Effects of Chain Stiffness on the Dynamics of Loop Formation in Polypeptides. Appendix: Testing a 1-Dimensional Diffusion Model for Peptide Dynamics

Lisa J. Lapidus,[†] Peter J. Steinbach,[‡] William A. Eaton,[†] Attila Szabo,[†] and James Hofrichter*,[†]

Laboratory of Chemical Physics, Building 5, National Institute of Diabetes and Digestive and Kidney Diseases, and Centers for Molecular Modeling and Information Technology, National Institutes of Health, Bethesda, Maryland 20892-0520

Received: March 28, 2002; In Final Form: June 24, 2002

Quenching of the triplet state of tryptophan by cysteine provides an important new tool for measuring the rate at which a specific intramolecular contact is formed in disordered polypeptides. By measuring the viscosity dependence of the quenching rate, both the reaction-limited and the diffusion-limited quenching rates can be determined. The diffusion-limited rate corresponds to the rate of forming a short-range contact. The reaction-limited rate, which depends solely on the equilibrium end-to-end distribution, becomes essentially length-independent for short chains, providing clear-cut evidence that the chain is stiff. The length dependence of the reaction-limited rate can be accurately calculated using the distance dependence of the quenching rate determined at room temperature in a rigid glass, together with the end-to-end distance distribution for a wormlike chain having a persistence length of 0.6−0.7 nm. In addition, the length dependence of the diffusion-limited rate can also be reproduced by treating the dynamics as diffusion on the 1D potential of mean force obtained from this distribution. The diffusion coefficient for the chain ends required to fit the diffusion-influenced rates is about 1.7×10^{-6} cm^2 s^{-1} at a viscosity of 1 cp and 293 K, almost 10 times smaller than the value expected for free diffusion of the contacting residues. Molecular dynamics simulations performed using a Ramachandran-like potential provide justification for this analysis. Diffusion-limited rates calculated from the trajectories are in excellent agreement with those calculated with the 1D diffusion model using the simulated end-to-end distance distributions.

Introduction

Understanding how proteins fold requires experimental and theoretical descriptions of the kinetics and dynamics of the elementary processes.[1] The simplest of these is intramolecular contact formation to form a loop. We have recently demonstrated that the long-lived triplet state of tryptophan formed by nanosecond optical excitation can be used as a probe for measuring the rate of loop formation.[2] To measure loop formation, a tryptophan is positioned at one end of a disordered polypeptide and cysteine at the other end. Because cysteine is the only naturally occurring amino acid that efficiently quenches the tryptophan triplet state,[3] measurement of the triplet decay monitors the formation of a contact between tryptophan and cysteine. However, quenching by cysteine at contact is not instantaneous, so the rate observed experimentally is less than the rate of contact formation.

In the previous study, the rate of contact formation was calculated from the experimental rate using a two-step kinetic model (Figure 1). In this model, the ends of the peptide first diffuse together with a rate k_{D+} to form an encounter complex or contact pair. The cysteine either quenches the tryptophan triplet state with a rate q or diffuses away with a rate k_{D-}. The observed quenching rate (k_{obs}) in this model is given by

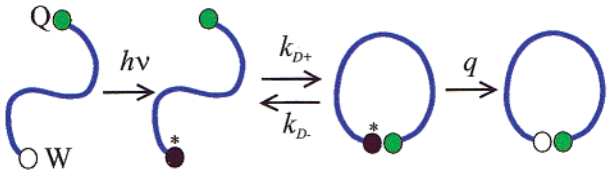

Figure 1. Scheme for measuring the rate of contact formation between tryptophan and cysteine at the ends of a random polypeptide. Pulsed excitation leads to population of the lowest excited triplet state of tryptophan. Tryptophan can then contact cysteine in a diffusion-limited process k_{D+}. During the lifetime of this contact pair, quenching of the triplet occurs with a probability determined by the quenching rate and the rate of diffusive dissociation of the contact pair.

$$k_{obs} = k_{D+}\frac{q}{k_{D-} + q} \equiv k_{D+}\phi \tag{1}$$

which can be rearranged to give

$$\frac{1}{k_{obs}} = \frac{1}{k_R} + \frac{1}{k_{D+}} \tag{2}$$

where $k_R = qK_{eq}$ is the reaction-limited rate, K_{eq} ($\equiv k_{D+}/k_{D-}$) is the equilibrium constant for forming the encounter complex, and ϕ is the probability of quenching during its lifetime. The diffusion-limited rate k_{D+} ($q \gg k_{D-}$) corresponds to the rate of bringing the ends together to form a short-range contact, whereas the reaction-limited rate, k_R, corresponds to the situation in which diffusion is sufficiently fast ($k_{D-} \gg q$) that the equilibrium end-to-end distance distribution is maintained at

* Corresponding authors. E-mail: jim@sunder.niddk.nih.gov. Tel: 301-496-6033. Fax: 301-496-0825.

[†] National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health.

[‡] Center for Molecular Modeling, Center for Information Technology, National Institutes of Health.

10.1021/jp020829v
Published on Web 10/16/2002

every instant. Comparison with the results for a diffusion-limited quencher (lipoate) showed that ϕ for quenching by cysteine is $\sim 1/2$. Assuming that ϕ is independent of chain length, the length dependence of the contact rate (k_{D+}) was found to decrease inversely as the $3/2$ power of the chain length for peptides of 10 to 20 residues, as predicted for a random-walk chain using the theory of Szabo, Schulten, and Schulten (SSS) described below. For shorter chains, the rate was found to be less dependent on length, and it was suggested that chain stiffness was responsible.

A potential problem with the triplet-quenching technique is that quenching is presumed to occur via electron transfer from the tryptophan triplet state to the sulfur of cysteine, and q is therefore expected to exhibit distance dependence. With a distant-dependent q, k_{D+} may differ significantly from the rate of forming a van der Waals contact. This prompted us to determine the distance dependence experimentally by measuring the distribution of quenching rates of tryptophan embedded at low concentration in a rigid glass of trehalose at room temperature containing high concentrations of cysteine. These experiments showed that the quenching rate $q(r)$ falls off as $q_0 \exp[-\beta(r - a_0)]$, where r is the separation and a_0 is the distance of closest approach. The trehalose experiments produced the values $q_0 = 4.2$ ns^{-1}, $\beta = 0.4$ nm^{-1}, and $a_0 = 0.4$ nm.[5] The large value of β indicates that quenching is indeed a very short-range process. Using these results, the reaction-limited bimolecular quenching rate, k_R^{bi} ($k_{D-} \gg q$), is

$$k_R^{bi} = \int_{\alpha_0}^{\infty} q(r)\, 4\pi r^2\, dr = 1.4 \times 10^8\ M^{-1}\ s^{-1} \qquad (3)$$

This rate is in excellent agreement with the rate of 1.3×10^8 M^{-1} s^{-1} measured in water at room temperature (the Smoluchowski diffusion-limited bimolecular rate is $4\pi D^{bi} a = 4 \times 10^9$ M^{-1} s^{-1}) and justifies our use of the rates and distance dependence from the glass experiments in a more detailed analysis of the end-to-end contact rates for the peptides.

The goal of the present study is to analyze additional experiments using a more rigorous description of the measured quenching rate, k_{obs}, in order to gain a quantitative understanding of its length dependence. The experiments consist of measuring the temperature and viscosity dependence of k_{obs} for peptides of varying length. Assuming that the equilibrium end-to-end distance distribution, $p_{eq}(r)$, and $q(r)$, are independent of the solution conditions used to alter the viscosity (η), the reaction-limited rate can be obtained from the intercept of a $1/k_{obs}$ versus η plot, and the diffusion-limited rate k_{D+}, from the slope, as indicated by eq 2. This approach has significant advantages over that used previously in which the quencher was varied at a single chain length.[2] In addition to yielding the reaction-limited rate k_R that provides information on the equilibrium end-to-end distribution, it permits the diffusion-limited rate, k_{D+}, to be determined without assuming that the quenching probability (ϕ of eq 1) is independent of chain length, as was done earlier.[2] In addition, it is not necessary to assume that the diffusion coefficient and the end-to-end distribution are unaffected by altering the quencher.

To explain the length dependence of the measured rates, we have used SSS theory, which treats the polymer dynamics as motion on a 1D potential of mean force corresponding to $p_{eq}(r)$. The question of whether diffusion on a 1D free-energy surface provides an accurate description of kinetics has become an important issue in statistical mechanical models of secondary structure formation and protein folding.[6] We have therefore explored the applicability of SSS theory by comparing its results with those obtained from molecular dynamics simulations of the peptides used in this work, carried out using a simplified potential. A comparison of SSS theory with simulations had been done previously for long Rouse chains, with mixed results.[7] In the Appendix, we present the results of simulations using a Ramachandran-like potential for a polypeptide in which nonbonded attractive interactions are eliminated but dihedral angle potentials and the excluded volume are retained. The rates directly calculated from the trajectories are in excellent agreement with those calculated from SSS theory using the $p_{eq}(r)$ and diffusion coefficients obtained from the trajectories, providing justification for its use in our analysis of the experimental data.

There are two major results. First, the end-to-end distribution $p_{eq}(r)$ for a wormlike chain, together with the experimentally determined $q(r)$, accurately predicts the length dependence of the reaction-limited rates. Second, using this $p_{eq}(r)$ and $q(r)$ in the SSS theory also reproduces the measured length dependence of the diffusion-limited rate k_{D+} with a diffusion coefficient that is almost 10-fold smaller than that for free diffusion of the terminal residues. The effect of backbone stiffness on polypeptide conformation has been considered in several previous studies. Thirumalai and co-workers estimated diffusion-limited rates for loop formation using an approximate description of the loop-formation probability.[8] More recently, this group has used the curvature to estimate the rigidity of loops in SH_3 domains and proteins with immunoglobulinlike folds from structures in the database,[9] and Zhou has reported that the distribution of end-to-end distances of loops in proteins can be described using a wormlike chain distribution.[10]

Theoretical Background

SSS Theory. In SSS theory, the observed quenching rate is calculated by solving a Smoluchowski diffusion equation for motion on a 1D potential of mean force defined as $U(r) = -\ln p_{eq}(r)/k_B T$, where $p_{eq}(r)$ is the equilibrium end-to-end distance distribution. In the simplest case, the quenching rate is a δ function (i.e., quenching occurs upon van der Waals contact of tryptophan and cysteine at a center-to-center distance a with a rate q and is zero for distances greater than a). For this case, SSS showed that the observed rate k_{obs} is given by

$$\frac{1}{k_{obs}} = \frac{1}{k_R} + \int_a^{l_c} (D(x)\, p_{eq}(x))^{-1} \left[\int_x^{l_c} p_{eq}(y)\, dy\right]^2 dx \qquad (4)$$

where the reaction-limited rate is $k_R = q\, p_{eq}(a)$, l_c is the contour length of the chain (for a polypeptide, $l_c = 0.38$ nm $\times$ number of peptide bonds), and D is the relative diffusion coefficient for the chain ends. The second term in this equation is the reciprocal of the diffusion-limited rate, k_{D+}, of eqs 1 and 2. The diffusion-limited rate is the rate of interest because it corresponds to the rate of contact formation. For a Gaussian (random-walk) chain, where $p_{eq}(r)$ is given by

$$p_{eq}(r) = 4\pi r^2 (2\pi \langle r^2 \rangle/3)^{-3/2} \exp(-3r^2/2\langle r^2 \rangle) \qquad (5)$$

the diffusion-limited rate for small a is given by

$$k_{D+} = \frac{4\pi D a}{(2\pi \langle r^2 \rangle/3)^{3/2}} = \frac{4\pi D a}{(2\pi C_n n l^2/3)^{3/2}} \qquad (6)$$

Thus, the contact rate for a random-walk chain scales inversely as the $3/2$ power of the chain length ($\langle r^2 \rangle = C_n n l^2$, where n is the number of peptide bonds, $l = 0.38$ nm is the projection of

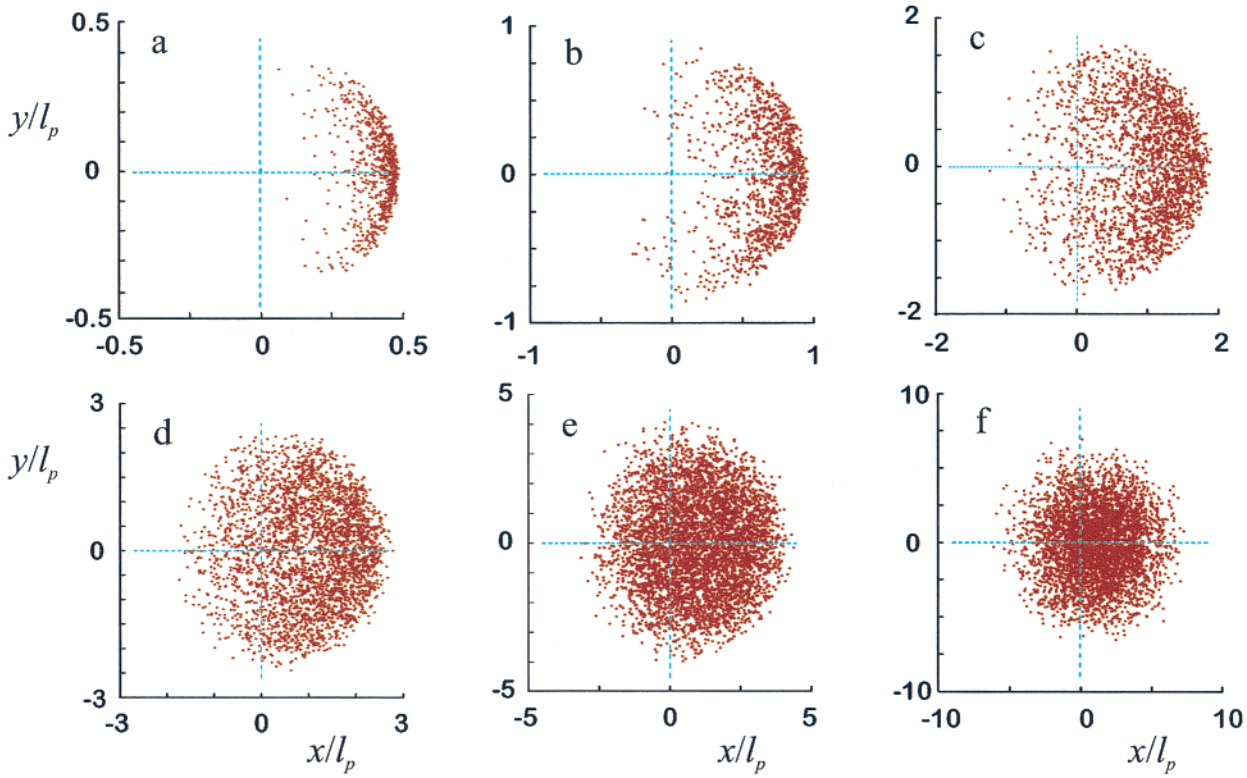

Figure 2. Distributions of positions for the ends of wormlike chains. Chains were started at the origin and were propagated in the $+x$ direction. The distribution of positions for the ends of 5000 chains selected from the simulations is shown for contour lengths of (a) $l_p/2$, (b) l_p, (c) $2l_p$, (d) $3l_p$, (e) $5l_p$, and (f) $10l_p$. The limits for the axes in each panel are determined by contour length.

the $C_\alpha - C_\alpha$ distance on the axis of the fully extended chain, and C_n is the Flory characteristic ratio).

With a distance-dependent quenching rate, Wilemski and Fixman showed that k_{obs} is approximately given by

$$\frac{1}{k_{obs}} = \frac{1}{k_R} + \frac{1}{k_R^2}\int_0^\infty \langle \delta q(t)\, \delta q(0)\rangle \, dt \tag{7}$$

where

$$k_R = \langle q \rangle = \int_a^{l_c} p_{eq}(x)\, q(x)\, dx \tag{8}$$

and $\delta q = q(r) - k_R$.[11] For 1D diffusion, the time integral of the sink−sink correlation function, $\langle \delta q(t)\delta q(0)\rangle$, can be evaluated analytically,[12] and eq 7 becomes

$$\frac{1}{k_{obs}} = \frac{1}{k_R} + \frac{1}{k_R^2}\int_a^{l_c} (D(x)\, p_{eq}(x))^{-1}\left[\int_x^{l_c} \delta q(y)\, p_{eq}(y)\, dy\right]^2 dx \tag{9}$$

This is the generalization of eq 4 and reduces to it when $q(x)$ is a δ function. We have used eqs 8 and 9 to fit our data. We have also shown that eq 9 is an excellent approximation for the systems studied here by numerically solving the diffusion equation with the measured quenching rates, $q(r)$, and $p_{eq}(r)$ obtained from the wormlike chain model (see Appendix).

Wormlike Chains. To account for the decrease in the length dependence of the rates measured for short chains, we utilize the wormlike chain model, originally proposed by Porod and Kratky,[13] as an approximate description of $p_{eq}(r)$. The wormlike chain, defined as an infinitely thin rod with bending rigidity κ, is the simplest model for a polymer chain that includes stiffness. This description accounts for the effects of local restraints on the curvature of the chain but ignores long-range excluded-volume interactions. The stiffness is conventionally characterized by the persistence length, l_p ($= \kappa/k_BT$). For contour lengths, l_c, that are much shorter than l_p, the wormlike chain becomes a rigid rod, whereas for $l_c \gg l_p$, the end-to-end distribution becomes Gaussian, as for a random-walk chain. For a wormlike chain, the bending energy to form a contact can be estimated by[14]

$$\frac{E_l}{k_BT} = \frac{k_l^2\, l_p l_c}{2} \approx \frac{2\pi^2 l_p l_c}{(l_c + a)^2} \tag{10}$$

where k_l is the curvature of the loop. The second expression is obtained if the loop is approximated as a circle of circumference $l_c + a$ and k_l is approximated by the reciprocal of the loop radius. Because the bending energy varies approximately as n^{-1}, it becomes significant as l_c approaches l_p. For example, using $l_p = 0.76$ nm and $a = 0.4$ nm, the bending energy increases from 4 to 6 k_BT nm as the peptide length decreases from 8 residues to 5 residues.

Fitting our experimental data with this model is significantly facilitated by having an analytic expression for the wormlike chain end-to-end distance distribution. To obtain this expression, numerical simulations of wormlike chains were carried out to produce distributions for short chains ($l_c \leq 10l_p$). The distribution of endpoints obtained from some of these simulations is shown in Figure 2. The results are shown for contour lengths varying from $l_p/2$ (Figure 2a) to $10l_p$ (Figure 2f). For visual orientation, the axes in each panel extend to the contour length. When $l_c = l_p/2$ (Figure 2a), the distribution of end positions extends over a small fraction of the $+x$ hemisphere, and the end-to-end distances are comparable to l_c. It is clear that the contour length must become significantly shorter than $l_p/2$ before the chain is a rigid rod. Note, however, that for $l_c \lesssim l_p$, the end-to-end distance distributions peak sharply at lengths close

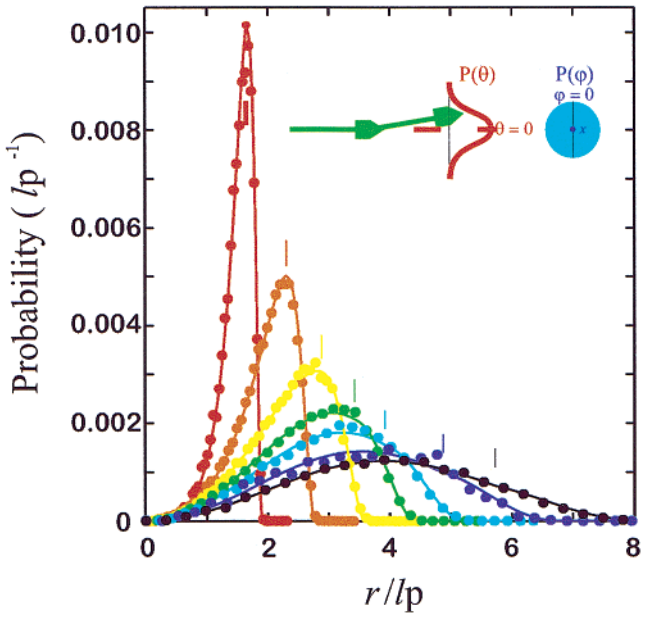

Figure 3. Comparison of simulated and interpolated wormlike chain distributions for short peptides. The cartoon describes the fact that the distribution of polar angles (in radians) was described by $P(\theta) = Z\theta \exp[-(\theta/\Delta)^2]$ and that the distribution of azimuthal angels was uniform on the interval $[0\ 2\pi]$. Z is a normalization constant, and $\Delta\theta^2$ was assigned a value of 0.102 corresponding to a persistence length of 20 segments. The results of simulations in which $2(10^4)$ chains were generated for contour lengths equal to 2.0 (red), 3.0 (orange), 4.0 (yellow), 5.0 (green), 6.0 (light blue), 8.0 (dark blue), and 10.0 (black) times the persistence length are shown as the data points. The interpolations between the derived functions for short chains and long chains, obtained as described in Methods, are shown as the lines. The switch points, r_s (eq 14), for each contour length are shown as the short vertical bars of the same colors.

to the contour length. As l_c approaches $10l_p$, the distribution becomes nearly Gaussian, and the end-to-end distances start to become significantly smaller than l_c (for a Gaussian chain, the end-to-end distance scales as $l_c^{1/2}$). The value of C_∞, which describes the number of chain segments that are equivalent to that of the equivalent random-walk chain, is $C_\infty = 2l_p/l_s$, where l_s is the segment length.[15] For the simulations, $l_p/l_s = 20$, so $C_\infty = 40$. Note that because the chains all initially propagate in the $+x$ direction the distribution is centered at $x = l_p$.

The distributions of end-to-end distances are shown in Figure 3, together with the fits described in Methods. As is evident from Figure 2, the peak of the distance distribution lies close to the contour length even for $l_c = 2l_p$ (red curve), but for $l_c = 10l_p$, the distance distribution becomes Gaussian.

Methods

Theoretical. To obtain an expression that describes the equilibrium length distribution for wormlike chains, a database was generated by carrying out simulations of 20 000 chains at values of l_c/l_p ranging from 1 to 10. The procedures described by Hagerman and Zimm[16] were used for the simulations. Briefly, for a chain of n segments, a set of $n - 1$ random values for the polar angles (θ, ϕ) are selected. The values of ϕ are uniformly distributed over the interval $[0, 2\pi]$, and the values of θ are described by the distribution $P(\theta) = Z\theta \exp[-(\theta/\Delta\theta)^2]$. Z is a normalization constant, and $\Delta\theta$ is a parameter that determines $\langle\cos\theta\rangle = \int \cos\theta\, p(\theta)\, d\theta$. The persistence length is related to $\Delta\theta^2$ by the relation $l_p = l_s/(1 - \langle\cos\theta\rangle)$, where l_s is the segment length (in the simulations, $l_s \equiv 1$). In the simulations described here, $\Delta\theta^2$ was assigned a value of 0.102, corresponding to a persistence length of 20 segments. To construct each chain, the first two points were positioned at [0 0 0] and [1 0 0]. Each additional segment, i, was generated by transforming the coordinates of the first i points using the rotation matrix $\mathbf{R}(\theta_i,\phi_i)$ and adding a unit vector in the $+x$ direction to the end of the chain. The product rotation matrix was accumulated for each chain to permit transformation back to the original coordinate frame.

To describe these distributions, an interpolation procedure was developed that utilizes analytical expressions for the limiting distributions at contour lengths much larger than the persistence length ($l_c > 10l_p$),[17]

$$p_l(r, l_c) = 4\pi r^2 ({}^3/_4\pi l_p l_c)^{3/2} \exp(-3r^2/4l_p l_c)(1 - 5l_p/4l_c + 2r^2/l_c^2 - 33r^4/80l_p l_c^3 - 79l_p^2/160l_c^2 - 329r^2 l_p/120l_c^3 + 6799r^4/1600l_c^4 - 3441r^6/2800l_p l_c^5 + 1089r^8/12\,800 l_p^2 l_c^6) \quad (11)$$

and for contour lengths that are shorter than the persistence length ($l_c \lesssim l_p$),[18]

$$p_s(r, l_c) = Nr^2(1/l_p(l_c - r))^{3/2} \sum_{k=1}^{\infty} H(2k-1)^2 l_c^2/4l_p(l_c - r) \quad (12)$$

where $H(x) = (4x - 2)\exp(-x)$ and N is a normalization factor. The interpolation procedure was optimized by comparing predicted distributions with those obtained from the simulations. Preliminary comparison showed that, for end-to-end distances less than the most probable distance, the simulated distribution was well described by p_l at $l_c \gtrsim 8l_p$ but well described by p_s only for $l_c \lesssim l_p$. To describe the simulated distributions for intermediate contour lengths $l_p < l_c < 8l_p$, a switch function of the form

$$SW(l_c) = \exp\left(-\left(\frac{l_c}{\delta l_p}\right)^2\right) \quad (13)$$

and a constant switch factor, SW_l, were introduced to interpolate between these two expressions at end-to-end distances below and above a switch point, r_s:

$$p_w(r, l_c) = SW(l_c)\, p_s(r, l_c) + [1 - SW(l_c)]p_l(r, l_c) \qquad r \le r_s \quad (14)$$

$$p_w(r, l_c) = Z\{SW_l\, p_s(r, l_c) + [1 - SW_l]\, p_l(r, l_c)\} \qquad r > r_s$$

r_s was chosen as the value of r at which the difference in the normalized distributions, $p_s(r, l_c) - p_l(r, l_c)$, is largest, and Z was chosen to match the values of the two expressions at $r = r_s$. The value of r_s is very close to the position of maximum probability for $l_c \lesssim 2l_p$ but becomes progressively longer than the peak position as l_c increases. This strategy improves the sum of the squares by more than a factor of 4 in comparison to using a value of r_s that is fixed relative to the most probable distance. The values of δ and SW_l obtained by optimizing the calculated distributions were $\delta = 3.70$ and $SW_l = 1.163$. Because SW_l is larger than unity, eq 14 produces negative values of $p_w(r, l_c)$ at large values of r that were set to 10^{-8} before the resulting functions were normalized. The fits and switch points are shown in Figure 3.

To fit the data, reaction-limited rates were calculated with eq 8 using $p_{eq}(r) = p_w(r, l_c)$ with l_p and a, the distance of closest

approach, as fitting parameters. In carrying out the fit, wormlike chain distributions, $p_w(r, l_c)$, were calculated for each peptide length with the current value of l_p using eqs 11−14. The reaction-limited rates were then calculated from eq 8.[5] Using the results of this fit, D/k_{D+} was calculated for each peptide from the second term of eq 9. The reciprocals of these values were least-squares fitted to the experimental values of ηk_D to obtain the value of D at $\eta = 1$ cp.

Experimental. The temperature and viscosity dependences of the quenching rates were measured for the peptides Cys-(Ala-Gly-Gln)$_j$-Trp where $j = (1:6)$. Experiments were carried out at temperatures between 0 and 40 °C using the instrumentation and procedures described previously,[2] except that the sample was thermostated with a Peltier-controlled cuvette holder (Quantum Northwest). The highest temperature at which experiments could be carried out was limited to 40 °C because rapid quenching of the tryptophan singlet state reduces the quantum yield for triplet formation at higher temperatures. Sucrose was used to change the solvent viscosity. Earlier work had shown that sucrose minimally perturbs the helix−coil equilibrium in an alanine peptide and the hairpin−coil equilibrium in the C-terminal fragment of protein GB1.[19] We reasoned from these results that sucrose should also minimally perturb the properties of unfolded polypeptide chains.

Results and Discussion

Dependence of Reaction- and Diffusion-Limited Rates on Chain Length, Viscosity, and Temperature. From our earlier studies, it was clear that the observed rates of quenching by cysteine were slower than the rate of forming a contact with tryptophan.[2] To more accurately determine this contact rate, we carried out experiments in which the solvent viscosity was varied by adding sucrose and changing the temperature. The measured triplet lifetimes for three peptide lengths are plotted as a function of solution viscosity in Figure 4. At each temperature, the measured values fall on a straight line with a unique slope and intercept. These results suggest that the data exhibit a similar dependence on viscosity at all temperatures but that there is also an intrinsic temperature dependence of the quenching rates. If $p_{eq}(r)$ and $q(r)$ are independent of temperature and sucrose concentration so that all of the viscosity dependence arises from the process of diffusion to form the contact pair, eq 2 indicates that the intercepts of these plots are the reaction-limited rate, k_R, and the slopes yield the diffusion-limited rates, k_{D+}, at each temperature. To fit the data, we assume that the rates k_R and k_{D+} can be described by

$$k_R = A \exp\left(\frac{E_0(T - T_0)}{RTT_0}\right) \qquad (15)$$

$$k_{D+} = \frac{BT}{\eta T_0} \exp(C(T - T_0))$$

where T is the temperature, η is the solution viscosity, and A, B, C, and E_0 are fitting parameters. A and B are the values of k_R and ηk_{D+} at the reference temperature $T_0 = 273$ K. The temperature dependence of k_{D+} is discussed in more detail below. Independent fits of the data for each peptide using eq 15 showed no systematic dependence of the parameters C and E_0 on peptide length, implying that, to within our experimental uncertainty, the persistence length is temperature-independent.

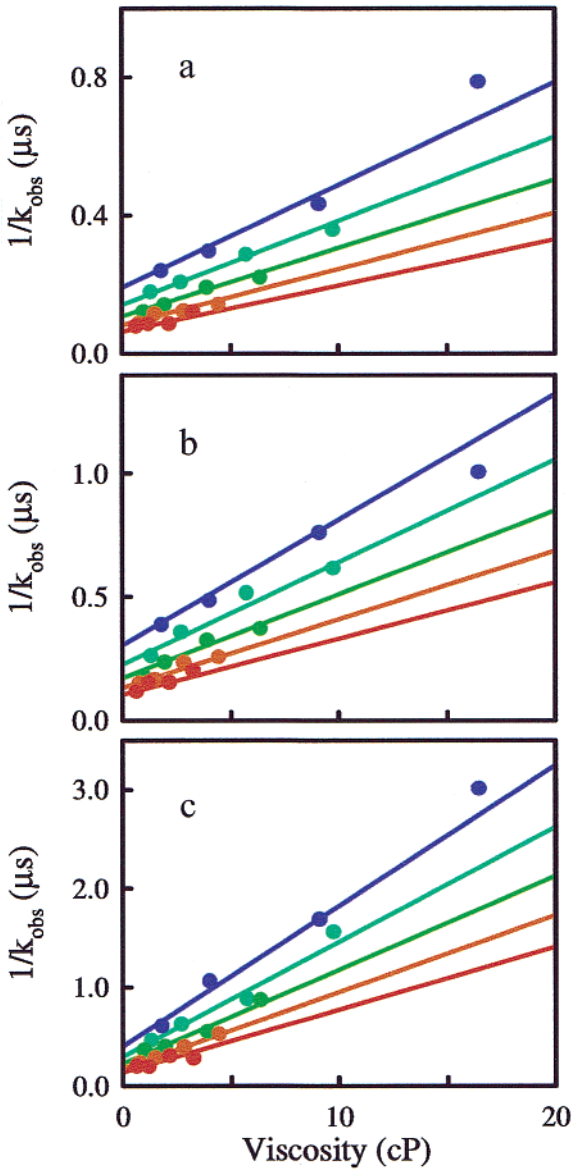

Figure 4. Temperature dependence and viscosity dependence of observed quenching rates, k_{obs}, for Cys-(Ala-Gly-Gln)$_j$-Trp. Data for the (a) $j = 2$, (b) $j = 4$, and (c) $j = 6$ peptides are plotted at 0 °C (blue), 10 °C (green), 20 °C (yellow), 30 °C (orange), and 40 °C (red) as a linear function of viscosity and fit to eq 15. The data were fit globally using eq 15 with independent values of A and B for each peptide length and a single value of the parameters E_0 and C for all peptide lengths. The fitted values are $E_0 = 4.33$ and $C = 0.0065$. The values of A and B are reported in Table 1.

Consequently, the data at all six lengths were fitted simultaneously with single values for C and E_0. The fits are plotted as solid lines, and the values of A and B for each peptide are given in Table 1.

The values of the diffusion-limited (k_{D+}) and reaction-limited (k_R) rates at 20 °C are shown in Figure 5. The observed rates at 20 °C in water ($\eta = 1$ cp) are included for comparison. The reaction-limited rates are very similar to the measured quenching rates in water at 20 °C and show significantly less dependence on peptide length than do the diffusion-limited rates for peptide lengths less than 10 residues.[20] If quenching occurs only at distance a, as in eq 4, the length dependence of k_R depends only on the probability of being at quenching distances $p_{eq}(a)$. Figure 5 shows that k_R (and hence $p_{eq}(a)$) becomes nearly independent of length for $n \lesssim 10$ peptide bonds. This flattening

TABLE 1: Parameters from Fits to the Temperature Dependence and Viscosity Dependence of Quenching Rates

peptide length	A (s^{-1})[a]	B (cp s^{-1})[b]
5	7.04 (10^6)	5.97 (10^7)
8	5.81 (10^6)	3.20 (10^7)
11	4.45 (10^6)	2.14 (10^7)
14	3.14 (10^6)	2.18 (10^7)
17	2.44 (10^6)	1.44 (10^7)
20	1.92 (10^6)	1.03 (10^7)

[a] A is the value of k_R at $T_0 = 273$ K. [b] B is the value of ηk_{D+} at $T_0 = 273$ K.

immediately rules out models such as random-walk chains that predict a $n^{-3/2}$ length dependence for all chain lengths and suggest that some energetic contribution is reducing the probability of short distances for short chains. We attribute the flattening of k_R (and $p_{eq}(a)$) to chain stiffness that reduces the probability of forming a loop for short chains.[8] In contrast, the diffusion-limited rate, k_{D+}, increases significantly at chain lengths below 10 peptide bonds. Whereas the value of k_{D+} also depends on $p_{eq}(a)$, which determines the free-energy barrier for diffusing to a, it also depends qualitatively on the distance from the point where the probability is maximal, r_{mp}, to a. Because $p_{eq}(a)$ is approximately constant, the increase in k_{D+} must result from a shift of the peak of the probability distribution to shorter distances as the chain length decreases, reducing the distance between a and r_{mp}.

We close this section with some comments on the use of eq 15 to describe the temperature dependence of the rates. The reaction-limited rate, k_R, depends on the equilibrium distribution of end-to-end distances, $p_{eq}(r)$, and the quenching rates, $q(r)$, both of which would be expected to be temperature-dependent. The equilibrium distribution, $p_{eq}(r)$, could include contributions from energetic terms that bias the distribution of end-to-end distances, and $q(r)$ could also be activated. The form of k_{D+} was motivated by similar considerations. The exponential temperature dependence of k_{D+} is required because the slopes in Figure 5 increase with decreasing temperature and the $1/T$ dependence in the Stokes−Einstein equation is too small to reproduce this dependence. Because this rate depends on temperature and viscosity as they affect the ratio D/V_{chain} where $V_{chain} = \int_a^\infty 4\pi r^2 p_{eq}(r)\, dr$, this result is not surprising. D is an *effective* diffusion coefficient, so it could also be modified by energetic interactions and could exhibit temperature dependence in addition to that resulting from T/η. In the presence of intramolecular interactions, V_{chain} would also be expected to be temperature-dependent.

Determination of the Persistence Length of the Wormlike Chain and the End-to-End Diffusion Coefficient. In the Appendix, we show that the SSS model provides an accurate description of end-to-end contact rates determined from atomic simulations using a simplified potential function. This result suggests that the SSS model can also be applied to even more realistic chains and therefore to the interpretation of our experimental values of the diffusion-limited rates. Application of the SSS model requires knowing three quantities: the end-to-end probability distribution $p_{eq}(r)$, the end-to-end diffusion coefficient D, and the distance dependence of the quenching rate $q(r)$. The quenching rate, $q(r)$, for quenching of tryptophan by cysteine in a trehalose glass has been determined and has been shown to be relevant to quenching in aqueous solutions because the correct bimolecular rate is predicted.[5] Our first problem was to determine $p_{eq}(r)$. Given $p_{eq}(r)$, D can be obtained from fitting the SSS model to the observed length dependence of the diffusion-limited rate, k_{D+}. Because the reaction-limited

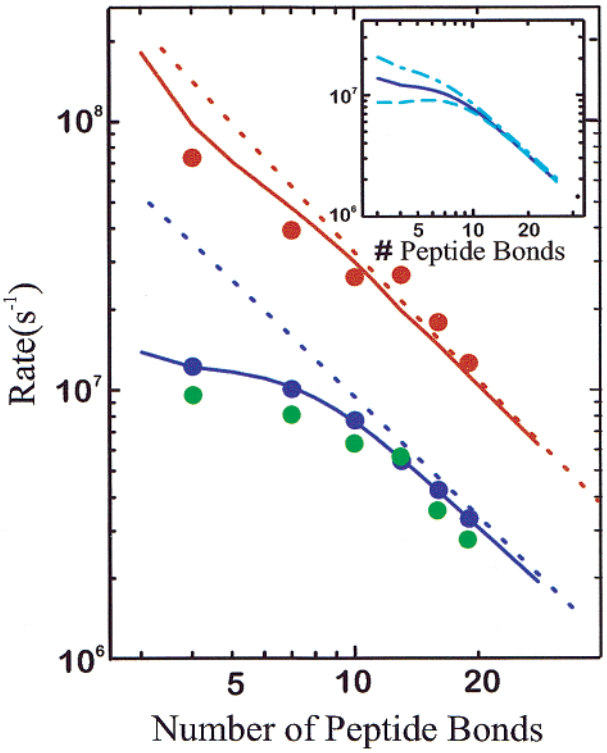

Figure 5. Length dependence of quenching rates for Cys-(Ala-Gly-Gln)$_j$-Trp peptides. The points are obtained from the analysis of the temperature dependence and viscosity dependence of the measured rates. The blue dots are the values of k_R at 293 K calculated from the parameters in Table 1 using eq 14. The solid blue line shows the fitted values of k_R. The values of the fit parameters are $l_p = 0.64$ nm and $a = 0.40$ nm. The dashed blue line shows the reaction-limited rates calculated for a Gaussian chain having $C_\infty = 4.0$ residues (selected to match the measured rates for $j = 4{:}6$). The red dots are the diffusion-limited loop-formation rates (ηk_{D+} from Table 1). The solid red line shows the diffusion-limited rates obtained from eq 7 using the same values of l_p and a and the fitted value of $\eta D = 1.7(10^{-6})$ cm^2 s^{-1} cp. The dashed red line shows the diffusion-limited rates calculated for a Gaussian chain with $C_\infty = 4.0$ residues and $\eta D = 1.9$ cm^2 s^{-1} cp. The experimental rates at 20 °C are shown as the green dots. Inset: Effect of persistence length on equilibrium rates. The solid blue line shows calculated equilibrium rates described above. The dotted−dashed blue line shows the equilibrium rates calculated using $l_p = 0.54$ nm, and the dashed blue line shows those calculated with $l_p = 0.74$ nm.

rates depend only on $p_{eq}(r)$ and $q(r)$, we needed a model for the chain that produces a $p_{eq}(r)$ that is consistent with the values of k_R in Figure 5. Preliminary calculations showed that the wormlike chain model met this requirement.

To fit the data, wormlike chain distributions $p_{eq}(r) = p_w(r, l_c)$ were calculated for each peptide using the analytical expressions described in Methods (eqs 11−14), and k_R was then calculated using eq 8. The calculated rates were optimized using l_p and the distance of closest approach, a, as fitting parameters. Using the fitted values, D/k_{D+} was then calculated for each peptide length from the double integral in the second term of eq 9. A linear least-squares fit of the reciprocals of these values to the values of ηk_{D+} in Figure 5 produced a value of $D/\eta = 1.7(10^{-6})$ cm^2 s^{-1} cp^{-1}. The fits, also shown in Figure 5, demonstrate that the wormlike chain model with $l_p \approx 0.64$ nm accurately reproduces the measured length-dependence of both the equilibrium and diffusion-limited rates. This value of the persistence length predicts a deviation from the $n^{-3/2}$ dependence for the reaction-limited rate of peptides having lengths less than 15 residues (14 peptide bonds) and corresponds to a value of $C_\infty = 3.4$.

The length dependence of the equilibrium rates is quite sensitive to the value of l_p. The inset to Figure 5 shows equilibrium rates calculated using values of a that match the rates for the longest chains and values of l_p that differ by ± 0.1 nm. For stiffer chains ($l_p \geq 0.8$ nm), k_R exhibits a maximum at a chain length of approximately eight residues. Because the reaction-limited rates are determined by $p_{eq}(r)$, where r is 0.4−0.5 nm, the equilibrium probability of forming a loop, $p_{eq}(r)/\int p_{eq}(r')dr'$, also shows a maximum at eight residues. The presence of a maximum is dependent on the value of r/l_p, and the peak is not observed when $r/l_p \gtrsim 0.5$. Significantly slower rates have been reported for threonine-rich peptides of comparable lengths, consistent with increased backbone stiffness, but diffusion-limited rates have not yet been determined.[21]

Other estimates of the persistence length for peptide chains have been obtained from the analysis of loops in proteins. Klimov and Thirumalai used the curvature to estimate the rigidity of loops in SH_3 domains and proteins with immunoglobulinlike folds from structures in the database.[9] The curvature varies from 1 to 8 nm^{-1}, with a mean value of about 3 nm^{-1}. If it is assumed that the strain energy stored in the loop is on the order of k_BT, then the reciprocal of the curvature provides a crude estimate of the persistence length (eq 10), which ranges from 0.2 to 1.5 nm. Zhou[10] has also recently noted that the distribution of end-to-end distances of loops in proteins can be described using a wormlike chain distribution with a persistence length l_p of 0.3 nm. This value is significantly smaller than that obtained from our fit, possibly because the loops sampled from the database include a high fraction of glycine and proline residues. The wormlike chain model has also been used to describe the force versus extension curves in single-molecule pulling experiments on titin and related proteins.[22] In these studies, the stretching and contraction of unfolded titin domains at relatively large extensions is described using a wormlike chain model with persistence lengths that range from 0.4 to 0.8 nm. Similar results have been obtained with both laser optical trapping and AFM methods.[22] These estimates are more in line with our result, despite the fact that the pulling experiments sense the distribution at large extensions whereas our experiment probes the distribution at almost zero extension.

In an earlier study, Bieri et al. reported an $n^{-1.4}$ dependence for peptides with three to nine peptide bonds in a study of triplet−triplet energy-transfer rates for $(gly\text{-}ser)_j$ peptides ($j = 2-4$) end-labeled with naphthalene and thioxanthone.[23] It is not clear whether the apparent absence of a turnover in this study results from the fact that the size of the probes permits "contact" over a wider range of end-to-end distances. Recently, Hudgins et al. have reported turnover in the length dependence of the fluorescence-quenching rates for a series of $(gly\text{-}ser)_j$ peptides labeled with 2,3-diazabicyclo[2,2,2]oct-2-ene (DBO) and tryptophan.[24] The ∼500-ns fluorescence lifetime of DBO, which is quenched by tryptophan with a rate of $2.0(10^9)$ M^{-1} s^{-1}, is well suited for studies of short, flexible polypeptides and permits the quenching rates to be measured accurately. The quenching rates observed (in D_2O at about 20 °C) in this study are quite similar to the values of k_{D+} in Figure 5, suggesting that they are close to diffusion limited values. The observed rates peak for $n \cong 4$ peptide bonds and decrease by almost a factor of 2 at $n = 2$. These results clearly warrant a more detailed analysis.

The value of D/η obtained by fitting the calculated rates, k_{D+}/D, to the measured rates, k_{D+}, is $1.7(10^{-6})$ cm^2 s^{-1} cp^{-1} for $l_p = 0.64$ nm. If D/η is calculated independently for each peptide length, it varies by ± 0.3 cm^2 s^{-1} cp^{-1}. Alternatively, the uncertainty in D/η can be estimated from the values obtained for persistence lengths that are 0.1 nm shorter and longer than the best-fit values, which are $1.4(10^{-6})$ and $2.4(10^{-6})$ cm^2 s^{-1} cp^{-1}. These values show that the fitted value is quite sensitive to the value of l_p, or, more generally, to $\langle r^2 \rangle$. Because the values of $\langle r^2 \rangle$ predicted by the wormlike chain model are comparable to those from the simulations described in the Appendix that include excluded volume but no attractive interactions (Figure A2), it is unlikely that the fitted values underestimate the true values of D/η for our peptides. These values are nearly an order of magnitude smaller than the sum of the diffusion coefficients for the terminal residues, which is estimated to be $1.5(10^{-5})$ cm^2 s^{-1} cp^{-1}.[25] This difference is much greater than that found in the Appendix, where we analyzed "experimental" results obtained from dynamics simulations using a Ramachandran-like potential (i.e., no attractive interactions) with the SSS model. In this case, the SSS model works very well for chain lengths shorter than 10 peptide bonds, and even for the longest peptide, it predicts rates that are only 60% larger than those calculated from the trajectory. The explanation of the slow diffusion coefficient obtained by analyzing the (real) experimental results (Figure 5) must thus arise from features omitted from our Ramachandran simulations. These include the presence of explicit solvent, attractive interactions, and hydrodynamic interactions. It is important to note that values of D which are 2−4 times smaller than our fitted value have been reported for longer polypeptide chains,[26,27] where chain stiffness is significantly less important.

To summarize, we have shown that a simple model that incorporates chain stiffness (the wormlike chain) provides a remarkably good description of the kinetic data. In addition to reproducing the reaction-limited rates, the wormlike chain distribution, $p_w(r)$, provides a good description of the diffusion-influenced rates using the SSS approximation. Using this model, the diffusion coefficient, D/η, is nearly10-fold slower than the sum of the diffusion coefficients of the amino acids that are at the chain ends. The chain diffusion coefficient is an important parameter in describing more complex properties of polypeptide chains. For example, it has been recently used, in conjunction with single-molecule experiments, to place an upper bound on the free-energy barrier for the folding of a small protein.[28] Studies of peptides that include bulky side chains or specific side chain−side chain interactions using the approach described here should provide some insight into the factors that determine the apparent end-to-end diffusion coefficient and persistence length of short peptides.

Acknowledgment. We thank Xiang Zhou for stimulating our interest in wormlike chains and for helpful discussions.

Appendix: Testing a 1D Diffusion Model for Peptide Dynamics

This Appendix presents results that provide support for the procedures used in this paper. We first describe a test of eq 9 that was used to fit the experimental data. To test this equation, which is based on the approximate result of Wilemski and Fixman,[11] we numerically solve the 1D diffusion equation for a system with a distance-dependent quenching rate $q(r)$ for a number of values of the diffusion coefficient, D. The resulting survival probabilities are analyzed using a procedure similar to that used to treat the experimental data to produce values of k_R and k_{D+}/D. The resulting values are very close to those obtained using eqs 8 and 9. We next describe simulations of polypeptide chains performed using a Ramachandran-like potential. The simulations, which were designed to explore the constraints

imposed by backbone geometry and excluded volume, were carried out using a modified potential in which attractive nonbonded interactions and all solvent interactions were intentionally excluded. Only bonded interactions, steric repulsions, and excluded volume were included. The simulated distance distributions, $p_{eq}(r)$, are compared with the wormlike chain distributions used to fit the experimental data, and the calculated values of k_{obs}, k_R, and k_{D+} are compared with the experimentally determined rates. The primary utility of the simulation results is that they can be used to examine the accuracy of the 1D diffusion model of Szabo, Schulten, and Schulten[4] in modeling the dynamics of short peptides by comparing distance−distance correlation functions and diffusion-controlled rates. The results show that the SSS model provides an excellent description of the dynamics of the chain ends at long times with a single adjustable parameter, D. Small deviations observed at short times probably arise from localized sidechain motions.

Testing Equation 9. Equation 9, which was used to fit the experimental data, is an approximate analytical solution to the 1D diffusion model of Szabo, Schulten, and Schulten[4] in the case where the reaction is distance-dependent. This diffusion problem can also be solved numerically, and we have used this approach to check the accuracy of eq 9. Tryptophan triplet survival probabilities that mimic the experimental data were calculated by numerically solving the diffusion equation. The distributions $p_w(r, l_c)$ obtained from the fits described in the body of this paper and the quenching rates $q(r)$ obtained from Lapidus et al.[5] were used to calculate survival probabilities for values of the diffusion coefficient D ranging from 10^{-7} to 10^{-5} cm^2 s^{-1}.

Lapidus et al.[5] found that the quenching rate decays exponentially with distance:

$$q(r) = q_0 e^{-\beta(r-a_0)} \tag{A1}$$

where $q_0 = 4.2$ ns^{-1}, $\beta = 0.4$ nm^{-1}, and $a_0 = 0.4$ nm. Survival probabilities were calculated by approximating the diffusion equation using rate equations to describe the time-dependent probabilities on a fixed grid of end-to-end distances. The survival probability, $S(t)$, is given by

$$S(t) = \int_a^{l_c} f(r, t)\, dr \tag{A2}$$

where $f(r, t)$ is the probability that the end-to-end distance is r at time t. $f(r, t)$ satisfies the equation

$$\frac{\partial f}{\partial t} = \frac{\partial}{\partial r} D(r)\, p_{eq}(r) \frac{\partial}{\partial r} \frac{f}{p_{eq}(r)} - q(r) f \tag{A3}$$

where $p_{eq}(r)$ is the r^2-weighted end-to-end distribution (e.g,. $p_w(r)$ discussed in the body of the paper) and $D(r)$ is the diffusion coefficient. We now must solve equation A3 with the initial condition that $f(r, 0) = p_{eq}(r)$ and the boundary conditions that f is reflecting at $r = a$ and $r = l_c$.

To solve this equation numerically, we consider a set of n spherical shells extending from a to l_c, having a thickness $\Delta r = (l_c - a)/n$ and approximate equation A3 by

$$\frac{d\mathbf{f}}{dt} = \mathbf{R}\mathbf{f} \tag{A4}$$

where $\mathbf{f}(t)$ is a vector of probabilities and $\mathbf{R}$ is an $n \times n$ rate matrix having elements

$$\begin{aligned}
2R_{i,i+1} &= D_i \frac{\pi}{p_{i+1}} + D_{i+1} \qquad \text{(upper diagonal)} \\
2R_{i+1,i} &= D_i + D_{i+1} \frac{p_{i+1}}{\pi} \qquad \text{(lower diagonal)} \\
R_{1,1} &= -R_{2,1} - q_1 \qquad \text{(diagonal elements)} \\
R_{i,i} &= -R_{i+1,i} - R_{i-1,i} - q_i \qquad (2 \le i \le n-1) \\
R_{n,n} &= -R_{n-1,n} - q_n
\end{aligned} \tag{A5}$$

where $p_i = p_{eq}(r_i)$, $q_i = q(r_i)$, and $D_i = D(r_i)/(\Delta r)^2$ are evaluated at $r_i = a + (i - 1/2)\Delta r$ (i.e., the midpoint of bin i). The diagonal elements include quenching of the triplet, which annihilates the population of each bin with a rate $q(r_i)$ given by eq A1. The matrix $\mathbf{R}$ was diagonalized to produce a set of n eigenvalues k_α and eigenvectors $\mathbf{u}^\alpha$. The survival probability $S(t)$ is given by

$$S(t) = \sum_{i=1}^{n} \sum_{\alpha=1}^{n} e^{k_\alpha t} u_i^\alpha a_i \tag{A6}$$

where the amplitudes, a_α, associated with each eigenvector were calculated by solving the set of linear equations $\mathbf{f}_0 = \sum_\alpha a_\alpha \mathbf{u}^\alpha$, where $\mathbf{f}_0$ is the end-to-end distance distribution vector at $t = 0$. In the calculations below, we assume that diffusion is independent of position, $D(r) = D$, and $(\mathbf{f}_0)_i = p_{eq}(r_i)$.

The survival probabilities were analyzed using procedures similar to those used to treat the experimental data. First, the relaxation rate for each set of conditions (D, l_c) was determined by fitting the survival probability to an exponential relaxation. The fits intentionally ignore a rapid relaxation that results from the initial population of ends at short distances (while this rapid phase is, in principle, experimentally observable, it is too fast to be detected in our instrument). The reciprocals of the resulting rates were then plotted as a function of $1/D$ and fit to straight lines to obtain values of the reaction-limited and diffusion-limited rates. The dependence of the calculated rates on D is shown in Figure A1. It is interesting that the dependence of the rates on chain length, shown in the inset to Figure A1, is quite dependent on the value of D. When $D = 10^{-5}$ cm^2 s^{-1}, the rates are reaction-limited and exhibit the same length dependence as k_R. As diffusion slows, the length dependence becomes more monotonic.

Figure A1 shows that the rates obtained from the survival probabilities are linear functions of $1/D$, confirming the validity of the comparable analysis used to obtain k_R and k_{D+} from the experimental data. The values of D/k_{D+} obtained from this procedure are 2% smaller than those calculated from eq 9 for the longest peptide, but the difference increases to about 12% for the shortest peptide. The values of k_R were about 6% greater than those calculated using eq 8 for the shortest peptide and about 6% smaller for the longest peptide. These results show that using eqs 8 and 9 to fit the data introduces errors that are significantly smaller than the uncertainties in the measured values of k_R and k_{D+}.

Langevin Simulations. The stochastic Langevin equation in one dimension is

$$\mu\ddot{x} = -\partial U(x)/\partial x - \zeta\dot{x} + R(t) \tag{A7}$$

Collisions of solvent atoms with the peptide influence its motion in two ways. They cause friction, which is an opposing force proportional to the velocity, $\dot{x}$, characterized by the coefficient ζ(g s^{-1}), and also produce a "random" force, $R(t)$. These de-energizing and energizing terms arise from collisions with the

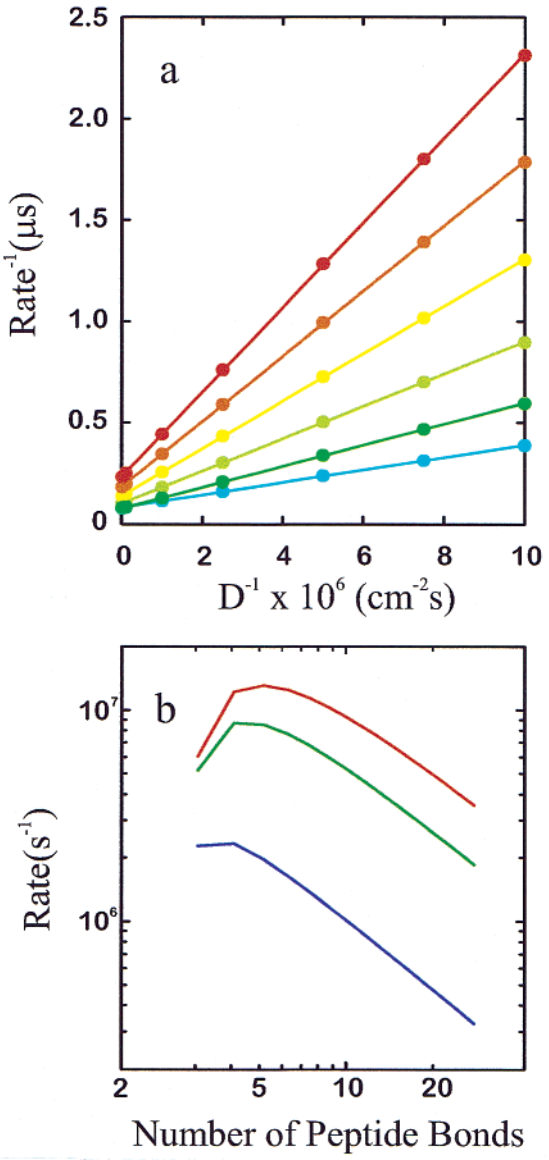

Figure A1. Determinations of equilibrium and diffusion-limited rates in the SSS approximation. (a) Rates obtained by fitting the survival probabilities calculated using equations A4−A6 are plotted as a function of the reciprocal of the diffusion coefficient, $1/D$. The intercept is equal to $1/k_{eq}$, and the slope is equal to D/k_{D+}. Results are shown for a persistence length of 0.85 nm and peptide of 4 (cyan), 7 (green), 10 (yellow-green), 13 (yellow), 16 (orange), and 19 (red) peptide bonds. (b) Dependence of quenching rate on peptide length for different values of D. Cross sections of the rates plotted in (a) are shown as a function of peptide length for $D = 10^{-5}$ cm^2 s^{-1} (red), $D = 10^{-6}$ cm^2 s^{-1} (green), and $D = 10^{-7}$ cm^2 s^{-1} (blue).

solvent, and the relation between them is described by the fluctuation−dissipation theorem. In implementing this equation, the random force, $R(t)$, whose fluctuations are related to the friction coefficient and the temperature, is modeled by applying random "kicks" to each atom of the peptide at each time step (1 fs). A collision frequency of $\gamma = \zeta/\mu = 50$ ps^{-1} was assumed to describe the dynamics in water at 293 K. In this representation, the diffusion coefficient of a heavy atom depends linearly on its mass, μ, as $D_L = k_B T N_0/\gamma\mu$.

The conventional van der Waals potential in the CHARMM19 force field was modified so that it became zero at distances greater than those dominated by the repulsive term, and all electrostatic interactions were set to zero. The modified van der

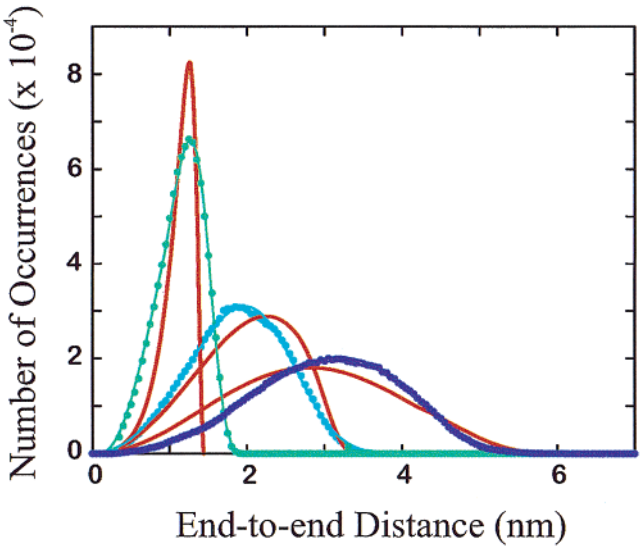

Figure A2. Distributions of end-to-end distances obtained from Langevin dynamics simulations for three peptides: Cys-(Ala-Gly-Gln)$_j$-Trp, where $j = \{1, 3, 6\}$. The distance distributions (blue points) were fitted to the function $P(r) = r^2 \exp(-[(r - r_m)/r_w]^4)$ (blue lines). The fitted parameters for $j = \{1, 3, 6\}$ are $r_w = \{0.845$ nm, 1.855 nm, 2.615 nm$\}$ and $r_m = \{0.557$ nm, 0.151 nm, 0.74 nm$\}$. The red lines are the predictions of the wormlike chain model with persistence length $l_p = 0.7$ nm.

Waals potential has the form

$$V_{vdw} = C(r^2 - m^2)^6 \qquad r \le m \tag{A8}$$

where r is the distance between atom centers and m is the minimum-energy (zero-force) separation for the standard Lennard-Jones potential, that is,

$$V_{LJ} = E\left(\left(\frac{m}{r}\right)^{12} - 2\left(\frac{m}{r}\right)^{6}\right) \tag{A9}$$

where E is the well depth at $r = m$. The constant C is proportional to $(E/m)^{12}$, with the proportionality constant chosen so that the modified potential agrees with V_{LJ} at $V_{vdw} = 85E$ (a good way up the wall). A total of 2 μs of dynamics were simulated for each of three peptides Cys-(Ala-Gly-Gln)$_j$-Trp with $j = (1, 3, 6)$. Coordinate files were saved every 2 ps (2000 steps), and the minimum distance between the sulfur of cysteine and any atom of the tryptophan ring was calculated from each coordinate file. These trajectories were used in the subsequent analysis.

Distance Distributions. The distributions of distances between tryptophan and cysteine obtained from the simulations are shown in Figure A2. The mean-square distances are 1.4 nm^2 for $j = 1$; 4.0 nm^2 for $j = 3$, and 10.1 nm^2 for $j = 6$. These values increase more rapidly than $n^{3/2}$, with the value of C_n in the expression $\langle r^2 \rangle = C_n n l^2$ increasing from 2.4 for $j = 1$ to 3.7 for $j = 6$. This is expected for nonideal chains that have no attractive interactions to compensate for excluded volume.[15] The distributions decay much more sharply at long distances than do those for Gaussian chains and can be reasonably well fit by the function $P(r) = r^2 \exp(-ar^4)$ and very well fit by the function $P(r) = r^2 \exp(-a\,(r - r_m)^4)$. When the r^2 contribution to these fitted distributions is removed, both functions produce flat distributions at short distances; the contribution of r_m simply increases the sharpness of the decay of the probability tail at large distances. The fitted parameters are given in the legend to Figure A2. Also shown in Figure A2 are the distributions

calculated for wormlike chains. The wormlike chain provides a reasonably good fit to the simulated distribution for $j = 3$, with the persistence length, $l_p = 0.6-0.7$ nm, as the only adjustable parameter. This persistence length also approximately describes the simulated distribution for $j = 1$, although the width of the predicted distribution is significantly sharper than that obtained from the simulations. A portion of this discrepancy could result from the fact that the observed distributions are broadened by side-chain degrees of freedom. These results suggest that the wormlike chain model correctly estimates the significant effects of chain stiffness on these distributions. However, the wormlike chain distribution predicted for $j = 6$ using $l_p = 0.6$ or 0.7 nm, which is nearly Gaussian, significantly underestimates the value of $\langle r^2 \rangle$, suggesting that additional effects that are not included in the wormlike chain model play a significant role in biasing the end-to-end distribution for longer peptides. The most obvious possibility is that the missing effects arise from the excluded volume.

Comparison of Simulated and Experimental Rates. It is possible to calculate three different rates from the trajectories to compare with those obtained from the experiments. The reaction-limited rates, k_R, predicted by the simulation were calculated by evaluating the average of $q(r)$ calculated on the distance trajectory. Diffusion-limited rates, k_{D+}, were calculated from the second term in eq 7 by evaluating the sink−sink correlation function

$$\langle \delta q(0)\, \delta q(t) \rangle = \frac{1}{t_{max} - t} \int_0^{t_{max}-t} \delta q(t_0)\, \delta q(t + t_0)\, dt_0 \quad \text{(A10)}$$

where $\delta q = q - k_R$ (eq 7) and t_{max} is the length of the trajectory. Simulated quenching rates, k_{obs}, were obtained by fitting the survival probabilities, $S(t)$, calculated as

$$S(t) = \langle \exp(-\int_0^t q(t')\, dt') \rangle = \frac{1}{t_{max} - t} \int_0^{t_{max}-t} dt_0 \exp(-\int_{t_0}^{t_0+t} q(t')\, dt') \quad \text{(A11)}$$

In practice, a large number of values of t_0 were sampled from the trajectory, and the maximum value of t was chosen to produce a decrease in $S(t)$ of at least a factor of 3. The survival probability was fitted to a single exponential to obtain the apparent rate for each trajectory. In calculating rates from the trajectories, $q(r)$ was set equal to $q(a)$ for $r \leq a$.

The calculated rates, shown in Figure A3, exhibit a stronger dependence on chain length than the measured rates. Because this behavior is seen for the reaction-limited rates, it must result from the length dependence of the probability distribution at distances where $q(r)$ is large ($r = 0.4-0.5$ nm). For the longer chains ($n > 10$), both k_R and k_{D+} scale as $\sim n^{-3}$, which is about twice as steep as predicted for Gaussian chains. This dependence can presumably be ascribed to the effect of excluded volume on the chain dimensions observed in Figure A2. Note that the dependence of k_R on chain length decreases at short chain lengths, but the rates do not become length-independent as they do for the experimental data and for the predictions of the wormlike chain model. The model used in the dynamics simulations may, therefore, also underestimate the chain stiffness of short peptides. The high flexibility must result from the absence of both van der Waals attractions and electrostatic interactions.

Testing the SSS Model. A principal objective for calculating the trajectories was to ask how well the dynamics of the end-to-end distance can be described by diffusion in a 1D potential

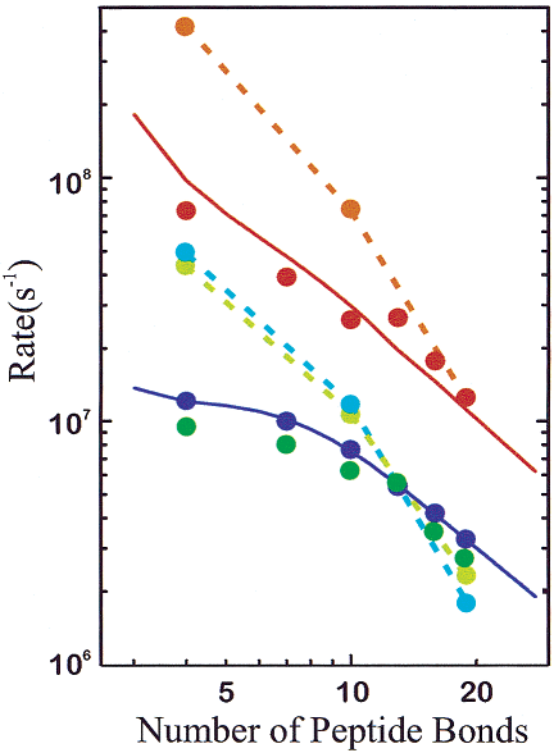

Figure A3. Rates calculated from Langevin dynamics simulations. The orange points (orange line) are the reciprocals of the first passage times to 0.4 nm, calculated using eq A12. The cyan points and line are the values of k_R calculated from the trajectory using eq 8 and the distance distributions in Figure A2, and the light green points and line show k_{obs} calculated with eq A11. In calculating the rates, eq A1 was used to calculate $q(r)$ for $r \geq 0.40$ nm, and the value of $q(r)$ at shorter distances was set equal to that at $r = 0.40$ nm. The dark green points are the experimental rates at 293 K. The dark blue points are the experimental values of k_R at 293 K, calculated from the results in Table 1, and the red points are the diffusion-limited rates at 1 cp. The solid lines are the fits to k_R (blue) and k_{D+} (red) obtained with the wormlike chain model.

of mean force (the SSS approximation). To explore this question, three calculations were carried out. (1) Values of k_{D+} calculated as the reciprocal of the mean first passage times were compared to those predicted in the SSS approximation using the sum of the diffusion coefficients for free tryptophan and cysteine, $D_L = 6.6\ (10^{-6})$ cm^2 s^{-1}. The SSS model describes the chain dynamics quite accurately for peptide lengths of less than ~10 peptide bonds but overestimates the rates for the longest peptide by a factor of about 1.6. (2) Distance−distance correlation functions were calculated from the trajectory and compared with the correlation functions predicted by the SSS approximation using D_L. We find that a value of D can be selected which reproduces both k_{D+} and the average correlation time for each peptide length. These calculations show that the SSS model works very well for these systems if D is regarded as an adjustable parameter. (3) The dynamics of localized populations selected from the trajectory were modeled using the SSS approximation. These calculations provide a more detailed picture of diffusion at times that are not accessible to the quenching experiments.

The dynamics were first examined by calculating k_{D+}, the diffusion-limited rates of contact formation from eq 4 with $D_L = 6.6 \times 10^{-6}$ cm^2 s^{-1} and $p_{eq}(r)$ calculated from the trajectory (Figure A2). These values were compared with values of k_{D+} calculated from the trajectories using Hummer's relation

$$k_{D+} = \frac{2\langle\tau\rangle}{\langle\tau^2\rangle} \tag{A12}$$

where the distribution of times τ, is defined as the time intervals between crossings of $r = a$ from below ($r < a$) to crossings from above ($r > a$). During this interval, the Cys−Trp distance, r, remains greater than a. In the SSS approximation, k_{D+} was calculated from eq 4 using the sum of the Langevin diffusion $D_L = 6.6 \times 10^{-6}$ cm^2 s^{-1}, and the $p_{eq}(r)$ calculated from the trajectory (Figure A2). The dependence of the quenching rates on the value of a is shown in Figure A4. The results show that the SSS approximation provides a remarkably accurate description of the simulation results for the $j = 1$ and 3 peptides with no adjustable parameters. The values of k_{D+} for these peptides are within a few percent of the values obtained from the trajectory. A significant discrepancy is found for the $j = 6$ peptide, where the rates predicted in the SSS approximation are about 60% larger than the simulation results. If optimal values of D, which we will refer to as D_S, are chosen to optimize the SSS predictions, we obtain $D_S = 6.3(10^{-6})$ cm^2 s^{-1} for the $j = 1$ peptide, $6.0(10^{-6})$ cm^2 s^{-1} for the $j = 3$ peptide, and $3.9(10^{-6})$ cm^2 s^{-1} for the $j = 6$ peptide, suggesting that D_S decreases with increasing peptide length.

The dynamics were next examined by calculating the correlation function

$$C_r(t) = \frac{\langle \delta r^2(0)\, \delta r^2(t)\rangle}{\langle (\delta r^2)^2\rangle} \tag{A13}$$

where $\delta r^2 = r^2 - \langle r^2\rangle$ and r is the end-to-end distance. The correlation functions for the three trajectories are compared in Figure A5. The functions are well fit by the sum of two exponential relaxations, with time constants that differ by about an order of magnitude. Interestingly, as the chain length increases, the functions are better approximated by a single-exponential relaxation. The amplitude of the fast phase decreases from about 15% for the $j = 1$ peptide to only about 3% for the $j = 6$ peptide.

Correlation functions for the SSS approximation were calculated from the rate matrix, **R**, using

$$\langle r^2(0)\, r^2(t)\rangle = \sum_{i,j=1}^{n} r_i^2 (e^{\mathbf{R}t})_{i,j} r_j^2 p_{eq}(r_j) \tag{A14}$$

where $p_{eq}(r)$ is the equilibrium distance distribution obtained from the trajectory. To calculate correlation times from the rate matrix, the diffusion coefficient, D, must be specified (eq A5). In Figure A5, we plot both the results obtained using D_L and those obtained when the values of D_S for each peptide were chosen so that the average relaxation times for the two correlation functions are identical. The SSS correlation functions are almost perfectly exponential, whereas those obtained from the trajectories are biexponential. This result shows that, as expected, the SSS approximation does not describe the dynamics on all time scales with a single diffusion coefficient. The values of D_S are $5.6(10^{-6})$ cm^2 s^{-1} for the $j = 1$ peptide, $4.7(10^{-6})$ cm^2 s^{-1} for the $j = 3$ peptide, and $4.0(10^{-6})$ cm^2 s^{-1} for the $j = 6$ peptide, very similar to those obtained from k_{D+} (Figure 4). This is an important result, because it shows that the SSS model provides a consistent description of different aspects of the dynamics when D_S is treated as an adjustable parameter.

The fitted values of D_S obtained from both k_{D+} and the correlation functions are smaller than D_L and decrease by about 30% as the peptide length increases. The dependence of D_S on chain length could result from an increase in the effective size of the chain ends as chain becomes longer, resulting in increased friction. The present results can be compared with those obtained by Pastor et al.[7] from Brownian dynamics simulations of long Rouse chains. For the $j = 1$ peptide, which might be compared to a Rouse chain composed of only three beads, the SSS approximation is exact, so the good agreement between the values of k_{D+} is consistent with their results (for long chains, the number of chain segments in the comparable Rouse chain is $N = n/C_n$, where n is the number of peptide bonds and $C_n \cong 3.4$ is estimated from the fitted persistence length). For much longer chains ($N = 50-100$) and values of a comparable to 0.4 nm ($a = 0.5$), k_{D+} was 2−3 times faster than the reciprocal of the first passage times calculated from their trajectories, with the ratio increasing with increasing chain length. The values of D_S required to reproduce the first passage times would be comparably smaller than the diffusion coefficient for the chain ends. These results suggest that the effective value of D_S decreases monotonically with increasing chain length. For the long-Rouse-chain simulations, the distance−distance correlation functions were highly nonexponential. The nearly exponential correlation functions for the peptide trajectories imply that the dynamics of these chains are simpler (i.e., take place over a narrower distribution of times) than those of the Rouse chains. This is the likely reason that the SSS approximation works so well for the peptides.

To explore the simulated dynamics in more detail, diffusion was also examined by monitoring the time evolution of populations initially located at a specified end-to-end distance

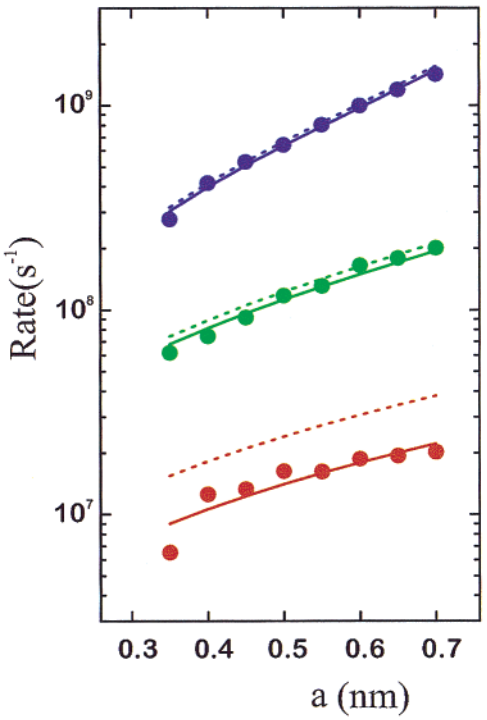

Figure A4. Values of k_{D+} calculated from Langevin dynamics simulations and from the SSS approximation. The diffusion-limited rates, k_{D+}, obtained for absorption at a fixed Trp−Cys distance, a, are compared. Rates calculated from the simulations of Cys-(Ala-Gly-Gln)$_j$-Trp using eq A12 are shown as the points for $j = 1$ (blue), $j = 3$ (green), and $j = 6$ (red). The rates calculated from eq 4 using the SSS approximation and the diffusion coefficient, $D_L = 6.6 \times 10^{-6}$ cm^2 s^{-1} are shown as dashed lines using the same color code. The rates calculated using optimal values of D are shown as solid lines. The fitted diffusion coefficients are 6.3×10^{-6} cm^2 s^{-1} for the $j = 1$ peptide, 6.0×10^{-6} cm^2 s^{-1} for the $j = 3$ peptide, and 3.9×10^{-6} cm^2 s^{-1} for the $j = 6$ peptide.

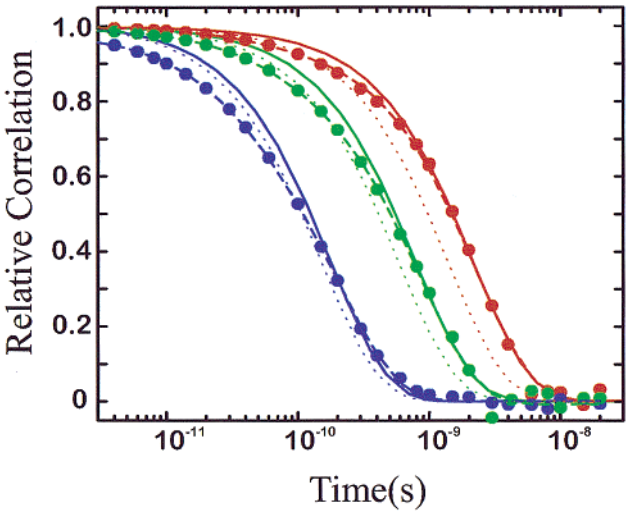

Figure A5. Distance correlation functions calculated from the simulations and the SSS approximation. The correlation functions calculated from the trajectories using eq A13 are shown as the points for $j = 1$ (blue), $j = 3$ (green), and $j = 6$ (red) peptides. The correlation functions were fitted with the sum of two exponential relaxations (dashed line). For the 5-mer, 15% of the amplitude decays with rate $k_f = 34$ ns^{-1}, and the remainder, with $k_s = 4.7$ ns^{-1}. For the 11-mer, 8% of the amplitude decays with rate $k_f = 16.1$ ns^{-1}, and the remainder, with $k_s = 1.2$ ns^{-1}. For the 20-mer, 3.7% of the amplitude decays with rate $k_f = 20$ ns^{-1}, and the remainder, with $k_s = 0.44$ ns^{-1}. The r^2 correlation functions calculated from eq A14 using $D_L = 6.6 \times 10^{-6}$ cm^2 s^{-1}, scaled for comparison with those calculated from the simulations, are plotted as dotted lines. These curves are indistinguishable from single-exponential relaxations. The r^2 correlation functions calculated from eq A14 using diffusion coefficients scaled to reproduce the average correlation time, $\int C_r(t)\,dt$, obtained from the simulation results are plotted as solid lines. For $j = 1$, the decay rate for the SSS correlation function is 5.58 ns^{-1} obtained with $D = 5.6 \times 10^{-6}$ cm^2 s^{-1}. For $j = 3$, the decay rate of the SSS correlation function is 1.256 ns^{-1} obtained with $D = 4.7 \times 10^{-6}$ cm^2 s^{-1}. For $j = 6$, the decay rate of the SSS correlation function is 0.47 ns^{-1} obtained with $D = 4.0(10^{-6})$ cm^2 s^{-1}.

r_s. The time-dependent distance distributions were modeled by solving the SSS diffusion equation in the potential defined by the equilibrium distribution, $p_{eq}(r)$, obtained from the trajectory (Figure A1). Preliminary analysis showed that the distributions were poorly fit using a constant diffusion coefficient. To describe the distributions accurately, we utilized a time-dependent diffusion coefficient, $D(t)$. In this case, the diffusion equation can be solved by substituting

$$\tau(t) = \int_0^t \frac{D(t')}{D_0}\,dt' \qquad \text{(A15)}$$

for the time. The results provide additional insight into the limitations of the SSS approximation.

To carry out this analysis, the set of points with a given end-to-end distance r_s was selected from the trajectory. If the times for these points are designated $\{t_0\}$, then the distributions of positions at times $\{t_0 \pm t\}$ were calculated from the trajectory for a set of times, t, spaced logarithmically from 2 ps to 10 ns. These distributions provided data sets that were fit using the SSS model, which was solved as described above (but without the sink term) to produce populations at times t. The initial distribution, $\mathbf{p}_0$, was selected to match the window of distances selected from the trajectory. To carry out the fits, t was transformed using $\tau(t) = t\{1 + \alpha[\log(t_{max}/t)]^\beta\}$, and the values of α, β, and a constant diffusion coefficient D_0 were fit to minimize the sum of squared residuals between the distribution calculated from the trajectory (at times t) and the calculated

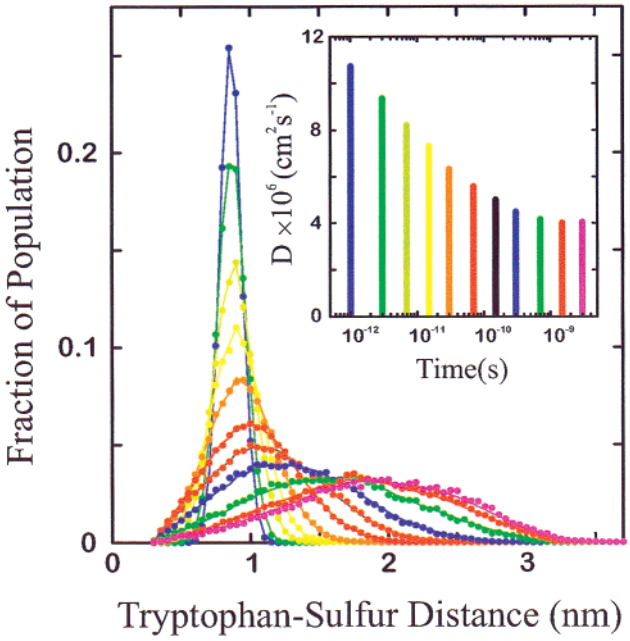

Figure A6. Time evolution of the distance distribution of a population that is initially separated by 0.8 nm for Cys-(Ala-Gly-Gln)$_3$-Trp. The points are the calculated distributions at 2 ps, 4 ps, 10 ps, 20 ps, 40 ps, 100 ps, 200 ps, 40 ps, 1 ns, 2 ns, and 10 ns, and the solid lines are the fits at times $\tau(t) = t\{1 + \alpha[\log(t_{max}/t)]^\beta\}$ calculated using the SSS approximation with parameters $D_0 = 4.16 \times 10^{-6}$ cm^2 s^{-1}, $\alpha = 0.0424$, and $\beta = 1.85$. Inset: Values of the time-dependent diffusion coefficient, $D(t)$, calculated as $D_0\,\Delta\tau/\Delta t$ at the midpoint of the time interval between each sequential pair of distributions. The colors correspond to the final distribution for each pair of points.

TABLE A1: Diffusion Coefficients Obtained from Langevin Simulations Using the SSS Approximation

peptide length	r_s (nm)	D_{150ps} (cm^2 s^{-1})	D_{15ps} (cm^2 s^{-1})	D_{1ps} (cm^2 s^{-1})
5	0.6	4.0 (10^{-6})	7.3 (10^{-6})	11.3(10^{-6})
	1.2	5.2 (10^{-6})	5.8 (10^{-6})	8.4 (10^{-6})
	1.6	3.5 (10^{-6})	3.7 (10^{-6})	5.2 (10^{-6})
11	0.8	5.1 (10^{-6})	7.4 (10^{-6})	10.7 (10^{-6})
	1.8	4.5 (10^{-6})	7.1 (10^{-6})	10.4 (10^{-6})
	2.8	4.2 (10^{-6})	5.7 (10^{-6})	8.5 (10^{-6})
19	0.7	4.6 (10^{-6})	6.8 (10^{-6})	10.4 (10^{-6})
	3.0	4.4 (10^{-6})	6.8 (10^{-6})	10.3 (10^{-6})
	4.5	3.8 (10^{-6})	5.8 (10^{-6})	9.1 (10^{-6})

distributions at times $\tau(t)$. The results of one such fit are shown in Figure A6. The fits for all peptide lengths and initial distributions are extremely good, with the population evolution modeled with accuracy that is better than a few percent. The fitted diffusion coefficients $D(t)$, calculated as $D_0\,d\tau/dt$ (cf. eq A15) for initial populations at three values of the initial distance r_s, are summarized in Table A1. The diffusion coefficients at 200 ps are similar for all three peptides, with a mean value of 4.3(10^{-6}) cm^2 s^{-1}, which is similar to the mean of the values obtained from the correlation functions in Figure A5. As shown in the inset to Figure A6, the diffusion coefficients at shorter times are larger by a factor of about 2. The fitted values of $D(t)$ depend not only on time but also on the initial end-to-end distance r_s (Table A1). For the $j = 1$ peptide, the increase is highly dependent on the initial position, with the largest dependence observed for short r_s, but the sensitivity to the initial position becomes less pronounced as the peptide length increases. The results show that, as anticipated from the correlation functions (Figure A4), it is primarily the dynamics at times less than ~100 ps that are poorly modeled using the SSS approximation with a fixed diffusion coefficient. One possible

explanation is that the faster dynamics result from local motions of the chain termini, such as side-chain rotations, that do not require significant backbone motion.

In summary, we find that the Langevin dynamics simulation qualitatively reproduces the features of the experimental rates: the values of k_R are very similar to the values of k_{obs}, whereas k_{D+} is significantly larger. Because our peptide model includes no attractive interactions, it is not surprising that the simulated kinetics exhibit a length dependence that is significantly steeper than that observed experimentally. The fact that the calculated rates are comparable to the experimental rates holds out hope that more accurate simulations will produce significantly better agreement with experiment.[29]

The dynamics of the ends of the simulated chains are remarkably well described by the diffusion in the potential of mean force that describes the equilibrium distribution of end-to-end distances (the SSS model). The diffusion-limited rates and correlation functions can be described by this model using an apparent diffusion coefficient that is slightly smaller than the Langevin diffusion coefficient, D_L, and decreases with peptide length. The dynamics are somewhat more complex than those predicted by the SSS model, as evidenced by a fast relaxation in the correlation functions and apparent diffusion coefficients that are approximately 2-fold larger at short times. The differences between the simulated dynamics of our Ramachandran chain and the SSS model are much too small to account for the decrease in D found in our analysis of the experimental data. This must mean that backbone connectivity and local excluded-volume constraints on backbone degrees of freedom are not the primary origins of the slow end-to-end diffusion coefficients obtained from the analysis of our experimental data in Figure 5. Other factors such as peptide−peptide and peptide−solvent interactions and possibly hydrodynamic interactions, which are not included in the Langevin dynamics model, must contribute to slowing diffusion.

References and Notes

(1) Eaton, W. A.; Muñoz, V.; Hagen, S. J.; Jas, G. S.; Lapidus, L. J.; Henry E. R.; Hofrichter, J. *Annu. Rev. Biophys. Biomol. Struct.* **2000**, *29*, 327−359.

(2) Lapidus, L. J.; Eaton, W. A.; Hofrichter, J. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 7220−7225.

(3) Gonnelli, M.; Strambini, G. B. *Biochemistry* **1995**, *34*, 13847−13857.

(4) Szabo, A.; Schulten, K.; Schulten, Z. *J. Chem. Phys.* **1980**, *72*, 4350−4357.

(5) Lapidus, L. J.; Eaton. W. A.; Hofrichter, J. *Phys. Rev. Lett.* **2001**, *87*, 258101-1−258101-4. To calculate rates for models in which there is no excluded volume, it is necessary to take into account that the quenching rates obtained by Lapidus et al. become very large at distances less than the 0.4-nm diameter assumed in fitting the trehalose data.[5] To overcome this problem, tryptophan−cysteine distances were constrained to be greater than a minimum distance, a, that was varied to produce optimal agreement between the calculated and experimental equilibrium rates. These distances are dependent on the persistence length, and for the best-fit value of $l_p = 0.64$ nm, a was 0.400 nm. The optimum values vary by ± 0.006 nm when l_p changes by ∓ 0.1 nm.

(6) Socci, N. D.; Onuchic, J. N.; Wolynes, P. G. *J. Chem. Phys.* **1996**, *104*, 5860−5868. Muñoz, V.; Henry, E. R.; Hofrichter, J.; Eaton, W. A. *Proc. Natl. Acad. Sci. U.S.A.*, **1998**, *95*, 5872−5879. Muñoz, V.; Eaton, W. A. *Proc. Natl. Acad. Sci. U.S.A.*, **1999**, *96*, 11311−11316. Portman, J. J.; Takada, S.; Wolynes, P. G. *J. Chem. Phys.* **2001**, *114*, 5082−5096.

(7) Pastor, R. W.; Zwanzig, R.; Szabo, A. *J. Chem. Phys.* **1996**, *105*, 3878−3882.

(8) Thirumalai, D. *J. Phys. Chem. B* **1999**, *103*, 608−610. Camacho, C. J.; Thirumalai, D. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 1277−1281.

(9) Klimov, D. K.; Thirumalai, D. *J. Mol. Biol.* **2002**, *315*, 721−737.

(10) Zhou, H.-X. *J. Phys. Chem. B* **2001**, *105*, 6763−6766

(11) Wilemski, G.; Fixman, M. *J. Chem. Phys.* **1974**, *60*, 866−877.

(12) Bicout, D. J.; Szabo, A. *J. Chem. Phys.* **1997**, *106*, 10292−10298.

(13) Kratky, O.; Porod, G. *Recl. Trav. Chim.* **1949**, *68*, 1106−1122.

(14) Doi, M.; Edwards, S. F. *The Theory of Polymer Dynamics*; Cambridge University Press: Cambridge, 1986; Section 8.8. Marko, J. F.; Siggia, E. D. *Macromolecules* **1995**, *28*, 8759−8770.

(15) Flory, P. J. *Statistical Mechanics of Chain Molecules;* Interscience Publishers: New York, 1969.

(16) Hagerman, P. J.; Zimm, B. H. *Biopolymers* **1981**, *20*, 1481−1502.

(17) Gobush, W.; Yamakawa, H.; Stockmayer, W. H.; Magee, W. S. *J. Chem. Phys.* **1972**, *57*, 2839−2843. Yamakawa, H.; Stockmayer, W. H. *J. Chem. Phys.* **1972**, *57*, 2843−2854.

(18) Wilhelm, J.; Frey, E. *Phys. Rev. Lett.* **1996**, *77*, 2581−2584.

(19) Jas, G. S.; Eaton, W. A.; Hofrichter, J. *J. Phys. Chem.* **2001**, *105*, 261−272.

(20) The values of k_{D+} are about 80% faster for the 5-residue peptide and 40% faster for the 11-residue peptide than those determined by Lapidus et al.[2] in water at room temperature, whereas the value for the $j = 6$ peptide is nearly identical. Most of the difference results from the fact the measured rates under these conditions are closer to the reaction limit for the shortest peptide than to the reaction limit for the 11-residue peptide, so the diffusion-limited rate decreases the relaxation time in eq 9 by less than a factor of 2. This result suggests that much of the flattening in the diffusion-limited rates reported by Lapidus et al.[2] can be attributed to the fact that the rates measured for short peptides are very close to the reaction-limited rates. Another contribution may result from the fact that, in the plots that compared the unimolecular and bimolecular rates (Figure 4 in Lapidus et al.[2]), the intercept is determined by peptides in which the cyclic disulfide, lipoate, is the quencher. The lipoate moiety has a larger molecular weight than cysteine, so it might be expected to diffuse more slowly.

(21) Lapidus, L. J.; Eaton. W. A.; Hofrichter, J. *J. Mol. Biol.* **2002**, *319*, 19−25.

(22) Kellermayer, M. S. Z.; Smith, S. B.; Ganzier, H. L.; Bustamante, C. *Science (Washington, D.C.)* **1997**, *276*, 1112−1116. Rief, M.; Gautel, M.; Schemmel, A.; Gaub, H. E. *Biophys. J.* **1998**, *75*, 3008−3014. Wang, K.; Forbes, J. G.; Jin, A. J. *Prog. Biophys. Mol. Biol.* **2001**, *77*, 1−44.

(23) Bieri, O.; Wirz, J.; Hellrung, B.; Schutkowski, M.; Drewello, M.; Kiefhaber, T. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 9597−9601.

(24) Hudgins, R. R.; Huang, F.; Gramlich, G.; Nau, W. M. *J. Am. Chem. Soc.* **2001**, *124*, 556−564.

(25) Creighton, T. E. *Proteins: Structures and Molecular Properties* Freeman: New York, 1993; pp 266, 269.

(26) Buckler, D. R.; Haas, E.; Scheraga, H. A. *Biochemistry* **1995**, *34*, 15965−15978.

(27) Hagen, S. J.; Hofrichter, J.; Szabo, A.; Eaton, W. A. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 11615−11617. Hagen, S. J.; Hofrichter, J.; Eaton, W. A. *J. Phys. Chem.* **1997**, *101*, 2352−2365.

(28) Schuler, B.; Lipman, E. A.; Eaton, W. A. *Nature (London)* **2002**, *419*, 743−747.

(29) Yeh, I.-C.; Hummer, G. *J. Am. Chem. Soc.* **2002**, *124*, 6563−6568.