *Af*ProX structure with the closed liganded structure of *Af*ProX (pink; PDB code: 1SW2). (c) Overlay of the Tyr111Ala *Af*ProX structure with the open unliganded structure of *Af*ProX (light green; PDB code: 1SW5). The overlays were calculated using LSQMAN.

forms a neither direct nor water-mediated interaction with the Hepes molecule. Therefore, the orientation of GB in the crystal structure of the *Af*ProX Tyr111Ala variant should not be influenced.

The binding sites in the closed liganded, open liganded, and open unliganded structures of *Af*ProX are closely related because Tyr63, Tyr214, Lys13, Thr66, and Asp109 are superimposable in all three structures (Fig. 4a and b). However, the side chain of Tyr190 is flipped in the closed structure by 130° when compared to both open structures. Interestingly, the position of the Tyr111 side chain in the closed liganded structure is located at the position of the side chain of Tyr190 in the open unliganded and open liganded structures (Fig. 4a and b). This implies that, during the opening and closing movements of *Af*ProX, the binding site does not undergo a major conformational change; rather, it is largely preformed in the open unliganded structure and, thus, pre-dispositioned to capture GB. Still, the

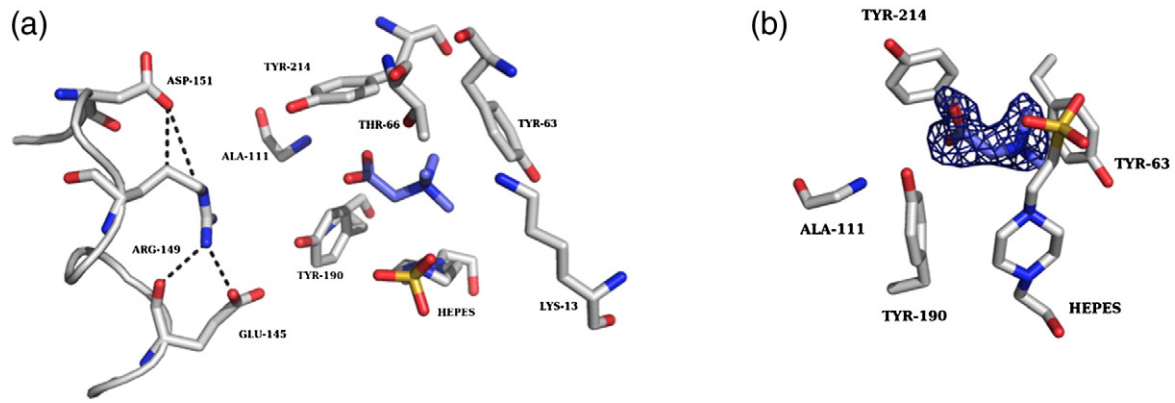

Fig. 3. The binding site of Tyr111Ala *Af*ProX. (a) The GB binding site of Tyr111Ala *Af*ProX is shown in sticks as is Arg149 with its neighboring residues Asp145 and Glu151. Broken lines depict distances <3.5 Å between Arg149 and Asp145 or Glu151. Bound GB is highlighted in blue. (b) A simulated annealing omit map is shown for GB. For clarity, only the tyrosine residues of the binding site are shown. Both figures have slightly different orientations in order to show also the density of the tail of GB.

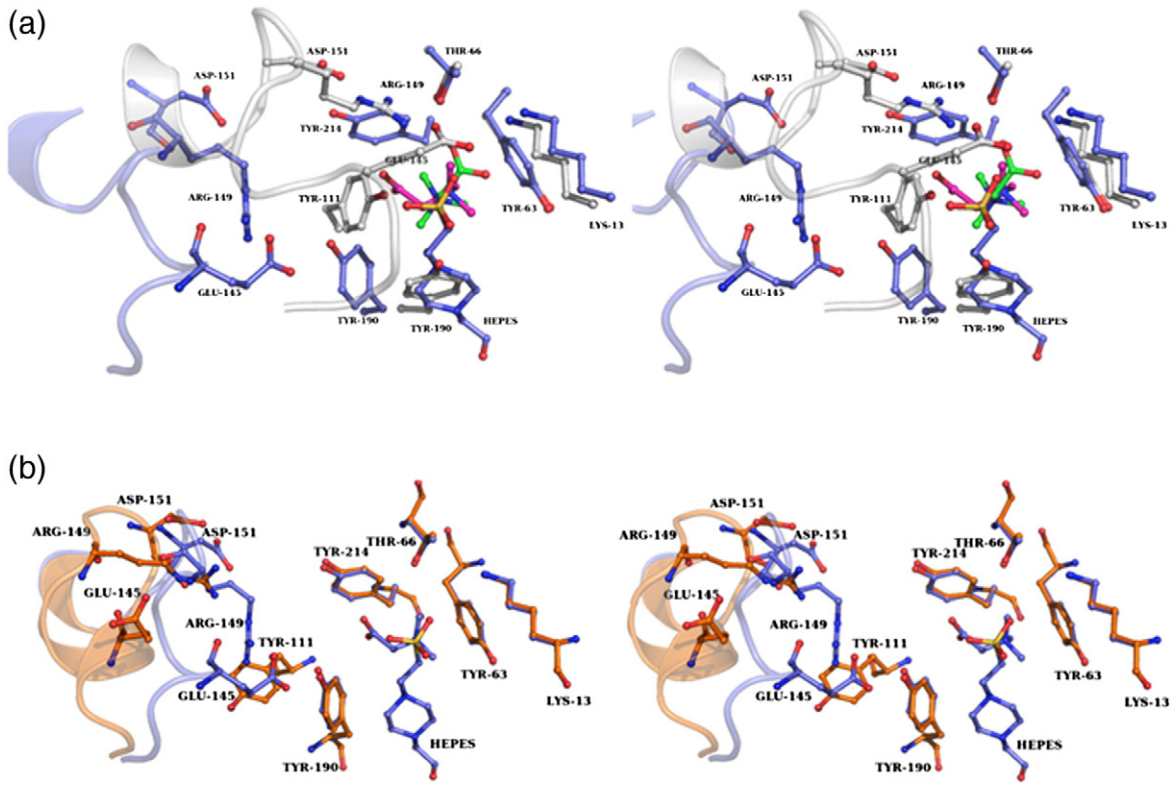

Fig. 4. Overlay of the Tyr111Ala *Af*ProX binding sites. (a) The binding site of Tyr111Ala *Af*ProX (shown in blue) is overlaid with the closed liganded structure (white). Residues involved in substrate binding are highlighted. In purple, the orientation of GB in the open liganded structure, and in green, the orientation as observed in the closed liganded structure. (b) The binding site of Tyr111Ala *Af*ProX (shown in blue) is superimposed with the open unliganded structure (orange). Residues involved in substrate binding are highlighted. For simplicity, Asp109 was not included in (a) and (b). We note that this residue superimposes perfectly in both structures.

side chains of Tyr111 and Tyr190 need to reorient. Furthermore, the largest movement is observed for the Arg loop, which moves by approximately 10 Å, starting from an outward-rotated orientation. This leads to side-chain interactions of Arg149 with GB, Tyr111, and Thr66, thereby complementing and locking the binding site.

The role of Arg149

In addition to the direct interaction with GB and residues that are part of the substrate binding pocket (Tyr111 and Thr66), Arg149 is a major determinant in domain–domain interactions in the closed structure (Table 3). As such, Arg149 interacts with Val70 (domain I) and Asp151 (domain II), thereby acting as a linking element between the two domains that enforces stable domain closure. These interactions complement those mediated by Pro172 of domain II, where Pro172 interacts via its C^{α} atom and a water molecule with Glu155 of domain II. Together, this provides a further explanation for the crucial role of Arg149 for the stability of the closed liganded state.

Table 3. Summary of the interactions of domain I and domain II in the closed form of *Af*ProX

Amino acid	Atom	Interaction partner	Atom	Distance (Å)
Pro172	C^{δ}	Pro14	O1	3.4
	C^{η}		O1	3.3
	C^{β}	Tyr190	OH	3.6
	C^{η}		OH	3.6
Glu155	NE1	Water–Pro76	CG	3.3
			C^{α}	3.3
Glu145	$O^{\varepsilon 1}$	Thr45	N	2.8
	$O^{\varepsilon 2}$	Thr45	C^{η}	2.5
Asp151	OD2	Tyr214	CZ	2.6
Arg149	NH1	Tyr111	OH	3.5
	NH1	Betaine	O2	3.1
	NH1	Glu145	OE2	3.0
	NH2	Tyr111	OH	3.8
	NH2	Thr66	OG1	2.8
	NH2	Thr66	C^{β}	3.5
	NH2	Thr66	C^{α}	3.6
	CZ	Val70	$C^{\eta 2}$	3.6
	NE	Val70	$C^{\eta 2}$	3.3
	NE	Asp151	OD1	2.8
	NH2	Tyr111	OH	3.8

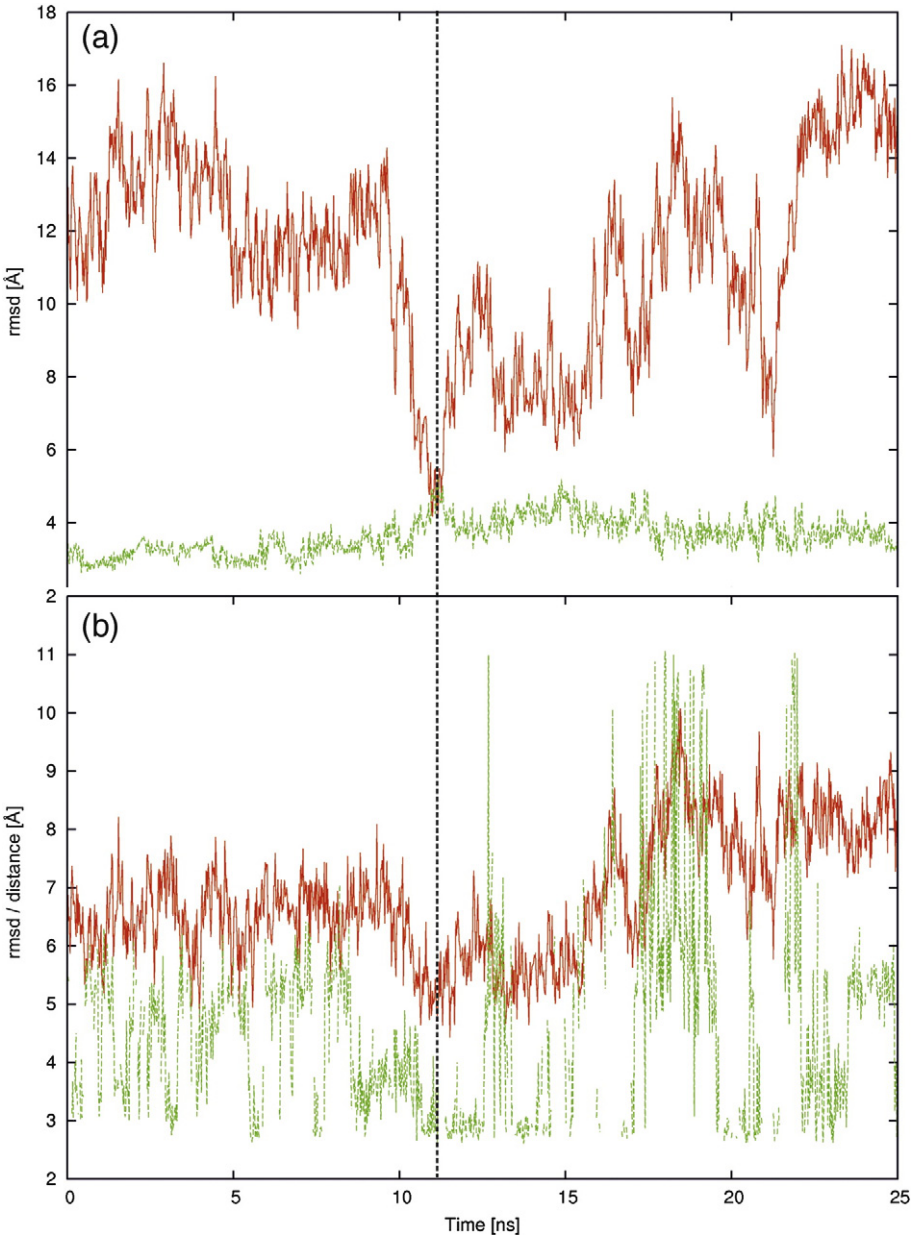

Fig. 5. MD simulation started from the open liganded wild-type *Af*ProX structure tethering GB in the binding site. (a) RMSD of domain II with respect to the closed liganded structure (red); RMSD of GB with respect to the starting structure (green); (b) RMSD of Tyr111 and Tyr190 with respect to the closed structure (red); minimal distance between O^{ε} atoms of Glu145 and NH atoms of Arg149 (green). The broken line indicates the simulation time at which the conformation shown in Fig. 6 was extracted.

Partial domain closure of *Af*ProX occurs even in the absence of GB

To study the role of Arg149 in domain closure of *Af*ProX in further detail, we performed MD simulations starting from the open liganded *Af*ProX structure with bound GB. Here, Ala111 was mutated *in silico* to Tyr in the Tyr111Ala *Af*ProX crystal structure. After 50 ps of MD simulation, GB reoriented from the configuration observed in the Tyr111Ala crystal structure to the configuration seen in the closed structure (Supplementary Fig. 4). In the initial configuration, only water-mediated interactions are formed between the carboxyl moiety of GB and *Af*ProX. After the reorientation, the carboxyl moiety of GB forms a hydrogen bond with Thr66 and a salt bridge with Lys13. Furthermore, after 4 ns of MD simulation, GB dissociates from *Af*ProX and does not re-associate during another 196 ns of simulation time (Supplementary Fig. 5). Notably,

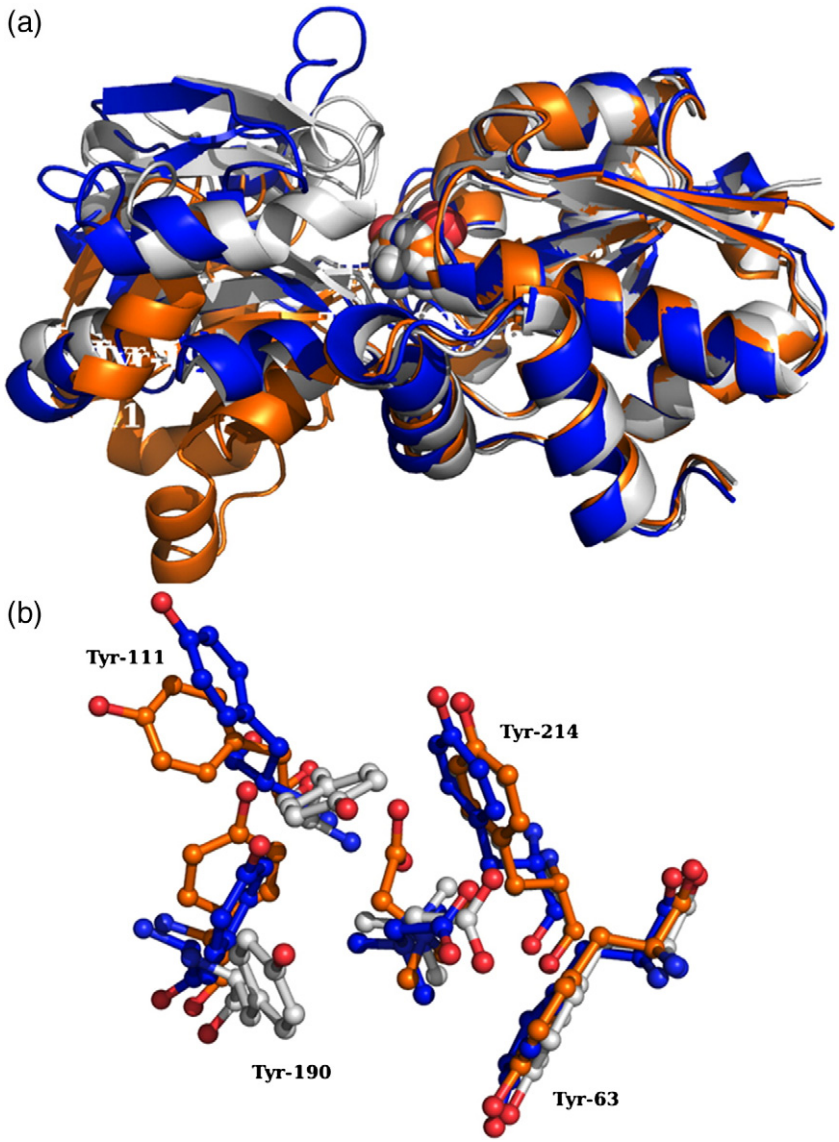

Fig. 6. Overlay of the open structure (orange), the closed structure (white), and the conformation extracted after 11 ns from the MD simulation started from the open liganded wild-type *Af*ProX structure, in which GB was tethered in the binding site (blue). (a) Domain II of the extracted conformation comes as close as 4.2 Å RMSD to the closed structure. (b) Close-up view of the tyrosine pocket.

after 50 ns of MD simulation, unliganded *Af*ProX undergoes a drastic conformational change: domain II comes as close as 2 Å RMSD to the configuration found in the closed *Af*ProX structure (Supplementary Fig. 5). During the remaining 150 ns of MD simulation, domain II repeatedly opens again but always returns to the configuration of the closed structure. In summary, MD simulations that started from the open *Af*ProX structure reveal that a tendency of domain closure exists even in the absence of GB. Such motions are inline with the "Venus flytrap mechanism"[4] and have been experimentally verified for MBP.[12]

Coupled structural reorganization in the binding site upon domain closure

In order to determine the coupling of conformational changes in the ligand binding site that occurs upon domain closure in the presence of GB, we repeated MD simulations starting from the open liganded *Af*ProX structure. However, this time, GB was restrained to the ligand binding site to prevent dissociation. Again, a domain closure was observed, with domain II coming as close to the ligand-bound state as 4.2 Å RMSD (Figs. 5a and 6a). Notably, this state is reached after 11 ns, about five times faster

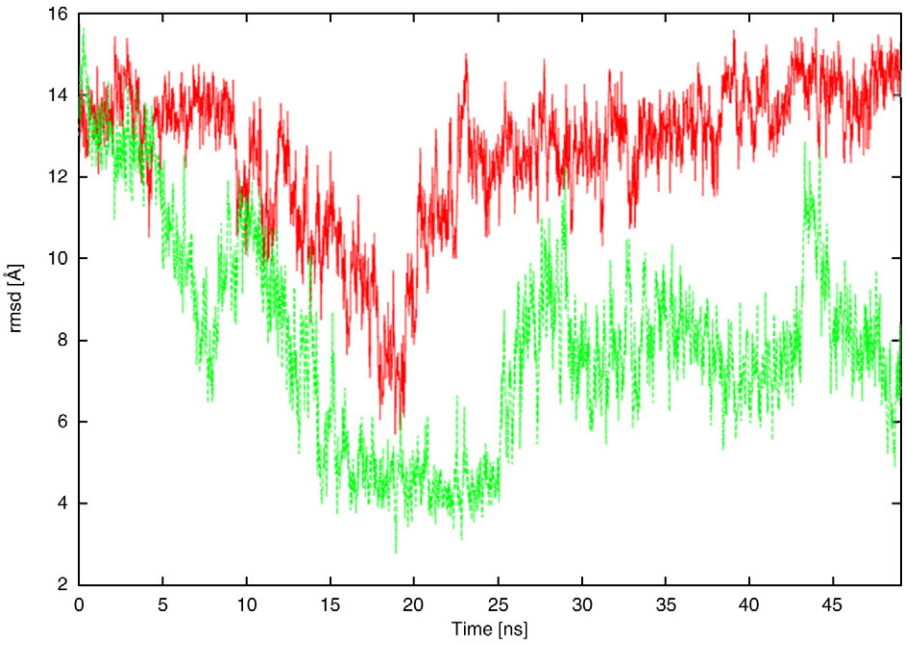

Fig. 7. MD simulation started from the closed liganded *Af*ProX structure for wild-type *Af*ProX (red) and Arg149Ala *Af*ProX (green). RMSD values of domain II with respect to the open *Af*ProX structure are shown.

than in the absence of GB (Fig. 5a). Although single events that occur only once during a MD simulation must be interpreted with caution, these results indicate that binding of GB facilitates conformational changes occurring during the transition from the open state to the closed state. Furthermore, domain closure is accompanied by a reorientation of GB (Fig. 5a). In parallel, the Tyr cage that coordinates the binding of the trimethylammonium moiety of GB via cation–π interactions starts to adopt the conformation observed in the closed structure (Fig. 5b). In this process, Tyr111 shifts about halfway between the conformations of the open and closed structures, whereas the conformation of Tyr190 is still close to the one observed in the open structure (Fig. 6b). For Tyr111 to reach its final position, Tyr190 must reorient. In addition, the minimal distance between any O^{ϵ} atom of Glu145 and any NH atom of Arg149 fluctuates between 2.8 Å and ~6 Å during the first 8 ns of simulation time (Fig. 5b), indicating the temporary formation of a salt bridge. This salt bridge consolidates during the next 2 ns, after which it is almost permanently formed for another 2 ns. Upon domain opening after about 12 ns of

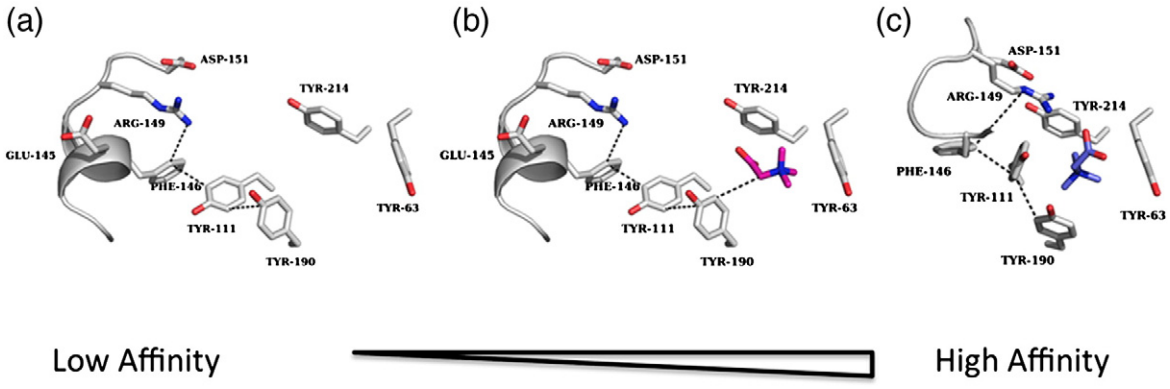

Fig. 8. Proposed mechanism of *Af*ProX during substrate capturing and high-affinity binding. (a) The open conformation of the binding site of *Af*ProX is shown. (b) The ligand GB is captured and bound in the binding site. (c) Binding of the substrate triggers closure of the two domains via the Tyr190-Tyr111-Phe146 network. In the closed structure, Arg149 interacts with GB (blue; broken line). All highlighted distances (broken lines) are between 3.2 Å and 3.8 Å.

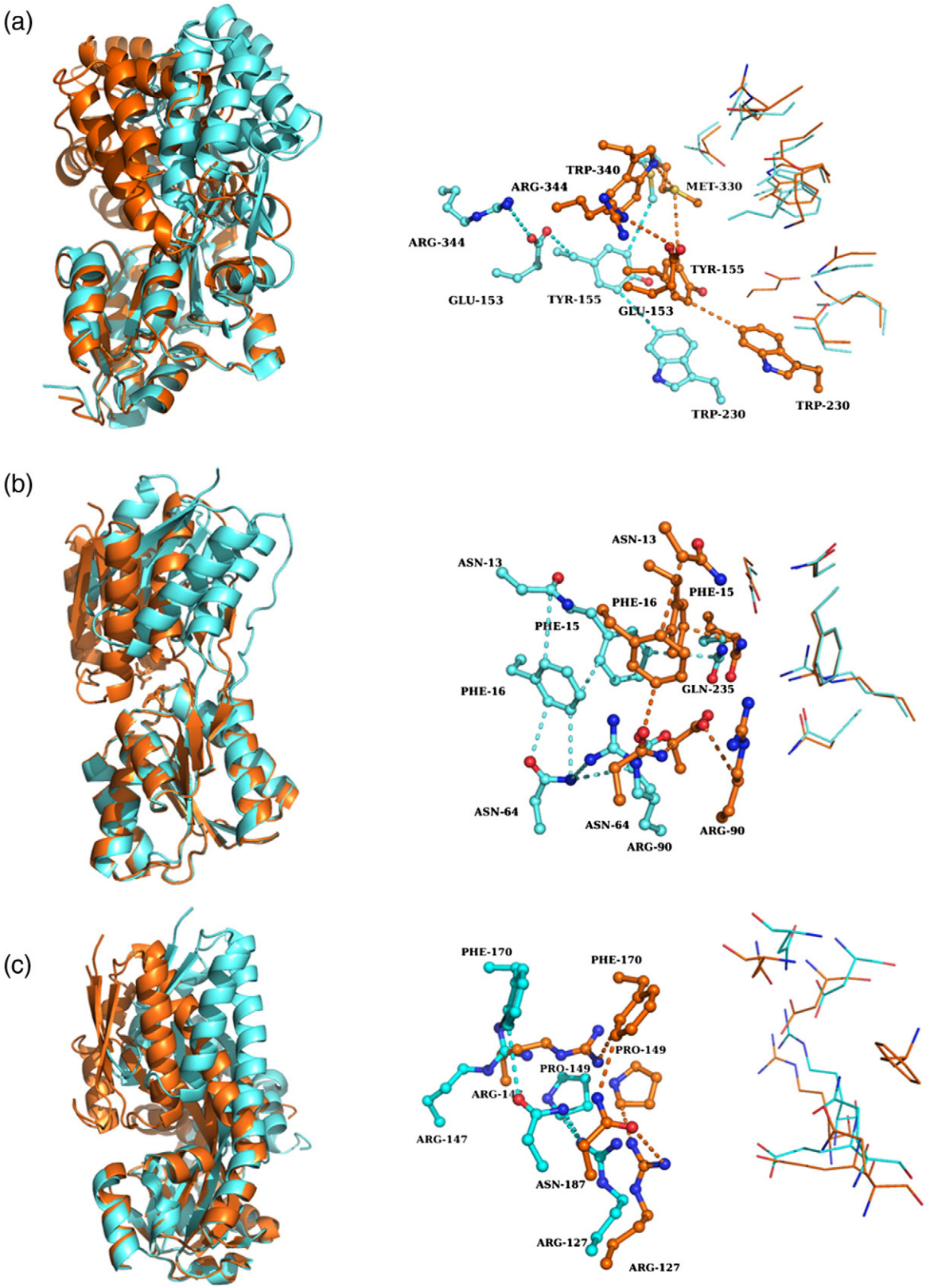

Fig. 9. (*legend on next page*)

simulation time, the interaction between Glu145 and Arg149 breaks up, and Glu145 and Arg149 remain fluctuating between a salt-bridge-formed state and a complete separation for the remainder of the simulation time. Thereby, maxima of the domain II–RMSD curve (Fig. 5a) are significantly paralleled by large Glu145–Arg149 distances (Fig. 5b), again demonstrating in a remarkable way the coupling of events that lead to domain closure and opening of the *Af*ProX protein.

To further investigate the central role of Arg149 for the domain closure, we performed unrestrained MD simulations starting from closed liganded wild-type and Arg149Ala *Af*ProX structures. Over the course of 50 ns of simulation time, the wild-type structure starts to open up once but immediately returns to the closed conformation (Fig. 7). In contrast, opening of the domains is observed earlier, more frequently, and more pronounced in the Arg149Ala structure. Most notably, after opening, the Arg149Ala structure never adopts its closed conformation again.

Substrate binding triggers domain closure to a closed state with a "loaded spring"

The above-described experiments unambiguously established the importance of Arg149 for the high-affinity binding of GB (or PB) to *Af*ProX. Furthermore, GB provides protection against thermal increases in the habitat of *A. fulgidus* (see above), and the response to such fluctuations must be fast. Thus, binding and release of GB (or PB) to and from *Af*ProX must be fast. This raises the question if an initial binding of GB (or PB) can trigger a domain closure leading to a closed conformation stabilized by Arg149. To identify such a possible triggering mechanism, we analyzed interactions of the bound substrate in the open liganded structure with Arg149. Already in this state, the presence of GB is communicated to Arg149 through interactions of the side chains of Tyr190, Tyr111, and Phe146 via a direct side-chain network. Interestingly, this network does not change during the transition from the open state to the closed state as judged by unaltered distances of these amino acids in both forms.

Together with the information from MD simulations, this suggests the following substrate-induced triggering mechanism in *Af*ProX. As highlighted in Fig. 8a, the open conformation appears to be stable and ready to bind GB (or PB) due to the largely preformed ligand binding site. The next state, after GB binding, is represented by the open liganded structure of *Af*ProX presented here (Fig. 8b). In this state, GB interacts with Tyr190 (distance between the C^{α} atom of GB and the $C^{\varepsilon 1}$ atom of Tyr190: 3.8 Å). Immediately after binding, GB reorients around its trimethylammonium group to the configuration seen in the closed structure and observed in the MD simulation. This reorientation leads to GB attracting the side chain of Tyr190 and Tyr111. The additional cation–π interactions between Tyr111 and Tyr190 and GB's trimethylammonium group are energetically favorable; their formation may thus be the cause why domain closure is facilitated in the ligand-bound state compared to the unliganded structure as observed in the MD simulations. The Tyr111 and Tyr190 movements are relayed to Arg149 via the above-described interactions. Ultimately, the reorientation of GB triggers domain closure and induces the high-affine closed liganded conformation in which Arg149 complements and locks the binding site and stabilizes the closed state (Fig. 8c).

During domain closure, part of the Arg loop changes from an α-helical conformation in the open structure to an unwound conformation in the closed liganded structure. This change leads to a "loaded spring", which is restrained by a salt bridge of Arg149 (domain II) with the carboxyl group of GB. GB itself forms additional salt bridges with Lys13 and Thr66 of domain I. If the Arg149–GB interaction is disrupted, the Arg loop will adopt its energetically favored, "relaxed" helical conformation. This results in opening of *Af*ProX and release of the substrate. As a consequence, *Af*ProX should not bind solutes that lack a carboxyl group (e.g., choline). Since the exact positioning of the carboxyl group is important, compatible solutes with increased size should not induce a stable closed conformation, and this is precisely what is observed in competition assays (Fig. 1a). Hence, our structural analysis provides a rational for the observed substrate specificity of *Af*ProX.

Discussion

High-affinity ligand binding by the *Af*ProX SBP

The crystal structures of wild-type *Af*ProX in complex with GB and PB and an open unliganded

Fig. 9. Substrate-induced triggering mechanism in SBPs. By comparing the open unliganded and closed liganded structures of SBP, we can detect a triggering mechanism. Left panel: the superimposition of the overall structure shown in cartoon representation using domain I as anchor point. The open unliganded structure is colored cyan, and the closed liganded structure is colored orange. Right panel: the binding site of both superimposed structures is shown and colored the same as above. The interactions are shown as broken lines with the same color coding. The amino acids involved in substrate binding were compared, and a network was deduced based on the interaction distances which are at most 4 Å. For clarity, the substrate in the closed binding site is not shown. (a) MBP from *E. coli*; (b) ribose binding protein from *E. coli*; (c) SIAP from *H. influenzae*.

configuration were previously reported.[19] The latter structure is thought to represent the open conformation of *Af*ProX as suggested by the "Venus flytrap model".[4] *Af*ProX exhibits high affinities for its substrates, and it is highly substrate specific for GB and PB (Fig. 1a), which serve as thermoprotectants for *A. fulgidus* (Fig. 1b). To our knowledge, thermoprotection of a hyperthermophilic archaeon by an externally provided compatible solute has never been described before, although both compatible solutes possess thermoprotective properties for mesophilic bacteria.[20]

We probed the contribution of individual residues within the substrate binding site of *Af*ProX to ligand binding and found that mutating the Tyr cage reduced the apparent ligand binding affinity between 60-fold (Tyr214Ala; apparent $K_d = 3.5 \pm 0.7$ μM) and 2400-fold (Tyr63Ala; apparent $K_d = 149 \pm 17$ μM). Mutating any two Tyr residues in the aromatic cage to Ala abrogated ligand binding (Supplementary Table 1). Similar effects have been observed for the ectoine binding protein EhuB from *Sinorhizobium meliloti*,[27] GB/PB-specific OpuAC from *Bacillus subtilis*,[28] and ProX from *Escherichia coli*.[29] In each of these SBPs, cation–π interactions also contribute strongly to ligand binding.

Switching *Af*ProX from low affinity to high affinity

In *Af*ProX, Arg149 plays a distinct role in substrate binding, as its substitution by an Ala residue resulted in a loss of the apparent binding affinity (apparent $K_d \approx 50$–70 nM to apparent $K_d \approx 300$–400 μM). In the open liganded structure, Arg149 is located ~10 Å away from its position in the closed liganded structure. Hence, the ligand binding site is laid open, and the substrate has free access to the binding pocket. Our data suggest that Arg149 is the final amino acid to interact with the substrate and, thereby, captures GB or PB in the high-affinity closed state of *Af*ProX.

Domain closure triggered by substrate binding

The "Venus flytrap" model describes the opening and closing of SBPs, where the equilibrium between these two conformations is shifted toward the closed state upon substrate binding.[30] Many crystal structures of SBPs have been solved in the open unliganded, closed liganded, and, more rarely, in the closed unliganded states.[4,10,19,31] In striking contrast, only a few proteins have been crystallized in an open liganded conformation. Examples are the LIV (*l*eucine-*i*soleucine-*v*aline) binding protein, where the substrate was soaked into preformed crystals of the unliganded protein[30] and the open liganded form of MBP.[13] In the case of MBP, the nonnatural substrate β-cyclodextrin is bound, which is not transported, however.[13] In the open liganded structures of LIV or MBP, the binding site is completely accessible, and the substrate is bound by amino acids that originate from only one domain. In the case of MBP, it was therefore postulated that maltose interacts first with aromatic groups of the C-terminal domain in open state.[32] Subsequently, the complex converts into the stable closed conformation. However, MD studies of MBP demonstrated that the unliganded protein can adopt a wider variety of conformations than the maltose-bound state, indicating that ligand binding not only affects the relative orientation of the two domains but also attenuates their movement relative toward each other.[33] In contrast to the situation described for MBP, we found that the substrate binding site of *Af*ProX is preformed in the open conformation. Thus, only a slight reorientation has to occur to optimize the binding site toward its closed conformation. Still, although the open conformation contains a largely assembled binding site, it exhibits only low binding affinity toward GB or PB. Our mutational analysis of *Af*ProX revealed a distinct role of Arg149 in substrate binding: the movement of this residue, which complements the binding site in the closed *Af*ProX structure, switching *Af*ProX from a low-affine state to a high-affine state by locking the binding site and mediating domain–domain interactions, thus finally trapping the ligand.

Comparison with other SBP

In this study, Arg149 plays an important role for the binding affinity of GB in *Af*ProX because it changes the binding site from low affinity to high affinity, and it is involved in domain–domain interaction and domain closure. Especially, the triggering mechanism postulated here is intriguing. The question arises whether this is a general phenomenon observed in other SBPs as well. Despite the fact that there are a large number of crystal structures known for SBP of ABC transporters and tripartite transporters, only a few have been crystallized in an open unliganded conformation and in a closed liganded conformation.

Therefore, we compared the open and closed structures of the available examples: the MBP and the ribose binding protein from *E. coli* and the *N*-acetyl-5-neuraminic acid binding protein (SIAP) from *Haemophilus influenzae*.[34] The first two binding proteins belong to the ABC transporter family, whereas the latter one belongs to the TRAP (*tr*ipartite *A*TP-independent *p*eriplasmic) transporter binding protein family. Domain I of each system was chosen as an anchor point for the superposition of both conformations; this domain showed an almost perfect overlay in all three cases (RMSD < 0.8 Å).

Comparing the substrate binding site of MBP in its open unliganded and closed liganded structures reveals that Met330 undergoes only a small movement of 2.5 Å. Similar to Arg149 in *Af*ProX, Met330 is

connected, via a network of side-chain interactions, to those amino acids that participate in substrate binding. This network forms through the following amino acids: Met330-Tyr155-Glu153-Arg344. Furthermore, Tyr155 interacts with Trp230 and Phe156. The distances within this network do not differ in the open and closed structures; however, the absolute positions of these amino acids change by 6–10 Å, suggesting a rigid-body movement. That way, the small movement of Met330 seems to trigger a pulling event toward all amino acids involved in substrate binding in the second MBP domain. In addition, Met330 also interacts with Phe258, which is located at the beginning of a β-sheet. It has been shown that these β-sheets are important for stabilizing the closed conformation as it contains Gly260, which is part of a salt bridge only observed in the closed conformation of MBP[33] (Fig. 9a).

In ribose binding protein, the substrate is bound by several amino acids originating from both domains. Here again, the superimposition of the open unliganded structure and the closed liganded structure revealed a perfect overlay of domain I. In domain II, however, the amino acids undergo a significant movement. Here, a network that starts at the side chain of Gln235 and, in addition, contains Phe15, Phe16, Asn64, Arg90, and Asp89 was identified. Gln235 is also in contact with the substrate in the liganded structure, similar to Met330 in MBP and to Tyr110 in *Af*ProX, which undergoes a small conformational change of 1.5 Å during closing and is located directly at the hinge between both domains. This explains the small movement. In total, this highlights the importance of this single amino acid for ligand binding, as seen for Met330 (MBP) and Arg149 (*Af*ProX), too (Fig. 9b).

SIAP is the binding protein of the *N*-acetyl-5-neuraminic acid TRAP transporter from *H. influenzae*. Here, a specific role in ligand binding can be assigned to Arg127. Besides the interaction with the substrate in the closed conformation, a side-chain interaction network is observed through interactions with Pro149 and Asn187. Asn187 itself is in contact with Phe170 and Arg147, which are the residues of domain II that build up the substrate binding site in SIAP. Here and in MBP and RBS, the distances in the network do not change in the open and closed forms of the structures. In SIAP, Arg127 has been mutated to Ala and Lys. In competition assays, the activity of SIAP has altered for Arg127Ala and Arg127Lys. Furthermore, no binding of a substrate to the Arg27Ala mutant could be observed anymore by isothermal calorimetry[35] (Fig. 9c).

Substrate release by external modulation of Arg149 interactions

The Arg loop of the *Af*ProX protein has an α-helical conformation in the open structure. In the closed liganded structure, however, this helix is unwound. The energetic cost of unwinding is diminished by multiple interactions of Arg149 with domain II of *Af*ProX and GB. Furthermore, in the closed liganded structure, the side chain of Arg149 interacts with Glu145, whereas no such interaction occurs in the open structure. Finally, salt-bridge formation between Arg149 and Glu145 upon domain closure is also observed during the MD simulations (see above). Interestingly, Glu145 is exposed at the membrane-facing side of *Af*ProX.

Taken together, our data suggest a model in which a conformational change in the membrane components (ProW-1, ProW-2) and the ATPase (ProV) of the ProU ABC transporter from *A. fulgidus* might result in the manipulation or even in the disruption of the Glu145–Arg149 salt bridge. This event is expected to weakened the interactions of Arg149 with GB or PB, which, in turn, would induce a reformation of the helical conformation of the Arg loop. Thereby, the binding site of the *Af*ProX protein would become accessible. Consequently, the affinity of the substrate binding site would change from low nanomolar to medium micromolar range, inducing the release of GB/PB into the substrate translocation pathway of the ProU transporter for ATP-dependent import.

In summary, Arg149 acts as a trigger to ensure a fast and efficient ligand capture. A preformed ligand binding site that only requires a single amino acid to switch the system from a low- to a high-affine state is a perfect solution to environmental restraints. Our data imply that distinct molecular events such as a single amino acid switch are operational in the overall "Venus flytrap" mechanism to ensure the specific biological function of SBPs.

Materials and Methods

Heterologous expression and purification of *Af*ProX

*Af*ProX was heterologously expressed using the *E. coli* strain BL21 CodonPlus RIL strain containing the *proX*$^+$ plasmid pHG26, a derivative of the pASKIBA6 expression vector (IBA, Göttingen, Germany). A 10-l glass flask containing 5 l of MMA supplemented with 150 μg ml^{-1} of ampicillin, 30 μg ml^{-1} chloramphenicol, 0.5% glucose, and 0.2% casamino acids was inoculated to an OD_{578} of 0.1 from an overnight culture. Cells were grown at 37 °C with vigorous stirring until the culture had reached OD_{578} = 0.6–0.7. Expression was induced by the addition of anhydrotetracycline (final concentration, 0.2 μg ml^{-1}). Cells were grown for an additional 2 h and were subsequently harvested by centrifugation (10 min, 3000*g*). To release the periplasmic proteins, we resuspended the cell paste in 50 ml ice-cold buffer P [100 mM Tris–HCl (pH 8), 500 mM sucrose, and 1 mM ethylenediaminetetraacetic acid]. After a 30-min incubation on ice, the periplasmic protein extract was harvested by

centrifugation (15 min, 21,000*g* at 4 °C), followed by an ultracentrifugation step (30 min, 120,000*g* at 4 °C) to remove insoluble material. The supernatant was then loaded onto a 10-ml Strep-tactin column (IBA) equilibrated with buffer W [100 mM Tris–HCl (pH 8)]. The column was washed with 5 column volumes of buffer W, and bound proteins were released from the affinity resin by washing the column with buffer E [100 mM Tris–HCl (pH 8) and 2.5 mM desthiobiotin]. Since a portion of the pro-OmpA signal sequence was not proteolytically removed from the ProX protein by the *E. coli* cells, the purified ProX protein was cleaved with factor Xa (1 μg factor Xa per 200 μg ProX) for 16 h at room temperature in a buffer containing 100 mM Tris–HCl (pH 8), 100 mM NaCl, and 1 mM $CaCl_2$. This proteolytic cleavage removed the OmpA signal sequence and the Strep-tag sequence, thereby resulting in *Af*ProX containing no N-terminal extensions. To remove factor Xa from the solution, we diluted the protein solution 1:2 with deionized water and loaded them on a UnoQ6 column (Bio-Rad, Muenchen, Germany) equilibrated with 50 mM Tris–HCl (pH 8) (buffer A). The column was washed with 20 ml buffer A, and the protein was eluted with a linear NaCl gradient. *Af*ProX was eluted at 250 mM NaCl. *Af*ProX was subsequently dialyzed overnight against 5 l of 10 mM Tris–HCl (pH 7.5) at 4 °C and stored until further use at 4 °C. In general, 2 mg of pure *Af*ProX protein was obtained per 1 l of overproducing cells.

Substrate binding of the purified *Af*ProX protein

To study substrate binding properties of *Af*ProX, we used the ammonium sulfate precipitation technique of Richarme and Kepes.[36] *Af*ProX (5 μM) was incubated at 85 °C (or at other temperatures as indicated) with 5 μM [1-^{14}C]GB in a volume of 100 μl of 10 mM Tris–HCl (pH 7.5) for 5 min. *Af*ProX was then precipitated by the addition of 900 μl of an ice-cold saturated ammonium sulfate solution. After a 5-min incubation on ice, the mixture was sucked through cellulose filters (0.45-μm pore size; Schleicher and Schuell, Dassel, Germany), and the radioactivity retained by the filters was determined by liquid scintillation counting. To determine the substrate specificity of *Af*ProX, we incubated the protein at 85 °C with 5 μM radiolabeled GB and a 40-fold excess of different unlabeled compatible solutes, and we then further treated the samples as described above.

Determination of the apparent binding affinity (K_d) of *Af*ProX using fluorescence spectroscopy

All spectra were obtained at room temperature with a VARIAN CARY Eclipse fluorometer (VARIAN, Darmstadt, Germany). Emission scans were collected at an excitation wavelength of 280 nm from 290 nm to 450 nm. Measurements were carried out in 1 ml of 10 mM Tris–HCl (pH 7.5). Five micrograms of *Af*ProX protein was added to the buffer, and the solution was mixed. To titrate the protein, we added 5 μl substrate from a 40-μM GB or PB stock solution. After substrate addition, the sample was mixed in the cuvette, and the change in the fluorescence was recorded. This step was repeated until a stable fluorescence signal was obtained. Assuming one binding site per monomer in *Af*ProX, the K_d for substrate binding was determined by nonlinear least-squares fitting of the data to the following equation: $F = F_0 + (\mathrm{DF}/2P_0)[(K_d + P_0 + L_0) - ((K_d + P_0 + L_0)^2 - 4L_0P_0)^{1/2}]$, which correct for the concentration of the receptor. F is the measured fluorescence; F_0 is the fluorescence of *Af*ProX; DF is the change in fluorescence at saturation; P_0 and L_0 are the total concentrations of protein and substrate. The change in fluorescence, however, gives no information about the open or closed state of the protein. We can only observe binding of the substrate without any information of the conformation of *Af*ProX.

Crystallization

*Af*ProX crystals were obtained using the hanging-drop method at 293 K against a reservoir solution of 100 mM Hepes (pH 7.0), 15–25% polyethylene glycol 6000, and 15–25% polyethylene glycol 8000. *Af*ProX was incubated with 10 mM GB prior to crystallization. Crystals grew slowly and appeared after 1 month. They were flash frozen in liquid nitrogen with mother liquor supplemented with 20–25% ethylene glycerol as cryoprotectant.

Data collection, refinement, and structure analysis

*Af*ProX crystals diffracted X-rays beyond 1.6 Å. However, for refinement purposes, the data set was truncated at a resolution of 2.0 Å. The data set was collected at the ID23-EH2 beamline at the European Synchrotron Radiation Facility and processed with XDS.[22] *Af*ProX in the ligand-free form (PDB code: 1SW5)[19] was used as a template to obtain initial phases using Phaser.[23] The structure was further refined using REFMAC5[25] and Coot.[37] Data set and refinement statistics are listed in Supplementary Table 2. As analyzed with PROCHECK,[38] the Ramachandran plot of *Af*ProX shows one residue in the disallowed region. However, this residue interacts with the bound Hepes molecule, which explains the rather unusual conformation. Figures of protein molecules were prepared using PyMOL‡.

MD simulations

MD simulations were performed with the Amber 10 suite of programs,[39] together with the force field as described by Cornell *et al.*[40] using modifications suggested by Simmerling *et al.*[41] In total, four different MD simulations were performed: (i) open liganded wild-type *Af*ProX. The starting structure was generated by mutating Ala111 in the crystal structure of Tyr111Ala *Af*ProX to Tyr. The simulation length is 200 ns; (ii) open liganded wild-type *Af*ProX. In contrast to (i), a distance restraint by means of a harmonic potential was applied between the ammonium nitrogen of GB and the N of Asp109 of *Af*ProX. The simulation length is 25 ns; (iii) closed liganded wild-type *Af*ProX. The starting structure was taken from PDB code: 1SW4.[19] The simulation length is 50 ns; (iv) closed liganded Arg149Ala *Af*ProX. The starting

‡ www.pymol.org

structure was generated from the closed liganded wild-type *Af*ProX structure by mutating Arg149 to Ala. The simulation length is 50 ns.

In all cases, the starting structure was placed into an octahedral periodic box of TIP3P water molecules.[42] The distance between the edges of the water box and the closest atom of the protein was at least 11 Å, resulting in a system of ~43,000 atoms. The system was minimized by 50 steps of steepest descent minimization followed by 450 steps of conjugate gradient minimization. The particle mesh Ewald method[43] was used to treat long-range electrostatic interactions, and bond lengths involving bonds to hydrogen atoms were constrained using SHAKE.[44] The time step for all MD simulations was 2 fs, with a direct-space nonbonded cutoff of 8 Å. Applying harmonic restraints with force constants of 5 kcal mol^{-1} $Å^{-2}$ to all solute atoms, we carried out canonical ensemble (NVT)-MD for 50 ps, during which the system was heated from 100 K to 300 K. Subsequent isothermal isobaric ensemble (NPT)-MD was used for 150 ps to adjust the solvent density. Finally, the force constants of the harmonic restraints on solute atom positions were gradually reduced to zero during 100 ps of NVT-MD. The following NVT-MD at 300 K with a time constant of 10 ps for heat-bath coupling was used for analysis, with conformations extracted every 20 ps. Atomic charges for the GB ligand were generated following the RESP procedure;[45] force field parameters for GB were taken from GAFF.[46]

For the analysis of the trajectories, conformations were superimposed with respect to the C^{α} atoms of domain I. This resulted in an almost perfect overlay in all cases. RMSD values with respect to the open or closed starting structures were then determined for C^{α} atoms of domain II or for atoms of GB.

Accession number

Coordinates and structure factors have been deposited in the PDB with accession number 3MAM.

Acknowledgements

We thank the members of the laboratory of L.S. and E.B. for stimulating discussion. We are indebted to Christoph Müller Dieckmann for excellent support at the beamlines of the European Molecular Biology Laboratory Outstation Grenoble (France) and to Dr. Paul Tucker and Dr. Matthew Groves at the BW7A beamline, European Molecular Biology Laboratory Outstation Hamburg. Financial support for this study was provided by the Fonds der Chemischen Industrie and the LOEWE program of the State of Hessen via the Centre for Synthetic Microbiology (SynMicro; Marburg) to E.B. B.T. and L.S. gratefully acknowledge financial support by the EDICT EU program. H.G. is grateful to the "Zentrum fuer Informations- und Medientechnologie" at the Heinrich Heine University Duesseldorf for computational support.

Supplementary Data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.jmb.2011.05.039

References

1. Kempf, B. & Bremer, E. (1998). Uptake and synthesis of compatible solutes as microbial stress responses to high-osmolality environments. *Arch. Microbiol.* **170**, 319–330.
2. Davidson, A. L., Dassa, E., Orelle, C. & Chen, J. (2008). Structure, function, and evolution of bacterial ATP-binding cassette systems. *Microbiol. Mol. Biol. Rev.* **72**, 317–364; Table of Contents.
3. Cui, J., Qasim, S. & Davidson, A. L. (2010). Uncoupling substrate transport from ATP hydrolysis in the *Escherichia coli* maltose transporter. *J. Biol. Chem.* **285**, 39986–39993.
4. Wilkinson, J. & Verschueren, K. H. G. (2003). Crystal structures of periplasmic solute-binding proteins in ABC transport complexes illuminate their function. In *ABC Proteins: From Bacteria to Man* (Holland, I. B., Cole, S. P. C., Kuchler, K. & Higgins, C. F., eds), pp. 187–208, Academic Press (Elsevier Science), London, UK.
5. Berntsson, R. P., Smits, S. H., Schmitt, L., Slotboom, D. J. & Poolman, B. (2010). A structural classification of substrate-binding proteins. *FEBS Lett.* **584**, 2606–2617.
6. Quiocho, F. A. & Ledvina, P. S. (1996). Atomic structure and specificity of bacterial periplasmic receptors for active transport and chemotaxis: variation of common themes. *Mol. Microbiol.* **20**, 17–25.
7. Mao, B., Pear, M. R., McCammon, J. A. & Quiocho, F. A. (1982). Hinge-bending in L-arabinose-binding protein. The "Venus's-flytrap" model. *J. Biol. Chem.* **257**, 1131–1133.
8. Sack, J. S., Saper, M. A. & Quiocho, F. A. (1989). Periplasmic binding protein structure and function. *J. Mol. Biol.* **206**, 171–191.
9. Oh, B. H., Pandit, J., Kang, C. H., Nikaido, K., Gokcen, S., Ames, G. F. & Kim, S. H. (1993). Three-dimensional structures of the periplasmic lysine/arginine/ornithine-binding protein with and without a ligand. *J. Biol. Chem.* **268**, 11348–11355.
10. Oswald, C., Smits, S. H., Höing, M., Sohn-Bösser, L., Dupont, L., Le Rudulier, D. *et al.* (2008). Crystal structures of the choline/acetylcholine substrate-binding protein ChoX from *Sinorhizobium meliloti* in the liganded and unliganded-closed states. *J. Biol. Chem.* **283**, 32848–32859.
11. Shilton, B. H. (2008). The dynamics of the MBP-MalFGK(2) interaction: a prototype for binding protein dependent ABC-transporter systems. *Biochim. Biophys. Acta*, **1778**, 1772–1780.
12. Tang, C., Schwieters, C. D. & Clore, G. M. (2007). Open-to-closed transition in apo maltose-binding protein observed by paramagnetic NMR. *Nature*, **449**, 1078–1082.

13. Sharff, A. J., Rodseth, L. E. & Quiocho, F. A. (1993). Refined 1.8-Å structure reveals the mode of binding of β-cyclodextrin to the maltodextrin binding protein. *Biochemistry*, **32**, 10553–10559.
14. Spurlino, J. C., Lu, G. Y. & Quiocho, F. A. (1991). The 2.3-Å resolution structure of the maltose- or maltodextrin-binding protein, a primary receptor of bacterial active transport and chemotaxis. *J. Biol. Chem.* **266**, 5202–5219.
15. Millet, O., Hudson, R. P. & Kay, L. E. (2003). The energetic cost of domain reorientation in maltose-binding protein as studied by NMR and fluorescence spectroscopy. *Proc. Natl Acad. Sci. USA*, **100**, 12700–12705.
16. Marvin, J. S. & Hellinga, H. W. (2001). Manipulation of ligand binding affinity by exploitation of conformational coupling. *Nat. Struct. Biol.* **8**, 795–798.
17. Duan, X., Hall, J. A., Nikaido, H. & Quiocho, F. A. (2001). Crystal structures of the maltodextrin/maltose-binding protein complexed with reduced oligosaccharides: flexibility of tertiary structure and ligand binding. *J. Mol. Biol.* **306**, 1115–1126.
18. Klenk, H. P., Clayton, R. A., Tomb, J. F., White, O., Nelson, K. E., Ketchum, K. A. *et al.* (1997). The complete genome sequence of the hyperthermophilic, sulphate-reducing archaeon *Archaeoglobus fulgidus*. *Nature*, **390**, 364–370.
19. Schiefner, A., Holtmann, G., Diederichs, K., Welte, W. & Bremer, E. (2004). Structural basis for the binding of compatible solutes by ProX from the hyperthermophilic archaeon *Archaeoglobus fulgidus*. *J. Biol. Chem.* **279**, 48270–48281.
20. Holtmann, G. & Bremer, E. (2004). Thermoprotection of *Bacillus subtilis* by exogenously provided glycine betaine and structurally related compatible solutes: involvement of Opu transporters. *J. Bacteriol.* **186**, 1683–1693.
21. Vahedi-Faridi, A., Eckey, V., Scheffel, F., Alings, C., Landmesser, H., Schneider, E. & Saenger, W. (2008). Crystal structures and mutational analysis of the arginine-, lysine-, histidine-binding protein ArtJ from *Geobacillus stearothermophilus*. Implications for interactions of ArtJ with its cognate ATP-binding cassette transporter, Art(MP)$_2$. *J. Mol. Biol.* **375**, 448–459.
22. Kabsch, W. (1993). Automatic processing of rotation diffraction data from crystals of initially unknown symmetry and cell constants. *J. Appl. Crystallogr.* **26**, 795–800.
23. McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. & Read, R. J. (2007). Phaser crystallographic software. *J. Appl. Crystallogr.* **40**, 658–674.
24. Otwinowski, Z. & Minor, W. (1997). Processing of X-ray diffraction data collected in oscillation mode. In *Methods in Enzymology* (Carter, C. W. & Sweet, R. M., eds), *Methods in Enzymology*, 276, pp. Academic Press, London, UK.
25. Murshudov, G., Vagin, A. A. & Dodson, E. J. (1997). Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **53**, 240–255.
26. Hayward, S. & Lee, R. A. (2002). Improvements in the analysis of domain motions in proteins from conformational change: DynDom version 1.50. *J. Mol. Graphics Modell.* **21**, 181–183.
27. Hanekop, N., Höing, M., Sohn-Bösser, L., Jebbar, M., Schmitt, L. & Bremer, E. (2007). Crystal structure of the ligand-binding protein EhuB from *Sinorhizobium meliloti* reveals substrate recognition of the compatible solutes ectoine and hydroxyectoine. *J. Mol. Biol.* **374**, 1237–1250.
28. Smits, S. H., Höing, M., Lecher, J., Jebbar, M., Schmitt, L. & Bremer, E. (2008). The compatible-solute-binding protein OpuAC from *Bacillus subtilis*: ligand binding, site-directed mutagenesis, and crystallographic studies. *J. Bacteriol.* **190**, 5663–5671.
29. Schiefner, A., Breed, J., Bösser, L., Kneip, S., Gade, J., Holtmann, G. *et al.* (2004). Cation–pi interactions as determinants for binding of the compatible solutes glycine betaine and proline betaine by the periplasmic ligand-binding protein ProX from *Escherichia coli*. *J. Biol. Chem.* **279**, 5588–5596.
30. Sack, J. S., Saper, M. A. & Quiocho, F. A. (1989). Periplasmic binding protein structure and function. Refined X-ray structures of the leucine/isoleucine/valine-binding protein and its complex with leucine. *J. Mol. Biol.* **206**, 171–191.
31. Oswald, C., Smits, S. H., Hoing, M., Bremer, E. & Schmitt, L. (2009). Structural analysis of the choline-binding protein ChoX in a semi-closed and ligand-free conformation. *Biol. Chem.* **390**, 1163–1170.
32. Diez, J., Diederichs, K., Greller, G., Horlacher, R., Boos, W. & Welte, W. (2001). The crystal structure of a liganded trehalose/maltose-binding protein from the hyperthermophilic archaeon *Thermococcus litoralis* at 1.85 Å. *J. Mol. Biol.* **305**, 905–915.
33. Stockner, T., Vogel, H. J. & Tieleman, D. P. (2005). A salt-bridge motif involved in ligand binding and large-scale domain motions of the maltose-binding protein. *Biophys. J.* **89**, 3362–3371.
34. Muller, A., Severi, E., Mulligan, C., Watts, A. G., Kelly, D. J., Wilson, K. S. *et al.* (2006). Conservation of structure and mechanism in primary and secondary transporters exemplified by SiaP, a sialic acid binding virulence factor from *Haemophilus influenzae*. *J. Biol. Chem.* **281**, 22212–22222.
35. Johnston, J. W., Coussens, N. P., Allen, S., Houtman, J. C., Turner, K. H., Zaleski, A. *et al.* (2008). Characterization of the *N*-acetyl-5-neuraminic acid-binding site of the extracytoplasmic solute receptor (SiaP) of nontypeable *Haemophilus influenzae* strain 2019. *J. Biol. Chem.* **283**, 855–865.
36. Richarme, G. & Kepes, A. (1983). Study of binding protein–ligand interaction by ammonium sulfate-assisted adsorption on cellulose esters filters. *Biochim. Biophys. Acta*, **742**, 16–24.
37. Emsley, P. & Cowtan, K. (2004). Coot: model-building tools for molecular graphics. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **60**, 2126–2132.
38. Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr.* **26**, 283–291.
39. Case, D. A., Cheatham, T. E., 3rd, Darden, T., Gohlke, H., Luo, R., Merz, K. M., Jr *et al.* (2005). The Amber biomolecular simulation programs. *J. Comput. Chem.* **26**, 1668–1688.
40. Cornell, W. D., Cieplak, P., Bayly, C. I., Gould, I. R., Merz, K. M., Jr, Ferguson, D. M. *et al.* (1995). A second

generation force field for the simulation of proteins, nucleic acids, and organic molecules. *J. Am. Chem. Soc.* **117**, 5179–5197.
41. Simmerling, C., Strockbine, B. & Roitberg, A. E. (2002). All-atom structure prediction and folding simulations of a stable protein. *J. Am. Chem. Soc.* **124**, 11258–11259.
42. Jorgensen, W. L., Chandrasekhar, J., Madura, J. & Klein, M. L. (1983). Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **79**, 926–935.
43. Darden, T., York, D. & Pedersen, L. (1993). Particle mesh Ewald: An $N\cdot\log(N)$ method for Ewald sums in large systems. *J. Chem. Phys.* **98**, 10089–10092.
44. Ryckaert, J. P., Ciccotti, G. & Berendsen, H. J. C. (1977). Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of *n*-alkanes. *J. Comput. Phys.* **23**, 327–341.
45. Bayly, C. I., Cieplak, P., Cornell, W. D. & Kollman, P. A. (1993). A well-behaved electrostatic potential based method using charge restraints for determining atom-centered charges: the RESP model. *J. Phys. Chem.* **97**, 10269–10280.
46. Wang, J., Wolf, R. M., Caldwell, J. W., Kollman, P. A. & Case, D. A. (2004). Development and testing of a general amber force field. *J. Comput. Chem.* **25**, 1157–1174.