

Conformational flexibility of the leucine binding protein examined by protein domain coarse-grained molecular dynamics

Iwona Siuda · Lea Thøgersen

Received: 5 July 2013 / Accepted: 29 August 2013 / Published online: 19 September 2013
© Springer-Verlag Berlin Heidelberg 2013

Abstract Periplasmic binding proteins are the initial receptors for the transport of various substrates over the inner membrane of gram-negative bacteria. The binding proteins are composed of two domains, and the substrate is entrapped between these domains. For several of the binding proteins it has been established that a closed-up conformation exists even without substrate present, suggesting a highly flexible apo-structure which would compete with the ligand-bound protein for the transporter interaction. For the leucine binding protein (LBP), structures of both open and closed conformations are known, but no closed-up structure without substrate has been reported. Here we present molecular dynamics simulations exploring the conformational flexibility of LBP. Coarse grained models based on the MARTINI force field are used to access the microsecond timescale. We show that a standard MARTINI model cannot maintain the structural stability of the protein whereas the ELNEDIN extension to MARTINI enables simulations showing a stable protein structure and nanosecond dynamics comparable to atomistic simulations, but does not allow the simulation of conformational flexibility. A modification to the MARTINI-ELNEDIN setup, referred to as domELNEDIN, is therefore presented. The domELNEDIN setup allows the protein domains to move independently and thus allows for the simulation of conformational changes. Microsecond domELNEDIN simulations starting from either

the open or the closed conformations consistently show that also for LBP, the apo-structure is flexible and can exist in a closed form.

Keywords Coarse-graining · domELNEDIN · Elastic network · ELNEDIN · LBP · MARTINI

Introduction

The leucine binding protein (LBP) belongs to the superfamily of binding proteins [1]. More than 100 crystal structures of periplasmic binding proteins (PBP) from different sources and in different conformations have been solved, showing a conserved structural fold, despite high sequence diversity [2]. The protein fold consists of two domains connected by a three-, two- or one-stranded hinge (group I, II and III, respectively) [1]. The movement of the domains is a classical example of hinge motion [3], and the mechanism of ligand entrapment has been referred to as a Venus flytrap model [4]. LBP is a group I PBP, and structures have been established for both open and closed conformations [5, 6], with the substrate leucine bound in the cleft between the domains in the closed conformation. In gram-negative bacteria, binding proteins found in the periplasm, act as primary receptors in adenosine triphosphate (ATP)-binding cassette (ABC) transport systems, trafficking various substrates across the plasma membrane at the cost of ATP hydrolysis [7]. Despite the great variety in substrates ranging from nutrients such as sugars and amino acids in prokaryotes to poly-saccharides, lipids and hormones in eukaryotes [8], the ABC transport systems seem to share overall mechanism (for recent reviews see ref [8–10]). For the ABC-importers present in gram-negative bacteria such as *E. coli*, a binding protein entraps the substrate in a closed conformation [11], and docks in this form to the membrane transporter, which is interacting closely with the ATP binding domains

Electronic supplementary material The online version of this article (doi:10.1007/s00894-013-1991-9) contains supplementary material, which is available to authorized users.

I. Siuda · L. Thøgersen
Centre for Membrane Pumps in Cells and Disease, Danish National Research Foundation, Bioinformatics Research Centre, Aarhus University, C. F. Møllers Alle 8, DK-8000 Aarhus C, Denmark

L. Thøgersen (✉)
CLC bio, Silkeborgvej 2, DK-8000 Aarhus C, Denmark
e-mail: lea@birc.au.dk

on the cytoplasmic side. The ATP-binding and hydrolysis then help fuel the conformational shifts necessary to open the binding protein, release the substrate to the membrane transporter, change the membrane transporter to an occluded state, and finally release the substrate to the cytoplasm [8].

The role of the substrate in the mechanism for the different ABC transporters, as well as the complex interplay between binding protein, conformational changes, and ATP interaction have still not been settled on [8–10]. Some of the unsettled issues regarding the substrate role are addressed in the present study of LBP.

Molecular dynamics (MD) simulations can provide valuable information to better understand the conformational changes of the PBPs. Previous MD studies of binding proteins have shown that conformational changes between open and closed states can be observed in unbiased simulations on the 10–50 ns timescale [12–18]. However, sampling of changes will in many cases be out of reach for unbiased MD simulations where all atoms (AA) of the biomolecular system are represented, and different biased MD approaches [19–22] and network models [23, 24] have also been applied to study the domain motions of binding proteins. To our knowledge, only one MD simulation of LBP has been published [25]. This 100 ns simulation of the open LBP apo-structure showed great flexibility in the structure, but no overall conformational shift [25]. To access longer timescales with MD simulations, coarse grained (CG) models can be applied. In CG models atoms are grouped in “beads” on the amino acid level for proteins [26–31], allowing for time steps in the 10s of femtoseconds and unbiased simulations on the microsecond time scale. The MARTINI CG force field (FF) represents such a model [29–31], where the amino acid representations have been fitted to match the experimentally observed partitioning between hydrophilic and hydrophobic environments [31]. The model has gained great popularity in recent years and lipids, water, carbohydrates as well as amino acids have been parameterized for the FF [30–32]. However, the MARTINI CG protein model contains a bias toward the initial structure. The protein secondary structure is stabilized using dihedral restraints on the backbone [30, 31]. This is sufficient for stabilizing the structure of peptides while for larger proteins extra restraints on the structure need to be introduced. One approach for doing this is the ELNEDIN model [33], where an elastic network (EN) is applied to the whole protein with bonds connecting all CG backbone beads which are within a specific distance in the initial structure.

In the present study we examine the conformational changes of LBP, using a modification to the MARTINI-ELNEDIN [33] CG method which is applicable to multi-domain proteins. The approach, referred to as domELNEDIN, keeps the residue-level coarse-graining while at the same time allowing the protein to change conformation in an unbiased manner during the MD simulation as ENs are not applied

globally, but only locally in the protein domains. We show that this is a sufficient degree of stabilization to avoid a collapse of the structure and keep it stable on the nanosecond timescale.

The progress of the paper is as follows. Atomistic simulations of 100 ns are carried out for both the open and closed structures, showing no overall conformational change. The ability to produce a stable structural scaffold with nanosecond dynamics comparable to the atomistic simulation is then established for the ELNEDIN and domELNEDIN models of LBP, while the standard MARTINI is seen to fail as expected. Following this, domELNEDIN simulations are carried out starting from both the open and closed conformations, to study the conformational flexibility of LBP in water on the microsecond timescale. To illustrate the differences between the ELNEDIN and domELNEDIN models on the long time scale, results from the corresponding ELNEDIN simulations are also reported, and to inspect the sensitivity of the domELNEDIN model toward the assigned protein domain boundaries, four different domain setups are tested. Finally, the model limitations as well as the biological implications of the results are discussed.

Methods

MD simulations of LBP were carried out starting from a closed (pdb 1USK [6]) or open (pdb 1USG [6]) conformation of the protein. For the closed conformation two different setups were applied; either with or without the leucine ligand present in the binding cleft. In all setups counter ions were added (9 Na⁺) and the protein was solvated with water in a cubic box with dimensions of 100 Å. The GROMACS package version 4.0.7 [34, 35] was used for all simulations and the pressure and temperature were kept constant at 1 bar and 300 K, respectively, using the Berendsen coupling algorithm [36]. For MARTINI CG simulations it has been shown that the simulated time typically should be multiplied by a factor of 4 to roughly account for the increase in diffusion observed for CG water beads [29, 30]. Throughout this paper, the simulated time multiplied by 4 will therefore be referred to as “effective” time for the CG simulations, and this is the time used in all figures.

Atomistic simulations The simulations were performed with the AMBER03 FF [37] for the protein and the SPC water model [38] for the solvent. Partial charges for the leucine ligand were derived from Antechamber [39, 40] (see supplementary material Table S1). The temperature and isotropic pressure were kept constant with time constants $\tau_T=0.1$ ps and $\tau_P=1$ ps. PME was used for the long-range electrostatics interactions and a cut-off of 10 Å was used for the short-range electrostatics contributions, while the van der Waals interactions

were cut off at 14 Å. Bond lengths were constrained using the LINCS algorithm [41] for the protein. The setups were energy-minimized followed by a relaxation of the solvent and ions, with position restraints ($1000 \text{ kJ}\cdot\text{mol}^{-1}\cdot\text{nm}^{-2}$) applied to all heavy atoms of the protein for 20 ps. Simulations starting from either open or closed conformations without the leucine ligand present, as well as the closed conformation including the ligand were carried out for 100 ns without any restraints.

MARTINI CG simulations The simulations were performed with version 2.1 of the MARTINI CG FF [31]. The coarse grained representation of LBP was generated using MARTINI scripts, topologies and parameters [31]. Standard MARTINI CG water beads were used to model the solvent [29]. The leucine ligand was represented by one backbone bead of type P5 and a side chain bead of type C1. The setups were energy-minimized and the solvent and ions were relaxed with position restraints ($1000 \text{ kJ}\cdot\text{mol}^{-1}\cdot\text{nm}^{-2}$) applied to all backbone beads of the protein for 1 ns. For the temperature and pressure settings, the time constants $\tau_T=1 \text{ ps}$ and $\tau_P=5 \text{ ps}$ were applied. Non-bonded interactions were cut-off at 1.2 nm and shifted from 0.9 nm for the Lennard-Jones potential and from 0.0 nm for the electrostatic potential. Neighbor lists were updated every 10 steps. Setups with the closed conformation including the ligand and the open conformation without the ligand were simulated for 25 ns (100 ns “effective” time) using a 25 fs time step.

ELNEDIN simulations The applied version of ELNEDIN [33] is based on modifications to version 2.1 of the MARTINI CG FF [31]. In ELNEDIN simulations, an EN is put on the backbone beads of a slightly modified MARTINI CG model of the protein, to maintain the initial tertiary structure [31]. The interaction between first and second neighbor backbone beads is defined by bond and angle parameters. All other pairs of backbone beads, for which the distance in the input structure is below some cut-off R_C , are assigned a harmonic network bond with the force constant K_S [33]. Thus, the two input parameters, R_C and K_S , define the network. For two different setups of LBP (the open form without leucine and the closed form with leucine present in the binding cleft) nine different ENs were tested in 25 ns simulations (100 ns “effective” time), varying the cut-off distance R_C (Å) $\in \{8, 9, 10\}$ and spring force constant K_S ($\text{kJ}\cdot\text{mol}^{-1}\cdot\text{nm}^{-2}$) $\in \{50, 500, 5,000\}$. For the EN providing the best overlap with the atomistic simulations, $\{R_C, K_S\} = \{8 \text{ Å}, 500 \text{ kJ}\cdot\text{mol}^{-1}\cdot\text{nm}^{-2}\}$, 1 μs simulations (4 μs “effective” time) were carried out for setups starting from either the closed or open conformations of LBP without ligand present.

All ELNEDIN setups were solvated as in the MARTINI CG simulations. Setups were energy-minimized with position restraints ($1000 \text{ kJ}\cdot\text{mol}^{-1}\cdot\text{nm}^{-2}$) and the solvent and ions

were relaxed while all protein beads were restrained with the same force constant for 50 ps using 1 fs time step followed by a 1 ns equilibration using 10 fs time step with restraints put only on the backbone beads of the protein. For the temperature and pressure the settings $\tau_T=0.5 \text{ ps}$ and $\tau_P=1.2 \text{ ps}$ were applied. Neighbor lists were updated every five steps. The non-bonded interactions were treated with the same shifts and cut-offs as applied for the MARTINI CG simulations. All simulations were run using a 10 fs time step.

domELNEDIN simulations There is no unique way to assign each residue in a protein to one or the other structural domain. Four different domain definitions for LBP were therefore tested in this study. Based on visual inspection of both the open and closed conformations it was decided to define domain 1 as residues 1–120 and 250–330 and domain 2 as residues 121–249 and 331–345. This is referred to as the “main” domain assignment. Apart from this, a “main loose” definition was applied, where four amino acids in each of the three domain linkers were released from the ENs. Automatic domain boundary assignments from the pDomains server [42] and the DomFOLDpdp server [43] were also tested, as described in the Results section and listed in Table 1.

Preparation of the setups as well as the production run parameters for the domELNEDIN simulations were the same as for the ELNEDIN model, except the ENs were applied within the protein domains only, as described in Results. For the open form of LBP, and closed form with leucine positioned in the binding cleft, 25 ns simulations (100 ns “effective” time) were carried out with the ENs $\{R_C, K_S\} = \{8 \text{ Å}, 500 \text{ kJ}\cdot\text{mol}^{-1}\cdot\text{nm}^{-2}\}$ and $\{9 \text{ Å}, 500 \text{ kJ}\cdot\text{mol}^{-1}\cdot\text{nm}^{-2}\}$. For the EN providing the best overlap with the atomistic simulations, $\{R_C, K_S\} = \{8 \text{ Å}, 500 \text{ kJ}\cdot\text{mol}^{-1}\cdot\text{nm}^{-2}\}$, 1 μs simulations (4 μs “effective” time) were carried out for setups starting from either the closed or open conformations of LBP, without the leucine ligand present.

Scripts and example input and output files for converting an ELNEDIN setup to a domELNEDIN setup can be found on <http://www.birc.au.dk/~leat/domELNEDIN>, and the protocol flow can be seen in supplementary material Fig. S1.

Analysis The conformational stability of LBP structures is examined by evaluating the root mean-square deviations

Table 1 Protein domain boundary assignments

	Domain 1	Domain 2
Main	1–120+250–330	121–249+331–345
Main loose	1–118+252–328	123–247+333–345
pDomains [42]	1–124+247–332	125–246+333–345
DomFOLDpdp [43]	1–119+249–332	120–248+333–345

(RMSDs) of the protein backbone based on C α atoms (for the AA simulations) or the backbone beads (for the CG simulations). The RMSD is compared to the structure in the first frame of the simulation, unless stated otherwise. The RMSD and root mean-square fluctuations (RMSF) of individual residues were examined for the last 80 ns of simulation for the AA simulations and last 80 ns of “effective” time for the CG simulations, for the protein with the backbone aligned as described above. To compare the large-amplitude fluctuations in the AA and CG (ELNEDIN or domELNEDIN) simulations, the covariance matrix of the positional fluctuations was constructed for the coordinates of C α atoms obtained from the last 80 ns of AA simulations or the backbone beads obtained from the last 80 ns of “effective” time of the CG simulations. Trajectories were fitted to the same reference structure, i.e., the X-ray structure of the closed conformation of LBP was used as a reference structure for the simulations starting from the closed conformation of LBP, and the X-ray structure of the open conformation of LBP was used as a reference structure for the simulations starting from the open conformation of LBP. The eigenvectors found when diagonalizing the covariance matrix were then used in the root mean-square inner-product (RMSIP) analysis, that quantifies the overlap between the essential subspaces (described by the 10 first eigenvectors) obtained from the AA and CG simulations, Eq. (1) [44–46]:

$$\text{RMSIP} = \sqrt{\frac{1}{10} \sum_{i=1}^{10} \sum_{j=1}^{10} (\eta_i^{\text{AA}} \cdot \eta_j^{\text{CG}})^2} \quad (1)$$

where η_i^{AA} η_j^{CG} are the i th and j th eigenvectors from the AA and various CG (ELNEDIN or domELNEDIN) simulations.

Results

Structure and dynamics on the nanosecond timescale

Atomistic simulations As seen in Fig. 1, the AA MD simulation starting from the open conformation shows more flexibility than the ones starting from the closed conformation with ligand present. However, both simulations clearly produce a structural ensemble around their starting conformations, and show no sign of conformational change. A 100 ns simulation starting from the closed conformation without ligand present was also carried out (closed AA w/o lig.), and it can be seen that also the ligand-free structure is highly stable (RMSD of 1.5 ± 0.1 Å over the last 50 ns), and even more so than when the ligand is present (RMSD of 1.9 ± 0.2 Å over the last 50 ns).

Standard MARTINI CG In the MARTINI CG model for proteins, each residue is mapped to a backbone bead and zero

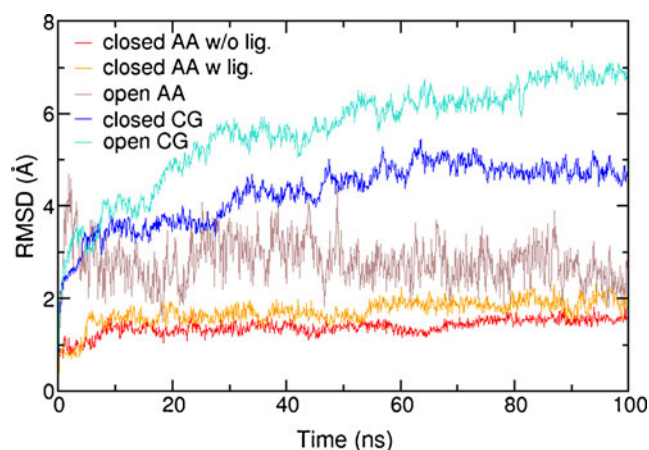


Fig. 1 RMSDs for the protein backbone in the AA and standard MARTINI CG simulations starting from the closed and open conformations of LBP. For the closed conformation, data from atomistic simulations both with (closed AA w lig.) and without (closed AA w/o lig.) the leucine ligand present are shown

to four side chain beads, and even though detailed interactions are lost in the coarse graining, the resolution is high enough to represent the particular physicochemical properties of the different amino acids. As directional hydrogen bonds are not possible to represent in the CG model, secondary structure elements are not self-contained, and the local structure is therefore predefined by restraining backbone angles and dihedrals to values supporting helix or extended structures, based on structural analysis of the initial atomic resolution structure [31]. For purely α -helical structures or membrane proteins completely surrounded by lipid membrane, MARTINI CG simulations have been reported where no further restraints on the structure have been imposed [47–50]. However, as β -sheets are defined by hydrogen bonds between sequence distant strands, this form of local structure is poorly described by the model. Moreover, as water does not stabilize protein structure in the same manner as a lipid bilayer, the standard MARTINI CG model does generally not succeed in maintaining the tertiary structure of globular proteins. The MARTINI CG simulations of LBP in water are no exception. For both the simulation starting from the open and the one starting from the closed conformation of the protein, the tertiary structure is observed to collapse into a packed globular structure in an unspecific manner (see Fig. 1 and 2), which does not resemble the known structure of the protein in a closed conformation. The collapse of tertiary structure resulting from MARTINI CG simulations has been the main motivation for the previously proposed ELNEDIN [33] extension to the MARTINI FF as well as the modification, domELNEDIN, presented in this work. The inability of the standard MARTINI CG model to keep the structure stable for even 100 ns makes this model irrelevant for further studies of LBP, and instead the focus will be on the above mentioned extensions to the MARTINI CG model, where the tertiary

structure of the protein is stabilized by the application of an EN to the backbone beads.

The ELNEDIN extension The ELNEDIN extension to MARTINI CG was intended to stabilize the overall protein structure, while at the same time allowing for structural fluctuations on the nanosecond timescale, comparable to that observed for AA simulations [33]. This was achieved through the application of an EN to the backbone beads of a MARTINI CG model. Minor modifications to the standard MARTINI CG model were introduced with the ELNEDIN extension, as the backbone beads now were placed in the $C\alpha$ positions instead of at the center-of-mass of the backbone atoms. Two parameters were used to define the network and tune the dynamics, namely the cut-off distance, R_C , for applying an EN bond between two beads, and the force constant, K_S , of the harmonic EN bonds forming the network.

For the purpose of establishing the optimal set of parameters to describe the dynamics of LBP using the ELNEDIN model, simulations were carried out with all possible combinations of three different force constants and three different cut-offs, as described in the Methods section. The choice of the scaffold parameters to use for further simulations is made based on a comparison with the protein dynamics observed in an atomistic simulation. Just as in the original ELNEDIN study, the RMSD as a function of time and the RMSD per residue are used to quantify the global and local structural deformations, while the RMSF per residue and the essential subspace overlap (quantified by the RMSIP) are used to describe the local and large-amplitude fluctuations, respectively. As expected, the RMSD and RMSF values decrease in a systematic manner with the increase of the cut-off and

force constant values (see supplementary material Figs. S2, S3, S4, S5, S6 and S7). The RMSIP cannot be expected to show such a systematic behavior, but in this case the overlap consistently increases when increasing the cut-off and force constant (see supplementary material Tables S2 and S3). The RMSIP values are in all cases found to be between 0.50 and 0.75, and the overlap is thus considered highly satisfactory for all the tested parameter sets. From Figs. S2, S3, S4, S5, S6 and S7 it is seen that the parameters giving the best overlap with the atomistic data is $K_S=500 \text{ kJ}\cdot\text{mol}^{-1}\cdot\text{nm}^{-2}$, suggested as default value for the ELNEDIN model [33], and a cut-off of $R_C=8 \text{ \AA}$. The cut-off suggested as default for the ELNEDIN model is $R_C=9 \text{ \AA}$, and the effect of changing the cut-off by 1 \AA is also seen in Fig. 3 for both the open and closed conformation.

As previously documented, the ELNEDIN extension to the standard MARTINI CG model succeeds in reproducing a structural scaffold similar to what is seen in atomistic simulations of 100 ns [33]. For the case of LBP this is achieved building the EN using the parameters $R_C=8 \text{ \AA}$ and $K_S=500 \text{ kJ}\cdot\text{mol}^{-1}\cdot\text{nm}^{-2}$. It should be recognized as a quality of the model that, despite quite different overall levels of dynamics, the ELNEDIN setup giving the best fit to the atomistic simulations is the same for both the closed and open conformations. A major drawback, though, is that the structural scaffold is limited to a single protein conformation, not allowing studies that involve rearrangements of the protein domains. To circumvent this bias toward the initial conformation, we have examined the effect on the protein structure and dynamics when removing the EN bonds between protein domains.

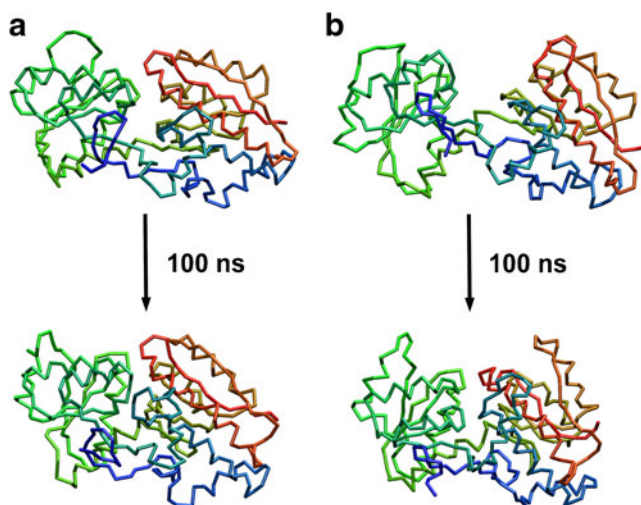


Fig. 2 Snapshots of the (a) closed and (b) open conformations of LBP at the first frame of the standard MARTINI CG simulation, with corresponding structures after 25 ns of CG simulation (100 ns of “effective” time). The protein is colored by residue number going from the N-terminus in red to the C-terminus in blue

The domELNEDIN extension There is no unique definition of how to divide a protein into domains, and various criteria have been used such as structure, function, folding units, sequence or evolution [51]. In the growing number of known protein structures, domains containing highly conserved combinations of secondary structure elements repeated in different proteins or within the same protein are observed [52]. The structural difference between different protein conformations is seen mainly to arise from hinge, shear, or rotational motions between these structural domains [3]. In the case of LBP, where the structure of both an open and closed conformation is known, it is clearly a two-domain protein. The RMSD between the crystal structures of the open and closed conformation of LBP used for this work is 7.0 \AA , while the RMSD between the same domains of these two conformations are 0.7 \AA for domain 1 and 0.6 \AA for domain 2, when using the main domain assignment (Table 1). The structural rearrangements between the conformations are thus mainly caused by movements in the hinge region of the protein. The individual LBP domains could therefore be restrained with EN bonds as in an ELNEDIN model, while still allowing for a

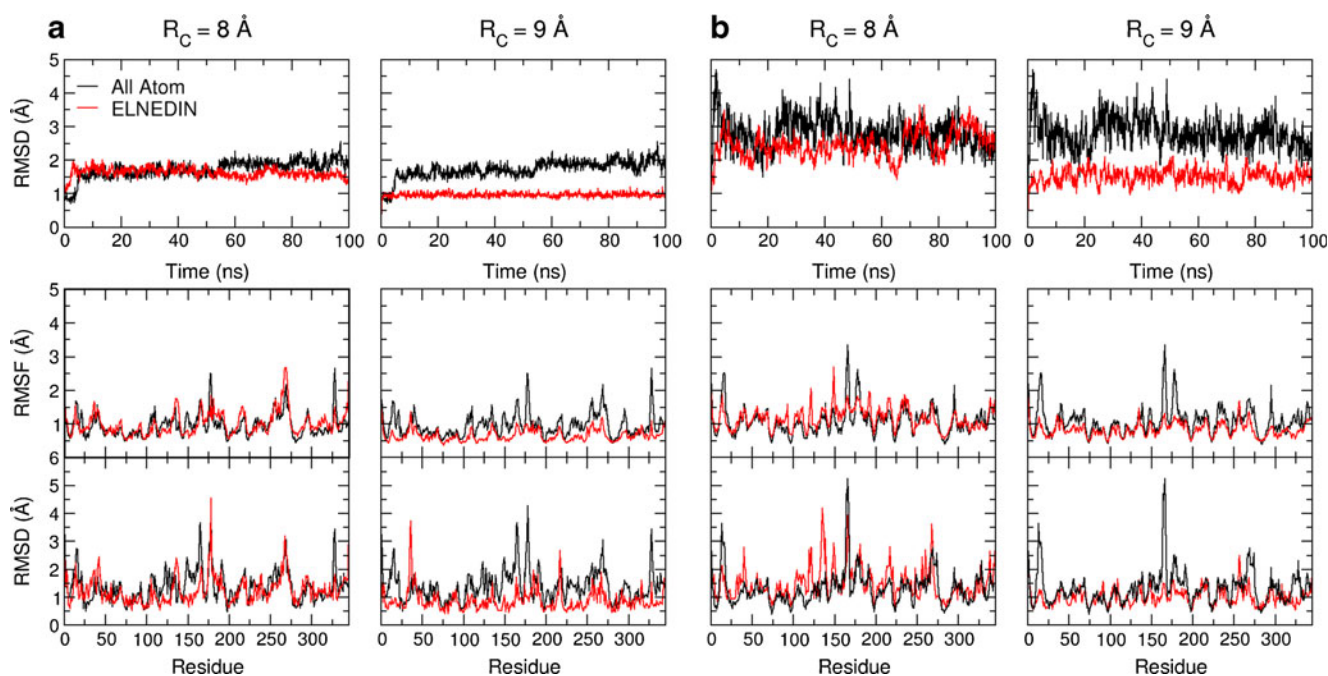


Fig. 3 RMSD as a function of simulation time and RMSF and RMSD per residue for the last 80 ns of simulation obtained from ELNEDIN simulations of the closed conformation of LBP with ligand (**a**) and the

open conformation without ligand (**b**). The force constant parameter, K_S , of the EN is in all cases $500 \text{ kJ} \cdot \text{mol}^{-1} \cdot \text{nm}^{-2}$ and the cut-off, R_C , is either 8 Å or 9 Å as indicated on the figure

full description of the structural ensemble covering the whole functional cycle. This is the concept of the domELNEDIN model; an EN is applied within each domain with the same type of parameter setup as for ELNEDIN, the difference being that no EN bonds are applied connecting the protein domains. For LBP, the differences between ELNEDIN and domELNEDIN are seen in Fig. 4 for both the open and the closed conformations. In order for the domELNEDIN model to be successful, the non-bonded interactions between protein domains should be well described by the MARTINI CG FF. MARTINI has previously successfully been used to study macro-assemblies of proteins [33] and the aggregation of proteins [33, 47, 48, 53, 54], and recently a thorough study of side chain dimerization showed that the MARTINI FF produced dimerization free energies which in aqueous solution correlated reasonably with those obtained with the atomic resolution FFs OPLS [55, 56] and GROMOS [57, 58]. It would therefore seem reasonable to expect that the correct packing and interaction of protein domains can be described within the FF as well, as domain packing holds many similarities to protein-protein interaction.

Just as for the ELNEDIN model, simulations were carried out to establish the most suitable set of EN parameters for a domELNEDIN simulation of LBP. If the annihilation of the inter-domain EN bonds does not completely destabilize the structure, it is to be expected that the dynamics on the 100 ns time scale is similar for the ELNEDIN and domELNEDIN models, and therefore only the parameters

$K_S=500 \text{ kJ} \cdot \text{mol}^{-1} \cdot \text{nm}^{-2}$ and $R_C=8 \text{ Å}$ and 9 Å were inspected. As the protein domains are allowed to move independently, it is not expected that the dynamics will be dampened in a systematic manner going toward a higher cut-off and force constant. Still, the overall best parameter fit is also for domELNEDIN achieved when applying the cut-off $R_C=8 \text{ Å}$ (see supplementary material Fig. S8 and Table S4). In Fig. 5, the RMSD and RMSF data are compared for the ELNEDIN and domELNEDIN simulations, and in Table S4 in the SI, the RMSIP values are compared. It is clear that even though the overall RMSD indicates more changes in the protein structure for the domELNEDIN simulations, the RMSD and RMSF per residue are very similar between the two models, and as well are the RMSIP values. The ELNEDIN and domELNEDIN extensions to the MARTINI CG FF thus show the same level of dynamics and structural stability on the 100 ns time scale while using the same EN parameter set. This is achieved for the domELNEDIN model despite the removal of the EN bonds stabilizing the domain interfaces, and thus conformational changes are allowed in an unbiased manner within this model. The difference between the ELNEDIN and domELNEDIN models should therefore be noticeable and important on the microsecond time scale.

Long time scale events

For a number of the binding proteins it has been established that the apo-structure is flexible and occupies a wide range of

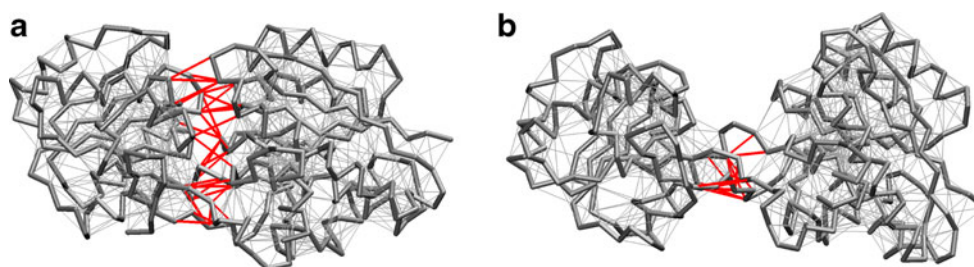


Fig. 4 Comparing the ELNEDIN and domELNEDIN models for ENs generated with $R_C=8$ Å, and using the main domain assignment. **a** Closed conformation. **b** Open conformation. Bonds are colored as

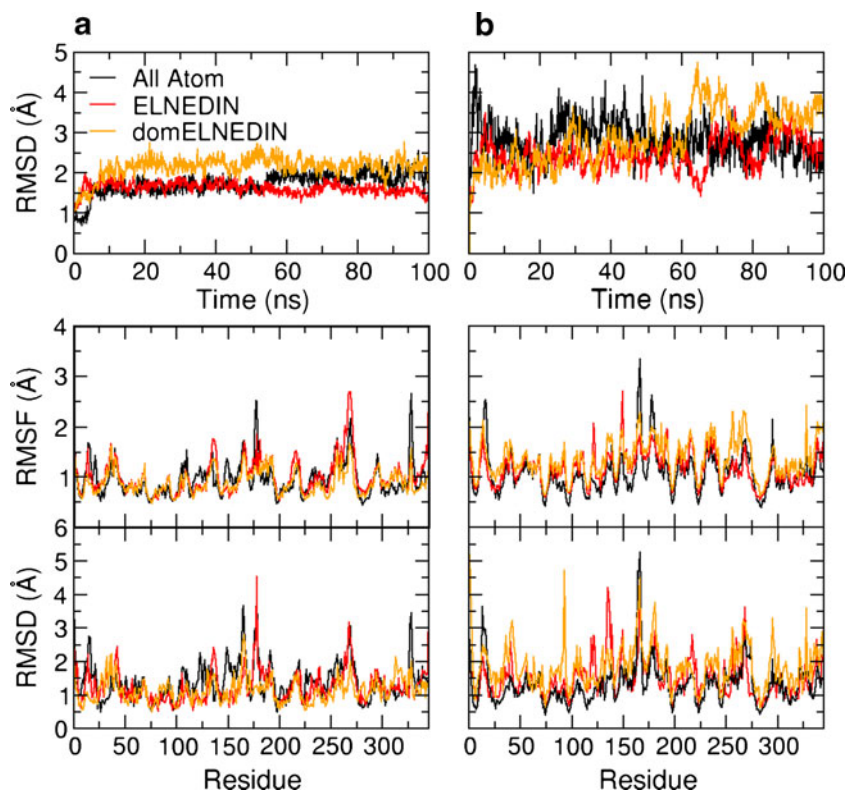
follow: gray – bonds present in both the ELNEDIN and domELNEDIN models, red – bonds present in the ELNEDIN model, but removed in the domELNEDIN setup

conformations ranging from a full opened to a closed conformation [15, 59–64] and the open-to-closed transition is expected to take place on the nanosecond to microsecond timescale [59]. A closed conformation without substrate has not yet been observed for LBP, and to study the conformational flexibility of LBP when substrate is not present, microsecond long domELNEDIN simulations have been carried out, starting from both the open and closed conformations. For both cases, the domELNEDIN simulations are compared to the corresponding ELNEDIN simulations. Furthermore, the sensitivity toward the domain boundary definitions and toward the choice of initial topology to use for the domELNEDIN setups are tested.

Starting from the open conformation In Fig. 6, the development of the protein backbone RMSDs over the full simulations are depicted compared to the crystal structures of both

the open and the closed conformations. The ELNEDIN model applied to the open conformation is expected to result in simulations producing a structural ensemble around the crystal structure of the open conformation. However, a change toward the closed conformation is observed on the microsecond time scale, going from an RMSD of 7.0 Å to 5–6 Å with respect to the crystal structure of the closed conformation. As is clear from Fig. 4b, there are only a few EN bonds connecting the domains in the open conformation (red lines), and they are all positioned in the hinge region. This is why an overall structural change is allowed without conflicting too much with the globally applied EN in the ELNEDIN setup. However, the conformational change has a limit, as further change toward the closed conformation would be energetically highly unfavorable, due to the necessary change in the EN bonds bridging the domains.

Fig. 5 Backbone RMSD as well as RMSF and RMSD per residue in the ELNEDIN and domELNEDIN simulations of closed form with ligand (**a**) and open form without ligand (**b**) of LBP. The EN scaffolds were parameterized with $R_C=8$ Å and $K_S=500$ kJ·mol⁻¹·nm⁻²



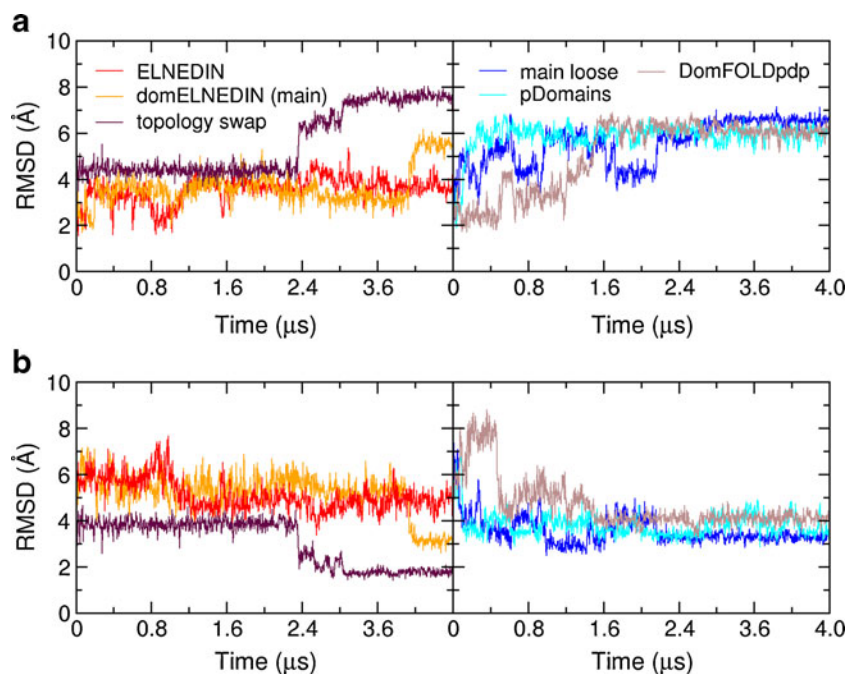


Fig. 6 Backbone RMSDs in ELNEDIN and domELNEDIN simulations starting from the open conformation of LBP. **a** Compared to the crystal structure of the open conformation. **b** Compared to the crystal structure of the closed conformation. Left-most graphs show results from the ELNEDIN simulation, the domELNEDIN simulation using the main domain assignment and the “topology swap” setup, where the topology

for the closed conformation is applied to the domELNEDIN simulation starting from the open conformation. The right-most graphs show results from domELNEDIN simulations using three alternative domain assignments in the setup. The protein domain boundaries corresponding to main, main loose, pDomains and DomFOLDpdp can be seen in Table 1

In the domELNEDIN simulations the open conformation is free to close up, and for all setups the structure clearly approaches the closed conformation, and to a significant higher degree than seen for the ELNEDIN simulation, resulting in RMSDs as low as 3 Å with respect to the crystal structure of the closed conformation, starting from an RMSD of 7 Å.

Even though the protein domains are structurally very similar between the conformations, there are a few differences in their topology parameters since equilibrium angle values for the backbone as well as the EN bonds are assigned based on the exact atom positions in the initial structures. Out of the ~1080 bonds forming the EN for the domELNEDIN simulations ~50 are unique for the closed or open structures (around 10 of these involve residues in the domain linkers) while the rest are shared between the setups for the two conformations. A setup was made named “topology swap”, where the topology input for the closed conformation was applied to the open conformation, enforcing the EN bonds and equilibrium backbone angles of the closed conformation. For this setup, a simulation starting from the open conformation is seen to reach a closed structure which has an RMSD as low as 1.6 Å compared to the crystal structure of the closed form. The application of local intra-domain structural parameters from the known closed structure thus allowed the domains to adapt to each other in the process of closing (induced fit), changes that could not be accommodated using the topology of the open conformation.

The division of a protein into structural domains will often be done in the most sensible way by a trained eye [65], and the well-established fold databases CATH [66] and SCOP [67] rely on human expertise. However, with the pace at which new protein structures are submitted to the protein data bank (PDB [68, 69]), the human expert assignments lag behind, and the development of methods for high-quality automatic assignment of domain boundaries is an active field of research [65, 70–72]. To inspect how the outcome of the domELNEDIN simulations of LBP depends on the domain boundary assignment, four different assignments have been tested (Table 1). The main domain assignment was based on our judgment from visual inspection of both the closed and open conformations, and it is identical to the assignment by CATH. The “main loose” assignment has the same boundaries as main, except four amino acids in each of the three domain linkers are released from the ENs. This setup was chosen to test if a more flexible linker region would alter the outcome. Two alternative protein domain boundaries were also acquired from two public available servers for automatic protein domain assignment. One is the DomFOLDpdp server [43], which based on the amino acid sequence predicts the protein fold using the nFOLD3 method [73, 74], and then use the Protein Domain Parser (PDP) program [75] to establish the number of domains and their boundaries. The other is the pDomains server [42], which for a protein PDB Id can provide

an overview of domain assignment results from seven different methods, as well as provide a consensus assignment based on these results. In the consensus assignment, the different methods contribute with a weight which is based on benchmarked knowledge of each method's performance for that particular type and size of protein [65, 70, 76]. For the closed conformation two consensus assignments were presented, the first consensus assigned it as a one-domain protein while the second consensus was a two-domain assignment. The two-domain consensus assignments for the open and closed structures used in this study were identical, except residue 333 was assigned to domain 1 in the closed conformation and domain 2 in the open conformation. As seen in Table 1, we used the consensus assignment as given to the open conformation.

The four simulations using different domain assignments show the same overall result; in all cases the LBP starts from an open conformation and ends up in a closed conformation after 4 μ s of simulation. However, as could be expected, the highest degree of flexibility is observed for the “main loose” setup. For this setup a number of structural changes between open and closed-like conformations are seen on the 4 μ s time scale, while the three other setups only show a single change toward the closed conformation in the same time frame. All four simulations differ in the time it takes before the protein makes the change to the closed conformation. For the simulation using the pDomains assignment, the conformational change is observed within the first 0.4 μ s, while it is observed after 3.8 μ s of simulation when the main assignment is applied. However, multiple repeat simulations of the setups are required, to conclude whether the observed differences are inherent to the domain assignments, or simply a result of the stochastic nature of the event. Furthermore, using the domELNEDIN model, the dynamics of the domain movements will very likely be dependent on how tight the linkers between the domains are bound to the domains themselves.

Starting from the closed conformation In all simulations of the closed conformation LBP stays more or less closed as seen in Fig. 7, even though the ligand is not present. In the ELNEDIN simulation it is under no circumstances expected to observe any opening of the protein, as several EN bonds are applied between the domains, keeping them tightly together (Fig. 4a). Also for the domELNEDIN simulations, the closed conformation is very stable and no opening is observed. Thus, even though the protein is free to change its conformation in a long time scale simulation, it stayed in the stable closed conformation.

Similar to the “topology swap” setup for the open conformation, a setup was made where the topology input for the open conformation was applied to the closed conformation, enforcing the EN bonds and equilibrium backbone angles of the open conformation onto the closed structure. In this

simulation, the protein clearly remained closed, while remodeling the structure locally, to better fit with the EN bonds and equilibrium angles specific for the open conformation, thus resulting in an elevated RMSD to the closed structure while keeping an even higher RMSD to the open structure. This observation corroborates that the application of the topology from a different conformation does not force a conformational change in itself, but merely alters the internal domain structure.

The different domain assignments, applied as described in the previous sub-section, do also in this case not alter the outcome of the simulations. Even the release of the linkers from the ENs of the protein domains in the “main loose” setup does not affect the overall structural changes within the 4 μ s time frame.

Discussion

As both standard MARTINI CG and ELNEDIN simulations do not allow the study of conformational flexibility, the discussion only concerns the AA and domELNEDIN simulations.

Flexible apo-structure moves from open to closed conformation

For several of the binding proteins, it has been established that the apo-structure is very flexible, and has a structural diversity which goes all the way from a completely open to a closed-up conformation [12, 15, 18, 59–64] and with an open-to-close exchange on the nanosecond or microsecond timescale [12, 15, 18, 59]. For both AA and domELNEDIN simulations starting from the open conformation, we also see a highly dynamical structure, which fluctuates around a fully open conformation on the 100 ns timescale (Fig. 1 and Fig. 6b). Then, in the domELNEDIN simulations extending to the microsecond timescale, conformational changes toward a closed conformation are clearly observed (Fig. 6). For some of the binding proteins, structural information on a closed conformation without ligand has been established. In most of these cases, the ligand bound and unbound structures are almost identical [61, 64]. However, a case has also been reported, where the closed conformation without substrate seems to close up in a distinguishable different conformation than when the substrate is present [59]. In the domELNEDIN simulations, the closed-up structure obtained from the simulations starting from the open conformation has a backbone RMSD of around 3 Å (Fig. 6b) compared to the crystal structure of the closed conformation with ligand bound. The degree of closing of the open apo-structure could be dependent on the exact domain definitions applied in the setup, but it seems not to be (Fig. 6). That the RMSD does not go below

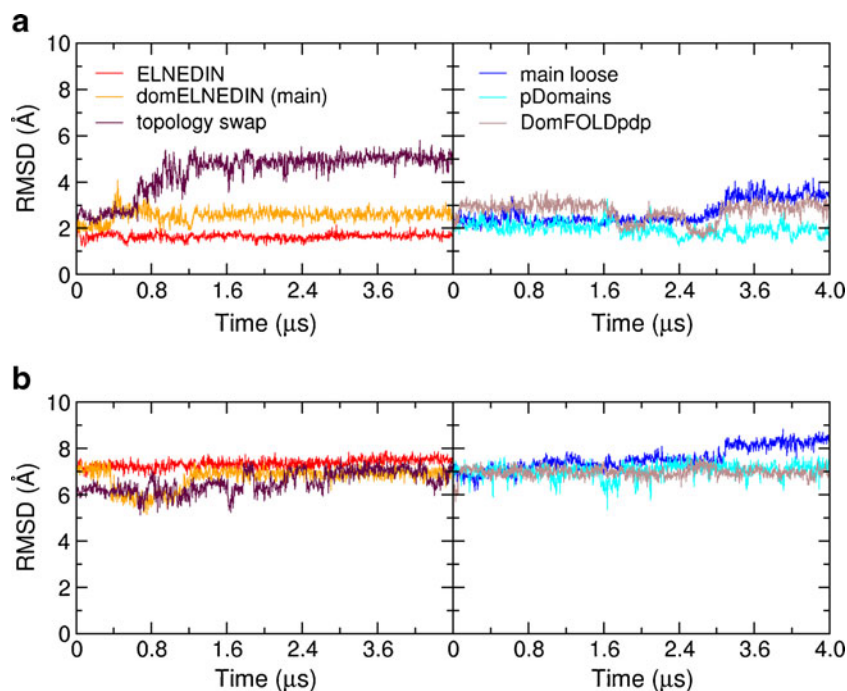


Fig. 7 Backbone RMSDs in ELNEDIN and domELNEDIN simulations starting from the closed conformation of LBP. **a** Compared to the crystal structure of the closed conformation. **b** Compared to the crystal structure of the open conformation. Left-most graphs show results from the ELNEDIN simulation, the domELNEDIN simulation using the main domain assignment and the “topology swap” setup, where the topology

for the open conformation is applied to the domELNEDIN simulation starting from the closed conformation. The right-most graphs show results from domELNEDIN simulations using three alternative domain assignments in the setup. The protein domain boundaries corresponding to main, main loose, pDomains and DomFOLDpdp can be seen in Table 1

3 Å can be because LBP actually closes up in a different manner when leucine is not present. Similarly, the stable domELNEDIN model of the closed conformation without leucine present also has an RMSD of 2.5–3 Å to the closed X-ray structure (Fig. 7a). However, the coarse model description may in itself be the reason for the observed RMSD, resulting in a stable closed conformation which is different from the X-ray structure.

A limitation of the domELNEDIN model is that the ENs set up inside the protein domains are based on the original structure, and any induced fit going from an open to a closed structure will thus not be supported. As the RMSDs between the domains in the closed and open structures are 0.7 Å for domain 1 and 0.6 Å for domain 2, this effect was not expected to be significant. However, if the simulation starting from the open conformation is applied the closed structure topology (topology swap), and thus, inside the domains, enforced the equilibrium distances between backbone beads corresponding to the closed structure, LBP was observed to undergo a conformational change all the way from the open conformation to a structure with RMSD of only 1.6 Å with respect to the crystal structure of the closed conformation (Fig. 6b). This shows that induced fit of the domains indeed play a role in the complete close-up of the protein, but from these simulations it cannot be determined if this fit is induced by interaction with the substrate or if it could also take place in the apo-situation.

To address this, atomistic simulations are needed, e.g., starting from a reverse coarse graining of the closed up apo-structure observed in the domELNEDIN simulations, and both with and without a leucine ligand included in the setup.

While the observation of induced fit here required the previous knowledge of the structure in both open and closed conformations, it could also be imagined that the introduction of a non-uniform stiffness of the ENs would allow for the simulation of induced fit between protein domains, e.g., by making the network weaker in areas based either on distance to the protein surface or based on knowledge from a short AA simulation. Structures from multiple protein conformations could also be used to derive a common domELNEDIN model for a particular protein, where the differences in the domELNEDIN models established from the individual structures are removed or modified, to allow for a model which is not biased by one structure in particular. It has to be tested, though, if such setups would still produce dynamics comparable to atomistic simulations on the nanosecond timescale.

Highly stable closed conformation

The simulations starting from the closed conformation with the leucine ligand removed from the setup show a very stable structure, also on the microsecond timescale (Fig. 7a). Even when the topology for the open conformation is applied in the

simulation starting from the closed conformation, the structure keeps a stable closed conformation, although the internal domain structure and backbone are changed to match those in the open conformation.

When interpreting the results, it should be kept in mind that the simulations starting from the closed apo-structure are artificial in the sense that the crystal structure used for the setup contained the substrate. It could be that an energetic barrier would keep the structure from closing this tightly *in vivo* without leucine present, and the leucine ligand would induce a fit in the structure, as discussed in the previous subsection.

Also, it cannot be ruled out that the observation of the microsecond stability of the closed conformation is due to an over-stabilization of the protein-protein interaction in the MARTINI CG FF. However, also the AA simulations of the closed ligand-free conformation show a very stable structure at the 100 ns timescale (Fig. 1), while, e.g., the maltose binding protein has been seen to go from closed to open conformation in a 30 ns simulation, when removing the ligand from the binding site [13]. As described in the Results section, MARTINI has previously successfully been used to study protein-protein interactions [33, 47, 48, 53, 54, 58], and it therefore seems reasonable to expect that the interaction of protein domains can be described within the FF as well. Naturally, the coarse model is a compromise, and in a very recent update to the MARTINI protein model (version 2.2

[77]), the description of side-chains have been further improved by adding particles with opposite charges on polar side chains as well as moving the charge on charged side-chains away from the van der Waals center of the charged bead, to allow for a higher resolution in the modeling of electrostatics interactions [77]. The water model has also been extended with particles with opposite and movable charges, to allow for polarization of the water beads [78]. Using this new version of the MARTINI CG protein and water model together with domELNEDIN on the closed conformation with the substrate removed, microsecond long simulations consistently show a stable structure (data not shown).

All microsecond domELNEDIN simulations in this study agree that LBP can close up without the leucine ligand present, as also observed for several other binding proteins [15, 59–64]. No matter if the closed-up conformation is identical to the one observed for the substrate-bound state, or if it deviates with an RMSD of 3–4 Å, as seen in the simulations starting from the open conformation, the existence of a closed conformation without substrate has implications for the ABC-transporter mechanism, that we will consider in the following.

The binding protein interacts with a transmembrane permease, and in this way helps ensure that the substrate transport is unidirectional [79]. In the few crystal structures of full ABC-transporter complexes, where the interaction between a transmembrane permease and its binding protein can be seen [80–83], the two subunits of the permease interact with each

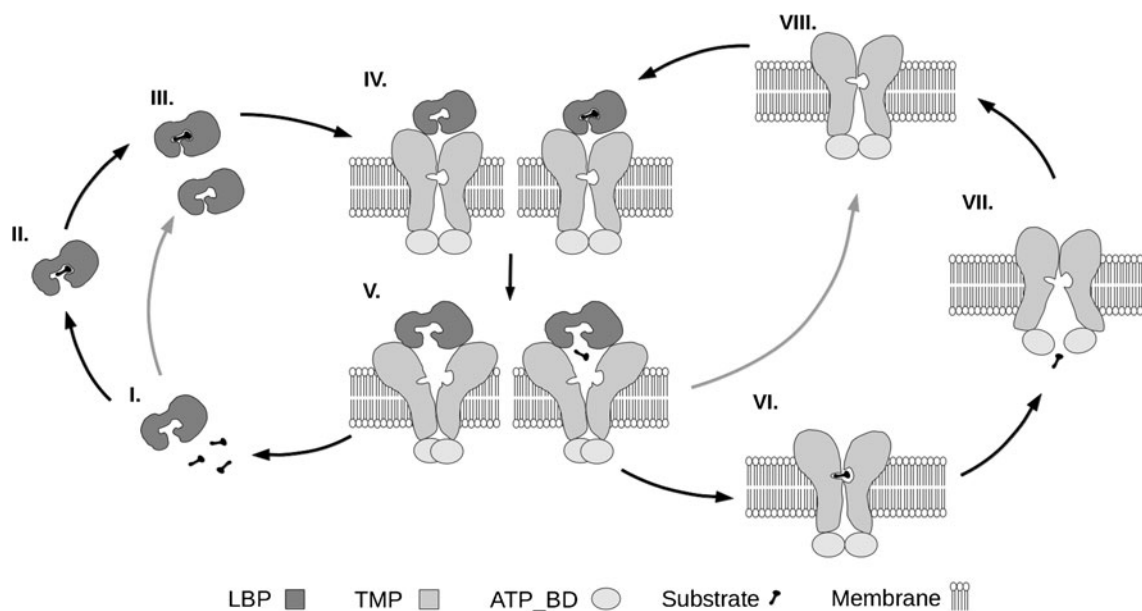


Fig. 8 Schematic of leucine transport through the membrane. The steps involved when LBP closes up without substrate bound, as suggested by the present study, is also included, and gray arrows indicate where it differs from the substrate-transport cycle. (I) LBP in the open, substrate-free form (*dark gray*) and unbound substrate (*black*) in the periplasm. (II) The substrate binds to LBP. (III) LBP closes up. (IV) LBP docks to the transmembrane permease (TMP)

(*light gray*), which is connected with the ATP-binding domains (ATP_BD) on the cytoplasmic side. (V) Binding triggers the opening of LBP and release of the substrate to the permease, if substrate is present, as well as release of the binding protein back to the periplasm. (VI–VII) The substrate is transported through the permease and released into the cytoplasm. (VIII) The permease is back in a state ready to interact with a closed up binding protein

their domain of the binding protein. It therefore seems plausible that the binding protein in its closed form docks to the permease while it is in an inward facing or occluded conformation (Fig. 8 IV). An ATP-driven conformational change to an outward facing form of the permease would then break open the binding protein, for the substrate to be released into the permease (Fig. 8 V). Yet another conformational change would convert the permease to an occluded conformation, releasing the binding protein in its open form into the periplasm again (Fig. 8 VI) [83]. The permease interaction would by this mechanism help the binding protein overcome the energy barrier going from the closed to the open conformation, ensuring that the ligand is not released erroneously when first captured by the binding protein.

At a first glance, this mechanism seems unlikely to be combined with the presence of a closed conformation without substrate, as this would result in seemingly futile ATPase activity. Nonetheless, experimental data and observations also support that empty binding protein interacts with the permease. Modeling the kinetics of a binding-protein dependent transport system it is clear that a model where only the substrate bound binding protein is recognized by the permease does not explain the experimental data, whereas a model where the substrate-bound and empty binding proteins compete for the permease interaction do [84, 85]. The histidine binding protein has been shown to have equal affinity for the transmembrane permease whether loaded with substrate or not [86], and the binding protein without substrate competes efficiently with the loaded binding protein, and thus inhibits the substrate transport [87]. It was also found that the ATPase activity is stimulated by the substrate-free binding protein, albeit at a lower level than when substrate is present [86]. Similar measurements have been carried out for the vitamin B12 importer. In this case the ATPase activity was equally stimulated by the binding protein with or without substrate, which suggests that the transporter is unable to distinguish between empty or B12-loaded binding protein [88]. These observations point to an overall inefficient substrate transport, as ATP hydrolysis does not seem tightly coupled to substrate transport. However, it has recently been recognized for other ATP-burning machines, that they too operate at efficiencies much lower than 100 % [89, 90], and it could be that it is rather an unlikely scenario to have biological machines running at maximal efficiency.

Conclusions

Using MD simulations we have studied the conformational flexibility of LBP when the substrate leucine is not present. CG models are used to access the dynamics on the microsecond timescale, and in this study three CG models based on the MARTINI CG FF have been evaluated for the purpose. The

standard MARTINI CG FF [29–31] is shown not to maintain a stable structure at the 100 ns timescale. The ELNEDIN model [33], where harmonic bonds are added between close-by backbone beads forming a global EN, is shown to provide simulations with a stable structure as well as dynamics comparable to atomistic simulations on the 100 ns timescale. However, due to the EN, this model does not allow the description of conformational flexibility. The domELNEDIN extension, presented in this work, is a modification to ELNEDIN where all EN bonds connecting protein domains are annihilated while keeping the bonds inside the protein domains. The domELNEDIN simulations of LBP show the same qualities as ELNEDIN on the 100 ns timescale, reproducing both the structural stability and the nanosecond dynamics comparable to atomistic simulations, while at the same time allowing for unbiased movements of the protein domains.

In microsecond long simulation, four different ways of assigning the 345 residues of LBP to the two protein domains were tested, and the outcome of the domELNEDIN simulations was seen not to depend significantly on the exact domain assignment. As the formation of harmonic bonds in the ENs is based on interatomic distances in the initial structure used for the setup, the domELNEDIN model topology established from respectively the open and closed conformations of LBP differ. Applying the closed conformation domELNEDIN model in a simulation starting from the open conformation and vice versa can therefore enable the modeling of a more complete conformational change. It is in this way observed that for LBP, an induced fit between the domains is needed for the open conformation to close up completely.

For several of the binding proteins, taking part in ABC transport systems, it has been established that a closed-up conformation exists without ligand present [15, 59–64]. All microsecond domELNEDIN simulations presented in this study show that this is also the case for LBP. Due to the coarseness of the model, it cannot be determined if the closed empty conformation is identical to the substrate-bound conformation, or if the substrate is what induces the fit in the structure. Whether the closed up apo-structure is identical or just very similar to the closed holo-structure, our results support previous experimental observations, which imply that the ABC-transport system is not exclusively triggered by a substrate-bound binding protein [86–88].

As it cannot be ruled out that protein-protein interaction, and thereby the closed conformation, is over-stabilized by the MARTINI CG FF, we will in a subsequent study test this hypothesis in more detail. Determining the free energy associated with the open-to-closed transition, e.g., by employing potential-of-mean-force calculations on the system, would give quantitative input to this discussion. Using the domELNEDIN approach together with the latest MARTINI CG FF improvements, we also plan to model how the conformational

flexibility of LBP is affected by the presence of substrate, and how the stability of the closed conformation of a binding protein, with and without the substrate bound, is affected by the interaction with the transmembrane permease.

Acknowledgments This work was supported by funds from the Danish National Research Foundation for the Pumpkin Centre, the Carlsberg Foundation (LT), and from the Aarhus Graduate School of Science and Technology (IS).

References

- Quiocho FA, Ledvina P (1996) Atomic structure and specificity of bacterial periplasmic receptors for active transport and chemotaxis: Variation of common themes. *Mol Microbiol* 20:17–25
- Dwyer MA, Hellinga HW (2004) Periplasmic binding proteins: A versatile superfamily for protein engineering. *Curr Opin Struct Biol* 14:495–504
- Gerstein M, Lesk AM, Chothia C (1994) Structural mechanisms for domain movements in proteins. *Biochemistry (N Y)* 33:6739–6749
- Sack JS, Saper M, Quiocho FA (1989) Structure of the L-leucine-binding protein refined at 2.4 Å resolution and comparison with the Leu/Ile/Val-binding protein structure. *J Mol Biol* 206:171–191
- Sack JS, Trakhanov SD, Tsigannik IH, Quiocho FA (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* 206:193–207
- Magnusson U, Salopek-Sondi B, Luck LA, Mowbray SL (2004) X-ray structures of the leucine-binding protein illustrate conformational changes and the basis of ligand specificity. *J Biol Chem* 279:8747–8752
- Quiocho FA (1991) Atomic structures and function of periplasmic receptors for active transport and chemotaxis. *Curr Opin Struct Biol* 1:922–933
- George AM, Jones PM (2012) Perspectives on the structure-function of ABC transporters: The switch and constant contact models. *Prog Biophys Mol Biol* 109:95–107
- Moussatova A, Kandt C, O'Mara ML, Tieleman DP (2008) ATP-binding cassette transporters in *Escherichia coli*. *Biochim Biophys Acta, Biomembr* 1778:1757–1771
- Oldham ML, Davidson AL, Chen J (2008) Structural insights into ABC transporter mechanism. *Curr Opin Struct Biol* 18:726–733
- Quiocho FA (1990) Atomic structures of periplasmic binding-proteins and the high-affinity active-transport systems in bacteria. *Philos Trans R Soc, B* 326:341–352
- Pang A, Arinaminpathy Y, Sansom MS, Biggin PC (2005) Comparative molecular dynamics—similar folds and similar motions? *Proteins: Struct, Funct, Bioinf* 61:809–822
- Stockner T, Vogel HJ, Tieleman DP (2005) A salt-bridge motif involved in ligand binding and large-scale domain motions of the maltose-binding protein. *Biophys J* 89:3362–3371
- Krewulak KD, Shepherd CM, Vogel HJ (2005) Molecular dynamics simulations of the Periplasmic Ferric-Hydroxamate binding protein FhuD. *Biometals* 18:375–386
- Kandt C, Xu Z, Tieleman DP (2006) Opening and closing motions in the Periplasmic vitamin B-12 binding protein BtuF. *Biochemistry (N Y)* 45:13284–13292
- Liu M, Su JG, Kong R, Sun TG, Tan JJ, Chen WZ, Wang CX (2008) Molecular dynamics simulations of the bacterial Periplasmic heme binding proteins ShuT and PhuT. *Biophys Chem* 138:42–49
- Silva D, Dominguez-Ramirez L, Rojo-Dominguez A, Sosa-Peinado A (2011) Conformational dynamics of L-lysine, L-arginine, L-ornithine binding protein reveals ligand-dependent plasticity. *Proteins: Struct, Funct, Bioinf* 79:2097–2108
- Loeffler HH, Kitao A (2009) Collective dynamics of periplasmic glutamine binding protein upon domain closure. *Biophys J* 97:2541–2549
- Trakhanov S, Vyas NK, Luecke H, Kristensen DM, Ma JP, Quiocho FA (2005) Ligand-free and bound structures of the binding protein (LivJ) of the *Escherichia coli* ABC leucine/isoleucine/valine transport system: trajectory and dynamics of the interdomain rotation and ligand specificity. *Biochemistry (N Y)* 44:6597–6608
- Perilla JR, Beckstein O, Denning EJ, Woolf TB (2011) Computing ensembles of transitions from stable states: dynamic importance sampling. *J Comput Chem* 32:196–209
- Bucher D, Grant BJ, Markwick PR, McCammon JA (2011) Accessing a hidden conformation of the maltose binding protein using accelerated molecular dynamics. *PLoS Comp Biol* 7:e1002034
- Ravindranathan KP, Gallicchio E, Levy RM (2005) Conformational equilibria and free energy profiles for the allosteric transition of the ribose-binding protein. *J Mol Biol* 353:196–210
- Su JG, Jiao X, Sun TG, Li CH, Chen WZ, Wang CX (2007) Analysis of domain movements in glutamine-binding protein with simple models. *Biophys J* 92:1326–1335
- Hall BA, Kaye SL, Pang A, Perera R, Biggin PC (2007) Characterization of protein conformational states by normal-mode frequencies. *J Am Chem Soc* 129:11394–11401
- Hub JS, de Groot BL (2009) Detection of functional modes in protein dynamics. *PLoS Comp Biol* 5:e1000480
- Tozzini V (2010) Multiscale modeling of proteins. *Acc Chem Res* 43:220–230
- Tozzini V, Rocchia W, McCammon JA (2006) Mapping all-atom models onto One-bead coarse-grained models: General properties and applications to a minimal polypeptide model. *J Chem Theory Comput* 2:667–673
- Yap E, Fawzi NL, Head-Gordon T (2008) A coarse-grained alpha-carbon protein model with anisotropic hydrogen-bonding. *Proteins: Struct, Funct, Bioinf* 70:626–638
- Marrink SJ, de Vries AH, Mark AE (2004) Coarse grained model for semiquantitative lipid simulations. *J Phys Chem B* 108:750–760
- Marrink SJ, Risselada HJ, Yefimov S, Tieleman DP, de Vries AH (2007) The MARTINI force field: Coarse grained model for biomolecular simulations. *J Phys Chem B* 111:7812–7824
- Monticelli L, Kandasamy SK, Periole X, Larson RG, Tieleman DP, Marrink S (2008) The MARTINI coarse-grained force field: extension to proteins. *J Chem Theory Comput* 4:819–834
- Lopez CA, Rzepiela AJ, de Vries AH, Dijkhuizen L, Huenenberger PH, Marrink SJ (2009) Martini coarse-grained force field: extension to carbohydrates. *J Chem Theory Comput* 5:3195–3210
- Periole X, Cavalli M, Marrink SJ, Ceruso MA (2009) Combining an elastic network with a coarse-grained molecular force field: structure, dynamics, and intermolecular recognition. *J Chem Theory Comput* 5:2531–2543
- Van der Spoel D, Lindahl E, Hess B, Groenhof G, Mark AE, Berendsen HJC (2005) GROMACS: fast, flexible, and free. *J Comput Chem* 16:1701–1718
- Hess B, Kutzner C, van der Spoel D, Lindahl E (2008) GROMACS 4: algorithms for highly efficient, load-balanced, and scalable molecular simulation. *J Chem Theory Comp* 4:435–447
- Berendsen HJC, Postma JPM, Vangunsteren WF, Dinola A, Haak JR (1984) Molecular-dynamics with coupling to an external bath. *J Chem Phys* 81:3684–3690
- Duan Y, Wu C, Chowdhury S, Lee MC, Xiong GM, Zhang W, Yang R, Cieplak P, Luo R, Lee T, Caldwell J, Wang JM, Kollman PJ (2003) A point-charge force field for molecular mechanics simulations of proteins based on condensed-phase quantum mechanical calculations. *Comput Chem* 24:1999–2012

38. Berendsen HJC, Postma JPM, van Gunsteren WF, Hermans J (1981) Interaction models for water in relation to protein hydration. In *Intermolecular forces*. Reidel, Dordrecht
39. Wang JM, Wolf RM, Caldwell JW, Kollman PA, Case DA (2004) Development and testing of a general amber force field. *J Comput Chem* 25:1157–1174
40. Wang JM, Wang W, Kollman PA, Case DA (2006) Automatic atom type and bond type perception in molecular mechanical calculations. *J Mol Graphics Modell* 25:247–260
41. Hinsen K (1998) Analysis of domain motions by approximate normal mode calculations. *Proteins* 33:417–429
42. The pDomains Server. <http://pdomains.sdsc.edu/v2/index.php>. Accessed 28 June 2013
43. The DomFOLD Server. <http://www.reading.ac.uk/bioinf/DomFOLD/>. Accessed 28 June 2013
44. Amadei A, Ceruso MA, Di Nola A (1999) On the convergence of the conformational coordinates basis set obtained by the essential dynamics analysis of proteins' molecular dynamics simulations. *Proteins: Struct, Funct, Genet* 36:419–424
45. Ceruso MA, Amadei A, Di Nola A (1999) Mechanics and dynamics of B1 domain of protein G: Role of packing and surface hydrophobic residues. *Protein Sci* 8:147–160
46. Ceruso MA, Grottesi A, Di Nola A (2003) Dynamic effects of mutations within two loops of cytochrome C(551) from *Pseudomonas aeruginosa*. *Proteins: Struct, Funct, Genet* 50:222–229
47. Sengupta D, Rampioni A, Marrink SJ (2009) Simulations of the C-subunit of ATP-synthase reveal helix rearrangements. *Mol Membr Biol* 26:422–434
48. Sengupta D, Marrink SJ (2010) Lipid-mediated interactions tune the association of Glycophorin A helix and its disruptive mutants in membranes. *Phys Chem Chem Phys* 12:12987–12996
49. Louhivuori M, Risselada HJ, van der Giessen E, Marrink SJ (2010) Release of content through mechano-sensitive gates in pressurized liposome. *Proc Natl Acad Sci U S A* 107:19856–19860
50. Yefimov S, Onck PR, van der Giessen E, Marrink SJ (2008) Mechanosensitive membrane channels in action. *Biophys J* 94:2994–3002
51. Ponting CP, Russell RR (2002) The natural history of protein domains. *Annu Rev Biophys Biomol Struct* 31:45–71
52. Richardson JS (1981) The anatomy and taxonomy of protein structure. *Adv Protein Chem* 34:167–339
53. Thogersen L, Schiott B, Vosegaard T, Nielsen NC, Tajkhorshid E (2008) Peptide aggregation and pore formation in a lipid bilayer: A combined coarse-grained and all atom molecular dynamics study. *Biophys J* 95:4337–4347
54. Periole X, Knepp AM, Sakmar TP, Marrink SJ, Huber T (2012) Structural determinants of the supramolecular organization of G protein-coupled receptors in bilayers. *J Am Chem Soc* 134:10959–10965
55. Jorgensen WL (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
56. Kaminski G (2001) Evaluation and reparametrization of the OPLS-AA force field for proteins via comparison with accurate quantum chemical calculations on peptides. *J Phys Chem B* 105:6474–6487
57. Oostenbrink C (2004) A biomolecular force field based on the free enthalpy of hydration and solvation: the GROMOS force-field parameter sets 53A5 and 53A6. *J Comput Chem* 25:1656–1676
58. de Jong DH, Periole X, Marrink SJ (2012) Dimerization of amino acid side chains: lessons from the comparison of different force fields. *J Chem Theory Comput* 8:1003–1014
59. Tang C, Schwieters CD, Clore GM (2007) Open-to-closed transition in apo maltose-binding protein observed by paramagnetic NMR. *Nature* 449:1078–1082
60. Walmsley AR, Shaw JG, Kelly DJ (1992) The mechanism of ligand-binding to the periplasmic C4-dicarboxylate binding-protein (Dctp) from *Rhodobacter capsulatus*. *J Biol Chem* 267:8064–8072
61. Flocco MM, Mowbray SL (1994) The 1.9 angstrom X-ray structure of a closed unliganded form of the periplasmic glucose/galactose receptor from *Salmonella typhimurium*. *J Biol Chem* 235:709–717
62. Wolf A, Shaw EW, Nikaido K, Ames GFL (1994) The histidine-binding protein undergoes conformational-changes in the absence of ligand as analyzed with conformation-specific monoclonal-antibodies. *J Biol Chem* 269:23051–23058
63. Bjorkman AJ, Mowbray SL (1998) Multiple open forms of ribose-binding protein trace the path of its conformational change. *J Mol Biol* 279:651–664
64. Oswald C, Smits SHJ, Höing M, Sohn-Boesser L, Dupont L, Le Rudulier D, Schmitt L, Bremer E (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
65. Holland TA, Veretnik S, Shindyalov IN, Bourne PE (2006) Partitioning protein structures into domains: why is it so difficult? *J Mol Biol* 361:562–590
66. Orengo CA, Michie AD, Jones S, Jones DT, Swindells MB, Thornton JM (1997) CATH—a hierarchic classification of protein domain structures. *Structure* 5:1093–1108
67. Andreeva A, Howorth D, Chandonia J, Brenner SE, Hubbard TJP, Chothia C, Murzin AG (2008) Data growth and its impact on the SCOP database: new developments. *Nucleic Acids Res* 36:D419–425
68. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE (2000) The protein data bank. *Nucleic Acids Res* 28:235–242
69. The Protein Data Bank. <http://www.pdb.org/pdb/home/home.do>. Accessed 28 June 2013
70. Veretnik S, Bourne PE, Alexandrov NN, Shindyalov IN (2004) Toward consistent assignment of structural domains in proteins. *J Mol Biol* 339:647–678
71. Hayward S, Berendsen HJC (1998) Systematic analysis of domain motions in proteins from conformational change: new results on citrate synthase and T4 lysozyme. *Proteins: Struct, Funct, Genet* 30:144–154
72. Zhang Z, Pfaendtner J, Grafmueller A, Voth GA (2009) Defining coarse-grained representations of large biomolecules and biomolecular complexes from elastic network models. *Biophys J* 97:2327–2337
73. McGuffin LJ (2008) The ModFOLD server for the quality assessment of protein structural models. *Bioinformatics* 24:586–587
74. Jones DT, Bryson K, Coleman A, McGuffin LJ, Sadowski MI, Sodhi JS, Ward JJ (2005) Prediction of novel and analogous folds using fragment assembly and fold recognition. *Proteins: Struct, Funct, Bioinf* 61(7):143–151
75. Alexandrov NN, Shindyalov IN (2003) PDP: Protein domain parser. *Bioinformatics* 19:429–430
76. Alden K, Veretnik S, Bourne PE (2010) dConsensus: a tool for displaying domain assignments by multiple structure-based algorithms and for construction of a consensus assignment. *BMC Bioinforma* 11:310–317
77. de Jong DH, Singh G, Bennett WFD, Arnarez C, Wassenaar TA, Schäfer LV, Periole X, Tieleman DP, Marrink SJ (2013) Improved parameters for the martini coarse-grained protein force field. *J Chem Theory Comput* 9:687–697
78. Yesylevskyy SO, Schäfer LV, Sengupta D, Marrink SJ (2010) Polarizable water model for the coarse-grained MARTINI force field. *PLoS Comp Biol* 6:e1000810
79. Krupka RM (1992) Testing models for transport-systems dependent on periplasmic binding-proteins. *Biochim Biophys Acta* 1110:11–19

80. Hollenstein K, Frei DC, Locher KP (2007) Structure of an ABC transporter in complex with its binding protein. *Nature* 446:213–216
81. Hvorup RN, Goetz BA, Niederer M, Hollenstein K, Perozo E, Locher KP (2007) Asymmetry in the structure of the ABC transporter-binding protein complex BtuCD-BtuF. *Science* 317:1387–1390
82. Oldham ML, Khare D, Quioco FA, Davidson AL, Chen J (2007) Crystal structure of a catalytic intermediate of the maltose transporter. *Nature* 450:515–521
83. Oldham ML, Chen J (2011) Crystal structure of the maltose transporter in a pretranslocation intermediate state. *Science* 332:1202–1205
84. Bohl E, Shuman HA, Boos W (1995) Mathematical treatment of the kinetics of binding-protein dependent transport-systems reveals that both the substrate loaded and unloaded binding-proteins interact with the membrane-components. *J Theor Biol* 172:83–94
85. Shilton BH, Mowbray SL (1995) Simple-models for the analysis of binding protein-dependent transport-systems. *Protein Sci* 4:1346–1355
86. Ames GFL, Liu CE, Joshi AK, Nikaido K (1996) Liganded and unliganded receptors interact with equal affinity with the membrane complex of periplasmic permeases, a subfamily of traffic ATPases. *J Biol Chem* 271:14264–14270
87. Prossnitz E, Gee A, Ames GF (1989) Reconstitution of the histidine periplasmic transport-system in membrane-vesicles—energy coupling and interaction between the binding-protein and the membrane complex. *J Biol Chem* 264:5006–5014
88. Borths EL, Poolman B, Hvorup RN, Locher KP, Rees DC (2005) In vitro functional characterization of BtuCD-F, the escherichia coli ABC transporter for vitamin B-12 uptake. *Biochemistry (N Y)* 44:16301–16309
89. Lervik A, Bresme F, Kjelstrup S, Miguel Rubi J (2012) On the thermodynamic efficiency of Ca²⁺-ATPase molecular machines. *Biophys J* 103:1218–1226
90. Kawaguchi K (2008) Energetics of kinesin-1 stepping mechanism. *FEBS Lett* 582:3719–3722