Electronic Supplementary Material (ESI) for Physical Chemistry Chemical Physics.

Electronic Supplementary Information for: "First-Passage Times in Protein Folding: Exploring the Native-Like States vs Overcoming the Free Energy Barrier"

Sergei F. Chekmarev*,†,‡

†*Institute of Thermophysics, SB RAS, 630090 Novosibirsk, Russia*

‡*Department of Physics, Novosibirsk State University, 630090 Novosibirsk, Russia*

E-mail: chekmarev@itp.nsc.ru

1 Deviation from the Transition State

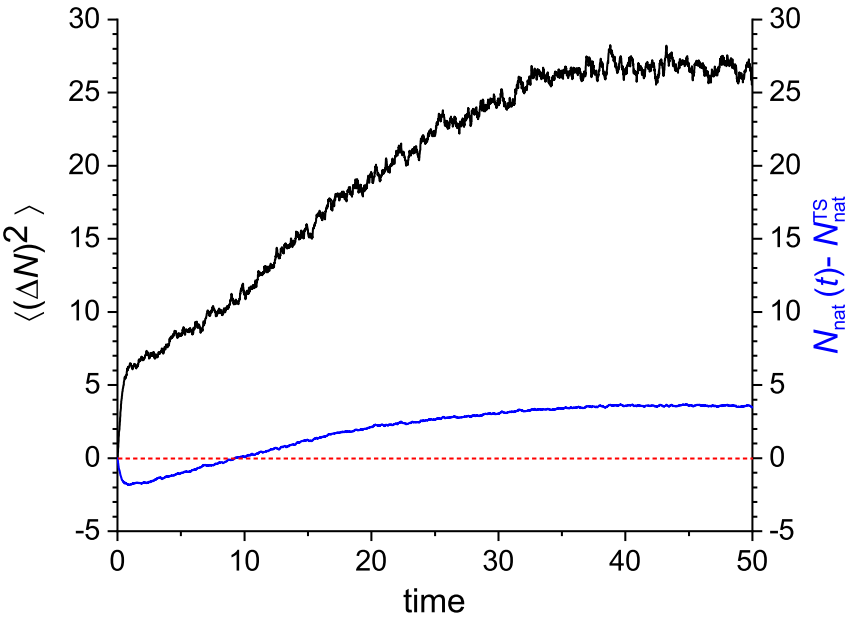

Figure S1: $T = 0.2$, $\gamma = 10M/\tau$, and ten thousand MD trajectories. The mean square (black curve) and mean (blue curve) deviations from the transition state.

2 A Simple Model for Single-Exponential First-Passage Time Distribution

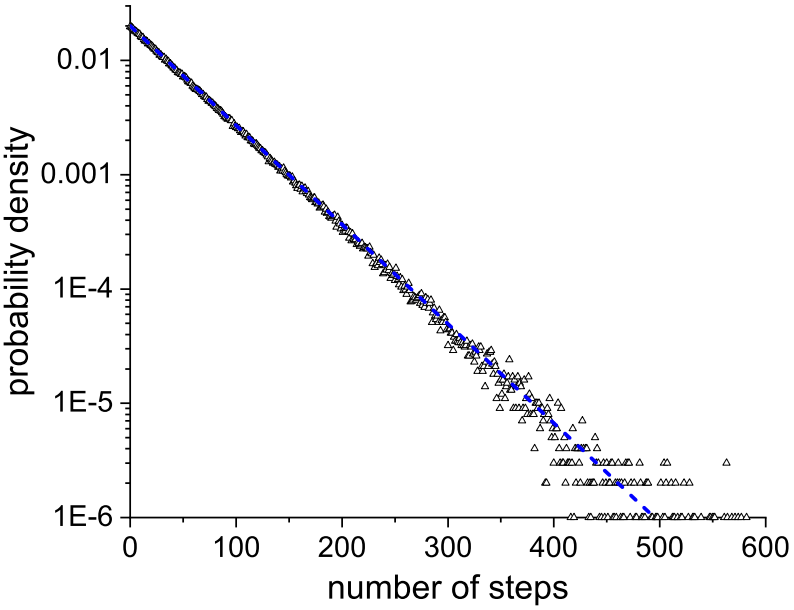

Figure S2: A simulated distribution of first-passage times. Random number generator with a uniform distribution of the numbers between 0 and 1 was used. In the ensemble of 10^6 trajectories, each trajectory was started from a random number and proceeded through the numbers until the value of 0.7 ± 0.01 was achieved. The label corresponds to the simulated trajectories, and the blue dashed line shows an exponential fit to the simulate distribution with the decay rate of 50.0.

3 Friction Constant $\gamma = 10M/\tau$

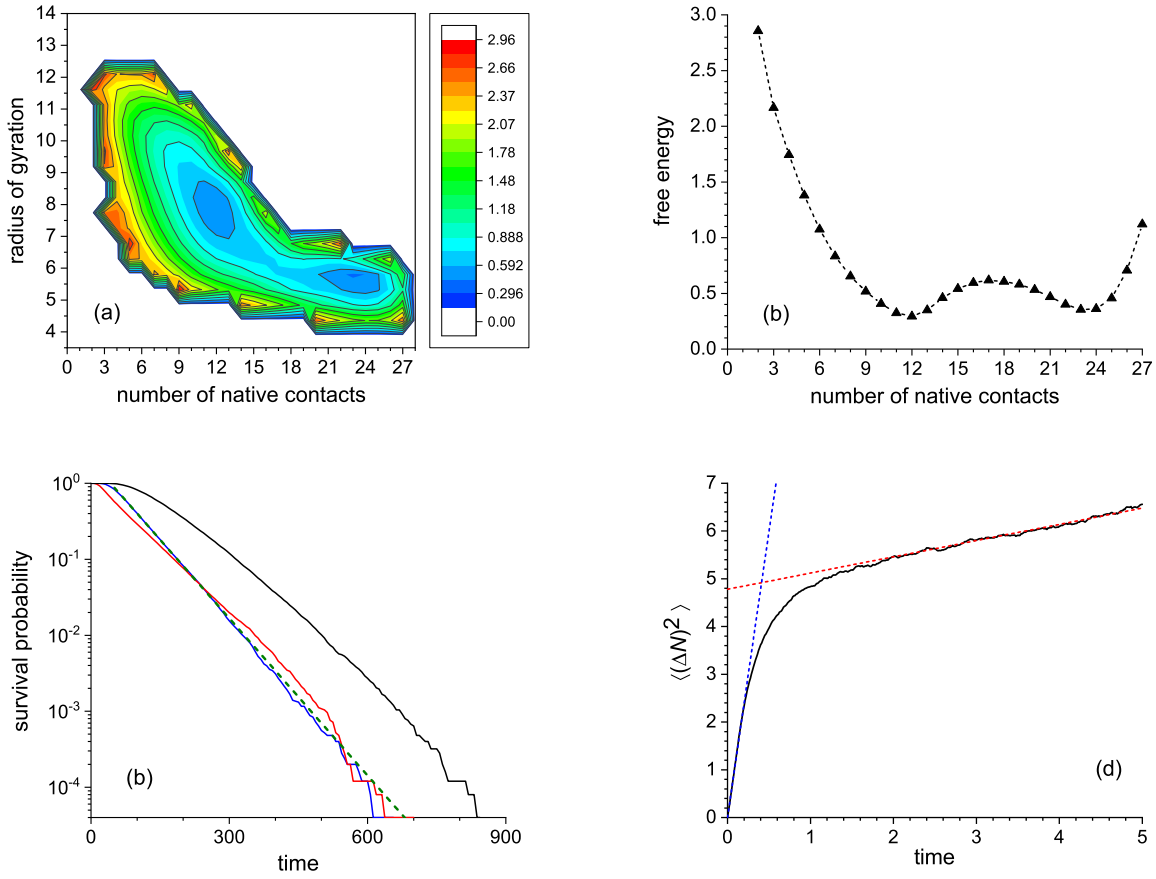

Figure S3: $T = 0.15$. (**a**) The free energy surface $F(N_{\rm nat}, R_{\rm g})$, and (**b**) free energy profile $F(N_{\rm nat})$. (**c**) First-passage time distributions in the form of survival probabilities: the U $\rightarrow$ NL trajectories (blue), the NL $\rightarrow$ N trajectories (red), and the U-N trajectories (black); the dashed green line denotes an exponential fit to the U $\rightarrow$ NL distribution. (**d**) The time-dependent mean-square deviation from the transition state in the number of native contacts (black curve); the blue and red dashed lines are the linear fits to the curve for short and long times, respectively.

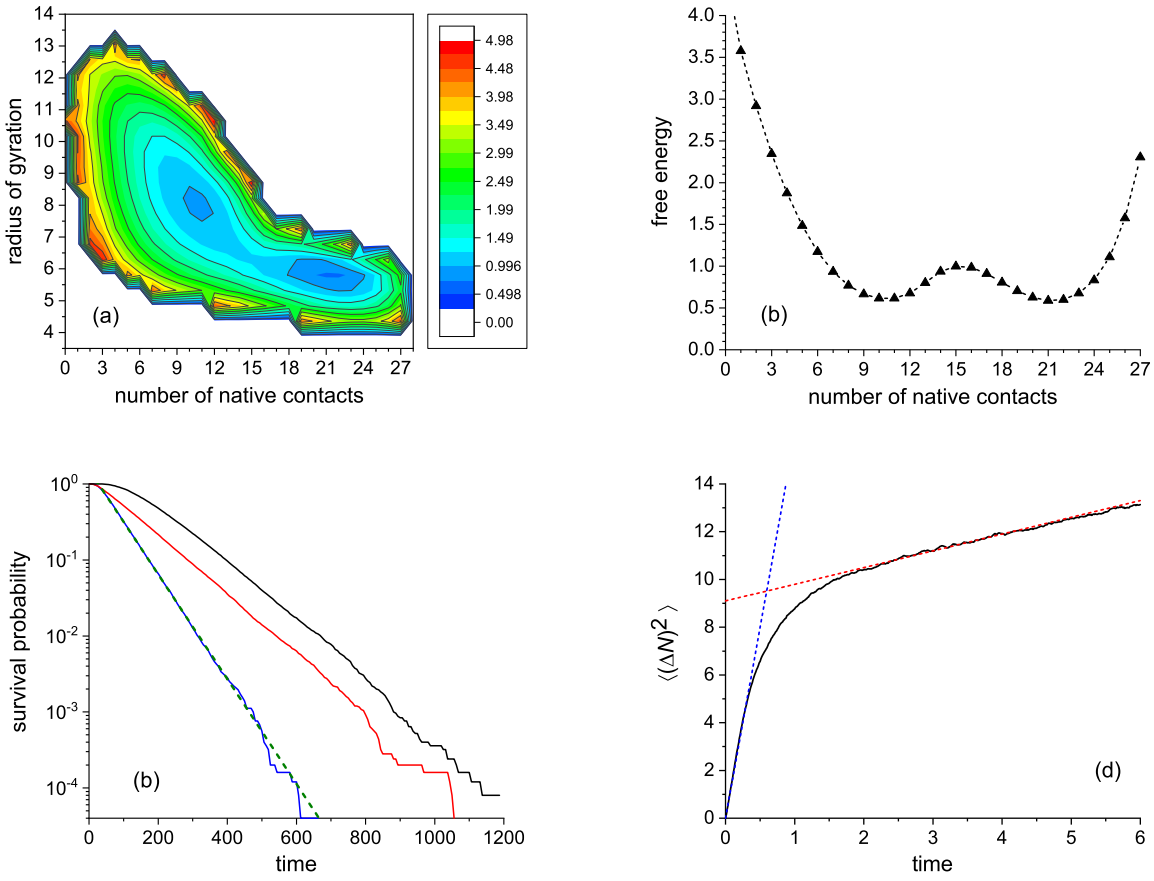

Figure S4: $T = 0.25$. The notations are as in Fig. S3.

Table S1: Parameters to calculate the U $\rightarrow$ NL transition time with the Kramers formula

T	0.1	0.15	0.2	0.25	0.3
ΔF	0.26	0.33	0.36	0.38	0.34
$F''_{\mathrm{U}}{}^{\underline{\mathrm{a}}}$	0.27	0.29	0.27	0.23	0.24
$F''_{\mathrm{U}}{}^{\underline{\mathrm{b}}}$	0.29	0.29	0.25	0.23	0.23
$F''_{\mathrm{TS}}{}^{\underline{\mathrm{a}}}$	0.14	0.19	0.21	0.26	0.32
$F''_{\mathrm{TS}}{}^{\underline{\mathrm{b}}}$	0.10	0.19	0.22	0.28	0.32
D_{TS}	4.5	6.0	6.5	8.0	9.0

$^{\underline{\mathrm{a}}}$ from the polynomial approximation.
$^{\underline{\mathrm{b}}}$ calculated as the three-point finite difference.

Table S2: Comparison of folding times

T	0.1	0.15	0.2	0.25	0.3
$\langle t_{\mathrm{U}\rightarrow\mathrm{NL}}\rangle^{\underline{\mathrm{a}}}$	93.0	63.0	59.0	63.0	63.0
$\langle t_{\mathrm{U}\rightarrow\mathrm{NL}}\rangle^{\underline{\mathrm{b}}}$	48.7	24.9	21.0	15.2	8.4
$\langle t_{\mathrm{U}\rightarrow\mathrm{NL}}\rangle^{\underline{\mathrm{c}}}$	144.6	103.4	92.5	89.9	85.0
$\langle t_{\mathrm{NL}\rightarrow\mathrm{N}}\rangle^{\underline{\mathrm{c}}}$	105.8	81.0	92.0	133.8	235.3
$\langle t_{\mathrm{U}\rightarrow\mathrm{N}}\rangle^{\underline{\mathrm{c}}}$	250.4	184.4	184.5	223.7	320.3

$^{\underline{\mathrm{a}}}$ calculated from the slope of the simulated U $\rightarrow$ NL decay curve.
$^{\underline{\mathrm{b}}}$ Kramers formula [Eq. (2), the main text] for the average values of F''_{U} and F''_{TS} (Table S1).
$^{\underline{\mathrm{c}}}$ simulated times.

4 Friction Constant $\gamma = 3M/\tau$

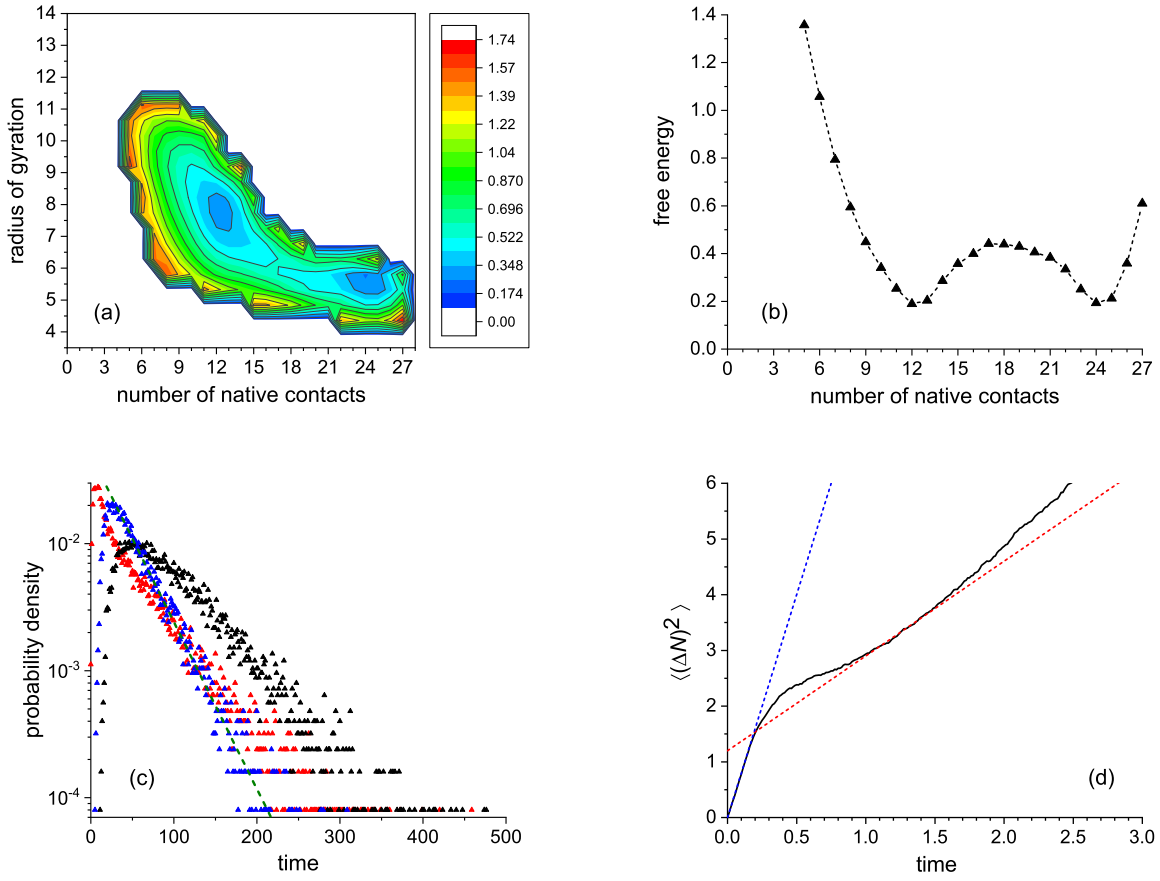

Figure S5: $T = 0.1$. (**a**) The free energy surface $F(N_{\rm nat}, R_{\rm g})$, and (**b**) free energy profile $F(N_{\rm nat})$. (**c**) First-passage time distributions: the U $\rightarrow$ NL trajectories (blue), the NL $\rightarrow$ N trajectories (red), and the U-N trajectories (black); the dashed green line denotes an exponential fit to the U $\rightarrow$ NL distribution. (**d**) The time-dependent mean-square deviation from the transition state in the number of native contacts (black curve); the blue and red dashed lines are the linear fits to the curve for short and long times, respectively.

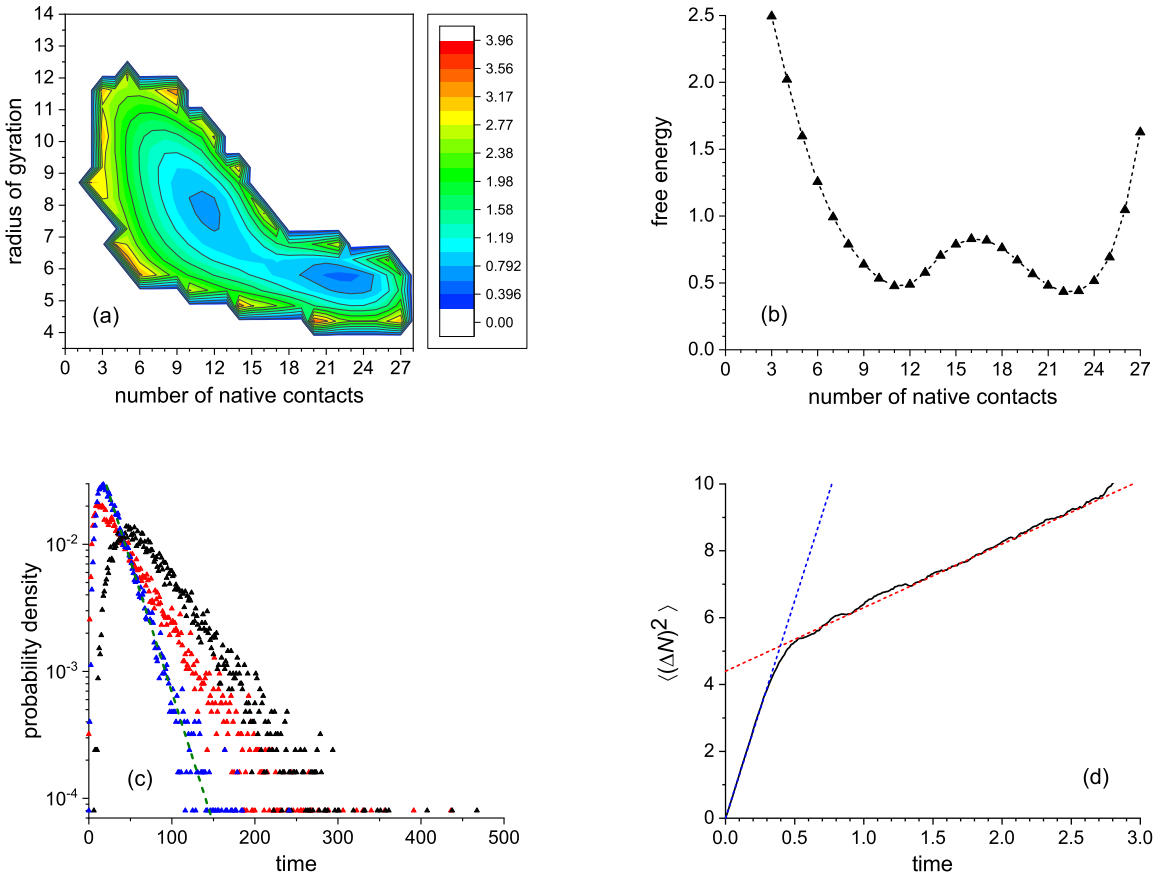

Figure S6: $T = 0.2$. The notations are as in Fig. S5.

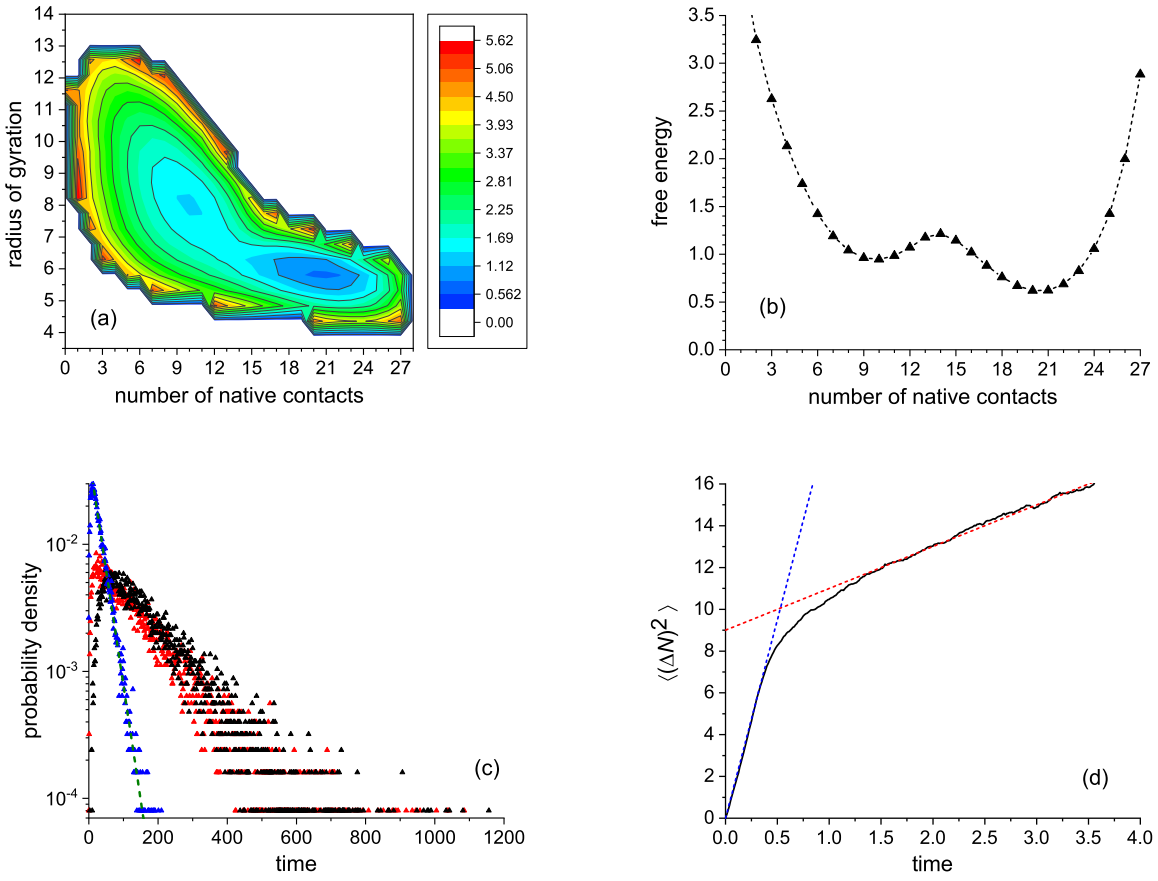

Figure S7: $T = 0.3$. The notations are as in Fig. S5.

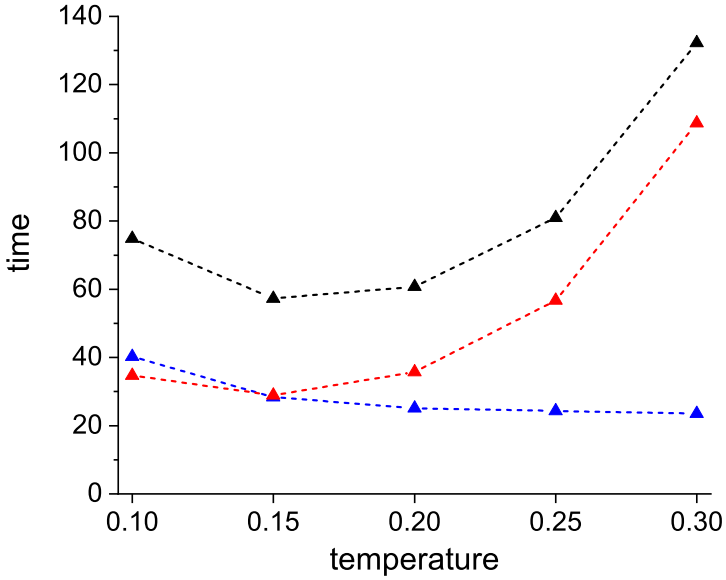

Figure S8: The simulated MFPT times: the U → NL times (blue), NL → N times (red), and U → N times (black). The dashed lines are to guide the eye.

Table S3: Parameters to calculate the U → NL transition time with the Kramers rate formula

T	0.1	0.15	0.2	0.25	0.3
ΔF	0.25	0.32	0.35	0.35	0.27
$F''_{\mathrm{U}}{}^{\underline{\mathrm{a}}}$	0.28	0.30	0.27	0.24	0.23
$F''_{\mathrm{U}}{}^{\underline{\mathrm{b}}}$	0.28	0.29	0.26	0.25	0.23
$F''_{\mathrm{TS}}{}^{\underline{\mathrm{a}}}$	0.13	0.17	0.23	0.28	0.33
$F''_{\mathrm{TS}}{}^{\underline{\mathrm{b}}}$	0.21	0.18	0.23	0.29	0.33
D_{TS}	4.0	5.8	6.5	8.5	9.5

$^{\underline{\mathrm{a}}}$ from the polynomial approximation.
$^{\underline{\mathrm{b}}}$ calculated as the three-point finite difference.

Table S4: Comparison of Folding Times

T	0.1	0.15	0.2	0.25	0.3
$\langle t_{\mathrm{U}\to\mathrm{NL}}\rangle^{\underline{a}}$	33.0	23.0	21.0	23.0	24.0
$\langle t_{\mathrm{U}\to\mathrm{NL}}\rangle^{\underline{b}}$	42.6	26.8	18.5	10.7	6.5
$\langle t_{\mathrm{U}\to\mathrm{NL}}\rangle^{\underline{c}}$	40.2	28.4	25.1	24.3	23.5
$\langle t_{\mathrm{NL}\to\mathrm{N}}\rangle^{\underline{c}}$	34.7	28.9	35.7	56.7	108.7
$\langle t_{\mathrm{U}\to\mathrm{N}}\rangle^{\underline{c}}$	74.9	57.3	60.8	81.0	132.2

$^{\underline{a}}$ calculated from the slope of the simulated U $\to$ NL decay curve.
$^{\underline{b}}$ Kramers formula [Eq. (2), the main text] for the average values of F''_{U} and F''_{TS} (Table S3).
$^{\underline{c}}$ simulated times.

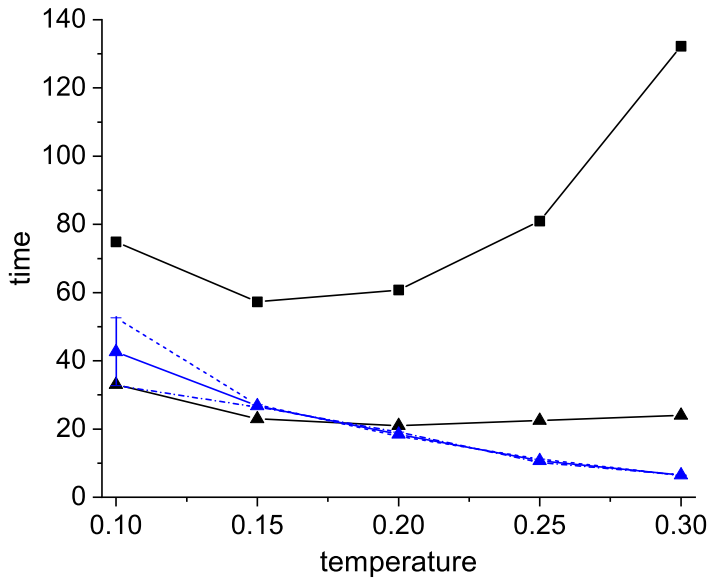

Figure S9: The black squares are for the $\langle t_{\mathrm{U}\to\mathrm{N}}\rangle$ times from simulations, the black triangles denote the $\langle t_{\mathrm{U}\to\mathrm{NL}}\rangle$ times calculated from the slopes of the simulated U $\to$ NL decay curves, and the blue triangles are for $\langle t_{\mathrm{U}\to\mathrm{NL}}\rangle$ times from Eq. (2) of the main text with the average values of F''_{U} and F''_{TS} (the dashed and dash-dotted blue lines indicate the results for F''_{U} and F''_{TS} obtained by the polynomial approximation of the FEP and calculated by finite-differences, respectively). In all cases, the lines are to guide the eye.

5 Friction Constant $\gamma = 50M/\tau$

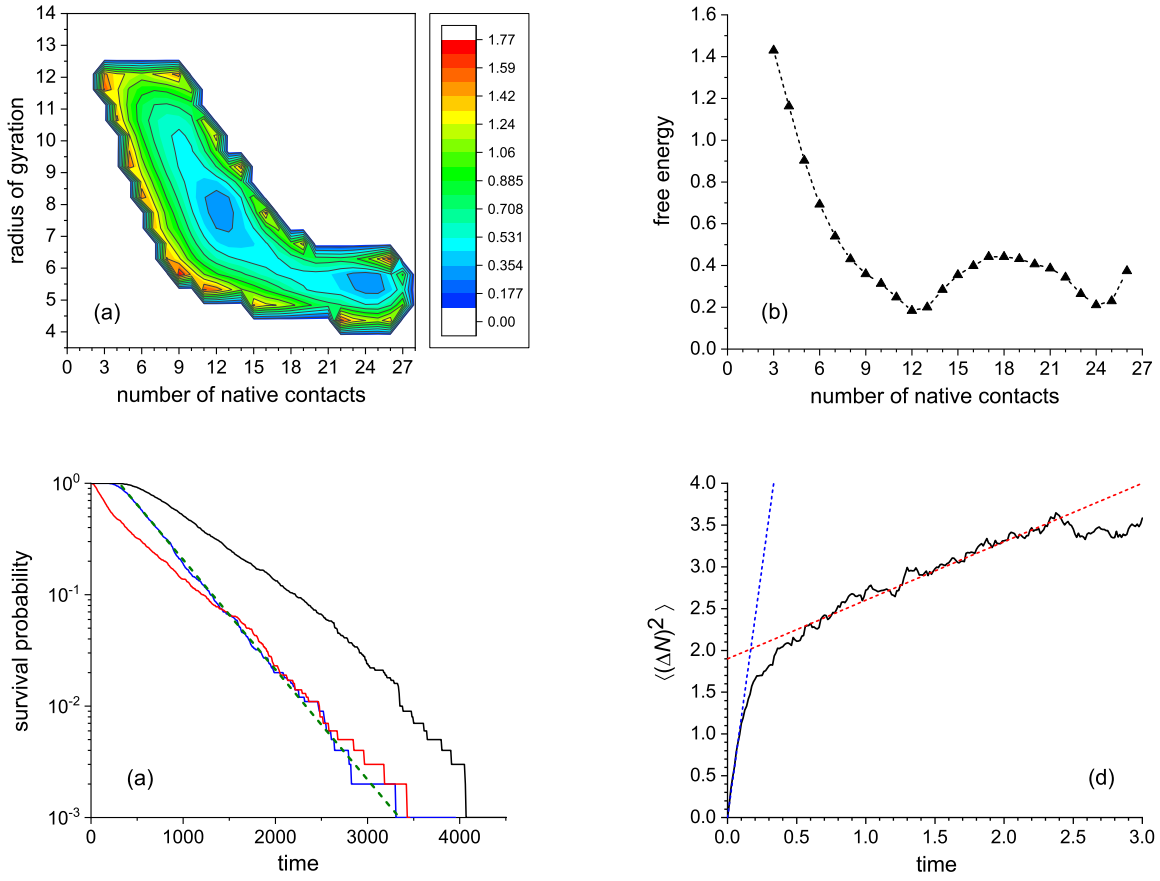

Figure S10: $T = 0.1$. (**a**) The free energy surface $F(N_{\rm nat}, R_{\rm g})$, and (**b**) free energy profile $F(N_{\rm nat})$. (**c**) First-passage time distributions in the form of survivaal probabilities: the U $\rightarrow$ NL trajectories (blue), the NL $\rightarrow$ N trajectories (red), and the U-N trajectories (black); the dashed green line denotes an exponential fit to the U $\rightarrow$ NL distribution. (**d**) The time-dependent mean-square deviation from the transition state in the number of native contacts (black curve); the blue and red dashed lines are the linear fits to the curve for short and long times, respectively.

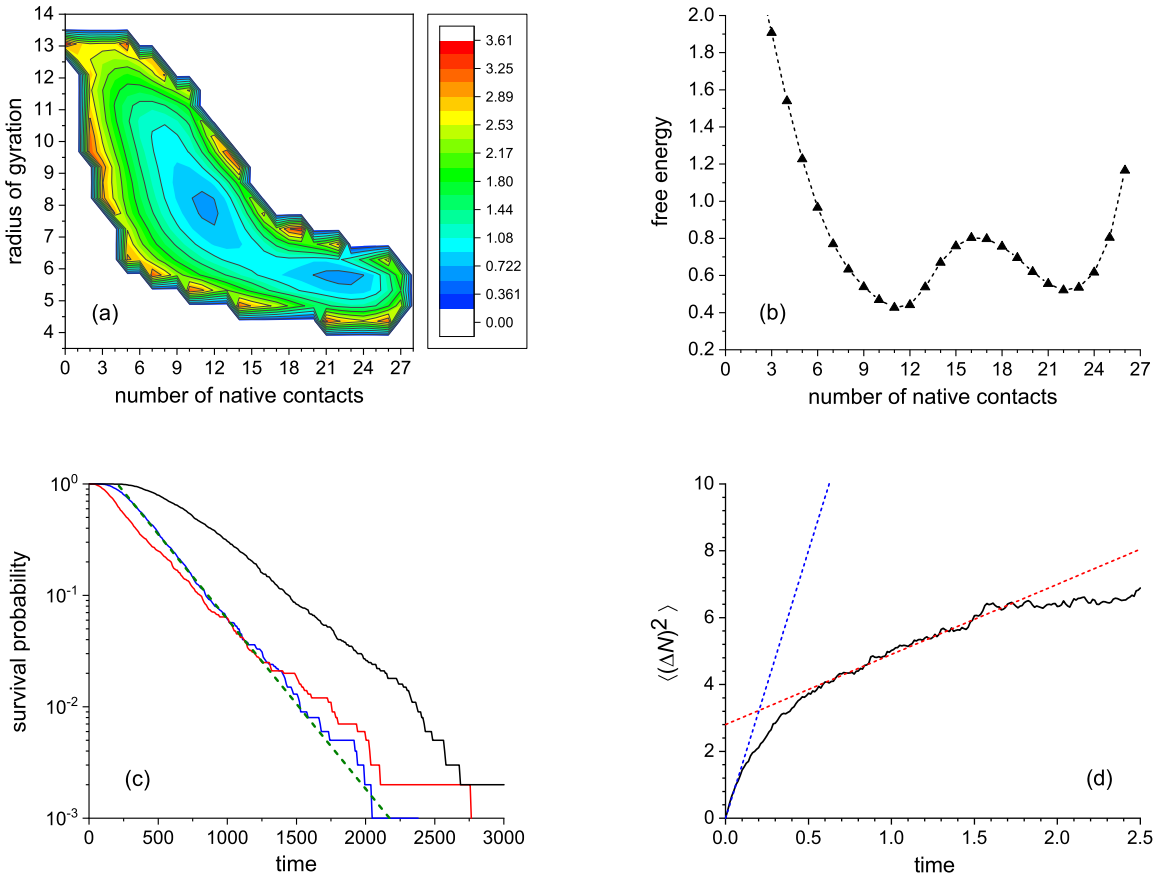

Figure S11: $T = 0.2$. The notations are as in Fig. S10.

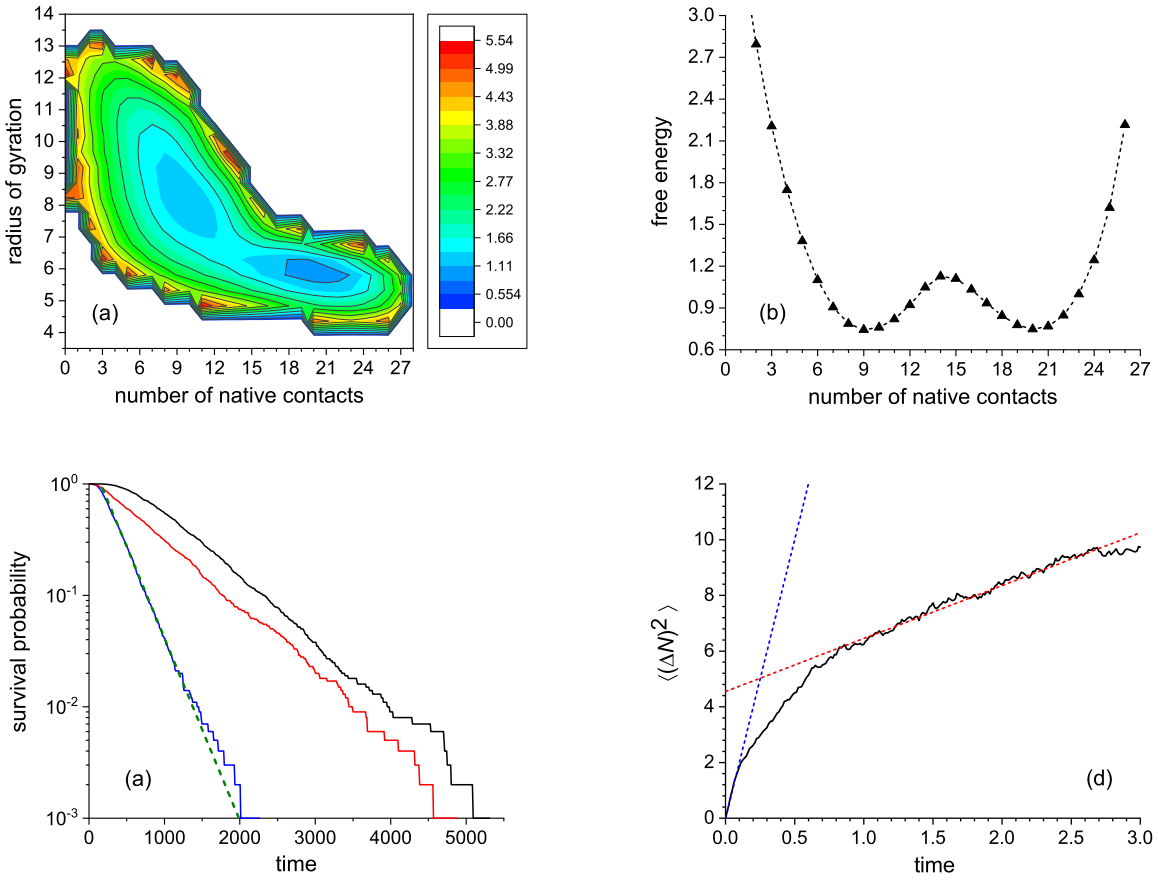

Figure S12: $T = 0.3$. The notations are as in Fig. S10.

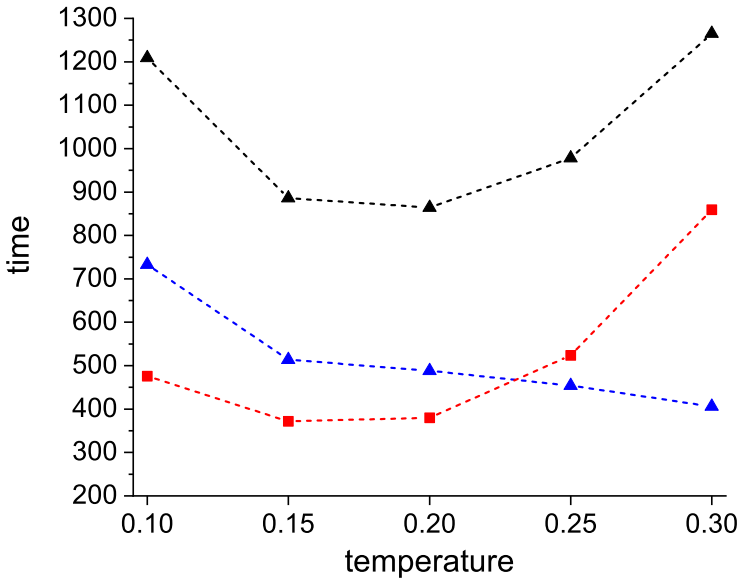

Figure S13: The simulated MFPT times: the U → NL times (blue), NL → N times (red), and U → N times (black). The dashed lines are to guide the eye.

Table S5: Parameters to calculate the U → NL transition time with the Kramers rate formula

T	0.1	0.15	0.2	0.25	0.3
ΔF	0.25	0.32	0.38	0.41	0.38
F''_{U}[a]	0.27	0.25	0.24	0.22	0.24
F''_{U}[b]	0.29	0.29	0.24	0.21	0.24
F''_{TS}[a]	0.14	0.16	0.21	0.26	0.28
F''_{TS}[b]	0.21	0.18	0.23	0.27	0.31
D_{TS}	0.35	0.50	1.1	0.90	0.95

[a] from the polynomial approximation.
[b] calculated as the three-point finite difference.

Table S6: Comparison of folding times

T	0.1	0.15	0.2	0.25	0.3
$\langle t_{\mathrm{U}\to\mathrm{NL}}\rangle^{\underline{\mathrm{a}}}$	440	280	285	310	265
$\langle t_{\mathrm{U}\to\mathrm{NL}}\rangle^{\underline{\mathrm{b}}}$	472	358	151	159	100
$\langle t_{\mathrm{U}\to\mathrm{NL}}\rangle_{\underline{\mathrm{c}}}$	733	514	484	454	406
$\langle t_{\mathrm{NL}\to\mathrm{N}}\rangle_{\underline{\mathrm{c}}}$	476	372	380	524	859
$\langle t_{\mathrm{U}\to\mathrm{N}}\rangle_{\underline{\mathrm{c}}}$	1209	886	864	978	1265

$^{\underline{\mathrm{a}}}$ calculated from the slope of the simulated U $\to$ NL decay curve.
$^{\underline{\mathrm{b}}}$ Kramers formula [Eq. (2), the main text] for the average values of F''_{U} and F''_{TS} (Table S5).
$^{\underline{\mathrm{c}}}$ simulated times.

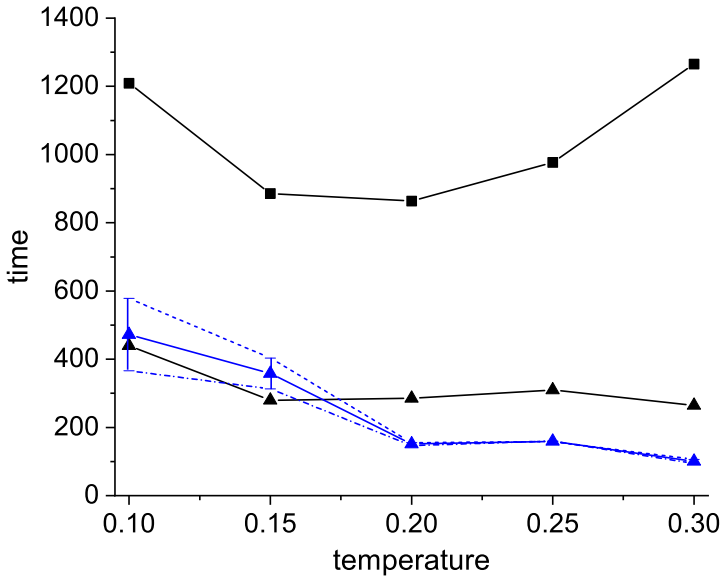

Figure S14: The black squares are for the $\langle t_{\mathrm{U}\to\mathrm{N}}\rangle$ times from simulations, the black triangles denote the $\langle t_{\mathrm{U}\to\mathrm{NL}}\rangle$ times calculated from the slopes of the simulated U $\to$ NL decay curves, and the blue triangles are for $\langle t_{\mathrm{U}\to\mathrm{NL}}\rangle$ times from Eq. (2) of the main text with the average values of $F_{\mathrm{U}}^{''}$ and $F_{\mathrm{TS}}^{''}$ (the dashed and dash-dotted blue lines indicate the results for $F_{\mathrm{U}}^{''}$ and $F_{\mathrm{TS}}^{''}$ obtained by the polynomial approximation of the FEP and calculated by finite-differences, respectively). In contrast to the cases of $\gamma = 3M/\tau$ and $\gamma = 10M/\tau$, where the diffusion coefficient was calculated from $\langle R^2(t)\rangle$ at short times, in the given case it was calculated at longer times where $\langle R^2(t)\rangle \sim t$ (the red dashed curves in Figs. S10 - S12). If the approximation of $\langle R^2(t)\rangle$ at short times is used (the blue dashed curves), the Kramers times are one order smaller. In all cases, the lines are to guide the eye.

6 Different Thresholds to Terminate the MD Trajectories. Friction Constant $\gamma = 10M/\tau$

.

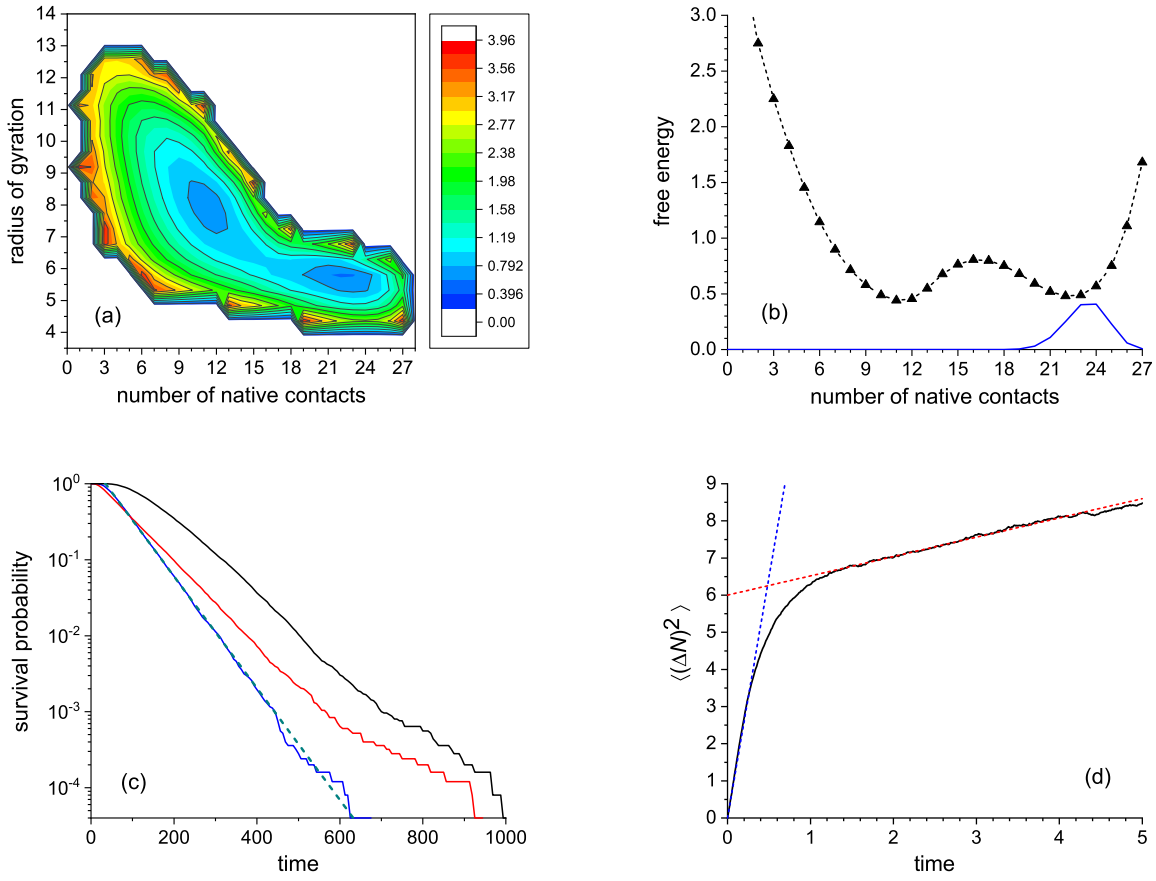

Figure S15: The trajectories were terminated as the RMSD from the native state was less than 1.0 Å; $T = 0.2$. (**a**) The free energy surface, and (**b**) free energy profile (black curve) with the normalized distributions of the protein states in the native-state ensemble (blue curve). (**c**) First-passage time distributions in the form of survival probabilities: the U $\rightarrow$ NL trajectories (blue), the NL $\rightarrow$ N trajectories (red), and the U-N trajectories (black); the dashed green line denotes an exponential fit to the U $\rightarrow$ NL distribution. (**d**) The time-dependent mean-square deviation from the transition state in the number of native contacts (black curve); the blue and red dashed lines are the linear fits to the curve for short and long times, respectively.

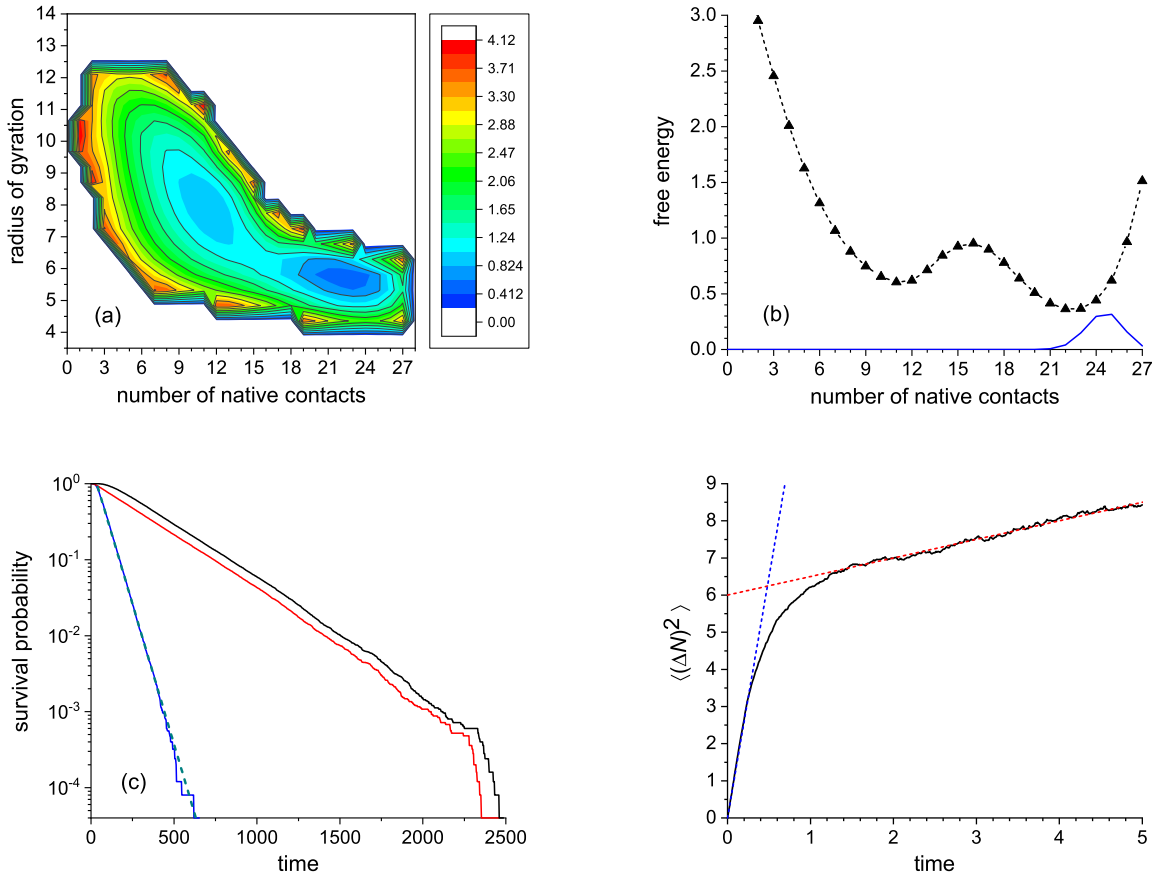

Figure S16: The trajectories were terminated as the RMSD from the native state was less than 0.65 Å; $T = 0.2$. The notations are as in Fig. S15.

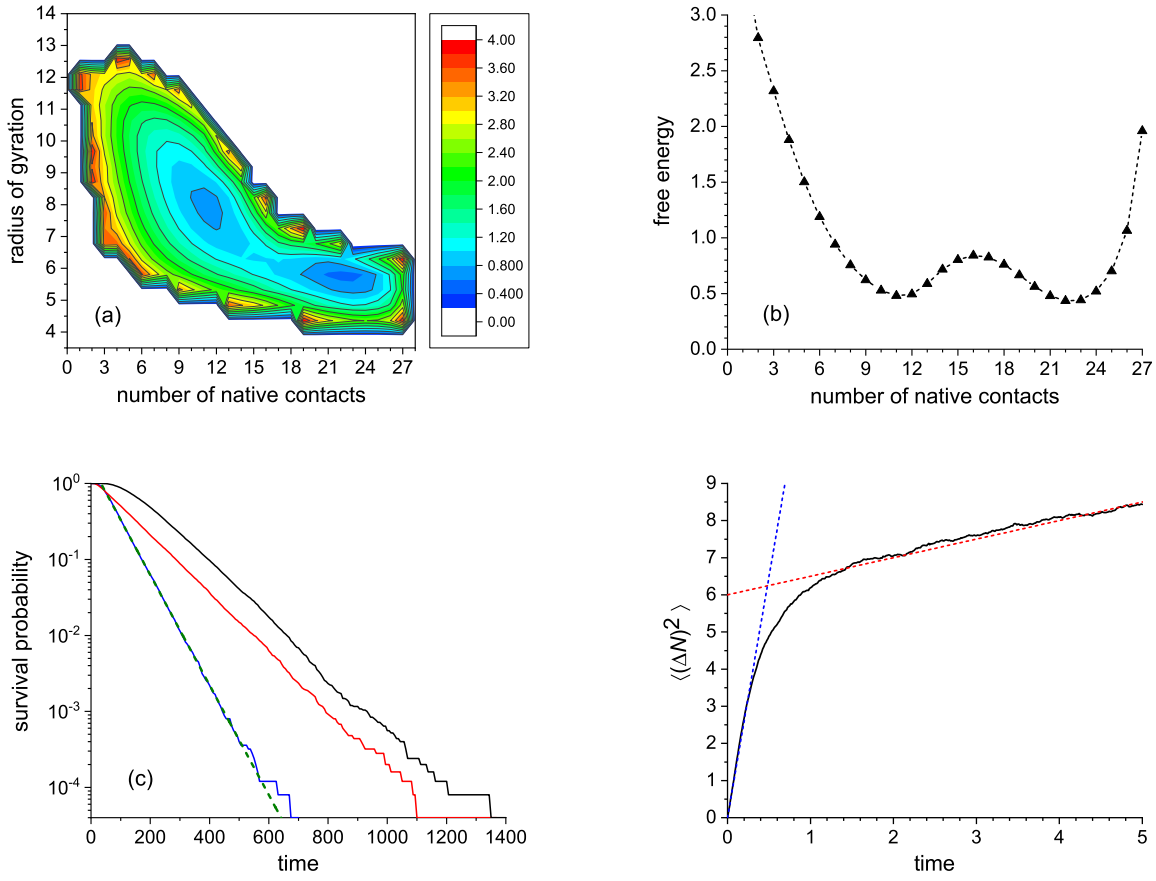

Figure S17: The trajectories were terminated as the number of native contacts $N_{\rm nat}$ was equal to the number of native contacts in the native state $N_{\rm nat}^{\rm NAT} = 27$; $T = 0.2$. The notations are as in Fig. S15.