PNAS

Supporting Information for

Fast protein folding is governed by memory-dependent friction

Benjamin A. Dalton, Cihan Ayaz, Henrik Kiefer, Anton Klimek, Lucas Tepper, and Roland R. Netz

Roland R. Netz.
E-mail: rnetz@physik.fu-berlin.de

This PDF file includes:

Supporting text
Figs. S1 to S15
Tables S1 to S7
SI References

Supporting Information Text

Contents

1 Simulations details and analysis 2

2 Free energy profiles for all proteins 5

3 The generalized Langevin equation and memory-kernel extraction 6

4 Memory kernels for all proteins 7

5 Effects of reaction-coordinate normalisation on friction and effective mass scaling 9

6 Additional analysis of folding times, friction, and barrier heights 11

7 Investigating discretisation effects with alanine-9 homo-peptide chain simulation 12

8 Prediction errors for barrier crossing times 14

9 Evaluating Markovian mean first-passage times: position-independent friction 16

10 Non-Markovian reaction-rate theory and Grote-Hynes theory 17

11 Comparison of various reaction coordinates 23

12 Evaluating Markovian mean first-passage times: position-dependent friction 26

13 Additional confirmation of the linear friction non-Markovian model for the Q reaction coordinate 28

1. Simulations details and analysis

All simulation data in the main manuscript was generated by the group of David E. Shaw and originally published by Lindorff-Larsen *et. al.* (1). The simulations were performed using the Anton special-purpose computer (2), a modified CHARMM force-field (3), and a modified TIP3P water model, with an integration time step of $dt = 2.5$ fs. Protein configurations were saved to file every 200 ps, reducing the time resolution of the trajectory data relative to the simulation time discretization. Effects due to low time-resolution discretization are discussed in Section 7. We only consider 8 out of the 12 proteins presented in the original publication, as BBA, BBL, Protein B, and Homeodomain either do not exhibit distinct energy barriers separating folded and unfolded states when projected onto the Q reaction coordinate, or have barriers that are less than 1 k_BT. Each protein was simulated at a unique temperature, identified as the appropriate melting temperature, which was chosen to maximize the number of folding and unfolding events. Table S1 shows relevant system parameters, including the total simulation time for each protein t_{sim}, the simulation temperature T, the number of amino acid residues N, the total number of folding events N_f and unfolding events N_u, all of which are taken from in the original publication (1). Additionally, we show properties of the free energy barriers, including the barrier heights faced by the folded and unfolded states, U_0^f and U_0^u, the distances from the unfolded and folded states minima to the barrier tops, L_f and L_u, and the effective reaction coordinate mass m.

Table S2 shows the time scales measured in our analysis, including the first-moment memory times τ_{mem}, inertial times τ_m, diffusion times in the folded and unfolded domains, τ_D^f and τ_D^u, and mean first-passage times for folding and unfolding extracted from MD simulations, $\tau_{MFP}^{MD,f}$ and $\tau_{MFP}^{MD,u}$, as well as the extracted mean transition-path times for folding and unfolding τ_{MTP}^f and τ_{MTP}^u.

Table S1. Table of relevant system parameters and properties

	t_{sim} [μs]	T[K]	N	N_f	N_u	U_0^f [k_BT]	U_0^u [k_BT]	L_f	L_u	m ($\times 10^9$) [unm^2]
Chignolin	106	300	10	39	38	6.0	1.6	0.042	0.53	0.17
Trp-Cage	208	290	20	12	12	3.9	3.1	0.075	0.22	0.31
Villin	125	360	35	34	34	2.0	1.6	0.043	0.17	0.62
WW Domain	1137	360	35	12	11	6.3	2.7	0.14	0.29	0.77
NTL9	2936	355	39	17	14	6.7	2.3	0.16	0.23	1.9
Protein-G	1154	350	56	12	13	4.5	2.3	0.10	0.16	2.3
α_3D	707	370	73	12	12	3.2	1.7	0.090	0.14	3.4
λ-Repressor	643	350	80	10	12	2.1	2.1	0.047	0.24	3.7

Table S2. Table of relevant time scales extracted from MD trajectories. All time scales are given in units of μs.

	τ_{mem}	τ_{m} $(\times 10^{-5})$	$\tau_{\mathrm{D}}^{\mathrm{f}}$	$\tau_{\mathrm{D}}^{\mathrm{u}}$	$\tau_{\mathrm{MFP}}^{\mathrm{MD,f}}$	$\tau_{\mathrm{MFP}}^{\mathrm{MD,u}}$	$\tau_{\mathrm{MTP}}^{\mathrm{f}}$ $(\times 10^{-2})$	$\tau_{\mathrm{MTP}}^{\mathrm{u}}$ $(\times 10^{-2})$
Chignolin	0.042	5.4	0.31	0.002	0.18	0.18	1.1	0.06
Trp-Cage	0.35	0.46	1.3	0.15	2.1	0.26	2.0	0.60
Villin	0.072	1.2	0.50	0.033	1.2	0.03	2.7	0.20
WW Domain	4.2	0.24	9.0	2.1	2.6	6.6	4.9	3.1
NTL9	0.44	0.48	7.1	3.4	2.2	37.8	1.2	6.6
Protein-G	0.33	0.32	6.4	2.4	9.5	10.1	12.0	8.3
α_3D	3.7	0.15	15.5	5.9	5.1	4.3	16.0	6.9
λ-Repressor	0.19	0.48	15.1	0.57	27.0	0.23	108.0	1.3

In Table S3, we compare our analysis results with those of Lindorff-Larsen *et. al.* Two important differences should be noted. Firstly, we use a slightly different reaction coordinate, which contains a sigmoidal function contribution for each residue pair, but with different exponents. Whereas the exponent of Best *et. al*, and which we use, is $\beta(s_{ij}(t) - \gamma s_{ij}^0)$ (see materials and methods), Lindorff-Larsen *et. al.* use $10(s_{ij}(t) - (s_{ij}^0 + 1))$. Secondly, our method for calculating reaction times differs from that of Lindorff-Larsen *et. al.*, who define the mean waiting time in the unfolded state as the folding reaction time and vice versa. Additionally, their definition of folded and unfolded states, based on the dual cut-off method of Northrup *et. al.* (4), further affects their calculation of transition path times. Therefore, we cannot directly equate our evaluations of reaction times and transition path times with those reported by Lindorff-Larsen *et. al.*. Throughout our paper, we only discuss folding and unfolding events as those that pass from their respective state minimum to the barrier top. To compare to Lindorff-Larsen *et. al.*, we modify our analysis to consider a folding transition as one that connects the unfolded state minimum to the folded state minimum, and an unfolding transition is one that connects the folded state minimum to the unfolded state minimum. Thus, for the comparison in Table S3, we introduce τ^{fold} and τ^{unfold} as our folding and unfolding reaction times, as extracted from MD simulations, that connect the state-minima, and τ^{TP} and as the transition path times. The results from Lindorff-Larsen *et. al.* are labelled the same way, except that they appear with the sub-script LL. This redefinition of folding and unfolding reactions applies only to the values presented in Table S3 and has no relevance elsewhere in our analysis. Overall, our results, including free-energy profiles and our determined native states, are in good agreement with those presented by Lindorff-Larsen *et. al.*.

Table S3. Comparison of reaction and transition path times to those of Lindorff-Larsen *et. al.* (1), and those reported for experiments. The different reaction times are described in the text. All times are presented in units of μs. Comparison to experimental values are done for folding times $\tau_{\mathrm{exp}}^{\mathrm{fold}}$.

	$\tau_{\mathrm{LL}}^{\mathrm{fold}}$	$\tau_{\mathrm{LL}}^{\mathrm{unfold}}$	$\tau_{\mathrm{LL}}^{\mathrm{TP}}$	τ^{fold}	τ^{unfold}	τ^{TP}	$\tau_{\mathrm{exp}}^{\mathrm{fold}}$
Chignolin	0.6	2.2	0.04	0.4	1.9	0.012	-
Trp-Cage	14	3	0.22	12.0	2.3	0.097	-
Villin	2.8	0.9	0.27	3.1	0.4	0.074	1.4 μs (300 K) (5)
WW Domain	21	80	0.5	38.0	44.0	0.35	5.7 μs (363 K) (6)
NTL9	29	175	0.9	20.0	104.0	0.58	730 μs (298 K, pH=5.4) (7)
Protein-G	56	37	1.8	63.0	28.0	0.62	57 μs (295 K) (8)
α_3D	27	31	0.9	32.0	25.0	0.91	12.5 μs (370 K, pH=2.2) (9)
λ-Repressor	49	13	3.1	33.0	8.6	1.4	10 μs (350 K) (10)

The experimental folding times $\tau_{\text{exp}}^{\text{fold}}$ in Table S3 are taken from the same sources as those quoted by (1). The temperatures and, where relevant, the pH are also given, in order to distinguish from the simulation values. In the following, we provide some additional experimental details.

- Villin: Folding rate is given as 0.7 μs^{-1}, measured using laser temperature-jump experiments (5).
- WW Domain: The refolding time of the GTT variant is reported to be 5.7 μs at 363 K, measured using temperature jump experiments (6).
- NTL9: The NTL9 simulations are for the K12M variant. (7) used jump stopped-flow measurements to determine the folding and unfolding rates of various NTL9 mutants over a range of urea concentrations. In (1), it is stated that the folding time is 730 μs, evaluated by extrapolating to 0 M urea for the K12M variant. Note that in these experiments, the pH and the temperature are significantly different from the values used in the simulation.
- Protein G: Simulations were performed for the N37A/A46D/D47A mutant of the NuG2 protein. (8) *et. al.* used stopped-flow experiments at 295 K to measure folding plus unfolding rates over a range of GuHCl concentrations. For NuG2, they quote two values for folding rates: $k_f = 160$ s^{-1} at 2.5 M GuHCl and $k_f = 1050$ s^{-1} at 1.5 M GuHCl. Extrapolation to 0 M gives a folding rate of approximately $k_f = 17500$ s^{-1}, and hence a folding time of approximately 57 μs.
- α_3D: Zhu *et. al.* provide a model for folding and unfolding times of α_3D, parametrized by measurements from temperature jump experiments. For a temperature of 370 K, we can read off a folding time of 12.5 μs. This data is only available at a low pH of 2.2.
- λ-repressor: From Yang *et. al.* (10), the λD14A variant has a folding rate of 0.1 μs^{-1} at 350 K., measured using laser-induced temperature-jump experiments.

2. Free energy profiles for all proteins

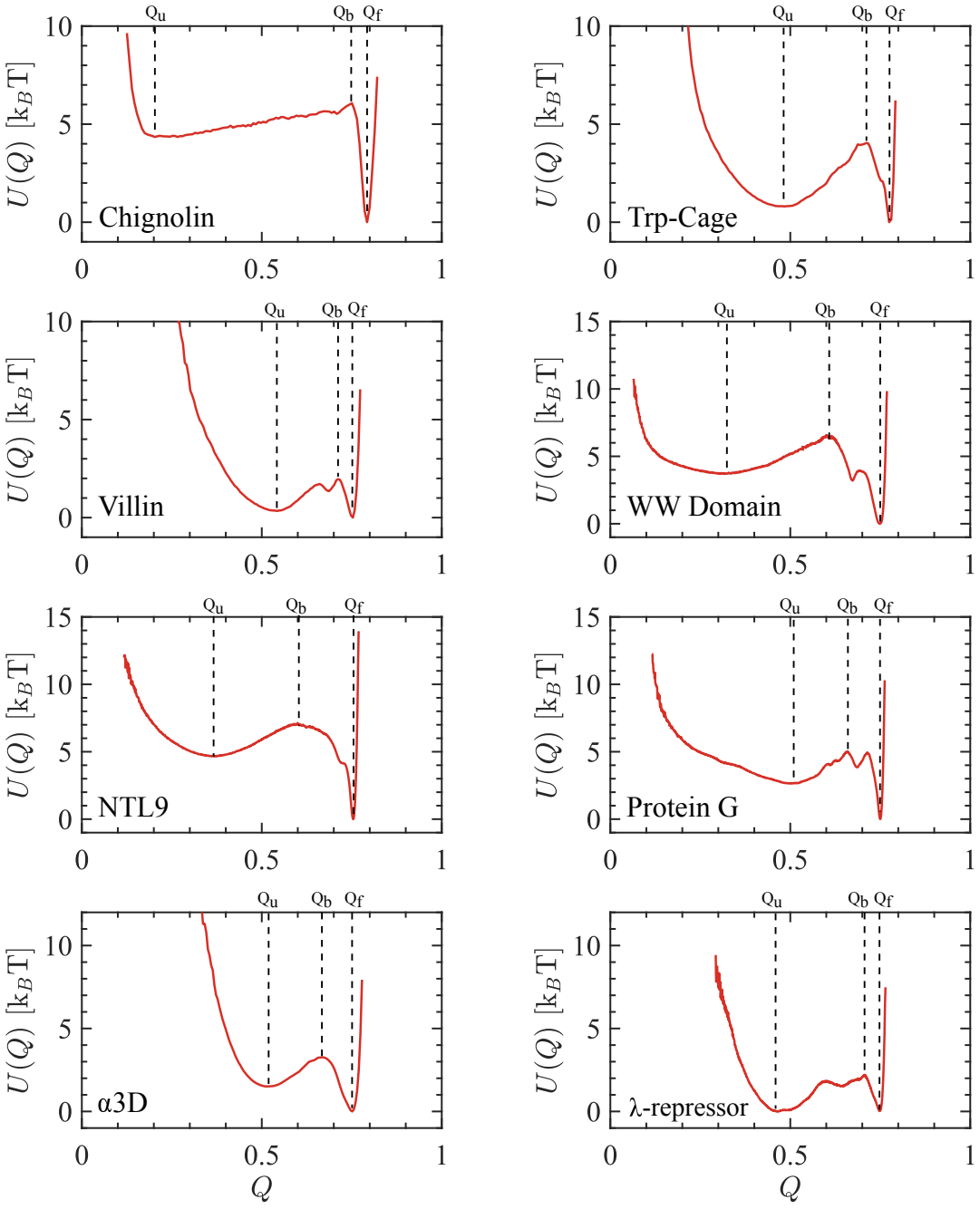

Fig. S1. Free energy profiles for the fraction of native contacts reaction coordinate, evaluated for the eight proteins. The unfolded Q_{u}, barrier Q_{b}, and folded Q_{f} states are indicated for each protein. Free energy profiles are calculated as $U(Q) = -k_{\mathrm{B}}T\log[\rho(\mathrm{Q})]$, where $\rho(Q)$ is the probability density for a given trajectory. Probability densities are calculated using normalized histograms, with 750 bins between Q_{min} and Q_{max} for each protein.

3. The generalized Langevin equation and memory-kernel extraction

For each protein, we project the all-atom trajectories onto the fraction of native contacts reaction coordinate $Q(t)$ and hence describe the trajectory of $Q(t)$ in terms of an approximate 1D GLE

$$m\ddot{Q}(t) = -\int_0^t \Gamma(t-t')\dot{Q}(t')dt' - \nabla U\big[Q(t)\big] + F_R(t). \quad [1]$$

The GLE is approximate in the sense that the memory kernel $\Gamma(t)$ is independent of position, as described in a recent publication (11). $F_R(t)$ is the random force term, which has a zero mean $\langle F_R(t)\rangle = 0$, and satisfies the fluctuation-dissipation theorem $\langle F_R(t)F_R(t')\rangle = k_{\mathrm{B}}T\Gamma(t-t')$. $U(Q)$ is the potential of mean force, which is extracted uniquely for each protein according to $U(Q) = -k_{\mathrm{B}}T\log\big[\rho(Q)\big]$, where $\rho(Q)$ is the probability density over $Q(t)$. The total friction on a reaction coordinate is given by the converged plateau value of the integrated memory kernel,

$$\gamma = G(t\to\infty) = \int_0^\infty \Gamma(t)dt, \quad [2]$$

where $G(t) = \int_0^t \Gamma(t')dt'$ is the running integral of the memory kernel.

To extract $\Gamma(t)$ for each system, we use the running integral extraction scheme for generalized potentials. The details of this extraction scheme can be found in references (12) and (13). In short, we correlate Eq. 1 with the initial position of the reaction coordinate $Q(0)$,

$$m\Big\langle Q(0)\ddot{Q}(t)\Big\rangle = -\int_0^t \Big\langle \Gamma(t-t')Q(0)\dot{Q}(t')dt'\Big\rangle - \Big\langle Q(0)\nabla U\big[Q(t)\big]\Big\rangle + \Big\langle Q(0)F_R(t)\Big\rangle, \quad [3]$$

and with the initial velocity of the reaction coordinate $\dot{Q}(0)$,

$$m\Big\langle \dot{Q}(0)\ddot{Q}(t)\Big\rangle = -\int_0^t \Big\langle \Gamma(t-t')\dot{Q}(0)\dot{Q}(t')dt'\Big\rangle - \Big\langle \dot{Q}(0)\nabla U\big[Q(t)\big]\Big\rangle + \Big\langle \dot{Q}(0)F_R(t)\Big\rangle. \quad [4]$$

Due to the orthogonality relations used to derive the GLE in Eq. 1, both $\big\langle Q(0)F_R(t)\big\rangle = 0$ and $\big\langle \dot{Q}(0)F_R(t)\big\rangle = 0$. Thus, we can write Eqs. 3 and 4 in terms of the position-velocity and velocity-velocity correlation functions $C^{Q\dot{Q}}(t) = \big\langle Q(0)\dot{Q}(t)\big\rangle$ and $C^{\dot{Q}\dot{Q}}(t) = \big\langle \dot{Q}(0)\dot{Q}(t)\big\rangle$, respectively, as well as the correlations between the reaction coordinate and the PMF gradients $C^{Q\nabla U}(t) = \big\langle Q(0)\nabla U\big[Q(t)\big]\big\rangle$, and the velocity of the reaction coordinate and the PMF gradients $C^{Q\nabla U}(t) = \big\langle \dot{Q}(0)\nabla U\big[Q(t)\big]\big\rangle$

$$m\frac{d}{dt}C^{Q\dot{Q}}(t) = -\int_0^t \Gamma(t')C^{Q\dot{Q}}(t-t')dt' - C^{Q\nabla U}(t), \quad [5]$$

$$m\frac{d}{dt}C^{\dot{Q}\dot{Q}}(t) = -\int_0^t \Gamma(t')C^{\dot{Q}\dot{Q}}(t-t')dt' - C^{\dot{Q}\nabla U}(t). \quad [6]$$

We integrate Eq. 6 in the time domain and obtain an equation in terms of $G(t) = \int_0^t \Gamma(t')dt'$,

$$mC^{\dot{Q}\dot{Q}}(t) - mC^{\dot{Q}\dot{Q}}(0) = -\int_0^t G(t-t')C^{\dot{Q}\dot{Q}}(t'')dt'' + C^{Q\nabla U}(t) - C^{Q\nabla U}(0). \quad [7]$$

Using the identity that $\frac{d}{dt}C^{Q\dot{Q}}(t) = C^{Q\ddot{Q}}(t) = -C^{\dot{Q}\dot{Q}}(t)$, we evaluate Eq. 5 at t=0, and obtain $mC^{\dot{Q}\dot{Q}}(0) = C^{Q\nabla U}(0)$. It follows that

$$\frac{C^{\dot{Q}\dot{Q}}(t)}{C^{\dot{Q}\dot{Q}}(0)}C^{Q\nabla U}(0) = C^{Q\nabla U}(t) - \int_0^t G(t-t')C^{\dot{Q}\dot{Q}}(t'')dt''. \quad [8]$$

Eq. 8 can be discretized using the trapezoidal rule for numerical integration. Using $G(0) = 0$, we arrive at the following numerical extraction scheme, which we use to extract the discrete representation of $G(t)$ directly from the trajectory of a given reaction coordinate,

$$G_i = \begin{cases} 0, & i = 0 \\ \frac{2}{\Delta t C_0^{\dot{Q}\dot{Q}}}\left[C_1^{\nabla UQ} - \frac{C_0^{\nabla UQ}}{C_0^{\dot{Q}\dot{Q}}}C_1^{\dot{Q}\dot{Q}}\right], & i = 1 \\ \frac{2}{\Delta t C_0^{\dot{Q}\dot{Q}}}\left[C_i^{\nabla UQ} - \frac{C_0^{\nabla UQ}}{C_0^{\dot{Q}\dot{Q}}}C_i^{\dot{Q}\dot{Q}} - \Delta t\sum_{j=1}^{i-1} G_j C_{i-j}^{\dot{Q}\dot{Q}}\right], & i > 1. \end{cases} \quad [9]$$

$C_i^{\dot{Q}\dot{Q}}$ and $C_i^{\nabla UQ}$ are the discretized representations of the velocity-velocity correlation function and the correlations between the reaction coordinate and the gradients of the PMF, respectively.

4. Memory kernels for all proteins

We extract the memory kernels for each protein for the fraction of native contacts reaction coordinate using the method given in Eq. 9. The results for the extractions are shown in Fig. S2. We expect that, due to the intrinsically low time resolution of the MD data, there exist decay modes that cannot be resolved by our extraction methods.

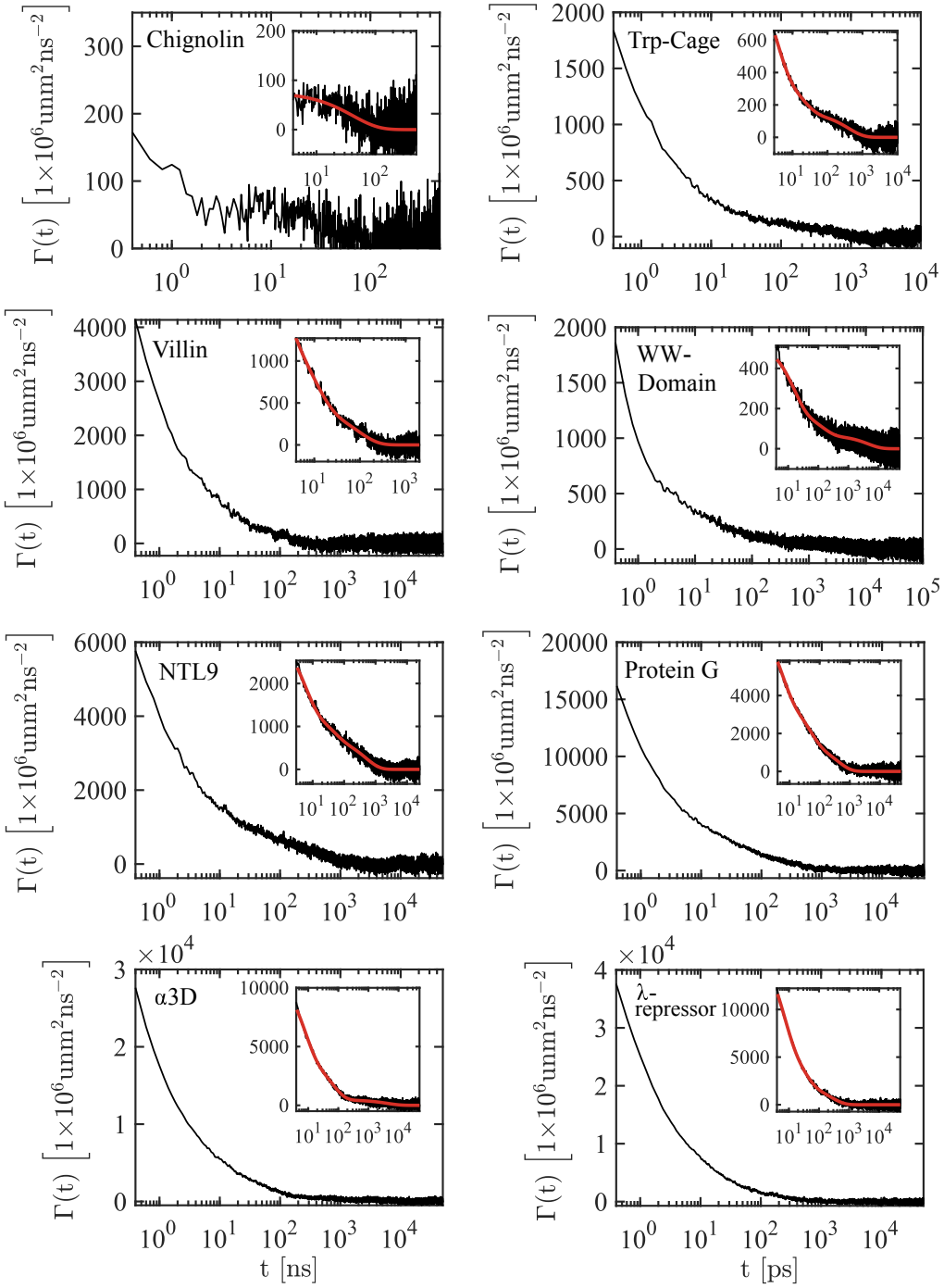

Fig. S2. Time-dependent memory kernels extracted via Eq. 9 for the 8 proteins. The main panel for each protein shows the full memory kernel. The inserts show the segment of the memory kernel that is used for the fit of the exponential series to the data, in addition to the result for the fitting (red curve). For Chignolin, the exponential series contains two terms. For all other proteins, there are three terms in the series.

In the inset for each panel, we show results for a three-component exponential fit to the memory kernels, and the corresponding region of fitting (note that for the case of chignolin, we only fit two exponential components). The fitting parameters are shown

in Table S4, along with the total friction γ and the 1st-moment memory time τ_{mem} for each protein. The total friction follows from $\gamma = \sum_{i=1}^{M} \gamma_i$.

Table S4. Table of fitting parameters for the memory kernels of the eight proteins. Memory kernels are given by $\Gamma(t) = \sum_{i=1}^{M} \gamma_i \exp(-\mathrm{t}/\tau_\mathrm{i})/\tau_\mathrm{i}$, where $M = 2$ for Chignolin, and $M = 3$ for all other proteins. Time scales τ_i are given in units of ns and amplitudes are in units of $[1\times10^6 \mathrm{unm^2 ns^{-1}}]$.

	γ_1 $(\times10^5)$	τ_1 $(\times10^3)$	γ_2 $(\times10^3)$	τ_2 $(\times10^1)$	γ_3 $(\times10^3)$	τ_3	γ $(\times10^5)$	τ_{mem} $(\times10^3)$
Chignolin	0.0318	0.041	0.0798	0.01	-	-	0.032	0.042
Trp-Cage	0.609	0.39	4.8	1.7	2.22	2.9	0.68	0.36
Villin	0.408	0.092	10.5	0.95	2.2	1.4	0.53	0.072
WW Domain	2.98	4.5	16.1	13.0	5.1	10.6	3.2	4.21
NTL9	3.55	0.49	28.7	4.2	10.1	3.3	3.9	0.45
Protein-G	5.88	0.41	113.5	4.3	22.4	2.8	7.3	0.33
α_3D	19.6	4.3	239.0	6.3	53.6	4.9	22.6	3.73
λ-Repressor	5.42	0.26	148.5	2.8	60.3	2.2	7.6	0.19

5. Effects of reaction-coordinate normalisation on friction and effective mass scaling

In Fig. 1E of the main manuscript, we show that for the fraction of native contacts reaction coordinate Q, the friction scales according to $\gamma/k_{\mathrm{B}}T \sim N^{2.8}$, where N is the number of residues. Q is a reaction coordinate that depends on the sum of individual atomic distances and is, furthermore, normalized by the number of native contacts N_{nc}. To investigate the role of normalisation on the N-scaling of the total friction γ, and the effective mass m of a reaction coordinate, we consider a linear reaction coordinate that is the sum of n Cartesian distances:

$$s(t) = \sum_{i=1}^{n} d_i(t), \qquad [10]$$

where $d_i(t) = |\mathbf{d}_i(t)|$ is the distance between two atomic positions, such as the distance between two C_α atoms in a protein chain, and $s(t)$ is the sum of all such distances. Note that the index i here implies a pair and there are n such pair separation distances. s is the non-normalized form of the reaction coordinate. We write the normalized form according to the following rescaling: $\tilde{s} = s/n$.

We know from Ayaz *et. al.* (11) that the generalized mass for a reaction coordinate corresponding to the scalar distance between two positions m_d is constant, satisfying the equipartition theorem $\langle \dot{d}_i^2 \rangle = k_{\mathrm{B}}T/m_d$. We note that this is for a single pair distance and that m_d is assumed to be the same for all pairs. Summing over all pairs results in $\sum_{i=1}^{n} \langle \dot{d}_i^2 \rangle = n k_{\mathrm{B}}T/m_d$. The collective coordinate $s(t)$ has collective mass m_s, satisfying $\langle \dot{s}^2 \rangle = k_{\mathrm{B}}T/m_s$. From Eq. 10, $\langle \dot{s}^2 \rangle = \langle \sum_{i=1}^{n} \sum_{j=1}^{n} \dot{d}_i \dot{d}_j \rangle = \sum_{i=1}^{n} \langle \dot{d}_i^2 \rangle$, resulting in $m_s = m_d/n$, where we have used that there is no off-diagonal velocity correlation. For the normalized collective variable $\tilde{s}$, the collective mass satisfies $\langle \dot{\tilde{s}}^2 \rangle = k_{\mathrm{B}}T/m$. Thus, $m = m_s n^2 = m_d n$. Since the number of native contacts n scales roughly linear in the number of residues N, as shown in Fig. S3A, we predict a linear scaling of the effective mass with number of residues.

The mean squared displacement (MSD) of the collective coordinate s is, in the appropriate diffusive regime, $\left\langle \big(s(t) - s(0)\big)^2 \right\rangle = 2D_s t$, where D_s is the diffusion coefficient for s. From Eq. 10, we can write:

$$\sum_{i=1}^{n} \sum_{j=1}^{n} \left\langle \big(d_i(t) - d_i(0)\big)\big(d_j(t) - d_j(0)\big) \right\rangle = \sum_{i=1}^{n} \left\langle \big(d_i(t) - d_i(0)\big)^2 \right\rangle = 2nD_d t, \qquad [11]$$

where we again assume no correlation between different distances. Thus, we see that $D_s = nD_d$, where D_d is the diffusion coefficient of the individual distances, and hence $\gamma_s = \gamma_d/n$, in line with the result for m_s in terms of m_d. Likewise, we can calculate $D = D_s/n^2$, where D is the diffusion coefficient for the normalized $\tilde{s}$, such that $\gamma = \gamma_s n^2 = \gamma_d n$.

Taken together, these results tell us that normalisation of a reaction coordinate that is linear in atomic distances results in an increase in both friction and mass by n^2. In the absence of normalisation, we expect the mass and friction to scale as $1/n$, and so normalisation would result in linear scaling. In Fig. 1 of the main manuscript, we see that both friction and mass scale super-linearly with the number of residues in the protein chain. The Q reaction coordinate is normalized by the number of native contacts N_{nc}, which we show in Fig. S3A scales approximately linearly with N. The strong super-linearities observed for the friction and the weak super-linearity observed for the mass in the main manuscript ($\gamma/k_{\mathrm{B}}T \sim N^{2.8}$ and $m \sim N^{1.5}$) result from non-linearities in the Q reaction coordinate and, in the case of friction, from the reptation-like reconfiguration dynamics of the protein chain (14, 15). By evaluating the friction and mass for other reaction coordinates, we can see how these effects manifest differently depending on the choice of the reaction coordinate. In Figs. S3B and C, we show the friction and mass respectively for Q, and three additional reaction coordinates (see section 11). We see a range of scaling relations. Importantly, for the case of the end-to-end distance, which is a linear reaction coordinate that does not depend on N, we see constant mass, as expected, but approximately linear scaling of friction. This latter result cannot be due to non-linearities in the reaction coordinate and is therefore solely due to the connectivity of the peptide chain that leads to reptation-like chain reconfiguration dynamics, as was previously seen for the reconfiguration dynamics of homopolymer globules (15).

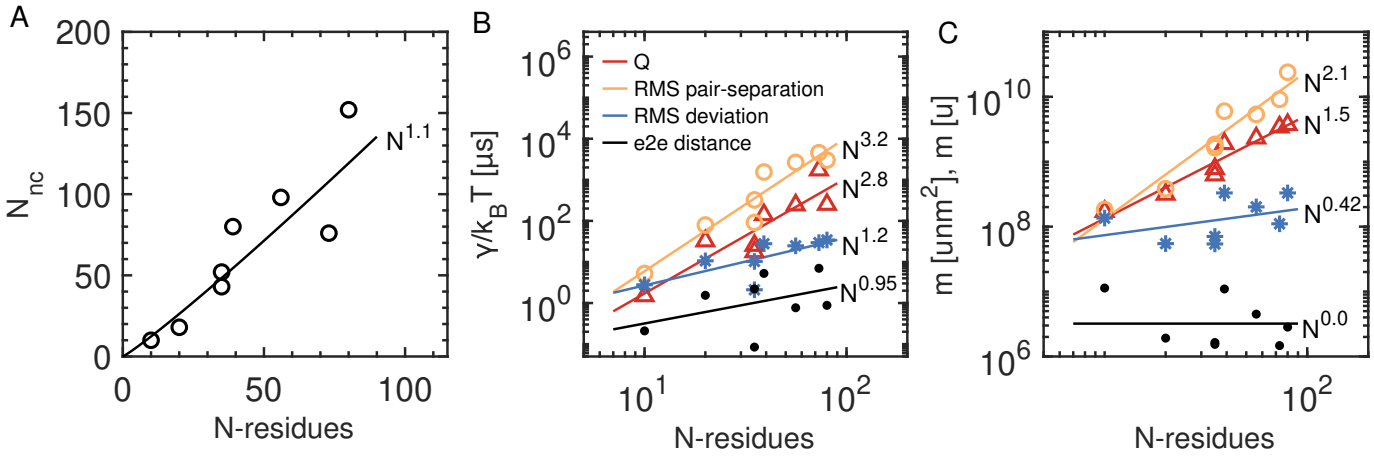

Fig. S3. Scaling of mass and friction with the length of protein chain. A) Number of native contacts $\mathrm{N_{nc}}$ as a function of the number of residues in each protein. A power law fit $\sim \alpha N^{\beta}$ reveals almost linear scaling with $\alpha = 1.0$ and $\beta = 1.1$. B) Scaling of total friction as a function of the number of residues for a range of reaction coordinates. Scaling relations are indicated such that $\gamma/k_{\mathrm{B}}T \sim N^{\beta}$. The legend is used for both B and C. C) Scaling of the effective reaction-coordinate mass as a function of chain-length for a range of reaction coordinates. Fraction of native contacts has units uns^2, all other reaction coordinates have units u.

6. Additional analysis of folding times, friction, and barrier heights

Fig. S4A compares the folding and unfolding times extracted from the MD simulations to the experimental results presented by Plaxco *et. al.* for a set of single domain proteins (16). We fit linear functions to the combined folding and unfolding times extracted from the MD simulations (red dashed line), the results presented by Plaxco *et. al.* (blue dashed line), and the two data sets combined (black dashed line) on a log-linear scale. We calculate the Pearson's correlation coefficient (r) and the corresponding p-values to assess the degree of correlation between the MD data, the Plaxco data, and the combined data set. The details of the calculation are described in Section 8 below. For the Plaxco data, we calculate the same r and p-value as the authors (16). Both the r and p values for the MD data show a higher degree of correlation than the Plaxco data, although the correlation is weak and not statistically significant at a conventional level of significance (i.e., $p^{\mathrm{MD}}_{\mathrm{MFP}} > 0.05$). For the MD folding processes only, i.e. neglecting unfolding (black plus symbols in Fig. S4A), we measure $r = 0.87$ and $p = 0.002$. We also show the exponential fits for folding only (exponent 0.053), and for unfolding only (exponent 0.025). When we combine the experimental and the MD simulation folding and unfolding data sets we find a high degree of correlation, characterized by $r_{\mathrm{all}} = 0.75$ and $p_{\mathrm{all}} < 0.0001$. We keep in mind, however, that this latter results combines simulation and experimental results. The small panel shows MD folding and unfolding times evaluated for transitions from minimum to minimum. We show the r and p-values for the combined set of folded and unfolding, and also just for the folding times. We also include the exponential fits for the combined folding and unfolding (exponent 0.053, thick dashed line), as well as for folding only (exponent 0.046) and unfolding only (exponent 0.03).

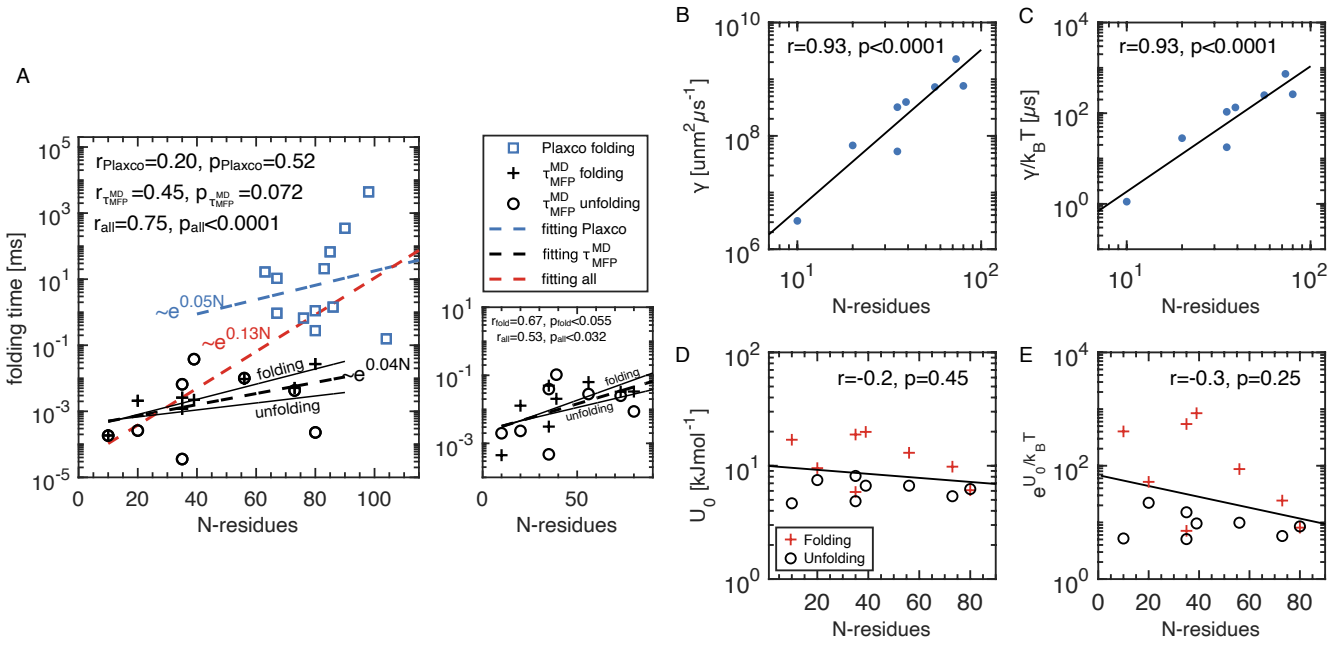

Fig. S4. A) Comparison of MD folding times with a set of experimental folding times measured for a set of single domain proteins. The red plus symbols and black circles represent folding and unfolding times from MD simulation, as presented in Fig. 1H of the main manuscript. Blue squares are taken from Plaxco *et. al.* (16). The Pearson's correlation coefficient (r) and p-value are obtained by evaluating a test statistic $t = (r - r_0)/(\sqrt{1 - r^2}/(n - 2))$, where $r_0 = 0$ is the null hypothesis and n is the number of data points in a given sample (see SI Section 8). The small panel shows folding and unfolding times for transitions from minimum to minimum. Exponential fits are shown for combined folding and unfolding (thick dashed line), as well as for folding only and unfolding only. B) and C) show that friction and mobility (inverse diffusivity) are highly correlated with chain length. The data is presented on a log-log scale since a power-law dependence is expected (15). The black lines indicate a linear fit to the log-log transformed data, giving a power-law exponent of 2.8 in both B) and C). D) and E) show that there is no correlation between the free energy barrier heights and protein chain length. The black lines indicate a linear fit to log-linear data, resulting in exponential scaling factors of -0.004 and -0.2 for D) and E), respectively.

In Figs. S4B-E, we examine the correlations between both friction and free energy barrier height with the length of protein chains. As shown in Fig. 1E, we observe a high correlation between mobility (inverse diffusivity $\gamma/k_{\mathrm{B}}T$) and chain length, which we confirm with the r and p-values. Furthermore, we demonstrate that high correlation is also evident in the friction coefficients γ. However, we note that the free energy barrier heights, expressed either as U_0 or the Arrhenius factor $\mathrm{e}^{U_0/k_{\mathrm{B}}T}$, do not display any correlation with chain length, and instead appear to decrease for larger N.

7. Investigating discretisation effects with alanine-9 homo-peptide chain simulation

The data analyzed in this paper have low time resolution since the protein configurations were saved at time intervals of Δt =0.2 ns, despite the original simulations being performed with time steps of dt =2.5 fs. Thus, five orders of magnitude of time resolution are not accessible for our analysis. To test the reliability of our results with this low-resolution representation, we compare it to a system with full-time-resolution data over long simulation times. We use a previously published trajectory of the 9-residue homo-alanine peptide (Ala9) (12) and investigate the accuracy of the memory time τ_{mem} by systematically discretizing a range of sub-sampled time steps Δt. The Ala_9 system is simulated using Gromacs with the Amber03 force field and 4023 explicit SPC/E water molecules, where the simulation time step is $dt = 1$ fs, and the total simulation time is 10 μs. We use the standard HB_4 reaction coordinate for α-helix forming chains, which is the average separation between the n^{th} residue nitrogen and n^{th}+4 residue oxygen for $n = 2$, 3, and 4.

We evenly sub-sample low-resolution trajectories at intervals of $\Delta t = 0.01$, 0.1, 1, 10, 100, and 200 ps to generate data. Running integrals of the memory kernels are extracted for each discretization, as described in Section 3. Fig.S5A shows the running integrals G(t) for the fully resolved data ($\Delta t = 0.001$ ps) and four other discretizations, while Fig.S5B displays the corresponding $\Gamma(t)$ functions. Due to the accurate extraction of the G(t) plateau-value, the extraction method produces reliable total friction γ over the entire range of discretization, as evident from Fig. S5C. Here, we compare γ evaluated in two ways. The first method fits an exponential series to $\Gamma(t)$ extracted from the simulations (Section 4) and reverts this $\Gamma(t)$ back to a running integral to determine the long-time limit of the running integral as γ. The second method calculates the average of the originally-extracted running integral (no exponential fit) over the plateau region for $t > 2000$ ps. The first method is representative of how γ is evaluated in the main manuscript, while the second method is purely complementary and provides an error estimate, which we here show as the standard deviation. Both methods are in agreement, implying that even under low resolution, we can evaluate the total friction using kernel extraction techniques.

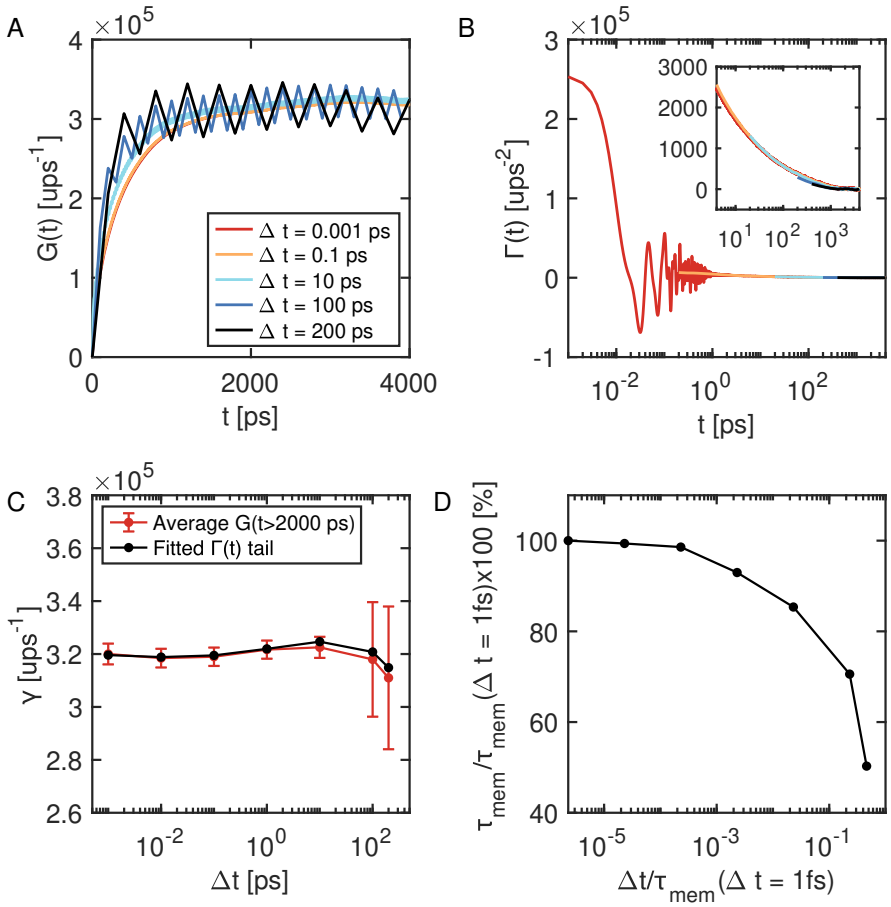

Fig. S5. Ala_9 discretisation and memory time extraction. A) Memory kernel running integrals G(t) extracted from HB_4 reaction coordinate trajectories at different time-series resolutions Δt, with discretisations spanning five orders of magnitude. B) Corresponding memory kernels $\Gamma(t)$. The inset is a magnification of the long-time tail. C) Total friction on the HB_4 reaction coordinate, evaluated via the memory kernel extraction, as a function of discretisation. The red points with error bars show the average of $G(t)$ for each discretisation, taken in the range 2 μs $\leq$ t $\leq$ 4 μs. The error bars show the standard deviations for $G(t > 2000$ ps). The black points show the plateau values for $G(t)$, taken as the running integral of $\Gamma(t)$, with a fitted single exponential long-time tail. D) Memory times τ_{mem}, evaluated via the 1st-moment of the memory kernel, as a function of rescaled discretisation step size, rescaled by the memory time evaluated for the fully resolved data $\tau_{\mathrm{mem}}(\Delta t = 1\mathrm{fs}) = 420$ ps.

In Fig. S5D, we plot the memory times (τ_{mem}) as a function of the discretization step size. The data is rescaled by the memory time, evaluated for the fully resolved trajectory with $\Delta t = 1$ fs. We also rescale the step size by the memory time of the trajectory at full resolution, $\tau_{\text{mem}}(\Delta t = 1 \text{ fs})$, to estimate the relative error expected for a discretized system. For the 8 fast-folding proteins, the fully resolved system has a step size of 2.5 fs. We observe a 20% reduction in the accuracy of the memory times for time-series discretization spanning four orders of magnitude. The accuracy drops more rapidly for lower resolutions. Based on the range of memory times observed for the 8 proteins and a step size of 0.2 ns, we expect a range of $1{\times}10^{-5} < 0.2 \text{ ns}/\tau_{\text{mem}}(\Delta\text{t} = 2.5 \text{ fs}) < 1{\times}10^{-3}$. We can conclude that the memory times presented in this paper for the low-time-resolution trajectories are inaccurate by at most 10%.

8. Prediction errors for barrier crossing times

We measure the deviation of three model predictions from observed MD values across all proteins. Fig. S6 shows these three model predictions compared to mean first-passage times from MD data. These plots are the same as Figs. 3A, B, and D in the main manuscript, here presented on a linear-linear scale. Deviations between folding times and theoretical predictions are greater for proteins with slower reaction times. For the fastest reactions, deviations cannot be visually resolved as they appear as a single data point at $\tau^{\mathrm{MD}}_{\mathrm{MFP},i} \approx 0$ and $\tau^{\mathrm{theo}}_{\mathrm{MFP},i} \approx 0$. This demonstrates that a comparison of linearly scaled MFPT data is meaningless as it is dominated by the proteins with the longest folding times.

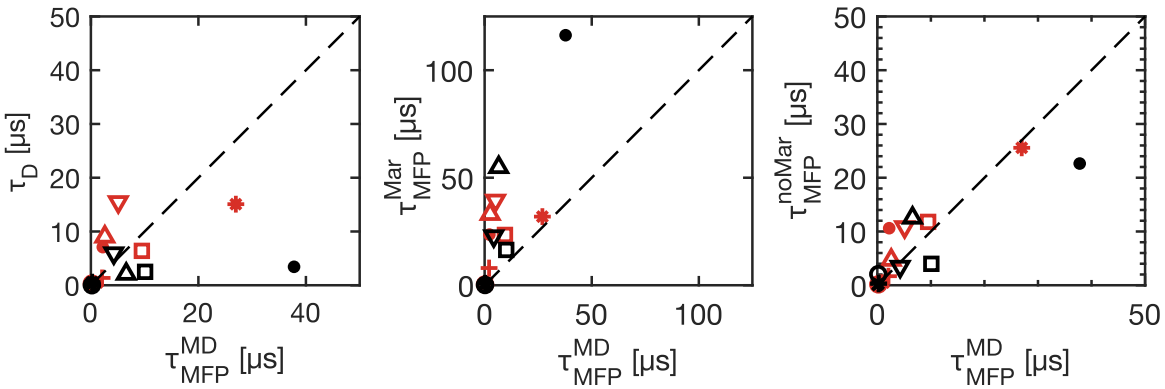

Fig. S6. Comparison of simulated protein folding and unfolding times $\tau^{\mathrm{MD}}_{\mathrm{MFP}}$ with predictions on different levels of theory, plotted on a linear scale (same data as is presented in the main manuscript Fig. 3A, B, and D). Presented on the linear scale, we see that deviations between the extracted MD mean first passage times $\tau^{\mathrm{MD}}_{\mathrm{MFP},i}$ and the various theoretical predictions $\tau^{\mathrm{theo}}_{\mathrm{MFP},i}$ are greater for proteins that fold and unfold slowly. The RMSD (Eq. 12) is dominated by these slow-folding proteins.

A common method for assessing the difference between observed experiment values and a predicted model is to compute the root-mean-square deviation (RMSD), as follows:

$$\mathrm{RMSD} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(\tau^{\mathrm{MD}}_{\mathrm{MFP},i} - \tau^{\mathrm{theo}}_{\mathrm{MFP},i}\right)^2} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(\frac{\tau^{\mathrm{MD}}_{\mathrm{MFP},i} - \tau^{\mathrm{theo}}_{\mathrm{MFP},i}}{\tau^{\mathrm{MD}}_{\mathrm{MFP},i}}\right)^2 (\tau^{\mathrm{MD}}_{\mathrm{MFP},i})^2}. \quad [12]$$

We combine folding and unfolding times into one set with $N = 16$. Figure S7A shows the distributions of $\tau^{\mathrm{MD}}_{\mathrm{MFP},i} - \tau^{\mathrm{theo}}_{\mathrm{MFP},i}$ for three theories. The Markovian model ($\tau^{\mathrm{Mar}}_{\mathrm{MD}}$) has a few outlying data points that contribute significantly to the RMSD. The last term in Eq. 12 shows that if we express the RMSD in terms of relative deviations, then each relative deviation is multiplied by a factor of $(\tau^{\mathrm{MD}}_{\mathrm{MFP},i})^2$. We use a logarithmic scaling scheme to account for relative contributions of data that spans multiple orders of magnitude. This gives us the root-mean-square logarithmic deviation (RMSLD), given by

$$\mathrm{RMSLD} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(\mathrm{Log}(\tau^{\mathrm{MD}}_{\mathrm{MFP},i}) - \mathrm{Log}(\tau^{\mathrm{theo}}_{\mathrm{MFP},i})\right)^2} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(\mathrm{Log}\left(\frac{\tau^{\mathrm{MD}}_{\mathrm{MFP},i} - \tau^{\mathrm{theo}}_{\mathrm{MFP},i}}{\tau^{\mathrm{MD}}_{\mathrm{MFP},i}} + 1\right)\right)^2}. \quad [13]$$

The last form in Eq. 13 shows that by calculating the RMSLD, we are actually calculating the logarithmic transformation of the relative deviations. The distributions of $\tau^{\mathrm{MD}}_{\mathrm{MFP},i} - \tau^{\mathrm{theo}}_{\mathrm{MFP},i}$ and $\log(\tau^{\mathrm{MD}}_{\mathrm{MFP,i}}) - \log(\tau^{\mathrm{theo}}_{\mathrm{MFP,i}})$ for the three different theories, along with the corresponding RMSD and RMSLD values, are shown in Fig. S7B.

We can quantify the correlations observed in Fig. 3 of the main manuscript and Fig. S6 using the Pearson's correlation coefficient r_{CC}, and likewise the Pearson's correlation coefficient for logarithmically transformed data r_{LogCC}. The Pearson's correlation coefficient r_{CC} is given by

$$r_{\mathrm{CC}} = \frac{\sum_{i=1}^{N}\left(\tau^{\mathrm{MD}}_{\mathrm{MFP},i} - \left\langle\tau^{\mathrm{MD}}_{\mathrm{MFP},i}\right\rangle\right)\left(\tau^{\mathrm{theo}}_{\mathrm{MFP},i} - \left\langle\tau^{\mathrm{theo}}_{\mathrm{MFP},i}\right\rangle\right)}{\sqrt{\sum_{i=1}^{N}\left(\tau^{\mathrm{MD}}_{\mathrm{MFP},i} - \left\langle\tau^{\mathrm{MD}}_{\mathrm{MFP},i}\right\rangle\right)^2}\sqrt{\sum_{i=1}^{N}\left(\tau^{\mathrm{theo}}_{\mathrm{MFP},i} - \left\langle\tau^{\mathrm{theo}}_{\mathrm{MFP},i}\right\rangle\right)^2}}. \quad [14]$$

This is a measure of the normalized bivariate correlation between the set of $\tau^{\mathrm{MD}}_{\mathrm{MFP},i}$ and $\tau^{\mathrm{theo}}_{\mathrm{MFP},i}$ values. For the logarithmically transformed data we calculate

$$r_{\mathrm{LogCC}} = \frac{\sum_{i=1}^{N}\left(\mathrm{Log}(\tau^{\mathrm{MD}}_{\mathrm{MFP},i}) - \left\langle\mathrm{Log}(\tau^{\mathrm{MD}}_{\mathrm{MFP},i})\right\rangle\right)\left(\mathrm{Log}(\tau^{\mathrm{theo}}_{\mathrm{MFP},i}) - \left\langle\mathrm{Log}(\tau^{\mathrm{theo}}_{\mathrm{MFP},i})\right\rangle\right)}{\sqrt{\sum_{i=1}^{N}\left(\mathrm{Log}(\tau^{\mathrm{MD}}_{\mathrm{MFP},i}) - \left\langle\mathrm{Log}(\tau^{\mathrm{MD}}_{\mathrm{MFP},i})\right\rangle\right)^2}\sqrt{\sum_{i=1}^{N}\left(\mathrm{Log}(\tau^{\mathrm{theo}}_{\mathrm{MFP},i}) - \left\langle\mathrm{Log}(\tau^{\mathrm{theo}}_{\mathrm{MFP},i})\right\rangle\right)^2}}. \quad [15]$$

Table S5. Comparison of various measures of prediction error for the four barrier-crossing-time models. Root-mean-squared deviations (RMSD) and root-mean-squared logarithmic deviations (RMSLD) are given by Eqs. 12 and 13, respectively. The Pearson's correlation coefficient (r_{CC}) and correlation coefficient for logarithmically transformed data (r_{LogCC}) are given by Eqs. 14 and 15, respectively. For the correlation coefficients, the values in parenthesis are the test-statistic values ($t = r/(\sqrt{1-r^2}/(n-2))$), which assumes a null-hypothesis of zero correlation, followed by the corresponding p-values. $n = 16$ is the number of degrees of freedom.

	Diffusion times (τ_D)	Markovian (τ_{MFP}^{Mar})	Non-Markovian (τ_{MFP}^{noMar})	Free energy factor ($\tau_D' \xi_U$)
RMSD [μs]	9.95	27.0	5.13	11.2
RMSLD	1.5	1.4	0.8	1.98
r_{CC}	0.37 (t=1.5, p=0.15)	0.81 (t=5.2, p<0.001)	0.87 (t=6.7, p<0.001)	0.19 (t=0.75, p=0.46)
r_{LogCC}	0.83 (t=5.2, p<0.001)	0.92 (t=8.4, p<0.001)	0.94 (t=9.4, p=<0.001)	0.48 (t=1.93, p=0.07)

The evaluated correlation coefficients r_{CC} and r_{LogCC}, as well as the RMSD and RMSLD values, for the four different prediction models are summarized in Table S5.

In Table S5, we also present the test statistics (t-values) and corresponding two-tailed p-values for the null hypothesis $r_0 = 0$, stating that the data is uncorrelated. The test statistic is the deviation of r_{CC} (or r_{LogCC}) from the null hypothesis, per standard error $t = r/(\sqrt{1-r^2}/(n-2))$, where $n = 16$ is the number of degrees of freedom. p-values of p<0.001 are considered extremely statistically significant. The Markovian model τ_{MFP}^{Mar} is an interesting case where RMSD and RMSLD suggest it is either a less accurate prediction than the diffusion time τD, or at best, a slight improvement. The Pearson's correlation coefficients, r_{CC} and r_{LogCC}, suggest that τ_{MFP}^{Mar} provides a significant improvement for prediction over τ_D and is almost as accurate as τ_{MFP}^{noMar}. The correlation coefficient reveals linear correlation regardless of the slope of the linear relation. The RMSD and RMSLD, however, penalize data that is linearly correlated but has a slope different from 1, as is the case when memory acceleration effects are present. The Markovian model typically predicts values slower than the MD data, despite being strongly linearly correlated. This relationship is not penalized by r, but is by the RMSD and RMSLD.

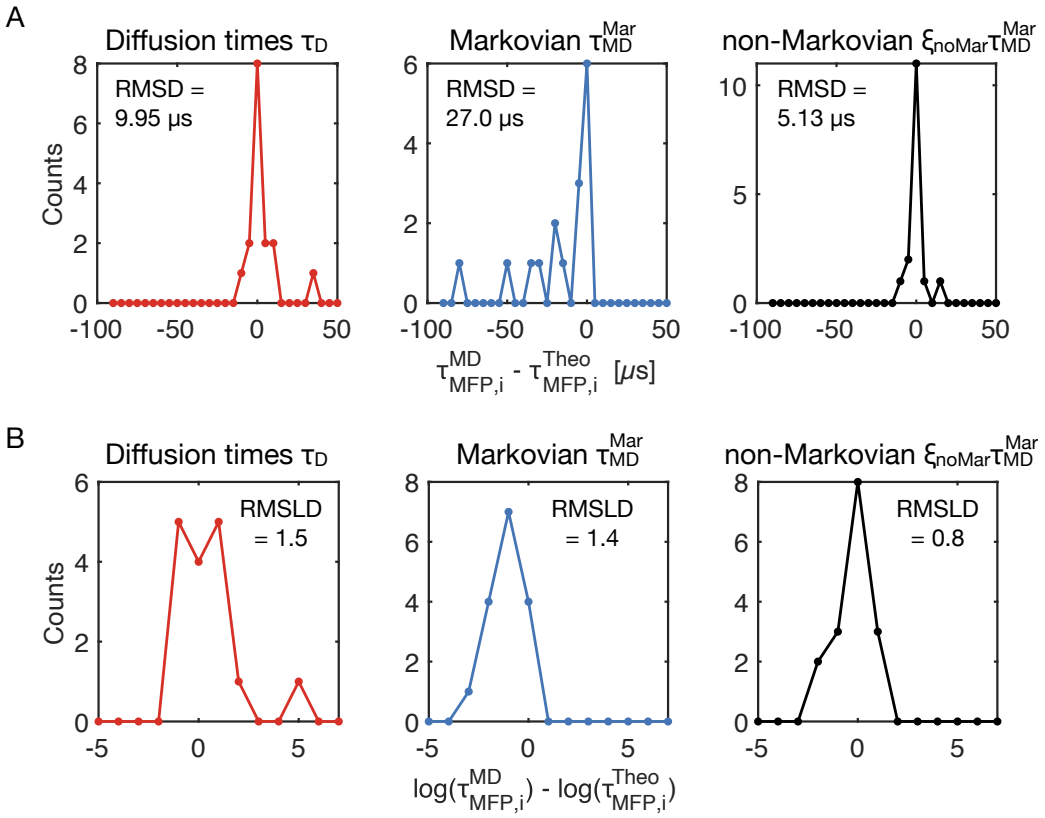

Fig. S7. Distributions for the deviations between theoretical predictions and simulated values of the barrier crossing times. Folding and unfolding reactions are collected into a single set, such that $N = 16$. A) Root-mean-square deviations (RMSD) and B) the root-mean-square logarithmic deviations (RMSLD).

9. Evaluating Markovian mean first-passage times: position-independent friction

For a purely Markovian system with constant friction, the mean first-passage time between reaction coordinate positions Q_s and Q_e is given by:

$$\tau_{\mathrm{MFP}}^{\mathrm{Mar}}(Q_\mathrm{s}, Q_\mathrm{e}) = \beta\gamma \int_{Q_\mathrm{s}}^{Q_\mathrm{e}} \mathrm{e}^{\beta U(x)} \left[\int_{-\infty}^{x} \mathrm{e}^{-\beta U(y)} dy \right] dx \quad [16]$$

where the subscript s indicates a start position and e indicates an end position.

In practice, the inner integral of Eq.16 will have a lower limit at the appropriate edge of the confining potential Q_0. The discrete form of the inner integral, using a rectangular discretization, will therefore be:

$$\int_{-\infty}^{x} \mathrm{e}^{-\beta U(y)} dy \;\; \rightarrow \;\; \sum_{i=0}^{j} \mathrm{e}^{-\beta U(Q_i)} \Delta y. \quad [17]$$

The composite integral becomes

$$\int_{Q_\mathrm{s}}^{Q_\mathrm{e}} \mathrm{e}^{\beta U(x)} \left[\int_{-\infty}^{x} \mathrm{e}^{-\beta U(y)} dy \right] dx \rightarrow \sum_{j=a}^{b} \mathrm{e}^{\beta U(Q_j)} \left[\sum_{i=0}^{j} \mathrm{e}^{-\beta U(Q_i)} \Delta Q \right] \Delta Q. \quad [18]$$

ΔQ is the discretisation step size, set by the free-energy histogram. a and b are the indices for the bins located at the beginning and the end of the transition, which might be, for example, the unfolded state minimum and the barrier top, respectively, for a folding transition. The discrete form of the mean first-passage time is then:

$$\tau_{\mathrm{MFP}}^{\mathrm{Mar}}(Q_\mathrm{s}, Q_\mathrm{e}) = \beta\gamma\Delta Q^2 \sum_{j=a}^{b} \sum_{i=0}^{j} \mathrm{e}^{\beta U(Q_j)} \mathrm{e}^{-\beta U(Q_i)}. \quad [19]$$

For the case of position-dependent friction, outlined in Eq. 49, $\gamma(Q)$ resides inside the outer integral, such that the composite integral becomes:

$$\tau_{\mathrm{MFP}}^{\mathrm{Mar}}(Q_\mathrm{s}, Q_\mathrm{e}) = \int_{Q_\mathrm{s}}^{Q_\mathrm{e}} \gamma(x) \mathrm{e}^{\beta U(x)} \left[\int_{-\infty}^{x} \mathrm{e}^{-\beta U(y)} dy \right] dx \rightarrow \sum_{j=a}^{b} \gamma_j \mathrm{e}^{\beta U(Q_j)} \left[\sum_{i=0}^{j} \mathrm{e}^{-\beta U(Q_i)} \Delta Q \right] \Delta Q \quad [20]$$

The discrete form is then written:

$$\tau_{\mathrm{MFP}}^{\mathrm{Mar}}(Q_\mathrm{s}, Q_\mathrm{e}) = \beta\Delta Q^2 \sum_{j=a}^{b} \sum_{i=0}^{j} \gamma_j \mathrm{e}^{\beta U(Q_j)} \mathrm{e}^{-\beta U(Q_i)}. \quad [21]$$

10. Non-Markovian reaction-rate theory and Grote-Hynes theory

In Fig. 4A of the main manuscript, we use the non-Markovian factor ξ_{noMar} to predict barrier crossing times, which agree well with the values obtained from MD simulations. For each protein, we evaluate ξ_{noMar} in both the folding and unfolding directions using parameters extracted from simulations, including mass m, free-energy barrier height U_0, characteristic length L, and the set of amplitudes γ_i and memory times τ_i comprising the extracted memory kernels $\Gamma(t) = \sum_{i=1}^{M} (\gamma_i/\tau_i) \exp(t/\tau_i)$. We use a previously published heuristic formula (17–19) to predict the barrier crossing time for a given protein, such that

$$\tau^{\mathrm{H}}_{\mathrm{MFP}}(\tau_{\mathrm{D}}, U_0, \{\gamma_i\}, \{\tau_i\}, \tau_{\mathrm{m}}) = \sum_{i=1}^{M} \tau^{i}_{\mathrm{OD}} + \left[\sum_{i=1}^{M} \frac{1}{\tau^{i}_{\mathrm{ED}}} \right]^{-1}, \quad [22]$$

where $\tau_{\mathrm{m}} = m/\gamma$, such that $\gamma = \sum_{i=1}^{M} \gamma_i$, and $\tau_{\mathrm{D}} = \gamma L^2/k_{\mathrm{B}}T$. For of the M components included in the memory kernel $\Gamma(t)$ there will be M over-damped contributions (τ^{i}_{OD}) to the barrier crossing times and M energy-diffusion contributions (τ^{i}_{ED}). The over-damped contributions are given by

$$\tau^{i}_{\mathrm{OD}} = \tau_{\mathrm{D}} \frac{\gamma_i}{\gamma} \frac{\mathrm{e}^{\beta U_0}}{\beta U_0} \left[\frac{\pi}{2\sqrt{2}} \frac{1}{1 + 10\beta U_0 \tau_i/\tau_{\mathrm{D}}} + \sqrt{\beta U_0 \frac{\tau_{\mathrm{m}}}{\tau_{\mathrm{D}}}} \right], \quad [23]$$

and the energy diffusion contributions are given by

$$\tau^{i}_{\mathrm{ED}} = \tau_{\mathrm{D}} \frac{\gamma}{\gamma_i} \frac{\mathrm{e}^{\beta U_0}}{\beta U_0} \left[\frac{\tau_{\mathrm{m}}}{\tau_{\mathrm{D}}} + 4\beta U_0 \left(\frac{\tau_i}{\tau_{\mathrm{D}}} \right)^2 + \sqrt{\beta U_0 \frac{\tau_{\mathrm{m}}}{\tau_{\mathrm{D}}}} \right], \quad [24]$$

as is also presented in the Methods section of the main manuscript. In the Markovian limit, i.e. $\tau_i \to 0$ for all memory time scales, Eq. 22 reduces to

$$\tau^{\mathrm{H}}_{\mathrm{Mark}} = \tau^{\mathrm{H}}_{\mathrm{MFP}}(\tau_{\mathrm{D}}, U_0, \{\gamma_i\}, 0, \tau_{\mathrm{m}}) = \tau_{\mathrm{D}} \frac{\mathrm{e}^{\beta U_0}}{\beta U_0} \left[\frac{\pi}{2\sqrt{2}} + 2\sqrt{\beta U_0 \frac{\tau_{\mathrm{m}}}{\tau_{\mathrm{D}}}} + \frac{\tau_{\mathrm{m}}}{\tau_{\mathrm{D}}} \right]. \quad [25]$$

Furthermore, we see that in the high-friction (HF) limit, i.e. when $\tau_{\mathrm{m}} \to 0$, we obtain

$$\tau^{\mathrm{H}}_{\mathrm{HF}} = \tau^{\mathrm{H}}_{\mathrm{MFP}}(\tau_{\mathrm{D}}, U_0, \{\gamma_i\}, 0, 0) = \tau_{\mathrm{D}} \pi \mathrm{e}^{\beta U_0}/2\beta U_0 \sqrt{2}. \quad [26]$$

Thus we see that the non-Markovian correction factor is given by the full non-Markovian prediction for barrier crossing times in the presence of multi-modal memory rescaled by the high-friction, memoryless limit

$$\xi_{\mathrm{noMar}} = \frac{\tau^{\mathrm{H}}_{\mathrm{MFP}}(\tau_{\mathrm{D}}, U_0, \{\gamma_i\}, \{\tau_i\}, \tau_{\mathrm{m}})}{\tau^{\mathrm{H}}_{\mathrm{MFP}}(\tau_{\mathrm{D}}, U_0, \{\gamma_i\}, 0, 0)} = \frac{\tau^{\mathrm{H}}_{\mathrm{MFP}}}{\tau^{\mathrm{H}}_{\mathrm{HF}}}, \quad [27]$$

which is Eq. 4 in the main manuscript.

Fig. 4B of the main manuscript introduces the memory-time-scaled heuristic curves that show the limits of vanishing memory time ($\tau_i \to 0$, $\forall\, i$) and long memory times ($\tau_i \to \infty$, $\forall\, i$). To achieve these limits, we uniformly rescale all memory times while keeping all other parameters, including the memory kernel amplitudes γ_i, fixed at the value corresponding to the MD data. α is the rescaling parameter, such that the rescaled memory kernel and the corresponding rescaled first-moment memory time are given by

$$\Gamma^{\alpha}(t) = \sum_{i=1}^{M} \frac{\gamma_i}{\alpha\tau_i} \exp(t/\alpha\tau_i), \quad \alpha\tau_{\mathrm{mem}} = \int_0^{\infty} t\Gamma^{\alpha}(t)dt \Big/ \int_0^{\infty} \Gamma^{\alpha}(t)dt, \quad [28]$$

where τ_{mem} denotes the memory time of the original unscaled (α=1) system. The $\xi_{\mathrm{noMar}}(\alpha)$ curves given in Fig. 4B of the main manuscript are generated by evaluating $\tau^{\mathrm{H},\alpha}_{\mathrm{MFP}} = \tau^{\mathrm{H}}_{\mathrm{MFP}}(\tau_{\mathrm{D}}, U_0, \{\gamma_i\}, \{\alpha\tau_i\}, \tau_{\mathrm{m}})$ across a range of $0 < \alpha < \infty$, such that the over-damped and energy diffusion components are given by

$$\tau^{\alpha,i}_{\mathrm{OD}} = \tau_{\mathrm{D}} \frac{\gamma_i}{\gamma} \frac{\mathrm{e}^{\beta U_0}}{\beta U_0} \left[\frac{\pi}{2\sqrt{2}} \frac{1}{1 + 10\beta U_0 \alpha\tau_i/\tau_{\mathrm{D}}} + \sqrt{\beta U_0 \frac{\tau_{\mathrm{m}}}{\tau_{\mathrm{D}}}} \right], \quad \tau^{\alpha,i}_{\mathrm{ED}} = \tau_{\mathrm{D}} \frac{\gamma}{\gamma_i} \frac{\mathrm{e}^{\beta U_0}}{\beta U_0} \left[\frac{\tau_{\mathrm{m}}}{\tau_{\mathrm{D}}} + 4\beta U_0 \left(\alpha \frac{\tau_i}{\tau_{\mathrm{D}}} \right)^2 + \sqrt{\beta U_0 \frac{\tau_{\mathrm{m}}}{\tau_{\mathrm{D}}}} \right]. \quad [29]$$

In doing so, it can be seen that neither the Markovian (τ_i=0) nor the high-friction ($\tau_{\mathrm{D}} \to \infty$) limit is affected by the α scaling. Thus,

$$\xi_{\mathrm{noMar}}(\alpha) = \frac{\tau^{\mathrm{H},\alpha}_{\mathrm{MFP}}}{\tau^{\mathrm{H}}_{\mathrm{HF}}}. \quad [30]$$

It is also clear that when $\alpha = 1$, we recover the predictions for the original MD simulations. In Fig. S8, we show the results for the α-scaled curves for all eight proteins. The vertical lines indicate the location corresponding to $\alpha = 1$, which is different for

the folding and unfolding processes since each process has a unique τ_{D}, which rescales the memory-time axis.

Grote-Hynes theory: The Grote-Hynes (GH) theory predictions presented in Fig. 4B of the main manuscript, as well as here in Fig. S8, are generated similarly. The GH theory corrects the classical transition state theory (TST) for the barrier-crossing rate by including frequency-dependent friction effects and particle mass. The TST barrier crossing time is given by

$$\tau^{\mathrm{TST}} = \frac{2\pi}{\omega_{\mathrm{min}}} \mathrm{e}^{\beta U_0} \tag{31}$$

where $\omega_{\mathrm{min}} = \sqrt{U''_{\mathrm{min}}/m}$ is the frequency at the free-energy minimum (well bottom). The GH prediction is therefore written as

$$\tau^{\mathrm{GH}} = \frac{\omega_{\mathrm{max}}}{\lambda} \tau^{\mathrm{TST}} = \frac{2\pi\omega_{\mathrm{max}}}{\lambda\omega_{\mathrm{min}}} \mathrm{e}^{\beta U_0}, \tag{32}$$

where $\omega_{\mathrm{max}} = \sqrt{|U''_{\mathrm{max}}|/m}$ is the frequency at the free-energy maximum (barrier top). U''_{min} and U''_{max} are the free energy curvatures at the minimum and maximum and m is the mass of the particle. λ is the barrier reactive frequency, which is determined by solving the Grote-Hynes equation

$$\lambda^2 + \lambda \frac{\tilde{\Gamma}(\lambda)}{m} = \omega^2_{\mathrm{max}}. \tag{33}$$

$\tilde{\Gamma}(\lambda)$ is the Laplace transform of the friction memory kernel, given by

$$\tilde{\Gamma}(\lambda) = \int_0^\infty \Gamma(t') \mathrm{e}^{-\lambda t'} dt', \tag{34}$$

which can be determined either analytically for simple memory kernels $\Gamma(t)$, or numerically for a more complicated $\Gamma(t)$. Having obtained $\tilde{\Gamma}(\lambda)$, one then solves Eq. 33, which will have one real and positive root. This root is assigned as the reactive frequency λ, which enters into Eq. 32. In the Markovian limit, i.e. $\Gamma(t) = \gamma\delta(t)$, and hence $\tilde{\Gamma}(\lambda) = \gamma$. τ^{GH} reduces to the Kramers barrier crossing time $\tau^{\mathrm{GH}}_{\mathrm{Kr}}$ for medium-to-high friction, which is given by

$$\tau^{\mathrm{GH}}_{\mathrm{Kr}} = \left[\sqrt{\frac{\gamma^2}{4m^2} + \omega^2_{\mathrm{max}}} - \frac{\gamma}{2m} \right]^{-1} \omega_{\mathrm{max}} \tau_{\mathrm{TST}}. \tag{35}$$

In the high-friction (HF) limit, we can further reduce the prediction Eq. 35 to

$$\tau^{\mathrm{GH}}_{\mathrm{HF}} = \frac{\gamma}{m} \tau_{\mathrm{TST}}, \tag{36}$$

which is achieved by expanding the term under the square-root in Eq. 35 as a Taylor series expansion for small $m^2\omega^2_{\mathrm{max}}/\gamma^2$. For the uniform α rescaling, we replace the Laplace transform in Eq. 34 with a Laplace transform of the rescaled memory kernel (Eq. 28)

$$\tilde{\Gamma}^\alpha(\lambda) = \int_0^\infty \Gamma^\alpha(t') \mathrm{e}^{-\lambda t'} dt', \tag{37}$$

and hence evaluate $\lambda^2 + \lambda\tilde{\Gamma}^\alpha(\lambda)/m = \omega^2_{\mathrm{max}}$ over the range of α. We indicate the α-scaled Grote-Hynes curves, as shown in Figs. 4B of the main manuscript and S8, as $\tau^{\mathrm{GH},\alpha}/\tau^{\mathrm{GH}}_{\mathrm{HF}}$. We give additional parameters required to evaluate the GH curves in Table S6.

In Fig. S8, we present the heuristic and GH curves for the folding and unfolding of all eight proteins, along with the MD folding and unfolding times $\tau^{\mathrm{MD}}_{\mathrm{MFP}}/\tau^{\mathrm{Mar}}_{\mathrm{MFP}}$ from Fig.4A of the main manuscript. The MD results correspond to $\alpha = 1$, which we indicate with vertical lines. The theoretical predictions corresponding to $\alpha = 1$ coincide the vertical dashed lines. All values for $\alpha = 1$ are shown in Fig. S9. We see that there is an agreement between the non-Markovian predictions (Eq. 22) and the GH theory for small memory times. However, for intermediate and large memory times, the predictions deviate such that the GH theory predicts barrier-crossing times that are orders of magnitude faster than the non-Markovian predictions. From Fig. 4A of the main manuscript, we know that the non-Markovian predictions agree well with the extracted MD results, in particular recapitulating the memory-induced slow-down regime, which is not well represented by the GH theory.

Grote-Hynes analysis of simple model system: We compare the Grote-Hynes theory to the heuristic formula for a simple model system and find interesting behaviour in the long memory-time limit as we approach the limit of vanishing mass. In Fig. S10, we show barrier crossing time predictions given by the two theories for a simple model consisting of a massive particle in 1D moving in a double-well potential with a single component exponential memory kernel. The double-well potential is given by $U(Q) = [(Q/L)^2 - 1]^2$, such that $U''_{\mathrm{min}} = 8U_0/L^2$ and $U''_{\mathrm{max}} = -4U_0/L^2$. For a single component memory kernel

Table S6. Free energy curvatures and barrier frequencies. Barrier curvatures (U'' - absolute values shown) are given in units of kJmol^{-1} = 1×10^6unm^{2}ns^{-2} and frequencies ω are in units of ns.

	$U''_{\mathrm{u,min}}$	$\lvert U''_{\mathrm{u,max}}\rvert$	$U''_{\mathrm{f,max}}$	$U''_{\mathrm{f,min}}$	$\omega_{\mathrm{u,min}}$	$\omega_{\mathrm{u,max}}$	$\omega_{\mathrm{f,max}}$	$\omega_{\mathrm{f,min}}$
Chignolin	66.3	60.2	10496	43210	0.62	0.59	7.84	15.91
Trp-Cage	211.2	939.8	3460.8	35873	0.81	1.72	3.31	10.68
Villin	553.2	3076.4	6593.5	12401	0.94	2.22	3.25	4.46
WW Domain	148.6	587.6	2247.5	17846	0.43	0.86	1.69	4.78
NTL9	249.1	492.6	338.7	87043	0.36	0.51	0.42	6.77
Protein-G	310.4	597.8	13502	66200	0.36	0.50	2.41	5.35
α_3D	410.1	625.9	1984.9	8719.8	0.34	0.42	0.76	1.59
λ-Repressor	264.2	481.4	7351.7	14732	0.26	0.35	1.40	1.99

$\Gamma(t) = \gamma \mathrm{e}^{-t/\tau_{\mathrm{mem}}}/\tau_{\mathrm{mem}}$, the Laplace transform (Eq. 34) required to solve Eq. 33 is simply given by $\tilde{\Gamma}(\lambda) = \gamma/(1+\tau_{\mathrm{mem}}\lambda)$. Note that since we have an analytical expression for $\tilde{\Gamma}(\lambda)$ in this simple system, we compare to results obtained using a numerical evaluation of $\tilde{\Gamma}(\lambda)$, which is the method that we use for the study of the protein systems discussed above. Using the analytical result for $\tilde{\Gamma}(\lambda)$, the Grote-Hynes equation (Eq. 33) is then given by a cubic polynomial: $\tau_{\mathrm{mem}}\lambda^3 + \lambda^2 + (\gamma/m - \omega^2_{\mathrm{max}}\tau_{\mathrm{mem}})\lambda - \omega^2_{\mathrm{max}} = 0$, which will always have one real and positive root.

The parameters used for Fig.S10A are chosen to match Kappler *et. al.* (17), as shown in the figure caption, with $\tau_{\mathrm{m}}/\tau_{\mathrm{D}} = 0.01$. GH agrees with the non-Markovian formula (Eq. 22) for small $\tau_{\mathrm{mem}}/\tau_{\mathrm{D}}$ but deviates for $\tau_{\mathrm{mem}}/\tau_{\mathrm{D}} > 0.1$. Both theories produce the memory-induced speed-up in the region of $1\times10^{-2} < \tau_{\mathrm{mem}}/\tau_{\mathrm{D}} < 1\times10^{0}$. Small deviations occur between the two theories in the memoryless limit ($\tau_{\mathrm{mem}}/\tau_{\mathrm{D}} \to 0$) due to finite-mass effects from the choice of $\tau_{\mathrm{m}}/\tau_{\mathrm{D}}$. In Fig.S10B, both theories reach their high-friction limit for small memory times as $\tau_{\mathrm{m}}/\tau_{\mathrm{D}} \to 0$. Interestingly, for the GH theory, as $\tau_{\mathrm{m}}/\tau_{\mathrm{D}}$ decreases, the long memory-time limit also decreases. This behavior is relevant for protein folding, as observed in Figs. S8 and S9, where the GH curves with $\alpha = 1$ are located in the GH curve region where the long memory-time behavior converges to a constant value, as seen in Trp-Cage, Villin, WW-domain, and Protein G. Therefore, it is interesting to consider the long memory-time behavior for both predictions, especially for small mass.

Limit analysis: In the small-mass (SM) limit, $\lambda(\lambda + \tilde{\Gamma}(\lambda)/m) \approx \lambda\tilde{\Gamma}(\lambda)/m$ (from Eq. 33), so we write $\lambda\tilde{\Gamma}(\lambda) = m\omega^2_{\mathrm{max}}$. For a non-Markovian system with single component memory kernels, we include the analytic form of $\tilde{\Gamma}(\lambda)$ from the previous section, which leads to a reactive frequency $\lambda = m\omega^2_{\mathrm{max}}/[\gamma(1 - \tau_{\mathrm{mem}} m\omega^2_{\mathrm{max}}/\gamma)]$. The prediction for the Grote-Hynes theory is then

$$\tau^{\mathrm{GH}}_{\mathrm{SM}} = \frac{2\pi\gamma}{m\omega_{\mathrm{min}}\omega_{\mathrm{max}}}\left(1 - \frac{\tau_{\mathrm{mem}} m\omega^2_{\mathrm{max}}}{\gamma}\right)\mathrm{e}^{\beta U_0} = \frac{2\pi\gamma}{\sqrt{|U''_{\mathrm{min}}U''_{\mathrm{max}}|}}\left(1 - \frac{U''_{\mathrm{max}}\tau_{\mathrm{mem}}}{\gamma}\right)\mathrm{e}^{\beta U_0}. \quad [38]$$

Here, we see that the memory-induced speed-up regime is recovered, but for larger τ_{mem} the massless limit is ill-defined and $\tau^{\mathrm{GH}}_{\mathrm{SM}}$ becomes negative.

When the memory time is large, $\tilde{\Gamma}(\lambda) \approx \frac{\gamma}{\tau_{\mathrm{mem}}\lambda}$ and hence $\lambda^2 + \frac{\gamma}{m\tau_{\mathrm{mem}}} = \omega^2_{\mathrm{max}}$. Taking the positive root, i.e. that $\lambda = \left[\omega^2_{\mathrm{max}} - \frac{\gamma}{m\tau_{\mathrm{mem}}}\right]^{1/2}$, the prediction for the Grote-Hynes theory is

$$\tau^{\mathrm{GH}} = \frac{2\pi}{\sqrt{\omega^2_{\mathrm{max}} - \frac{\gamma}{m\tau_{\mathrm{mem}}}}}\frac{\omega_{\mathrm{max}}}{\omega_{\mathrm{min}}}\mathrm{e}^{\beta U_0}. \quad [39]$$

For the long memory (LM) time limit ($\tau_{\mathrm{mem}} \to \infty$), we recover the result for the transition state theory

$$\tau^{\mathrm{GH}}_{\mathrm{LM}} = \frac{2\pi}{\omega_{\mathrm{min}}}\mathrm{e}^{\beta U_0} = \frac{2\pi\sqrt{m}}{\sqrt{U''_{\mathrm{min}}}}\mathrm{e}^{\beta U_0}. \quad [40]$$

This final result is interesting for two reasons. Firstly, it is independent of τ_{mem}. This can be seen in Fig. S10B, where for each $\tau_{\mathrm{m}}/\tau_{\mathrm{D}}$, the GH curve reaches a constant value for $\tau_{\mathrm{mem}} \to \infty$. Secondly, this result tells us that as $m \to 0$, τ^{GH} goes to zero for large memory times. This can also be seen in Fig. S10B since as $\tau_{\mathrm{m}}/\tau_{\mathrm{D}} \to 0$, the limiting value of τ^{GH} approaches zero.

The non-Markovian prediction has different limiting behaviour for large memory times, irrespective of mass. For a single component memory kernel in the limit of $\tau_{\mathrm{m}}/\tau_{\mathrm{D}} \to 0$, it is given by

$$\tau^{\mathrm{H}}_{\mathrm{MPT}} = \tau_{\mathrm{D}}\frac{\mathrm{e}^{\beta U_0}}{\beta U_0}\left[\frac{\pi}{2\sqrt{2}}\frac{1}{1 + 10\beta U_0\tau_{\mathrm{mem}}/\tau_{\mathrm{D}}} + 4\beta U_0\left(\frac{\tau_{\mathrm{mem}}}{\tau_{\mathrm{D}}}\right)^2\right]. \quad [41]$$

It should be noted that the quadratic asymptotic limit for memory times has been derived analytically and confirmed numerically (17–19). Assuming that $10\beta U_0\tau_{\mathrm{mem}}/\tau_{\mathrm{D}} \gg 1$, we minimize Eq. 41 such that

$$\frac{\partial}{\partial\tau_{\mathrm{mem}}}\left[\frac{\pi}{2\sqrt{2}}\frac{1}{10\beta U_0\tau_{\mathrm{mem}}/\tau_{\mathrm{D}}}+4\beta U_0\left(\frac{\tau_{\mathrm{mem}}}{\tau_{\mathrm{D}}}\right)^2\right]=0, \qquad [42]$$

which leads to a minimum in $\tau^{\mathrm{H}}_{\mathrm{MPT}}$ at τ^*_{mem}, which is given by

$$\tau^*_{\mathrm{mem}}=\left[\frac{\pi\tau_{\mathrm{D}}^3}{160\sqrt{2}(\beta U_0)^2}\right]^{\frac{1}{3}}. \qquad [43]$$

The minimum in the non-Markovian formula is then

$$\tau^{\mathrm{H}*}_{\mathrm{MPT}}=3\tau_{\mathrm{D}}\frac{\mathrm{e}^{\beta U_0}}{(\beta U_0)^{\frac{4}{3}}}\left(\frac{\pi}{20\sqrt{2}}\right)^{\frac{2}{3}}. \qquad [44]$$

For $\tau_{\mathrm{mem}} > \tau^*_{\mathrm{mem}}$, we enter the memory-induced slow-down regime that is not accounted for by the GH theory. The minima for each $\tau_{\mathrm{m}}/\tau_{\mathrm{D}}$ can be seen in Fig. S10B, as well as for all proteins in Fig. S8. From Eq. 41, we can see that, for negligible mass contribution, the second term with quadratic scaling dominates as $\tau_{\mathrm{mem}}/\tau_{\mathrm{D}} \to \infty$, which is again shown in Fig. S10B. This divergence of behaviours between the Grote-Hynes theory and the non-Markovian prediction for long memory times and small masses is relevant for protein folding kinetics, as seen in Fig. 4 of the main manuscript.

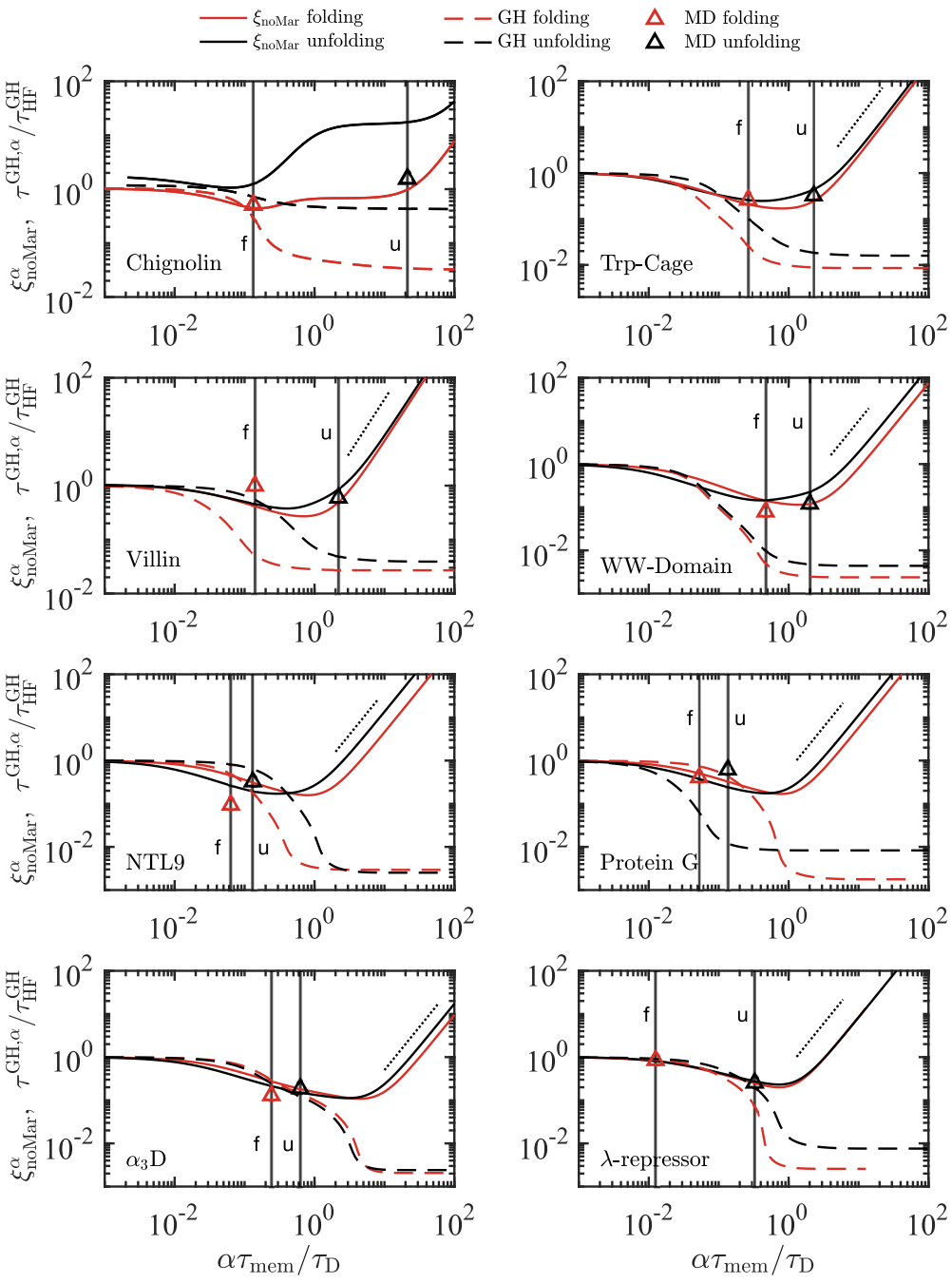

Fig. S8. The non-Markovian correction factor ξ_{noMar} (solid lines) and the Grote-Hynes theory prediction (dashed lines) for the folding (red) and unfolding (black) reactions of eight proteins are compared. Memory times are uniformly scaled by a factor of α to explore the small and long memory-time limits. As $\tau_{\mathrm{mem}}/\tau_{\mathrm{D}} \to 0$, all curves, except for Chignolin unfolding, reach their high-friction limit, indicating that finite mass effects are irrelevant for proteins. In the long memory time limit, $\xi_{\mathrm{noMar}} \sim (\tau_{\mathrm{mem}}/\tau_{\mathrm{D}})^2$ (dotted lines show quadratic scaling), and the GH curves reach a constant value, demonstrating disagreement between the two theories in the long memory-time limit. Triangles represent the barrier crossing times for MD simulations, coinciding with $\alpha = 1$, indicated with vertical lines. The theoretical predictions for $\alpha = 1$ for the two theories are given by the curve-intercepts with the vertical lines. The corresponding $\alpha = 1$ vertical lines for folding (f) and unfolding (u) differ because τ_{D} is different for the folding and unfolding processes.

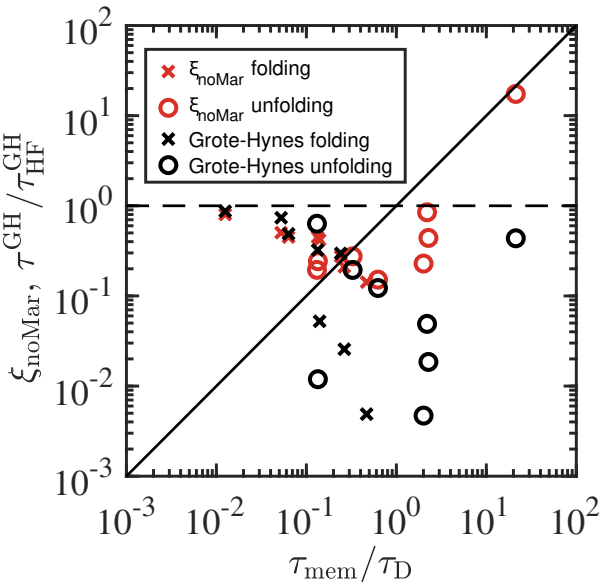

Fig. S9. Comparison of the Grote-Hynes theory to the non-Markovian correction factor ξ_{noMar} for protein folding and unfolding. Results for ξ_{noMar} are already presented in Fig. 4A of the main manuscript where they are compared to the barrier crossing times extracted from the MD simulations $\tau^{\mathrm{MD}}_{\mathrm{MFP}}/\tau^{\mathrm{Mar}}_{\mathrm{MFP}}$.

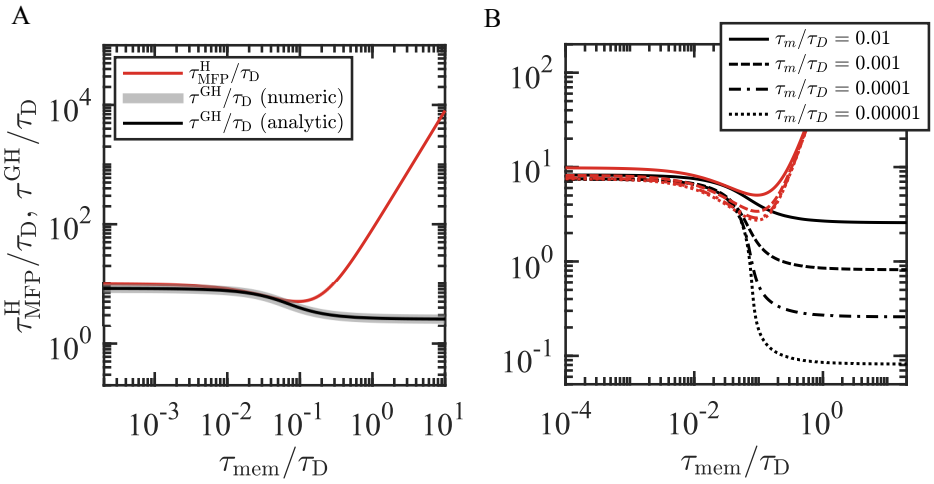

Fig. S10. Comparing the Grote-Hynes theory to the non-Markovian formula (Eq. 4 in the main manuscript, Eq. 22 in this SI) for a simple model system. Theoretical curves for a GLE with a single component exponential memory kernel $\Gamma(t) = \gamma \mathrm{e}^{-t/\tau}/\tau$ in a double-well potential $U(Q) = [(Q/L)^2 - 1]^2$. A) For comparison, parameters are taken from Kappler *et. al.* (17). $L = 1$, $\tau_{\mathrm{m}}/\tau_{\mathrm{D}} = 0.01$, $U_0/k_{\mathrm{B}}T = 3$. $U''_{\mathrm{min}} = 8U_0$ and $U''_{\mathrm{max}} = -4U_0$. The black curve is given by solving the cubic equation following from Eq. 33 analytically. The thick, grey curve is given by solving Eq. 33 numerically, which is suitable for more complicated memory kernels. B) The long memory limit for the GH theory decreases for vanishing mass ($\tau_{\mathrm{m}}/\tau_{\mathrm{D}} \to 0$), which is relevant for protein folding since finite mass effects are negligible.

11. Comparison of various reaction coordinates

In the main text, we only consider the fraction of native contacts reaction coordinate Q. Alternative standard reaction coordinates for the analysis of MD protein simulation trajectories include:

1. the end-to-end distance (R_{e2e}): the magnitude of the Cartesian vector separating the first and last α-carbon atoms in a chain. For a chain constructed of N, the end-to-end distance is given by:

$$R_{\mathrm{e2e}}(t) = |\mathbf{r}_N(t) - \mathbf{r}_1(t)|. \quad [45]$$

2. the radius of gyration (R_{rgy}): for a peptide chain with centre of mass $\mathbf{r}_{\mathrm{com}}(t)$ and set of particle masses m_i, the radius of gyration is given by:

$$R_{\mathrm{rgy}}(t) = \sqrt{\frac{\sum_{i=1}^{N} m_i |\mathbf{r}_i(t) - \mathbf{r}_{\mathrm{com}}(t)|^2}{\sum_{i=1}^{N} m_i}}, \quad [46]$$

where the summation is over all N α-carbon atoms.

3. the root-mean-squared deviation ($R^{\mathrm{d}}_{\mathrm{rms}}$): $\mathbf{r}_i^0$ is the position of the i^{th} α-carbon in the native state (the native state is defined in the Materials and Methods). The root-mean-squared deviation for the configuration of a protein at time t, deviating from the native state, is given by:

$$R^{\mathrm{d}}_{\mathrm{rms}}(t) = \sqrt{\frac{1}{N}\sum_{i=1}^{N} \left(\tilde{\mathbf{r}}_i(t) - \mathbf{r}_i^0\right)^2}, \quad [47]$$

where $\tilde{\mathbf{r}}_i(t)$ is the position vector of the i^{th} α-carbon at time t, for a configuration that has been uniformly translated and rotated to minimize the RMS deviation from the native state.

4. the root-mean-squared pair-separation ($R^{\mathrm{ps}}_{\mathrm{rms}}$): $\mathbf{s}^0_{ij}$ is the pair-separation vector connecting the i^{th} and j^{th} residues in the native state, where i and j are indices from the native contacts list. $\mathbf{s}_{ij}(t) = \mathbf{r}_j(t) - \mathbf{r}_i(t)$ is the pair-separation vector

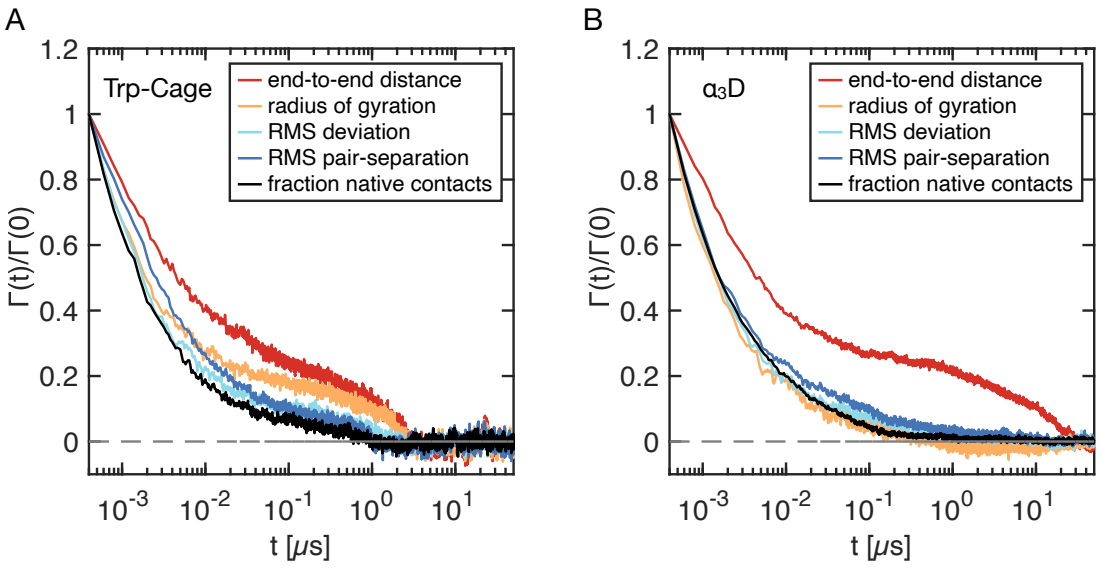

Fig. S11. Comparison of memory kernels for five different reaction coordinates. A) Memory kernels for Trp-Cage protein. B) Memory kernels for α_3D protein. Memory times, evaluate as the first moments of the memory kernels, are given in Table S7.

connecting the i^{th} and j^{th} residues at time t. $\mathbf{S}_{ij}(t,t') = \mathbf{s}_{ij}(t) - \mathbf{s}^0_{ij}$, then, is the time evolution of pair-separations, compared to native separations. The root-mean-squared pair-separation is given by:

$$R^{\text{ps}}_{\text{rms}}(t) = \sqrt{\frac{1}{N_{\text{nc}}} \sum_{i<j} \left(s_{ij}(t) - s^0_{ij}\right)^2}, \qquad [48]$$

where N_{nc} is the number of native contacts and $s_{ij} = |\mathbf{s}_{ij}|$ is the magnitude of the corresponding pair-separation vector. Note that here, the summation indices i and j run just over those pairs that form the native contacts list, of which there will be a total of N_{nc}.

The normalized memory kernels for these four reaction coordinates, plus the fraction of native contacts reaction coordinate Q (described in the Materials and Methods), as shown in Fig. S11, reveal large variation, depending on the choice of the reaction coordinate, and on the protein. For the Trp-Cage protein, the fraction of native contacts reaction coordinate has the shortest memory time-scale, a factor of 3 less than the end-to-end distance, which has the longest. For the α_3D protein, however, the radius of gyration memory kernel decays the fastest, two orders of magnitude faster than the end-to-end distance. The memory times for each reaction coordinate, evaluated as the first moments for each memory kernel, are given in Table S7.

Table S7. Memory times τ_{mem} for five different reaction coordinates, all given in units of μs. R_{e2e}, R_{rgy}, $R^{\text{d}}_{\text{rms}}$, **and** $R^{\text{ps}}_{\text{rms}}$ **are given by Eqs. 45-48.** Q **is given by Eq. M1 in the methods section of the main text.**

	R_{e2e}	R_{rgy}	$R^{\text{d}}_{\text{rms}}$	$R^{\text{ps}}_{\text{rms}}$	Q
Trp-Cage	1.44	1.65	0.9	0.65	0.48
α_3D	11.8	0.18	5.7	7.6	2.9

Heavy-atom Q reaction coordinates: We compare two variations of the fraction of native contacts reaction coordinate Q. In (20), Q was calculated using all heavy atoms in each residue, instead of just the Cα atoms, as we do in the main text. To test if this affects the measured memory times and folding kinetics, we recalculate the memory times τ_{mem} and the folding and unfolding mean first-passage times $\tau_{\text{MFP}}^{\text{MD}}$ for each protein. Additionally, we consider Q constructed using the minimum heavy atom distance between two residues to avoid ambiguity in defining contacts when using only C_α atom distances. The functional form for the reaction coordinates is the same as described in the methods section of the main manuscript (Eq. M1), with the only difference being the choice of atoms that comprise the list of native contacts.

In Fig.S12A, we compare memory times for Q constructed using all heavy atoms in each residue (Q-Heavy) to those determined using only the C_α atoms (Q-Cα) as presented in the main manuscript. We observe that the memory times are effectively equivalent for both constructions of Q. Similarly, for Q determined using the minimum heavy atom distances (Q-Heavy (minimum)), we find little variation in memory times (Fig.S12B), with a slight increase in overall memory times for Q-Heavy (minimum). In Figs. S12C and D, we compare memory times for Q-Heavy and Q-Heavy (minimum) to the folding and unfolding times measured in each reaction coordinate $\tau_{\text{MFP}}^{\text{MD}}$. We find that the overall behavior is similar to that presented in the main manuscript for Q-Cα, with folding times typically shorter than, but comparable to, the corresponding reaction times for each protein.

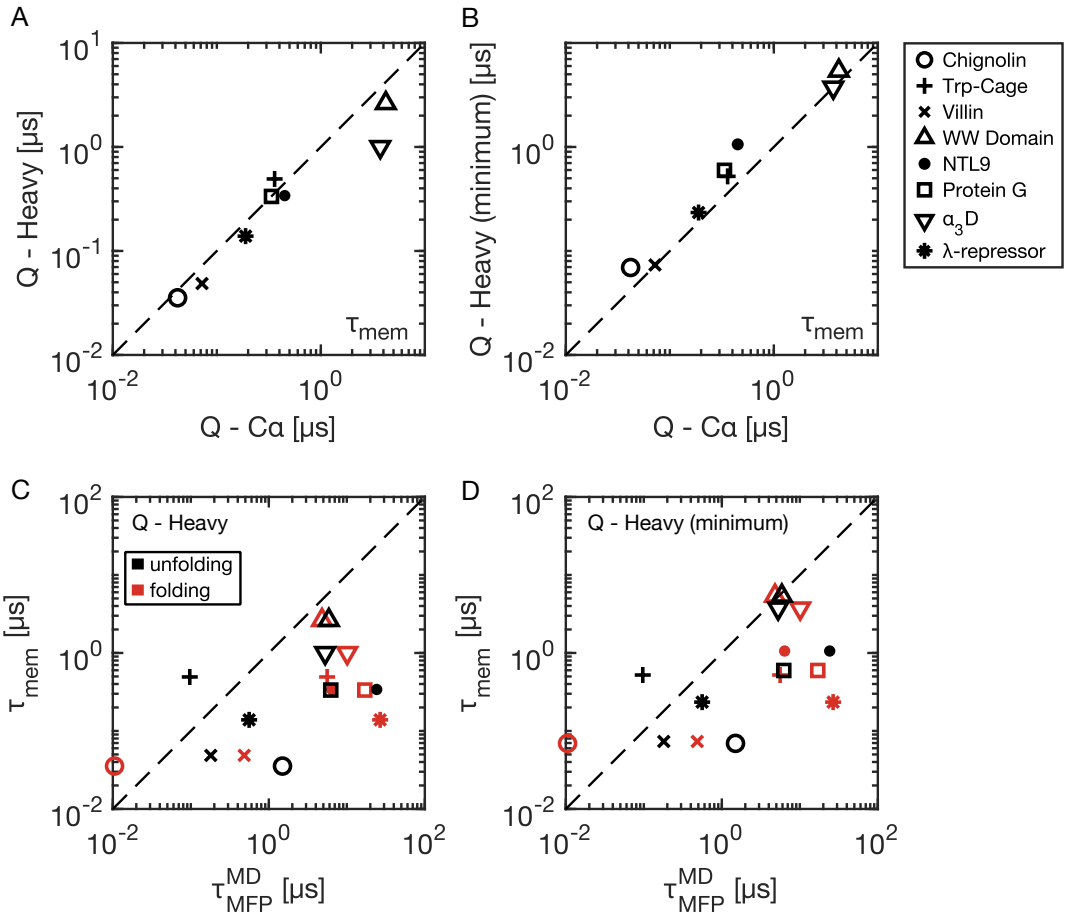

Fig. S12. Comparing the fraction of native contacts reaction coordinate Q constructed using only C_α atoms (Q - Cα) to Q constructed using all heavy atoms (Q - Heavy), or the minimal heavy-atom separation (Q - Heavy (minimum)). A) Memory times τ_{mem} for trajectories generated by Q - Cα compared to memory times for trajectories generated by Q - Heavy. Individual proteins are indicated by the symbols given in the figure legend. B) Memory times for Q - Cα compared to memory times evaluated for Q - Heavy (minimum), which is a Q reaction coordinate constructed using only the minimal heavy atom distance between the native contact residues. C) and D) the memory times for the Q - Heavy and Q - Heavy (minimum) reaction coordinates, compared against the folding and unfolding mean first-passage times $\tau_{\text{MFP}}^{\text{MD}}$ recalculated for the protein trajectories corresponding to these reaction coordinates.

12. Evaluating Markovian mean first-passage times: position-dependent friction

In Eq. 3 of the main manuscript, we show an equation to predict barrier crossing times using memoryless constant friction. One can add position-dependent friction, such that (21)

$$\tau_{\mathrm{MFP}}^{\mathrm{Mar}}\left(Q_{\mathrm{s}},Q_{\mathrm{e}}\right)=\beta\int_{Q_{\mathrm{s}}}^{Q_{\mathrm{e}}}\gamma(x)\mathrm{e}^{\beta U(x)}\left[\int_{Q_{\mathrm{min}}}^{x}\mathrm{e}^{-\beta U(y)}dy\right]dx, \quad [49]$$

where the superscript *Mar* indicates that this is a purely Markovian prediction. The quantity $\tau_{\mathrm{MFP}}^{\mathrm{Mar}}(Q_{\mathrm{s}},Q_{\mathrm{e}})$ can be extracted from the $Q(t)$ trajectories for each protein, where Q_{s}, and Q_{e} are appropriately chosen for both folding and unfolding processes. Eq. 49 is suitable when $Q_{\mathrm{s}}<Q_{\mathrm{e}}$, which is the case for folding in Q. When $Q_{\mathrm{e}}>Q_{\mathrm{s}}$, i.e. unfolding, we write

$$\tau_{\mathrm{MFP}}^{\mathrm{Mar}}\left(Q_{\mathrm{s}},Q_{\mathrm{e}}\right)=\beta\int_{Q_{\mathrm{e}}}^{Q_{\mathrm{s}}}\gamma(x)\mathrm{e}^{\beta U(x)}\left[\int_{x}^{Q_{\mathrm{max}}}\mathrm{e}^{-\beta U(y)}dy\right]dx. \quad [50]$$

Here, Q_{min} and Q_{max} are the left and right bounds of the free-energy profile $U(Q)$, respectively.

Having extracted $\tau_{\mathrm{MFP}}^{\mathrm{Mar}}(Q_{\mathrm{s}},Q_{\mathrm{e}})$, we can evaluate $\gamma(Q_{\mathrm{e}})$ in both the folding and unfolding directions. For $Q_{\mathrm{s}}<Q_{\mathrm{e}}$, we take the derivative of Eq. 49 with respect to Q_{e}:

$$\frac{\partial\tau_{\mathrm{MFP}}^{\mathrm{Mar}}\left(Q_{\mathrm{s}},Q_{\mathrm{e}}\right)}{\partial Q_{\mathrm{e}}}=\beta\gamma(Q_{\mathrm{e}})\mathrm{e}^{\beta U(Q_{\mathrm{e}})}\left[\int_{Q_{\mathrm{min}}}^{Q_{\mathrm{e}}}\mathrm{e}^{-\beta U(y)}dy\right], \quad [51]$$

which we then solve for $\gamma(Q_{\mathrm{e}})$:

$$\gamma(Q_{\mathrm{e}})=\frac{1}{\beta Z_{\mathrm{f}}}\frac{\partial\tau_{\mathrm{MFP}}^{\mathrm{Mar}}(Q_{\mathrm{s}},Q_{\mathrm{e}})}{\partial Q_{\mathrm{e}}}\mathrm{e}^{-\beta U(Q_{\mathrm{e}})} \quad [52]$$

where Z_{f} is given by:

$$Z_{\mathrm{f}}=\int_{Q_{\mathrm{min}}}^{Q_{\mathrm{e}}}\mathrm{e}^{-\beta U(y)}dy. \quad [53]$$

For $Q_{\mathrm{s}}>Q_{\mathrm{e}}$, we take the derivative of Eq. 50 also with respect to Q_{e}:

$$\frac{\partial\tau_{\mathrm{MFP}}^{\mathrm{Mar}}\left(Q_{\mathrm{s}},Q_{\mathrm{e}}\right)}{\partial Q_{\mathrm{e}}}=-\beta\gamma(Q_{\mathrm{e}})\mathrm{e}^{\beta U(Q_{\mathrm{e}})}\left[\int_{Q_{\mathrm{e}}}^{Q_{\mathrm{max}}}\mathrm{e}^{-\beta U(y)}dy\right], \quad [54]$$

which we then solve for $\gamma(Q_{\mathrm{e}})$:

$$\gamma(Q_{\mathrm{e}})=-\frac{1}{\beta Z_{\mathrm{u}}}\frac{\partial\tau_{\mathrm{MFP}}^{\mathrm{Mar}}(Q_{\mathrm{s}},Q_{\mathrm{e}})}{\partial Q_{\mathrm{e}}}\mathrm{e}^{-\beta U(Q_{\mathrm{e}})} \quad [55]$$

where Z_{u} is given by:

$$Z_{\mathrm{u}}=\int_{Q_{\mathrm{e}}}^{Q_{\mathrm{max}}}\mathrm{e}^{-\beta U(y)}dy. \quad [56]$$

In Figs. S13A and 1B, we present the friction profiles $\gamma(Q_{\mathrm{e}})$ for two example proteins evaluated in both the folding and unfolding directions. The profiles are evaluated between the folded and unfolded state minima, as indicated by the free-energy profile overlaid on the figures. The Markovian model Eq. 49 predicts that the friction experienced by each protein during folding is different from that experienced during unfolding, as previously demonstrated for the 9-residue alanine homo-peptide (12). Thus, there is no consistent way to describe the folding and unfolding dynamics simultaneously using a Markovian model. We verify this by reconstructing the barrier-crossing-time profiles for folding and unfolding while permuting the friction. The corresponding extracted friction profiles reproduce the folding and unfolding time profiles (Figs. S13C and 1D), serving as a consistency check. However, we cannot predict folding times across Q using the unfolding friction, validating that there is no unique friction profile that belongs to the reaction coordinate. Therefore, there is no consistent way to predict folding kinetics with a position-dependent Markovian friction model.

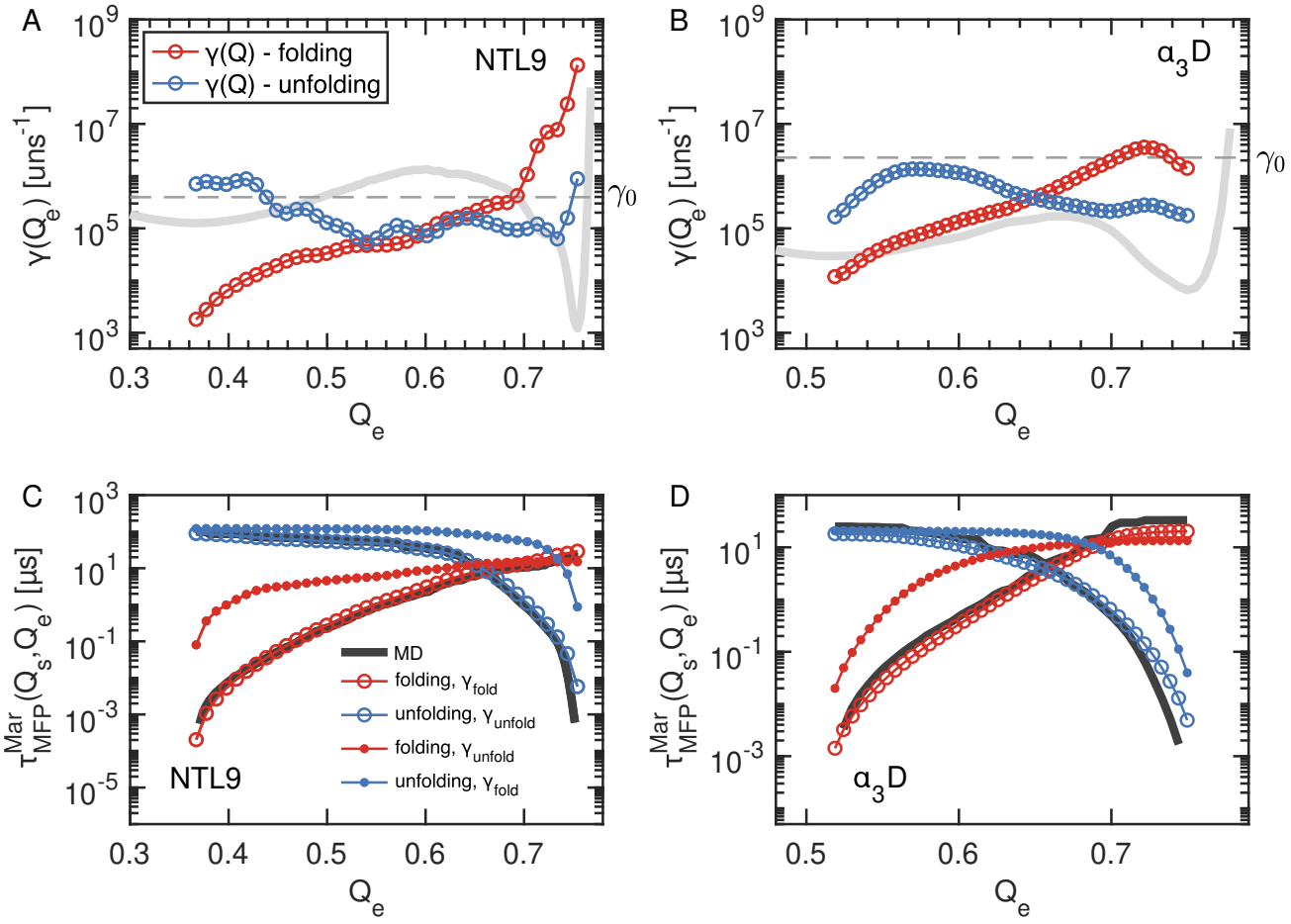

Fig. S13. Friction that varies with position is direction-dependent. A) and B) display the dependence of memoryless friction on the reaction coordinate position, calculated for the folding (red) or unfolding (blue) direction, for the NTL9 and α_3D proteins, respectively. The start-point of a transition, Q_s, is fixed, while Q_e, the end-point, serves as the independent variable. The light gray free-energy profile shows the minima and barriers of interest. The dashed lines represent the value of total friction for each protein, evaluated by the memory kernel extractions (Table S4). C) and D) illustrate various calculations of barrier crossing time profiles for the NTL9 and α_3D proteins, respectively. The bold black lines exhibit $\tau_{\mathrm{MFP}}(Q)$, derived from the MD trajectories. The open circles show a validation calculation of the crossing times with the corresponding $\gamma(Q)$ profiles from A) and B). The filled circles present an attempted construction of $\tau_{\mathrm{MFP}}(Q)$, using the anti-corresponding $\gamma(Q)$.

13. Additional confirmation of the linear friction non-Markovian model for the Q reaction coordinate

We simulate the GLE described by Eq. 1 using the Markov embedding methods described in (12) for a given memory kernel $\Gamma(t)$ and free-energy profile $U(Q)$. In Fig. S14, we show that the GLE reproduces the dynamics of the fraction of native contacts reaction coordinate $Q(t)$ for two example proteins (α_3D and the λ-repressor), as confirmed by the agreement of the mean square displacements (MSD) evaluated from the GLE simulation (red lines) with those extracted from the MD data (open circles). Notably, the GLE simulation, which we parametrize with the numerically extracted memory kernels, reproduces the same sub-diffusive scaling observed in the MD-extracted MSDs. Pure-Markovian Langevin simulations using position-independent friction (with values of γ from Table S4 - blue lines) over the same potential of mean force $U(Q)$ do not reproduce the sub-diffusive scaling, which indicates that the memory effects, rather than the free energy profile, cause the observed sub-diffusive behavior in the MD MSD profiles. Both the GLE and Markovian simulations reproduce the plateau caused by the confinement due to the free-energy profile.

We used a fourth-order Runge-Kutta integrator and a time step of $100\tau_m$ for both Langevin and GLE simulations. The value of τ_m was 1.5 ps for α_3D and 4.9 ps for λ-repressor. We averaged the MSD over 1000 independent trajectories, each containing 10^8 steps. The MSD of the GLE simulation exhibits oscillations at short times due to the absence of short time scale components in our extracted memory kernels. We could not extract these components due to discretization effects in the original MD data (see Section 7).

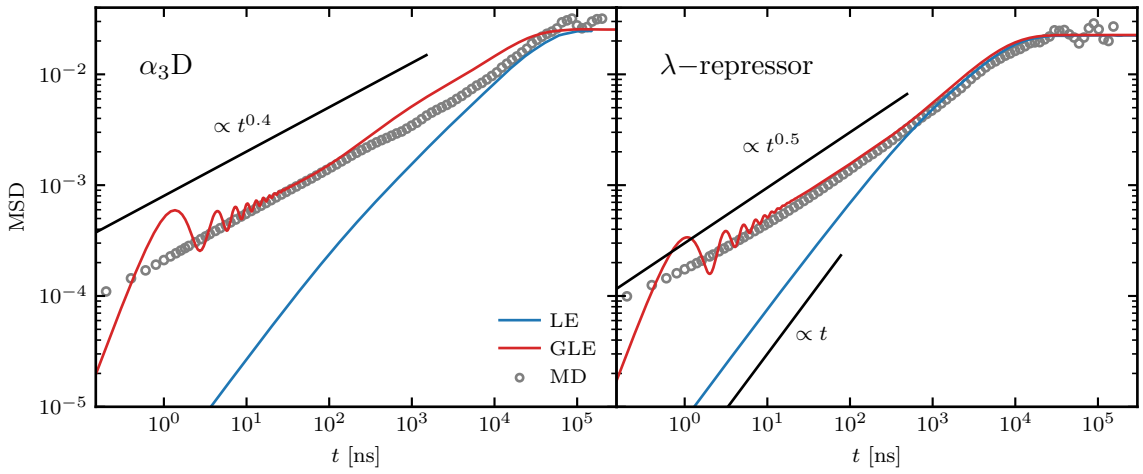

Fig. S14. Mean square displacements (MSD) for the fraction of native contacts trajectories extracted from the MD data of the α_3D and λ-repressor proteins (grey circles) agree well with the Markovian embedding GLE simulation using the fitted memory kernel (red lines). The simulation of the Markovian Langevin equation (LE) using the position-independent friction (blue line) does not reproduce the correct dynamics, but only the long-time behaviour.

Nonlinear friction effects: We use the GLE in Eq.1 of the main manuscript and Eq.1, derived by following the projection formalism of Ayaz et al. (11), which neglects non-linear friction effects. The two-projector formalism suggested by Vroylandt *et. al.* (22) can rigorously derive a GLE with friction linear in the velocity, as in Eq.1. In this formulation, all non-linear friction or position-dependent memory effects are absorbed in the random force, with the consequence that the fluctuation-dissipation theorem (FDT), i.e., $\langle F_R(t)F_R(t')\rangle = k_BT\Gamma(t-t')$, is violated. A modified version of the linear-friction GLE with a decomposition of the friction memory kernel, i.e., $\Gamma(t) = \Gamma^R(t) + \Gamma^\Delta(t)$, is useful, where $\Gamma^R(t)$ is the autocorrelator of the random force, i.e., $\langle F_R(t)F_R(t')\rangle = k_BT\Gamma^R(t-t')$. The FDT for the GLE in Eq.1 is valid if $\Gamma^\Delta(t) = 0$. Otherwise, a violation of the FDT indicates position-dependent memory effects. We check the FDT for the $Q(t)$ reaction coordinate of the α_3D protein in Fig.S15. In Fig.S15A, we show the probability distribution of the random force $F_R(t)$, obtained by inverting the GLE in Eq. 1

$$F_R(t) = m\ddot{Q}(t) + \int_0^t \Gamma(t-t')\dot{Q}(t')dt' + \nabla U\left[Q(t)\right], \qquad [57]$$

where we use the extracted free energy profile $U(Q)$ and memory kernel $\Gamma(t)$ from the data shown in Fig. 1 in the main text, and we discretized Eq. 57 to obtain the trajectory of the random force. Comparing the trajectory distribution with a Gaussian with zero mean and standard deviation $\sigma = \sqrt{k_BT\Gamma(0)}$, we see very weak non-Gaussian behaviour, with a probability below 1% of the maximum value. Fig. S15B shows the comparison between the extracted memory kernel $\Gamma(t)$ (solid black line) and the autocorrelation of the random force trajectory $\Gamma^R(t)$ (dashed red line, divided by k_BT), where we observe a very good agreement between both functions. The solid green line denotes the difference $\Gamma^\Delta(t)$, where in the inset, we see that values relative to the memory kernel always lie below 1%, a clear sign that the FDT between the random force and the memory kernel $\Gamma(t)$ is fulfilled and the GLE with linear friction in Eq. 1 is an accurate model for the Q reaction coordinate.

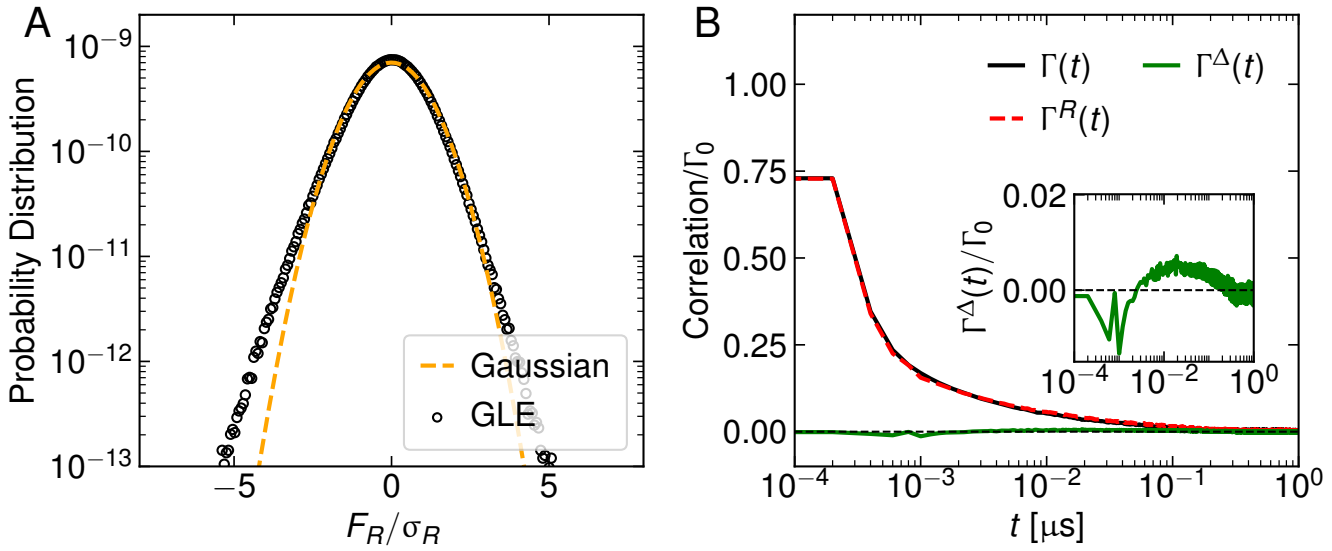

Fig. S15. Check for non-linear friction effects in the $Q(t)$ reaction coordinate of the α_3D protein. A) Probability distribution of the random force trajectory computed by the discretized version of Eq. 57 compared with a Gaussian with zero mean and standard deviation $\sigma = \sqrt{k_BT\Gamma_0}$. B) Comparison between the extracted memory kernel $\Gamma(t)$ and the autocorrelation of the random force trajectory ($\Gamma^R(t)$). The solid green line denotes the difference $\Gamma^\Delta(t) = \Gamma(t) - \Gamma^R(t)$ (22).

References

1. K Lindorff-Larsen, S Piana, RO Dror, DE Shaw, How Fast-Folding Proteins Fold. *Science* **334**, 517 LP – 520 (2011).
2. DE Shaw, et al., Millisecond-scale molecular dynamics simulations on Anton in *Proceedings of the Conference on High Performance Computing Networking, Storage and Analysis*. pp. 1–11 (2009).
3. S Piana, K Lindorff-Larsen, DE Shaw, How robust are protein folding simulations with respect to force field parameterization? *Biophys. journal* **100**, L47–9 (2011).
4. SH Northrup, JT Hynes, The stable states picture of chemical reactions. I. Formulation for rate constants and initial condition effects. *The J. Chem. Phys.* **73**, 2700–2714 (1980).
5. J Kubelka, TK Chiu, DR Davies, WA Eaton, J Hofrichter, Sub-microsecond protein folding. *J. molecular biology* **359**, 546–553 (2006).
6. S Piana, et al., Computational design and experimental testing of the fastest-folding β-sheet protein. *J. molecular biology* **405**, 43–48 (2011).
7. JC Horng, V Moroz, DP Raleigh, Rapid cooperative two-state folding of a miniature alpha-beta protein and design of a thermostable variant. *J. molecular biology* **326**, 1261–1270 (2003).
8. S Nauli, B Kuhlman, D Baker, Computer-based redesign of a protein folding pathway. *Nat. Struct. Biol.* **8**, 602–605 (2001).
9. Y Zhu, et al., Ultrafast folding of α3D: A de novo designed three-helix bundle protein. *Proc. Natl. Acad. Sci.* **100**, 15486–15491 (2003).
10. WY Yang, M Gruebele, Folding at the speed limit. *Nature* **423**, 193–197 (2003).
11. C Ayaz, L Scalfi, BA Dalton, RR Netz, Generalized Langevin equation with a nonlinear potential of mean force and nonlinear memory friction from a hybrid projection scheme. *Phys. Rev. E* **105**, 54138 (2022).
12. C Ayaz, et al., Non-Markovian modeling of protein folding. *Proc. Natl. Acad. Sci.* **118**, e2023856118 (2021).
13. B Kowalik, et al., Memory-kernel extraction for different molecular solutes in solvents of varying viscosity in confinement. *Phys. Rev. E* **100**, 12126 (2019).
14. PG de Gennes, Reptation of a Polymer Chain in the Presence of Fixed Obstacles. *The J. Chem. Phys.* **55**, 572–579 (1971).
15. A Alexander-Katz, H Wada, RR Netz, Internal Friction and Nonequilibrium Unfolding of Polymeric Globules. *Phys. Rev. Lett.* **103**, 28102 (2009).
16. KW Plaxco, KT Simons, D Baker, Contact order, transition state placement and the refolding rates of single domain proteins11Edited by P. E. Wright. *J. Mol. Biol.* **277**, 985–994 (1998).
17. J Kappler, JO Daldrop, FN Brünig, MD Boehle, RR Netz, Memory-induced acceleration and slowdown of barrier crossing. *The J. Chem. Phys.* **148**, 14903 (2018).
18. J Kappler, VB Hinrichsen, RR Netz, Non-Markovian barrier crossing with two-time-scale memory is dominated by the faster memory component. *The Eur. Phys. J. E* **42**, 119 (2019).
19. L Lavacchi, J Kappler, RR Netz, Barrier crossing in the presence of multi-exponential memory functions with unequal friction amplitudes and memory times. *EPL* **131** (2020).
20. RB Best, G Hummer, WA Eaton, Native contacts determine protein folding mechanisms in atomistic simulations. *Proc. Natl. Acad. Sci.* **110**, 17874 LP – 17879 (2013).
21. R Zwanzig, *Nonequilibrium Statistical Mechanics*. (Oxford University Press), (2001).
22. H Vroylandt, On the derivation of the generalized Langevin equation and the fluctuation-dissipation theorem. *Europhys. Lett.* **140**, 62003 (2022).