J|A|C|S
ARTICLES
Published on Web 12/23/2006

# End-to-End vs Interior Loop Formation Kinetics in Unfolded Polypeptide Chains

Beat Fierz and Thomas Kiefhaber*

*Contribution from the Division of Biophysical Chemistry, Biozentrum der Universität Basel, Klingelbergstrasse 70, CH-4056 Basel, Switzerland*

Received September 14, 2006; E-mail: t.kiefhaber@unibas.ch

***Abstract:*** The conformational search for favorable intramolecular interactions during protein folding is limited by intrachain diffusion processes. Recent studies on the dynamics of loop formation in unfolded polypeptide chains have focused on loops involving residues near the chain ends. During protein folding, however, most contacts are formed between residues in the interior of the chain. We compared the kinetics of end-to-end loop formation (type I loops) to the formation of end-to-interior (type II loops) and interior-to-interior loops (type III loops) using triplet−triplet energy transfer from xanthone to naphthylalanine. The results show that formation of type II and type III loops is slower compared to type I loops of the same size and amino acid sequence. The rate constant for type II loop formation decreases with increasing overall chain dimensions up to a limiting value, at which loop formation is about 2.5-fold slower for type II loops compared to type I loops. Comparing type II loops of different loop size and amino acid sequence shows that the ratio of loop dimension over total chain dimension determines the rate constant for loop formation. Formation of type III loops is 1.7-fold slower than formation of type II loops, indicating that local chain motions are strongly coupled to motions of other chain segments which leads to faster dynamics toward the chain ends. Our results show that differences in the kinetics of formation of type I, type II, and type III loops are mainly caused by differences in internal flexibility at the different positions in the polypeptide chain. Interactions of the polypeptide chain with the solvent contribute to the kinetics of loop formation, which are strongly viscosity-dependent. However, the observed differences in the kinetics of formation of type I, type II, and type III loops are not due to the increased number of peptide−solvent interactions in type II and type III loops compared to type I loops as indicated by identical viscosity dependencies for the kinetics of formation of the different types of loops.

## Introduction

During protein folding many short-range and long-range intrachain interactions have to be established. Intrachain diffusion allows the chain to bring interacting groups together and to probe for favorable intramolecular interactions. Intrachain loop formation can thus be viewed as a fundamental step of protein folding and sets an upper limit for the rate at which the free energy surface can be explored. It was shown that inserting long loops into protein sequences reduces the stability of the native state[1,2] and slows down the folding kinetics,[3,4] which indicates that the size of loops is an important determinant of protein folding and stability. Three categories of intramolecular loops can be distinguished (Figure 1A). Type I loops are formed between the chain ends. The kinetics of formation of these loops has been studied in different polypeptide chains and with different experimental systems[5−10] (for a review see ref 11). However, during protein folding loop formation between an end and the interior of the chain (type II loops) or between two internal residues (type III loops) is much more common. In both cases, loop formation involves motions of additional chain segments (tails) beyond the region of contact formation. Polymer theory predicts that loop closure reactions become slower with increasing tail length until a limiting value is reached.[12−15] The effects of additional tails were predicted to depend on the size and on the stiffness of both the loop and the tails.[12] Experimental data on loop formation in long organic homopolymers showed slower dynamics for formation of type III loops compared to type I loops, but no clear dependency on the tail length was

(1) Nagi, A. D.; Regan, L. *Folding Des.* **1997**, *2*, 67−75.
(2) Wang, L.; Rivera, E. V.; Benavides-Garcia, M. G.; Nall, B. T. *J. Mol. Biol.* **2005**, *353*, 719−729.
(3) Viguera, A. R.; Serrano, L. *Nat. Struct. Biol.* **1997**, *4*, 939−946.
(4) Ladurner, A. G.; Fersht, A. R. *J. Mol. Biol.* **1997**, *273*, 330−337.
(5) Bieri, O.; Wirz, J.; Hellrung, B.; Schutkowski, M.; Drewello, M.; Kiefhaber, T. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 9597−9601.
(6) Krieger, F.; Fierz, B.; Bieri, O.; Drewello, M.; Kiefhaber, T. *J. Mol. Biol.* **2003**, *332*, 265−274.
(7) Krieger, F.; Fierz, B.; Axthelm, F.; Joder, K.; Meyer, D.; Kiefhaber, T. *Chem. Phys.* **2004**, *307*, 209−215.
(8) Chang, I.-J.; Lee, J. C.; Winkler, J. R.; Gray, H. B. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 3838−3840.
(9) Hudgins, R. R.; Huang, F.; Gramlich, G.; Nau, W. M. *J. Am. Chem. Soc.* **2002**, *124*, 556−564.
(10) Lapidus, L. J.; Eaton, W. A.; Hofrichter, J. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 7220−7225.
(11) Fierz, B.; Kiefhaber, T. Dynamics of unfolded polypeptide chains. In *Protein Folding Handbook*; Buchner, J., Kiefhaber, T., Eds.; WILEY-VCH: Weinheim, 2005; pp 805−851.
(12) Perico, A.; Beggiato, M. *Macromolecules* **1990**, *23*, 797−803.
(13) Friedman, B.; O'Shaughnessy, B. *Macromolecules* **1993**, *26*, 4888−4898.
(14) Ortiz-Repiso, M.; Rey, A. *Macromolecules* **1998**, *31*, 8356−8362.
(15) Ortiz-Repiso, M.; Rey, A. *Macromolecules* **1998**, *31*, 8363−8369.


10.1021/ja0666396 CCC: $37.00 

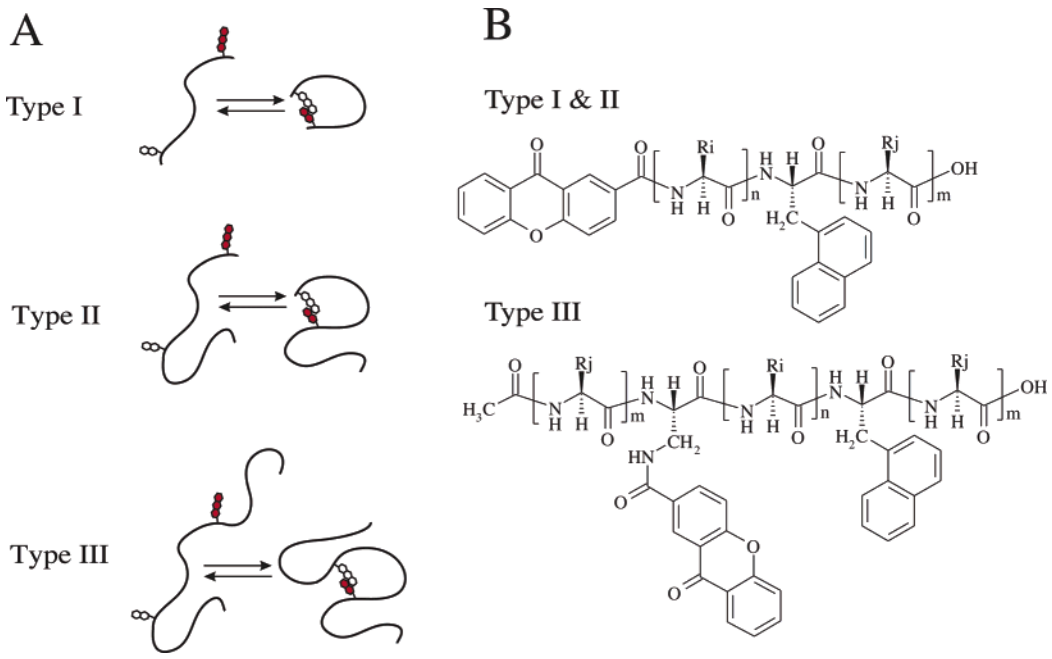


***Figure 1.*** (A) Schematic representation of type I, and type II and type II loop formation reactions. (B) General structures of peptides used in TTET experiments. In type I and type II loops the xanthone label is attached to the N-terminus and naphthylalanine (NAla) defines the C-terminus of the loop region. For type III loop peptides xanthone is attached to an α,β-diaminopropionic acid side chain at the N-terminal end of the loop region.

observed.[16] Measurements of loop formation between two internal residues in unfolded cytochrome c yielded slower kinetics[8] than expected from results on type I loop formation in model peptides.[6,11] However, to date it has not been systematically investigated whether end extensions affect the kinetics of loop formation in polypeptide chains.

We have studied the kinetics of end-to-end loop formation in unfolded polypeptide chains using triplet−triplet energy transfer (TTET) between a xanthone (Xan) moiety and a naphthylalanine (NAla) group.[6,7,11,17−19] TTET between Xan and NAla requires van der Waals contact between donor and acceptor, and all photophysical processes involved in this process occur on the picoseconds time scale.[6,20,21] This allows measurements of diffusion controlled intrachain loop formation reactions on an absolute time scale. In previous studies triplet donor and acceptor groups were attached at the ends of poly-serine and poly-(glycine-serine) chains of different lengths[6] and fragments from natural proteins.[7] In all investigated chains single-exponential kinetics for loop formation were observed with rate constants, $k_c$, around $10^7-10^8$ $s^{-1}$. These studies allowed us to elucidate the effect of amino acid sequence on loop formation and to derive scaling laws. In flexible glycine-serine copolymers it was found that for sufficiently long chains (number of loop residues $N > 20$) $k_c$ decreases with $N^{-1.7\pm0.1}$, which is the expected behavior for an excluded volume chain.[6,11] For shorter loops, the dynamics reached an upper limit of $k_c = 1.9 \times 10^8$ $s^{-1}$. Stiffer chains devoid of glycine residues showed similar scaling laws with only slightly slower kinetics compared to poly(glycine-serine).[6] Here we test whether additional tails beyond the loop region influence the kinetics of loop formation. We used TTET between xanthone and naphthylalanine to measure formation of type II and type III loops in several model polypeptide chains and compared the results to kinetics of type I loops with the same loop region. In addition, we studied the effect of length and amino acid sequence of the additional tails on the dynamics of loop formation.

## Results

**Type II vs Type I Loop Formation.** To investigate the effect of an additional tail at one end of a loop on the kinetics of contact formation we measured TTET between Xan and NAla attached to the ends of a loop consisting of four Gly-Ser pairs (see Figure 1B). This represents a model for an average size protein loop with high flexibility.[22] We used two different series of C-terminal extensions (see Table 1). In series 1, the extension is composed of $(Ser\text{-}Gly)_n$ units ($n = 1, 4, 7, 12$). In series 2 the tail consists of the sequence Ser-Gly-$(Thr\text{-}Gly\text{-}Gln\text{-}Ala)_n$-Gln-Ala-Ser-Gly ($n = 2, 4$). All peptides from both series are free of charged side-chains and unstructured as judged from circular dichroism spectra (data not shown). Figure 2 compares kinetics of contact formation in the *Xan*-$(Gly\text{-}Ser)_4$-*NAla*-Ser-Gly peptide and in a peptide with the same loop sequence but with the amino acids Ser-Gly-$(Thr\text{-}Gly\text{-}Gln\text{-}Ala)_4$-Gln-Ala-Ser-Gly attached C-terminal from the NAla moiety. The reaction was initiated by a 4 ns laserflash at 355 nm, which produces xanthone triplet states within the duration of the laserflash.[6,20,21] Loop formation between xanthone and naphthylalanine was monitored by the decrease in xanthone triplet absorbance at 590 nm.[6] Loop formation is slowed down 1.6-fold in the presence of the additional 20 C-terminal tail residues in the longer peptide. Varying the length of the C-terminal extension reveals that $\log(k_c)$ decreases linearly with increasing tail length (Figure 3A). However, the effect of increasing tail length on $k_c$ is not identical

(16) Lee, S.; Winnik, M. A. *Macromolecules* **1997**, *30*, 2633−2641.
(17) Bieri, O.; Kiefhaber, T. *Biol. Chem.* **1999**, *380*, 923−929.
(18) Krieger, F.; Möglich, A.; Kiefhaber, T. *J. Am. Chem. Soc.* **2005**, *127*, 3346−3352.
(19) Möglich, A.; Krieger, F.; Kiefhaber, T. *J. Mol. Biol.* **2005**, *345*, 153−162.
(20) Satzger, H.; Schmidt, B.; Root, C.; Zinth, W.; Fierz, B.; Krieger, F.; Kiefhaber, T.; Gilch, P. *J. Phys. Chem. A* **2004**, *108*, 10072−10079.
(21) Heinz, B.; B., S.; Root, C.; Satzger, H.; Milota, F.; Fierz, B.; Kiefhaber, T.; Zinth, W.; Gilch, P. *Phys. Chem. Chem. Phys.* **2006**, *8*, 3432−3439.
(22) Leszcynski, J. F.; Rose, G. D. *Science* **1986**, *234*, 849−855.

**Table 1.** Sequences of Peptides Used for TTET Experiments[a]

| peptides | sequence | *n* |
|---|---|---|
| series 1 | **Xan-(Gly-Ser)$_4$-NAla**-(Ser-Gly)$_n$-OH | 1, 4, 7, 12 |
| series 2 | **Xan-(Gly-Ser)$_4$-NAla**-Ser-Gly-(Thr-Gly-Gln-Ala)$_n$-Gln-Ala-Ser-Gly-OH | 2, 4 |
| series 3 | **Xan-Gly-Ser-NAla**-(Ser-Gly)$_n$-OH | 1, 6 |
| series 4 | **Xan-Gly-Ser-NAla**-(Thr-Gly-Gln-Ala)$_n$-Gln-Ala-Ser-Gly-OH | 2, 4 |
| series 5 | **Xan-Ser-Ser-Ser-NAla**-(Ser-Gly)$_n$-OH | 1, 6 |
| series 6 | **Xan-Ser-Ser-Ser-NAla**-(Thr-Gly-Gln-Ala)$_n$-Gln-Ala-Ser-Gly-OH | 2, 4 |
| series 7 | **Xan-(Gly-Ser)$_{10}$-NAla**-(Ser-Gly)$_n$-OH | 1, 6 |
| series 8 | **Xan-(Gly-Ser)$_{10}$-NAla**-(Thr-Gly-Gln-Ala)$_n$-Gln-Ala-Ser-Gly-OH | 2 |
| series 9 | Ac-(Gly-Ser)$_n$-**Dpr(Xan)-(Gly-Ser)$_4$-NAla**-(Ser-Gly)$_n$-OH | 2, 4, 6 |
| series 10 | Ac-Gly-Ser-Ala-Gln-(Ala-Gln-Gly-Thr)$_n$-**Dpr(Xan)-(Gly-Ser)$_4$-NAla**-(Thr-Gly-Gln-Ala)$_n$-Gln-Ala-Ser-Gly-OH | 2 |
| biotin | **Xan-(Gly-Ser)$_4$-NAla**-(Ser-Gly)$_8$-Lys(Biotin)-Ser-Gly-OH | |

[a] Ac: acetylated N-terminus. Xan: xanthonic acid. NAla: naphthylalanine. Dpr: α,β-diaminopropionic acid. Loop regions are printed in boldface.

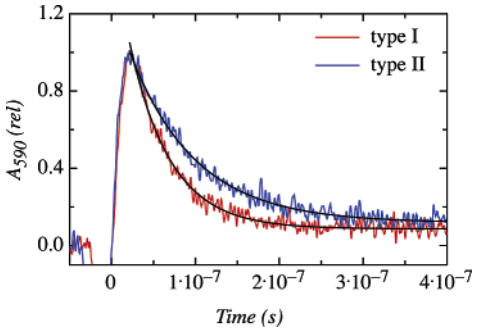


***Figure 2.*** Comparison of the kinetics of formation of a type I loop (Xan-(Gly-Ser)$_4$-NAla-Ser-Gly) and a type II loop (Xan-(Gly-Ser)$_4$-NAla-Ser-Gly-(Thr-Gly-Gln-Ala)$_4$-Gln-Ala-Ser-Gly-OH) with identical loop regions monitored by the decay in xanthone triplet absorbance at 590 nm. Single-exponential fits to the data yield rate constants of $(2.1 \pm 0.1) \times 10^7$ s$^{-1}$ and $(1.3 \pm 0.1) \times 10^7$ s$^{-1}$ for the type I and type II loop, respectively. The experiments were carried out in a 44% (v/v) glycerol/water mixture corresponding to $\eta = 3$ cP. Data for the type I loop were taken from ref 6.

for the two series of extensions when the number of amino acids in the tail, *N*, is compared. The rate constants are less affected by the (Ser-Gly)$_n$ extension compared to the (Thr-Gly-Gln-Ala)$_n$ series of equal length (Figure 3A) indicating that the effect on $k_c$ does not correlate with the number of chain segments in the tail.

To elucidate the origin of the observed effect of end extensions on the kinetics of type II loop formation we tested the correlation of $k_c$ with various parameters. Peptide motions in water are overdamped, and thus inertial forces should not contribute to the dynamics. The kinetics of loop formation should thus not correlate with the mass of the tail. Intrachain motions are rather coupled to solvent motion[23,24] as indicated by the observed strong viscosity-dependence of the kinetics of end-to-end loop formation.[5,19] The increased number of interactions of the polypeptide chain with the solvent in type II loops compared to type I loops may slow down chain dynamics. In this case the effect of the additional tail on $k_c$ should correlate with the solvent accessible surface area (ASA) of the tail. The exact ASA of an unfolded polypeptide chain is typically approximated by calculating the ASA of a fully extended chain.[25] Figure 3B shows log($k_c$) as a function of the ASA of the tail (ASA$_{tail}$) for both series of extension. The effect of ASA$_{tail}$ on $k_c$ is similar for both series, but there is still a slight deviation between the two series of extensions for long tails. The data can be fitted according to

$$k_c = k_0 \exp(\alpha \mathrm{ASA}_{tail}) \quad (1)$$

where $k_0$ represents the rate constants for type I loop formation. A fit to the data from both series of extensions gave values of $k_0 = (8.5 \pm 0.1) \times 10^7$ s$^{-1}$ and $\alpha = (-1.5 \pm 0.1) \times 10^{-4}$ Å$^{-2}$.

Theoretical considerations based on polymer theory suggested that the effect of additional tails on the dynamics of loop formation is a function of tail length and chain stiffness, i.e., of the dimensions of the tail.[12] Flory introduced the characteristic ratio, $C_N$, to calculate dimensions of polymer chains[26] which includes contributions from the number of chain segments and from chain stiffness

$$C_N = \frac{\langle r^2 \rangle}{Nl^2} \quad (2)$$

where $\langle r^2 \rangle$ denotes the mean square end-to-end distance, *N* is the number of chain segments, and $l^2$ is the squared segment length. $C_N$ for real chains is larger than unity; i.e., the dimensions of a real chain grow more strongly with increasing number of chain residues compared to an ideal chain where the end-to-end distance grows with $\langle r^2 \rangle = Nl^2$. $C_N$ is a function of *N* and increases with increasing chain length up to a limiting value, $C_\infty$, which is related to the persistence length of the chain.[26] In polypeptide chains $C_\infty$ is 9.27 for poly-alanine, and a similar value was proposed for all other amino acids except glycine and proline. Introduction of glycine residues leads to more flexible chains and thus to a smaller $C_N$ and decreased chain dimensions. Using the transformation matrices to calculate characteristic ratios for polypeptide chains determined by Flory[26] we calculated mean square end-to-end distances for all investigated polypeptide chains according to eq 2. This allows us to test for the effect of changes in chain dimensions caused by additional tails on the rate constant of loop formation. The absolute values of $\langle r^2 \rangle$ are most likely inaccurate, since only nearest neighbor potentials were considered by Flory. Long-range intrachain interactions[27] and peptide−solvent interactions, which are neglected in these calculations, were also shown to affect the chain dimensions. However, the calculated $\langle r^2 \rangle$-values allow an estimate of the relative dimensions of the different loops. Figure 3C shows log($k_c$) as a function of the average

(23) Kramers, H. A. *Physica* **1940**, *4*, 284−304.
(24) Szabo, A.; Schulten, K.; Schulten, Z. *J. Chem. Phys.* **1980**, *72*, 4350−4357.
(25) Myers, J. K.; Pace, C. N.; Scholtz, J. M. *Protein Sci.* **1995**, *4*, 2138−2148.
(26) Flory, P. J. *Statistical Mechanics of Chain Molecules*; Hanser Publishers: Munich, 1969.
(27) Pappu, R. V.; Srinivasan, R.; Rose, G. D. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 12565−12570.

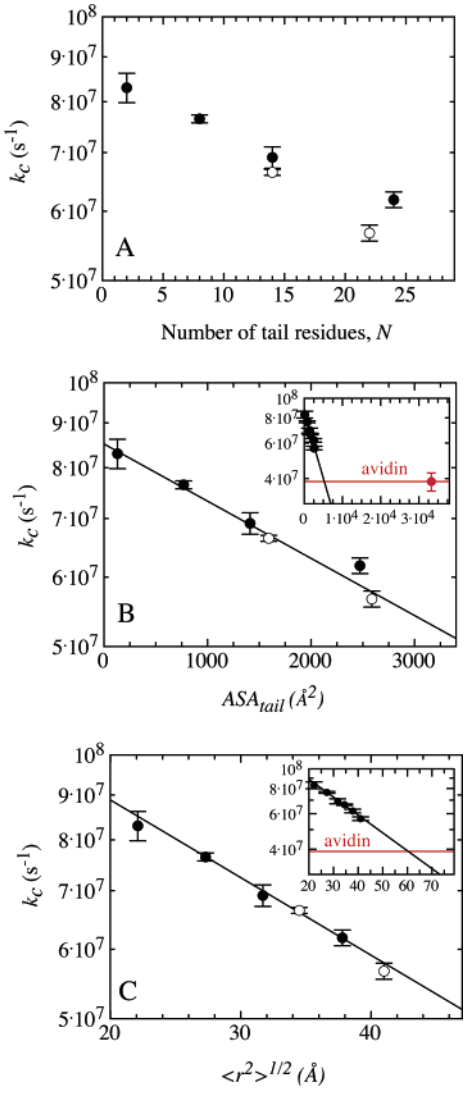


***Figure 3.*** (A) Effect of the number of tail residues, *N*, on the rate constants for type II loop formation, $k_c$, in a Xan-(Gly-Ser)$_4$-NAla loop with (Ser-Gly)$_n$ (●, series 1 peptides) or with Ser-Gly-(Thr-Gly-Gln-Ala)$_n$-Gln-Ala-Ser-Gly (○, series 2 peptides) C-terminal extensions. (B) Effect of the accessible surface area (ASA) of the C-terminal tail on the rate constants for type II loop formation. Peptides and symbols are the same as those in panel A. The solid line is an exponential fit of eq 1 to the data with $k_0 = (8.5 \pm 0.1) \times 10^7$ s$^{-1}$ and $\alpha = (-1.5 \pm 0.1) \times 10^{-4}$ Å$^{-2}$. (C) Effect of the calculated average end-to-end distance, $\sqrt{\langle r^2 \rangle}$, on the rate constants for type II loop formation. Peptides and symbols are the same as those in panel A. The solid line is an exponential fit of eq 3 to the data with $k_0 = (8.9 \pm 0.2) \times 10^7$ s$^{-1}$ and $\alpha = (-2.0 \pm 0.1) \times 10^{-2}$ Å$^{-1}$. The insets in panels B and C show $k_c$ for the biotinylated peptide bound to avidin. Rate constants are given for 22.5° in water.

end-to-end distance of the whole chain ($\sqrt{\langle r^2 \rangle}$) for both series of *Xan*-(Gly-Ser)$_4$-*NAla* type II loops. Changes in $\sqrt{\langle r^2 \rangle}$ have the same effect on $k_c$ in the (Ser-Gly)$_n$ and (Thr-Gly-Gln-Ala)$_n$ extension series, which can be described by

$$k_c = k_0 \exp(\alpha \sqrt{\langle r^2 \rangle}) \qquad (3)$$

where $k_0$ denotes the rate constant of the type I loop formation at $\sqrt{\langle r^2 \rangle_{loop}}$, which is 20 Å for the *Xan*-(Gly-Ser)$_4$-*NAla* loop. A fit of eq 3 to the data obtained from both series of peptides gives values of $k_0 = (8.9 \pm 0.2) \times 10^7$ s$^{-1}$ and $\alpha = (-2.0 \pm 0.1) \times 10^{-2}$ Å$^{-1}$ (see Table 2). This suggests that the effect of an additional tail on the dynamics of loop formation depends on the overall chain dimensions. This finding supports the results from polymer theory which suggested that both chain length and chain stiffness influence loop formation. The agreement of data obtained from the two series of chain extensions seems to be slightly better when $k_c$ is correlated with $\sqrt{\langle r^2 \rangle}$ (Figure 3C) compared to the correlation between $k_c$ and ASA$_{tail}$ (Figure 3B). However, $\sqrt{\langle r^2 \rangle}$ is correlated with ASA$_{tail}$, which makes it difficult to evaluate the major determinant for the effect of end extensions on chain dynamics.

***Table 2.*** Dimensions and Rate Constants for Loop Formation for the Different Type I Loops and Effect of a C-Terminal Extension on the Kinetics of Loop Formation

| loop sequence | $\sqrt{\langle r^2 \rangle_{loop}}$[a] (Å) | $k_0$[b] (s$^{-1}$) | $\alpha$[b] (Å$^{-1}$) |
|---|---|---|---|
| Xan-Gly-Ser-NAla | 11.3 | $(1.9 \pm 0.1) \times 10^8$ | $(-4.6 \pm 0.2) \times 10^{-2}$ |
| Xan-(Gly-Ser)$_4$-NAla | 20.0 | $(8.9 \pm 0.2) \times 10^7$ | $(-2.0 \pm 0.1) \times 10^{-2}$ |
| Xan-(Gly-Ser)$_{10}$-NAla | 30.3 | $(3.0 \pm 0.2) \times 10^6$ | $(-2.1 \pm 0.8) \times 10^{-2}$ |
| Xan-Ser-Ser-Ser-NAla | 14.8 | $(8.1 \pm 0.2) \times 10^6$ | $(-3.0 \pm 0.2) \times 10^{-2}$ |

[a] Calculated using the parameters given by Flory.[26] [b] $k_0$ and $\alpha$ were determined by a fit of eq 3 to the data shown in Figure 4A.

Polymer theory predicts that the effect of end extensions on the kinetics of loop formation should reach a limiting value for long tails.[12] The effect of a C-terminal extension on the kinetics of formation of the *Xan*-(Gly-Ser)$_4$-*NAla* loop does not reach a limiting value in the range of extensions accessible with standard peptide synthesis. To test whether a very large polypeptide chain attached to the C-terminus has a strong effect on the kinetics of loop formation, we synthesized a type II *Xan*-(Gly-Ser)$_4$-*NAla* loop containing a biotinylated lysine residue near the end of the C-terminal extension (Table 1). This allowed us to bind the peptide to avidin, a tetrameric glycoprotein from hen egg white ($M_r = 62\,400$, ASA = 21 000 Å$^2$), which strongly binds a biotin group per monomer with a dissociation constant of $1.3 \times 10^{-15}$ M.[28] Considering the dimensions of an avidin tetramer this leads to a very large increase in chain dimensions and should allow us to answer the question of a limiting value for type II loop dynamics. The rate constant for loop formation in the biotinylated peptide in the absence of avidin is $(6.3 \pm 0.2) \times 10^7$ s$^{-1}$. Upon addition of avidin, the rate constant decreases to $(3.9 \pm 0.4) \times 10^7$ s$^{-1}$, which corresponds to a 2.3-fold slower loop formation compared to the corresponding type I loop. The effect of avidin is much smaller than expected from the surface area (Figure 3B, inset) or the size (Figure 3C, inset) of the avidin tetramer with four biotinylated polypeptide chains bound (ASA ≈ 33 000 Å$^2$). This result shows that the kinetics of loop formation in the type II Xan-(Gly-Ser)$_4$-NAla loop become independent of tail length when the overall chain dimensions are larger than $\sqrt{\langle r^2 \rangle} = 62$ Å (Figure 3C, inset), i.e. when the overall chain dimensions are about three times the dimensions of the loop region. This limiting value can be explained by an increased flexibility of polypeptide chains toward the chain ends, which was directly observed in NMR experiments.[29,30] In type

(28) Green, N. M. *Adv. Protein Chem.* **1975**, *29*, 85−133.
(29) Schwalbe, H.; Fiebig, K. M.; Buck, M.; Jones, J. A.; Grimshaw, S. B.; Spencer, A.; Glaser, S. J.; Smith, L. J.; Dobson, C. M. *Biochemistry* **1997**, *36*, 8977−8991.
(30) Klein-Seetharaman, J.; Oikawa, M.; Grimshaw, S. B.; Wirmer, J.; Duchardt, E.; Ueda, T.; Imoto, T.; Smith, L. J.; Dobson, C. M.; Schwalbe, H. *Science* **2002**, *295*, 1719−1722.

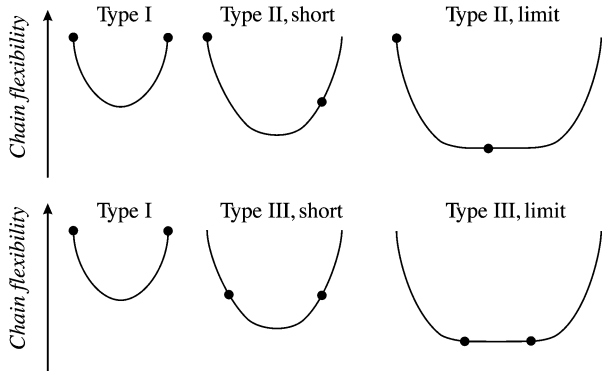


***Figure 4.*** Schematic representation of the position-dependence of chain flexibility and the resulting position of the points of contact formation in type I, type II, and type III loops. The scheme shows that a limiting value for the chain stiffness is reached for long type II and type III loops. The position-dependence of chain stiffness is based on theoretical work[31] and on results from NMR experiments.[29,30]

I loops, both contact points are located in flexible parts of the chain ends, which leads to faster dynamics compared to type II loops, where one of the contact points is located in a stiffer chain segment of the chain (Figure 4). This effect should be saturated for longer type II loops, when chain stiffness at the interior contact point has reached a limiting value. The observation of a limiting value for the effect of an end extension supports the model that the ASA of the tail is not the major determinant for the effect of a type II tail on $k_c$. In this case, a much larger effect of avidin on $k_c$ would be expected (Figure 3B, inset). In the following, we will therefore correlate the rate constants of loop formation with changes in $\langle r^2 \rangle$. This correlation allows us to compare our results with theoretical results from polymer theory.

**Effect of Loop Length and Loop Sequence on the Kinetics of Type II Loop Formation.** To test the effect of loop length and loop sequence on the kinetics of type II loop formation we investigated type II loops with different loop regions (see Table 1). A loop consisting of *Xan*-Gly-Ser-*NAla* (series 3, 4) was used as a model for an $i$, $i + 3$ contact, which is frequently observed in $\beta$-turn sequences. A *Xan*-Ser-Ser-Ser-*NAla* loop (series 5, 6) was used as a model for short and stiff loops, and a *Xan*-$(Gly\text{-}Ser)_{10}$-*NAla* loop (series 7, 8) served as a model for long and flexible loops. The effect of the additional tail on the kinetics of loop closure in the different peptides is shown in Figure 5. In agreement with our earlier results,[6] a large variation in $k_c$ is observed for the different type I loops, but also the effect of end extensions is different for the different loops. Contact formation in the short *Xan*-Gly-Ser-*NAla* and *Xan*-Ser-Ser-Ser-*NAla* loops is more strongly influenced by additional tails compared to the *Xan*-$(Gly\text{-}Ser)_4$-*NAla* and the *Xan*-$(Gly\text{-}Ser)_{10}$-*NAla* loops (Figure 5A). In the short *Xan*-Gly-Ser-*NAla* loop the two longest extensions show the same $k_c$ values within experimental error. For these peptides the total peptide dimensions are 2.6 times and 3.3 times larger, respectively, than the dimensions of the loop region and the rate constant for contact formation is 2.3-fold smaller than that in the corresponding type I reference loop. The relative size of the additional tail and the decrease in $k_c$ are comparable to the respective values found for the limiting regime in the longer *Xan*-$(Gly\text{-}Ser)_4$-*NAla* type II loop (Figure 3C, inset). This indicates that the rate constants for loop formation in the two

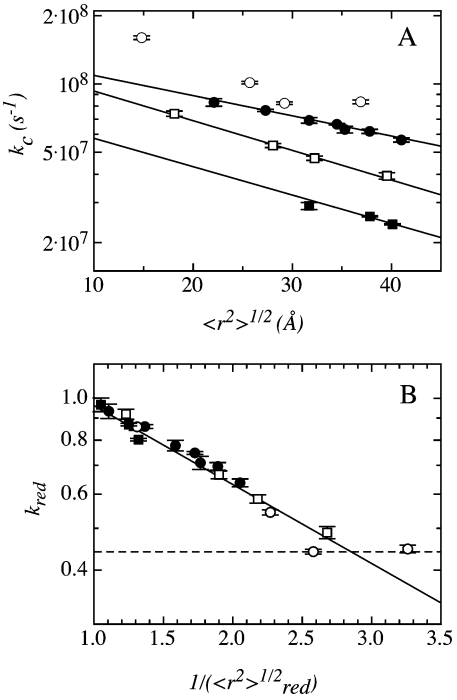


***Figure 5.*** (A) Effect of $\sqrt{\langle r^2 \rangle}$ on the rate constants for type II loop formation in different peptide systems with varying loop sequence: (○) Gly-Ser loop (series 3 and 4 peptides); (●) $(Gly\text{-}Ser)_4$ loop (series 1 and 2 peptides); (□) $Ser_3$ loop (series 5 and 6 peptides), (■) $(Gly\text{-}Ser)_{10}$ loop (series 7 and 8 peptides). For peptide sequences see Table 1. (B) Effect of the inverse of the reduced loop size, $\sqrt{\langle r^2 \rangle}_{red} = \sqrt{\langle r^2 \rangle}_{loop}/\sqrt{\langle r^2 \rangle}$, on the reduced rate constants, $k_{red} = k_c/k_0$ for type II loop formation. Peptide and symbols are the same as those in panel A. Values for $k_0$ were obtained from the fit to the data shown in panel A using eq. 3 and are given in Table 2. Values of $\sqrt{\langle r^2 \rangle}_{loop}$ for the different loops are also given in Table 2. A fit of eq 6 to the data gives $\alpha = (-4.5 \pm 0.1) \times 10^{-1}$.

longest peptides of the type II *Xan*-Gly-Ser-*NAla* loops may have already reached the limiting value.

**Effect of Relative Loop Size on Kinetics of Type II Loop Formation.** To directly compare the results from the different type II loops and to test whether the ratio of total peptide dimensions, $\sqrt{\langle r^2 \rangle}$, over loop dimensions, $\sqrt{\langle r^2 \rangle}_{loop}$, is rate-determining for loop formation, we calculated the reduced loop sizes, $\sqrt{\langle r^2 \rangle}_{red}$, for all type II loops according to

$$\sqrt{\langle r^2 \rangle}_{red} = \sqrt{\langle r^2 \rangle}_{loop}/\sqrt{\langle r^2 \rangle} \quad (4)$$

The $\sqrt{\langle r^2 \rangle}_{red}$ values were compared to reduced rate constants for loop formation, $k_{red}$, calculated according to

$$k_{red} = k_c/k_0 \quad (5)$$

which allows a comparison of the different type II loops by accounting for the differences in $k_0$. Figure 5B shows that $\log(k_{red})$ is correlated with the inverse of the reduced loop size, $1/\sqrt{\langle r^2 \rangle}_{red}$ and that the $k_{red}$ values of all type II loops show the same dependence on $1/\sqrt{\langle r^2 \rangle}_{red}$ which can be described by

$$k_{red} = \exp(\alpha/\sqrt{\langle r^2 \rangle}_{red}) \quad (6)$$

with $\alpha = (-4.5 \pm 0.1) \times 10^{-1}$ (eq. 6). Figure 5B further shows that the limiting value for $k_{red}$ and $\sqrt{\langle r^2 \rangle}_{red}$ are very similar for

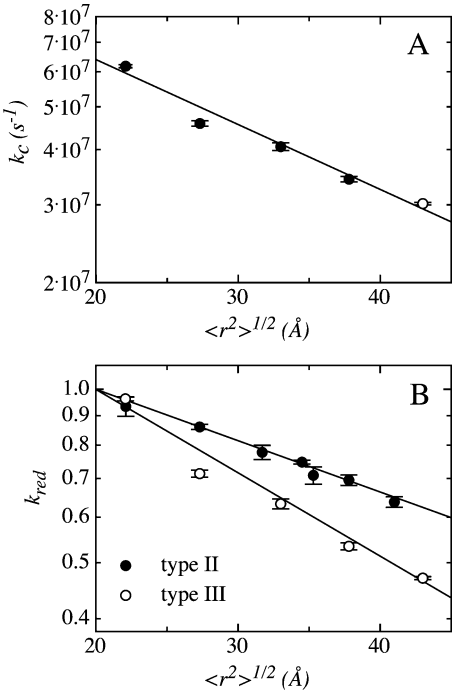


***Figure 6.*** (A) Effect of the calculated average end-to-end distance, $\sqrt{\langle r^2 \rangle}$, on the rate constants for type III loop formation in a *Dpr(Xan)*-(Gly-Ser)$_4$-*NAla* loop with (Ser-Gly)$_n$ (●, series 9 peptides) or with (Thr-Gly-Gln-Ala)$_n$ (○, series 10 peptides) C- and N-terminal extensions. The solid line is an exponential fit of eq 3 to the data with $k_0 = (6.4 \pm 0.1) \times 10^7$ s$^{-1}$ and $\alpha = (-3.3 \pm 0.1) \times 10^{-2}$ Å$^{-1}$. (B) Comparison of the reduced rate constants, $k_{red}$, for type II and type III loop formation with the identical loop sequence Xan-(Gly-Ser)$_4$-NAla. A fit of eq 6 to the data gives $\alpha = (-2.0 \pm 0.1) \times 10^{-2}$ for type II loops (see Figure 5B) and $\alpha = (-3.3 \pm 0.2) \times 10^{-2}$ for type III loops.

the type II loop sequences *Xan*-Gly-Ser-*NAla* and *Xan*-(Gly-Ser)$_4$-*NAla*. These results suggest that the reduced loop size $\sqrt{\langle r^2 \rangle}_{red}$ determines the effect of tails on contact formation in type II systems and that the limiting value for short and long loops is reached at similar values of $\sqrt{\langle r^2 \rangle}_{red} = 0.35$ with $k_{red} = 0.43$.

**Type III Loop Formation.** Interior-to-interior loop formation reactions (type III loops, see Figure 1A) are most relevant during protein folding. To probe the dynamics of type III loop formation we synthesized different series of peptides with extensions at both ends of the loop region (Table 1). To compare the dynamics of type III and type II loops we chose a *Xan*-(Gly-Ser)$_4$-*NAla* loop sequence. The xanthone moiety was attached to a diaminopropionic acid side chain at the N-terminal end of the loop region. This allows the addition of amino acids N-terminal from the loop regions in order to produce type III loops (see Figure 1B). The C-terminal end of the loop was again defined by a NAla residue. The C-terminal chain extensions were composed of (Ser-Gly)$_n$ ($n = 2, 4, 6$) or of (Thr-Gly-Gln-Ala)$_2$ sequences, whereas for the N-terminal extensions the sequences were inverted (series 9, 10 peptides). To compare these peptides to a type I system, a reference peptide with the same loop sequence but without extensions was produced. Due to the attachment of the xanthone moiety to a side chain via an amide bond in this series of peptides the kinetics were slightly slower than those in previously used type I systems with the xanthone attached directly to the N-terminus via an amide bond.

Figure 6A compares the rate constant of contact formation in the *Xan*-(Gly-Ser)$_4$-*NAla* type III loop with different extensions. Contact formation becomes slower when tails are attached to both sides of the loop region. A comparison of $k_c$ in peptides bearing (Ser-Gly)$_n$ or (Thr-Gly-Gln-Ala)$_n$ extensions shows that log($k_c$) linearly decreases with increasing overall chain dimensions, $\sqrt{\langle r^2 \rangle}$, as observed for type II loops (see Figure 3). A fit of eq 3 to the data yields $\alpha = (-3.3 \pm 0.1) \times 10^{-2}$ Å$^{-1}$ for the type III loops compared to $\alpha = (-2.0 \pm 0.1) \times 10^{-2}$ Å$^{-1}$ observed for type II loops with the same loop sequence (Figure 3C). To directly compare kinetics of *Xan*-(Gly-Ser)$_4$-*NAla* type II and III loops we calculated reduced rate constants according to eq 5, which eliminates the differences in $k_0$ caused by the different attachments of the xanthone labels. Figure 6B shows that an increase in the overall chain dimensions has a 1.7-fold stronger effect on the formation of type III loops compared to type II loops, but there is no evidence for a limiting value for the effect of end extensions on the dynamics of type III loops. However, the reduced dimension, $\sqrt{\langle r^2 \rangle}_{red}$, of the longest investigated type III loop is 0.47 whereas saturation in type II loops was not observed above $\sqrt{\langle r^2 \rangle}_{red} = 0.35$. It is expected that the kinetics for type III loop formation also reach a limiting value in longer chains, when both points of contact formation are located in the stiff central part of the chain (Figure 4B). If the limiting value would be reached at the same $\sqrt{\langle r^2 \rangle}_{red}$-value as in type II loops the maximum effect of tails in type III loops would result in a 3.8-fold decrease in $k_c$.

## Discussion

Contact formation between two amino acid side chains on a polypeptide chain is an elementary reaction in protein folding. To gain insight into the mechanism of loop formation we have previously studied end-to-end diffusion in model polypeptide chains and in fragments from natural proteins. In proteins, however, the majority of contacts involve internal residues. Local motions of a chain segment depend on the motions of its nearest neighbors which was predicted to lead to faster motions at the ends of the chain (see Figure 4).[31] This has been directly observed using NMR spin−lattice relaxation measurements in organic polymers[32−34] and in unfolded proteins.[29,30] Increased chain flexibility at the ends implies that formation of intrachain interactions is slower in the interior of a chain compared to the ends of the chain. Our results show that intrachain loop formation is slowed down by extending the chain beyond the loop region on either one side (type II loops) or on both sides (type III loops). Above a certain chain size the kinetics of type II loop formation become independent of chain dimensions. In this limit a 2.3-fold reduction in $k_c$ is observed. The limit is reached when the overall chain dimensions are about 3 times larger than the dimensions of the loop region, independent of loop size and amino acid sequence. The effect of end extensions is independent of the size and the type of the loops. The longest investigated loops were shown to behave like Gaussian chains,[6,11] whereas the shortest loops are similar in length or even shorter

(31) Perico, A. *Biopolymers* **1989**, *28*, 1527−1540.
(32) Allerhand, A.; Hailstone, R. *J. Chem. Phys.* **1972**, *56*, 3718−3720.
(33) Logan, T. M.; Theriault, Y.; Fesik, S. W. *J. Mol. Biol.* **1994**, *236*, 637−648.
(34) Frank, M. K.; Clore, G. M.; Gronenborn, A. M. *Protein Sci.* **1995**, *4*, 2605−2615.

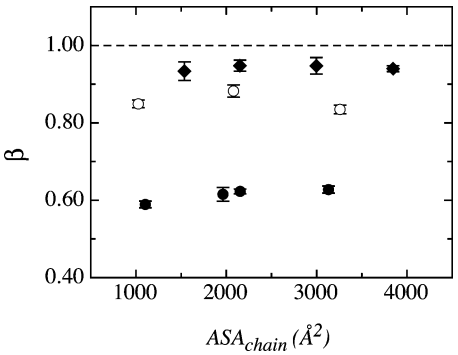


***Figure 7.*** Dependence of viscosity-dependence of the rate constants of loop formation ($\beta$) on the $ASA_{chain}$ in type II loops with the loop sequences *Xan*-(Gly-Ser)-*NAla* (●; series 3 and 4 peptides) and *Xan*-Ser-Ser-Ser-*NAla* (○; series 5 and 6 peptides). In addition, the results for the type III loop with the loops sequence *Xan*-(Gly-Ser)$_4$-*NAla* (◆; series 9, 10 peptides) are shown. The $\beta$-values are the results of the fits of the original data to eq 7 and reflect the sensitivity of the loop formation reactions toward changes in solvent viscosity.

than the persistence length of the chain. Despite these different chain properties, the same effect of $\sqrt{\langle r^2\rangle}_{red}$ on $k_c$ is observed for all loops (Figure 5B), which shows that only the chain dimensions determine the internal polypeptide chain dynamics. Comparison of type II and type III loops with identical loop regions and with identical overall chain dimensions reveals that extending the chain at both ends has a stronger effect on the kinetics of loop formation than extensions at only one end (Figure 6). The rate constants for type III loop formation also scale with the overall chain dimensions, but the decrease in $k_c$ with increasing $\sqrt{\langle r^2\rangle}$ is about 1.7 times larger for type III loops. This is compatible with an increased flexibility toward the ends of a polypeptide chain,[29,31] which allows faster formation of type II loops, which have one flexible end (Figure 4). These results indicate that local internal chain motions are strongly coupled to motions of other chain segments and imply that the internal flexibility at different positions along a polypeptide chain is the major origin for the differences in contact formation in type I, type II, and type III loops.

If the observed differences in the kinetics of loop formation would contain major contributions from the additional ASA of the tails, a difference in the viscosity-dependence for the different types of loops would be expected. We therefore measured the effect of solvent viscosity, $\eta$, on $k_c$ for all peptides. In all cases log($k_c$) was linearly dependent on log $\eta$ and the data could be fitted according to

$$k_c = k_c^0\left(\frac{\eta}{\eta_0}\right)^{-\beta} \qquad (7)$$

where $\eta_0$ is the reference solvent viscosity of water at 22.5 °C ($\eta_0 = 0.94$ cP), $k_c^0$ is the rate constant of end-to-end diffusion at $\eta_0$, $k_c$ is the rate constant at a given viscosity $\eta$, and the exponent $\beta$ reflects the sensitivity of the reaction to solvent viscosity. A $\beta$-value of 1 indicates that $k_c$ is inversely proportional to solvent viscosity, and a $\beta$-value of 0 indicates that $k_c$ is independent of solvent viscosity. For all peptides the $\beta$-values are independent of tail length and type of tails (Figure 7). For longer type I loops $\beta$-values around 1 are observed[19] and an additional increase due to the addition of the tail is not expected. However, for short loops the $\beta$-values are significantly smaller than unity and the increased number of solvent−peptide interactions in type II and type III loops should increase the $\beta$-values. Figure 7 shows that the $\beta$-values of the different type II and type III loops are independent of tail length and that identical overall ASA results in different $\beta$-values for short and long loop regions. This result indicates that the additional peptide−solvent interactions introduced by the tails have only minor contributions to the kinetics of loop formation.

We cannot exclude that the decreased rate constants for loop formation in type II and type III loops contain contributions from steric effects of the additional tails, which decrease the accessibility of the interacting groups. However, the observed difference in the effect of end extensions on the kinetics of formation of different loops (Figure 5A) argues against major contributions from local steric effects. This is supported by the observation that the limiting value for the effect of a type II extension is reached for rather long tails (Figure 3).

Results from theoretical work proposed that the effect of additional tails on the kinetics of loop III formation is a function of chain stiffness and reaches a limiting value when the length of the end extensions is around 0.5 to 2 times the loop length.[12,14] Due to restrictions in chain length given by peptide synthesis we could not investigate whether a limit for the rate constants of formation of type III loops exists. The longest synthesized type III chain consists of 35 amino acids with a calculated $\sqrt{\langle r^2\rangle}$-value of 43.0 Å. This is similar to the size of the smallest known independently folding protein subdomains. We could not attach avidin to both ends of the chain, since this leads to cross-linking of the tetramers and to precipitation.

Our results allow an extrapolation of data obtained from type I loops to kinetics of formation of type III loops in longer chains. The investigated *Xan*-(Gly-Ser)$_4$-*NAla* loop located in the interior of a typical 100 amino acid globular protein would form about 10 times slower compared to end-to-end contact formation in the same type I loop, if the extrapolations hold over the complete range. If a limiting value would be reached at the same reduced loop size, $\sqrt{\langle r^2\rangle}_{red}$, as observed for type II loops, the kinetics would only be slowed down about 4-fold in larger proteins. Type III loop formation was measured in GdmCl-unfolded Zn-cyt c by electron transfer between a $Ru(NH_3)_5^{3+}$ moiety attached to His33 and the Zn-porphyrine-coordinated His-18 residue,[8] which corresponds to an ~15 residue type III loop. Electron transfer occurred with a rate constant of $k_c = 4.0 \times 10^6$ s$^{-1}$ in the presence of 5.4 M GdmCl. Using TTET we measured a rate constant of $(1.9 \pm 0.1) \times 10^7$ s$^{-1}$ for end-to-end contact formation in 18-residue fragments from carp parvalbumin.[7] If this rate constant is rescaled to the type III limit and corrected for the effect of GdmCl on loop formation,[19] a value of $2.0 \times 10^6$ s$^{-1}$ is obtained, which is similar to $k_c$ measured in unfolded cytochrome c.

## Conclusions

We compared the kinetics of end-to-end contact formation (type I loops) to the formation of end-to-interior (type II loops) and interior-to-interior loops (type III loops) using triplet−triplet energy transfer from xanthone to naphthylalanine. The results show that formation of type II and type III loops compared to type I loops of the same loop size and amino acid sequence is slower until a limiting value is reached. The limiting regime is reached when the overall chain dimensions are about 3 times

larger than the loop dimensions. For these chains loop formation is about 2.5 times slower for type II loops compared to type I loops. Formation of type III loops is about 1.7 times slower than formation of type II loops of the same loop sequence and overall chain dimensions. These results show that local chain motions are strongly coupled to motions of other chain segments and are faster at the chain ends than in the interior of the chain. Our results suggest that differences in the kinetics of formation of type I, II, and II loops are caused by differences in flexibility at the different positions in a polypeptide chain (Figure 4). Interactions of the whole chain with the solvent contribute to the kinetics of loop formation as indicated by their strong viscosity-dependence. However, the increased number of peptide−solvent interactions in type II and type III loops compared to type I loops does not contribute significantly to the observed differences in the kinetics of loop formation for the different types of loops.

## Experimental Section

**Peptide Synthesis, Modification, and Purification.** All peptides were synthesized using standard fluorenylmethoxycarbonyl (Fmoc) chemistry on an Applied Biosystems 433A peptide synthesizer as described before.[6] The naphthalene moiety was included via 1-naphthylalanine (NAla) by peptide synthesis. The xanthone derivative 9-oxoxanthen-2-carboxylic acid was prepared according to Graham and Lewis[35] and introduced on resin by PyBOP mediated coupling in DMF to either the N-terminus or to the free amino group of an α,β-diaminopropionic acid residue. The biotin labeled peptides were prepared with biotinylated Fmoc-lysine (Novabiochem) which was introduced by standard peptide synthesis. All peptides were purified to >95% purity by preparative HPLC on an RP-8 column. The purity of the peptides was checked by analytical HPLC, and the mass was determined by MALDI or ESI mass spectrometry. Avidin (Fluka chemicals) was added to the purified peptide in buffered solution in slight excess.

**Triplet−Triplet Energy Transfer Measurements.** All measurements were performed on a Laser Flash Photolysis Reaction Analyzer (LKS.60) from Applied Photophysics. A Quantel Nd:YAG Brilliant laser (354.6 nm, 4 ns pulse width, 50 mJ) was used to selectively excite the xanthone moiety. Transient absorbance traces were recorded at 590 nm for the xanthone and at 420 nm for the naphthalene triplet band. All measurements were performed at 22.5 °C. For each experiment at least four traces were recorded, averaged, and fitted to exponential functions. For all peptides loop formation was measured in water and in the presence of different concentrations of glycerol to increase solvent viscosity. Measuring the viscosity-dependence of loop formation allows a more reliable determination of $k_c$ in water when loop formation is fast and allowed the determination of the viscosity-dependence of $k_c$ according to eq. 7. For all peptides $\ln(k_c)$ was linearly dependent on $\ln(\eta)$. Peptide concentrations were 20−100 μM and were determined by UV-absorption at 343 nm using a molar extinction coefficient of 3900 $M^{-1}$ $cm^{-1}$ for xanthone. All solutions were degassed prior to measurements. Viscosities were determined by a falling ball viscosimeter (Haake, Germany).

To test the effect of avidin on the xanthone triplet lifetime a donor-only control was synthesized with NAla replaced by phenylalanine which is not able to quench xanthone triplets.[7] This peptide revealed that binding of avidin does not influence the lifetime of the xanthone triplet state.

**Calculation of ASA.** Accessible surface areas (ASAs) of the peptides were calculated using the program MOLMOL[36] using a sphere of 1.4 Å diameter to probe the ASA.

**Data Analysis.** For data evaluation the ProFit and Matlab software packages were used.

**Acknowledgment.** We thank Annett Bachmann and Andreas Reiner for comments on the manuscript, Florian Krieger for discussion, and Josef Wey for synthesis of 9-oxoxanthen-2-carboxylic acid. This work was funded by a grant from the Schweizerische Nationalfonds and from the Volkswagen Stiftung.

JA0666396

(35) Graham, R.; Lewis, J. R. *J. Chem. Soc., Perkin Trans. I* **1978**, 876−881.
(36) Koradi, R.; Billeter, M.; Wüthrich, K. *J. Mol. Graphics* **1996**, *14*, 51−55.