

## Supporting material

### Additional models considered

In addition to the common  $G\bar{o}$  model and the two  $G\bar{o}$ +HP models described in the maintext, we considered two other  $C_\alpha$  coarse-grained models with varied interaction schemes. The first variation, the “ $G\bar{o}$ +MJ” model, is purely native-centric. It uses the Miyazawa-Jernigan residue-residue statistical potential<sup>1</sup> for native contact strengths. In this scheme, instead of using a uniform, homogeneous native contact strength as in the common  $G\bar{o}$ -model approach, pairwise native-centric energies depend on the types of amino acid residues that are in contact. We have recently applied such an approach in a theoretical analysis of  $\phi$ -values<sup>2</sup>. The heterogeneous native contact energies used in the present  $G\bar{o}$ +MJ model are the same as those in ref.2. The second variation is referred to as the “ $G\bar{o}$ +mb” model. This interaction scheme also uses Miyazawa-Jernigan statistical potential for pairwise native contact energies as in the “ $G\bar{o}$ +MJ” model; but in addition, a many-body local-nonlocal coupling effect is added, such that the total potential energy of the model is given by

$$E = E_{bonded} + \sum_{i < j-3}^{native} E_{mb} + \sum_{i < j-3}^{nonnative} f_{rep} \varepsilon \left( \frac{r_{rep}}{r_{ij}} \right)^2 + \sum_{i < j-3}^{nonnative} E_{nonnative} \quad (S1)$$

where  $E_{bonded}$  corresponds to the terms for  $C_\alpha$ - $C_\alpha$  virtual bond lengths, bond angles, and torsion angles as given by the first three summations on the right-hand side of Eq.(1) of ref. 3, with energy parameters  $K_r = 100$ ,  $K_\theta = 20$ ,  $K_\phi^{(1)} = 0.5$ , and  $K_\phi^{(3)} = 0.25$ . The many-body energy  $E_{mb}$  term with local-nonlocal coupling now takes the form

$$E_{mb} = \{1 - \alpha + \alpha \exp[-(A + B)]\} E_{db} \quad (S2)$$

in which

$$A = [(\Delta_1)^2 + (\Delta_2)^2]/(\Delta_c)^2 \quad (S3)$$

$$B = (\psi - \psi_0)^2/B_0 \quad (S4)$$

Here,  $\alpha$  is an attenuation factor that weakens the native contact energy if the local conformation of one or both sets of five contiguous residues is nonnative; we set  $\alpha = 0.2$  in the present study. For any given conformation, the factors  $\Delta_1$  and  $\Delta_2$  are the minimum root-mean-square-deviations, in Å, of the spatial coordinates of two 5-residue segments centered around each of the contacting residues (i.e., residues  $i-2, i-1, i, i+1, i+2$  and residues  $j-2, j-1, j, j+1, j+2$ ) from their respective best-superposed local conformations in the native PDB structure.  $\psi$  is the cross angle (in radian) of the two segments, which is defined as the cross angle of the two lines connected by the terminal residue positions of each segment, and  $\psi_0$  is the corresponding cross angle of the two segments in the PDB structure. The parameters  $\Delta_c$  and  $B_0$  were introduced to tune the effect of local conformational variation, here we set  $(\Delta_c)^2 = 0.4 \text{ \AA}^2$  and  $B_0 = 0.4 \text{ (radian)}^2$ .  $E_{db}$  is the desolvation potential defined in reference<sup>3</sup>, but now with the  $\varepsilon$  in that reference modified to the Miyazawa-Jernigan potential  $\varepsilon_{MJ}$ . In the nonnative repulsive term,  $f_{rep} = 0.8$ , and  $r_{rep} = \sigma_i + \sigma_j$ , where  $\sigma_i$  and  $\sigma_j$  are the experimental radii of residue  $i$  and  $j$ , respectively<sup>4</sup>. The last summation of  $E_{nonnative}$  over

nonnative residue pairs accounts for sequence-dependent nonnative interactions, and the term for residue pair  $i,j$  is given by

$$E_{nonnative} = E_{nonnative}(i,j) = -K_{nonnative} (\varepsilon_{MJ})_{\mu\nu} \exp[-(r_{ij} - \sigma_i - \sigma_j)/2] \quad (S5)$$

where  $\mu, \nu$  are the amino acid residue types of  $i, j$ , respectively. Here  $(\varepsilon_{MJ})_{\mu\nu} \equiv e_{\mu\nu}/(\text{constant})$ , where  $e_{\mu\nu}$  corresponds to the 210 “ $e_{ij}$ ” quantities provided in the upper-triangular part of Table 3 in ref.1 (in which residue types were labeled by  $i,j$ ); the rescaling factor, (constant), is negative, and is taken to be 1.465/210 times the average of all 210  $e_{ij}$ s such that the most favorable (most strongly negative)

$(\varepsilon_{MJ})_{\mu\nu}$  is about -1.5 (i.e.,  $constant = 1.465 \sum_{\mu=1}^{20} \sum_{\nu=1}^{20} e_{\mu\nu}/210$ ). For the present study, we chose an overall nonnative interaction strength  $K_{nonnative} = 0.3$ .

## Supporting Figures

Figure S1. A control computation to gain further insight into the dependence of transition path time and mean folding first passage time on folding free energy barrier height. The plotted quantity here are the essentially the same as those represented by the red circles in Figure 6 (A), except the one-dimensional diffusion dynamics is now computed using an *ad hoc* free energy profile obtained by uniformly scaling the free energy profile generated by explicit-chain dynamics in all-atom MD simulations by a factor of two such that the overall folding barrier height in the present *ad hoc* profile is twice that of the original. The Pearson correlation coefficient for the present data points is  $r = 0.74$  and the slope of fitted dash line is 1.02.

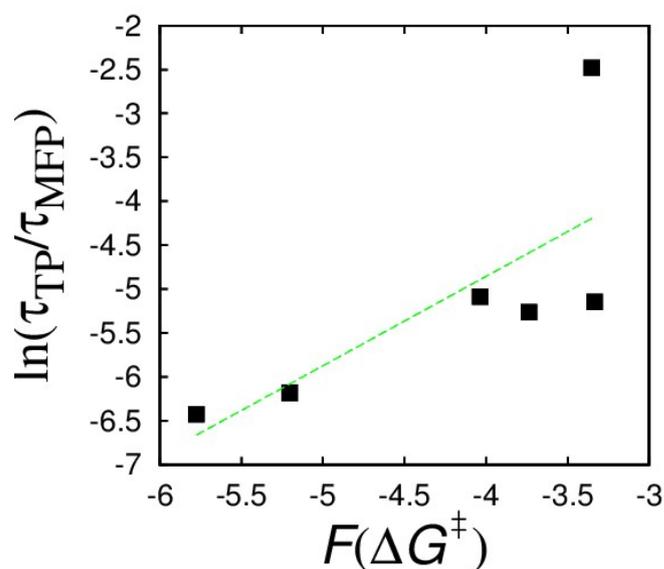


Figure S2. Contact patterns of transition path ensembles. The predicted native contact probability maps of transition path (TP) computed by the common G $\bar{o}$  model (top-left of each panel) and by all-atom MD simulations (bottom-right of each panel) are shown for the four studied proteins. The plotted native contact probabilities are computed from all conformations visited, irrespective of their  $Q$  values, along all simulated transition paths.

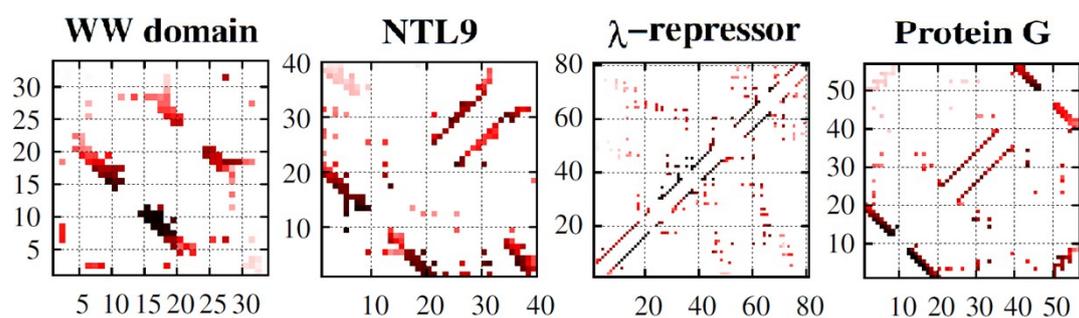


Figure S3. Computed  $\phi$ -values of the four studied proteins predicted by all-atom MD simulations (black triangles), the  $G\bar{o}+MJ$  (red squares), and the  $G\bar{o}+mb$  (blue circles) models. Lines connecting plotted symbols are merely guides for the eye.

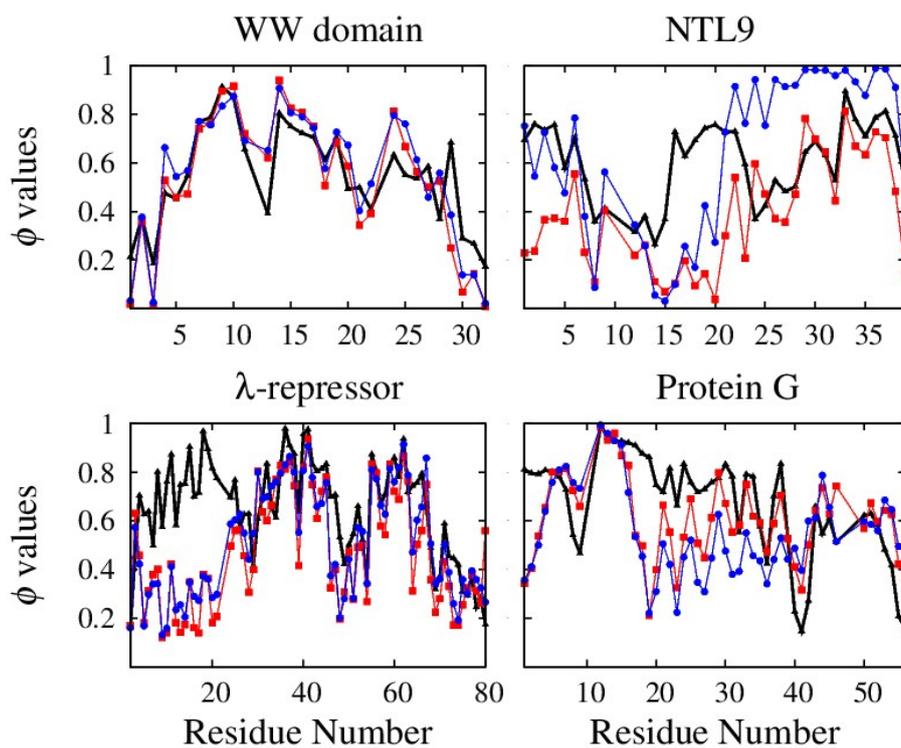


Figure S4. Correlation between mean folding first passage time  $\tau_{\text{MFP}}$  predicted by all-atom MD simulations and  $\tau_{\text{MFP}}(\text{MC})$ , which is the corresponding quantity modeled as one-dimensional diffusion dynamics using the free energy profile generated by all-atom MD simulations, wherein each Kawasaki Monte Carlo step can either add or subtract one native contact or leaving the number of native contacts unchanged, i.e. the step size in  $Q$  is given by  $\delta Q = 1/Q_n$ , where  $Q_n$  is the number of native contacts in the PDB structure of the given protein (black squares, left vertical scale; Pearson coefficient  $r = 0.311$ ). If the data point for  $\lambda$  repressor (black square marked by blue circle) is excluded, the correlation coefficient improves significantly to  $r = 0.867$ . Alternate Kawasaki dynamics results  $\tau_{\text{MFP}}(\text{MC2})$  using a constant  $\delta Q = 1/80$  irrespective of the protein are shown by the red circles (right vertical scale); with  $\tau_{\text{MFP}}(\text{MD}) - \tau_{\text{MFP}}(\text{MC2})$  correlation coefficient  $r \approx 0$ . Details of our Kawasaki formulation is provided in ref. 5 and the Supporting Information of this reference.

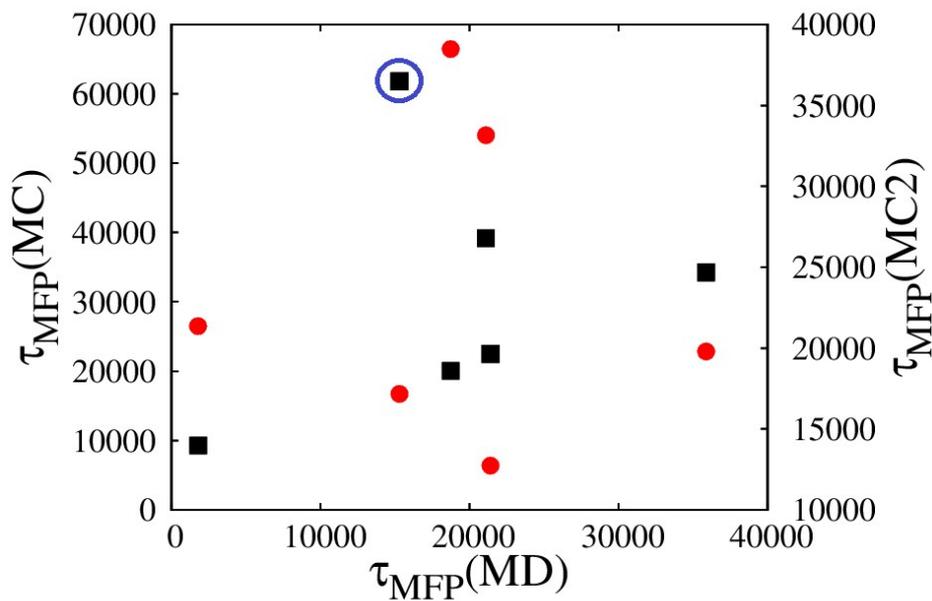
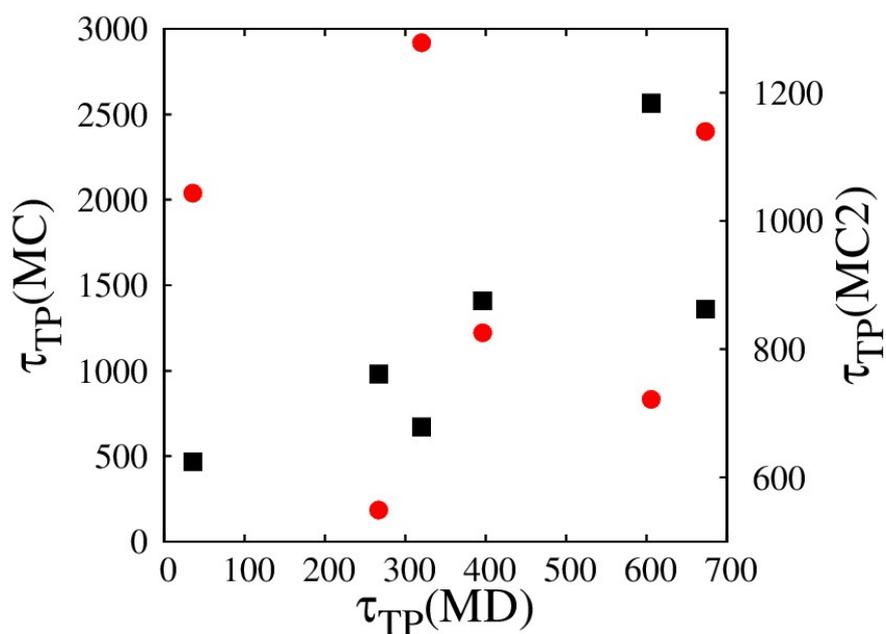


Figure S5. Correlation between transition path time  $\tau_{\text{TP}}$  predicted by all-atom MD simulations and  $\tau_{\text{TP}}$  (MC), which is the corresponding quantity computed by one-dimensional Kawasaki diffusion dynamics ( $\delta Q = 1/Q_n$ ) using the free energy profile generated by all-atom MD simulations (black squares, the left vertical scale;  $r = 0.765$ ). The transition path time  $\tau_{\text{TP}}$  (MC2) computed using the alternate protein-independent  $\delta Q = 1/80$  Kawasaki dynamics is shown by the red circles (right vertical scale), with  $\tau_{\text{TP}}$  (MD)— $\tau_{\text{TP}}$  (MC2) correlation coefficient  $r \approx 0$ .



## References

1. S. Miyazawa and R.L. Jernigan, *J. Mol. Biol.*, 1996, **256**, 623-644.
2. Z. Zhang, Y. Ouyang and T. Chen, *Phys. Chem. Chem. Phys.*, 2016, **18**, 31304-31311
3. Z. Liu and H.S. Chan, *Phys. Biol.*, 2005, **2**, S75-S78
4. A.A. Zamyatnin, *Ann. Rev. Biophys. Bioeng.*, 1984, **13**, 145-165
5. Z. Zhang, H.S. Chan, *Proc. Natl. Acad. Sci. USA*. 2012, **109**, 20919-20924