Mechanistic Picture for Conformational Transition of a Membrane Transporter at Atomic Resolution (Supporting Information)

Mahmoud Moradi * and **Emad Tajkhorshid** *

* Center for Biophysics and Computational Biology, Department of Biochemistry, College of Medicine, and Beckman Institute for Advanced Science and Technology, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801, U.S.A.

Methods

In the following sections we discuss our novel use of nonequilibrium work relations to compare different transition paths, give a detailed description of the reaction coordinates and collective variables used to design our system-specific biasing protocols, provide all the MD simulation details, and explain all the technical details regarding the methods used to quantify the results.

Nonequilibrium Work Relations. "Nonequilibrium work relations" [1, 2, 3, 4] are powerful tools in extracting information about the equilibrium behavior (*e.g.*, free energies) of a system from its nonequilibrium counterpart, *i.e.*, a "nonequilibrium driven system" that is driven away from an equilibrium state by the variation of an external parameter [5]. Generally, the distribution of nonequilibrium trajectories can be connected to that of the equilibrium trajectories via a statistical weight calculated from the work performed on the system [3, 6, 4, 7, 5] (see Connecting Nonequilibrium and Equilibrium Ensembles). In an MD simulation, a time-dependent biasing potential defined in terms of a collective variable can be used to drive the system from an initial to a final state. Here the collective variable acts as a control parameter, varied according to a protocol. This approach is employed most prominently in methods such as steered [8] and targeted [9] MD in which the biasing potential is a harmonic constraint with a center in the collective variable space, moving linearly in time.

Nonequilibrium work relations have been used in many applications to numerically estimate the free energies based on the nonequilibrium work measurements [10, 11]; however, the use of these relations is not limited to free energy calculations. In principle, nonequilibrium work relations can be used to estimate any equilibrium macroscopic quantity from nonequilibrium driven trajectories [2, 12]. Due to the sampling limitations such generalizations are not necessarily practical but one may find certain quantities whose estimate converge faster than free energy itself. Estimating these quantities could thus be computationally less costly when compared to often prohibitively expensive free energy calculations. As an example, one may estimate the relative transition rates of competing pathways using nonequilibrium work measurements [13]. The relative transition rates of different paths can be used to estimate their relative importance without requiring an accurate estimate of the whole free energy landscape.

In this study, we measure work along different transition pathways, not for a quantitative description of the energetics of the system, but rather to compare the feasibility of different pathways obtained from nonequilibrium simulations [14, 11, 15]. Suppose that there are different hypothetical mechanisms for a transition that can be expressed as distinct transition tubes in a particular collective variable space. If one designs biasing protocols that guide the system via these transition tubes, nonequilibrium work measurements can be used to estimate the relative importance of each tube [13]. Thus, by sampling many different pathways one may establish a reliable understanding of the overall trend of a transition and intermediate structures, without the need for expensive calculation of accurate quantities, such as free energies.

In principle, the probability of each transition tube (which is proportional to its associated transition rate) can be measured accurately given an adequate sampling. However, the presence of distinct work trends between the actual transition tube (which is the dominant one when the system is not biased) and other hypothetical transition tubes (which are disfavored energetically) can simplify the calculations. In order to explore the transition tubes, one may define a relevant set of reaction coordinates to reduce the phase space to a reaction coordinate space with a clear distinction between different states of the system including initial, final, and different hypothetical intermediate states. Let us assume we run several nonequilibrium simulations with different biasing protocols and measure the work values. The optimal transition tube (within the tubes sampled) can be identified if there is a clear difference in the trend of the work between different classes of pathways.

Note that due to the nonequilibrium feature of the simulations, these work profiles are associated with a dissipative term that is stochastic in nature; thus, one cannot make reliable statements based on single trajectories. However, the trend of the work (determined by repeating the simulations) can be used to compare different transition paths/mechanisms. We also note that any parameter involved in the biasing protocol (*e.g.*, simulation time) can influence the trend of the work. One can simplify the comparison by keeping some of these parameters the same in different protocols associated with different paths. Ideally, the protocols should be designed in a way that focuses on one "explanatory" variable to avoid complications in the comparison. Here we consider the trend of the work as a "response" variable; any parameter that is different in the biasing protocols could be generally considered a candidate "explanatory" variable which explains the difference in the trends of the work resulted from different protocols. For instance, if two protocols use two different collective variables that are essentially different (*e.g.*, a distance versus an angle) the comparison will be nontrivial; the different work trends could be due to the way the collective variables are defined (not due to the difference of the paths taken). In order to simplify the problem, we designed our protocols such that only the order of biasing stages is varied in different protocols while everything else is kept the same.

More quantitative analyses such as free energy calculations or accurate transition path optimizations can be generally performed, once a practical biasing protocol is found that does not require a large amount of work when used for inducing the transition. Thus, this study not only sheds light on the mechanistic features of the IF-OF transition of MsbA but also provides a good framework for more expensive/accurate calculations that can be carried out in future studies.

Reaction Coordinates. Here we introduce several reaction coordinates that describe the global conformational features of the MsbA transporter, and are used as intuitive metrics/measures to identify different states. These coordinates are mainly designed as analysis tools to monitor the conformational changes but they are also closely related to the collective variables used in the biasing protocols to impose a certain conformational change by applying forces on the system (see Collective Variables). We generally use the terms, reaction coordinates and collective variables interchangeably in the paper.

We define a three-dimensional space (α, β, γ) that describes the conformation of MsbA in a reduced holonomic coordinate space. α and β are defined on the TMD helices while γ is defined on the NBDs. We define TM_i^{cis} and TM_i^{trans} as the the ith transmembrane helix of the two monomers (labeled *cis* and *trans*, arbitrarily) and consider four relatively rigid bundles, B_1 ($\text{TM}_{1,2}^{cis}$,$\text{TM}_{4,5}^{trans}$ helices), B_2 ($\text{TM}_{1,2}^{trans}$,$\text{TM}_{4,5}^{cis}$ helices), B_3 ($\text{TM}_{3,6}^{cis}$ helices), and B_4 ($\text{TM}_{3,6}^{trans}$ helices), colored in Fig. 1, blue, red, yellow, and green, respectively. α describes the angle between two groups of bundles B_1/B_3 and B_2/B_4. On the other hand β describes the angle between B_1/B_4 and B_2/B_3. In both cases, the C^α atoms of each group was used to find the principal axes and the direction of the roll axis was used to measure the angle between the two groups. Finally, γ was defined as the angle between the roll axes of the two NBDs constructed using the C^α atoms of each NBD. α and β as defined here intuitively describe the opening/closure of the cytoplasmic and periplasmic sides, respectively. γ, on the other hand, describes the relative orientation of the two NBDs.

While the (α, β, γ) space is the main focus of our analysis, more conventional metrics including distance and RMSD were used along with these angles. d_{NBD}, the distance between the C^α mass centers of the two NBDs is an intuitive representative of the NBD dissociation/dimerization. $RMSD_{OF}$, $RMSD_{IF-c}$, and $RMSD_{IF-o}$ are C^α RMSDs from the crystal structures of OF, IF-c, and IF-o, respectively.

Collective Variables. In order to induce a conformational transition, it is often relevant to define a set of collective variables that describe different states of the system, such that by applying appropriate forces on the system, one can vary these collective variables and change the conformation of the system. One particular collective variable that is widely used in the context of structural transition of proteins is the RMSD. Although using the RMSD from a target structure as a collective variable (*e.g.*, in a targeted MD simulation) has proved useful, the method has its own pitfalls and limitations. Targeted MD requires a target structure whose quality is a determining factor in the reliability of the results. This reduces the flexibility of the method to a great extent. Ironically, RMSD is associated with both extreme degeneracy and large entropy loss (for large and small values of RMSD). The trajectory generated by targeted MD represents a pathway along which the RMSD decreases almost monotonically and nearly linearly. Due to such reasons, a targeted MD simulation typically requires a large amount of work to induce a transition, thus making the interpretation of its results difficult in the context of nonequilibrium work relations.

Other conventional collective variables such as distance and radius of gyration have their own limitations that make them impractical for inducing a global large-scale structural transition such as the IF↔OF conformational change in ABC transporters. One particular collective variable that seems to best reflect the nature of the conformational changes of MsbA intuitively is angles such α, β, and γ as defined above. The simplest way to define a collective variable associated with one of these angles is to use the mass centers of three groups of atoms. Unfortunately, this simple definition often results in structural deformation of the protein. Another approach is to use the principal axes as used in defining the measures above. However, in order to define a practical set of collective variables to induce the desired transitions associated with the global angles such as α, β, and γ, here we use the "orientation quaternions" [16, 17, 18].

The orientation quaternion [17], often used for "optimal superposition" in computational biology [16], is a tool to deal with the so-called "absolute orientation" problem. Suppose for a set of N atoms (labeled $1 \leq k \leq N$), we have two different sets of measurements: $\{\mathbf{x}_k\}$ and $\{\mathbf{y}_k\}$. To simplify the problem we assume both sets have been already shifted to bring their barycenters to the origin (optimum translation). To find the optimum rotation to superimpose $\{\mathbf{y}_k\}$ on $\{\mathbf{x}_k\}$, we introduce "pure quaternions" x_k and y_k whose vector parts are $\mathbf{x}_k$ and $\mathbf{y}_k$, respectively. A quaternion can be thought of as a vec-

tor with four components, as a composite of a scalar and an ordinary vector, or as a complex number with three different imaginary parts. A quaternion whose scalar part is zero is called pure (reminiscent of pure imaginary numbers). The optimal rotation can be parametrized by a unit quaternion, $\hat{q}$ that minimizes $\left\langle \|\hat{q} x_k \hat{q}^* - y_k\|^2 \right\rangle$ in which $\langle . \rangle$ denotes an average over k, q^* is the conjugate of q, and $\|q\|^2 \equiv qq^*$ (see Ref. [16] for more details). The optimal rotation unit quaternion (or orientation quaternion) $\hat{q}$ can be written as $(\cos\frac{\theta}{2}, \sin\frac{\theta}{2}\hat{u})$ in which θ and $\hat{u}$ (a unit vector) are the optimum angle and axis of rotation, respectively.

As a collective variable, orientation quaternion can be used not only to monitor the rotational changes but also to apply forces (that are proportional to the derivatives of the orientation quaternions) on the system in a practical way to induce the desired rotational changes. Suppose that we are interested in inducing a particular rotation – given by its axis of rotation (unit vector $\hat{u}$) and its target angle of rotation θ_{target} – on a particular segment of a biomolecule (*e.g.*, part of a helix, a helix, or a bundle of helices). One simple way is to use a time-dependent harmonic potential (similar to steered MD in spirit) [18]:

$$U_B(q_{ref}(\{\mathbf{x}_k\}), t) = \frac{1}{2}k\Omega^2(q_{ref}(\{\mathbf{x}_k\}), Q(t)). \qquad \textbf{[1]}$$

Here $q_{ref}(\{\mathbf{x}_k\})$ is the optimum orientation quaternion to superimpose $\{\mathbf{x}_k\}$ on a reference set $\{\mathbf{x}_k^{ref}\}$. The reference could be the initial, target, or any other structure; Here to simplify the notations we assume the reference is the same as the initial structure. $Q(t) \equiv (\cos(\frac{\theta(t)}{2}), \sin(\frac{\theta(t)}{2})\hat{u})$ is a unit quaternion that is varied externally, providing the center of our harmonic potential at time t during a simulation ($0 \leq t \leq T$). If the reference is the same as the initial structure, $\theta(0)$ and $\theta(T)$ can be set to 0 and θ_{target}, respectively. Now that we have $Q(0)$ and $Q(T)$ we can use different interpolation methods to determine $Q(t)$. A simple method is varying $\theta(t)$ linearly that is a special case of spherical linear interpolation (Slerp) method [19]. The particular method used here (*i.e.*, NAMD implementation [18]), is based on the linear interpolation of the quaternion $Q(t)$ (based on the current and final targets) followed by its normalization at each *timestep* (see Nonequilibrium Work Measurements). Finally $\Omega(\hat{p}, \hat{q})$ is the length of the geodesic between two points on the unit sphere, transformed by $\hat{p}$ and $\hat{q}$ from an arbitrary point on the unit sphere. One can show $\cos(\Omega(\hat{p}, \hat{q})) = \hat{p} \cdot \hat{q}$.

We use six orientation quaternions as collective variables to describe the three-dimensional (α, β, γ) space including $(q_\alpha^{cis}, q_\alpha^{trans}, q_\beta^{cis}, q_\beta^{trans}, q_\gamma^{cis}, q_\gamma^{trans})$. For each angle two collective variables are defined on the two sides of the angle. q_α^{cis} and q_α^{trans} are the orientations of the two groups of bundles B_1/B_3 and B_2/B_4, respectively. Similarly, q_β^{cis} and q_β^{trans} are the orientations of B_1/B_4 and B_2/B_3, respectively, while $q_\gamma^{cis/trans}$ is the orientation of $\text{NBD}^{cis/trans}$. For the quaternion-based collective variables, the harmonic constant was set to 10^5 kcal/mol. Along with these collective variables, we also used $RMSD_{IF-o}$ (for targeted MD simulations) and d_{NBD} (see Reaction Coordinates) with a harmonic constant of 40 and 100 $\frac{kcal}{mol \cdot \text{Å}^2}$, respectively.

System-Specific Biasing Protocols. We used different combinations of collective variables associated with α, β, and γ reaction coordinates to induce the OF→IF transition. In addition, d_{NBD} was used to dissociate NBDs in some cases. One can generate many distinct biasing protocols by using these collective variables. Let us assume that main changes in α, β, and γ occur in 3 discrete stages. Here the presumptive explanatory variable is the order of events, resulting in six possible classes of protocols: $(\alpha \to \beta \to \gamma)$, $(\alpha \to \gamma \to \beta)$, $(\beta \to \alpha \to \gamma)$, $(\beta \to \gamma \to \alpha)$, $(\gamma \to \alpha \to \beta)$, and $(\gamma \to \beta \to \alpha)$. One may add d_{NBD} as an additional collective variable and an additional stage in the biasing protocol. We use d_{NBD} only for an initial dissociation of NBDs (going from about 30 to 40 Å: d_{NBD} associated

with OF and IF-c states, respectively). Targeting a larger distance is practically problematic since results in the deformation of the system (unless accompanied by other biases). Our results show that varying γ and/or α results in the NBD dissociation thus varying d_{NBD} for an initial dissociation of NBDs was not used after the stages that involved changing γ and/or α. One can show the total number of protocols satisfying these criteria with four stages of varying α, β, γ, and d_{NBD} in different orders is 8. Figures 3A and S5 show the work profiles and the trajectories resulted from all the 14 protocols described above including 6 three-stage and 8 four-stage $(\alpha, \beta, \gamma, d_{NBD})$ based protocols, each carried out for 160 ns in total. In three- and four-stage simulations, each stage was carried out for 40 ns except for the γ-based stages not preceded by pushing d_{NBD} that were carried out for 80 ns (since changing γ induces the change in d_{NBD} as well).

Connecting Nonequilibrium and Equilibrium Ensembles. Consider a system governed in equilibrium by Hamiltonian $H(\mathbf{x})$ in which $\mathbf{x}$ is a point in the phase space. Suppose $\mathbf{x}^t$ is a trajectory of this system driven out of equilibrium over a time interval $[0, T]$ using a time-dependent biasing potential $U(\mathbf{x}^t, t)$ added to the Hamiltonian. A common choice for biasing potential is $U(\mathbf{x}^t, t) = U(\xi^t, t) = U_{\xi(t)}(\xi^t)$ in which ξ is a collective variable, ξ^t is the projection of $\mathbf{x}^t$ onto the ξ space, and $\xi(t)$ is a parameter controlled externally. Although there is no particular restriction, but for simplicity $U_{\xi(t)}(\xi^t)$ and $\xi(t)$ are often chosen to be harmonic (*e.g.*, $k(\xi^t - \xi(t))^2/2)$ and linear, respectively. A statistical ensemble of trajectories $\mathbf{x}_1^t, \mathbf{x}_2^t, \ldots$ is described by its time-dependent distribution in phase space $f^{dr}(\mathbf{x}, t) = \langle \delta(\mathbf{x} - \mathbf{x}^t) \rangle^{dr}$, in which $\langle . \rangle^{dr}$ denotes an ensemble average over the driven trajectories and $f^{dr}(\mathbf{x}, t)$ gives the probability of finding the system at microstate $\mathbf{x}$ at time t. If the process starts from equilibrium state A, biased only by $U_{\xi_A}(\xi^t)$ with $\xi(0) = \xi_A$, one can show [3] the distribution of states associated with ξ in equilibrium may be represented by the driven ensemble in which each trajectory $\mathbf{x}^t$ carries a time-dependent statistical weight $\exp(-\beta w_d^t)$, *i.e.*,

$$p_\xi^{eq}(\mathbf{x}) = \left\langle \delta(\mathbf{x} - \mathbf{x}^t) \right\rangle_\xi^{eq} = \left\langle \delta(\mathbf{x} - \mathbf{x}^t) \exp(-\beta w_d^t) \right\rangle^{dr}, \quad [2]$$

in which β is the inverse temperature, $w_d^t = w^t - (\mathcal{F}(\xi) - \mathcal{F}(\xi_A))$ is the dissipative work, $w^t = \int_0^t \frac{\partial}{\partial t'} U(\xi^{t'}, t') \, dt'$ is the total nonequilibrium (accumulative) work over $[0, t]$ time interval, $\langle . \rangle_\xi^{eq}$ describes an average over an ensemble of equilibrium trajectories governed by $H + U_\xi$, and $\mathcal{F}(\xi)$ is the free energy of the *perturbed* system. In the stiff-spring limit (large force constant) [20] the potential of mean force (PMF) of the *unperturbed* system, $F(\xi)$, can be approximated as the free energy of the *perturbed* system, $\mathcal{F}(\xi)$, but in a more general case, $F(\xi)$ can be reconstructed via [3, 6]:

$$F(\xi) = F(\xi_A) - \beta^{-1} \log \left\langle \delta(\xi - \xi^t) \exp(-\beta \Delta w^t) \right\rangle^{dr}, \quad [3]$$

in which $\Delta w^t = w^t - U(\xi^t, t)$. More generally, the *unperturbed* equilibrium distribution of $\mathbf{x}$ can be connected to the nonequilibrium driven trajectories via:

$$p^{eq}(\mathbf{x}) \propto \left\langle \delta(\mathbf{x} - \mathbf{x}^t) \exp(-\beta \Delta w^t) \right\rangle^{dr}. \quad [4]$$

Now if λ is a collective variable defined on $\mathbf{x}$ that may or may not be the same as ξ, one can write:

$$\exp(-\beta \Delta F(\lambda)) = \left\langle \delta(\lambda - \lambda^t) \exp(-\beta \Delta w^t) \right\rangle^{dr}. \quad [5]$$

in which $\Delta F(\lambda) = F(\lambda) - F(0)$.

Estimating the free energy for a particular λ requires sampling an adequate number of trajectories that visit λ particularly those associated with a small amount of work that dominate the RHS of (5). This is not always feasible, particularly when large-scale transitions are

targeted. Typically, work values in such simulations (*e.g.*, a targeted MD simulation on a membrane transporter) are on the order of hundreds of $kcal/mol$, much greater than the actual free energies. One can try to use longer simulations to decrease the dissipation; however, if the path taken is not close to the minimum free energy path, the results will never converge to the relevant free energy values.

Although accurate free energy calculations are not always feasible due to the practical reasons, one may make qualitative statements about the transition paths based on nonequilibrium work measurements. The relation (5) describes how a particular driven ensemble – with a particular biasing protocol – is related to the equilibrium ensemble. Here for the main conclusions of the paper we have used a rough comparison of the work values. Fortunately, there is a clear difference in the trend of work between different classes of pathways and the optimum transition tube can be identified.

Nonequilibrium Work Measurements. The accumulative work at time t along a trajectory generated by the quaternion-based biasing protocol (1) can be measured via,

$$w^t = \int_0^t \frac{\partial}{\partial t'} U_B(q_{ref}(\{\mathbf{x}_k\}), t') \, dt'. \quad [6]$$

One can collect the biasing potential ($U_B(q_{ref}, t)$) and its partial time derivative ($\partial U_B(q_{ref}, t)/\partial t$) based on the instantaneous q_{ref} at time t. For the particular q_{ref} schedule, $Q(t)$ that comes from the linear interpolation of the quaternion ($Q'(t + \Delta t) = Q(t) + \frac{Q(T) - Q(t)}{T - t} \Delta t$) followed by its normalization ($Q(t + \Delta t) = \frac{Q'(t + \Delta t)}{||Q'(t + \Delta t)||}$), one can show:

$$\frac{\partial}{\partial t} U_B(q_{ref}, t) = k \, \Omega(q_{ref}, Q(t)) \frac{\partial}{\partial t} \Omega(q_{ref}, Q(t)), \quad [7]$$

$$\frac{\partial}{\partial t} U_B(q_{ref}, t) = -\frac{\Omega}{\sin(\Omega)} (q - Q(t) \cos(\Omega)) \cdot \frac{Q(T) - Q(t)}{T - t}. \quad [8]$$

We collected $\partial U_B/\partial t$ every 0.2 picosecond and estimated the accumulative work at each time t from the relation (6). The work profiles plotted against time show $\Delta w^t - U_B(0) = w^t - (U_B(t) - U_B(0))$ whose physical meaning was discussed above. This is the so-called transferable work [3]. The work profiles reported in terms of the reaction coordinates is also the transferable work associated with a bin in the reaction coordinate space averaged over all the observations in which the bin has been visited using the generalized implementation of the weighted histogram method described in the reference [3]. Note that this algorithm is designed to reconstruct the free energies but what we obtain includes a dissipative term.

Free Energy Calculations. For free energy calculations, we employed umbrella sampling (US) [21] in conjunction with a replica-exchange scheme [22, 23], termed here bias-exchange umbrella sampling (BEUS) (also known as window-exchange or replica-exchange umbrella sampling [23, 24, 25]), to efficiently sample a continuous portion of the phase space along a reaction coordinate.

US [21] combined with the weighted histogram analysis method (WHAM) [26] is a standard free energy calculation scheme for reconstructing the PMF along a given reaction coordinate. By biasing the system using a known potential (*e.g.*, harmonic), one may sample high-energy states, allowing for an accurate reconstruction of free energy landscape of the unbiased system when used along a reweighting scheme such as WHAM. Employing the method to large-scale transitions is often challenging and simple biasing protocols (*e.g.*, using RMSD from a target structure as the reaction coordinate) usually produce unreliable estimates for free energies. By using system-specific reaction coordinates and sampling around relatively reliable transition pathways (obtained using methods discussed above), one may significantly improve the sampling of the regions of the phase space relevant to a transition of interest. Replica-exchange MD [22, 23] is a Monte Carlo algorithm that couples multiple MD simulations in

order to enhance the sampling. Each replica is associated with a different value of a given property whose periodic exchange between the replicas based on an "exchange rule" accelerates the exploration of the phase space. Temperature is the most typical property to exchange between the replicas which accelerates the sampling of all degrees of freedom somewhat blindly. An alternative is to exchange (time-dependent [27] or time-independent [24]) biasing potentials in a "bias-exchange" scheme to specifically accelerate the sampling of the degrees of freedom most relevant to a transition of interest.

Integrating US into the replica-exchange scheme results in an exceptionally practical enhanced sampling approach that allows for an accurate reconstruction of rugged free energy landscapes [23, 25]. The mixing of the replicas in the bias-exchange method guarantees the continuity of the conformational space sampled, yielding a more reliable free energy estimate. Note that due to the presence of a large number of degrees of freedom in a large protein system such as membrane transporters, it is virtually impossible to sample a continuous conformational space if the simulations were to run independently as in a conventional US scheme.

The efficiency of the BEUS simulations (in terms of sampling) relies on the definition of collective variables and the distribution of the replicas in the collective variable space. The choice of the collective variable can be improved in the procedure discussed in detail above (*i.e.*, using nonequilibrium simulations and fine-tuning the biasing protocol to lower the amount of nonequilibrium work). In order to optimize the sampling, the distribution of the replicas in the collective variable space can be adjusted iteratively using short runs to result in a roughly similar rate of exchange between all neighboring replicas.

Molecular Dynamics Simulation Details. The initial model used for all the MD simulations is based on the crystal structure of the *salmonella typhimurium* MsbA in its OF conformation (PDB entry: 3B60) [28]. The unresolved N- and C-terminal residues (M1-T9 and Q582) were not modeled, and the nucleotides were removed to generate a nucleotide-free *apo* state.

All MD simulations were performed using NAMD 2.8 and NAMD 2.9 [29]. The CHARMM27 force field [30, 31, 32], including the ϕ/ψ cross-term map (CMAP) correction for the proteins [30] was used for all the simulations. Water molecules were described with the TIP3P model [33]. The protein was energy-minimized *in vacuo* for 3000 steps using conjugate gradient algorithm [34].

Simulations were carried out using a 2 fs timestep at 310 K constant temperature using Langevin dynamics with a damping coefficient γ of $0.5\,\mathrm{ps}^{-1}$. The pressure along the membrane normal (the z-axis of the simulation system) was maintained at 1 atm using the Nosé-Hoover Langevin piston method [35, 36], with a constant cross-sectional area imposed on the xy-plane unless specified otherwise. The smoothed cutoff distance for non-bonded interactions was set to $10-12\,\text{Å}$, and long-range electrostatic interactions were computed with the particle mesh Ewald (PME) method [37].

The protein was embedded in a lipid bilayer consisting of 470 POPC molecules (237 and 233 lipids in the periplasmic and the cytoplasmic leaflets, respectively), and solvated in a periodic TIP3P [33] water box with 100 mM of NaCl, resulting in a simulation system of $\sim$250,000 atoms, with approximate dimensions of $140 \times 140 \times 150\,\text{Å}^3$ before equilibration.

The relaxation of the system started with the acyl chains of the lipid molecules under constant volume conditions for 0.5 ns, with all other atoms fixed. The system was then further equilibrated with all protein atoms, all protein heavy atoms, and all protein C^α atoms restrained ($k=5\,\mathrm{kcal/mol \cdot \text{Å}^2}$) in 0.5, 1.5, and 3 ns runs, respectively, followed by a 5 ns unrestrained simulation, all in the constant-pressure (NPT) conditions, in order to allow the lipid molecules to pack against the protein surface, and for the area of the lipid bilayer to adjust accordingly. Once the system area stabilized, a 5 ns relaxation run was performed (Table S1: Simulation 0) under constant area and normal pressure conditions (1 atm; $\mathrm{NP}_n\mathrm{AT}$ ensemble). All the production

runs including the unbiased and biased simulations used the $\mathrm{NP}_n\mathrm{AT}$ conditions.

First we performed a 150-ns unbiased equilibrium simulation (Table S1: Simulations 1 and 2). We used three structures from t=0, 75, and 150 ns of this equilibrium trajectory (Table S1: Conformations 0, 1, and 2, respectively), to initiate several nonequilibrium driven MD simulations that were carried out using different time-dependent biasing protocols in which the system was driven away from the initial OF state toward an IF state. These protocols include conventional steered and targeted MD simulations (Table S2: Simulations 33-37 and 183-187, respectively) as well as non-conventional protocols (see Tables S2-S6) that use different combinations of collective variables (see Reaction Coordinates and Collective Variables).

Select number of these nonequilibrium simulations were followed by restrained MD (RMD) simulations in which the system is subject to a time-independent biasing potential centered at the final target (Table S3: Simulations 188, 190, and 192; Table S4: Simulations 193, 195, 197, and 199; and Table S6: Simulations 234-238 and 242-246). Select number of the conformations resulted from the biased simulations were further equilibrated with no bias (Table S1: Simulations 247-251). We also performed BEUS MD simulations (Table S7) to quantify the free energies associated with different IF conformations (see Sampling Protocol for Free Energy Calculations). Collectively, we have performed more than 5 μs of unbiased and biased simulations (0.545 and 4.803 μs, respectively). For a complete list of these simulations, see Tables S1-S6.

Sampling Protocol for Free Energy Calculations. Prior to production runs for BEUS MD simulations we used the following protocol to prepare the initial conformations and umbrella potentials:

1. 22 initial conformations were taken from the last stage of the optimized pathway in which the system is pushed along α (*i.e.*, Table S4: Simulation 198).
2. Based on each conformation i (selected above), two quaternion-based collective variables $q_{\alpha,i}^{cis}$ and $q_{\alpha,i}^{trans}$ were defined (conformation i used as the reference structure; for the definition of $q_\alpha^{cis/trans}$ see Collective Variables). For each umbrella $i = 1, \ldots, 22$, a biasing potential (or umbrella potential) was designed with two harmonic terms restraining $q_{\alpha,i}^{cis}$ and $q_{\alpha,i}^{trans}$ around the unity quaternion $1 \equiv (1, 0, 0, 0)$ with harmonic constant k_i (see Collective Variables). The biasing potential for umbrella i can be simplified to: $U_i = \frac{k_i}{2}((\frac{\theta_i^{cis}}{2})^2 + (\frac{\theta_i^{trans}}{2})^2)$ in which $\theta_i^{cis/trans} \equiv 2\cos^{-1}(q_{\alpha,i}^{cis/trans}.1)$ is the angular deviation from the reference i in cis/trans wing of α.
3. BEUS MD simulations were performed for 10-100 ps (for each replica) starting with the initial conformations and using the umbrella potentials obtained from (1) and (2), respectively. An exchange between any two neighboring replicas was attempted every 1 ps.
4. The steps (1) to (3) were iterated with different initial conformations and harmonic constants (identified by trial and error) until (i) the exchange rate between any two neighboring replicas was estimated to be in the 20-40% range, and (ii) the α space (in a given continuous range) was expected to be sampled without any gap.

Production runs: The initial conformations and umbrella potentials satisfying the criteria above were used to perform 24 ns of BEUS MD simulations (total simulation time $22 \times 24 = 528\,ns$).

The conformations used for the production runs are associated with the α's ranging roughly from $13°$ to $49°$ (shown in Fig. S12B,C). See Table S7 for the centers and force constants and average exchange rates based on the entire simulations. Figure S12A,B shows how each replica has covered a large portion of the α space. Note that Fig. S12A show select trajectories in $(\theta_1^{cis}, \theta_1^{trans})$ space in which $\theta_1^{cis/trans}$

represents the angular deviation of α angle in its cis/trans wing from the initial conformation of replica 1 (with $\alpha \approx 13°$).

Reweighting Scheme for Free Energy Reconstructions. The 22 trajectories of BEUS MD simulations were collected every 4 ps. The first 4 ns of each trajectory was discarded (as equilibration phase) and the rest were used to generate an ensemble of 11×10^4 (22 replicas $\times$ 20 ns / 4 ps) configurations $\{\mathbf{X}_i^t\}$ grouped according to the umbrella potential based on which the system was biased ($i = 1, \dots, 22$) and indexed arbitrarily within its umbrella group ($t = 1, \dots, N_i$ and $N_i = 5000$ for all i). The reweighting scheme used is a method originally proposed in Ref. [38] which can be considered to be a generalization of weighted histogram analysis method [26] (with bin size $\rightarrow 0$). This scheme is closely related to multistate Bennet acceptance ratio method (MBAR) [39]. The weight of each configuration $\mathbf{X}_i^t$, p_i^t is determined by self-consistently solving the equations [38]:

$$\begin{cases} 1/p_i^t = \sum_j N_j f_j \exp(-\beta U_j(\mathbf{X}_i^t)), \\ 1/f_j = \sum_i \sum_{t=1}^{N_i} p_i^t \exp(-\beta U_j(\mathbf{X}_i^t)). \end{cases} \qquad \textbf{[9]}$$

in which $U_i(\mathbf{X})$ is the biasing potential for configuration $\mathbf{X}$ according to umbrella i.

The samples were reweighted according to normalized p_i^t values to reconstruct the PMF in terms of a given reaction coordinate ζ (including α, $RMSD_{IF-c}$, and $RMSD_{IF_o}$). The kernel density estimation [40] method was used to reconstruct the unbiased probabilities $p(\zeta)$ based on which the PMF (*i.e.*, $-\beta^{-1} \log(p(\zeta))$) was estimated. A Gaussian kernel was used with a bandwidth selected according to the least-squares cross validation criterion [41]. The bandwidth selection was independently repeated for each reaction coordinate.

A Bayesian bootstrapping technique was used to estimate the statistical error associated with the PMFs [42]. Each trajectory was partitioned into five 4-ns pieces forming $M = 22 \times 5 = 110$ groups of samples. These groups were assumed to be independent and identically distributed (i.i.d) data points with assigned weights $w_l = \xi_l - \xi_{l-1}$ ($l = 1, \dots, M$ and $\{\xi_l\}$ is a set of low-to-high ordered random numbers on $[0, 1]$ with $\xi_0 = 0$ and $\xi_M = 1$). Each configuration $\mathbf{X}_i^t$ takes a weight $\omega_i^t = w_l M$ (l is determined by i and t) prior to be plugged in the equations (9). This can be done implicitly by modifying the first equation in (9):

$$\omega_i^t / p_i^t = \sum_j N_j' f_j \exp(-\beta U_j(\mathbf{X}_i^t)). \qquad \textbf{[10]}$$

in which $N_j' = \sum_{t=1}^{N_j} \omega_j^t$.

The normalized probabilities $\{p_i^t\}$ can be used to estimate the PMFs as discussed above. The procedure was repeated 100 times with different random values $\{\xi_l\}$ to generate a set of PMFs in terms of ζ whose average and standard deviation at each ζ was used as an estimate of free energy and the associated error. Figures 5 and S13A,B show the average PMFs and their associated errors in terms of α, $RMSD_{IF-c}$, and $RMSD_{IF_o}$ reaction coordinates, respectively, while 100 bootstrapped PMFs in terms of α are shown in Fig. S12.

Analysis and Plotting. The trajectories were analyzed using the VMD [43] and the ProDy packages [44]. All the projected trajectories (excluding Figs. 2B and 10, and Figs. S2 and S4) and work profiles are smoothed using a rational Bezier curve [45] while the time series in Figs. S9C-D and S10, and S2, are smoothed using a running average. The PMFs in Figs. 5, S12, and S13 are generated using a kernel density estimation method as discussed above. The protein structures were plotted using VMD [43] v. 1.9.

1. Jarzynski C (1997) Nonequilibrium equality for free energy differences. *Phys Rev Lett* 78:2690–2693.
2. Crooks G E (2000) Path-ensemble averages in systems driven far from equilibrium. *Phys Rev E* 61:2361–2366.
3. Hummer G, Szabo A (2001) Free energy reconstruction from nonequilibrium single-molecule pulling experiments. *Proc Natl Acad Sci USA* 98:3658–3661.
4. Jarzynski C (2001) How does a system respond when driven away from thermal equilibrium? *Proc Natl Acad Sci USA* 98:3636 – 3638.
5. Jarzynski C (2008) Nonequilibrium work relations: foundations and applications. *Eur Phys J B* 64:331 – 340.
6. Hummer G, Szabo A (2005) Free energy surfaces from single-molecule force spectroscopy. *Acc Chem Res* 38:504 – 513.
7. Jarzynski C (2002) What is the microscopic response of a system driven far from equilibrium? In *Dynamics of Dissipation (Lecture Notes in Physics)* (Springer), pages 63 – 82.
8. Izrailev S, Stepaniants S, Balsera M, Oono Y, Schulten K (1997) Molecular dynamics study of unbinding of the avidin-biotin complex. *Biophys J* 72:1568–1581.
9. Schlitter J, Engels M, Krüger P, Jacoby E, Wollmer A (1993) Targeted molecular dynamics simulation of conformational change — application to the T ↔ R transition in insulin. *Mol Sim* 10:291–308.
10. Jensen M Ø, Park S, Tajkhorshid E, Schulten K (2002) Energetics of glycerol conduction through aquaglyceroporin GlpF. *Proc Natl Acad Sci USA* 99:6731–6736.
11. Moradi M, Babin V, Roland C, Sagui C (2010) A classical molecular dynamics investigation of the free energy and structure of short polyproline conformers. *J Chem Phys* 133:125104.
12. Minh D D L, Chodera J D (2011) Estimating equilibrium ensemble averages using multiple time slices from driven nonequilibrium processes: Theory and application to free energies, moments, and thermodynamic length in single-molecule pulling experiments. *J Chem Phys* 134:024111–024119.
13. Moradi M, Sagui C, Roland C (2011) Calculating relative transition rates with driven nonequilibrium simulations. *Chem Phys Lett* 518:109–113.
14. Moradi M, Babin V, Roland C, Darden T, Sagui C (2009) Conformations and free energy landscapes of polyproline peptides. *Proc Natl Acad Sci USA* 106:20746.
15. Moradi M, Babin V, Roland C, Sagui C (2013) Reaction path ensemble of the B−Z−DNA transition: a comprehensive atomistic study. *Nucleic Acids Res* 41:33–43.
16. Coutsias E A, Seok C, Dill K A (2004) Using quaternions to calculate rmsd. *J Comp Chem* 25:1849–1857.
17. Horn B K P (1987) Closed-form solution of absolute orientation using unit quaternions. *J Opt Soc Am A* 4:629–642.
18. Fiorin G, Klein M L, Hénin J (2013) Using collective variables to drive molecular dynamics simulations. *Molecular Physics* doi:10.1080/00268976.2013.813594.
19. Shoemake K (1985) Animating rotation with quaternion curves. *SIGGRAPH Comput Graph* 19:245–254. ISSN 0097-8930.
20. Park S, Khalili-Araghi F, Tajkhorshid E, Schulten K (2003) Free energy calculation from steered molecular dynamics simulations using Jarzynski's equality. *J Chem Phys* 119:3559–3566.
21. Torrie, Valleau (1977) Nonphysical sampling distributions in monte carlo free-energy estimation: Umbrella sampling. *Journal of Computational Physics* 23:187–199.
22. Sugita Y, Okamoto Y (1999) Replica-exchange molecular dynamics method for protein folding. *Chem Phys Lett* 314:141–151.
23. Sugita Y, Kitao A, Okamoto Y (2000) Multidimensional replica-exchange method for free-energy calculations. *J Chem Phys* 113:6042–6051.
24. Park S, Kim T, Im W (2013) Transmembrane helix assembly by window exchnage umbrella sampling. *Phys Rev Lett* 108:108102–108105.
25. Park S, Im W (2013) Two dimensional window exchange umbrella sampling for transmembrane helix assembly. *J Chem Theor Comp* 9:13–17.
26. Kumar S, Bouzida D, Swendsen R H, Kollman P A, Rosenberg J M (1992) The weighted histogram analysis method for free-energy calculations on biomolecules. I. The method. *J Comp Chem* 13:1011–1021.
27. Piana S, Laio A (2007) A bias-exchange approach to protein folding. *J Phys Chem B* 111:4553–4559.
28. Ward A, Reyes C L, Yu J, Roth C B, Chang G (2007) Flexibility in the ABC transporter MsbA: Alternating access with a twist. *Proc Natl Acad Sci USA* 104:19005–19010.
29. Phillips J C, *et al.* (2005) Scalable molecular dynamics with NAMD. *J Comp Chem* 26:1781–1802.
30. MacKerell, Jr A D, Feig M, Brooks, III C L (2004) Extending the treatment of backbone energetics in protein force fields: Limitations of gas-phase quantum mechanics in reproducing protein conformational distributions in molecular dynamics simulations. *J Comp Chem* 25:1400–1415.
31. Feller S E, MacKerell, Jr A D (2000) An improved empirical potential energy function for molecular simulations of phospholipids. *J Phys Chem B* 104:7510–7515.
32. Foloppe N, MacKerrell Jr A D (2000) All-atom empirical force field for nucleic acids: I. parameter optimization based on small molecule and condensed phase macromolecular target data. *J Comp Chem* 21:86–104.
33. Jorgensen W L, Chandrasekhar J, Madura J D, Impey R W, Klein M L (1983) Comparison of simple potential functions for simulating liquid water. *J Chem Phys* 79:926–935.
34. Reid J K (1971) On the method of conjugate gradients for the solution of large sparse systems of linear equations. In Reid J K, ed., *Large Sparse Sets of Linear Equations* (Academic Press, London), pages 231–254.
35. Martyna G J, Tobias D J, Klein M L (1994) Constant pressure molecular dynamics algorithms. *J Chem Phys* 101:4177–4189.

36. Feller S E, Zhang Y, Pastor R W, Brooks B R (1995) Constant pressure molecular dynamics simulation: The Langevin piston method. *J Chem Phys* 103:4613–4621.

37. Darden T, York D, Pedersen L G (1993) Particle mesh Ewald: An $N \cdot \log(N)$ method for Ewald sums in large systems. *J Chem Phys* 98:10089–10092.

38. Bartels C (2000) Analyzing biased monte carlo and molecular dynamics simulations. *Chem Phys Lett* 331:446–454.

39. Shirts M R, Chodera J D (2008) Statistically optimal analysis of samples from multiple equilibrium states. *J Chem Phys* 129:124105.

40. Silverman B W (1986) *Density Estimation for Statistics and Data Analysis*. Monographs on statistics and applied probability (Chapman and Hall).

41. Chiu S T (1991) Bandwidth selection for kernel density estimation. *The Annals of Statistics* 19:1883–1905.

42. Hub J S, de Groot B L, van der Spoel D (2010) g_wham-a free weighted histogram analysis implementation including robust error and autocorrelation estimates. *J Chem Theor Comp* 6:3713–3720. doi:10.1021/ct100494z.

43. Humphrey W, Dalke A, Schulten K (1996) VMD – Visual Molecular Dynamics. *J Mol Graphics* 14:33–38.

44. Bakan A, Meireles L M, Bahar I (2011) ProDy: Protein dynamics inferred from theory and experiments. *Bioinformatics* 27:1575–1577.

45. Foley J, van Dam A, Feiner S, Hughes J (1990) *Computer Graphics, Principles and Practices* (Addison-Wesley, New York).

46. E W, Ren W, Vanden-Eijnden E (2005) Transition pathways in complex systems: Reaction coordinates, isocommittor surfaces, and transition tubes. *Chemical Physics Letters* 413:242 – 247. ISSN 0009-2614.
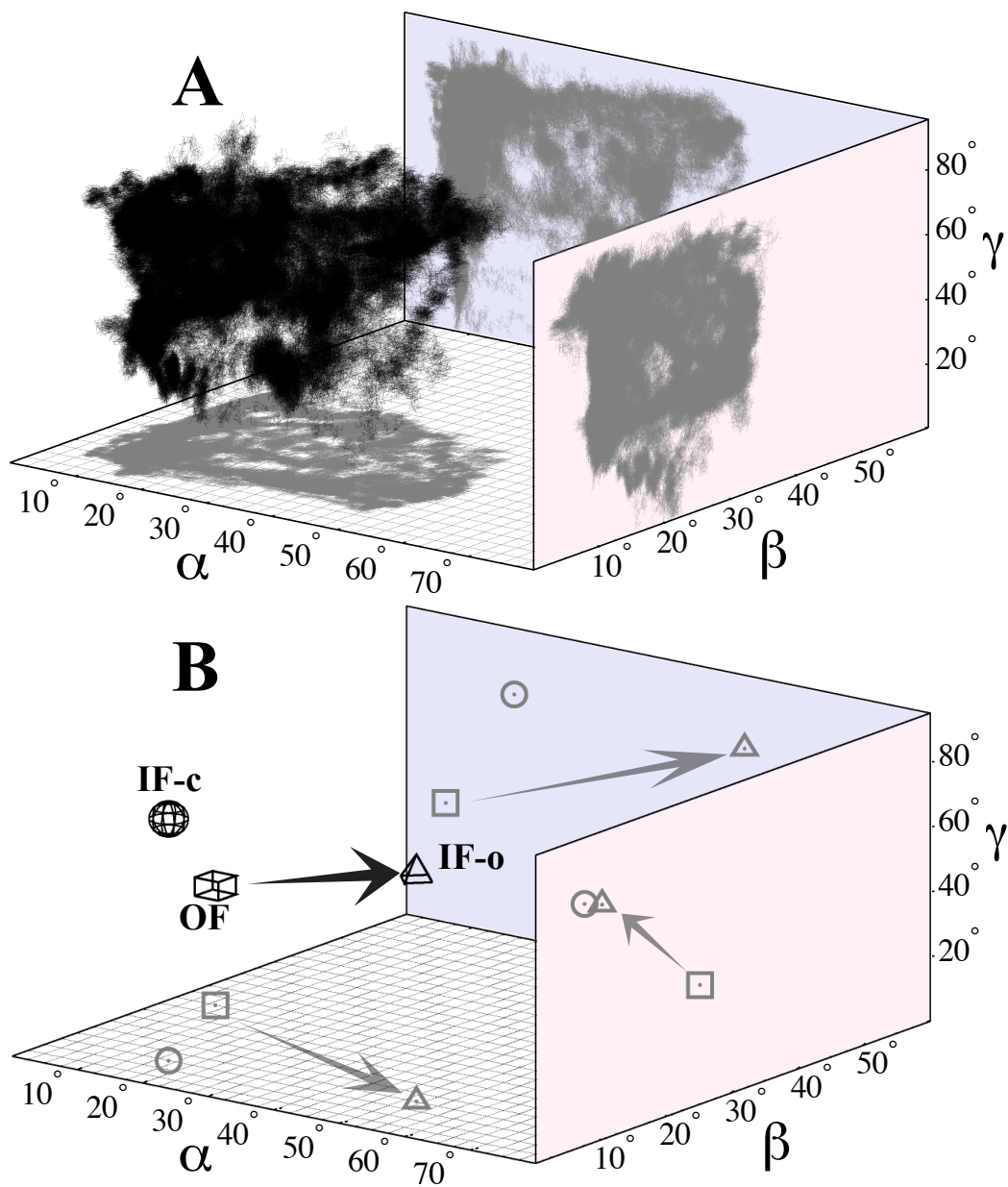
Figure S1. A Projection of ~250 biased and unbiased MD trajectories onto the (α, β, γ) space along with their projections onto the two-dimensional spaces (α, β), (α, γ), and (β, γ). **B** OF (cube), IF-c (sphere), and IF-o (pyramid) crystal structures are also shown in the (α, β, γ) space whose two-dimensional projections are given by square, circle, and triangle, respectively. The arrows illustrate the OF→IF-o transition (the main subject of this study) projected in these 3D (black) and 2D (grey) spaces. Also see Fig. 2.
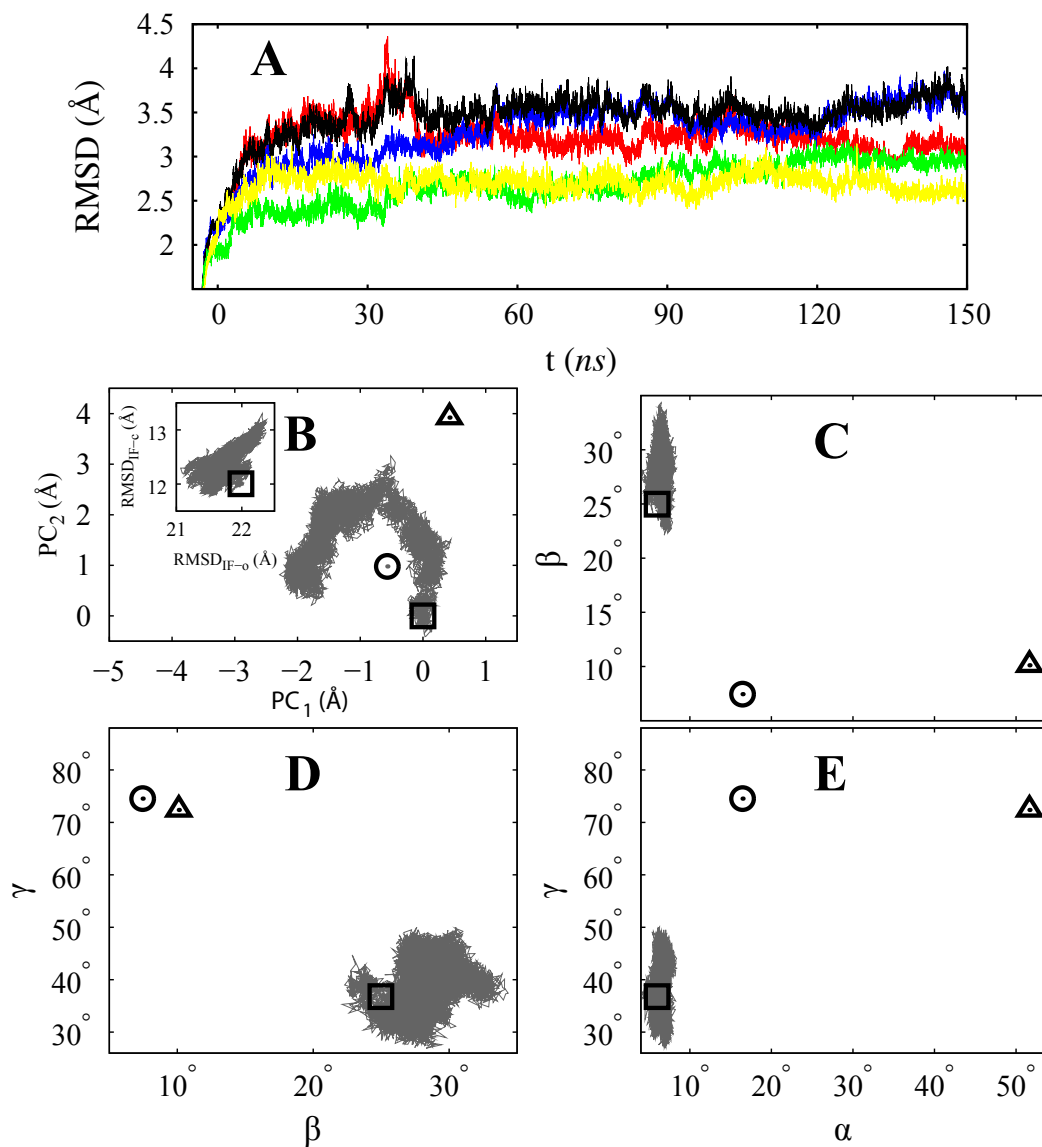
Figure S2. Unbiased equilibrium simulation of OF conformation: **A** The RMSD of a 155-ns trajectory (Table S1: Simulations 0-2, combined) from the OF crystal structure calculated based on the heavy atoms of the entire protein (black), NBDs (red), TMDs (blue), NBD^{cis} (yellow), and NBD^{trans} (green). The *apo* system stays close to the nucleotide-bound OF crystal structure during the simulations. **B** Projection of the 150-ns equilibrium trajectory of OF structure (Table S1: Simulations 1 and 2) onto its first and second principal components PC_1 and PC_2 (constructed based on the C^{α} atoms of the protein). The Projection of the same trajectory onto the $(RMSD_{IF-o}, RMSD_{IF-c})$, (α, β), (β, γ), and (α, γ) spaces are given in **B**-inset, **C**, **D**, and **E**, respectively. Square, circle, and triangle represent OF, IF-c and IF-o crystal structures. The structure firmly stays close to the known OF state throughout the simulations.
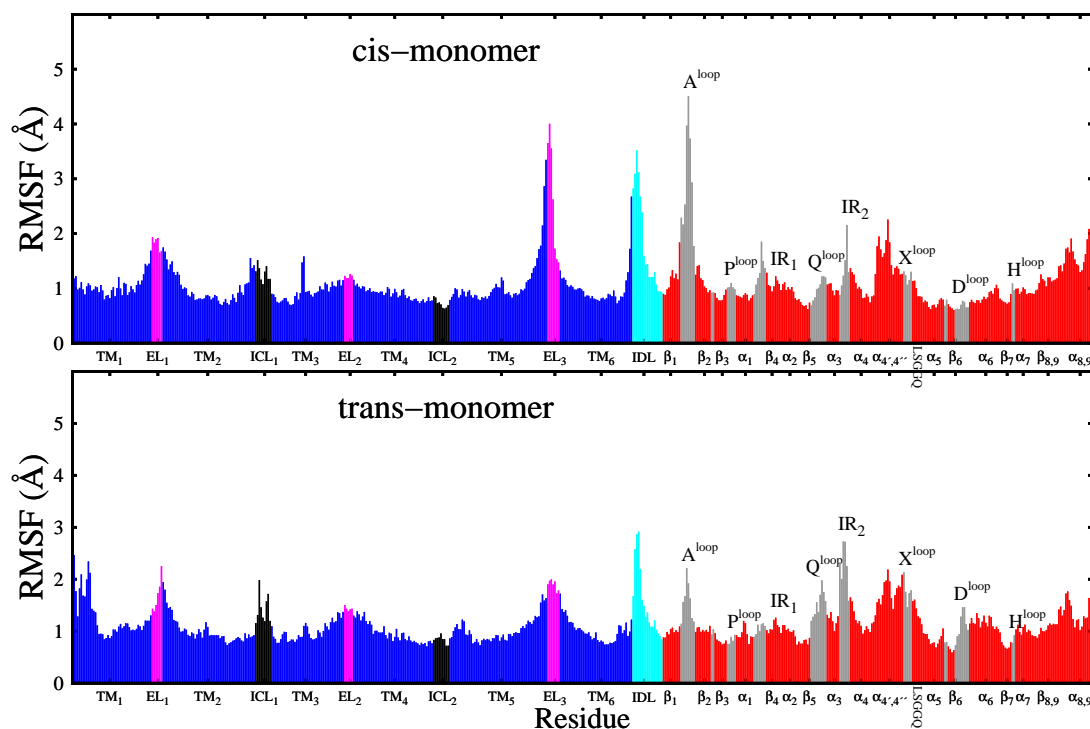
Figure S3. Root mean square fluctuations (RMSF) of the C^α atoms of the two MsbA monomers (termed *cis* and *trans*) as obtained from the last 75 ns of the equilibrium trajectory of the OF structure (Table S1: Simulation 2). Different regions of the protein are marked according to the following definitions. TMD segments: TM_1 (Residues 10-53), EL_1 (54-59), TM_2 (60-111), ICL_1 (112-120), TM_3 (121-161), EL_2 (162-166), TM_4 (167-211), ICL_2 (212-220), TM_5 (221-275), EL_3 (276-282), TM_6 (283-323); inter-domain loop (IDL) (324-340); and NBD segments: β_1 (341-350), A loop (351-358), β_2(359-367), β_3 (370-376), P loop (377-381), α_1 (382-391), β_4 (399-403), α_2 (413-418), β_5 (419-423), Q loop (424-432), α_3 (433-439), $\alpha_{4,4',4''}$ (446-475), X loop (476-480), LSGGQ signature (481-485), α_5 (486-498), Pro loop (499-500), β_6 (501-504), D loop (505-512), α_6 (513-528), β_7 (532-536), H loop (537-538), α_7 (539-545), β_8 (548-553), β_9 (555-561), $\alpha_{8,9}$ (562-581). Here TM, EL, and ICL stand for transmembrane helix, extracellular loop, and intracellular coupling loop, respectively. The NBD-TMD interaction regions IR_1 (404-418) and IR_2 (440-445) are also marked. In both monomers, the largest RMSF in TMDs belongs to the extracellular loops, particularly EL3 and EL1 while the lowest fluctuation in TMDs belongs to the intracellular coupling loop ICL2. In the NBDs, the NBD-TMD interaction region IR2 and the A loop fluctuate more than the other loops although these fluctuations are asymmetric. From these short simulations, it is not clear whether or not this asymmetric behavior is mechanistically relevant.
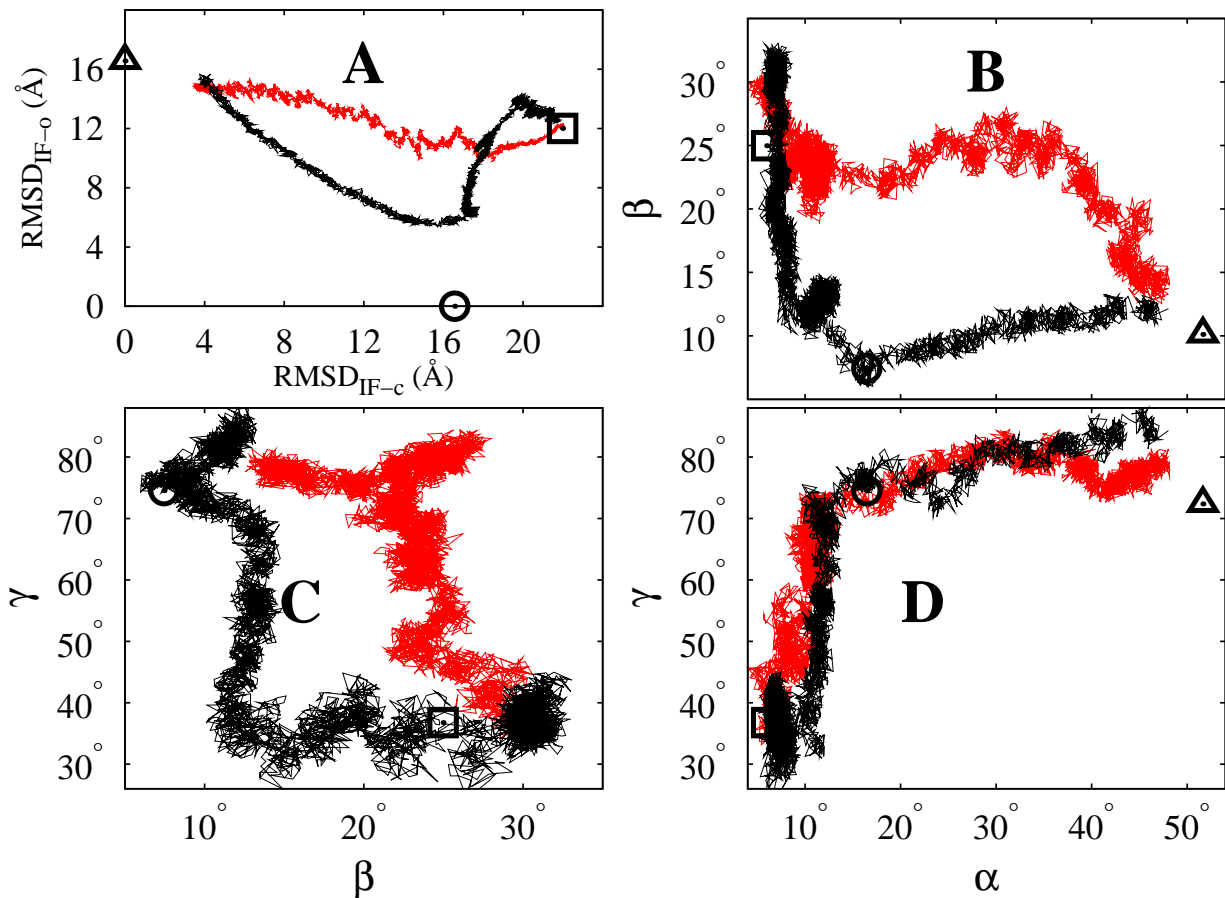
Figure S4. Comparison of the 160-ns targeted MD trajectory (red) (Table S2: Simulation 183) and the 160-ns optimized trajectory (Table S2: Simulations 33, 11, 173, and 178) projected onto the $(RMSD_{IF-o}, RMSD_{IF-c})$, (α, β), (β, γ), and (α, γ) spaces are given in **A**, **B**, **C**, and **D**, respectively. Square, circle, and triangle represent OF, IF-c and IF-o crystal structures. One particular measure to assess the relevance of these trajectories is the order of cytoplasmic opening (change in α) and periplasmic closing (change in β). The targeted MD simulations result in intermediate conformations that are open to both cyto- and periplasm – that is inconsistent with the alternating access mechanism. One interesting feature in these trajectories – that is in agreement with our optimized pathway – is that γ collective variable changes prior to the change in α, *i.e.*, the twisting motion of the NBDs occur prior to the opening of the cytoplasmic side.
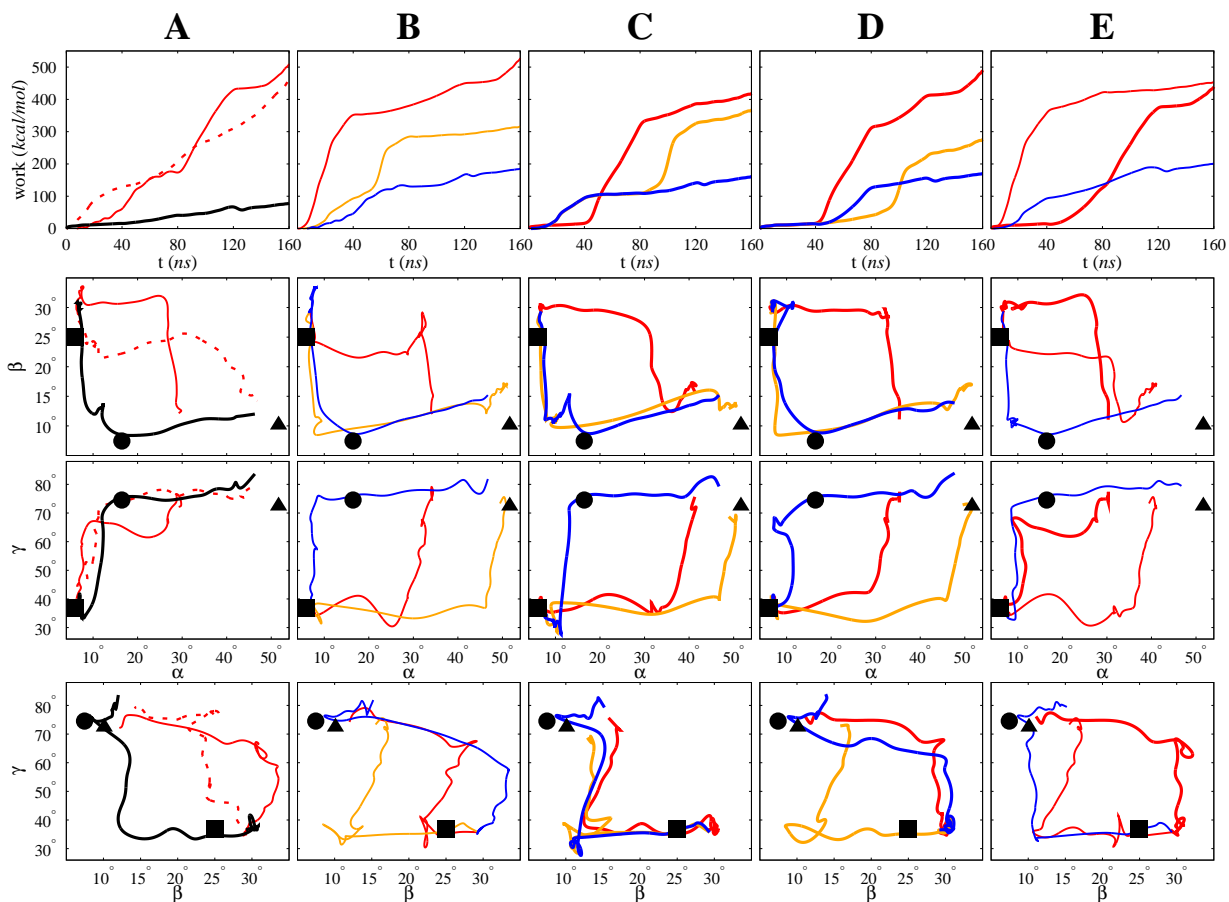
Figure S5. Work profiles and 2D traces of 160 ns driven pathways: **A** targeted MD (red, dashed line), $(d_{NBD} \rightarrow \beta \rightarrow \gamma \rightarrow \alpha)$ [the optimized pathway] (black), and $(\gamma \rightarrow \alpha \rightarrow \beta)$ (red); **B** $(\alpha \rightarrow \gamma \rightarrow \beta)$ (red), $(\beta \rightarrow \alpha \rightarrow \gamma)$ (orange), and $(\gamma \rightarrow \beta \rightarrow \alpha)$ (blue); **C** $(d_{NBD} \rightarrow \alpha \rightarrow \beta \rightarrow \gamma)$ (red), $(\beta \rightarrow d_{NBD} \rightarrow \alpha \rightarrow \gamma)$ (orange), and $(\beta \rightarrow d_{NBD} \rightarrow \gamma \rightarrow \alpha)$ (blue); **D** $(d_{NBD} \rightarrow \alpha \rightarrow \gamma \rightarrow \beta)$ (red), $(d_{NBD} \rightarrow \beta \rightarrow \alpha \rightarrow \gamma)$ (orange), $(d_{NBD} \rightarrow \gamma \rightarrow \beta \rightarrow \alpha)$ (blue); **E** $(d_{NBD} \rightarrow \gamma \rightarrow \alpha \rightarrow \beta)$ (red, thick line), $(\alpha \rightarrow \beta \rightarrow \gamma)$ (red, thin line), and $(\beta \rightarrow \gamma \rightarrow \alpha)$ (blue), The 2D traces are the projection of these trajectories onto the (α, β), (α, γ), and (β, γ) spaces. The projection of OF, IF-c, and IF-o crystal structures on the 2D spaces are given by square, circle, and triangle, respectively. The trajectories that use d_{NBD} collective variable in their protocol are represented by thick lines.
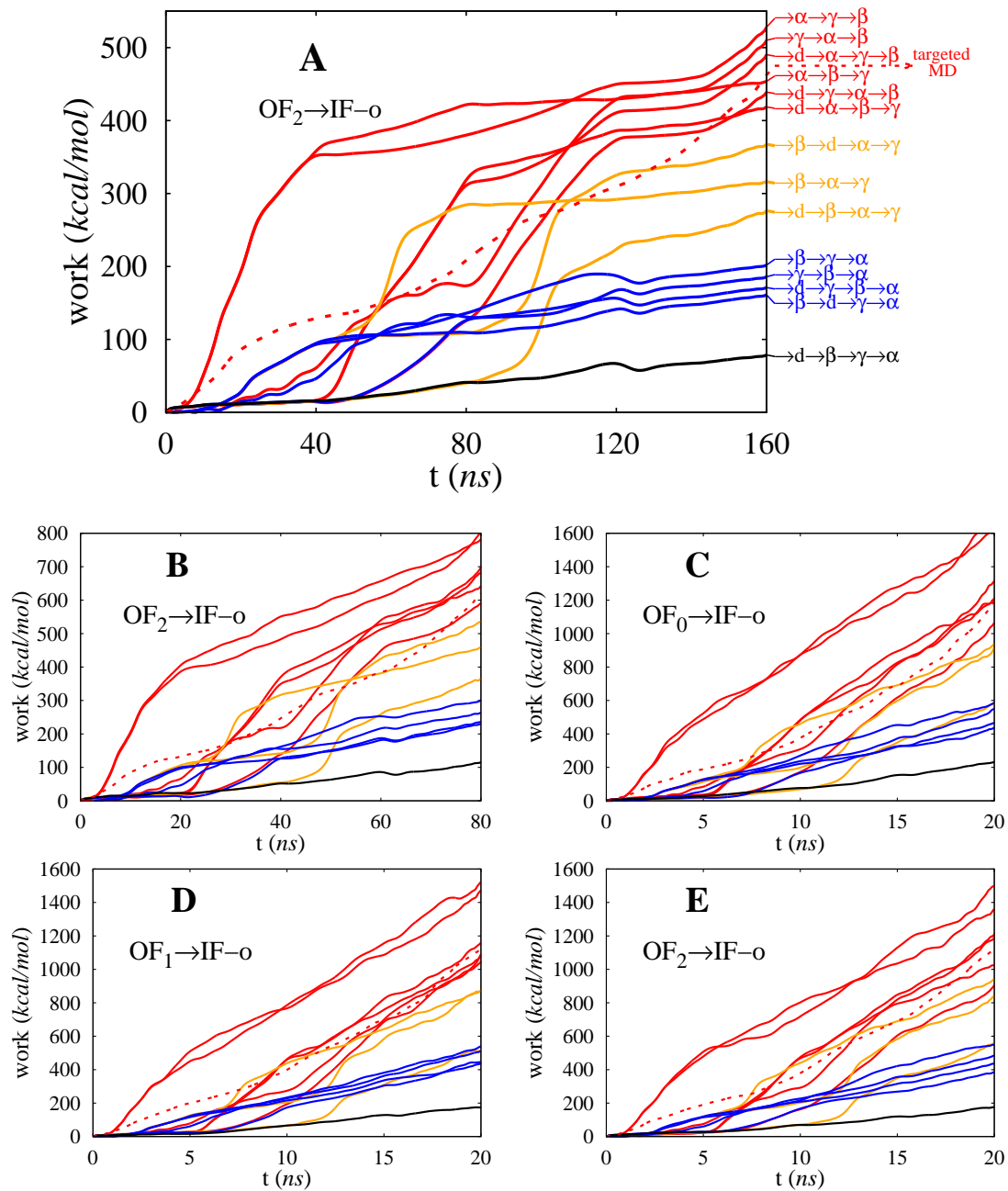
Figure S6. Work profiles of **A** 160-ns **B** 80-ns and **C-E** 20-ns driven trajectories in which OF→IF-o transition has been induced using different biasing protocols (Table S2). **Solid lines:** in each protocol, the system was driven along α, β, and γ (and in some cases d_{NBD}) in different orders. **Dashed lines:** targeted MD protocols. OF_i indicates the initial conformation used in the simulations associated with each plot; OF_1, OF_2, and OF_3 correspond to Table S1: Conformations 0, 1, and 2, respectively.
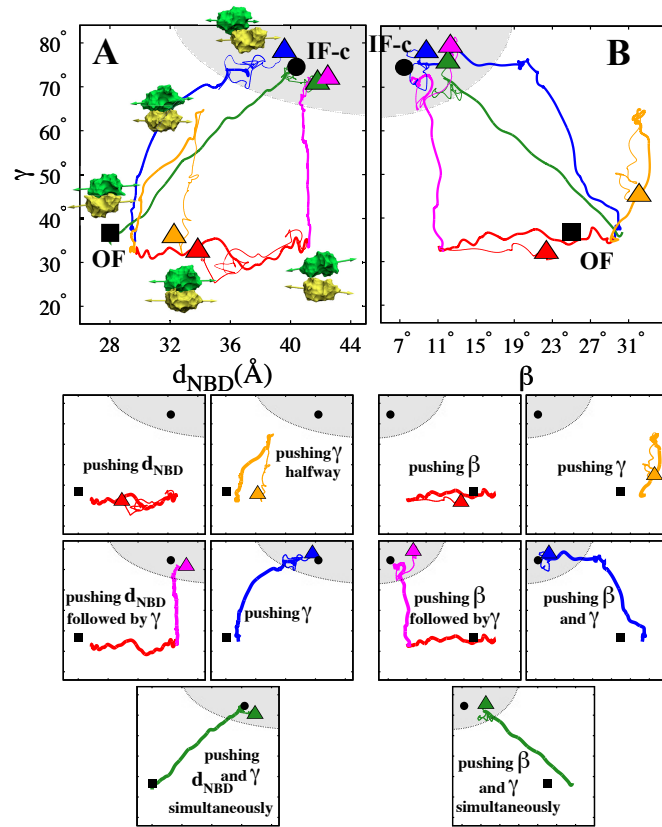
Figure S7. The role of NBD twist in NBD/TMD conformational changes. **A** The system is pushed along d_{NBD} and/or γ toward the IF-c conformation according to different protocols (thick lines), then released (thin lines). β is restrained at $9°$ along the simulations in both stages to keep the periplasmic side closed. **B** The system is pushed along γ and/or β toward the IF-c conformation according to different protocols (thick lines), then released (thin lines). The d_{NBD} is restrained at 41 Å along the simulations to keep the NBDs dissociated. See Table S6 for the list of simulations.
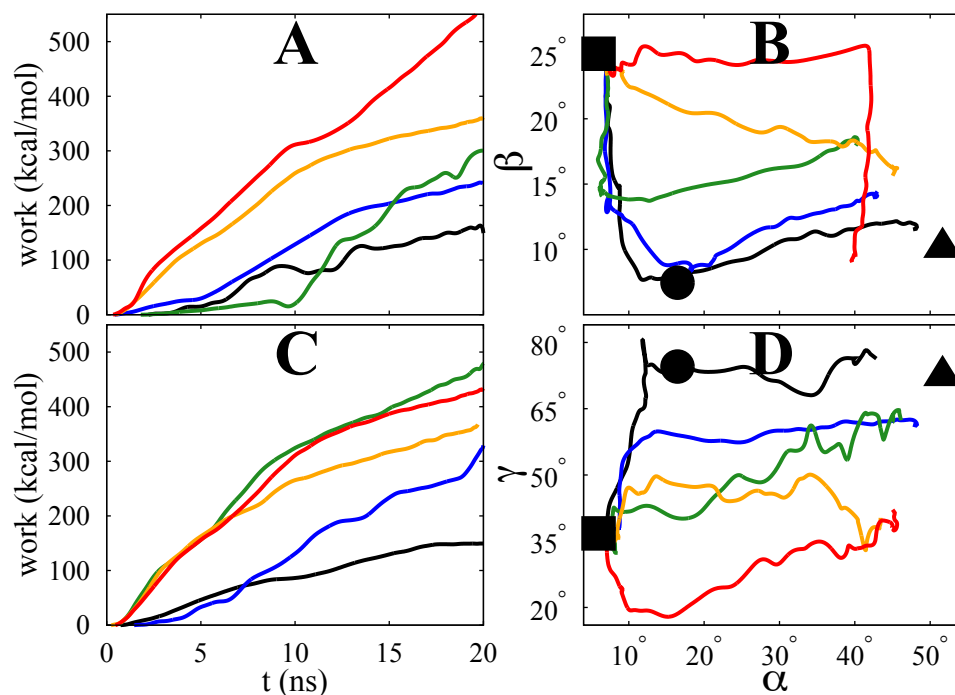
Figure S8. Exploring non-discrete protocols: **A** Work profile of a select number of 20-ns trajectories resulted from steering the system along different paths **B** in the (α, β) space. The NBDs are dissociated by pushing γ toward 75° prior to these simulations. **C** Work profile of a select number of 20-ns trajectories resulted from steering the system along different paths **D** in the (α, γ) space. The NBDs are dissociated by pushing d_{NBD} toward 41 Å prior to these simulations. See Table S5 for the complete list of simulations. **Discussion:** In the main paper, we made an assumption that the α-, β-, and γ-related conformational changes occur in discrete stages. This simplification allows us to explore several major pathways in a systematic manner. The significant drop of work required for driving the α variable when the β and/or γ variables have already changed supports this assumption to a great extent. To further examine this assumption, we performed two sets of simulations exploring the (α, β) and (α, γ) spaces. If the system is steered along different paths in the (α, β) space (**A** and **B**), the discrete $\beta \rightarrow \alpha$ protocol (involving the closure of the periplasmic side prior to the opening of the cytoplasmic side) is found to require the least amount of work. To make the comparison easier, the NBDs are dissociated and twisted (using γ) prior to these simulations. The results further support our assumption on breaking down the TMD conformational changes into two α- and β-based stages. Similarly, if the system is steered along different paths in the (α, γ) space (**C** and **D**), the discrete $\gamma \rightarrow \alpha$ protocol (involving the closure of the NBD twist prior to the TMD cytoplasmic opening) requires the least amount of work. The NBDs are dissociated by pushing d_{NBD} prior to these simulations. The results are in agreement with our conclusions (based on discrete simulations) that the γ-related conformational changes must occur prior to any significant α-related conformational change. Although (α, β) and (α, γ) spaces can be used to identify the OF→IF transition pathway of MsbA in a relatively simple manner, the exact transition pathway in (β, γ) or (γ, d_{NBD}) spaces is not clear and requires more accurate methods such as free energy calculation techniques to derive a reliable conclusion.
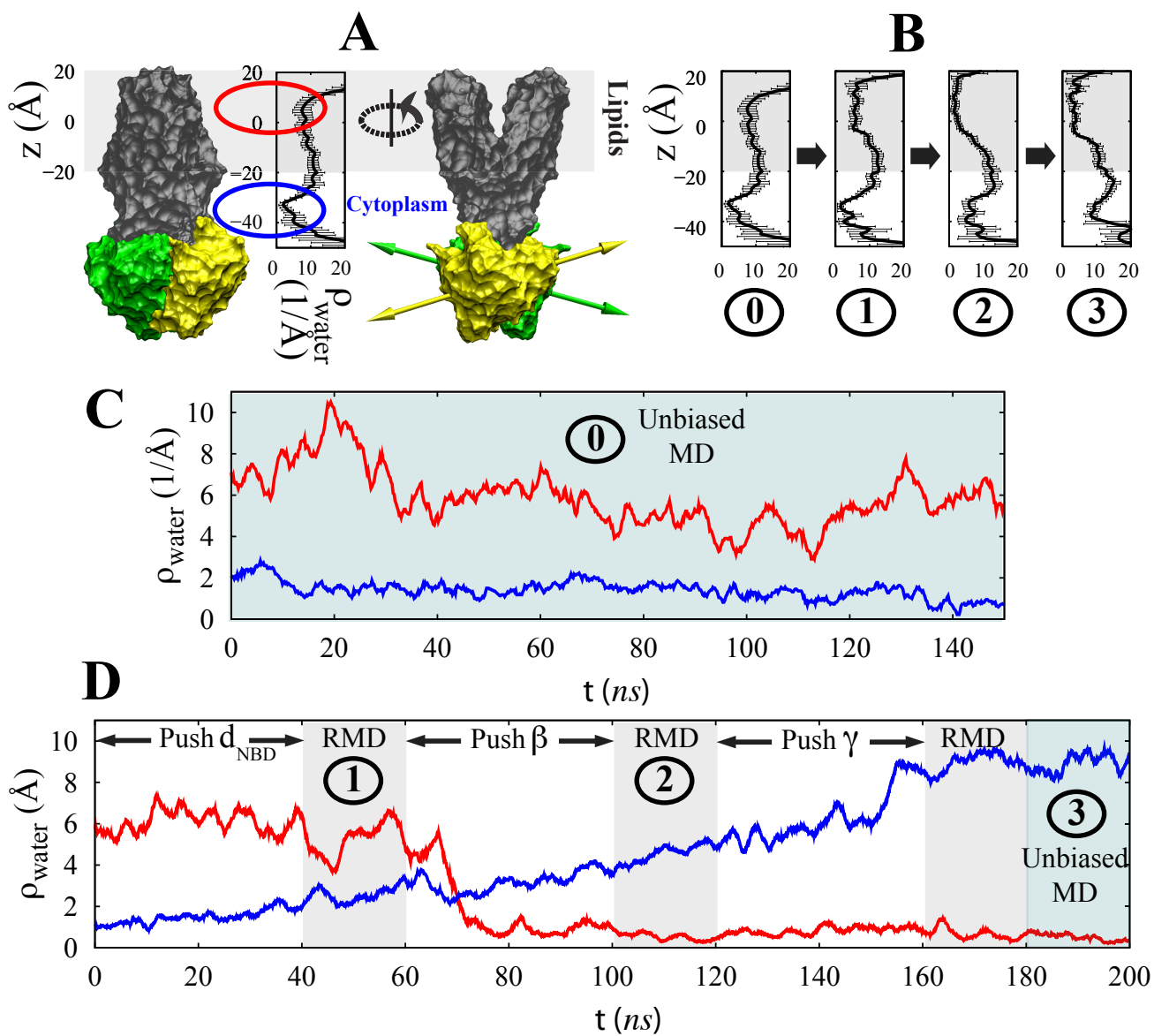
Figure S9. Water accessibility along the pore: **A** The (linear) water density along the pore, ρ_{water} as estimated from the equilibrium trajectory of OF conformation. **B** ρ_{water} associated with different conformations of MsbA as shown in **C** and **D**. **C** Time-series of ρ_{water} at the cytoplasmic (blue) and periplasmic (red) gates along the equilibrium trajectory of OF conformation. **D** Cyto- and periplasmic ρ_{water} along the OF→IF-c trajectory generated according to the optimum protocol with interstage RMD simulations (see Table S4).
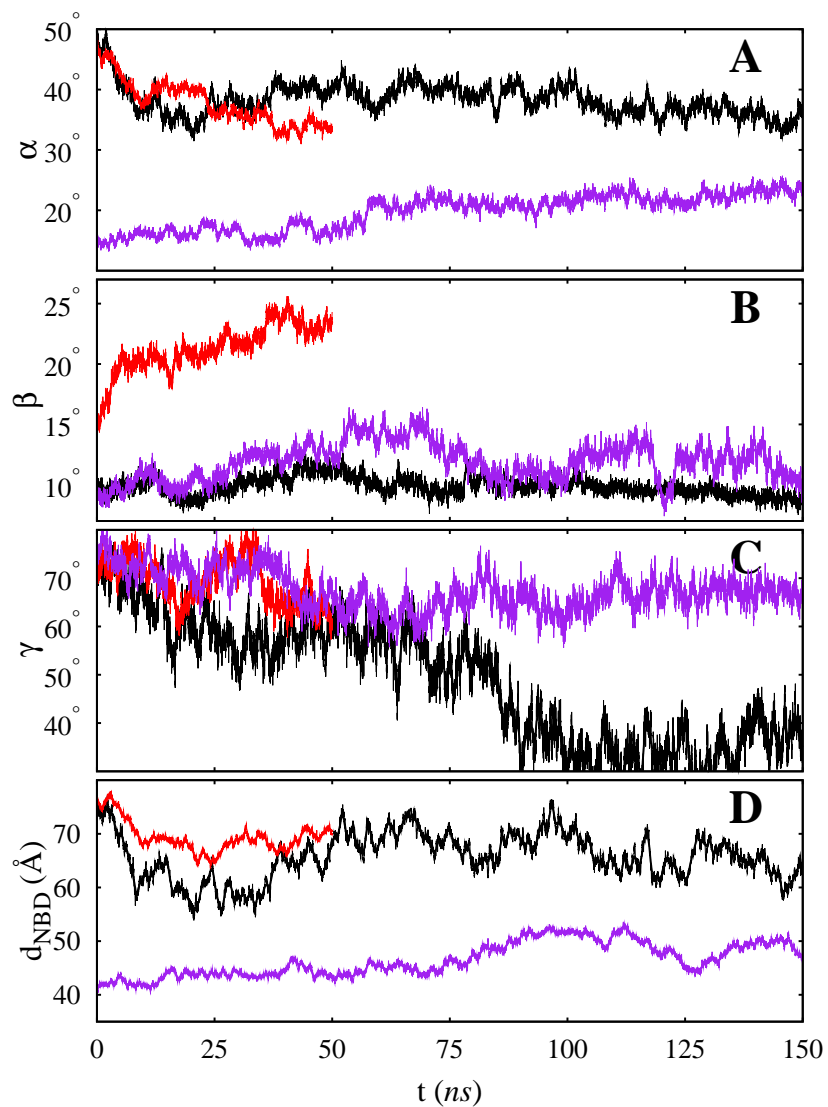
Figure S10. Resting state of MsbA in IF conformation. **A** α, **B** β, **C** γ, and **D** d_{NBD} along the unbiased equilibrium trajectories of IF-c (violet) and IF-o (black) conformations resulted from the optimum protocol along with the unbiased equilibrium trajectory of IF-o conformation resulted from the targeted MD protocol (red). See Table S1 for the list of simulations.
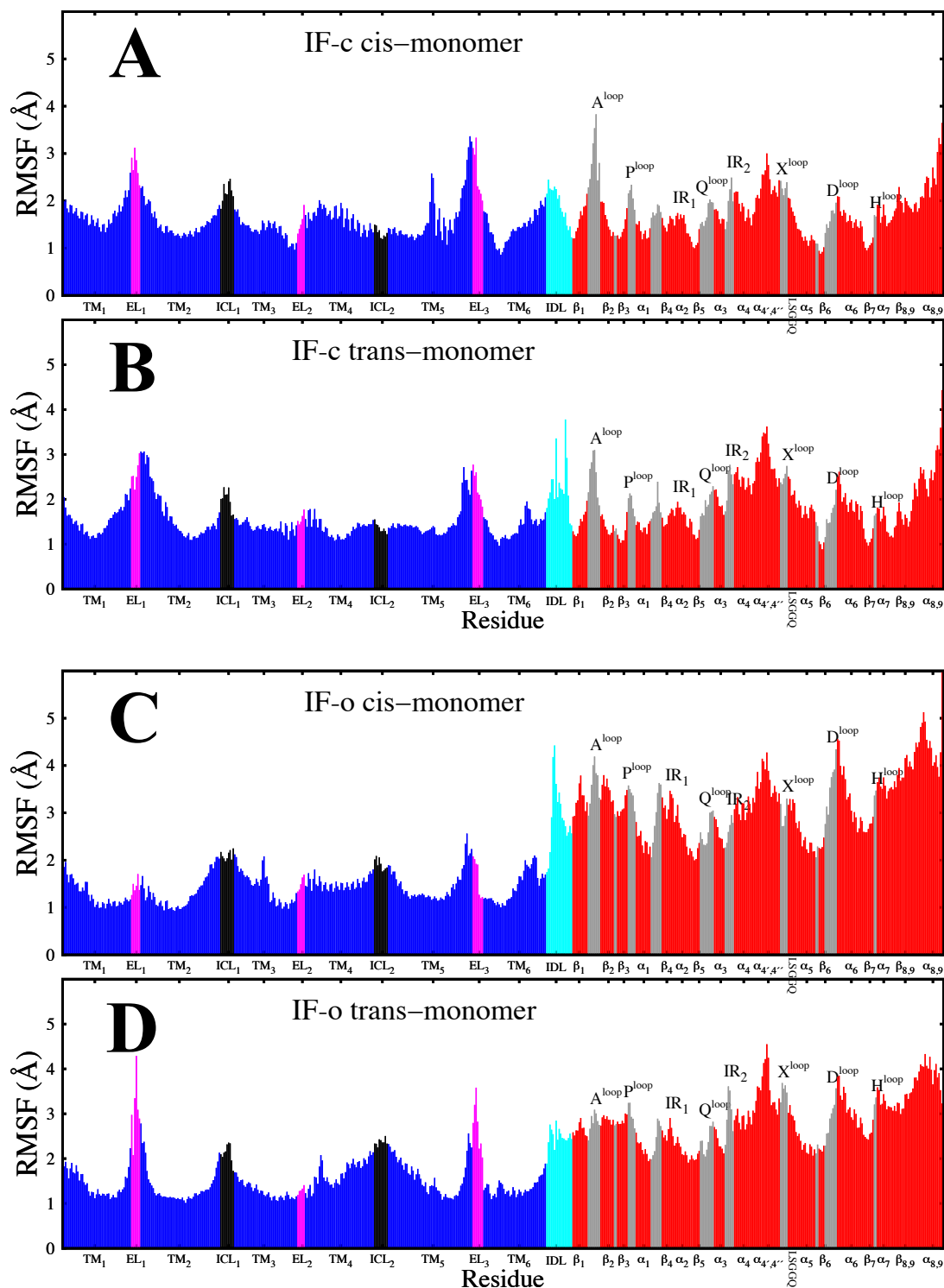
Figure S11. RMSF of C^α atoms of the two MsbA monomers (termed *cis* and *trans*) as obtained from the the last 75 ns of an unbiased equilibrium trajectory of IF-c (**A** and **B**) and IF-o (**C** and **D**) structures (Table S1: Simulations 249 and 247, respectively). See Fig. S3 for definitions.
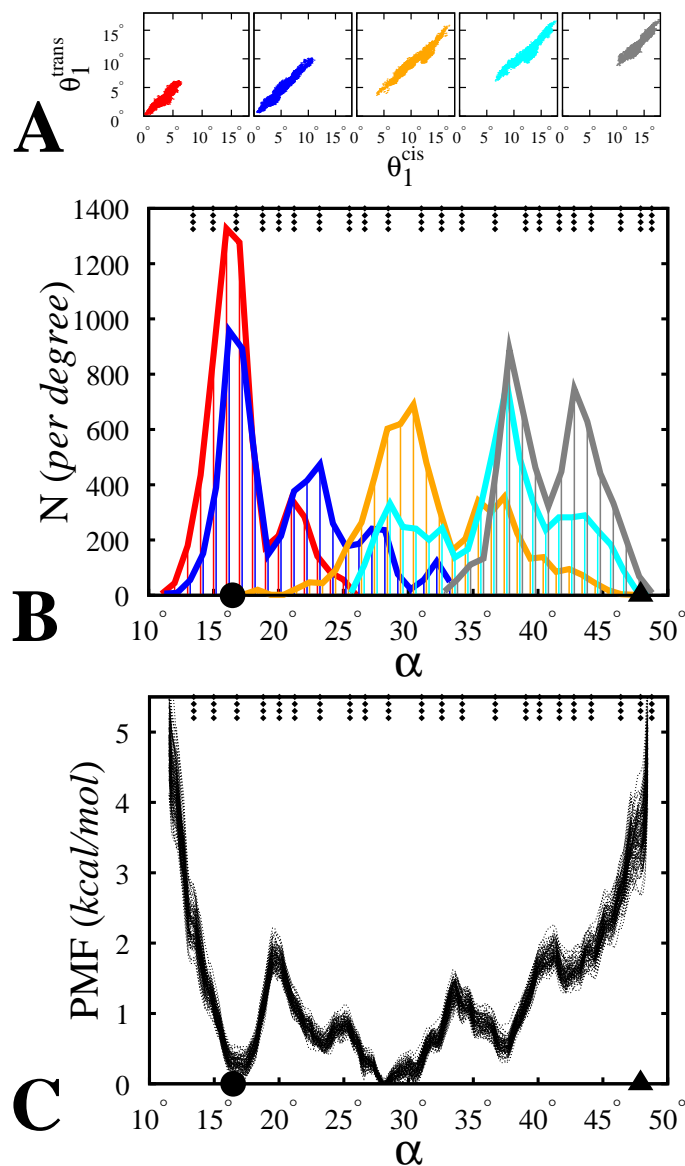
Figure S12. Sampling of the α space using BEUS MD simulations. **A** Trajectories of select replicas (5 out of 22) projected onto the $(\theta_1^{cis}, \theta_1^{trans})$ space. **B** Distribution of α as sampled by the replicas shown in **A**. **C** PMF along α as estimated using 100 sets of samples generated using a bootstrapping algorithm. The average PMF and error bars in Fig. 5 are based on the statistics shown here. Note that the centers of all 22 umbrella potentials (projected onto the α space) are marked on the upper x-axis of panels **B** and **C** (see Table S7) while the values of α associated with the IF-c (circle) and IF-o (triangle) crystal structures are marked on the lower x-axis of the same panels. It is also important to note that, although α turns out to be a good reaction coordinate to sample the configuration space of MsbA in the IF conformation, it is not to be confused with an ideal reaction coordinate (*i.e.*, the committor function [46]), thus the presence of an approximately 2-kcal/mol barrier around $\alpha \approx 20°$ is not necessarily a good representative of the actual transition barrier. Nonetheless, the estimated free energies reveal a great conformational flexibility in the resting state of *apo* MsbA.
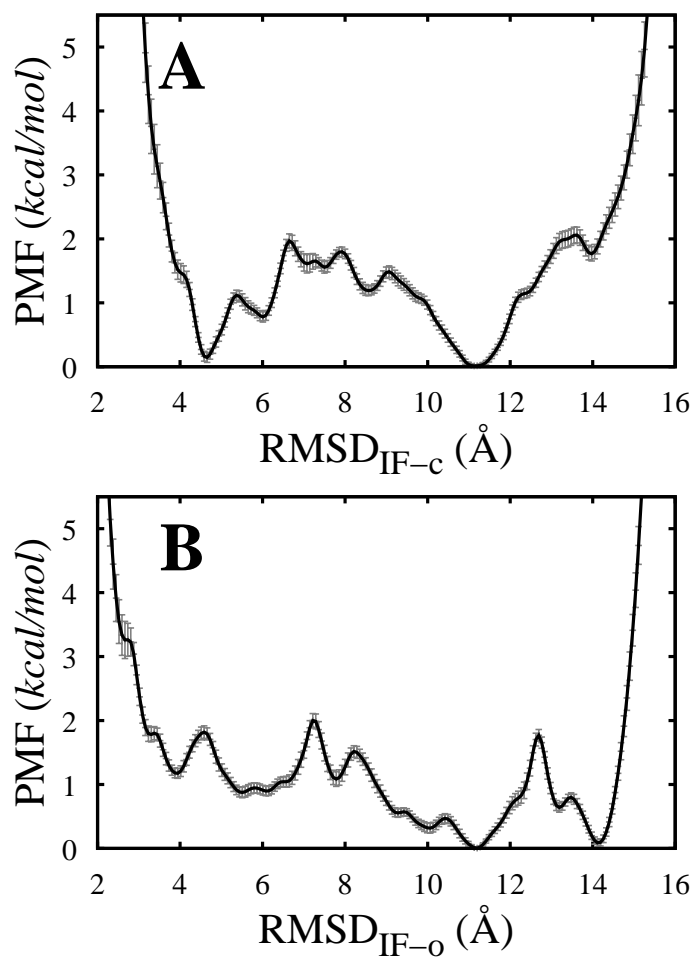
Figure S13. PMF along C'^α RMSD of the entire protein from **A** IF-c and **B** IF-o crystal structures. The PMF and the error bars are estimated using the same data and algorithm used for estimating the PMF in α space (see Figs. 5 and S12). There is a deep minimum around 4.5 Å away from the IF-c crystal structure (**A**) and a thermally accessible minimum around 4 Å away from the IF-o crystal structure (**B**). Note that the resolution of both crystal structures is 4.5 Å. One can conclude that the IF-c and IF-o are two (out of many) accessible conformations of *apo* MsbA; however, the most significant free energy basin is more than 10 Å away from both IF-c and IF-o crystal structures representing an IF conformation less open than IF-o and more open that IF-c.
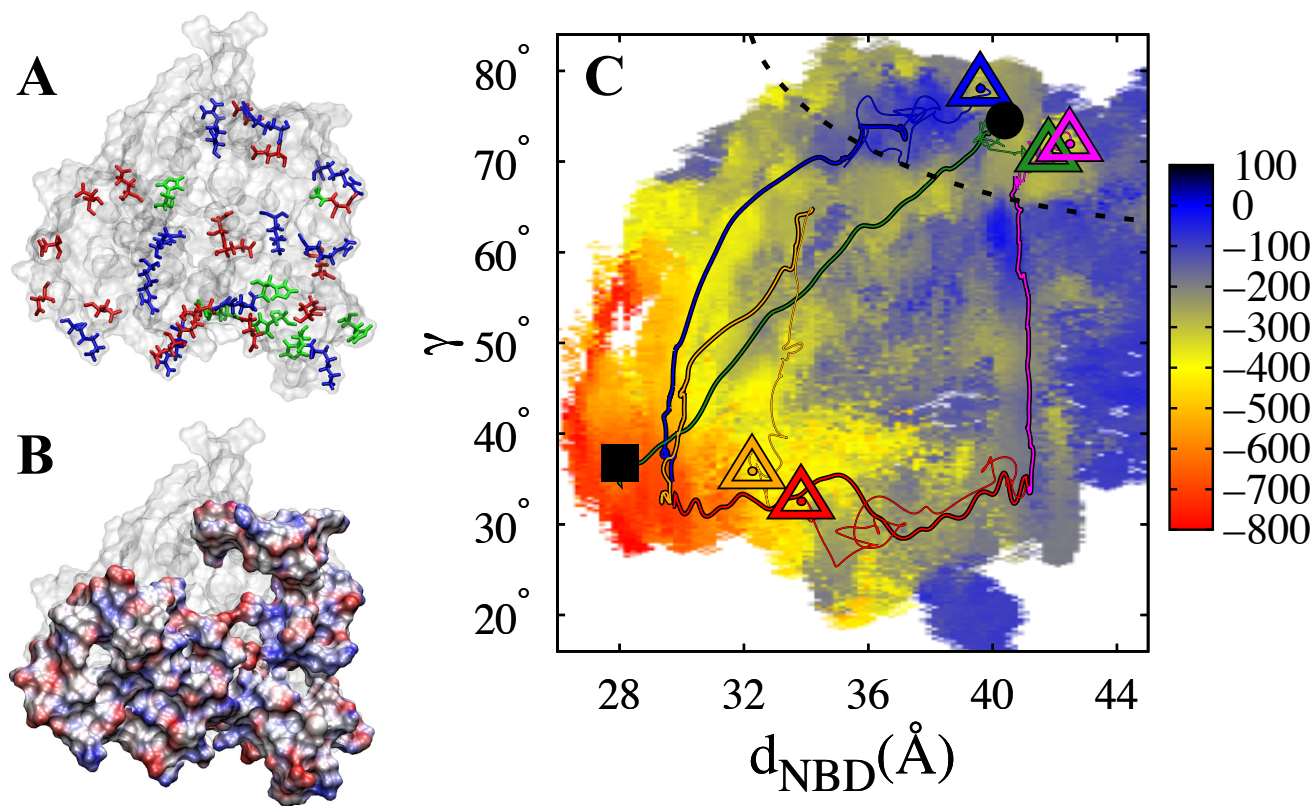
Figure S14. NBD-NBD interface: **A** The positively and negatively charged residues on the dimeric interface (highlighted as blue and red, respectively). Histidines are also shown (green) alongside. **B** The surface charge distribution of NBD on the NBD-NBD interface as obtained from a snapshot of the MsbA structure in the OF conformation (Table S1: Conformation 0). Blue/red represents positive/negative charges. **C** The interaction energy (in kcal/mol) between the two NBDs in the (d_{NBD}, γ) space along with the simulation trajectories previously shown in Fig. S7A. We used all the unbiased and biased simulations performed (see Tables S1-S6) (without any reweighting) to reconstruct the energy landscape. This energy landscape approximates the NBD-NBD interaction without considering the environment or entropic effects. One can identify a region around the IF-c crystal structure (associated with a twisted NBD conformation) that is surrounded by several regions with positive (repulsive) interaction energies. These repulsive interactions are due to the proximity of several positively-charged subdomains (*e.g.*, P-loop and H-loop) from the *cis*- and *trans*-NBDs in the absence of enough attractive interaction between the other subdomains.

Table S1. List of the unbiased simulations.

index[*]	state[†]	initial conformation	runtime (ns)
0	OF	crystal structure[‡]	5
1	OF	0	75
2	OF	1	75
247	IF-o	178	150
248	IF-o	183	50
249	IF-c	173	150
250	IF-c	192	20
251	IF-c	197	20

[*] The index by which the resulting trajectory/conformation will be referred to (*e.g.*, as an "initial conformation" for another simulation).

[†] The state associated with the initial conformation. The system can be generally considered to be in this state throughout the simulation.

[‡] See Molecular Dynamics Simulation Details.

Table S2. List of the nonequilibrium driven simulations involved in the OF→IF-o transition, induced using different biasing protocols. The work profiles associated with these simulations are shown in Fig. S6.

index[*]	initial conformation	protocol[†]	runtime (ns)	stage work (kcal/mol)[§]	total work (kcal/mol)[‡]
3,4,5,6,7	2,2,0,1,2	$\alpha \to 47°$	40,20,5,5,5	342,392,497,467,467	342,392,497,467,467
8,9,10,11,12	3,4,5,6,7	$\gamma \to 75°, \alpha$at$47°$	80,40,10,10,10	101,262,665,607,645	444,656,1164,1075,1113
13,14,15,16,17	8,9,10,11,12	$\beta \to 9°, \gamma$at$75°, \alpha$at$47°$	40,20,5,5,5	81,186,482,374,397	526,843,1648,1452,1513
18,19,20,21,22	2,2,0,1,2	$\gamma \to 75°$	80,40,10,10,10	185,238,320,308,317	185,238,320,308,317
23,24,25,26,27	18,19,20,21,22	$\alpha \to 47°, \gamma$at$75°$	40,20,5,5,5	236,294,439,503,445	422,533,760,813,763
28,29,30,31,32	23,24,25,26,27	$\beta \to 9°, \alpha$at$47°, \gamma$at$75°$	40,20,5,5,5	88,185,476,404,394	510,719,1239,1219,1159
33,34,35,36,37	2,2,0,1,2	$d_{NBD} \to 41$Å	40,20,5,5,5	24,32,36,34,33	24,32,36,34,33
38,39,40,41,42	33,34,35,36,37	$\alpha \to 47°$	40,20,5,5,5	273,324,399,402,417	298,357,436,437,451
43,44,45,46,47	38,39,40,41,42	$\gamma \to 75°, \alpha$at$47°$	40,20,5,5,5	105,173,414,401,382	405,532,852,840,835
48,49,50,51,52	43,44,45,46,47	$\beta \to 9°, \gamma$at$75°, \alpha$at$47°$	40,20,5,5,5	85,178,448,387,400	491,711,1302,1229,1237
53,54,55,56,57	3,4,5,6,7	$\beta \to 9°, \alpha$at$47°$	40,20,5,5,5	67,146,370,305,354	418,549,878,782,832
58,59,60,61,62	53,54,55,56,57	$\gamma \to 75°, \beta$at$9°, \alpha$at$47°$	80,40,10,10,10	35,224,706,605,621	454,774,1587,1389,1455
63,64,65,66,67	33,34,35,36,37	$\gamma \to 75°, d_{NBD}$at41Å	40,20,5,5,5	114,136,158,161,161	128,159,183,184,188
68,69,70,71,72	63,64,65,66,67	$\alpha \to 47°, \gamma$at$75°$	40,20,5,5,5	236,284,459,461,411	366,444,644,647,601
73,74,75,76,77	68,69,70,71,72	$\beta \to 9°, \alpha$at$47°, \gamma$at$75°$	40,20,5,5,5	72,157,402,347,338	439,602,1047,996,941
78,79,80,81,82	38,39,40,41,42	$\beta \to 9°, \alpha$at$47°$	40,20,5,5,5	76,156,388,373,354	383,507,835,820,816
83,84,85,86,87	78,79,80,81,82	$\gamma \to 75°, \beta$at$9°, \alpha$at$47°$	40,20,5,5,5	35,140,370,316,326	419,648,1207,1138,1144
88,89,90,91,92	2,2,0,1,2	$\beta \to 9°$	40,20,5,5,5	90,104,126,119,120	90,104,126,119,120
93,94,95,96,97	88,89,90,91,92	$d_{NBD} \to 41$Å, βat$9°$	40,20,5,5,5	21,54,103,100,102	112,159,230,221,223
98,99,100,101,102	93,94,95,96,97	$\alpha \to 47°, d_{NBD}$at41Å, βat$9°$	40,20,5,5,5	209,248,350,399,354	321,408,581,621,579
103,104,105,106,107	98,99,100,101,102	$\gamma \to 75°, \alpha$at$47°, \beta$at$9°$	40,20,5,5,5	46,120,341,287,279	368,529,923,910,860
108,109,110,111,112	88,89,90,91,92	$\alpha \to 47°, \beta$at$9°$	40,20,5,5,5	189,244,344,310,341	280,350,472,429,463
113,114,115,116,117	108,109,110,111,112	$\gamma \to 75°, \alpha$at$47°, \beta$at$9°$	80,40,10,10,10	36,158,496,423,438	317,510,970,854,902
118,119,120,121,122	33,34,35,36,37	$\beta \to 9°, d_{NBD}$at41Å	40,20,5,5,5	27,41,48,51,50	42,64,77,75,78
123,124,125,126,127	118,119,120,121,122	$\alpha \to 47°, \beta$at$9°, d_{NBD}$at41Å	40,20,5,5,5	179,201,263,295,268	221,266,340,370,347
128,129,130,131,132	123,124,125,126,127	$\gamma \to 75°, \alpha$at$47°, \beta$at$9°$	40,20,5,5,5	55,109,269,229,224	277,375,611,601,572
133,134,135,136,137	88,89,90,91,92	$\gamma \to 75°, \beta$at$9°$	80,40,10,10,10	94,152,332,298,319	183,259,450,412,434
138,139,140,141,142	133,134,135,136,137	$\alpha \to 47°, \gamma$at$75°, \beta$at$9°$	40,20,5,5,5	16,63,202,156,151	200,323,652,569,586
143,144,145,146,147	18,19,20,21,22	$\beta \to 9°, \gamma$at$75°$	40,20,5,5,5	30,57,154,121,140	160,224,383,329,372
148,149,150,151,152	143,144,145,146,147	$\alpha \to 47°, \beta$at$9°, \gamma$at$75°$	40,20,5,5,5	23,66,194,131,146	184,291,578,461,518
153,154,155,156,157	63,64,65,66,67	$\beta \to 9°, \gamma$at$75°, d_{NBD}$at41Å	40,20,5,5,5	32,59,126,124,142	151,203,304,285,321
158,159,160,161,162	153,154,155,156,157	$\alpha \to 47°, \beta$at$9°, \gamma$at$75°$	40,20,5,5,5	18,58,155,117,130	169,262,461,402,452
163,164,165,166,167	93,94,95,96,97	$\gamma \to 75°, d_{NBD}$at41Å, βat$9°$	40,20,5,5,5	27,50,134,105,122	136,201,356,304,345
168,169,170,171,172	163,164,165,166,167	$\alpha \to 47°, \gamma$at$75°, \beta$at$9°$	40,20,5,5,5	23,61,171,115,127	159,262,528,420,473
173,174,175,176,177	118,119,120,121,122	$\gamma \to 75°, \beta$at$9°, d_{NBD}$at41Å	40,20,5,5,5	22,30,68,58,62	61,89,142,124,134
178,179,180,181,182	173,174,175,176,177	$\alpha \to 47°, \gamma$at$75°, \beta$at$9°$	40,20,5,5,5	15,36,91,67,62	77,126,234,191,196
183,184,185,186,187	2,2,0,1,2	$RMSD_{IF-o} : 21$Å $\to 0$Å	160,80,20,20,20	460,638,1254,1111,1144	460,638,1254,1111,1144

[*] The index by which the resulting trajectory/conformation will be referred to (*e.g.*, as an "initial conformation" for another simulation). Each index in a row corresponds to a simulation with an "initial conformation" and a "runtime" specified in the same order.

[†] The center of the harmonic bias is moving toward ($\to$) a target point or is fixed "at" a constant point in the collective variable space.

[‡] Work accumulated only during the current stage ("stage work") or during the current and all previous stages ("total work").

Table S3. List of the simulations involved in a OF→IF-c transition biased according to the protocol $(\beta \to d_{NBD} \to \gamma)$ with interstage relaxations (*i.e.*, RMD).

index	initial conformation	protocol	runtime (ns)	stage work (kcal/mol)	total work (kcal/mol)
88*	2	$\beta \to 9°$	40	90	90
188	88	βat9°	20	-	-
189	188	$d_{NBD} \to 41$Å, βat9°	40	11	117
190	189	d_{NBD}at41Å, βat9°	20	-	-
191	190	$\gamma \to 75°$, d_{NBD}at41Å, βat9°	40	23	159
192	191	γat75°, d_{NBD}at41Å, βat9°	20	-	-

*This simulation was also listed in Table S2.

Table S4. List of the simulations involved in a OF→IF-o transition biased according to the "optimum protocol" $(d_{NBD} \to \beta \to \gamma \to \alpha)$ with interstage RMD relaxations (see Figs. 4 and S9).

index	initial conformation	protocol	runtime (ns)	stage work (kcal/mol)	total work (kcal/mol)
33*	2	$d_{NBD} \to 41$Å	40	24	24
193	33	d_{NBD}at41Å	20	-	-
194	193	$\beta \to 9°$, d_{NBD}at41Å	40	22	47
195	194	βat9°, d_{NBD}at41Å	20	-	-
196	195	$\gamma \to 75°$, βat9°, d_{NBD}at41Å	40	15	62
197	196	γat75°, βat9°, d_{NBD}at41Å	20	-	-
198	197	$\alpha \to 47°$, γat75°, βat9°	40	15	77
199	198	αat47°, γat75°, βat9°	20	-	-

*This simulation was also listed in Table S2.

Table S5. List of the biased simulations which explore the (α, β) and the (α, γ) spaces without assuming discrete stages (discussed* in Fig. S8).

index	initial conformation	protocol	runtime (ns)	stage work (kcal/mol)	total work (kcal/mol)
200,201	22,22	$\beta \to 9°$, γat75°	10,10	84,94	404,414
202,203	200,201	$\alpha \to 47°$, βat9°, γat75°	10,10	65,61	469,475
204,205	22,22	$\beta \to 14°$, γat75°	5,5	29,34	349,354
206,207	204,205	$\alpha \to 16°$, $\beta \to 9°$, γat75°	10,10	172,183	521,537
208,209	206,207	$\alpha \to 47°$, γat75°	5,5	37,52	558,589
210,211	22,22	$\beta \to 15°$, γat75°	10,10	21,32	341,352
212,213	210,211	$\alpha \to 47°$, γat75°	10,10	179,165	520,517
214,215	22,22	$\alpha \to 47°$, $\beta \to 9°$, γat75°	20,20	361,382	682,701
216,217	22,22	$\alpha \to 47°$, γat75°	10,10	311,334	631,654
218,219	216,217	$\beta \to 9°$, γat75°	10,10	254,286	885,940
220,221	97,97	$\gamma \to 75°$, βat9°, d_{NBD}at41Å	10,10	86,92	316,322
222,223	220,221	$\alpha \to 47°$, βat9°	10,10	63,82	379,404
224,225	97,97	$\gamma \to 60°$, βat9°, d_{NBD}at41Å	5,5	33,37	263,267
226,227	224,225	$\alpha \to 47°$, βat9°	15,15	306,330	569,597
228,229	97,97	$\alpha \to 47°$, βat9°	20,20	367,433	597,663
230,231	97,97	$\alpha \to 47°$, $\gamma \to 75°$, βat9°	20,20	474,520	704,750

*Figure S8A,B: Simulations 200 and 202 (black), Simulations 204, 206, and 208 (blue), Simulations 210 and 212 (green), Simulation 214 (orange), and Simulations 216 and 218 (red). Figure S8C,D: Simulations 220 and 222 (black), Simulations 224 and 226 (blue), Simulation 228 (orange), Simulation 229 (red), and Simulation 230 (green).

Table S6. List of the biased simulations which examine the role of the NBD twist (*i.e.*, change in γ) in other NBD and TMD conformational changes associated with OF→IF transition (discussed* in Fig. S7).

index	initial conformation	protocol	runtime (ns)	stage work (kcal/mol)	total work (kcal/mol)
232	89	$\gamma \to 65°, \beta$at$9°$	20	198	302
233	90	$\gamma \to 75°, d_{NBD} \to 41$Å$, \beta$at$9°$	20	180	306
234,235,236,237,238	94,134,164,232,233	βat$9°$	10,10,10,10,10	-	-
239	34	$\gamma \to 75°, \beta \to 9°, d_{NBD}at41$Å	20	176	208
240	34	$\gamma \to 75°, \beta \to 20°, d_{NBD}at41$Å	20	122	154
241	240	$\beta \to 9°, \gamma$at$75°, d_{NBD}$at41Å	20	29	183
242,243,244,245,246	64,119,174,239,240	d_{NBD}at41Å	10,10,10,10,10	-	-

*Figure S7A: Simulations 94 and 234 (red), Simulations 134 and 235 (blue), Simulations 164 and 236 (magenta), Simulations 232 and 237 (orange), and Simulations 233 and 238 (green). Figure S7B: Simulations 64 and 242 (orange), Simulations 119 and 243 (red), Simulations 174 and 244 (magenta), Simulations 239 and 245 (green), and Simulations 240, 241, and 246 (blue).

Table S7. List of the umbrella potentials in the BEUS simulations. The initial conformations were selected from Table S4: Simulation 198. The simulations were carried out for 24 ns (for each replica).

index	1	2	3	4	5	6	7	8	9	10	11
umbrella center (degrees)*	13.4	14.9	16.7	18.7	20.0	21.2	23.1	25.4	26.6	28.4	31.0
harmonic constant (kcal/mol)	10000	10000	5000	5000	10000	8000	4000	2500	2500	2500	2500
exchange rate (percentage) †	26	25	31	22	20	20	20	35	35	30	28
index	12	13	14	15	16	17	18	19	20	21	22
umbrella center (degrees) *	32.5	34.1	36.6	39.0	40.1	41.6	42.7	44.1	46.3	47.9	48.7
harmonic constant (kcal/mol)	5000	5000	2500	2500	5000	5000	5000	5000	5000	7500	10000
exchange rate (percentage) †	24	22	31	33	26	32	32	25	23	30	-

*The centers are projected onto the α space. The actual centers of the biasing potentials are in the $(q_{\alpha,i}^{cis}, q_{\alpha,i}^{trans})$ space (see Sampling Protocol for Free Energy Calculations).

†The reported exchnage rate for each umbrella is with its following umbrella.

Supplementary Movie S1. IF↔OF conformational transition of *apo* MsbA: The OF→IF transition was induced using the optimized protocol $(d_{NBD} \rightarrow \beta \rightarrow \gamma \rightarrow \alpha)$. The backward process was not simulated (although shown here by reversing the movie). The total simulation time was 220 ns including four 40-ns "transition" stages (using nonequilibrium driven MD) and three 20-ns interstage "relaxation" simulations (using restrained MD). The NBD/TMD conformational changes associated with each transition stage are given.