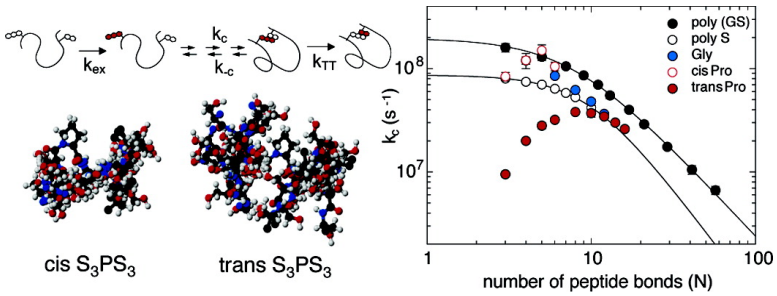
Article

Effect of Proline and Glycine Residues on Dynamics and Barriers of Loop Formation in Polypeptide Chains

Florian Krieger, Andreas Mglich, and Thomas Kiefhaber

J. Am. Chem. Soc., **2005**, 127 (10), 3346-3352• DOI: 10.1021/ja042798i • Publication Date (Web): 18 February 2005



More About This Article

Additional resources and features associated with this article are available within the HTML version:

- Supporting Information
- Links to the 13 articles that cite this article, as of the time of this article download
- Access to high resolution figures
- Links to articles and content related to this article
- Copyright permission to reproduce figures and/or text from this article

View the Full Text HTML

J|A|C|S
ARTICLES
Published on Web 02/18/2005

Effect of Proline and Glycine Residues on Dynamics and Barriers of Loop Formation in Polypeptide Chains

Florian Krieger, Andreas Möglich, and Thomas Kiefhaber*

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

Received November 30, 2004; E-mail: t.kiefhaber@unibas.ch

Abstract: Glycine and proline residues are frequently found in turn and loop structures of proteins and are believed to play an important role during chain compaction early in folding. We investigated their effect on the dynamics of intrachain loop formation in various unstructured polypeptide chains. Loop formation is significantly slower around trans prolyl peptide bonds and faster around glycine residues compared to any other amino acid. However, short loops are formed fastest around cis prolyl bonds with a time constant of 6 ns for end-to-end contact formation in a four-residue loop. Formation of short loops encounters activation energies in the range of 15 to 30 kJ/mol. The altered dynamics around glycine and trans prolyl bonds can be mainly ascribed to their effects on the activation energy. The fast dynamics around cis prolyl bonds, in contrast, originate in a higher Arrhenius pre-exponential factor, which compensates for an increased activation energy for loop formation compared to trans isomers. All-atom simulations of proline-containing peptides indicate that the conformational space for cis prolyl isomers is largely restricted compared to trans isomers. This leads to decreased average end-to-end distances and to a smaller loss in conformational entropy upon loop formation in cis isomers. The results further show that glycine and proline residues only influence formation of short loops containing between 2 and 10 residues, which is the typical loop size in native proteins. Formation of larger loops is not affected by the presence of a single glycine or proline residue.

Introduction

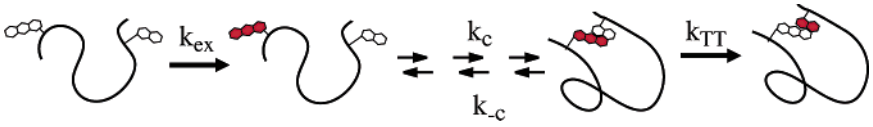
A major field of study in biophysical chemistry is the elucidation of the protein folding process. Starting from an ensemble of unfolded states a polypeptide chain has to form a large number of specific noncovalent intramolecular interactions during folding. To understand the complete folding reaction at the molecular level it is essential to characterize the structural and dynamic properties of the free energy landscape between the unfolded and the native state. Of particular interest is the ensemble of unfolded conformations as the starting point of the reaction. Several experimental techniques such as nuclear magnetic resonance spectroscopy[1−4] and small-angle X-ray scattering[5−7] have been applied to analyze structural properties of unfolded proteins. Less is known about the dynamics in unfolded polypeptide chains, which determine the formation of interactions during the early stages of the folding process. Especially the formation of loop and turn structures is important early in folding, since it allows the polypeptide chain to explore energetically favorable interactions and leads to chain compaction. Here we investigate the effect of glycyl and prolyl residues on intrachain contact formation in unfolded polypeptide chains. These residues frequently occur in turn and loop sequences[8,9] and may thus play an important role early in folding.

In recent years several experimental systems have been applied to obtain information on the time scale of loop formation in polypeptides.[10−16] We have used triplet−triplet energy transfer (TTET) from xanthone (Xan) to naphthylalanine (NAla) to directly measure the kinetics of intrachain contact formation in polypeptide chains.[12,15,17] TTET from a triplet donor to a triplet acceptor group involves transfer of two electrons (Dexter mechanism) and requires van der Waals contact between the

(1) Neri, D.; Billeter, M.; Wider, G.; Wüthrich, K. *Science* **1992**, *257*, 1559−1563.
(2) Logan, T. M.; Theriault, Y.; Fesik, S. W. *J. Mol. Biol.* **1994**, *236*, 637−648.
(3) Kortemme, T.; Kelly, M. J.; Kay, L. E.; Forman-Kay, J. D.; Serrano, L. *J. Mol. Biol.* **2000**, *297*, 1217−1229.
(4) 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.
(5) Damaschun, G.; Damaschun, H.; Gast, K.; Zirwer, D. *Biochemistry (Moscow)* **1998**, *63*, 259−275.
(6) Bachmann, A.; Kiefhaber, T. *Biophys. Chem.* **2002**, *96*, 141−151.
(7) Millett, I. S.; Doniach, S.; Plaxco, K. W. *Adv. Protein Chem.* **2002**, *62*, 241−262.
(8) Richardson, J. S. *Adv. Protein Chem.* **1981**, *34*, 167−339.
(9) Wilmot, C. M.; Thornton, J. M. *J. Mol. Biol.* **1988**, *203*, 221−232.
(10) Haas, E.; Katchalski-Katzir, E.; Steinberg, I. Z. *Biopolymers* **1978**, *17*, 11−31.
(11) Hagen, S. J.; Hofrichter, J.; Szabo, A.; Eaton, W. A. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 11615−11617.
(12) Bieri, O.; Wirz, J.; Hellrung, B.; Schutkowski, M.; Drewello, M.; Kiefhaber, T. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 9597−9601.
(13) Lapidus, L. J.; Eaton, W. A.; Hofrichter, J. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 7220−7225.
(14) Hudgins, R. R.; Huang, F.; Gramlich, G.; Nau, W. M. *J. Am. Chem. Soc.* **2002**, *124*, 556−564.
(15) Krieger, F.; Fierz, B.; Bieri, O.; Drewello, M.; Kiefhaber, T. *J. Mol. Biol.* **2003**, *332*, 265−274.
(16) Chang, I.-J.; Lee, J. C.; Winkler, J. R.; Gray, H. B. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 3838−3840.
(17) Krieger, F.; Fierz, B.; Axthelm, F.; Joder, K.; Meyer, D.; Kiefhaber, T. *Chem. Phys.* **2004**, *307*, 209−215.

10.1021/ja042798i CCC: $30.25

Scheme 1. Schematic Representation of Triplet−Triplet Energy Transfer Measurements in Unfolded Polypeptide Chains



two groups.[18] Formation of the xanthone triplet state and electron transfer between xanthone and naphthalene both occur on the 1−2 ps time scale,[19] and the transfer process is diffusion-controlled.[15,17] TTET between xanthone and naphthalene thus allows measurements of absolute rate constants for diffusional processes slower than about 10 ps (Scheme 1).[19,20] In earlier experiments the triplet donor and acceptor groups were introduced near the ends of various repetitive sequences or of fragments from natural proteins.[12,15,17] This enabled us to determine the effects of chain length and amino acid sequence on the dynamics. Single-exponential kinetics on the nanoseconds time scale were observed for contact formation in all peptides. In host−guest studies on short peptides only small effects of the amino acid sequence on local chain dynamics were observed. The only exceptions were glycine and proline, which show significantly different dynamics.[15] Glycine showed faster rate constants for contact formation than any other amino acid as expected from its increased backbone flexibility due to the lack of a C_β-atom. The presence of a proline residue leads to more complex dynamics with a slow and a very fast process of contact formation.[15] Here we investigate the molecular origins of the effects of prolyl and glycyl residues on the kinetics of loop formation. We show that the heterogeneity in the prolyl containing peptides can be assigned to dynamics of cis and trans Xaa-Pro isomers. Loops of varying length are used to assess the distance dependence of the effect exerted by a single glycine, cis proline, or trans proline residue. Measuring the temperature dependence of loop formation allows us to evaluate changes in the entropy and enthalpy of activation caused by the different residues. These results are compared to chain properties of the various peptides derived from all-atom simulations.

Materials and Methods

9-Oxoxanthen-2-carboxylic acid (xanthone acid) synthesis and peptide synthesis were carried out as described previously.[15] Pseudoproline (ΨPro) peptides were synthesized using a Fmoc-Val-[$\Psi^{Me,Me}$Ser] building block (Novabiochem) on an acid sensitive preloaded trityl resin (TentaGel S, Rapp Polymere) and trityl protecting groups. Cleavage conditions were 1% (v/v) trifluoroacetic acid (Fluka) in dichloromethane (Fluka). All peptides were purified by preparative HPLC. Purity and mass of the peptides were tested by analytical HPLC and mass spectrometry, respectively. Laserflash experiments were performed as described.[12,15] At the end of the fast exponential decay corresponding to TTET through intrachain contact formation a small amount of xanthone triplets (<10%) remains in the triplet state. These triplets decay on a much slower time scale with a rate corresponding to the intrinsic lifetime of xanthone triplets under the given conditions ($\tau > 20\ \mu$s) and may be due to oligomeric structures which do not allow TTET.[15] NMR spectra were recorded on a Bruker ARX600 spectrometer. The cis content of the proline-containing peptides was determined according to the method described by Reimer et al.[21]

Viscosity Dependence. All measurements were carried out in aqueous solutions. The viscosity was varied by adding 0−70% (v/v) glycerol (Aldrich). For all solutions, the solvent viscosity, η, was measured by using a falling sphere viscometer (Haake) at 22.5 °C. The viscosity dependencies of the rate constants for loop formation were analyzed using the empirical equation

$$k_c = k_c^0\left(\frac{\eta}{\eta_0}\right)^\beta \tag{1}$$

where k_c is the observed contact formation rate constant, η_0 is the reference viscosity of water at 22.5 °C (0.94 cP), k_c^0 is the rate constant for intrachain contact formation at η_0, and the exponent β reflects the sensitivity of the reaction to solvent viscosity. The β-values were independent of the cosolutes used to modify solvent viscosities. We typically used water/glycerol solutions to modify solvent viscosity. However, identical β-values were obtained for glycerol, glucose, sucrose, and ethylene glycol, indicating a genuine effect of solvent viscosity on the dynamics.

Temperature Dependence. The rate constants measured at different temperatures were corrected for the effect of temperature on water viscosity. Equation 2 was used to determine the viscosity corrected activation energies (E_a):

$$k_c^0(T) = k_c(T)\left(\frac{\eta_0}{\eta(T)}\right)^\beta = A\exp\left(\frac{-E_a}{RT}\right) \tag{2}$$

where A is the Arrhenius pre-exponential factor at 22.5 °C ($\eta_0 = 0.94$ cP). The β-values were taken from the viscosity dependence of k_c (eq 1; see Table 1). The viscosity of water at various temperatures was taken from ref 22.

Simulations of Chain Conformations. All-atom simulations of oligopeptides were carried out using a hard-sphere model. Contact radii, bond lengths, and bond angles corresponded to the values described by Rose and co-workers.[23] Atom contact radii were taken from Hopfinger[24] and softened by multiplying with a factor of 0.95. Bond lengths and angles were obtained from Engh and Huber.[25] The dihedral angle ω was fixed at 179.5° for trans peptide bonds and at −0.5° for a cis prolyl peptide bond. For proline residues the dihedral angle ϕ was fixed at −60°.[26] Peptide conformations were generated by randomly varying the values of the backbone dihedral angles ϕ and ψ and the side-chain dihedral angle χ_1 of serine residues. Calculated peptide conformations were checked for steric overlaps using the above hard-sphere contact radii. Following Pappu et al.[23] only pairs of atoms separated by at least four covalent bonds were tested for steric clashes. In addition, N-terminal amide proton and C-terminal carbonyl oxygen atoms were not checked for steric interactions. For each peptide at least 10^5 valid conformations without steric overlaps were generated by this procedure. End-to-end distances reflect the distance between the N-terminal amide nitrogen and the C-terminal carbonyl carbon atom.

(18) Klessinger, M.; Michl, J. *Excited states and photochemistry of organic molecules*; VCH: Weinheim, 1995.
(19) Satzger, H.; Schmidt, B.; Root, C.; Zinth, W.; Fierz, B.; Krieger, F.; Kiefhaber, T.; Gilch, P. *J. Phys. Chem. A* **2004**, *108*, 10072−10079.
(20) Fierz, B.; Kiefhaber, T. In *Protein Folding Handbook*; Buchner, J., Kiefhaber, T., Eds.; WILEY-VCH Verlag GmbH & Co KGaA: Weinheim, 2005; pp 805−851.
(21) Reimer, U.; Scherer, G.; Drewello, M.; Kruber, S.; Schutkowski, M.; Fischer, G. *J. Mol. Biol.* **1998**, *279*, 449−460.
(22) *Handbook of Chemistry and Physics*, 53rd ed.; CRC: Cleveland, OH, 1972.
(23) Pappu, R. V.; Srinavasan, R.; Rose, G. D. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 12565−12570.
(24) Hopfinger, A. J. *Conformational Properties of Macromolecules*; Academic Press: New York, 1973.
(25) Engh, R. A.; Huber, R. *Acta Crystallogr.* **1991**, *47*, 392−400.
(26) Schimmel, P. R.; Flory, P. J. *Proc. Natl. Acad. Sci. U.S.A.* **1967**, *58*, 52−59.

Table 1. Rate Constants and Activation Parameters for Intrachain Contact Formation in Various Peptides

sequence[a]	k_1[b] (10^6 s^{-1})	A_1[c] (s^{-1})	E_a^1[d] (kJ/mol)	β_1[e]	k_2[b] (10^6 s^{-1})	A_2[c] (s^{-1})	E_a^2[a] (kJ/mol)	β_2[e]	% cis[f]
SP	9.5 ± 0.5	(1.7 ± 0.8) × 10^{10}	18.3 ± 1.2	−0.75 ± 0.03	84 ± 8	(2.7 ± 4.0) × 10^{15}	42.9 ± 3.4	−0.69 ± 0.04	15 ± 1
SS	75 ± 4	(5.1 ± 1.0) × 10^{10}	16.1 ± 0.3	−0.83 ± 0.03					
SPS	20 ± 1	(4.1 ± 1.8) × 10^{10}	18.8 ± 1.2	−0.75 ± 0.02	120 ± 20	nd	nd	−0.77 ± 0.04	17 ± 2
SSS	72 ± 8	(2.8 ± 1.0) × 10^{10}	14.7 ± 0.8	−0.86 ± 0.06					
SGS	120 ± 10	(1.1 ± 1.3) × 10^{10}	11.4 ± 3.0	−0.83 ± 0.03					
VPS	9.5 ± 1.0	(2.9 ± 1.4) × 10^{10}	20.0 ± 1.2	−0.81 ± 0.05	92 ± 8	(1.6 ± 1.3) × 10^{13}	29.8 ± 3.4	−0.77 ± 0.04	16 ± 2
V(ΨP)S	9.7 ± 0.4	nd	nd	−0.80 ± 0.05	98 ± 5	nd	nd	−0.78 ± 0.06	72 ± 5
SSPS	28 ± 2	nd	nd	−0.83 ± 0.04	150 ± 20	nd	nd	−0.87 ± 0.06	18 ± 4
SVPS	16 ± 1	nd	nd	−0.86 ± 0.02	71 ± 5	nd	nd	−0.96 ± 0.04	18 ± 2
SV(ΨP)S	15 ± 1	nd	nd	−0.83 ± 0.06	75 ± 5	nd	nd	−0.98 ± 0.03	78 ± 6
S_4PS_4	37 ± 2	(2.4 ± 0.7) × 10^9	10.3 ± 0.7	−0.93 ± 0.03					
S_9	40 ± 2	(1.7 ± 0.8) × 10^9	9.2 ± 1.0	−0.92 ± 0.03					
S_4GS_4	48 ± 3	(1.6 ± 0.8) × 10^9	8.7 ± 1.2	−0.93 ± 0.02					

[a] Loop sequence, X, between the labels for TTET in peptides of the canonical sequence Xan-X-NAla-Ser-Gly. [b] k_1 and k_2 are the rate constants of the slower and the faster kinetic phase, respectively, at 22.5 °C in water (η = 0.94 cP). [c] Arrhenius pre-exponential factor in water at 22.5 °C (η = 0.94 cP) corrected for temperature effects on solvent viscosity according to eq 2. [d] Arrhenius activation energy corrected for temperature effects on solvent viscosity according to eq 2. The apparent activation energies before viscosity correction can be calculated by adding ($-\beta \times 17.7$) kJ/mol to the given values. [e] β-values reflect the sensitivity of the respective kinetic phase to solvent viscosity and were determined by fitting the viscosity dependence of the rate constants according to eq 1. The origin of the fractional β-values will be discussed elsewhere (F.K. and T.K., in preparation). [f] The cis content reflects the relative amplitude of the fast kinetic phases. The values obtained from measurements at different viscosities were averaged. [g] Loop formation in the cis isomer of the SPS peptide was too fast to reliably determine activation energies.

Ramachandran maps for cis and trans prolyl isomers were calculated for the central serine residue in Ser-Ser-Pro peptides[27] as described above. Conformational free energies were calculated by dividing the resulting Ramachandran maps into bins of 5 × 5 degrees. The number of conformations per bin was used to calculate relative free energies according to the Boltzmann equation. The most frequent conformation was arbitrarily assigned a value of 0; all other free energies are relative to that value.

The simulation program was written in ANSI C and compiled using the Bloodshed Dev-C++ development environment (http://www.bloodshed.net). Compiled programs were executed on IBM-compatible computers running Windows.

Results and Discussion

Kinetics of Loop Formation in Short Proline- and Glycine-Containing Peptides. To investigate the effect of prolyl and glycyl residues on chain dynamics in more detail we measured the kinetics of loop formation in host−guest peptides of the structure Xan-(Ser)$_x$-**Xaa**-(Ser)$_y$-NAla-Ser-Gly with $x,y = 0-7$ and Xaa = Gly, Ser, or Pro. Figure 1 compares TTET kinetics in the Xan-Ser-**Ser**-Ser-NAla-Ser-Gly peptide (SSS peptide)[28] with Xan-Ser-**Pro**-Ser-NAla-Ser-Gly (SPS peptide) and Xan-Ser-**Gly**-Ser-NAla-Ser-Gly (SGS peptide). These peptides serve as good models for tight protein turns with i, i+3 interactions which frequently occur at the end of β-hairpins.[29] Intrachain loop formation in the SSS and SGS peptides can be described by single-exponential kinetics with rate constants of (1.6 ± 0.1) × 10^7 s^{-1} and (2.5 ± 0.1) × 10^7 s^{-1}, respectively, in the presence of 52% glycerol. Under the same conditions, double exponential kinetics are observed for loop formation in the SPS peptide with rate constants of (4.7 ± 0.2) × 10^6 s^{-1} and (2.7 ± 0.3) × 10^7 s^{-1} and respective amplitudes of 80 ± 5% and 20 ± 5%. The faster kinetic phase in the SPS peptides is slightly faster than loop formation in the SGS peptide, whereas the slower phase is the slowest reaction observed in the three peptides. Since the faster kinetic phase of the proline peptide in water is near the resolution limit of the measurements, the experiments were carried out in the presence of increasing amounts of glycerol between 0% and 70% to slow chain dynamics due to increased solvent viscosity. The experiments on the water/glycerol mixtures allow a more reliable determination of the rate constant and amplitude of the fastest reaction via extrapolation to pure water (Figure 2A). The rate constants for loop formation in the SPS peptide in water at 22.5 °C are

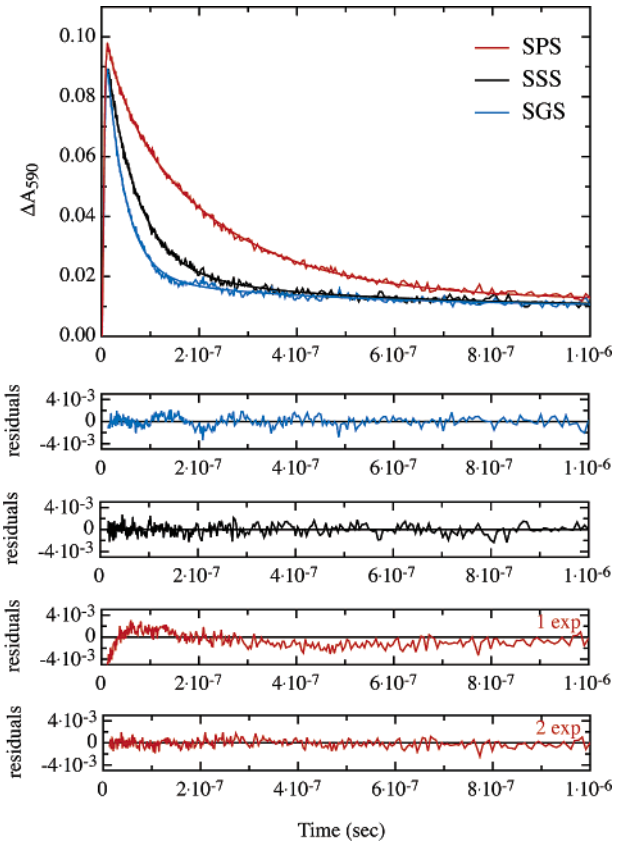


Figure 1. Effect of proline and glycine residues on the kinetics of loop formation in peptides of the sequence Xan-Ser-Xaa-Ser-NAla-Ser-Gly with Xaa = Ser, Gly, or Pro.

(27) Schimmel, P. R.; Flory, P. J. *J. Mol. Biol.* **1968**, *34*, 105−120.
(28) The peptides are abbreviated using one letter codes for the amino acid sequence (X) between the xanthone (Xan) and the naphthylalanine (NAla) labels. The peptides have the general sequence Xan-X-NAla-Ser-Gly. Thus, an SPS peptide denotes the peptide sequence Xan-Ser-Pro-Ser-NAla-Ser-Gly.
(29) Sibanda, B. L.; Thornton, J. M. *Nature* **1985**, *316*, 170−174.

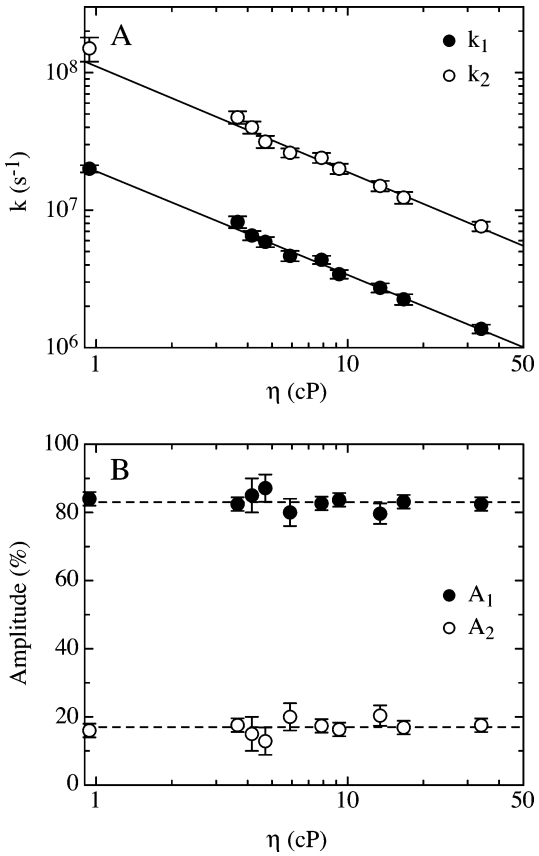
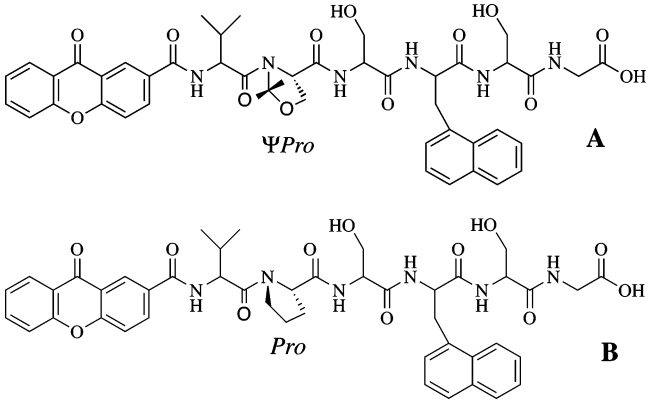


Figure 2. Viscosity dependence of the rate constants (A) and their respective amplitudes (B) for the kinetics of loop formation in a Xan-Ser-Pro-Ser-NAla-Ser-Gly peptide at 22.5 °C. Glycerol was used as cosolute to adjust solvent viscosity. The dashed lines in part B indicate the average amplitudes.

$(2.0 \pm 0.1) \times 10^7$ s^{-1} and $(1.2 \pm 0.2) \times 10^8$ s^{-1} (see Figure 2A and Table 1). For the SSS and SGS peptides rate constants of $(7.2 \pm 0.8) \times 10^7$ s^{-1} and $(1.2 \pm 0.1) \times 10^8$ s^{-1}, respectively, were determined (Table 1). Glycerol has no effect on the amplitudes of the two kinetic phases. The average amplitudes for the faster and slower process measured at various glycerol concentrations are 17.0 ± 2.2% and 83.0 ± 2.2%, respectively (Figure 2B).

A likely origin for the double exponential kinetics in the SPS peptide are different dynamics in cis and trans isomers of the Ser-Pro peptide bond. Xaa-Pro peptide bonds populate a large fraction of cis isomer due to the cyclic structure of the prolyl moiety. Studies on model peptides showed that the cis content depends on the preceding amino acid and varies between 7% and 38% [21] in contrast to about 0.15%−0.5% cis isomer at nonprolyl peptide bonds.[30] The observed 17% fast phase in the loop closure dynamics in the SPS peptide essentially reflects the cis−trans ratio at the Ser-Pro peptide bond in the SPS peptide, which has a cis content of 16 ± 2% as determined by 1D ^{1}H NMR spectroscopy (data not shown). Since cis−trans isomerization is slow ($1/k_{app}$ ∼20 s at 22.5 °C),[21,31,32] no equilibration between the two isomers occurs on the time scale of the TTET experiments. Thus loop formation in both isomers should be observable if they have different dynamics. However, also rotation around the ψ-angle of proline was shown to be slow with a rotation barrier around 60 kJ/mol.[33] This might give rise to the observed heterogeneity in the kinetics of loop formation in prolyl-containing peptides. To discriminate between cis/trans isomers of Xaa-Pro bonds and slow rotation around the proline ψ-angle as the molecular origin for the complex dynamics, we investigated the effect of the cis content on the kinetics of loop formation. Pseudo-proline (ΨPro) is known to increase the cis content of a Val-Pro peptide bond to about 80% due to steric effects induced by the methyl groups in the trans conformation.[32,34] We introduced ΨPro into two different peptides of the sequence Xan-Val-ΨPro-Ser-NAla-Ser-Gly (VΨPS-peptide **A**) and Xan-Ser-Val-ΨPro-Ser-NAla-Ser-Gly (SVΨPS-peptide) and compared the resulting kinetics of loop formation to the kinetics of the corresponding proline-containing peptides VPS (peptide **B**) and SVPS.



Both ΨPro-containing peptides show double exponential kinetics of contact formation with significantly increased amplitude of the faster reaction (70%−80%) compared to the corresponding prolyl-containing peptides (Figure 3A and Table 1). This shows that the faster reaction corresponds to the dynamics of peptides with a cis isomer and is not caused by slow rotation around the proline ψ-angle.

Temperature Dependence of Loop Formation. Information on the origin of the faster dynamics in chains with cis prolyl isomers was obtained from the temperature dependence of loop formation in the Xan-Val-Pro-Ser-NAla-Ser-Pro peptide (VPS peptide). Fitting the data to the Arrhenius equation yielded apparent activation energies of 43.2 ± 2.0 kJ/mol for the cis isomer and 35.9 ± 1.2 kJ/mol for the trans isomer. Since the dynamics were shown to be sensitive to solvent viscosity (Figure 2), the rate constants measured at different temperatures had to be corrected for contributions from the effect of temperature on water viscosity to obtain the actual activation energies. Since, the kinetics of loop formation do not correlate with $1/\eta$ but show fractional viscosity dependencies ($k \sim 1/(\eta)^{\beta}$; see eq 1 and Table 1), we used eq 2 to determine viscosity-corrected activation energies. The resulting activation energies (E_a) were 29.8 ± 3.4 and 20.0 ± 1.2 kJ/mol for loop formation in peptides with a cis and a trans Val-Pro peptide bond, respectively (Figure 3B, Table 1). The results further revealed largely different

(30) Scherer, G.; Kramer, M. L.; Schutkowski, M.; Reimer, U.; Fischer, G. *J. Am. Chem. Soc.* **1998**, *120*, 5568−5574.
(31) Brandts, J. F.; Halvorson, H. R.; Brennan, M. *Biochemistry* **1975**, *14*, 4953−4963.
(32) Keller, M.; Sager, C.; Dumy, P.; Schutkowski, M.; Fischer, G. S.; Mutter, M. *J. Am. Chem. Soc.* **1998**, *120*, 2714−2720.
(33) Nagaraj, R.; Venkatachalapathi, Y. V.; Balaram, P. *Int. J. Pept. Protein Res.* **1980**, *16*, 291−298.
(34) Dumy, P.; Keller, M.; D. E., R.; Rohwedder, B.; Wohr, T.; Mutter, M. *J. Am. Chem. Soc.* **1997**, *119*, 918−925.

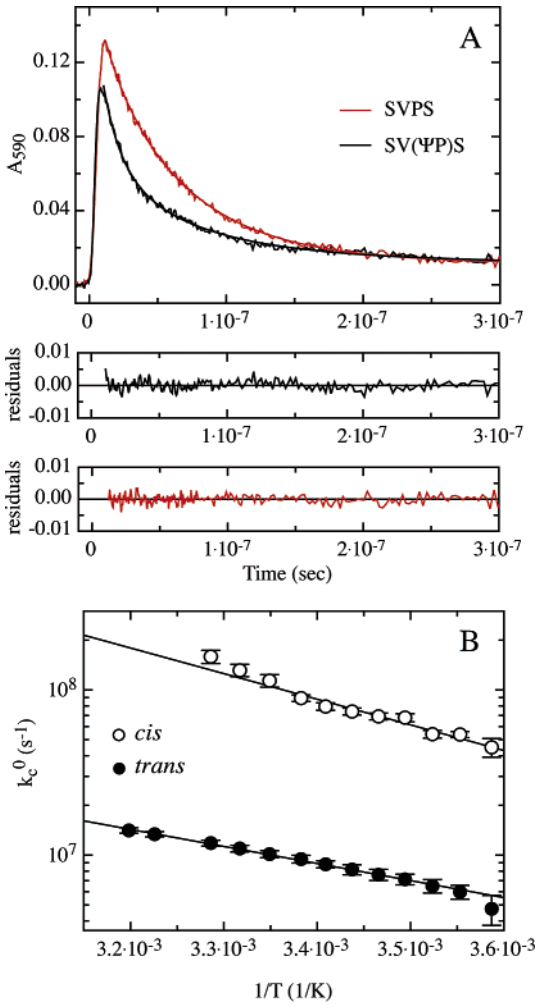


Figure 3. (A) Kinetics of intrachain contact formation in Xan-Val-**Xaa**-Ser-NAla-Ser-Gly peptides with Xaa = ΨPro (peptide **A**) or Pro (peptide **B**). (B) Temperature dependence of intrachain contact formation in cis and trans isomers of the Xan-Val-Pro-Ser-NAla-Ser-Gly peptide. A fit to the Arrhenius equation (eq 2) gives viscosity-corrected activation energies and pre-exponential factors (*A*) indicated in Table 1.

Arrhenius pre-exponential factors (*A*) of $(1.6 \pm 1.3) \times 10^{13}$ s^{-1} and $(2.9 \pm 1.4) \times 10^{10}$ s^{-1} for loop formation in the cis and trans isomer, respectively. Obviously, the fast loop formation of the cis isomer is the result of a significantly higher pre-exponential factor despite a higher activation energy compared to the trans peptide bonds. This indicates favorable entropic contributions to the rate constant for loop closure in the cis isomer. A significantly higher activation energy of 42.9 ± 3.4 kJ/mol and a higher pre-exponential factor of $(2.7 \pm 4) \times 10^{15}$ s^{-1} are found for the cis isomer in the shorter Xan-Ser-Pro-NAla-Ser-Gly peptide (SP peptide; Table 1), whereas the activation parameters for dynamics in the trans isomer of this peptide are similar to the respective values in the trans isomer of the VPS peptide (Table 1). For all proline-containing peptides the pre-exponential factors for the loop closure dynamics in the trans Pro isomers are similar to the values in polyserine peptides of similar length, whereas the activation energies in the polyserine peptides are significantly lower (Table 1). This indicates that the presence of a trans prolyl isomer mainly influences the height of the energy barrier for loop formation but not the entropic contributions. The activation parameters for loop formation in the SGS peptide reveal that

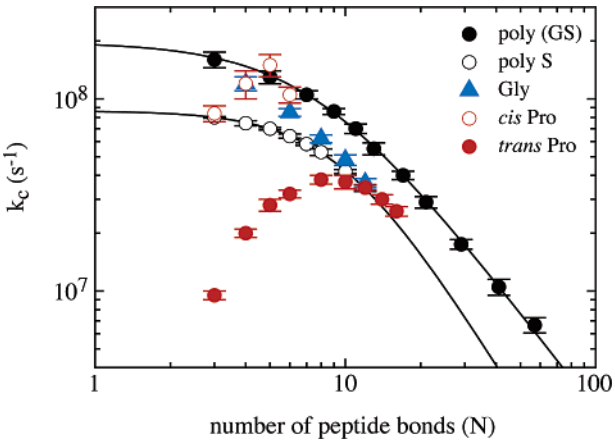


Figure 4. Effect of a single proline or glycine residue on the rate constant of loop formation in peptides of varying length compared to polyserine and poly(glycine-serine) peptides. Data for polyserine and poly(Gly-Ser) were taken from ref 15. The lines describe the loop size dependence of intrachain loop formation in the polyserine and poly(glycine-serine) series as described in ref 15.

glycine decreases the activation energy for loop formation compared to serine- and proline-containing peptides (Table 1). A likely origin of this effect is the lower rotation barrier around the ϕ, ψ angles in glycine. This accelerates loop formation despite a decreased pre-exponential factor, which can be explained by increased conformational space around glycyl residues (Table 1).

Linear Arrhenius plots were observed for all peptides, which shows that loop formation is not associated with significant changes in heat capacity ($\Delta C_p \approx 0$). This argues against major changes in solvation during the rate-limiting motions for loop formation. The results on the dynamics in the different proline-containing peptides further reveal that the amino acid preceding the prolyl residue significantly influences the dynamics of both the cis and the trans isomer. In all short peptides Val-Pro bonds exhibit about 2-fold slower loop formation kinetics compared to Ser-Pro bonds (Table 1). This indicates steric effects on loop formation imposed by the amino acid preceding the prolyl residue. Interestingly, the bulkier ΨPro residue gives virtually the same rate constant for loop formation as Pro in all peptides.

Effect of Loop Size. The results show that proline and glycine residues strongly influence the kinetics of formation of short loops. To test whether these effects are also observed over longer distances we measured loop formation in peptides containing a single prolyl or glycyl residue at a central position and varied the loop size. Figure 4 compares the kinetics of loop formation in peptides of the sequence Xan-(Ser)$_x$-**Xaa**-(Ser)$_y$-NAla-Ser-Gly (Xaa = Gly or Pro, $x,y = 0-7$) with the results from polyserine and poly(glycine-serine) loops. The effect of a central prolyl and glycyl residue decreases with increasing loop size. A glycyl residue has no effect on loop formation if the sites of contact formation are further than five amino acids from both sides of the glycin. Peptides with $N \geq 11$ (N = number of peptide bonds between the two sites of contact formation) show the same rate constants for contact formation as polyserine chains of the same length (Figure 4). A more complex behavior is observed for prolyl residues. Only for $x,y \leq 2$ the dynamics of peptides with cis and trans isomers can be resolved, suggesting that the dynamics of two isomers are similar in longer peptides (Figure 4). In short peptides the cis isomer always shows faster

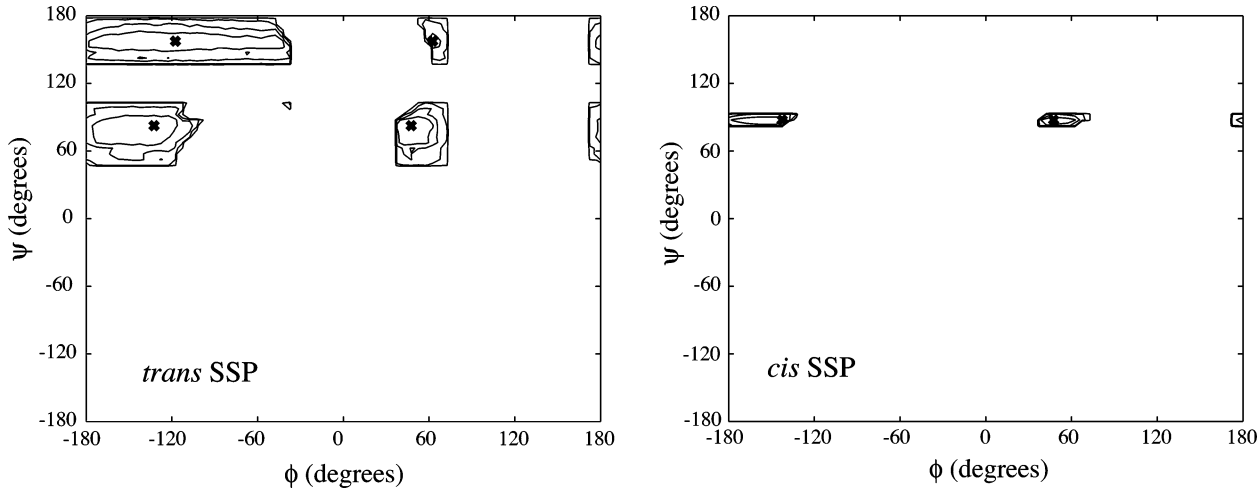


Figure 5. Ramachandran maps for the serine residue in either the cis or the trans Ser-Pro peptide bond in Ser-Ser-Pro peptides. The contour levels denote free energies of 0.5, 1, 2, and $4k_BT$. The crosses indicate the conformation of lowest free energy.

loop formation than the trans isomer, irrespective of whether donor and acceptor are separated by an even or an uneven number of peptide bonds (Figure 4 and Table 1). Formation of short loops becomes faster with increasing loop size in chains with cis Pro isomers and the rate constant reaches a maximum of $k_c = 1.5 \times 10^8$ s^{-1} for $N = 5$, i.e., if two amino acids are located on both sides of the cis prolyl bond. The rate constant for loop formation in peptides with trans prolyl isomers also increases with increasing chain length until a loop size of 9−10 residues and matches dynamics of a polyserine chain at $N = 11$. Formation of single proline-containing loops with $N \geq 11$ behave similar to polyserine chains and show decreasing rate constants with increasing loop size (Figure 4). These results show that single glycine and proline residues affect chain dynamics only very locally and do not influence formation of loops if the contact sites are further than five amino acids on either side of the glycyl or prolyl residue.

Effect of Cis and Trans Prolyl Peptide Bonds on Chain Dimensions. The temperature dependence of the loop closure reaction showed that the faster dynamics in cis prolyl isomers are due to a significantly increased pre-exponential factor compared to trans isomers indicating entropically favored loop formation in cis isomers. A likely origin of this effect are differences in the equilibrium properties of chains containing cis or trans isomers. This prompted us to evaluate the sterically allowed conformations for the model polypeptide chains. Using all-atom Monte Carlo simulations of conformational space introduced by Rose and co-workers[23,35] we calculated the probability distribution of end-to-end distances for different chains containing either cis or trans Ser-Pro isomers and compared the results to the dimensions of polyserine chains. In these calculations all bond lengths, bond angles, and peptide bond torsion angles, ω, were held at the values given by Pappu et al.[23] Sterically allowed ϕ,ψ pairs were evaluated for each amino acid, and the resulting allowed regions in the ϕ,ψ space were used to generate energetically favorable conformations for each peptide.[23] The results show that the energetically favorable conformational space of a cis prolyl isomer in a Ser-Pro bond is drastically reduced compared to the trans isomer (Figure 5). This is in agreement with an analysis of nonprolyl cis peptide bonds in protein structures, which revealed that also in native proteins the conformational space of cis peptide bonds is significantly restricted compared to trans peptide bonds.[36] Our calculations further revealed that a central cis Ser-Pro peptide bond leads to a distance distribution function that is significantly shifted to shorter end-to-end distances compared to a trans bond in all peptides (Figure 6). This is consistent with faster rate constants for intrachain contact formation in chains with a cis peptide bond due to a smaller average distance between the sites of contact formation. However, the relative difference in end-to-end distance between cis and trans isomers is small for longer chains, which might explain the inability to resolve both kinetic phases for contact formation in longer peptides (Figure 6A). These results show that a single cis Xaa-Pro peptide bond significantly influences the available conformational space of a polypeptide chain and leads to reduced chain dimensions. This is compatible with earlier calculations on poly(L-proline), which also showed that the presence of a small fraction of cis prolyl isomers drastically reduces the chain dimensions.[37] A similar behavior was observed for stereoirregular vinyl polymers[38] and in stereoirregular 1,4-polybutadiene and 1,4-polyisoprene,[39] where the presence of a small fraction of stereoirregularity significantly reduces chain dimension.

The restricted conformational space and the shorter end-to-end distances found for cis Xaa-Pro peptide bonds explain the higher pre-exponential factor for loop closure reactions due to a smaller loss of conformational entropy. The results from the simulations are further compatible with higher activation energies for chain motions in cis isomers due to increased barriers for bond rotation caused by steric clashes during bond rotations on a restricted conformational space (Figure 6B). The simulations also show that trans prolyl isomers have similar chain dimensions as the corresponding serine-conatinaing peptides, which is in agreement with virtually identical pre-exponential factors for loop formation for these peptides.

(35) Fitzkee, N. C.; Rose, G. D. *Protein Sci.* **2004**, *13*, 633−639.
(36) Jabs, A.; Weiss, M. S.; Hilgenfeld, R. *J. Mol. Biol.* **1999**, *286*, 291−304.
(37) Tanaka, S.; Scheraga, H. A. *Macromolecules* **1975**, *8*, 623−631.
(38) Flory, P. J. *Statistical Mechanics of Chain Molecules*; Hanser Publishers: Munich, 1969.
(39) Tanaka, S.; Nakajima, A. *Polym. J.* **1972**, *3*, 600−506.

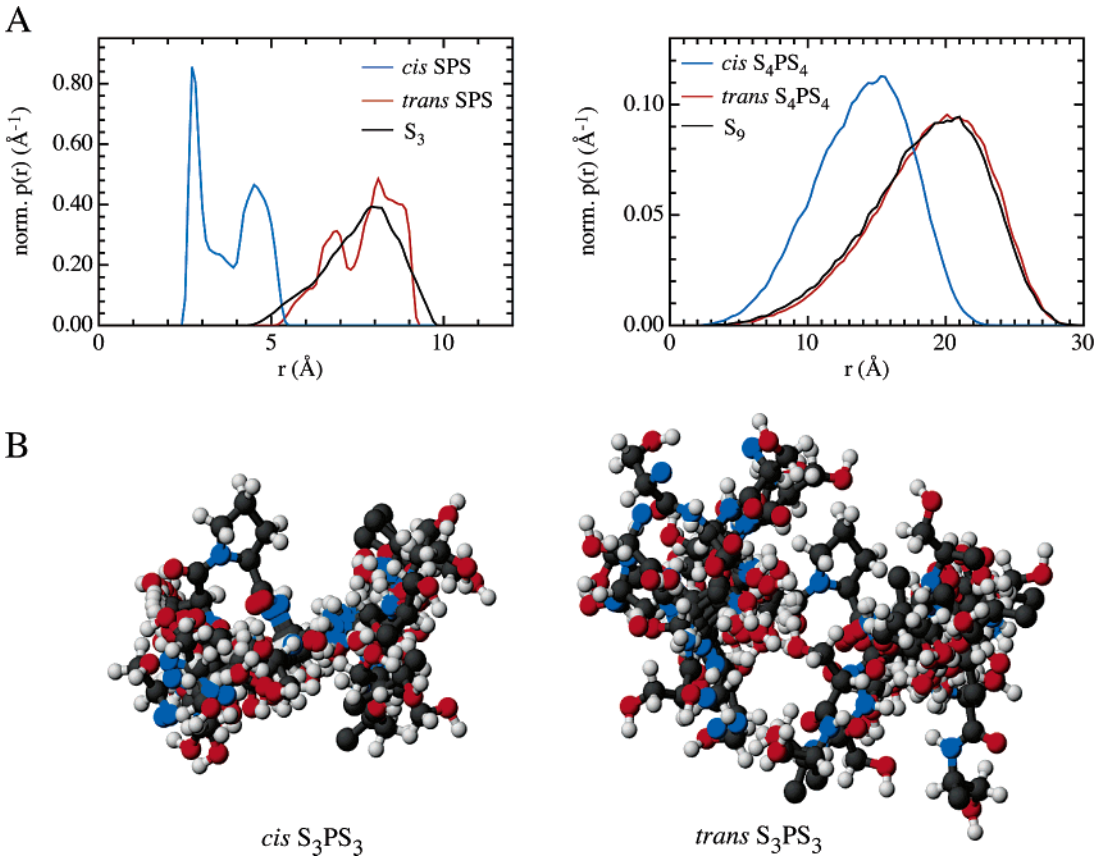


Figure 6. (A) Normalized end-to-end distance distribution in cis and trans prolyl isomers in various peptides compared to polyserine peptides of the same length (B) Representative structures (10) of cis and trans prolyl isomers of a Ser_3-Pro-Ser_3 peptide. The structures are overlaid with respect to identical orientation of the central prolyl residue. The structures are displayed using the program MOLMOL.[41]

Conclusions

Glycine and proline residues have a major influence on the kinetics of loop formation in proteins. Glycine accelerates loop formation by decreasing the activation energy, whereas trans prolyl bonds slow loop formation by increasing the barrier height. The most interesting effect is observed for loop formation around cis prolyl bonds, which show the fastest kinetics of all sequences despite an increased activation energy. The fast loop formation around cis prolyl isomers is due to largely restricted conformational space and shorter end-to-end distances compared to trans isomers which results in a drastically increased pre-exponential factor for loop formation. A single prolyl or glycyl residue influences chain dynamics only locally and has no effect on formation of loops longer than 11 residues. The most abundant loop and turn structures in natural proteins have average sizes of 6−10 residues (ω-loops)[40] and of 2−7 residues (β-hairpin loops)[29] and are especially rich in proline and glycine. Our results show that the rate constants for formation of these loops during protein folding can be modulated by a factor of 10 by the presence of either glycine/cis proline or trans proline, which might help to direct the folding polypeptide chain to productive folding routes. The significantly different dynamic and conformational properties of loops containing cis and trans prolyl peptide bonds may further be important for the function of prolines as molecular switches in regulation of cellular processes.

Acknowledgment. We thank Gerd Scherer for recording NMR spectra, Thiery Mini and Paul Jenö for mass spectroscopy, Annett Bachmann and Andreas Reiner for comments on the manuscript, and Beat Fierz for discussions. This work was funded by a grant from the Schweizerischen Nationalfonds (SNF).

JA042798I

(40) Leszcynski, J. F.; Rose, G. D. *Science* **1986**, *234*, 849−855.
(41) Koradi, R.; Billeter, M.; Wüthrich, K. *J. Mol. Graphics* **1996**, *14*, 51−55.