Supporting Information

Persson and Halle 10.1073/pnas.1506079112

SI Materials and Methods

Simulation Data. The analysis was based on a previously reported 1.031-ms MD simulation at 300 K of BPTI solvated by 4,215 water molecules (1). The trajectory was sampled with a resolution of $\Delta\tau = 0.25$ ns, yielding 4,125,000 frames. Using VMD (2), we aligned all protein atoms in each frame with the first frame in the trajectory. For the Amber ff99SB-I/TIP4P-Ew force field used in the simulation, the rotational isomer populations for the C14–C38 disulfide bond are known (3–6) to differ substantially from the experimentally deduced ones (7, 8). The analysis was therefore restricted to the subset of $N_F = 1{,}048{,}349$ frames (corresponding to a 262.1-μs-long trajectory) where the 14–38 disulfide bond is in the experimentally dominant (95%) M1 conformation. We refer to this subset as "the trajectory." If this constraint is applied only to the amide C state, we find that the disulfide remains in the M1 conformation also in the amide O state, except for two amides (Gly37 and Cys38) near the disulfide bond (Fig. S7). For these two amides, 60–70% of the O-state frames coincide with the disulfide M2 conformation, which also dominates for the transient water tunnel that constitutes the open state for exchange of the internal water molecule (W122) buried under the 14–38 disulfide bond (6).

Experimental Data. To minimize errors introduced by temperature corrections, we gave preference to k_{HX} data measured as close as possible to the simulation temperature, 300 K. First we included data for 19 amides measured at 30° C (9) and for the amide of Gly37 reported at 10° C but back-corrected to the measurement temperature 30° C using an Arrhenius activation energy $E_A = 30$ kcal·mol^{-1} (10). Then we added data for nine amides that were measured at 36° C and extrapolated to 30° C with $E_A = 35$ kcal·mol^{-1} (9) and for the three most slowly exchanging amides (in residues 21–23) measured at high temperature and extrapolated to 30° C with $E_A = 78$ kcal·mol^{-1} (9). (The latter three amides were not used for the quantitative comparison because they do not access the O state in the simulation.) The data for these 32 amides were then corrected from 30° C to 300 K using the activation energies quoted above and, for the first set of 19 amides, $E_A(\text{kcal}\cdot\text{mol}^{-1}) = -1.17 - 9.306\,\log[k_{HX}(30\,°\text{C}, \text{pH*}\,3.5)]$, a relationship obtained from a linear fit for 11 amides that were measured at both 10 and 30° C (9, 10). Finally, we included data for nine amides that were only measured at 10° C (10), extrapolated to 300 K with $E_A(\text{kcal}\cdot\text{mol}^{-1}) = -0.73 - 8.072\,\log[k_{HX}(10\,°\text{C}, \text{pH*}\,4.6)]$. The HX rate constants for these 41 amides, corrected to 300 K, are given in Table S1.

The PF was obtained from the experimental HX rate constant, k_{HX}, as $\kappa_{exp} = k_{int}/k_{HX}$, where $k_{int} = k_0 + k_1\,[D_3O^+] + k_2\,[OD^-] = k_0 + k_1\,10^{-pD} + k_2\,10^{pD-pK_W^D}$, where k_0, k_1, and k_2 are, respectively, the rate constants for uncatalyzed, acid-catalyzed, and base-catalyzed HX. Furthermore, $pD = pH^* + 0.4$ (11) and pK_W^D is the negative logarithm of the ionization constant for D_2O (12). Reference rate constants for poly-DL-alanine in D_2O were used with values at 20° C of $k_0^0 = 10^{-1.5}$ min^{-1}, $k_1^0 = 10^{1.62}$ min^{-1}, and $k_2^0 = 10^{10.05}$ min^{-1} (13), and with activation energies $E_{A,0} = 13$ kcal·mol^{-1}, $E_{A,1} = 15$ kcal·mol^{-1}, and $E_{A,2} = 2.6$ kcal·mol^{-1} taken from Roder's Sphere program (www.fccc.edu/research/labs/roder/sphere/sphere.html). Nearest-neighbor effects were corrected for according to $k_0 = B_L\,B_R\,k_0^0$, $k_1 = A_L\,A_R\,k_1^0$, and $k_2 = B_L\,B_R\,k_2^0$, where, for example, B_L refers to the side chain that has the considered amide group to its left. The correction factors were taken from ref. 13, except for the basic forms of Asp and Glu, where the factors were taken from ref. 14. The C-terminal carboxyl group (which is involved in a salt bridge) was taken to be fully deprotonated at the experimental pH* values, whereas for the other carboxyl groups we used experimentally determined pK_a^D values: 3.9 (Asp3), 3.5 (Asp50), 4.2 (Glu7), and 4.3 (Glu49) (15). The temperature dependence of these pK_a^D values is negligibly small.

In this way, we obtained k_{int} at 300 K and at the experimental pH* values of 3.5 or 4.6. Even at the lower pH* value, the basic mechanism dominates strongly over the acidic mechanism. The uncatalyzed mechanism makes a minor but significant contribution. The 41 PFs obtained by combining these k_{int} values with the experimental k_{HX} values are listed in Table S1. The experimental uncertainties in k_{int} and k_{HX} are not known, but we note that ±10% error propagates to ±0.14 in $\beta\,\Delta G_{exp}$ shown in Fig. 2*A*.

O-State Correlation Matrix. To investigate the temporal correlation of the O state of different amides, we consider the joint probability, $P(n,n')$, that amides n and n' are in the O state at the same time. Defining an indicator function, $h(n,k)$, such that $h = 1$ if amide n is in the O state in frame k and $h = 0$ otherwise, we compute $P(n,n') = \langle h(n,k)\,h(n',k)\rangle$ by averaging over all frames in the trajectory. We then compute the O-state correlation matrix $C(n,n') = [P(n,n') - P(n)P(n')]\{P(n)P(n')[1-P(n)][1-P(n')]\}^{-1/2}$, where $P(n) = \langle h(n,k)\rangle$ is the a priori probability that amide n is in state O. For uncorrelated amides, $P(n,n') = P(n)\,P(n')$ so that $C(n,n') = 0$; for perfectly correlated amides, $P(n,n') = P(n) = P(n')$ so that $C(n,n') = 1$; and for perfectly anticorrelated amides $C(n,n') = -1$.

Binning Error. Fig. S8 shows a small part of a trajectory sampled in two different ways. On the upper continuous time line, O → C and C → O transitions are indicated by tick marks. On the lower discretely sampled time line, the observation time points, that is, the frames saved for analysis, are indicated by dots separated by the time interval $\Delta\tau$, which is the sampling resolution. For each frame, the observed state is indicated.

The distribution of state residence times (RTs) in the continuous-time trajectory, that is, the separation of adjacent tick marks on the upper time line in Fig. S8, is described by the continuous probability densities $\psi_O^0(\tau)$ and $\psi_C^0(\tau)$, such that $\psi_O^0(\tau)d\tau$ is the fraction of all O-state RTs that are within $d\tau$ of τ in length. Here, and in the following, quantities pertaining to the continuous-time trajectory are indicated by a zero superscript.

The total length T of the trajectory is the sum of the total time T_O^0 spent in state O and the total time T_C^0 spent in state C,

$$T = T_O^0 + T_C^0. \qquad \text{[S1]}$$

These parts may be written as

$$T_O^0 = N_O^0\,\tau_O^0, \qquad \text{[S2a]}$$

$$T_C^0 = N_C^0\,\tau_C^0, \qquad \text{[S2b]}$$

where N_O^0 is the total number of O-state residences in the continuous-time trajectory and τ_O^0 is the mean RT in the O state, that is,

$$\tau_O^0 = \int_0^\infty d\tau\;\tau\;\psi_O^0(\tau), \qquad \text{[S3]}$$

and similarly for state C.

Analogous relations hold for the discretely sampled trajectory, which is of the same total length T as the continuous trajectory, namely,

$$T = T_O + T_C, \quad \text{[S4]}$$

with

$$T_O = N_O\ \tau_O = N_{FO}\ \Delta\tau, \quad \text{[S5a]}$$

$$T_C = N_C\ \tau_C = N_{FC}\ \Delta\tau, \quad \text{[S5b]}$$

where N_{FO} and N_{FC} are the total number of frames in states O and C, respectively. The mean RT in state O, as judged from the discrete trajectory, is

$$\tau_O = \Delta\tau\langle n_O\rangle = \Delta\tau\ \frac{1}{N_O}\ \sum_{\alpha=1}^{N_O} n_{O,\alpha} = \Delta\tau\ \frac{N_{FO}}{N_O}, \quad \text{[S6]}$$

and similarly for τ_C. Here, $n_{O,\alpha}$ is the number of frames in the α^{th} visit to state O.

The PF is estimated from the discrete trajectory as $\kappa = N_{FC}/N_{FO}$, which, in view of Eq. **S5**, can also be expressed as

$$\kappa = \frac{T_C}{T_O} = \frac{N_C\ \tau_C}{N_O\ \tau_O}. \quad \text{[S7]}$$

Because of the finite sampling resolution $\Delta\tau$, this PF estimate is subject to a systematic binning error. Ideally, we would like to compare the experimental PF with a theoretical PF based on the continuous trajectory (in the limit $\Delta\tau \to 0$), that is,

$$\kappa^0 = \frac{T_C^0}{T_O^0} = \frac{N_C^0\ \tau_C^0}{N_O^0\ \tau_O^0}. \quad \text{[S8]}$$

We would also like to know the true mean RTs τ_O^0 and τ_C^0, rather than the estimates τ_O and τ_C based on the discrete trajectory.

The origin of the binning error is evident from Fig. S8, which shows two short O-state residences, with $\tau_{O,\alpha} < \Delta\tau$. Some of these short residences will be assigned a too-long RT, namely $\tau_{O,\alpha} = \Delta\tau$ (this applies to the RT detected in frame 12), whereas others will not be detected at all (this applies to the RT between frames 14 and 15). However, some of the continuous residences between $n\Delta\tau$ and $(n+1)\Delta\tau$ will be assigned a too-short RT of $n\Delta\tau$. This random lengthening and shortening of RTs cancels out to first order, but the omission of some RTs gives rise to a systematic error of order $\Delta\tau/\tau_O^0$. The net effect of these binning errors is therefore to make τ_O longer than τ_O^0. However, the omission of some RTs makes N_O smaller than N_O^0. As we shall see, for a Poisson process, these effects precisely cancel out, so that $T_O = N_O\ \tau_O$ is equal to $T_O^0 = N_O^0\ \tau_O^0$, unaffected by the binning error. Similar considerations apply to the C state. Omissions of short O-state RTs makes N_C smaller than N_C^0, but the merging of the flanking C-state RTs as well as the "conversion" of the omitted O-state RT to the C state makes τ_C longer than τ_C^0. Again, for a Poisson process, these effects precisely cancel out, making $T_C = T_C^0$. Therefore, for a Poisson process, the PF is unaffected by the binning error, that is, $\kappa = \kappa^0$. We shall now demonstrate this explicitly and show how the mean RTs are affected by the binning error.

To correct quantitatively for the binning error, we assume that the conformational fluctuations that interconvert a given amide between the C and O states can be modeled as an alternating Poisson process (sometimes called a dual Poisson process). In other words, the RTs are independently and exponentially distributed with normalized probability densities

$$\psi_O^0(\tau) = \frac{1}{\tau_O^0}\exp\left(-\frac{\tau}{\tau_O^0}\right), \quad \text{[S9a]}$$

$$\psi_C^0(\tau) = \frac{1}{\tau_C^0}\exp\left(-\frac{\tau}{\tau_C^0}\right). \quad \text{[S9b]}$$

If all frames in the trajectory are used, the first and last RTs will in general be truncated. It would thus seem that the terminal RTs differ from the internal RTs, being shorter on average. However, for a Poisson process, the sampling bias (that a randomly selected time point is likely to fall in a long RT) precisely cancels the truncation effect. Therefore, we can describe all RTs, both internal and terminal, with the same probability density **S9**. Because it does not matter whether or not we prune the trajectory by removing the terminal RTs, they should be included to improve the statistical accuracy.

The RT histogram $F_O(n)$ is obtained by substituting $\psi_O^0(\tau)$ from Eq. **S9a** into equation S5 of ref. 16 and peforming the integrations. The result is, for $n \geq 1$,

$$F_O(n) = \frac{N_O^0}{x}e^{(1-n)x}(1-e^{-x})^2, \quad \text{[S10]}$$

where we have defined

$$x \equiv \frac{\Delta\tau}{\tau_O^0}. \quad \text{[S11]}$$

The number of (detected) O-state RTs can then be obtained as

$$N_O = \sum_{n=1}^{\infty} F_O(n) = N_O^0\ \frac{(1-e^{-x})}{x}, \quad \text{[S12]}$$

where, in the second step, we have summed the geometric series. Similarly, the number of (detected) O-state frames is obtained as

$$N_{FO} = \sum_{n=1}^{\infty} n\ F_O(n) = \frac{N_O^0}{x}. \quad \text{[S13]}$$

Combining Eqs. **S2a**, **S11**, and **S13**, we obtain

$$T_O = N_{FO}\ \Delta\tau = N_O^0\ \tau_O^0 = T_O^0. \quad \text{[S14]}$$

According to Eqs. **S1** and **S4**, $T_O + T_C = T_O^0 + T_C^0$. It therefore follows from Eq. **S14** that $T_C = T_C^0$ and, in view of Eqs. **S7** and **S8**, that

$$\kappa = \kappa^0. \quad \text{[S15]}$$

The PF is thus immune to the binning error if the O/C transitions can be described as an alternating Poisson process.

We now consider the effect of the binning error on the mean RTs. Using Eqs. **S11**–**S13**, we can write the O-state mean RT in the following ways:

$$\tau_O = \Delta\tau\ \frac{N_{FO}}{N_O} = \tau_O^0\ \frac{N_O^0}{N_O} = \tau_O^0\ \frac{x}{(1-e^{-x})} = \frac{\Delta\tau}{(1-e^{-x})}. \quad \text{[S16]}$$

Combining the second and fifth members of Eq. **S16** with Eq. **S11**, we can express τ_O^0 in terms of simulation-derived quantities as

$$\tau_O^0 = -\frac{\Delta\tau}{\ln(1-N_O/N_{FO})}. \quad \text{[S17]}$$

This result shows that for an amide where all O-state RTs are a single frame, so that $N_{FO} = N_O$, we cannot obtain τ_O^0. However, κ^0 (which is equal to κ) can still be obtained.

A Taylor expansion of the fourth member of Eq. **S16** yields

$$\frac{\tau_O}{\tau_O^0} = 1 + \frac{1}{2}x + \frac{1}{12}x^2 - \frac{1}{720}x^4 + \frac{1}{30{,}240}x^6 \cdots, \quad \textbf{[S18]}$$

showing that the binning error on the mean RT is of first order in $x = \Delta\tau/\tau_O^0$, as we have shown more generally (without assuming a Poisson process) in another context (16).

We now turn to the mean RT τ_C of the C state. Because the problem is symmetric in the two states, it may be thought that we could proceed in complete analogy with the preceding treatment of the O state. However, this is not so, because equation S5 of Ref. 16 is not strictly valid for an alternating (Poisson) process. That equation assigns with a certain probability RTs in the interval $(n-1)\Delta\tau < \tau \le (n+1)\Delta\tau$ to the n^{th} bin of the RT histogram $F(n)$. This takes care of the discretization error, which can make the discrete RT shorter or longer than the continuous RT, and it also allows for the possibility that some RTs (with $\tau < \Delta\tau$) escape detection. However, equation S5 of ref. 16 does not describe the other consequence of an omitted RT, namely, that the two flanking RTs are merged. In the present HX context, the O-state RTs are always short whereas the C-state RTs are always long. Therefore, the merging effect is only relevant for the C state. We can therefore use equation S5 of ref. 16 for the O state, but not for the C state. Instead, we shall exploit the complementary nature of the two states in the alternating process.

We established below Eq. **S14** that $T_C = T_C^0$. Combining this result with Eqs. **S2b** and **S5b**, we obtain

$$\tau_C^0 = \tau_C \frac{N_C}{N_C^0} = \Delta\tau \frac{N_{FC}}{N_C^0}, \quad \textbf{[S19]}$$

where, in the second step, we have used the second equality in Eq. **S5b**. Next, we note that for every O-state RT that is lost due to the finite sampling resolution, a C-state RT is also lost (when the two flanking C-state RTs are merged into one, together with the lost O-state RT). The number ΔN of undetected (or lost) RT pairs can therefore be written as

$$\Delta N = N_C^0 - N_C = N_O^0 - N_O. \quad \textbf{[S20]}$$

From Eqs. **S13** and **S17**, we find that

$$N_O^0 = -N_{FO} \ln(1 - N_O/N_{FO}), \quad \textbf{[S21]}$$

which is combined with Eq. **S20** to yield

$$N_C^0 = N_C - N_O\left[1 + \frac{N_{FO}}{N_O} \ln(1 - N_O/N_{FO})\right]. \quad \textbf{[S22]}$$

By inserting this result into Eq. **S19** we obtain an expression for the true mean RT for state C in terms of quantities derived from the MD trajectory. It is seen that for an amide where all O-state RTs are a single frame, so that $N_{FO} = N_O$, we cannot obtain τ_C^0. As noted above, this is also true for τ_O^0. In both cases, the reason is that, when all O-state RTs are a single frame, we have no quantitative information about the number ΔN of lost RT pairs (except that they are many). However, when $N_O \ll N_{FO}$, then Eq. **S22** shows that $N_C^0 = N_C$, as expected. If $N_{FC} \gg N_{FO}$ (meaning that $\kappa \gg 1$), as is the case for the amides examined here, then both of the terminal (truncated) RTs are with high probability associated with the C state. Then $N_C = N_O + 1$ so Eq. **S22** reduces to

$$N_C^0 = 1 - N_{FO} \ln(1 - N_O/N_{FO}). \quad \textbf{[S23]}$$

Of course, the distinction between N_C and N_O only matters if they are small. For most amides, $N_O \gg 1$ and then we can set $N_C = N_O$ irrespective of the state associated with the terminal RTs. If $N_C = N_O$ it follows from Eq. **S20** that also $N_C^0 = N_O^0$. Comparing Eqs. **S16** and **S19**, we then find that

$$\frac{\tau_C}{\tau_C^0} = \frac{\tau_O}{\tau_O^0}, \quad \textbf{[S24]}$$

showing that the binning error lengthens τ_C and τ_O by the same factor. The Taylor expansion in Eq. **S18** is thus valid also for τ_C/τ_C^0.

Statistical Error. We now consider the statistical uncertainties (or SEs) in the PF and mean RTs, resulting from the finite length of the MD trajectory. These errors will be significant if the number of detected RTs (N_O) is small. As in our treatment of the binning error, we model the O/C fluctuations by an alternating Poisson process, with the RT probability densities given by Eq. **S9**. Because the individual RTs are then independent, we can treat the O-state RTs separately from the C-state RTs. For a Poisson process, the variance is the square of the mean so the SEM RT is

$$\sigma(\tau_O) = \frac{\tau_O}{N_O^{1/2}}, \quad \textbf{[S25]}$$

with an analogous expression for $\sigma(\tau_C)$. The SE of the PF $\kappa = \tau_C/\tau_O$ is

$$\sigma(\kappa) = \left\{\left[\frac{\tau_C\ \sigma(\tau_O)}{\tau_O^2}\right]^2 + \left[\frac{\sigma(\tau_C)}{\tau_O}\right]^2\right\}^{1/2}, \quad \textbf{[S26]}$$

which is combined with Eq. **S25** to give

$$\sigma(\kappa) = \kappa\left[\frac{1}{N_O} + \frac{1}{N_C}\right]^{1/2}. \quad \textbf{[S27]}$$

1. Shaw DE, et al. (2010) Atomic-level characterization of the structural dynamics of proteins. *Science* 330(6002):341–346.
2. Humphrey W, Dalke A, Schulten K (1996) VMD: Visual molecular dynamics. *J Mol Graph* 14(1):33–38, 27–28.
3. Long D, Brüschweiler R (2011) Atomistic kinetic model for population shift and allostery in biomolecules. *J Am Chem Soc* 133(46):18999–19005.
4. Xue Y, Ward JM, Yuwen T, Podkorytov IS, Skrynnikov NR (2012) Microsecond time-scale conformational exchange in proteins: Using long molecular dynamics trajectory to simulate NMR relaxation dispersion data. *J Am Chem Soc* 134(5):2555–2562.
5. Pierce LCT, Salomon-Ferrer R, Augusto F de Oliveira C, McCammon JA, Walker RC (2012) Routine access to millisecond time scale events with accelerated molecular dynamics. *J Chem Theory Comput* 8(9):2997–3002.
6. Persson F, Halle B (2013) Transient access to the protein interior: Simulation versus NMR. *J Am Chem Soc* 135(23):8735–8748.
7. Otting G, Liepinsh E, Wüthrich K (1993) Disulfide bond isomerization in BPTI and BPTI(G36S): An NMR study of correlated mobility in proteins. *Biochemistry* 32(14):3571–3582.
8. Grey MJ, Wang C, Palmer AG (2003) Disulfide bond isomerization in basic pancreatic trypsin inhibitor: Multisite chemical exchange quantified by CPMG relaxation dispersion and chemical shift modeling. *J Am Chem Soc* 125(47):14324–14335.
9. Kim K-S, Fuchs JA, Woodward CK (1993) Hydrogen exchange identifies native-state motional domains important in protein folding. *Biochemistry* 32(37): 9600–9608.
10. Battiste JL, Li R, Woodward C (2002) A highly destabilizing mutation, G37A, of the bovine pancreatic trypsin inhibitor retains the average native conformation but greatly increases local flexibility. *Biochemistry* 41(7):2237–2245.
11. Covington AK, Paabo M, Robinson RA, Bates RG (1968) Use of the glass electrode in deuterium oxide and the relation between the standardized pD (pa_D) scale and the operational pH in heavy water. *Anal Chem* 40(4):700–706.
12. Covington AK, Robinson RA, Bates RG (1966) The ionization constant of deuterium oxide from 5 to 50°. *J Phys Chem* 70(12):3820–3824.
13. Bai Y, Milne JS, Mayne L, Englander SW (1993) Primary structure effects on peptide group hydrogen exchange. *Proteins* 17(1):75–86.

14. Mori S, van Zijl PCM, Shortle D (1997) Measurement of water-amide proton exchange rates in the denatured state of staphylococcal nuclease by a magnetization transfer technique. *Proteins* 28(3):325–332.
15. Richarz R, Wüthrich K (1978) High-field ^{13}C nuclear magnetic resonance studies at 90.5 MHz of the basic pancreatic trypsin inhibitor. *Biochemistry* 17(12):2263–2269.
16. Halle B, Persson F (2013) Analysis of protein dynamics simulations by a stochastic point process approach. *J Chem Theory Comput* 9(6):2838–2848.

Table S1. Experimental and simulated protection factors at 300 K

Residue	No.	$^{10}\log(k_{HX}/s^{-1})$	$^{10}\log\kappa_{exp}$	$^{10}\log\kappa_{sim}$	$\beta\,\Delta\Delta G$
Asp	3	—	—	1.047	—
Phe	4	—	—	3.337	—
Cys	5	−5.556	3.704	3.844	0.324
Leu	6	−5.810	3.229	3.149	−0.184
Glu	7	−6.092	3.574	3.561	−0.030
Tyr	10	−5.457	2.500	2.437	−0.146
Thr	11	—	—	1.046	—
Gly	12	−3.659	1.668	2.168	1.152
Cys	14	−4.551	2.400	—	—
Lys	15	—	—	1.114	—
Ala	16	−5.359	3.024	2.623	−0.925
Arg	17	−3.139	0.761	1.670	2.094
Ile	18	−8.344	5.375	4.573	−1.847
Ile	19	−4.396	0.980	2.002	2.355
Arg	20	−9.089	6.478	—	—
Tyr	21	−11.068	8.559	—	—
Phe	22	−11.068	8.419	—	—
Tyr	23	−11.068	8.400	—	—
Asn	24	−8.198	6.276	—	—
Ala	25	−3.458	1.323	0.925	−0.916
Lys	26	—	—	3.396	—
Ala	27	−5.021	2.686	—	—
Gly	28	−5.520	3.334	4.367	2.378
Leu	29	−7.566	4.716	5.322	1.395
Cys	30	−3.718	1.597	2.015	0.962
Gln	31	−8.344	6.402	—	—
Thr	32	−4.808	2.476	1.763	−1.643
Phe	33	−8.265	5.765	—	—
Val	34	−4.719	1.622	1.817	0.450
Tyr	35	−7.742	4.879	6.020	2.627
Gly	36	−6.298	4.158	4.029	−0.296
Gly	37	−6.984	4.968	4.112	−1.972
Cys	38	−3.697	1.956	3.979	4.658
Arg	39	—	—	0.968	—
Ala	40	—	—	2.098	—
Lys	41	−5.532	3.035	3.510	1.092
Arg	42	−3.191	0.930	1.617	1.581
Asn	43	−3.978	2.226	1.687	−1.240
Asn	44	−7.742	6.090	4.659	−3.296
Phe	45	−8.265	5.886	—	—
Lys	46	—	—	1.681	—
Ser	47	−3.927	1.956	4.429	5.695
Ala	48	−3.202	1.044	1.494	1.037
Glu	49	—	—	3.535	—
Asp	50	—	—	—	—
Cys	51	−6.489	4.788	—	—
Met	52	−6.803	4.790	5.322	1.223
Arg	53	−6.537	4.267	4.339	0.167
Thr	54	−5.769	3.458	3.874	0.959
Cys	55	−6.974	5.261	4.235	−2.363
Gly	56	−5.210	3.478	2.621	−1.974
Gly	57	—	—	1.044	—
Ala	58	—	—	1.121	—

Table S2. Amide O/C state statistics from simulation

Residue	No.	N_O	N_{FO}	max(n_O)	max(n_C)
Asp	3	78,768	86,307	5	576
Phe	4	459	482	3	22,212
Cys	5	145	150	2	104,293
Leu	6	729	743	3	13,312
Glu	7	280	288	2	117,517
Tyr	10	3,754	3,821	3	2,921
Thr	11	78,782	86,593	6	294
Gly	12	7,013	7,065	2	2,003
Cys	14	0	0	—	1,048,349
Lys	15	69,107	74,871	4	410
Ala	16	2,425	2,493	3	14,459
Arg	17	21,075	21,933	6	10,498
Ile	18	28	28	1	121,241
Ile	19	10,217	10,326	3	1,119
Arg	20	0	0	—	1,048,349
Tyr	21	0	0	—	1,048,349
Phe	22	0	0	—	1,048,349
Tyr	23	0	0	—	1,048,349
Asn	24	0	0	—	1,048,349
Ala	25	99,171	111,430	6	194
Lys	26	420	421	2	20,898
Ala	27	0	0	—	1,048,349
Gly	28	42	45	2	291,667
Leu	29	5	5	1	776,924
Cys	30	9,841	10,032	4	1,366
Gln	31	0	0	—	1,048,349
Thr	32	17,462	17,799	3	540
Phe	33	0	0	—	1,048,349
Val	34	15,410	15,722	3	672
Tyr	35	1	1	1	937,750
Gly	36	90	98	3	463,949
Gly	37	76	81	2	439,529
Cys	38	104	110	2	439,154
Arg	39	91,524	101,809	5	4,756
Ala	40	8,151	8,302	4	1,433
Lys	41	301	324	2	37,624
Arg	42	23,414	24,733	5	1,343
Asn	43	19,973	21,120	5	2,498
Asn	44	23	23	1	143,438
Phe	45	0	0	—	1,048,349
Lys	46	20,456	21,394	5	1,119
Ser	47	39	39	1	107,407
Ala	48	31,451	32,537	3	378
Glu	49	306	306	1	16,036
Asp	50	0	0	—	1,048,349
Cys	51	0	0	—	1,048,349
Met	52	5	5	1	440,280
Arg	53	45	48	3	339,677
Thr	54	129	140	3	249,296
Cys	55	60	61	2	634,880
Gly	56	2,333	2,505	4	4,954
Gly	57	78,840	86,797	6	340
Ala	58	68,407	73,791	5	341

Fig. S1. Pair correlation function, $g(r)$, for water oxygens relative to the amide hydrogen, averaged over all amides and over all O-state frames (blue solid curve), over all C-state frames (red solid curve), or over all frames (gray dotted curve). (*Inset*) The corresponding running coordination number, $N(r)$. The primary coordination number, N_W, including water oxygens within the 2.6-Å cutoff (vertical gray line), is given for the two states.

Fig. S1

Fig. S2. Pair correlation function, $g(r)$ (*Left*) and running coordination number, $N(r)$ (*Right*) for water oxygens relative to the indicated amide hydrogen, averaged over all O-state frames (blue curve) or all C-state frames (red curve). The primary coordination number, N_W, including water oxygens within the 2.6-Å cutoff (vertical gray line), is given for the two states.

Fig. S2

Fig. S3. Correlation of O/C RMS deviation σ_{loc} from simulation (*Left*) and inverse RMS fluctuation $1/\sigma_B$ from crystal structure (*Right*), in both cases averaged over atoms within 7 Å from amide nitrogen, with O/C free energy difference $\beta\,\Delta G_{exp}$ from experimental HX rates.

Fig. S3

Fig. S4. Snapshots of O states for nine selected amides with the N–H group and the two (or three) primary waters in space-filling and other waters within a 7-Å sphere (and in tunnels or pores) in stick representation. To the left, the backbone conformation is shown for the selected O-state frame (dark gray) and for the first C-state frame in the trajectory (light gray). To the right, the molecular surface (1.4 Å probe radius, standard vdW radii) is shown. For Gly36 and Lys41, several snapshots (with and without water tunnels) are shown. Click on a structure to activate the interactive mode.

Fig. S4

Fig. S5. Normalized C-state residence time distribution, $\psi_C(\tau)$, for the indicated amides (gray bars) and single-exponential fit (red curve). The number, N_C, of visits to the C state during the trajectory is indicated for each amide as is the C-state MRT, τ_C, obtained from the fit. For 16 of the 20 amides shown, the fitted τ_C is within 10% of the value calculated (without binning-error correction) from the trajectory statistics as $\tau_C = \Delta\tau\, N_{FC}/N_C$.

Fig. S5

Fig. S6. Schematic representation of a proton transfer within an encounter complex with two water molecules directly coordinated to the N–H group. At the endpoints of the proton transfer process, at least one of these waters accepts a H-bond (blue dashed lines) from the N–H or N–D group. In the intermediate structure, both water molecules donate H-bonds (red dashed lines) to the imidate nitrogen. More than two water molecules may participate in the more or less concerted proton transfer (a third water molecule is shown to the right in the configuration at the top).

Fig. S6

Fig. S7. Fractional populations of C14–C38 disulfide rotamers (M1 blue, M2 magenta, other gray) in the O state for the amides of BPTI when the C state is constrained to the experimentally dominant M1 conformation.

Fig. S7

Fig. S8. Continuous (*Top*) and discretely sampled (*Bottom*) trajectory of alternating residences of an amide in the O and C states.

Fig. S8