

Supporting Online Material for

Molecular Basis of Alternating Access Membrane Transport by the Sodium-Hydantoin Transporter Mhp1

Tatsuro Shimamura, Simone Weyand, Oliver Beckstein,
Nicholas G. Rutherford, Jonathan M. Hadden, David Sharples, Mark S. P. Sansom,
So Iwata,* Peter J. F. Henderson,* Alexander D. Cameron*

*To whom correspondence should be addressed. E-mail: s.iwata@imperial.ac.uk (S.I.);
p.j.f.henderson@leeds.ac.uk (P.J.F.H.); a.cameron@imperial.ac.uk (A.D.C.)

Published 23 April 2010, *Science* **328**, 470 (2010)
DOI: 10.1126/science.1186303

This PDF file includes:

Materials and Methods
SOM Text
Figs. S1 to S9
Tables S1 to S3
References

Other Supporting Online Material for this manuscript includes the following:
(available at www.sciencemag.org/cgi/content/full/328/5977/470/DC1)

Movies S1 to S3

Supporting Online Material

Materials and Methods

Protein Expression and Purification

SeMet-substituted Mhp1 was expressed in *E. coli* B834 as described (S1), except that cells were grown in 2l flasks each containing 500 ml of a defined medium supplemented with 50 mg/l seleno-L-methionine (S2). The protein was purified as before (S1, S3) with 10 mM DTT added to all buffers.

Crystallisation

Crystals were grown at 4°C using the hanging drop vapor diffusion technique. To form the drops, 1.5 µl of protein solution at a concentration of ~10 mg/ml was mixed with an equal volume of reservoir solution containing 24-28% PEG 300, 100 mM NaCl and 100 mM Bicine pH 9.0 and equilibrated against 400 µl of reservoir solution. Crystals appeared after one day and grew to full size within a week. Two types of crystals grew in these drops, some rod-like with cell dimensions similar to the crystals of the previously reported structures (S1) and others hexagonal.

Data collection

Data were collected from the hexagonal crystals. Prior to data collection, cryo-solution (30% PEG 300, 100 mM NaCl and 100 mM Bicine pH 9.0) was gradually added to the drop and left to equilibrate for six to seven days. The crystals were then mounted in loops and flash-cooled in liquid nitrogen. X-ray diffraction data were collected at 100 K at a wavelength of 0.979147 Å on beamline ID29 at the European Synchrotron Radiation Facility using an ADSC Q210 CCD detector and processed using the HKL suite of programs (S4). Further processing was carried out using programs from the CCP4 package (S5). The space group was determined to be *P*61 with one molecule in the asymmetric unit.

Structure Solution and Refinement

The structure was first solved by molecular replacement in Phaser (S6) using a trimmed down model (TMs 1,2, N-term of 3, 5, 6, 7, C-term of 10 and 12) of the outward-facing structure (S1). The resulting phases were improved by density modification in RESOLVE (S7) using the prime-and-switch option (S8). Anomalous difference Fourier maps calculated using these phases showed clearly the positions of the selenium atoms in the structure. The positions of these selenium atoms were then used to solve the structure by SAD. The selenium positions were refined and phases calculated in SHARP (S9). After density modification with a solvent content of 75% the protein could clearly be traced in the electron density maps. The map

was slightly improved by adding the phases derived from the outward-facing model into the initial cycles of density modification. Model building was performed in O (S10) starting from the higher resolution outward-facing Mhp1 model, including only the parts of the structure that could be seen in the electron density maps. The model building was aided by the positions of the selenium peaks in the anomalous difference maps (Fig S1B). Refinement was carried out in Phenix (S11) at a resolution of 3.8 Å. Additional geometrical restraints were put on the hydrogen bonding distances in the alpha helices. B-factors were not refined but the use of TLS (S12) resulted in improved electron density maps after B-factor sharpening. After each round of refinement the improvement of the resulting electron density maps allowed additional residues to be inserted into the model. The density was also sufficiently good to be able to correct errors in model building. As the refinement proceeded and the maps improved it became evident that the density could be satisfied by rigid body movements of segments of the outward-facing structure rather than moving amino acids individually. Thus the outward-facing structure was rebuilt into the improved maps by minimizing the movement of individual side chains while satisfying the associated electron density and positions of the selenomethionines. In the final cycle of refinement methionines were replaced by selenomethionines, F' and F'' refined against the anomalous pairs and the single TLS group was replaced by four groups: the bundle; the hash motif; helix 5 and the three C-terminal helices, as suggested by an analysis of the structure. The final structure contains all residues from 6 to 470 and has an R_{factor} of 27.3% and a corresponding R_{free} (based on 10% of the data) of 31.3%

Superpositions were carried out using Lsqman (S13) such that all matching C_{α} pairs were less than 3.8 Å apart after superposition. The figures were made in Pymol (S14) and VMD (S15).

Simulations

Structure preparation

The three states of Mhp1 are represented by the crystal structures of the outward-facing state (PDB accession code 2JLN) (S1), the occluded state, (2JLO) (S1) and the inward-facing state reported here. Equilibrium molecular dynamics (MD) simulations (Table S2) for the apo transporter used the full length structures (outward-facing: Glu 8 – Gly470; occluded: Arg 10 – Gly 470; inward-facing: Ile 6 Gly 470). For the DIMS simulations and all other equilibrium MD with ligands, proteins were prepared as the shortest common sequence Arg 10 – Gly 470.

The equilibrium MD simulations in the apo state (Table S2) and all dynamic importance sampling (DIMS) simulations were based on crystal structures with all ligands removed. For simulations with a sodium ion in the Na site (Table S2), the crystallographic Na^+ ion was retained but the benzylhydantoin ligand, where present, was deleted (2JLN and 2JLO). Na^+ was modelled into the inward-facing

structure by superimposing the occluded state structure (2JLO) by a backbone RMSD fit onto the inward-facing structure. Starting structures for simulations with Na⁺ and benzylhydantoin (Table S2) were prepared in a similar manner by superimposing 2JLO onto 2JLN or the inward-facing structure respectively.

DIMS transitions

Transitions were simulated with the DIMS method (S16-S18). For computational efficiency the protein was simulated in an implicit solvent using the ACE2 generalized Born model (S19). Langevin dynamics at 300 K with an integration time step of 2 fs was used while all hydrogen bonds were constrained with the SHAKE algorithm (S20). Simulations were performed with Charmm 36a1 (S21) and the Charmm27 all-atom force field (S22). The membrane was omitted; this gross approximation is only possible because the target conformation implies implicit restraints on the protein motions that result in a stable protein trajectory on the time scale of the simulations.

DIMS requires a progress variable, a metric that indicates the distance of a protein conformation from the target, and a biasing scheme that alters the dynamics so that transitions are preferentially sampled. As a robust progress variable the root mean square distance (RMSD) from the target, computed from backbone and C_β atoms, was used. The progress variable was evaluated for the non-hydrogen atoms of the structural elements TM1, TM3, TM4, TM5, TM6, TM8, TM9, and TM10. The target structures were reoriented relative to the current conformation every ten MD steps; only the backbone and C_β atoms of the static core elements of TM1, TM2, TM6, TM7, TM11, and TM12 were used for the translational and rotational fit. Transitions were computed between the three known states of Mhp1 ($A \rightarrow Q \rightarrow B$ where Q was the occluded state and A and B could each be either the outward-facing or the inward-facing state). The first half of the transition ($A \rightarrow Q$) was considered complete if the RMSD between the simulated conformation $X(t)$ and the intermediate target structure Q fell below 0.3 Å. Then the target was switched from Q to B and the transition continued seamlessly until the RMSD between $X(t)$ and B dropped below the cut-off of 0.3 Å.

Here we used the soft ratcheting biasing scheme (S17, S23), which implements a “Maxwell's daemon” that checks for every trial MD step if the system moves towards the target, as measured by the progress variable. If a move would lead away from the target then the step is only accepted with a probability that decreases with the magnitude of the deviation; the details of the decrease are controlled by the single tuning parameter $\Delta\varphi_0$. All MD steps towards the target are always accepted. This leads to a biased dynamics in which every single, microscopic MD step evolves naturally and in which no additional forces are exerted on the system. Thus the resulting trajectories are presumed close to ones that would have been obtained from unbiased equilibrium simulations (S18). Because the system can become trapped, any trajectory was aborted in which more than 106 consecutive attempts were rejected. A soft ratcheting parameter of $\Delta\varphi_0 = 10^{-5}$ Å was found to provide a sufficient number of successful transitions while maintaining trajectory diversity.

Three hundred trajectories were initialized in each direction, using the energy minimized crystal structures as starting and ending conformations. The completion rate for transitions from the outward open to the inward open state was 51% versus 58% in the reverse direction, producing 152 and 176 complete transitions.

Coarse grained MD simulations

In order to find the position of Mhp1 in the membrane, coarse grained (CG) self assembly simulations were performed (S24). A local modification of the MARTINI coarse grained force field (S25) was used to model lipids, the protein, water and ions. As initial protein structures, we used the Mhp1 crystal structures after removal of any ligands. The protein conformation was maintained closely to the crystallographic structure with a strong Gaussian network model with harmonic springs of force constant of $15 \times 10^3 \text{ kJ}\cdot\text{mol}^{-1}\cdot\text{nm}^{-2}$ connecting all residues within a distance cut-off of 0.9 nm.

Simulations were performed with Gromacs 4 (S26). The systems were run in the $NP_zP_{xy}T$ ensemble at $T = 323 \text{ K}$ and $P = 1 \text{ bar}$ using weak coupling algorithms. The time step was typically 20 fs. A 4:1 mixture of 1-palmitoyl-2-oleoylphosphatidylethanolamine (POPE) and 1-palmitoyl-2-oleoylglycerol-3-phosphoglycerol (POPG) lipids assembled around Mhp1 within less than 50 ns and was fully equilibrated after 200 ns. The coarse grained system was translated back into an atomistic system using the CG2AT protocol (S27). Because Mhp1 is a compact molecule and does not undergo sizable conformational changes when being embedded in the membrane, we substituted the original crystal structure in place of the CG2AT-back-translated protein to obtain a starting point for further simulations.

Atomistic Molecular Dynamics Simulations

Atomistic MD simulations used the well-established Gromos96 43a1 force field (S28) for the protein and Åqvist's parameters for ions (S29), combined with the appropriate united-atom parameters for lipids (S30) and the simple point charge (SPC) water model (S31). Parameters for benzylhydantoin were based on an initial PRODRG topology (S32) and refined with the parameters for the peptide bond and backbone (for the hydantoin ring) and phenylalanine, using quantum mechanical calculations of the partial charges with Gaussian03 (Gaussian, Inc. www.gaussian.com) as guidance.

Atomistic MD simulations were also performed with Gromacs 4. The classical equations of motions were computed with a leap-frog integrator at a time step of 2 fs for $\sim 100 \text{ ns}$ unless otherwise noted (Table S2). Bonds with hydrogen atoms were constrained with the P-LINCS algorithm (S33). Periodic boundary simulations were set up to resemble conditions of the uptake assay experiments (S3) at room temperature and pressure; hence semi-isotropic pressure coupling ($P_z = 1 \text{ bar}$, weak coupling algorithm) and constant velocity rescaling $T = 298.15 \text{ K}$ (S34) were used to simulate a $NP_zP_{xy}T$ ensemble. Van der Waals interactions were cut off at 1 nm; long range Coulomb interactions were computed with the SPME method (S35) (fourth

order spline interpolation on a 0.12 nm grid with a 1 nm real space cut-off).

Each system contained about 65,000 atoms. The system measured approximately 10 nm × 10 nm in the plane of the membrane and about 9 nm normal to the membrane. All ionisable residues of Mhp1 were simulated in their default charge state, as predicted by PROPKA (S36).

The membrane contains a 4:1 mixture of POPE and POPG lipids to approximate the *Escherichia coli* membranes (S37) used in experiments on Mhp1 (S1, S38). The multi-scale approach (CG self assembly MD followed by CG2AT conversion and atomistic simulations) ensures full mixing of POPE and POPG lipids. Typically, bilayers contained about 240 POPE and 60 POPG molecules in total; numbers varied slightly between simulations of different crystal structures because each system was built to fit the particular conformation of Mhp1.

The simulation systems contain free sodium and chloride ions at a physiological concentration of 100 mM with additional counter ions to make the system charge neutral. Initial structures were obtained from the coarse grained simulations, energy minimized, and relaxed for 5 ns with position restraints (isotropic force constant $k = 1 \times 10^3 \text{ kJ} \cdot \text{mol}^{-1} \cdot \text{nm}^{-2}$) on all protein non-hydrogen atoms to allow repacking of lipids and equilibration of water and ions around the protein. For simulations with a Na⁺ ion in its binding site, the original sodium ion positions in the outward-facing (2JLN) and occluded (2JLO) structures were retained; for the inward-facing structure, the Na⁺ ion position in the occluded structure was used as described above. All Na⁺ site-occupied simulations included an additional equilibration phase during which only the Na⁺ site-ion was restrained; this equilibration phase lasted 5 ns (except for the occluded state for which 1 ns was deemed sufficient).

Simulations of the inward-facing state with Na⁺ in the Na site were repeated six times for 10 ns each. The starting conformations for these repeats were picked from the Na⁺-restrained equilibration phase. Other simulations were run for ~100 ns, as shown in Table S2. In total, over 1.6 μs were simulated.

Analysis

Order parameters

In order to quantify and easily visualize functionally relevant conformational changes, we define a number of order parameters that characterize functionally relevant degrees of freedom. Here we are interested in the movements of the gating systems in Mhp1, and hence we define our order parameters to be roughly the extent of opening of these gates. Each one consists of a distance between the centers of mass of two groups of C_α atoms. These groups were chosen as to span the largest range of movement exhibited by the thick and thin gate components. The movement of the extracellular (EC) thin gate was described by the distance q_{EC} between the

mobile TM9/TM10 loop (N360, T361, F362) and the bundle stationary residues I47, A48, A49 on TM1. The intracellular (IC) thin gate distance q_{IC} was calculated between I161, T162, F163 (TM5) and D229, I230, V231 (TM6). The thick gate was characterized by the extent of the Na sodium binding site as the distance q_{Na} between A309, S312, T313 on TM8 and A38, I41 on TM1. The values of the order parameters in the crystal structures are listed in Table S3 for reference.

The order parameter q_{Na} measures the distance of TM8 from TM1 across the Na-binding site, i.e. the opening of the thick gate. Similarly, q_{EC} and q_{IC} measure the opening of the EC and IC gates.

Plotting one of the gate order parameters over time along a trajectory such as $q_{\text{EC}}(t)$ shows the time evolution of the gate. When two parameters such as $q_{\text{EC}}(t)$ and $q_{\text{Na}}(t)$ are plotted against each other for each time step then the resulting graph reveals correlations between the two gates. In a more general sense such a graph is a projection of the high-dimensional conformational space of the protein onto a low-dimensional space spanned by the two order parameters. For gate order parameters we call this “gating space”.

Progress measure $\Delta\Delta\text{RMSD}_{A,B}$

DIMS trajectory ensembles can be analyzed by computing averages of observables of interest as a function of an arbitrary progress variable. In order to be able to directly compare $A \rightarrow Q \rightarrow B$ transitions (A =outward facing (pdb code 2JLN), Q =occluded (2JLO), B =inward facing) with the reverse transitions ($B \rightarrow Q \rightarrow A$), we employed a generalization of the “Delta-RMSD” measure that was used previously (S18). We define the “Delta Delta RMSD” of a conformation X and two endpoint conformations A and B with one intermediate Q as

$$\Delta\Delta\text{RMSD}_{A,B}(X) = [\rho_A(X) - \rho_Q(X)] - [\rho_B(X) - \rho_Q(X)] = \rho_A(X) - \rho_B(X)$$

where

$$\rho_\alpha(X) = \sqrt{\frac{1}{3N} \sum_{i=1}^{3N} (X_i - X_{\alpha,i})^2}$$

is the root mean square distance (RMSD) of a trajectory coordinate frame $X(t)$ from the reference coordinates X_α for N atoms. As shown, $\Delta\Delta\text{RMSD}_{A,B}$ is formally equivalent to the “Delta RMSD” between the endpoints. Negative values of $\Delta\Delta\text{RMSD}_{A,B}$ indicate that a structure is closer to the reference structure A , whereas positive values are to be found in the vicinity of B . The $\Delta\Delta\text{RMSD}_{A,B}$ is anti-symmetric in the sense that $\Delta\Delta\text{RMSD}_{A,B}(X_A) = -\Delta\Delta\text{RMSD}_{A,B}(X_B)$, but this does not imply any particular value for the intermediate $\Delta\Delta\text{RMSD}_{A,B}(X_Q)$. For Mhp1, the outward facing state has $\Delta\Delta\text{RMSD}_{A,B} = -3.3 \text{ \AA}$, the occluded state -1.7 \AA , and the inward facing one $+3.3 \text{ \AA}$. $\Delta\Delta\text{RMSD}_{A,B}$ was chosen from a number of different trial progress variables because it captures both parts of the transition in comparable detail and illustrates the fact that the transition between occluded and inward facing state is a larger

conformational change than the one between outward facing and occluded. It also has the useful property that it can be computed across a continuous $A \rightarrow Q \rightarrow B$ trajectory unlike piece-wise progress measures such as using ρ_Q for the $A \rightarrow Q$ part of the transition and ρ_B for $Q \rightarrow B$.

Supporting Online Text

Stability of Equilibrium Molecular Dynamics Simulations

The equilibrium molecular dynamics (MD) simulations were set up and conducted as described in Methods. The quality of the simulations (S39) was assessed by calculation of the root mean square fluctuations (RMSF) of the protein C_α atoms and the root mean square distance (RMSD) of the protein structure over time (C_α atoms) from the initial structure. The RMSF for atoms in secondary structure elements tends to be <1 Å (Fig. S4). Mobile elements such as the termini or solvent-exposed loops show higher fluctuations, up to about 4 Å. These values are very similar for all simulations, regardless of the presence or absence of substrates in the starting structure (data not shown). The consistent RMSF profiles indicate that the secondary structure mobility is not influenced by the resolution of the three starting structures. The outward-facing conformation (2JLN) was refined at 2.85 Å, the occluded state (2JLO) at 4 Å (S1), and the inward-facing structure at 3.8 Å. More subtle differences exist in the thin gates [around T162 (intracellular (IC) gate) and T361 (extracellular (EC) gate)] between simulations, which reflect different degree of flexibility in the different functional states. The overall pattern of RMSF values is also consistent with the RMSF derived from the refined B-factors of the high-resolution Mhp1 structure 2JLN; RMSF = $(3B/8\pi^2)^{1/2}$. The simulations starting from the lower resolution structures typically have RMSD values between 2 Å and 3 Å after 100 ns whereas the higher resolution simulations stabilize between 1.5 Å and 2.3 Å (data not shown). RMSF and RMSD are similar to what has been observed for simulations of relatively high resolution membrane proteins (S40). These measures indicate that the structural models for Mhp1 used for the simulations are stable and that the MD trajectories are sampling experimentally relevant conformations.

Dynamic Importance Sampling-MD simulations

Because the equilibrium MD simulations do not connect the inward facing with the outward facing state, we employed dynamic importance sampling MD (DIMS-MD, (S16-S18)) simulations (see Methods) in order to sample these transitions. Fig. S6 shows the region of gating space covered by the ensemble of DIMS transitions. The trajectories are not shown explicitly. Instead, the density of points (q_{EC}, q_{Na}) was computed and its negative logarithm displayed in order to clearly indicate the regions of gating space that are visited by the DIMS trajectories. Fig. S6 shows that intermediate states between the crystal structures are visited. The DIMS simulations are traversing the barrier between the occluded and inward facing state that was

not sampled during the equilibrium simulations. The dynamic importance sampling method is designed to only focus computing time on sampling “important” events, i.e. transitions between states. Hence, very few states are visited that lie outside the regions between the end states of the DIMS simulations.

The evolution of the individual order parameters along simulated transitions are depicted in Fig S7A. The progress of the transition is measured by comparing the RMSD of frames along the transition to the outward and the inward facing state ($\Delta\text{RMSD}_{\text{outward,inward}}$) as described in Methods; it tracks the structural changes during a transition and allows a direct comparison of trajectories in both directions. For example, in Fig S7A the outward facing state at the beginning of the transition from the outward facing conformation to the inward facing one is characterized by an open EC gate ($q_{\text{EC}}=1.6 \text{ nm}$) and closed thick gate ($q_{\text{Na}}=1.1 \text{ nm}$). The transition proceeds to the occluded state in which all three gates are closed. Finally, the thick gate opens together with the IC gate. (The values for the gates in the crystal states are stated in Table S3.) The concomitant opening of these two gates is a consequence of the relatively simple RMSD progress variable that was employed for the present DIMS simulations. RMSD is probably too coarse a measure to resolve the exact order of gating events. On an alternative pathway, the thick gate would open first and reach a state with the IC gate closed (gray cartoon in Fig. 5); subsequent opening of the thin IC gate would lead to the inward facing, open conformation.

The trajectories are independent and stochastic and hence the order parameters vary between runs, as evidenced by the one-standard deviation error bars. The range of observed values mostly contains values near the transition. This is not surprising, given that DIMS avoids extensive sampling of the stable states in order to focus on the productive transition events. Transitions in forward and backward direction follow slightly different pathways as seen by the difference in order parameters as a function of $\Delta\text{RMSD}_{\text{outward,inward}}$, especially of the thick gate distance.

The average potential energy of the system, U , along the transition (Fig. S7B) increases over the first few frames by as the system heats up from the initial structure that corresponds to a 0-K conformation due to energy minimization. All trajectories obtain a temperature of $\sim 300 \text{ K}$ within the first 2–3 ps (data not shown). During this initial heating phase the overall structure changes little ($\Delta\text{RMSD}_{\text{outward,inward}} < 0.1 \text{ \AA}$ and change in RMSD towards the target of the DIMS run $< 0.05 \text{ \AA}$; data not shown) although the internal energy increases by $\sim 400 \text{ kcal/mol}$. Inspection of the individual energy terms shows that this increase is mainly due to the bonded interactions, for which small changes in distances from the energy-minimized structure due to thermal vibrations translate to large increases in internal energy. The average potential energy varies smoothly and displays fluctuations on the order of 50 kcal/mol (standard deviation from the mean). A small barrier of $\sim 200 \text{ kcal/mol}$ is seen after the initial heating phase, which was also observed for other DIMS transitions with soft-ratcheting on an RMSD progress variable (S17). The transition ensembles do not exhibit other barriers larger than the fluctuations in U . The absence of barriers demonstrates that the computed transition paths are accessible to the protein; there are no steric

clashes or physically un-attainable conformations along the pathway. The potential energies of the endpoints are of comparable magnitude (within one standard deviation) once the heating phase is discounted. The energy profiles along the path differ, indicating that forward and backward transitions sample pathways that are not identical, which was also borne out by the gate distances (Fig S7A). Because the internal energy is very sensitive to small structural differences, it is generally less useful in order to evaluate the true energetics of a transition (which are determined by a potential of mean force) and, in our case, mainly serves to ascertain that the transition proceeds without generating clashes or unphysical conformations or running into dead ends.

The soft-ratcheting algorithm (S23) employed in the present application of DIMS appears to let the system evolve in steps close to equilibrium (S17, S18). These generated transitions are, thus, candidates for the conformational transition that connects the outward-facing and the inward-facing conformations.

Supplementary Tables

Table S1. Data collection, phasing and refinement statistics.

Resolution (Å) ^a	40–3.8 (3.94–3.80)
Cell dimensions (Å)	$a=b=173.9$ $c=74.5$
Number of measured reflections	47,973
Number of unique reflections	12,492
Completeness (%)	97 (96)
Redundancy	3.8 (2.8)
$I/\sigma(I)$	17.6 (1.4)
R_{merge} (%)	8.2 (69.4)
R_{cullis} (%) (33–4Å) ^b	0.96 (1.00)
Phasing Power (33–4Å) ^c	0.46 (0.19)
$R_{\text{factor}}^{\text{d}}$ (%)	27.3
$R_{\text{free}}^{\text{d}}$ (%)	31.3

rms deviations from ideal values

Bonds (Å)	0.006
Angle (°)	1.038
Ramachandran plot outliers (%) ^e	1.1

^a Values in parentheses refer to data in the highest resolution shell.

^b $R_{\text{cullis}} = \sum |F_{PH} - |F_P + F_H|| / \sum |F_{PH} - F_P|$

^c Phasing power = rms ($|F_H| / |F_{PH} - |F_P + F_H||$)

^d $R_{\text{factor}} = \sum |F_{\text{obs}} - F_{\text{calc}}| / \sum F_{\text{obs}}$. The R_{free} , is the same as the R_{factor} but for the 10% of test reflections.

^e as defined in MolProbity (S41)

Table S2. Equilibrium MD simulations performed.

Conformation	Na site	S1 site	Simulations
Outward-facing	–	–	100 ns
Occluded	–	–	100 ns
Inward-facing	–	–	100 ns
Outward-facing	Na ⁺	–	3×100 ns
Occluded	Na ⁺	–	3×100 ns
Inward-facing	Na ⁺	–	100 ns, 6×10 ns
Outward-facing	Na ⁺	Benzylhydantoin	105 ns
Occluded	Na ⁺	Benzylhydantoin	100 ns
Inward-facing	Na ⁺	Benzylhydantoin	4×105 ns

Table S3. Gate order parameters of energy minimized crystal structures.

Conformation	PDB code	EC gate q_{EC} (nm)	IC gate q_{IC} (nm)	Thick gate q_{Na} (nm)
Outward-facing	2JLN	1.60	1.12	0.56
Occluded	2JLO	0.83	1.10	0.55
Inward-facing	2X79	0.89	2.27	1.07

Supplementary Figures

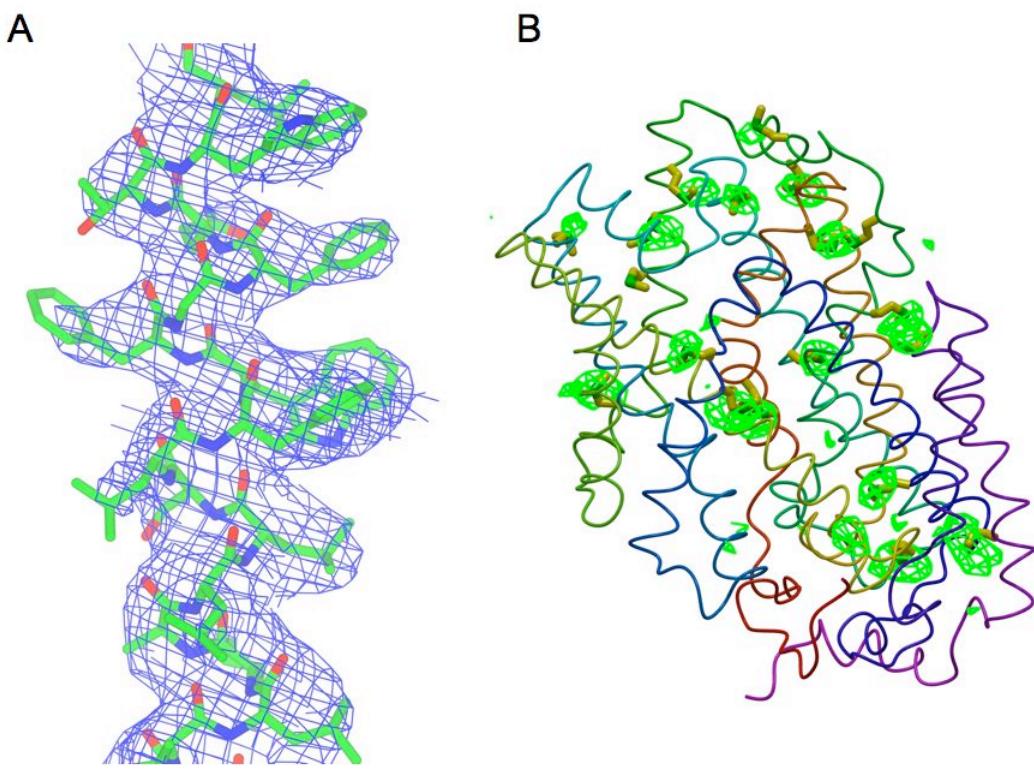


Figure S1. Electron density associated with Mhp1. **(A)** Refined electron density map ($2mF_{\text{obs}} - DF_{\text{calc}}$) for the inward-facing Mhp1 at 3.8 \AA resolution. The map is contoured at 1.0σ . **(B)** Anomalous difference map calculated using phases from the refined model and contoured at 3.1σ . The positions of the selenomethionines are shown on the ribbon representation, which has been color ramped according to residue number, starting with red at the N-terminus and finishing with purple at the C-terminus.

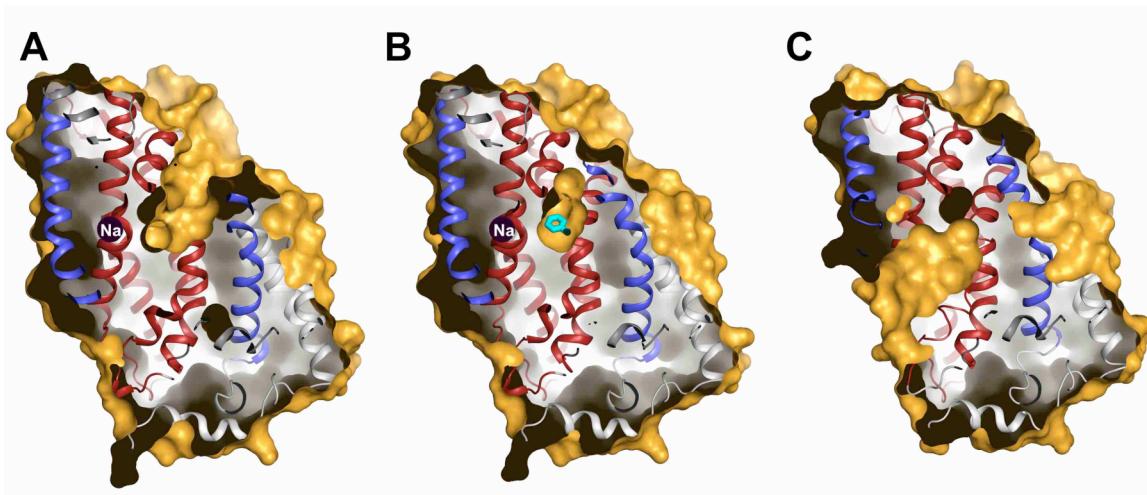


Figure S2. Cavities in Mhp1. Surface representations of **(A)** the outward facing structure, **(B)** the occluded structure and **(C)** the inward facing structure, sliced to show the respective cavities. The figure has been adapted from (1). The ribbon representation of Mhp1 is coloured as in Fig. 1. The benzylhydantoin is shown in cyan.

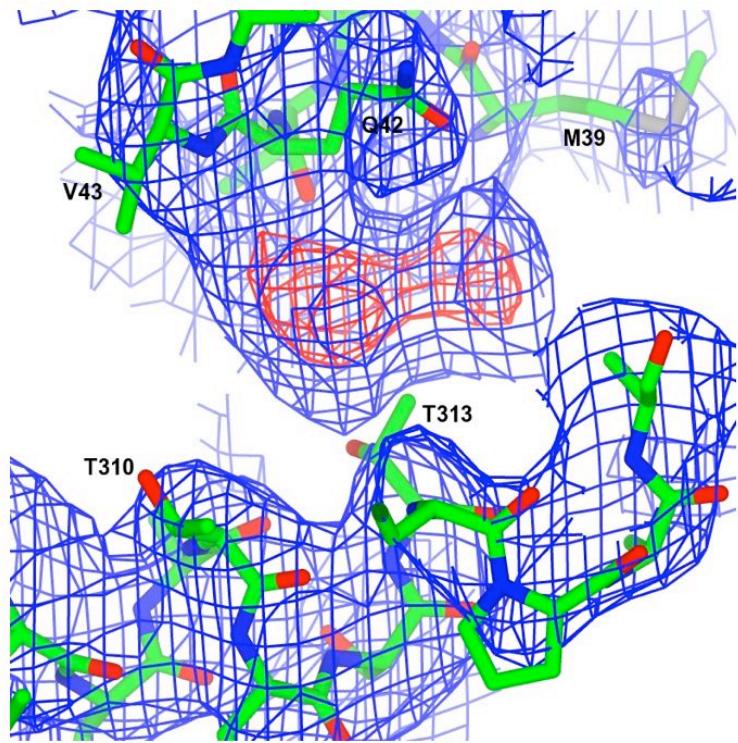


Figure S3. Electron density in the region of the sodium binding site. A strong feature is present in both $2mF_{\text{obs}} - DF_{\text{calc}}$ (blue, contoured at 1σ) and $mF_{\text{obs}} - DF_{\text{calc}}$ (red, contoured at 6σ) maps calculated using the refined structure. The electron density is not consistent with any of the ingredients in the crystallization mixture, nor was it possible to model the protein differently to satisfy the density. The feature was also present in maps calculated using data collected from similar crystals grown using different batches of protein. It seems likely that a ligand has bound during cell growth.

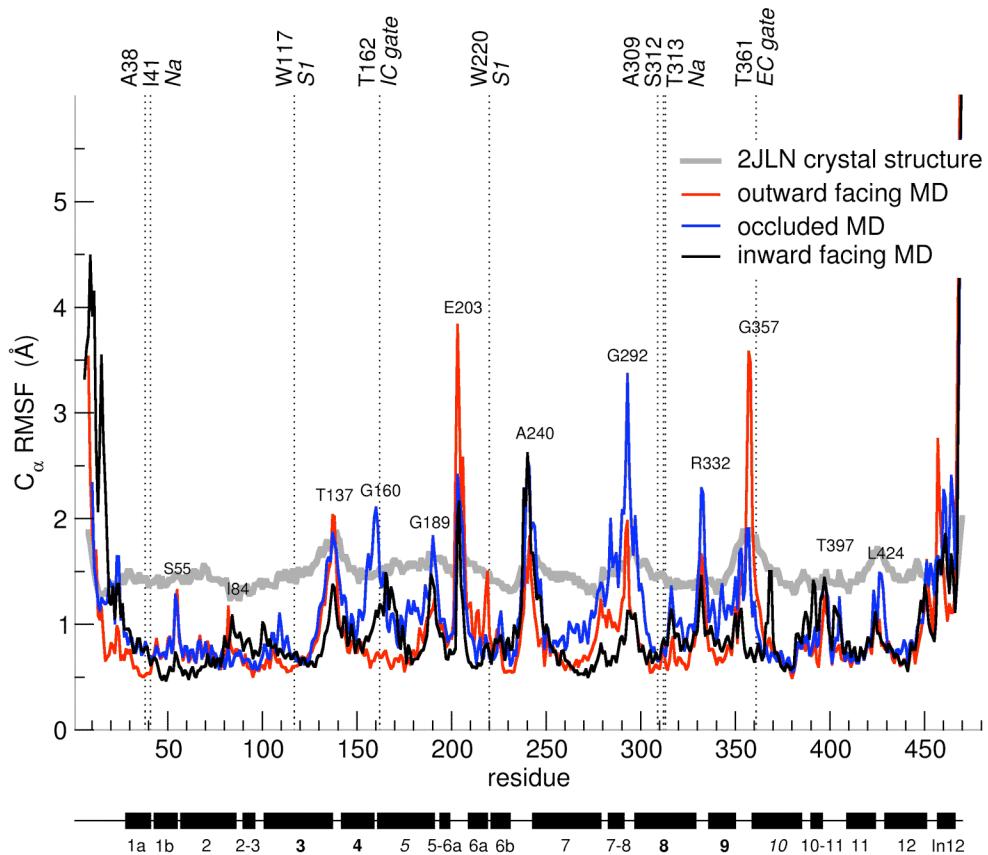


Figure S4. C_α RMSF from refined crystallographic *B*-factors and *apo* equilibrium MD simulations. Secondary structure elements (helices) are indicated below the graph. The thick gate helices (“hash motif”) are bold-faced, the thin gates (“flexible helices”) italicized. Residues involved in substrate binding and gating are indicated by dashed lines. Individual peaks are labelled with the Mhp1 residue. *Holo* simulations (not shown) exhibit similar patterns.

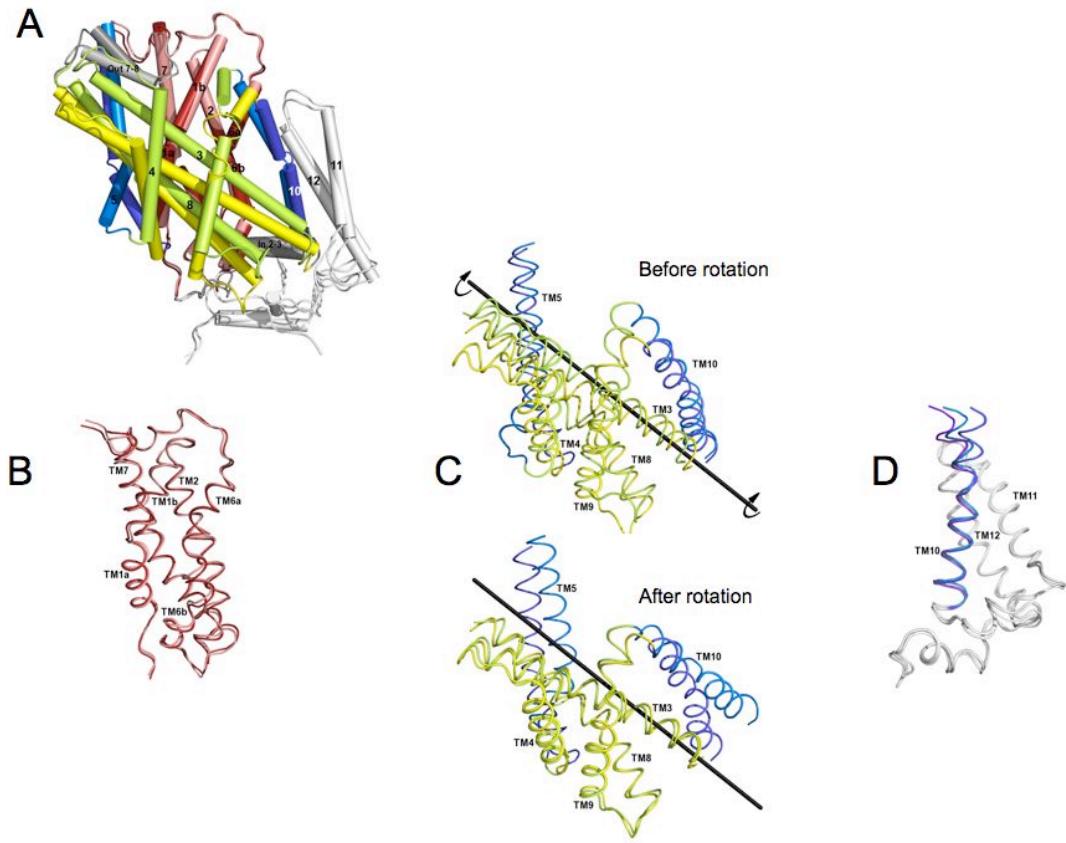


Figure S5. Superpositions of the inward- and outward-facing structures. **(A)** Overall superposition of the two structures (see Methods). The outward-facing structure has been coloured as in Fig 1 with red for the bundle, yellow for the hash motif, and blue for TMs 5 and 10. The inward-facing structure is shown in salmon for the bundle, light green for the hash motif and light blue for TMs 5 and 10. **(B)** Superposition of the bundles of the two structures. **(C)** Superposition of the hash motifs along with TMs 5 and 10: in the upper panel the two structures are shown using the superposition shown in (A); in the lower panel the two motifs have been rotated around the axis shown. It is clear that TMs 3, 4, 8, and 9 (the hash motif) superpose very well but TMs 5 and 10 vary. **(D)** Superposition of the C-terminal TMs. In this view TMs 10, 11 and 12 from the outward-facing (blue), occluded (purple) and inward facing (light blue) structures are shown. The conformation of TM10 in the inward-facing structure is very similar to that in the occluded state.

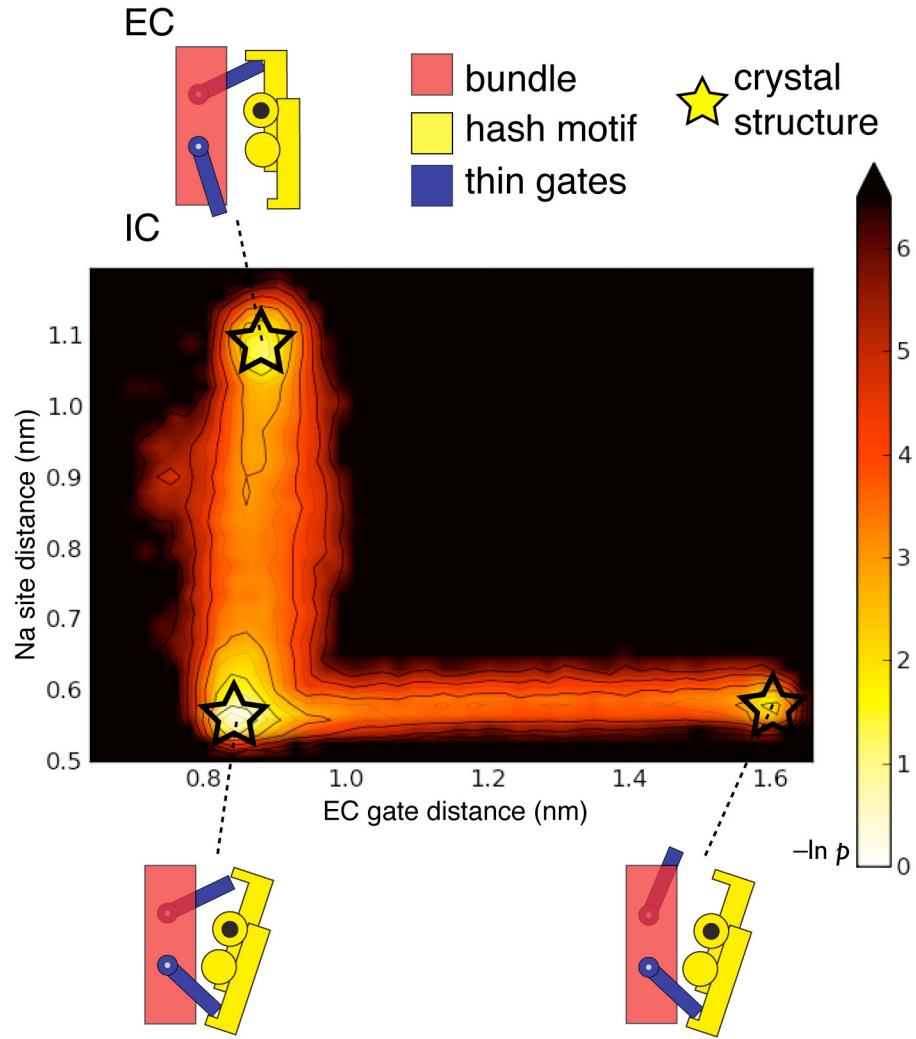


Figure S6. DIMS simulations. Trajectories are projected on the EC gate distance and thick gate (“Na site”) distance. Instead of plotting about 160 trajectories in each direction, the logarithmic density $-\ln p$ of trajectories is shown. White regions contain a high number of conformations whereas black regions are not sampled during the transitions. The values of the crystal structures are indicated by stars, with cartoon models (see Fig. 5 in the article) describing the functional state of the three gating elements. Note that the IC gate distance is not shown here; it moves parallel to thick gate in the DIMS simulations as in the hypothetical “direct” occluded-to-inward facing pathway in Fig. 5.

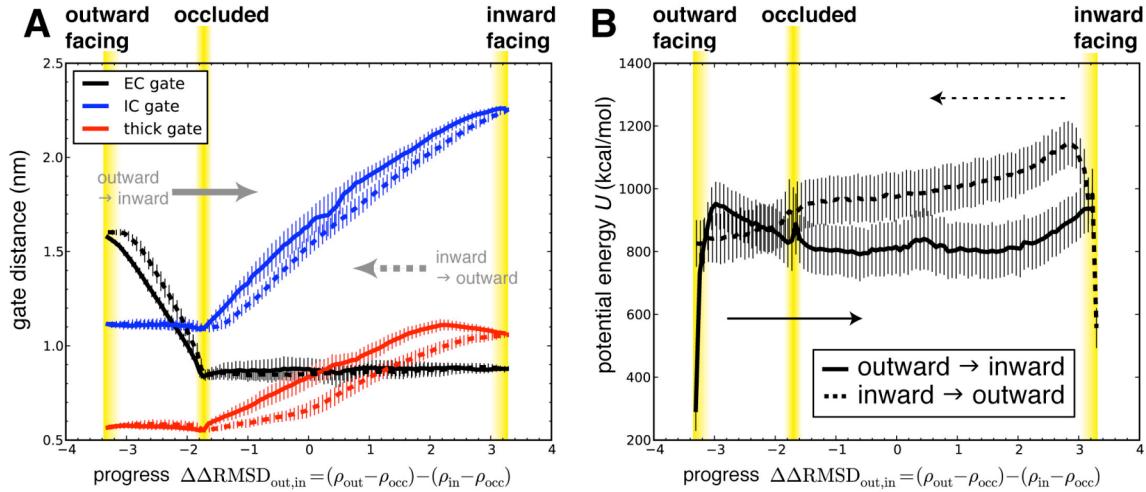


Figure S7. Gate order parameters and potential energy from DIMS-MD generated transitions. Gate distances **(A)** and the force field energy U **(B)** are shown as functions of the structural progress along the transition, measured by $\Delta\Delta\text{RMSD}_{\text{out,in}}$ (see supplementary Methods). Yellow bars indicate the conformations of the crystal structures along the transition. Solid lines show data from transitions starting from the outward facing structure (2JLN), through the occluded state (2JLO), towards the inward facing conformation; these transitions progress from $\Delta\Delta\text{RMSD}_{\text{out,in}} = -3.3 \text{ \AA}$ to $+3.3 \text{ \AA}$ (solid arrows). Dashed lines describe the reverse transition and progress from positive to negative $\Delta\Delta\text{RMSD}_{\text{out,in}}$ -values as indicated by the dashed arrows. Error bars are derived as one standard deviation from the mean (solid/dashed lines) of ~ 160 trajectories. The initial $\sim 400 \text{ kcal/mol}$ increase in U over the first and second data points reflects a $\sim 2 \text{ ps}$ -heating phase from an initial energy minimized (i.e. 0 K-structure) to 300 K; the change in $\Delta\Delta\text{RMSD}_{\text{out,in}}$ is $< 0.1 \text{ \AA}$ over these first 2 ps (and $< 0.05 \text{ \AA}$ in RMSD towards the target conformation).

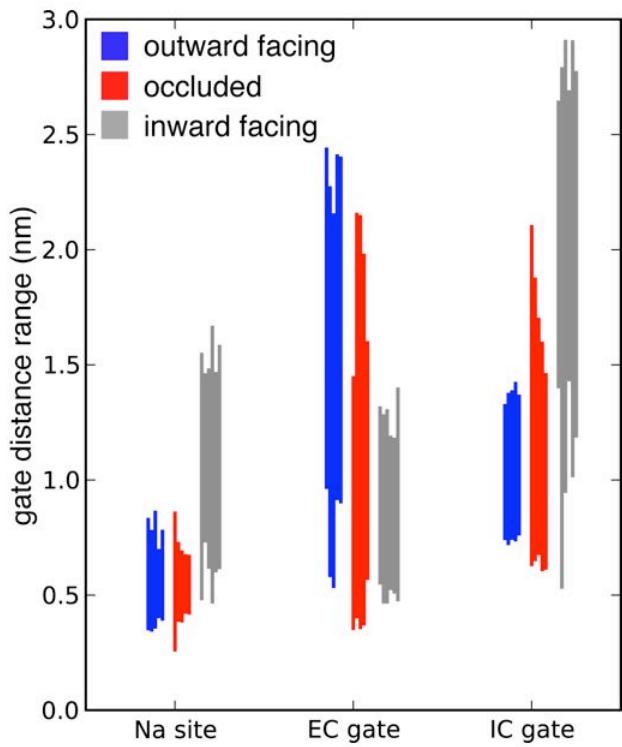


Figure S8. Range of gate distances seen in equilibrium MD simulations. The ranges of values of the three gate distances are shown as bars that stretch from the minimum to the maximum of the distances observed in a single equilibrium MD simulation. Simulations are color-coded depending on the initial conformation. Data are shown for all simulations in Table S2 of length 100 ns or greater. The limited number of simulations prevents us from concluding whether the presence or absence of a sodium ion in the Na site or benzylhydantoin in the S1 site influences the range of gate values; hence, all simulations for a single conformation are shown together. See Methods for a definition of the thick gate (Na site) and thin gate (EC and IC gate) distances.

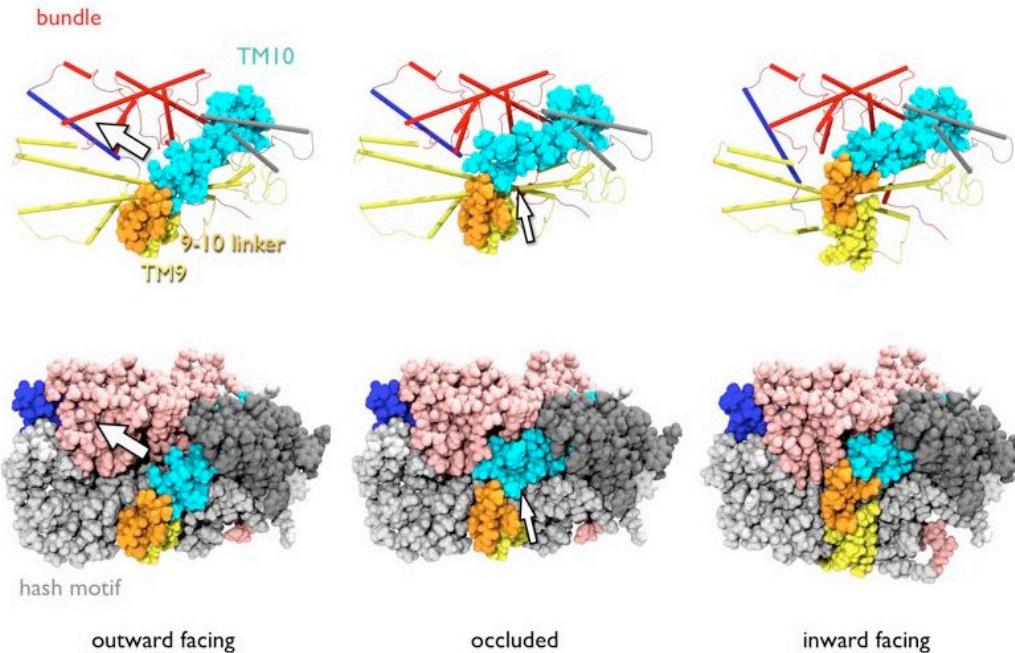


Figure S9. Arrest of the EC gate in the inward facing conformation. The three crystal structures are compared with a focus on the EC gate formed by TM10 (cyan) closing up against the bundle (red/pink). The main closing movement of the EC gate during the outward-facing to occluded transition is indicated with a large arrow for the outward facing conformation. Once the thick gate opens via the rotation of the hash motif (yellow/white), TM9 (yellow) moves upwards towards the EC side and pushes down on TM10 (cyan) via the 9-10 linker (M351-V358, orange) as shown by the smaller arrow for the occluded conformation. The tighter packing and partial burial of the TM10 N-terminus restricts the conformational freedom of TM10 and thus “latches” the EC gate. Top and bottom row show the same views but in the top row TM9 to TM10 are shown in a space-filling representation whereas the rest of Mhp1 is shown as a thin cartoon. The bottom row uses space filling for all atoms in order to illustrate the tighter packing of the EC gate in the inward facing conformation compared to the occluded one. Based on the pseudo twofold-symmetry of Mhp1 and the MD results in Fig. S8, a similar mechanism appears to arrest the IC gate (TM5, blue) in the outward facing/occluded conformation.

References

- S1. S. Weyand *et al.*, *Science* **322**, 709 (2008).
- S2. J. M. Hadden, M. A. Convery, A. C. Declais, D. M. Lilley, S. E. Phillips, *Nat Struct Biol* **8**, 62 (Jan, 2001).
- S3. T. Shimamura *et al.*, *Acta Crystallogr F* **64**, 1172 (Dec 1, 2008).
- S4. Z. Otwinowski, W. Minor, Charles W. Carter, Jr., *Methods in Enzymology* **276**, 307 (1997).
- S5. Collaborative Computational Project Number 4, *Acta Crystallogr D* **50**, 760 (1994).
- S6. A. J. McCoy *et al.*, *J Appl Crystallogr* **40**, 658 (2007).
- S7. T. C. Terwilliger, *Acta Crystallogr D* **56**, 965 (2000).
- S8. T. Terwilliger, *Acta Crystallogr D* **60**, 2144 (2004).
- S9. E. de La Fortelle, G. Bricogne, Charles W. Carter, Jr., *Methods in Enzymology* **276**, 472 (1997).
- S10. T. A. Jones, M. Kjeldgaard, W. C. Charles, Jr., M. S. Robert, *Methods in Enzymology* **277**, 173 (1997).
- S11. P. D. Adams *et al.*, *Acta Crystallogr D* **58**, 1948 (2002).
- S12. M. D. Winn, M. N. Isupov, G. N. Murshudov, *Acta Crystallogr D* **57**, 122 (2001).
- S13. G. J. Kleywegt, T. A. Jones, *ESF/CCP4 Newsletter* **31**, 9 (1994).
- S14. W. L. Delano, *DeLano Scientific, Palo Alto, CA, USA*, (2002).
- S15. W. Humphrey, A. Dalke, K. Schulten, *J Mol Graph* **14**, 33 (1996).
- S16. T. B. Woolf, *Chem. Phys. Lett.* **289**, 433 (1998).
- S17. J. R. Perilla, O. Beckstein, E. J. Denning, T. B. Woolf, *J. Comp. Chem.* submitted, (2009).
- S18. O. Beckstein, E. J. Denning, J. R. Perilla, T. B. Woolf, *J. Mol. Biol.* **394**, 160 (2009).
- S19. M. Schaefer, C. Bartels, F. Leclerc, M. Karplus, *J. Comp. Chem.* **22**, 1857 (2001).
- S20. J. P. Ryckaert, G. Ciccotti, H. J. C. Berendsen, *J Comp Phys* **23**, 327 (1977).
- S21. B. R. Brooks *et al.*, *J Comp Chem* **30**, 1545 (2009).
- S22. A. MacKerell *et al.*, *J Phys Chem B* **102**, 3586 (1998).
- S23. D. M. Zuckerman, T. B. Woolf. (2006), arXiv:physics/0209098.
- S24. K. A. Scott *et al.*, *Structure* **16**, 621 (2008).
- S25. P. J. Bond, C. L. Wee, M. S. P. Sansom, *Biochemistry* **47**, 11321 (2008).
- S26. B. Hess, C. Kutzner, D. van der Spoel, E. Lindahl, *J Chem Theo Comp* **4**, 435

- (2008).
- S27. P. J. Stansfeld, R. Hopkinson, F. M. Ashcroft, M. S. P. Sansom, *Biochemistry* **48**, 10926 (2009).
- S28. W. F. van Gunsteren *et al.* (Biomos and Hochschulverlag AG an der ETH Zürich, Zürich and Groningen, 1996).
- S29. J. Åqvist, *J. Phys. Chem. B* **94**, 8021 (1990).
- S30. O. Berger, O. Edholm, F. Jähnig, *Biophys J* **72**, 2002 (1997).
- S31. J. Hermans, H. J. C. Berendsen, W. F. van Gunsteren, J. P. M. Postma, *Biopolymers* **23**, 1513 (1984).
- S32. A. W. Schüttelkopf, D. M. F. v. Aalten, *Acta Crystallographica D* **60**, 1355 (2004).
- S33. B. Hess, *J Chem Theo Comp* **4**, 116 (2008).
- S34. G. Bussi, D. Donadio, M. Parrinello, *J Chem Phys* **126**, 01410 (2007).
- S35. U. Essman *et al.*, *J Chem Phys* **103**, 8577 (1995).
- S36. H. Li, A. D. Robertson, J. H. Jensen, *Proteins: Struct Funct Genet* **61**, 21 (2005).
- S37. C. R. Raetz, *Microbiol Rev* **42**, 614 (1978).
- S38. S. Suzuki, P. J. F. Henderson, *J Bacteriol* **188**, 3329 (2006).
- S39. S. E. Murdock *et al.*, *J. Chem. Theory Comp.*, 1477 (10, 2006).
- S40. W. L. Ash, M. R. Zlomislic, E. O. Oloo, D. P. Tielemans, *Biochim Biophys Acta* **1666**, 158 (Nov, 2004).
- S41. I. W. Davis *et al.*, *Nucleic Acids Res* **35**, W375 (2007).