

SI Appendix

Method details

Reference structures

We used the 2 reference structures for simulations: the open-up (OU) form with the PDB code 1do2(1) and the closed-down (CD) form with 1g4b(2). The 1do2 structure is the trimer of the dimer that contains monomers in the nucleotide-free state and the nucleotide-bound (AMP-PNP) state. The 1g4b structure is most likely to bind 6 nucleotides, perhaps ADPs.

Fully atomistic simulation

We performed 5 types of MD simulations with the fully atomistic model: 1) the CD form of HslU with the substrate, the N-terminal 12 residues of Arc repressor (N-Arc) threaded into the GYVG pore; 2) the OU form of HslU with N-Arc threaded into the GYVG pore; 3) N-Arc alone; 4) the CD form of HslU with N-Arc pulled from the HslV side; 5) the OU form with N-Arc pulled from the HslV side.

All of the simulations were performed using the GROMACS 3.3.1(3, 4). HslU and N-Arc was described by the OPLS/AA force field(5) and solvated in a rectangular or dodecahedron box with TIP4P water molecules(6). The total charge of the system was neutralized by the addition of counter ion, Na⁺ or Cl⁻. The electrostatic interactions were described by the particle mesh Ewald (PME) method(7) with 8 Å cutoff and 1.2 Å grid spacing for the Fourier transform in the reciprocal space. The Lennard-Jones interactions were treated by the twin range cutoffs, 8 Å and 12 Å. Neighbor lists for the nonbonded interactions were updated every 10fs. During the simulations, the covalent bonds involving hydrogen atoms of HslU and N-Arc were constrained by LINCS(8), and the ones of water molecules were done by SETTLE(9). The MD time step was 2fs. The temperature and pressure were controlled using Berendsen algorithm(10).

The details of each simulation type: the number of trajectories, simulation box and the number of atoms are summarized in the following table.

type	#traj	Simulation box (x, y, z-axis) (Å)	#atoms: HslU	#atoms: N-Arc	#atoms: ion	#molecules: water
1	3	Rectangular: 145, 145, 130	39024 (CD)	206	58 (Na ⁺)	90734
2	3	Rectangular: 150, 150, 140	33870 (OU)	206	88 (Na ⁺)	78749
3	3	Dodecahedron: 55.0, 51.8, 44.9	0(not included)	206	2 (Cl ⁻)	4057
4	5	Rectangular:	39024	206	58	90734

		145, 145, 130	(CD)		(Na ⁺)	
5	5	Rectangular: 150, 150, 140	33870 (OU)	206	88 (Na ⁺)	78749

The initial states of the simulations in the types 1), 2), 4), and 5) were prepared in the following way. First, the substrate, N-Arc, was forcibly threaded into the CD (OU) form of the hexameric HslU ring, where the N terminus of the substrate was directed to the HslV side (HslV was not included in the simulation, though) and the N-terminal 3 residues were out of the HslU pore. Second, the system was minimized with the harmonic positional constraints ($k=10$ kJ/mol/nm²) on the C_α atoms of HslU and the substrate. Third, the system was heated in 4 steps from 0 K to 100, 200, 300 and 350K with 50ps equilibration at each temperature with the same harmonic constraints. Finally, the system was equilibrated at 350K for 100ps with the harmonic restraints for the C_α atoms of HslU. The initial setting of the simulation type 3) was identical to the above procedure without the setting of HslU.

The production runs were conducted at 350K for 4ns. The initial velocity of each atom was set up to satisfy Boltzmann distribution at 350K. HslU was constrained by the harmonic forces ($k=10$ kJ/mol/nm²) for all the C_α atoms of HslU except for 78-100 residues, which were located near the pore region. In the simulation types 4) and 5), the N-terminal acetyl carbon of the substrate was pulled from the HslV side (-z direction). The pulling force was applied by the spring with harmonic constant $k=50$ (kJ/mol/nm²), and the origin of the spring was located at -50 Å distant from the initial position of the acetyl carbon.

Coarse-grained (CG) model

HslU was represented by beads located at the C_β atoms, in contrast to the previous works(11-13) where the C_α atoms were used. Since the atomistic simulations show the side chains of Tyr-91s and Val-92s grip the substrate, we tried to incorporate the side-chain structures into the CG model by using the positional information of the C_β atoms. The substrate used as a rail was represented by one bead per residue (the details are described later.)

The total energy function of the CG model is represented by

$$V_{\text{tot}} = V_{\text{MB}}^{\text{HslU}} + V^{\text{substrate}} + V_{\text{exvol}}^{\text{substrate-HslU}} + V_{\text{HP}}^{\text{substrate-YV(HslU)}} + V_{\text{fix}}$$

where $V_{\text{MB}}^{\text{HslU}}$ is the multiple-basin energy function of HslU, $V^{\text{substrate}}$ is the energy function of the substrate, $V_{\text{exvol}}^{\text{substrate-HslU}}$ is the excluded volume effect between HslU and the substrate,

$V_{\text{HP}}^{\text{substrate-YV(HslU)}}$ is the hydrophobic interaction between the substrate and the residues of

Tyr-91s, Val-92s in the GYVG motifs of HslU, and V_{fix} is the potential of constraining the substrate rail or HslU onto their initial positions.

$V_{\text{MB}}^{\text{HslU}}$ is constructed by combining the two funnel-shaped single-basin potentials $V(R|X_{\text{OU}})$ and $V(R|X_{\text{CD}})$ (13):

$$V_{\text{MB}}^{\text{HslU}}(R|X_{\text{OU}}, X_{\text{CD}}, \Delta V_{\text{bias}}) = \frac{V(R|X_{\text{OU}}) + V(R|X_{\text{CD}}) + \Delta V_{\text{bias}}}{2} - \sqrt{\left(\frac{V(R|X_{\text{OU}}) - V(R|X_{\text{CD}}) - \Delta V_{\text{bias}}}{2}\right)^2 + C^2},$$

where R indicates the simulated structure of HslU, and X_{OU} and X_{CD} denote the reference structures in the open-up (OU) and the closed-down (CD) forms (PDB code: 1do2(1) and 1g4b(2)), respectively, which correspond to the structures at the two basins of the energy function. C is a coupling constant between the two potentials and is set to 125.0. ΔV_{bias} modulates the relative stability of the two basins. When we conduct the simulations of conformational transitions from the OU to the CD forms, we switch ΔV_{bias} from 3000, to -3000 . For the reverse transitions, we switch it back from -3000 to 3000.

The funnel-shaped single-basin potential $V(R|X_v)$ (X_v denotes the reference structure) is based on the off-lattice Gō model that was developed by Clementi *et al.* for studying protein folding(11). The potential function of this model consists of the local and nonlocal interactions, $V_{\text{local}}(R|X_v)$ and $V_{\text{non-local}}(R|X_v)$.

The local interactions are represented by

$$V_{\text{local}}(R|X_v) = \sum_{\text{bonds}} K_{b,i}(b_i - b_{v,i})^2 + \sum_{\text{angles}} K_{\theta,i}(\theta_i - \theta_{v,i})^2 + \sum_{\text{dihedral}} \{K_{\phi,i}^{(1)}[1 - \cos(\phi_i - \phi_{v,i})] + K_{\phi,i}^{(3)}[1 - \cos 3(\phi_i - \phi_{v,i})]\},$$

where b_i is the bond length between the i th and the $(i+1)$ th C_{β} atoms, θ_i is the i th bond angle between the i th and the $(i+1)$ th bonds, and ϕ_i is the dihedral angle around the $(i+1)$ th bond. Parameters with the subscript v indicate the values of the corresponding variables in the reference structure v ($v = \text{OU}, \text{CD}$). The site-specific constants $K_{b,i}$, $K_{\theta,i}$, $K_{\phi,i}^{(1)}$ and $K_{\phi,i}^{(3)}$ represent the stiffness at a site. If a site undergoes large conformational changes, the values at the site are set to be small, which assumes that the structural strain at the site upon conformational changes is released by breaking fragile local interactions(13).

They are given by

$$K_{b,i} = \min[K_b, \varepsilon_b / (b_{OU,i} - b_{CD,i})^2],$$

$$K_{\theta,i} = \min[K_{\theta}, \varepsilon_{\theta} / (\theta_{OU,i} - \theta_{CD,i})^2],$$

$$K_{\varphi,i}^{(1)} = K_{\varphi}^{(1)} \times \min[1, \varepsilon_{\varphi} / \{K_{\varphi}^{(1)}[1 - \cos(\phi_{OU,i} - \phi_{CD,i})] + K_{\varphi}^{(3)}[1 - \cos 3(\phi_{OU,i} - \phi_{CD,i})]\}],$$

and,

$$K_{\varphi,i}^{(3)} = K_{\varphi}^{(3)} \times \min[1, \varepsilon_{\varphi} / \{K_{\varphi}^{(1)}[1 - \cos(\phi_{OU,i} - \phi_{CD,i})] + K_{\varphi}^{(3)}[1 - \cos 3(\phi_{OU,i} - \phi_{CD,i})]\}],$$

where the constant values are independent of residue i , $K_b=100.0$, $K_{\theta}=20.0$, $K_{\varphi}^{(1)}=1.0$,

$$K_{\varphi}^{(3)}=0.5, \varepsilon_b=1.0, \varepsilon_{\theta}=2.0, \text{ and } \varepsilon_{\varphi}=2.0.$$

The nonlocal interactions of the single-basin potential are represented by

$$V_{\text{nonlocal}}(R | X_v) = \varepsilon \sum_{i < j-3}^{\text{native contacts}, X_v} \min \left\{ 1, 5 \left(\frac{r_{v,ij}}{r_{ij}} \right)^{12} - 6 \left(\frac{r_{v,ij}}{r_{ij}} \right)^{10} + 1 \right\} + \varepsilon \sum_{i < j-3}^{\text{all native contacts}} \max \left\{ 0, 5 \left(\frac{r_{0ij}^{\text{min}}}{r_{ij}} \right)^{12} - 6 \left(\frac{r_{0ij}^{\text{min}}}{r_{ij}} \right)^{10} \right\} + \varepsilon \sum_{i < j-3}^{\text{non native contacts}} \left(\frac{D}{r_{ij}} \right)^{12},$$

where the r_{ij} is the distance between the C_{β} atoms of the i th and the j th amino acids. The

constants are the following: $r_{0ij}^{\text{min}} = \min[r_{OU,ij}, r_{CD,ij}]$, $D=4.0\text{\AA}$, and $\varepsilon=0.18$. The first term is the

attractive interactions for “native-contact” residue pairs at the reference structure, X_v , which are defined if one of the nonhydrogen atoms in the i th amino acid is within a certain cutoff distance from one and more of the nonhydrogen atoms in the j th amino acid at the reference structure, X_v . The cutoff distance is set to 5.5\AA for the native contacts within subunits, and 6.5\AA for those between subunits so that the conformational changes between the OU and the CD forms proceed smoothly. The second term represents repulsive interactions for all native contact pairs, which consist of the native contact pairs at X_{CD} and those at X_{OU} . The last term is repulsion for nonnative pairs: the pairs that are not included in the native contact pairs at any reference structures.

$V^{\text{substrate}}$ is the funnel-shaped single-basin potential consisting of only local interactions and is given by

$$V^{\text{substrate}}(R | X_{\text{substrate}}) = \sum_{\text{bonds}} K_b (b_i - b_0)^2 + \sum_{\text{angles}} K_\theta (\theta_i - \theta_0)^2 + \sum_{\text{dihedral}} \{K_\phi^{(1)} [1 - \cos(\phi_i - \phi_0)] + K_\phi^{(3)} [1 - \cos 3(\phi_i - \phi_0)]\}$$

$X_{\text{substrate}}$ represents the reference structure of the substrate. The bond lengths and the dihedral angles are set up so that the structure forms a stretched chain; $b_0=3.8 \text{ \AA}$, $\phi_0 = 230^\circ$. The bond angles are set to $\theta_0=130^\circ$ ($\theta_0=150^\circ$ for the thinner substrate). Other constant parameters are the same as those in $V_{\text{local}}(R | X_v)$.

$V_{\text{exvol}}^{\text{substrate-HslU}}$ is the excluded volume effect between HslU and the substrate and is defined

by

$$V_{\text{exvol}}^{\text{substrate-HslU}} = \varepsilon \sum_{\substack{i \in \text{HslU} \\ j \in \text{substrate}}} \left(\frac{D}{r_{ij}} \right)^{12},$$

where $\varepsilon=0.18$, and $D=6.0\text{\AA}$, (8.0\AA only for Tyr-91). Since the atomistic simulations show that the bulky side chains of Tyr-91s gripped the substrate firmly in the CD form, the distance parameter D was set to the maximal value, $D=8.0 \text{ \AA}$, that does not lead to the fatal collision in the CD form. Since valine is smaller than tyrosine, we set $D=6.0\text{\AA}$ for Val-92. For all the other residues, we also set $D=6.0\text{\AA}$.

The hydrophobic interaction between the substrate and Tyr-91s and Val-92s of HslU is taken from (14) and is given by

$$V_{\text{HP}}^{\text{substrate-YV(HslU)}} = \sum_{\substack{i \in \text{substrate}, \\ \text{YV(HslU)}}} \varepsilon_{\text{HP}} S_{\text{HP}}(\rho_i),$$

where the constant ε_{HP} is equal to -1.0 . $S_{\text{HP}}(\rho_i)$ represents the buriedness of a residue number i and is defined as

$$S_{\text{HP}}(\rho_i) = \begin{cases} 1.0 & (\rho_i \geq 1.0) \\ c_{\text{linear}} \rho_i + 0.5(1.0 - c_{\text{linear}}) \left[1.0 + \cos \left(\frac{\pi(1.0 - \rho_i)}{1.0 - \rho_{\text{min}}} \right) \right] & (\rho_{\text{min}} < \rho_i < 1.0) \\ c_{\text{linear}} \rho_i & (\rho_i \leq \rho_{\text{min}}) \end{cases},$$

where $c_{\text{linear}} = 0.2$, ρ_i is the local density of atoms around the i th residue, and $\rho_{\text{min}}=0.3$. ρ_i is expressed as

$$\rho_i = \frac{\sum_{\substack{j \in \text{substrate when } i \in Y, V(\text{HslU}) \\ j \in Y, V(\text{HslU}) \text{ when } i \in \text{substrate}}} u_{\text{HP}}(i, j)}{n},$$

where n is the maximum coordination number of the i th residue and is set to 3.0. The function u_{HP} is the degree of the contact between the i th and the j th residues. Here, the summation is taken only when one of the i - j pair is the substrate residue and the other one is Tyr-91 or Val-92 of HslU. The function u_{HP} is defined by

$$u_{\text{HP}}(i, j) = \begin{cases} 1.0 & (r_{ij} \leq d_{\min}) \\ 0.5 \left[1 + \cos \left(\frac{\pi(r_{ij} - d_{\min})}{d_{\max} - d_{\min}} \right) \right] & (d_{\min} < r_{ij} < d_{\max}) \\ 0.0 & (d_{\max} \leq r_{ij}) \end{cases},$$

where $d_{\min} = 6.0 \text{ \AA}$ and $d_{\max} = 10.0 \text{ \AA}$ for Tyr-91 and $d_{\min} = 4.0 \text{ \AA}$ and $d_{\max} = 8.0 \text{ \AA}$ for Val-92. The atomistic simulations elucidated the difference of the grip strength between the CD and the OU forms, thus d_{\max} and d_{\min} were chosen to make the grip at the CD form firm and to make the grip at the OU form loose simultaneously. For this purpose, these parameters of Tyr-91 were determined to be $d_{\min} = 6.0 \text{ \AA}$ and $d_{\max} = 10.0 \text{ \AA}$, since the pore radius measured by six Tyr-91s were 6.7 \AA in the CD and 10 \AA in the OU forms (Fig.4C). The determined d_{\max} and d_{\min} of Tyr-91 were $\pm 2 \text{ \AA}$ from the excluded volume distance ($D = 8.0 \text{ \AA}$), thus the parameters of Val-92 were also determined to be $\pm 2 \text{ \AA}$ from $D = 6.0 \text{ \AA}$.

V_{fix} is represented by

$$V_{\text{fix}} = \begin{cases} c_{\text{fix}} \sum_i (L_i - d)^2 & (L_i \geq d) \\ 0 & (L_i < d) \end{cases},$$

where L_i is the distance of the i th residue from its initial position, and $c_{\text{fix}} = 1.0$. For the simulations where the substrate rail was used, the substrate was constrained at every 10 residues and d was set to 1.0. For the simulations where the substrate was translocated, the 214th residues of HslU were fixed and d was set to 3.0.

To perform conformational transitions for four subunits out of the six subunits, we redefined the multiple-basin potential as follows. The hexameric structure of 1do2(1), the OU form, is a trigonal crystal. Based on this symmetry, the 6 subunits, S_i ($i=1,2,3,4,5$, and 6) are divided into the 3 groups, S_1S_2 , S_3S_4 , and S_5S_6 . Then, the multiple-basin potential is given by

$$V_{\text{MB}}^{\text{HslU}} = V_{\text{MB}}^{S_1S_2} + V_{\text{MB}}^{S_3S_4} + V_{\text{MB}}^{S_5S_6},$$

$$V_{\text{MB}}^{S_iS_{i+1}} = V_{\text{MB}}^{S_iS_{i+1}}(R_{S_iS_{i+1}} | X_{\text{OU}}^{S_iS_{i+1}}, X_{\text{CD}}^{S_iS_{i+1}}, \Delta V_{\text{bias}}^{S_iS_{i+1}}) \quad i = 1, 3, 5.$$

$V_{\text{MB}}^{S_iS_{i+1}}$ is the multiple-basin potential of the subunits S_i and S_{i+1} , $R_{S_iS_{i+1}}$ is the simulated

structure of them. $X_{OU}^{S_i S_{i+1}}$ and $X_{CD}^{S_i S_{i+1}}$ are the reference structures of the subunits S_i and S_{i+1} at the OU and the CD forms, respectively.

$\Delta V_{bias}^{S_1 S_2}$ is the constant value, -1000 (the CD form is stable), and $\Delta V_{bias}^{S_3 S_4}$ and $\Delta V_{bias}^{S_5 S_6}$ are simultaneously switched from -1000 to 1000 or from 1000 to -1000 at every 4×10^4 time step to conduct conformational transitions for the four subunits, S_3 , S_4 , S_5 , and S_6 .

In the simulations where the intra- and inter-subunits conformational changes are separately and sequentially induced upon the $OU \rightarrow CD$ and the $CD \rightarrow OU$ transitions of hexameric HslU, the multiple-basin potential is set up as follows.

$$V_{MB}^{HslU} = V_{MB}^{Intra}(R | X_{OU}, X_{CD}, \Delta V_{bias}^{Intra}) + V_{MB}^{Inter}(R | X_{OU}, X_{CD}, \Delta V_{bias}^{Inter}),$$

where the V_{MB}^{Intra} , V_{MB}^{Inter} indicate the multiple-basin potentials for the intra- and inter-subunits

interactions, respectively. V_{MB}^{Intra} consists of the two single-basin potentials containing the local and the nonlocal interactions as described above, but the nonlocal interactions are calculated only for the residue pairs within the subunits, $V^{Intra}(R | X_\nu) = V_{local}^{Intra}(R | X_\nu) + V_{nonlocal}^{Intra}(R | X_\nu)$.

On the other hand, V_{MB}^{Inter} consists of the two single-basin potentials containing only the nonlocal interactions that are calculated for the pairs between the subunit, $V^{Inter}(R | X_\nu) = V_{nonlocal}^{Inter}(R | X_\nu)$. ΔV_{bias}^{Intra} and ΔV_{bias}^{Inter} are switched between 3000, and -3000 depending on the order of conformational transitions. For example, when we perform the $OU \rightarrow CD$ transitions by the order of 1) intra- and 2) inter-subunits conformational changes, the simulations start with $\Delta V_{bias}^{Intra} = \Delta V_{bias}^{Inter} = 3000$. Then the values of the parameters are changed to $\Delta V_{bias}^{Intra} = -3000$ and $\Delta V_{bias}^{Inter} = 3000$, and finally changed to $\Delta V_{bias}^{Intra} = \Delta V_{bias}^{Inter} = -3000$.

Pore radius defined by Tyr-91s of the six subunits

The pore radius at a time t , $Radius(t)$, is defined by the C_β atoms of Tyr-91s of the six subunits,

$$Radius(t) = \frac{\sum_{i \in Y91}^6 \sqrt{(x_i(t) - M_x(t))^2 + (y_i(t) - M_y(t))^2}}{6},$$

where x_i and y_i are the x - and y -coordinates of the C_β atoms of Tyr-91s, and M_x and M_y are the x - and y -coordinates of the center of mass of the C_β atoms of six Tyr-91s.

The CG simulation protocol

MD simulations are conducted by the underdamped Langevin equation at a constant temperature T ,

$$m\dot{\mathbf{v}}_i = \mathbf{F}_i - \gamma\mathbf{v}_i + \hat{\mathbf{I}}_i$$

where \mathbf{v}_i is the velocity of the i th residue and a dot indicates the derivative with respect to time. m is the mass, and γ is the friction coefficient. \mathbf{F}_i and $\hat{\mathbf{I}}_i$ are systematic and random forces on the i th residue. The white Gaussian random forces $\hat{\mathbf{I}}_i$ satisfy $\langle \hat{\mathbf{I}}_i \rangle = 0$ and its variance is $\langle \hat{\mathbf{I}}_i(t)\hat{\mathbf{I}}_i(t') \rangle = 2\gamma k_B T \delta(t-t')\mathbf{1}$, where $\mathbf{1}$ is a 3×3 -unit matrix. For numerical integration of the Langevin equation, we used the algorithm proposed by Honeycutt and Thirumalai (15). We used $m = 10.0$, $\gamma = 0.01$, $k_B T = 0.16$ and the finite time step $\Delta t = 0.2$.

1. Bochtler, M., Hartmann, C., Song, H. K., Bourenkov, G. P., Bartunik, H. D. & Huber, R. (2000) *Nature* **403**, 800-5.
2. Wang, J., Song, J. J., Franklin, M. C., Kamtekar, S., Im, Y. J., Rho, S. H., Seong, I. S., Lee, C. S., Chung, C. H. & Eom, S. H. (2001) *Structure* **9**, 177-84.
3. Berendsen, H. J. C., Vanderspoel, D. & Vandrunen, R. (1995) *Comput Phys Commun* **91**, 43-56.
4. Lindahl, E., Hess, B. & van der Spoel, D. (2001) *J Mol Mod* **7**, 306-317.
5. Kaminski, G. A., Friesner, R. A., Tirado-Rives, J. & Jorgensen, W. L. (2001) *J Phys Chem B* **105**, 6474-6487.
6. Jorgensen, W. L. & Madura, J. D. (1985) *Mol Phys* **56**, 1381-1392.
7. Essmann, U., Perera, L., Berkowitz, M. L., Darden, T., Lee, H. & Pedersen, L. G. (1995) *J Chem Phys* **103**, 8577-8593.
8. Hess, B., Bekker, H., Berendsen, H. J. C. & Fraaije, J. G. E. M. (1997) *J Comp Chem* **18**, 1463-1472.
9. Miyamoto, S. & Kollman, P. A. (1992) *J Comp Chem* **13**, 952-962.
10. Berendsen, H. J., Postma, J. P. M., van Gunsteren, W. F., DiNola, A. & Haak, J. R. (1984) *J. Chem. Phys.* **81**, 3684-3690.
11. Clementi, C., Nymeyer, H. & Onuchic, J. N. (2000) *J Mol Biol* **298**, 937-53.
12. Koga, N. & Takada, S. (2006) *Proc Natl Acad Sci U S A* **103**, 5367-72.
13. Okazaki, K., Koga, N., Takada, S., Onuchic, J. N. & Wolynes, P. G. (2006) *Proc Natl Acad Sci U S A* **103**, 11844-9.
14. Fujitsuka, Y., Chikenji, G. & Takada, S. (2006) *Proteins* **62**, 381-98.
15. Honeycutt, J. D. & Thirumalai, D. (1992) *Biopolymers* **32**, 695-709.