

**Dynamics in Unfolded Polypeptide Chains as Model
for Elementary Steps in Protein Folding**

INAUGURALDISSERTATION

zur

Erlangung der Würde eines Doktors der Philosophie

vorgelegt der

Philosophisch-Naturwissenschaftlichen Fakultät

der Universität Basel

von

Florian Krieger

aus

Deutschland

Basel, 2004

Genehmigt von der Philosophisch-Naturwissenschaftlichen Fakultät auf Antrag von

Prof. Dr. Thomas Kiefhaber

Prof. Dr. Joachim Seelig

Basel, den 16.11.2004

Prof. Dr. Hans-Jakob Wirz

(Dekan)

Table of contents

1 INTRODUCTION.....	1
1.1 Proteins as the major functional biomacromolecules	1
1.2 The protein folding problem	1
1.3 Structural heterogeneity in the unfolded state of proteins	5
1.4 Elementary steps in protein folding.....	6
1.5 Intrachain loop formation.....	8
1.6 Measuring intrachain contact formation in unfolded peptide chains.....	10
1.7 Triplet-Triplet Energy Transfer.....	14
2 AIMS OF RESEARCH	18
3. RESULTS	20
3.1 The xanthone/naphthalene system as triplet donor/acceptor pair for studying intrachain contact formation in polypeptides	20
3.1.1 Tests for diffusion-controlled triplet-triplet energy transfer from xanthone to naphthalene	20
3.1.2 Influence of natural amino acids on xanthone triplet lifetime.....	24
3.2 Effect of loop length on end-to-end contact formation in flexible polypeptides	25
3.3 Effect of amino acid sequence on intrachain contact formation in peptides.....	29
3.4 Barrier-limited and diffusive chain dynamics.....	35
3.4.1 Activation barriers and viscosity dependence of intrachain diffusion in polypeptides	35
3.4.2 Kinetic isotope effects on the dynamics of loop formation in short peptides ...	44
3.5 Solvent effects on intrachain contact formation in unstructured peptides.....	47
3.5.1 Guanidine hydrochloride and urea.....	47
3.5.2 Other co-solutes	50

3.6 Applying triplet-triplet energy transfer to measure loop formation in natural protein sequences	53
3.7 Intrachain diffusion of α -helices in helix-turn-helix motifs	58
3.7.1 Induction of secondary structure in the Ca^{2+} -binding EF-hand motif from carp muscle β -parvalbumin	58
3.7.2 Design of peptides with de novo helix-turn-helix motifs.....	61
3.7.3 Position effect of charges on the structure of alanine based peptides.....	65
4 SUMMARY	72
5 ACKNOWLEDGEMENTS.....	75
6 REFERENCES.....	76
7. PUBLISHED AND SUBMITTED WORK.....	86
7.1 Dynamics of Unfolded Polypeptide Chains as Model for the Earliest Steps in Protein Folding (Krieger et al., 2003)	86
7.2 Intrachain diffusion in a protein loop fragment from carp parvalbumin (Krieger al., 2004)	96
7.3 Ultrafast Quenching of the Xanthone Triplet by Energy Transfer: New Insight into the Intersystem Crossing Kinetics (Satzger et al., 2004)	103
7.4 Molecular Basis of the Effect of Urea and Guanidinium Chloride on the Dynamics of Unfolded Polypeptide Chains (Möglich et al., 2004)	111
7.5 Effect of Proline and Glycine Residue on the Dynamics and Barriers of Loop Formation in Polypeptide Chains (Krieger et al., submitted).....	121
7.6 Kinetics and Barriers of Loop Formation in Natural Protein Sequences (Krieger and Kiefhaber, to be submitted).....	147
8. CURRICULUM VITAE.....	175

1 Introduction

1.1 Proteins as the major functional biomacromolecules

Among biological macromolecules proteins are the major functional molecules of life with diverse properties in structure and function. They aid and control chemical reactions that make the cell work. They receive and transmit signals from outside of the cell. They control the processes by which proteins are made from the instructions in the genes. They form the scaffolding that gives cells their shape, and they are parts of the linkages that stick cells together into tissues and organs. Protein folding is the process by which a protein assumes its functional conformation. The manner how a newly synthesized polypeptide chain adopts the native folded structure depends on the intrinsic properties of the linear amino acid sequence and on the contributing influences from the surrounding milieu. Folding and misfolding are critical ways of regulating biological activity, targeting proteins to different cellular locations and causing neurodegenerative diseases, like prion diseases or Alzheimer disease. Folding is a spontaneous process that is mainly guided by non-covalent interactions and entropic contributions to the Gibbs free energy. The molecular origins of folding and misfolding processes are not yet understood. A main goal in protein science is to understand the mechanism how a polypeptide chain reaches the native, stable state during a folding reaction and how proteins fulfil their biological functions.

1.2 The protein folding problem

Starting in the 1930's, first experiments on protein stability and folding revealed that the denaturation of proteins is a reversible process^{1;2}. It was recognized that the native state of a protein is characterized by a defined configuration, whereas the denatured state is characterized by the absence of a defined configuration³. In this time the investigation of

protein folding reactions resembled an unsolvable puzzle, since both, the structures of the native and the unfolded states were not known. Later, with the determination of the first protein structures due to X-ray crystallography^{4; 5} investigation of the mechanism of protein folding become more successful.

Studies on the renaturation of fully denatured and reduced ribonuclease showed that the amino acid sequence contains all information needed to form the native state with all correct disulfide bonds in a given environment^{6; 7}. The results showed that the sequence determined the native structure, which was interpreted in terms of a thermodynamic hypothesis, that an unfolded polypeptide chain folds under a driving force and reaches thermodynamically the most stable conformation⁷. However the results could not explain which forces govern the folding process and how a chain finds the native state. At the same time Levinthal pointed out that a folded protein chain cannot sample all of its conformations randomly to find the native state on an appropriate time scale using simple consideration on the statistics of unstructured chain molecules^{8; 9}. He argued that a protein folding reaction cannot be a fully entropic driven process, but the folding process must be directed in some way by free energy barriers to restrict conformational space.

Physico-chemical and physical investigations on protein folding mechanisms deal with the description of all states, barriers and pathways on the free energy landscape between the unfolded and native states. Folding pathways are usually presented on reaction coordinates from the unfolded state over a separating barrier to the native state in analogy to simple reactions in chemistry. In the simplest case the folding reaction can be described with a two-state model and a single energy barrier, although the unfolded state comprises a large ensemble of different conformational states¹⁰. Theoretical considerations reveal that this approximation is valid if the interconversion of different

unfolded conformations is fast compared to the barrier crossing process¹¹, which indeed was shown by applying fast spectroscopic techniques to the dynamics of unstructured peptides¹².

X-ray diffraction and NMR structure determination allow the characterization of the native states of many proteins. Most proteins adopt a well-defined, biologically active structure under physiological conditions¹³. Conformational fluctuations and dynamics in the native state of proteins and enzymes could be detected and it is believed that such dynamics are a key event in biological function^{14; 15; 16}. In contrast, the unfolded state is disordered and represents an ensemble of different conformations, which are under physiological conditions less populated. Recently also methods have become available to characterize the structure and dynamics of unfolded states as initial point of protein refolding reactions^{12; 17; 18; 19; 20; 21}. NMR spectroscopy revealed that even under high denaturing conditions local residual structure is preserved^{17; 19; 21; 22}, contradicting that unfolded conformations can be regarded as an unstructured polymer, where the polypeptide chain behaves like a random coil chain²³. As shown by fast spectroscopic techniques the formation of first contacts on an unfolded polypeptide chain occurs in nanosecond time scale^{12; 18; 24}. With the formation of energetically favourable contacts the folding polypeptide chain becomes more and more structured and finally reaches the native state.

Several models were developed to explain the phenomenon of protein folding from the random coil state to the native structure. The “framework model” describes a step-wise mechanism to reduce the conformational search. This involves a hierarchical assembly whereby local elements of secondary structures are formed according to the primary sequence, but independently from tertiary structure. These elements then diffuse until they collide, whereupon they combine to form the tertiary structure^{25; 26; 27}.

The “hydrophobic collapse model” assumes that the native protein conformation is formed by rearrangement of a compact collapsed structure²⁸. Hydrophobic collapse to form a molten globule therefore represents an early step in the folding pathway.

The “nucleation model” suggests that tertiary structure forms as an immediate consequence of the formation of secondary structure^{29; 30}. Nucleation occurs through the formation of native secondary structure by only a few residues (e.g. a beta-turn, or the first turn of an alpha-helix), and structure propagates out cooperatively from this nucleus. All three models have common that the earliest steps of structure formation are determined by intrachain diffusion, which can therefore be regarded as elementary step in protein folding (Figure 1).

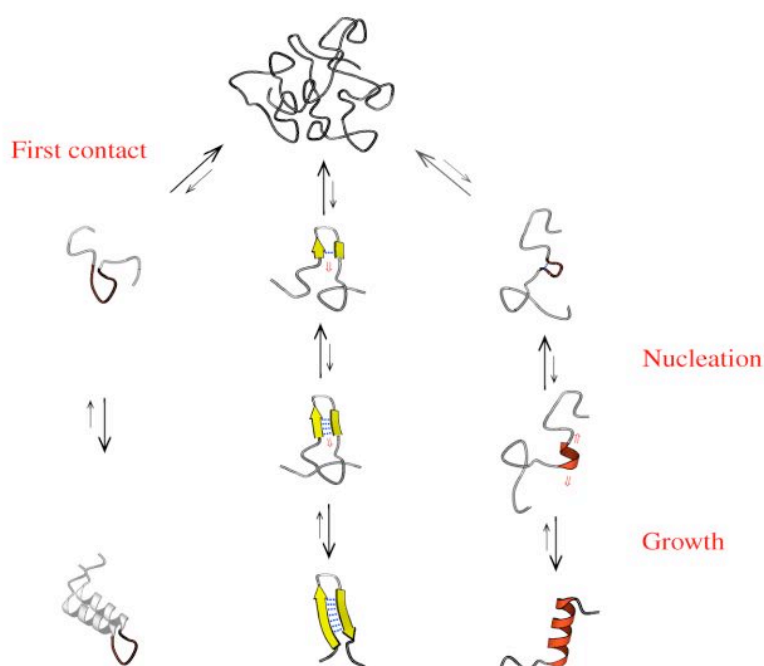


Figure 1: Intrachain diffusion as elementary step of structure formation in peptides and proteins (from ref.³¹).

Since the dynamics of unfolded polypeptides chains govern the formation of the first structural organisation during protein folding, the characterization of the dynamics of unfolded states will give insight into the earliest steps of protein folding and allows to

answer questions (i) how fast are intramolecular contacts formed on an unfolded polypeptide chain, (ii) what determines intrachain diffusion in polypeptides and (iii) how does intrachain diffusion contribute to the overall folding process?

1.3 Structural heterogeneity in the unfolded state of proteins

First knowledge about the properties of the unfolded state of proteins was theoretically derived using statistical models of chain molecules³². The dimensions of an ideal or freely jointed chain is described by the mean-square end-to-end distance $\langle r^2 \rangle$,

$$\langle r^2 \rangle = nl^2 \quad (1.1)$$

with n as the number of bonds and l as the bond lengths³². The mean-square end-to-end distance $\langle r^2 \rangle$ of an ideal chain scales linearly with the number of segments n . However, in real chains bond and torsion angles are strongly restricted. Flory introduced the characteristic ratio C_n , which relates the mean-square end-to-end distance of an unperturbed real chain compared to the mean-square end-to-end distance of a freely jointed chain:

$$C_n = \frac{\langle r^2 \rangle_0}{nl^2}. \quad (1.2)$$

The index 0 refers to the dimension of an unperturbed polymer chain. Due to restricted bond angles and steric hindrance the mean-square end-to-end distance becomes larger compared to a freely jointed chain. The characteristic ratio C_n converges to a finite value for long polymer chains. For long polypeptide chains without glycine or proline residues a characteristic ratio C_n of around 9 was calculated theoretically^{32; 33; 34; 35; 36} and determined experimentally^{23; 33; 37}. The obtained results suggest that unstructured polypeptide chains longer than 50 residues behave like random coils and that the dimension of shorter peptide chains is significantly influenced by chain stiffness³².

Experimental studies on the unfolded state of proteins revealed that the dimension of most urea- and guanidine hydrochloride denatured proteins apparently obey the theoretically expected random-coil scaling for real chains with excluded volume^{23; 38; 39}. These results suggested, that proteins are unstructured at high denaturant concentrations and behave like random coil polymers. The observation of random-coil dimensional scaling contradicts however recent spectroscopic work, which revealed that many denatured proteins contain significant residual structure^{17; 19; 20}. Furthermore exhaustive enumeration of all sterically accessible conformations of short hepta-alanyl peptide chains revealed that steric repulsion of neighbour residues restrict the conformational space contradicting Flory's hypothesis that each residue on a polymer chain can be regarded as isolated⁴⁰. These results are expected since natural polypeptide chains are hetero-polymers consisting of diverse residues with a variety of side chains of quite different chemical reactivities, leading to heterogeneous local structures in aqueous solutions, even in presence of denaturants. Therefore unfolded polypeptide chains might have locally definite backbone conformations and the concept of a denatured protein as a structureless random chain breaks down, when local interactions of individual residues are described, although the random chain concept may still be useful to describe the overall chain conformation⁴¹.

1.4 Elementary steps in protein folding

Starting from an ensemble of denatured chain conformations a folding polypeptide chain has to explore the vast conformational space in search for energetically favourable conformations. The molecular structures of native proteins consist of secondary structures such as α -helices, β -sheets, loops and turns, which represent the basic structural elements. To form secondary structure elements a first intramolecular contact

on the linear polypeptide chain is essential (Figure 1). The rate at which two points on a folding polypeptide chain make contact is limited by intrachain diffusion. Intrachain diffusion as part of the formation of secondary structure represents therefore an elementary step in the formation of the overall molecular structure⁴². Until recently mainly fast relaxation techniques were applied to study α -helix^{43; 44; 45} or β -hairpin⁴⁶ formation and to get knowledge about the time scales of the earliest steps in protein folding.

The α -helix formed by polypeptide chains is one of the best studied secondary structured elements in biology. Since the first structural description of the formation of polypeptide chains into helical conformations⁴⁷ the structure and stability of α -helices were well studied⁴⁸. Additionally a number of theoretical models were developed to successfully describe the helix-coil transition^{49; 50; 51}. Helix formation is generally regarded as a sequential process with two elementary reactions, a slow nucleation step and fast propagation steps. Nucleation is believed to represent the formation of the first helical turn, i.e. the interactions of four consecutive chain monomers with the right backbone hydrogen bonds and the right dihedral angles. After formation of the first helical turn the helix propagation follows. Using fast relaxation techniques, such as dielectric relaxation^{43; 51}, ultrasonic absorption⁴⁴ or nanosecond temperature laser jump⁴⁵, the determination of rate constants for the propagation process gave a better insight into the dynamics of the helix-coil transition. However, reliable rate constants for helix nucleation are not available.

Due to the tendency to aggregate there are less data available about the mechanism of β -hairpin folding of isolated β -hairpin peptides. β -Hairpins consist of two antiparallel β -strands, which are connected by a turn or a loop. The first hairpin-forming peptide to be characterized in equilibrium and kinetic experiments was the C-terminal fragment of the

B1 domain of protein G (residues 41-56). This peptide forms a stable β -hairpin structure in solution^{52; 53}. Kinetic experiments suggested an upper limit in the microsecond range for β -hairpin formation⁴⁶. Using a statistical mechanical model to interpret β -hairpin folding Munoz et al. suggested that the most probable way to initiate β -hairpin folding is to form the β -turn region⁵⁴.

1.5 Intrachain loop formation

Diffusion of free molecules in solution has been extensively investigated for 100 years^{55; 56; 57}. However, little is known about absolute diffusion rates of two defined points on a polymer chain. First approaches were based on the calculation of the probability of end-to-end cyclization in ideal polymer chains (Figure 2)^{32; 58}, later for polymer chain with hindered rotation⁵⁹.

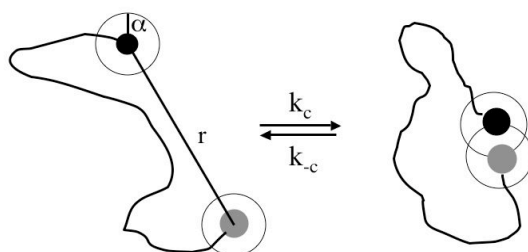


Figure 2: Schematic view of end-to-end cyclization in a polymer chain, with α as reactive boundary and r as the end-to-end distance. To react both ends have to make contact within the boundary α . k_c and k_{-c} are the rate constants of the loop closure reaction and back reaction, respectively.

Jacobson and Stockmayer calculated the end-to-end ring closure probability as function of the monomer number in a freely jointed chain⁵⁸. As result they obtained for an ideal chain of sufficient length that the probability of ring closure reaction p_{ring} scales with

$$P_{ring} \propto \langle r^2 \rangle_0^{-3/2} \propto n^{-3/2} \quad (1.3)$$

where $\langle r^2 \rangle_0$ is the mean-square end-to-end distance averaged over all configurations of an ideal chain with n monomers. Equation 1.3 fails if polymer chains are short or stiff and become non-ideal³². Similar results could be obtained for chains with hindered rotation⁵⁹. Application of the elaborated form of Jacobsen-Stockmayer theory to the determination of cyclization constants of L-polypeptides (poly(L-Ala)) revealed that especially flexible glycine and D-amino acids simplify ring closure reactions⁶⁰.

Later Szabo, Schulten and Schulten applied the first passage time approach to investigate the averaged reaction time for intramolecular diffusion reaction in a flexible ideal chain^{61; 62}. The loop closure reaction was treated as an intramolecular diffusion process under the influence of a potential. It could be shown that the kinetics of such a loop closure reaction can be approximated by a single exponential decay of the reaction probability $\Sigma(t)$ that the system is still unreacted at time t :

$$\Sigma(t) \approx \Sigma_{approx}(t) = \exp\left(\frac{-t}{\tau}\right) \quad (1.4).$$

The approximation is applicable to determine end-to-end contact formation rate constants in polypeptides if the averaged end-to-end distribution of the chain is equilibrated. The local equilibrium condition can be satisfied⁶² if

- (i) the interconversion between chain conformations is fast compared to the loop closure reaction,
- (ii) only a small fraction of conformations make contact at equilibrium conditions, i.e. the distance of the van-der-Waals contact radius between the two ends is small compared to the averaged end-to-end distance, that is, if $\alpha/r \ll 1$ (Figure 2).

For a small reactive boundary α the rate constant k_c for loop closure reaction of a Gaussian polymer with an equilibrated end-to-end distribution is approximately given by

$$\tau^{-1} = k_c \propto \frac{D_0}{\langle r^2 \rangle^{3/2}} \times \alpha \quad (1.5).$$

The rate constant of loop closure reactions of ideal polymers depends on⁶²:

- (i) the chain length independent diffusion coefficient D_0 , which is therefore proportional to T/η ,
- (ii) the averaged end-to-end distance of the polymer,
- (iii) the equilibrium distribution of the end-to-end distance,
- (iv) the reactive boundary α of reacting ends of the polymer.

Accounting for excluded volume and solvent effects, the rate constant of end-to-end contact formation k_c for real polymer chains of sufficient length depends on the monomer number n as

$$k_c \propto n^{-m} \text{ with } m = 1 - 1.8 \quad (1.6).$$

The exponent m depends strongly on excluded volume and solvent effects and a value of 1.5 was calculated for an unperturbed or θ -chain. For good solvents it is expected that the exponent adopts the value of 1.8 and in poor solvents the value of around one^{32; 61; 63}.

1.6 Measuring intrachain contact formation in unfolded peptide chains

Intramolecular reactions on polymer chains in various systems have been studied experimentally, for a long time. Early studies used fluorescence resonance energy transfer (FRET)^{24; 64}, excimer formation⁶⁵, or triplet-triplet annihilation⁶⁶ to investigate the dynamics of intramolecular reactions on polypeptide chains or synthetic polymers. Using FRET Haas et al. determined the diffusion coefficients for short peptide chains

describing the Brownian motion of the molecular ends of the peptide chains^{24; 64}. These studies gave a first insight on the mechanism of diffusional motion of two defined points on a peptide chain, suggesting that internal friction impedes intrachain diffusion in short peptides. However the determination of absolute contact formation rate constants for a specific chemical contact between two points was not possible with this method.

Later fast spectroscopic methods such as triplet-triplet energy transfer (TTET)¹², triplet^{18; 67} or fluorescence^{68; 69} quenching were used to determine end-to-end formation rate constants in unfolded polypeptide chains and proteins. Applying nanosecond-resolved laser flash spectroscopy Hagen et al. measured intrachain contact formation in GdnHCl unfolded cytochrome C. The bond formation reaction between a methionine residue and the heme group separated by around 50 residues⁷⁰ gave time constants around 35-40 μ s. The observed time constants and polymer theoretical approaches^{58; 61} were used to estimate the speed limit of protein folding reactions with a time constant $\tau = 1 \mu$ s for contact formation between ten peptide bonds in an unfolded protein. Later Bieri et al. introduced the method of TTET into the field of protein folding¹². Initially a derivative of thioxanthone (Txan) as triplet donor and the non-natural amino acid naphthylalanine (NAIa) as triplet acceptor were attached at the ends of the peptide chains (Figure 3).

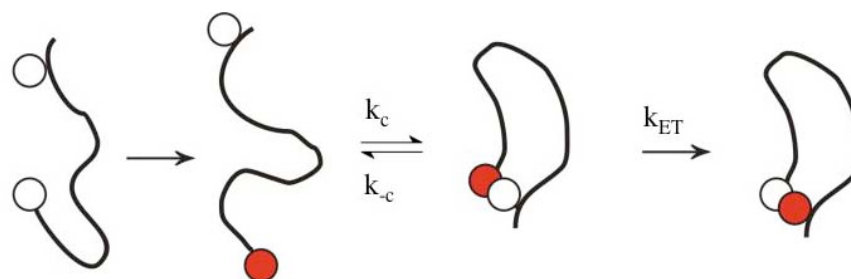


Figure 3: Principal scheme to determine absolute intrachain contact formation rate constants by TTET. The spectroscopically labeled ends of a molecule chain diffuse and collide. Due to non-covalent interactions a signal is changed by a photochemical reaction. Absolute rate constants are determined, if $k_c \gg k_c$ and $k_{ET} \gg k_c, k_c$ (see also chapter 1.7).

The measurements were carried out in ethanol, since under these conditions TTET from thioxanthone to naphthylalanine becomes diffusion-controlled. For short (GS)-peptides end-to-end contact formation could be well described by single-exponential kinetics, as predicted by theoretical investigations^{11; 61; 62}. A minimum time constant for end-to-end contact formation could be determined with $\tau_{\text{obs}} = 20$ ns for the shortest peptide with three peptide bonds between the triplet labels. TTET through bond electron transfer can be ruled out, since this process cannot occur over distances beyond eight bonds⁷¹. In the shortest glycine-serine peptide used in TTET studies the triplet labels were separated by 11 bonds. With increasing loop length up to eleven peptide bonds the end-to-end contact formation rate constant decreases monotonously, contradicting theoretical approaches. The time constant of contact formation at a distance of ten peptide bonds was predicted to be the distance for the maximum rate of contact formation due to chain stiffness slowing down intrachain diffusion over shorter distances⁷². The triplet donor thioxanthone, initially used in TTET experiments, was later replaced by xanthone, allowing measurements in water⁷³. First characterization of end-to-end contact formation in a (GS)₄-peptide in water using the xanthone/NAIa system revealed 3-4 times higher rate constants compared to the thioxanthone/NAIa system. The slower kinetics in ethanol could be attributed to solvent effects⁷³.

In the last few years several different systems have been used to study intrachain contact formation for unfolded peptides and protein fragments of various lengths. Several other spectroscopic techniques, such as tryptophane triplet state quenching by cysteine¹⁸, dye-fluorescence quenching by tryptophane^{68; 69} and electron transfer from a triplet excited Zn-porphyrine group to a Ru-complex⁶⁷ were applied to determine end-to-end contact formation rate constants in synthetic polypeptides and unfolded proteins. Generally it is

found that the time constant of intrachain contact formation over ten residues will be faster than 100 ns, although the time constants for intrachain contact formation have been found to vary significantly depending on the applied method, on the peptide sequence and on solvent conditions. The direct comparison of all data is therefore quite complicated. The crucial point in determining absolute rate constants for intrachain contact formation is the use of electron donor and acceptor pairs, which exchange electrons in a diffusion-controlled reaction.

Lapidus et al. used tryptophane triplet quenching by cysteine to determine end-to-end contact formation rate constants in synthetic peptides¹⁸. Tryptophane as triplet donor has the advantage that this amino acid can be easily incorporated in peptide and protein polypeptide sequences. However, the slow formation of tryptophane triplets ($\tau=3$ ns)⁷⁴ and the inefficient triplet quenching by natural amino acids such as cysteine⁷⁴ limit the determination of absolute contact formation rate constants in peptides, resulting in slower kinetics for end-to-end contact formation in peptides¹⁸ compared to results using TTET from xanthone to NAla in similar peptides³¹.

Other approaches applied fluorescence quenching techniques to determine intrachain contact formation rate constants. The fluorescence of 2,3-Diazabicyclo[2,2,2]oct-2-ene (DBO) is long-lived ($\tau \approx 1\mu\text{s}$) and is quenched by tryptophane close to the diffusion-controlled limit^{68; 75}. However, obtained rate constants for intrachain contact formation in (GS)-peptides are three times slower than the ones obtained with the xanthone/NAla system in the same peptides, suggesting that electron transfer in the chosen chromophore pair is not fully diffusion-controlled. The fluorescence dye (MR121) exhibits also long-lived fluorescence and can also be quenched by tryptophane⁶⁸. Although the electron transfer occurs only via van-der-Waals contact the quenching mechanism of this electron transfer is complex. The calculated rate constants for

intrachain contact formation are, however, significantly lower than those determined by other methods.

Recently a new method was presented to measure intrachain diffusion in partly unfolded cytochrome C using electron transfer from a triplet excited Zn-porphyrine group to a Ru-complex, which was bound selectively on a histidine residue⁶⁷. The determined rate constants agree well with ones obtained from TTET measurements if rate constants are corrected against solvent effects³¹.

The results revealed that only methods, which include very fast electron transfer reactions on the sub-nanosecond time scale, satisfy the requirements to determine absolute intrachain contact formation rate constants.

1.7 Triplet-Triplet Energy Transfer

Triplet-triplet energy transfer (TTET) is defined as energy transfer from an electronically excited triplet donor that produces an electronically excited acceptor in its triplet state⁷⁶. TTET occurs via collisions between the reacting partners⁷⁷. Collision means that the electron clouds of the reacting species overlap significantly in space. For ideal spectroscopic labels without steric hindrance the rate constant for energy transfer by electron exchange is expected to fall off exponentially as the distance of triplet donor and acceptor increases. Dexter proposed that energy transfer occurs via two electron-transfer exchanges and that the rate constant of TTET is expressed as

$$k_{TTET} = A * e^{-\frac{2R_{DA}}{L}} \quad (1.7)$$

where A is the pre-exponential factor and R_{DA} the donor-acceptor distance relative to their *van-der-Waals* radii (L) which is smaller than 1.5 nm for TTET reactions^{76; 78}. The rate constant of TTET within an encounter complex depends on the energetics of the triplet pair. Figure 4 shows a schematic view of the relative energetics of triplet donor

and triplet acceptor (Jablonski diagram). The observed rate constant of TTET depends at least on three chemical processes, the formation of the S_1 state of the triplet donor, the intersystem crossing rate constant k_{ISC} from S_1 to T_1 and the TTET rate constant k_{TTET} itself, which depends on the energetics of the triplet pairs. Triplet states show characteristic absorption spectra and can be monitored transiently due to a relatively long lifetime up to the microsecond time scale.

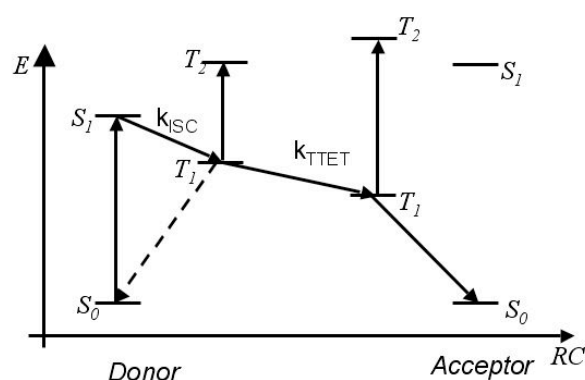


Figure 4: Jablonski diagram with triplet-triplet energy transfer. S_0 , S_1 , T_1 and T_2 are ground state, first excited singlet state, first excited triplet state and second excited triplet state, respectively. k_{ISC} and k_{TTET} are rate constants for intersystem crossing and triplet-triplet energy transfer, respectively. RC: reaction coordinate.

Generally it is found that energy transfer processes, which are endothermic by more than several kJ/mol are inefficient, even if they are spin-allowed ($\Delta E_{DA} < 0$). Spin-allowed, exothermic energy-transfer processes are generally efficient ($\Delta E_{DA} > 0$)⁷⁶. For several triplet pairs it could be shown that TTET is diffusion-controlled under appropriate conditions⁷⁹.

Intermolecular TTET in solutions can be described in a general case like in figure 5. An excited triplet donor (D^*) diffuses through the solution until it encounters a triplet acceptor molecule and both form an encounter complex (D^*A). The collision between

the energy donor D^* and A may lead to TTET and generation of a new encounter complex (DA^*). The encounter complex (DA^*) then disperses and the excited acceptor relaxes to the ground state (A). The observed TTET transfer rate constant in solution additionally depends on the diffusion of the triplet chromophores. It is empirically found that if $E(D^*) > E(A^*)$ TTET is faster than the diffusion process and the rate constant becomes generally close to the rate constant for diffusion-controlled bimolecular reactions.

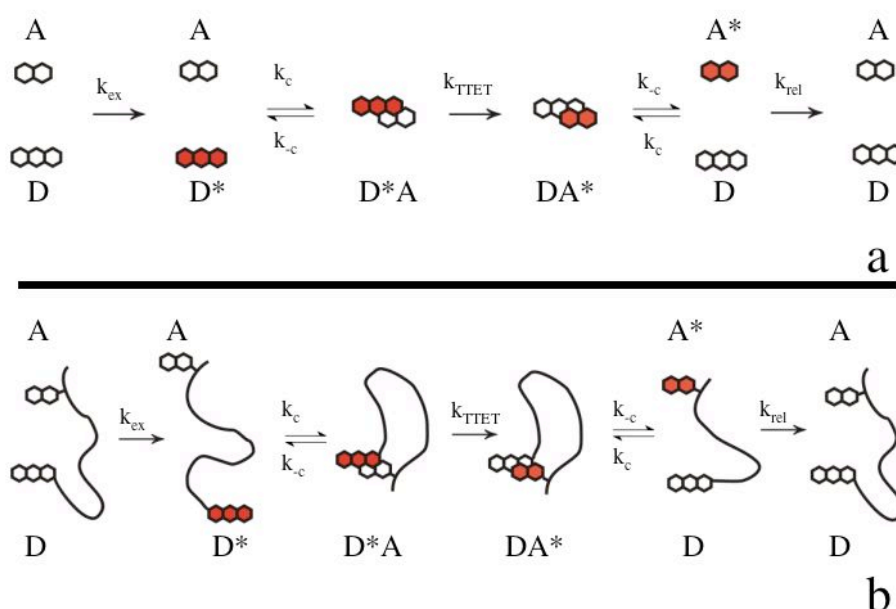


Figure 5: Schematic picture of the formation of encounter complexes of a triplet donor D and a triplet acceptor A ⁷⁶. The starlet and red colour indicate excited states. k_{ex} , k_{rel} are excitation and relaxation rate constants, respectively. k_c , k_c and k_{TTET} are the rate constants for contact formation and decomposition of the encounter complex and TTET, respectively.

The irreversible TTET, formation and decay of the encounter complex, can be described as a three-state model (Figure 5). The observed rate constants k_{obs} for TTET are functions of all microscopic rate constants k_c , k_c and k_{TTET} with k_{TTET} as the absolute rate constant of triplet electron transfer, k_c and k_c as the diffusion rate constant and the

decomposition rate constant of the encounter complex, respectively and can be

expressed¹⁴ as

$$k_{obs_{1,2}} = \frac{p \pm \sqrt{p^2 - 4q}}{2}$$

$$p = k_c + k_{-c} + k_{TTET}$$

$$q = k_c \cdot k_{TTET} \quad (1.8).$$

TTET becomes efficient if $k_{-c} \gg k_c$ and $k_{TTET} \gg k_c, k_{-c}$ and the observed rate constant simplifies to:

$$k_{obs} = k_c \quad (1.9).$$

Applying the Stokes-Einstein relation it is expected that the observed rate constants are directly proportional to the ratio T/η , with T as the absolute temperature and η as the solvent viscosity⁷⁶. The observed rate constant for intramolecular quenching is then fully influenced by the solvent motion.

If energy transfer is inefficient ($k_{TTET} \ll k_{-c}$) the observed rate constant is expressed as

$$k_{obs} = \frac{k_c}{k_{-c}} k_{TTET} \quad (1.10).$$

The rate constant is now a composite of the “pseudo”-equilibrium for the encounter complex formation and the rate constant of energy transfer.

2 Aims of Research

The rate how fast intramolecular contacts are formed during a protein folding reaction is limited by intrachain diffusion, which can be regarded as the elementary process of structure formation. The crucial point in the spectroscopic determination of absolute rate constants for intrachain contact formation is the application of electron donor and acceptor pairs, which exchange electrons in a diffusion-controlled reaction. We want to verify that triplet-triplet energy transfer (TTET) of an appropriate triplet pair is diffusion-controlled in aqueous solutions and provides the requirements to measure absolute rate constants for end-to-end contact formation in polypeptides and proteins. First, the photophysical processes of TTET from xanthone to naphthalene were investigated by temperature and viscosity dependencies. Further we wanted to study the influences of various amino acids on the triplet lifetime of xanthone to be able to test for possible limitations of the method in the determination of chain dynamics of natural protein sequences.

We used homo-polypeptide chains to study the loop-length dependence on intrachain contact formation and to establish scaling laws. In host-guest experiments we determined the effect of the amino acid sequence on intrachain contact formation in peptides.

The extent and magnitude of intrachain dynamics is limited by the molecular connectivity and excluded volume of the chain. To know how chain and solvent properties influence the dynamics in polypeptides we determined the effect of temperature, viscosity and solvent quality on intrachain contact formation.

Natural protein sequences consist of diverse amino acids with a variety of side-chains of quite different chemical reactivities. To test if experimental results obtained from the

studies on model homo-polypeptides can be generally applied to dynamics of protein polypeptide chains we characterized the chain dynamics of various natural sequences.

The diffusion-collision model²⁵ assumes that the formation of α -helices or β -sheets is rapid relative to the rate-limiting steps in protein folding. The formed secondary structures have to interact by chain diffusion processes. We want to know, how fast folded secondary structures, which are separated by an unstructured loop, collide on a polypeptide chain.

3. Results

3.1 The xanthone/naphthalene system as triplet donor/acceptor pair for studying intrachain contact formation in polypeptides

3.1.1 Tests for diffusion-controlled triplet-triplet energy transfer from xanthone to naphthalene

Triplet-triplet energy transfer (TTET) between a triplet donor and acceptor has recently been applied to determine end-to-end contact formation rate constants in peptide chains to gain information on the dynamics of the earliest steps in protein folding^{12; 73}. In these experiments derivatives of xanthone and naphthalene were chosen as triplet donor and acceptor, respectively. To test if TTET from xanthone acid (Xan) to naphthylalanine (NAla) in aqueous solution is diffusion-controlled (Figure 6A) we measured the effect of solvent viscosity and temperature on the bimolecular TTET reaction⁷⁶.

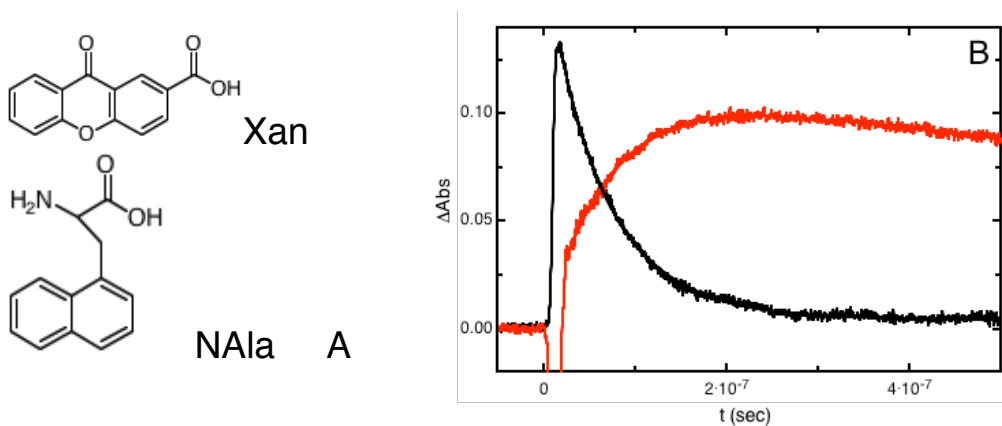


Figure 6: (A) Molecular structure of xanthone acid (Xan) and naphthylalanine (NAla). (B): Time course of formation and decay of Xan (—) triplets at 590 nm and NAla (—) triplets at 420 nm in a bimolecular TTET reaction. Xanthone triplets are produced by a short laser pulse at 355 nm and transferred to NAla with an observed rate constant of $1.4 \cdot 10^7 \text{ s}^{-1}$. NAla triplets decay with a time constant of $1 \mu\text{s}$ because of oxygen quenching. The negative amplitude during the laser flash at 420 nm (—) is due to xanthone fluorescence⁸⁰. The concentrations of Xan and NAla were $50 \mu\text{M}$ and 5 mM , respectively.

The formation and decay of xanthone triplets in aqueous solution can be monitored by transient absorption technique at 590 nm⁸¹ (Figure 6B). In presence of naphthalene acceptor molecules xanthone triplet electrons are transferred via two-electron transfer mechanism⁷⁷ to form naphthalene triplets, which have a characteristic absorption around 420 nm⁸² (Figure 6B). Xanthone triplets are long-lived and display lifetimes in water in the microsecond time-range. The observed bimolecular TTET rate constant scales proportionally to $1/\eta$ at constant temperature independent of chosen viscous co-solutes (Figure 7A) and is proportional to T/η (Figure 7B) as determined with the temperature dependence. The results show that TTET from xanthone acid to naphthylalanine in aqueous solution is fully diffusion-controlled. The results indicate that the chosen triplet pair provides the requirements to measure dynamics of molecular processes, which are governed by the solvent motion.

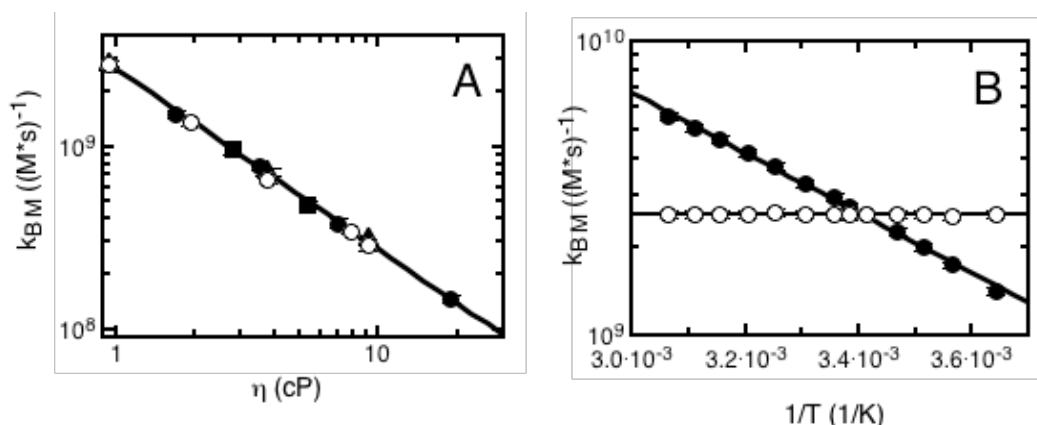


Figure 7: (A). Effect of solvent viscosity η on the bimolecular TTET rate constant k_{BM} from Xan to NAla in aqueous solutions. Following viscous co-solutes were used: (●) ethylene glycol, (○) glycerol, (■) glucose, and (▲) sucrose. The rate constants scale with $1/\eta$ (solid lines). (B) Temperature dependence of the bimolecular TTET rate constant k_{BM} from Xan to NAla in water, (C). The black dots (●) represent observed rate constants, the white dots (○) are rate constants corrected against viscosity and temperature effects on the reaction according the equation $k_{corr} = \frac{T_0 \cdot \eta}{T \cdot \eta_0} \cdot k_{obs}$ ⁷⁶. The uncorrected data can be analyzed by a Vogel-Fulcher type of equation, which describes the temperature dependence of the viscosity of water (—)⁸³. After the correction the rate constants become virtually independent of temperature (—). The viscosity dependence was carried out at 22.5°C.

A diffusion-controlled bimolecular reaction is fully governed by the diffusion coefficient⁷⁸. Therefore it is expected that the bimolecular TTET rate constants decrease, if the molecule dimension increases. Indeed, it could be shown that the observed bimolecular TTET rate constants are in the range between $(1-7) \cdot 10^9 \text{ (Ms)}^{-1}$ as expected for diffusion-controlled bimolecular reactions⁷⁶ and decrease if the dimensions of triplet donor and acceptor molecules are increased (Table 1). The change in the molecule dimension might explain the difference in observed bimolecular TTET rate constants of different naphthalene derivatives.

Table 1: Rate constants of bimolecular TTET (k_{BM}) from xanthone to naphthalene derivatives of different size. All measurements were carried out in water, pH 7 at 22.5°C.

Triplet donor	MW	Triplet acceptor	MW	$k_{BM}(\text{Ms}^{-1})$	Cond.
Xanthone	196	1-Methylnaphthalene	142	$6.7 \cdot 10^9$	EtOH, rt ^a
Xanthone acid	240	Naphthyl acetic acid	186	$4.0 \cdot 10^9$	water, rt ^b
Xanthone acid	240	Naphthylalanine (NAla)	216	$2.8 \cdot 10^9$	water, rt
Xan(S) ₃ ASG	716	Ac(S) ₄ NAlaSG	748	$1.4 \cdot 10^9$	water, rt

a: adapted from ref.⁸⁴,

b: adapted from ref.⁷³

The above-described experiments give a first insight into the time scales of xanthone triplet formation and TTET from xanthone to naphthalene derivatives in aqueous solutions, which are suggested to be faster than the bimolecular diffusion of small organic molecules. However with these experiments it is not possible to determine absolute rate constants for photophysical processes, such as intersystem crossing reaction or TTET. The rate constant of xanthone triplet formation in aqueous solutions can be estimated from investigations on xanthone fluorescence lifetimes, which revealed to be in the sub-nanosecond time range⁸⁰. To directly determine the absolute triplet-triplet energy transfer rate constants from xanthone to a naphthalene derivative, we measured transient absorption kinetics of TTET from xanthone in liquid, neat 1-methyl-

naphthalene similar to experiments with aromatic ketones⁷⁹. As shown in figures 8A, B, xanthone triplets are transferred to 1-methyl-naphthalene during the laser flash within 10 ns. No triplet bands of xanthone could be detected, which would normally appear around 600 nm (Figure 6A). Instead a strong transient absorption band appears at 420 nm, which is typical for naphthalene triplets⁸². The result suggests that the time constants for xanthone triplet formation and TTET from Xan to NAla are shorter than 10 ns, which is well in agreement with TTET rate constants of other aromatic ketones in 1-methyl-naphthalene⁷⁹.

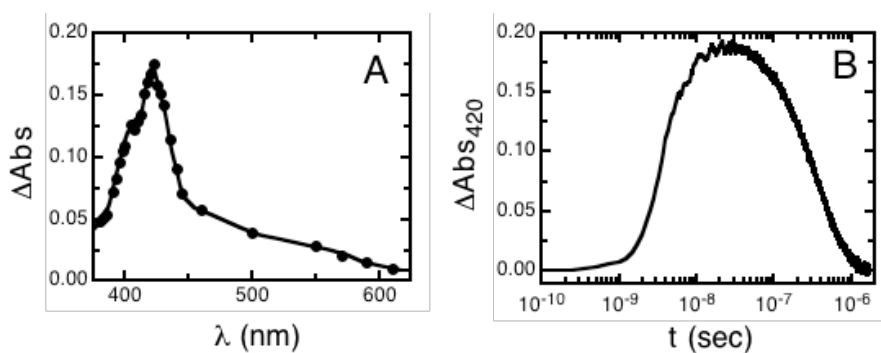


Figure 8: TTET from xanthone to 1-methyl-naphthalene. (A) Transient absorption spectrum, 10 ns after the laser flash at 355 nm. At this wavelength 1-methyl-naphthalene is not excited and triplet electrons must be transferred from xanthone to 1-methyl-naphthalene. (B) Time course of formation and decay of 1-methyl-naphthalene triplets after a 5 ns laser flash.

Because of the long laser pulse of 5 ns it was not possible to directly measure rate constants of xanthone triplet formation and of TTET from xanthone to 1-methyl-naphthalene.

In collaboration with the group of Prof. W. Zinth (LMU München) it could be shown that xanthone triplet formation and TTET from xanthone to naphthalene occur in the sub-picosecond time range using femtosecond laser flash spectroscopy (H. Satzger, B. Fierz, personal communication)⁸⁴. These results verify that the photochemical and photophysical processes of TTET from xanthone to 1-methyl-naphthalene are

significantly faster than end-to-end contact formation in short peptides, which occur in the nanosecond time scale^{31; 73}.

Viscosity and temperature dependencies and femtosecond laser flash spectroscopic measurements revealed, that TTET rate constants from xanthone to naphthalene derivatives are proportional to the ratio T/η and the photophysical processes occur on the picosecond time scale. The triplet pair xanthone/naphthalene provides the requirements to determine absolute rate constants of inter- and intramolecular diffusion processes.

Parts of these results are described in detail in the following manuscript:

“Ultrafast Quenching of the Xanthone by Triplet Energy Transfer: New Insight into the Intersystem Crossing Kinetics.” H. Satzger, B. Schmidt, C. Root, W. Zinth, B. Fierz, F. Krieger, T. Kiefhaber & P. Gilch (2004) *J. Phys. Chem. A*, **108**, 10072-10079.

3.1.2 Influence of natural amino acids on xanthone triplet lifetime

Protein polypeptide sequences consist of twenty different amino acids with quite distinct reactivities. To measure rate constants of intrachain contact formation in polypeptide chains we first determined the possible influence of amino acid side chains on xanthone triplet lifetime⁸⁵. It could be shown, that the aromatic amino acids tryptophane and tyrosine efficiently quench the xanthone triplet state with rate constants near the diffusion-controlled reaction rate limit. However, TTET from xanthone to tryptophane and tyrosine leads to complex reactions, with at least two observable rate constants. For both amino acids TTET is accompanied by radical formation of tryptophane and tyrosine triplet states^{74; 86}. Methionine and deprotonated histidine quench xanthone triplets partly diffusion-controlled. All other amino acid side-chains quench xanthone

triplets very inefficiently (Cys, His⁺, N-terminus) or not at all (Ala, Arg, Asn, Asp, Gly, Lys, Ser, Phe). Histidine containing sequences can be used if the pH of the aqueous solution is below 5.5. Below pH 3 generally no TTET experiments are possible since xanthone triplets are quenched due to proton abstraction like for triplet states of other aromatic ketones⁸⁷.

The results revealed that TTET from xanthone to naphthylalanine can be applied to investigate intrachain contact formation in polypeptide sequences, if they are free of tryptophane, tyrosine, methionine and histidine.

These results are described in detail in the following publication:

“Intrachain diffusion in a protein loop fragment from carp parvalbumin.” Florian Krieger, Beat Fierz, Fabian Axthelm, Karin Joder, Dominique Meyer and Thomas Kiefhaber (2004) *Chem. Phys.* **307**, 209-215.

3.2 Effect of loop length on end-to-end contact formation in flexible polypeptides

To reach the native three-dimensional structure a folding polypeptide chain has to form specific intramolecular interactions. The rate, at which two points on an unstructured polypeptide chain make contact is limited by intrachain diffusion. Intrachain diffusion is particularly important during protein folding, which requires formation of a large number of specific long-range and short-range interactions. Characterization of the rates of intrachain contact formation is therefore essential for the understanding of the earliest steps in protein folding. Natural polypeptide chains represent complex hetero-polymers due to the incorporation of twenty different amino acids of diverse chemical properties. The chemical heterogeneity of protein polypeptide chains leads to local changes in chain properties⁸⁸. The usage of homo-polypeptide chains such as alternating glycine-serine

repeats (Figure 9) has the advantage of being able to establish scaling laws⁸⁹ and of avoiding difficulties in peptide synthesis. In addition, the results obtained from homopolypeptide chains and repetitive sequences can be compared to results from polymer theory^{32; 89}.

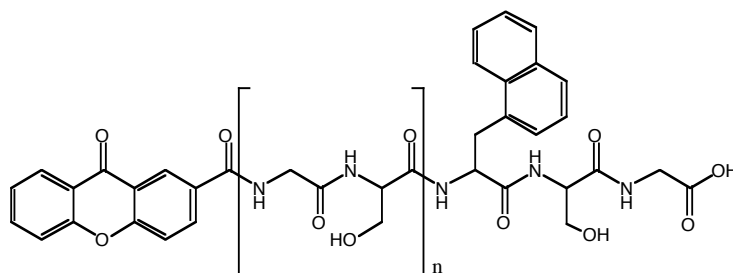


Figure 9: Molecular structure of Xan-(Gly-Ser)_n-NAla-Ser-Gly peptides with $n = 1 - 28$. Xanthone acid (Xan) and naphthylalanine (NAla) were linked at N- and C-terminus, respectively. In the following we will give the amino acid sequences as short cuts to facilitate the reading of this thesis.

Measurements of intrachain dynamics in flexible (GS)-peptide chains revealed single exponential kinetics for end-to-end contact formation in all peptides (Figure 10A)^{12; 31}. The observation of single-exponential kinetics for the loop-closure reaction of unstructured polypeptides indicates fast interconversion between different conformations^{61; 62} leading to an equilibrated distribution of the ensemble of conformations. Indeed, ultrafast spectroscopic investigations revealed that conformational dynamics in small peptides occur in the picosecond time range^{90; 91; 92} and are much faster at the chosen conditions than the measured rate constants for loop closure reaction in flexible peptides^{12; 18; 31; 68}. Additionally, the observation of single-exponential kinetics suggests that the contact radius of the Van-der-Waals interaction is small compared to the averaged peptide chain length, resulting that only a small fraction of conformations makes contact at equilibrium conditions⁶².

The loop-size dependence of end-to-end contact formation of poly(glycine-serine) peptide chains with up to 57 peptide bonds between the triplet labels showed that

diffusion over long and over short distances is limited by different chemical processes (Figure 10B).

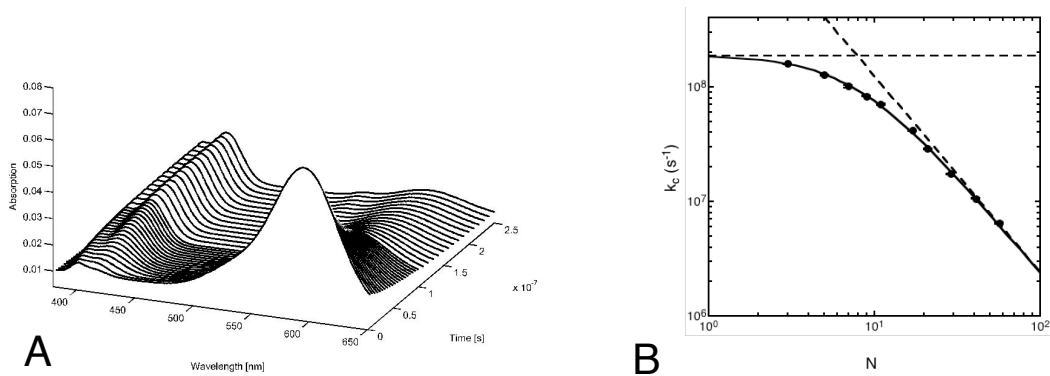


Figure 10: (A) Transient absorbance spectra of TTET in a (GS)₁₄-peptide. TTET from xanthone to naphthalene is observed at 590 nm for xanthone triplet decay and 420 nm for the formation of naphthalene triplets. Intrachain contact formation is well described as a single exponential process. (B) Effect of increasing chain length on the rate constants of contact formation k_c in a series of peptides with alternating glycine-serine repeats between triplet donor and acceptor. N represents the number of peptide bonds between donor and acceptor. The continuous line represents fits to equation 3.1 and give values of $k_0 = 1.8 \pm 0.2 \times 10^8 \text{ s}^{-1}$, $k_1 = 6.7 \pm 1.6 \times 10^9 \text{ s}^{-1}$ and $m = -1.72 \pm 0.08$. The broken lines represent the limiting regimes for end-to-end contact formation in poly(glycine-serine) peptides with a length-independent upper limit for contact formation of k_0 and length-dependent parameters $k_1/N^{m \ 31}$.

Rate constants of intrachain contact formation over short distances are virtually independent of chain length with a limiting time constant ($\tau_0 = 1/k_0$) of $5 \pm 1 \text{ ns}$ for (GS)-peptides³¹. With increasing chain length the rate constant of end-to-end contact formation decreases with $k \sim N^{-1.7 \pm 0.1}$ for long (GS)-peptides with N being the number of peptide bonds between triplet donor and triplet acceptor. This behaviour indicates a stronger distance dependence than calculated theoretically for ideal chains ($k \sim N^{-1.5}$)⁵⁸;⁶¹. Solvent and intramolecular interactions significantly influence the dimensions of real polymer chains³². Accounting for these effects⁶¹; ⁶³ it is expected that in good solvents the rate constants for end-to-end contact formation scale with $k \sim N^{-1.8}$, which is nearly the value, obtained for long (GS)-peptide chains³¹. The complete length dependence of

intrachain diffusion in unstructured polypeptides could thus be described by the following equation (Figure 10B):

$$k_c = \frac{1}{1/k_0 + 1/(k_l \cdot N^m)} \quad (3.1).$$

This suggests that chain stiffness governs local chain dynamics whereas entropy-limited conformational search sets the limit for contact formation over longer chain distances. The results set an upper time scale for intramolecular interactions on peptide chains and reveal that an $i, i + 3$ contact in short glycine containing turns is not formed faster than in 5 ns^{31} .

Table 2: Comparison of end-to-end contact formation rate constants in different model systems.

Sequence	Loop size	System	Method	k_c (s^{-1})	Conditions	Ref.
unfolded Cyt C	~ 50	Heme/Met	bond formation of heme-ligand	$2.8 \cdot 10^4$	5.6 M Gdn HCl, 40°C	⁷⁰
(GlySer) ₁₋₄	3 – 9	Txan/NAla	TTET	$5.0 - 1.4 \cdot 10^7$	EtOH, rt	¹²
(AlaGlyGln) ₁₋₆	4 – 19	Trp/Cys	triplet quenching	$2.7 - 0.7 \cdot 10^7$	10 mM P _i , 22.5°C	¹⁸
(GS) ₄	9	Xan/NAla	TTET	$1.0 \cdot 10^8$	water, rt	⁷³
(Ser) ₅	6	Xan/NAla	TTET	$5.2 \cdot 10^7$	water, rt	⁷³
(GlySer) ₁₋₁₀	3 – 21	DBO/Trp	fluorescence quenching	$6.8 - 1.1 \cdot 10^7$	D ₂ O, rt	⁶⁸
human p53	8 – 9	MR121/Trp	single molecule fluorescence quenching	$8.3 - 6.6 \cdot 10^6$	0.1 M P _i , 25°C	⁶⁹
unfolded Cyt C	15	Zn-porphyrine/Ru-complex	triplet quenching	$4.0 \cdot 10^6$	5.4 M Gdn HCl, 22°C	⁶⁷
(GlySer) ₁₋₂₇	3 – 57	Xan/NAla	TTET	$16 - 0.6 \cdot 10^7$	water, rt	this work
(Ser) ₂₋₁₁	3 – 12	Xan/NAla	TTET	$7.2 - 3.4 \cdot 10^7$	water, rt	this work

Txan: Thioxanthone, NAla: Naphthylalanine, Xan: Xanthone, DBO: 2,3-Diazabicyclo[2,2,2]oct-2-ene, GdnHCl: Guanidine hydrochloride, EtOH: Ethanol, rt: room temperature.

In comparison with results obtained from other methods, the experimentally determined rate constants for intrachain contact formation are the fastest ones (Table 2). All other determined rate constants are at least 2-3 times slower. However, the comparison of the rate constants obtained from different methods is complicated, since the systems were

investigated at different conditions. The sequences and solvent conditions are varied in all cases. The crucial point in the spectroscopic determination of absolute rate constants for intrachain contact formation is the application of electron donor and acceptor pairs, which exchange electrons in a diffusion-controlled reaction. Investigating the electron transfer reaction with femtosecond laser spectroscopy⁸⁴, viscosity and temperature dependencies, only for the xanthone/naphthalene system it could be shown that the spectroscopic method provides the requirements to determine absolute rate constants for intrachain contact formation.

The dynamics of end-to-end contact formation in unstructured polypeptide chains are several time decades faster than barrier crossing reactions in protein folding⁹³ and protein dynamics in the native state^{14; 94}, which occur in the microsecond to second time range. The results indicate, that the unfolded state can be regarded as a state consisting of an ensemble of different conformations, which interconvert fast compared to barrier crossing events¹¹.

These results have been published in the following article:

“Dynamics of Unfolded Polypeptide Chains as Model for the Earliest Steps in Protein Folding.” Florian Krieger, Beat Fierz, Oliver Bieri, Mario Drewello & Thomas Kiefhaber (2003) *J. Mol. Biol.* **332**, 265-274.

3.3 Effect of amino acid sequence on intrachain contact formation in peptides

Both theoretical calculations^{34; 35; 36; 60} and experimental investigations²² showed that glycine and proline residues influence the flexibility of polypeptide chains. All other amino acids residues are predicted to not change significantly the chain dimension of long polypeptides. First investigations on intrachain dynamics of a S₄GS₄- and a (GS)₄-

peptides revealed that replacement of glycine residues by serine residues significantly slows down end-to-end contact formation⁷³. Further experiments were carried out on polyserine peptides of up to 11 residues length and compared to the behaviour of poly(glycine-serine) chains³¹. For all polyserine peptides intrachain contact formation can be well described by single exponential kinetics (Figure 11A). The observed rate constants of the polyserine chains are 2-3 fold slower than the ones observed in poly(glycine-serine) peptides (Figure 11B). For short chains with less than five serines contact formation is virtually independent of the chain length and reaches a limiting value of $8.7 \cdot 10^7 \text{ s}^{-1}$. With increasing chain length the rate constant of contact formation decreases slightly stronger than observed for poly(glycine-serine) peptide chains. Due to limitations in peptide synthesis, it was not possible to investigate chains longer than (Ser)₁₁, which prevented an accurate determination of the scaling exponent for long polyserine peptide chains.

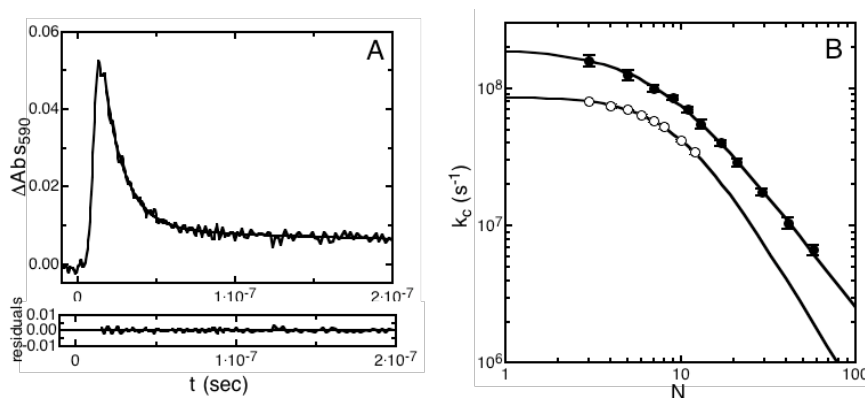


Figure 11: (A) Time course of formation and decay in a Xan-(Ser)₂-NAla-SerGly-peptide with a time constant of $\tau = 14 \pm 1 \text{ ns}$ for end-to-end contact formation. (B) Comparison of the end-to-end contact formation rate constants k_c of polyserine peptides (○) with rate constants of poly(glycine-serine) peptides (●) and the corresponding fits (solid lines). N represents the number of peptide bonds between triplet donor and acceptor. For polyserine peptide chains experimental data are described by equation 3.1 with $k_0 = 8.7 \pm 0.8 \cdot 10^7 \text{ s}^{-1}$, $m = 2.1 \pm 0.3$ and $k_1 = 1.0 \pm 0.8 \cdot 10^7 \text{ s}^{-1}$. The fit for the data of (GS)-peptides, see figure 10.

The strong effect of the replacement of glycine residues by serine residues stimulated us to investigate the effect of other amino acids on chain dynamics using host-guest

peptides of the structure XanSer-Xaa-SerNAlaSerGly with Xaa = Gly, Ser, Ala, Ile, His, Glu, Arg, and Pro. All amino acids, except proline and glycine show very similar rate constants of end-to-end diffusion with time constants between 12-25 ns for the formation of $i,i+3$ contacts⁶³. There are small, but significant differences in rate constants between amino acids with short side-chains (Ala, Ser) and longer side-chains (Ile, Glu, Arg) suggesting that larger side-chains slow down slightly local intrachain dynamics. The glycine containing peptide is significantly more flexible and reveals faster intrachain contact formation rate constants compared to polyserine peptides ($\tau = 8 \pm 1$ ns). Incorporation of a proline residue into the host-guest peptide leads to a more complex intrachain diffusion reaction, with an additional fast kinetic phase of 20 % amplitude ($\tau = 8 \pm 2$ ns) and a slow kinetic component of 80 % amplitude ($\tau = 50 \pm 3$ ns). This essentially reflects the *cis-trans* ratio at the Ser-Pro peptide bond in the host-guest peptides as determined by ¹H-NMR spectroscopy⁹⁵. A possible molecular origin of the bi-phasic kinetics in short proline containing peptides might be differences in the distribution of the end-to-end distances of *cis* and *trans* proline isomers. Both the distribution of *cis* and *trans* proline isomers are in equilibrium during end-to-end contact formation, since peptidyl-prolyl isomerization reactions occur on the second time range⁹⁶. Therefore an interconversion of the *cis* and *trans* population can be ruled out on the nanosecond time scale. Glycine and proline are frequently found in sequences of hairpin loops and turns⁹⁷. The results indicate that loops with glycine or a *cis* proline can form a first intramolecular contact within 5-10 ns. In glycine- and proline-free sequences intramolecular contacts will be formed within 12-25 ns⁶³.

Proline and glycine residues show the largest effect on intrachain dynamics in short peptides. To determine the effect of a single proline and glycine residue on local chain dynamics more accurately we synthesized host-guest peptides of the structure

$Xan(Ser)_x-Xaa-(Ser)_yNAlaSerGly$ with $x,y = 1 - 7$ and $Xaa = Gly, Ser$ or Pro . For short proline containing peptides with $x,y \leq 2$ local chain dynamics for peptides with *cis* and *trans* proline isomers could be distinguished leading to more complex kinetics of contact formation with two relaxation times (Figure 12).

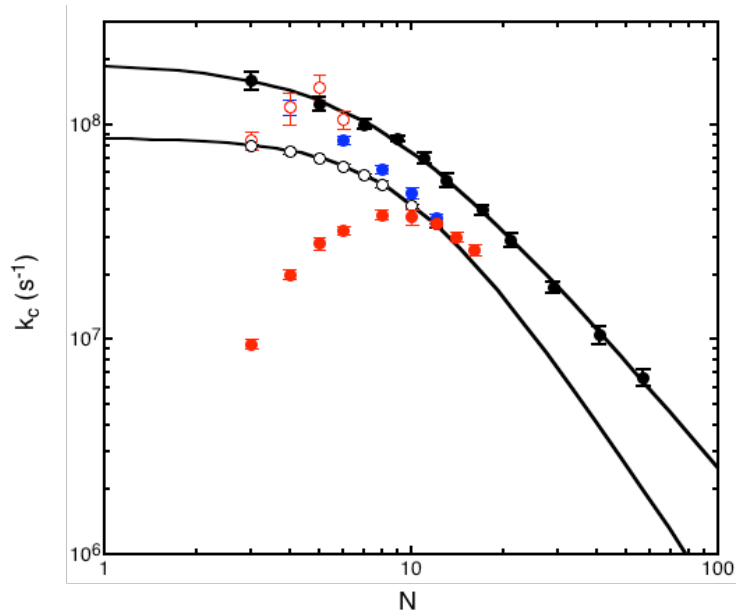


Figure 12: Effect of increasing chain length (N) on the rate constant of contact formation in various series of peptides. The effect of loop-size on the kinetics of contact formation in poly(glycine-serine) (\bullet) and polyserine (\circ) peptides, is compared to the kinetic rate constants in $(Ser)_xGly(Ser)_y$ (\bullet), $(Ser)_xtransPro(Ser)_y$ (\bullet) and $(Ser)_xcisPro(Ser)_y$ (\circ) with x,y varying from 0 to 7. N represents the number of peptide bonds between triplet donor and acceptor. The lines describe the loop-size dependence of intrachain contact formation in homo-polypeptides as described in ref³¹.

Short peptides with a *cis* proline show faster contact formation rate constants than peptides with a *trans* proline in all cases, suggesting that the stereoirregularity of the peptide bond decreases the end-to-end distance³² leading to faster end-to-end contact formation rate constants. In peptide chains with $x,y \geq 5$ ($N = 11$) the effect of proline and glycine residues on local chain dynamics vanishes (Figure 12).

To evaluate whether the double exponential kinetics in short proline containing peptides are indeed due to different dynamics in peptide chains with *cis* and *trans* peptidyl-prolyl

bonds we tested the effect of pseudoproline (psPro) containing peptides (Val-[$\Psi^{\text{Me,Me}}\text{Ser}$], Figure 13A, B) on chain dynamics in two different peptides of the sequence XanSerVal-Xaa-SerNAlaSerGly with Xaa = Pro and [$\Psi^{\text{Me,Me}}\text{Ser}$]. It is known that pseudoprolines increase selectively the *cis* content of the peptidyl-prolyl bond in unstructured peptides up to 90 %^{98;99}.

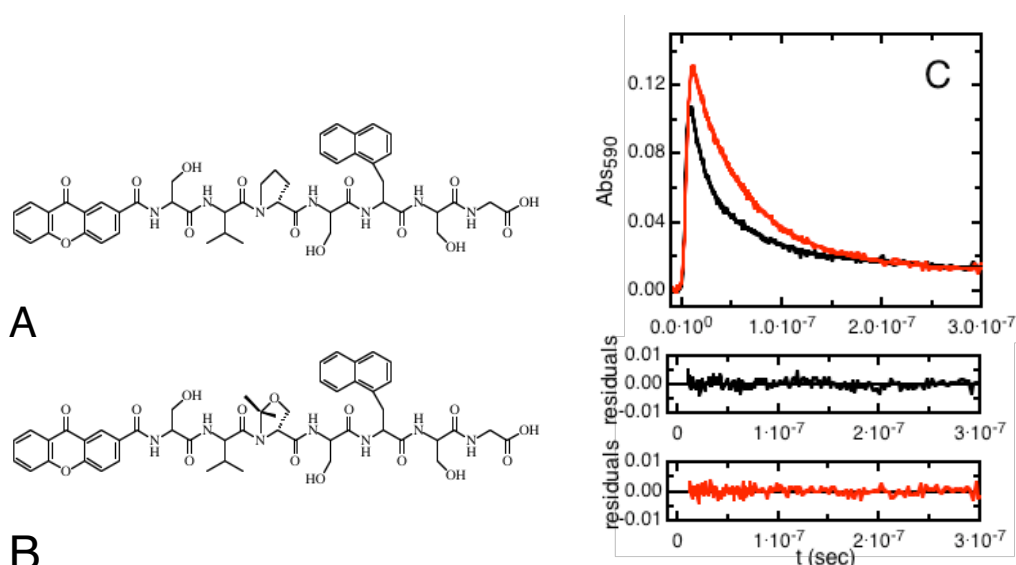


Figure 13: Molecular structure of XanSerVal-Pro-SerNAlaSerGly (A) and XanSerVal-psPro-SerNAlaSerGly (B). Panel C compares the kinetics of intrachain contact formation in peptide A (—) and peptide B (—) in 10 mM potassium phosphate, pH 7.0 at 22.5 °C.

The pseudoproline containing peptides show also double exponential kinetics for end-to-end contact formation with a significant increased amplitude of the faster reaction compared to the corresponding peptide containing a natural proline residue (Figure 13C). This confirms that the faster reaction corresponds to the dynamics of peptides with a peptidyl-prolyl bond in *cis* conformation. Furthermore the temperature dependencies of contact formation in short proline containing peptides revealed significantly higher activation barrier of 10-20 kJ/mol for the *cis* isomer, suggesting that a *cis* proline residue in a peptide chain strongly decreases the conformational space of the chain.

In summary, it could be shown that glycine and proline residues influence mainly local chain dynamics of intramolecular interaction on unfolded polypeptide, such as the dynamics of loop and β -turn formations, whereas a single proline or glycine residue does not affect end-to-end contact formation over distances with five residues on each side of the glycine or proline. For an $i, i + 4$ - and shorter interactions of proline containing sequences the *cis* proline conformer exhibits significant faster intrachain dynamic than the *trans* conformer. The interpretation of these results is supported by conformational and statistical analyses of allowed conformations in proline containing sequences. The analyses indicated that the average end-to-end distances of *cis* proline containing peptides are significantly shorter compared to *trans* proline containing peptides. Additionally the conformational space is stronger restricted in the *cis* proline isomer than in the *trans* proline isomer (A. Möglich, personal communication).

These results are described in detail in the following manuscripts:

“Dynamics of Unfolded Polypeptide Chains as Model for the Earliest Steps in Protein Folding.” Florian Krieger, Beat Fierz, Oliver Bieri, Mario Drewello & Thomas Kiefhaber (2003) *J. Mol. Biol.* **332**, 265-274.

“Effect of Proline and Glycine Residue on Dynamics and Barriers of Loop Formation in Polypeptide Chains” Krieger F., Möglich A. and Kiefhaber T. (2004), submitted.

Position and direction of the triplet donor/acceptor pair

In all synthesized peptides the triplet donor xanthone acid and triplet acceptor naphthylalanine were placed at the N- and C-terminus, respectively. To test, if the position and direction of the triplet chromophores influences the measured rate constants for intrachain contact formation we synthesized a (GS)₄-peptide with 1-

naphthoic acid (Naph) as triplet acceptor at the N-terminus and xanthone acid (Xan) as triplet donor at the C-terminus and compared to rate constants of the standard (GS)₄ peptide (Table 3). Xanthone acid was introduced into the peptide sequence using the non-natural amino acid 2,3-diaminopropionic acid (Dpr). Both peptides show virtually identical rate constants for end-to-end contact formation, suggesting that the position of the triplet donor/acceptor pair at the N- and C-terminus does not influence the dynamics. The comparison was performed on a (GS)₄-peptide. However we cannot rule out, that local chain dynamics over two, three peptide bonds are influenced by the position of the triplet chromophores.

Table 3: Comparison of the sequence and rate constants of (GS)₄-peptides with triplet donor and acceptor at the N- or C-terminus of the peptide.

Peptide	k_c (s ⁻¹)
Xan-(GS) ₄ -NAla-SG-COOH	$8.3 \pm 0.4 \cdot 10^7$
Naph-(GS) ₄ -Xan(Dpr)-SG-COOH	$8.1 \pm 0.4 \cdot 10^7$

Xan: xanthone acid, NAla: 1-naphthylalanine,
 Naph: 1-naphthoic acid, Dpr: 1,2-diaminopropionic acid,
 Rate constants were determined in H₂O, pH 7 at 22.5°C

3.4 Barrier-limited and diffusive chain dynamics

3.4.1 Activation barriers and viscosity dependence of intrachain diffusion in polypeptides

The effect of viscosity and temperature on a chemical reaction allows to estimate the contributions arising from solvent motion and limiting activation barriers. Experimental⁶⁶ and theoretical^{61;62} approaches suggest that diffusion determines intrachain contact formation in long synthetic polymers. Due to Brownian motion the conformations of unstructured polypeptide chains diffuse relative to the solvent molecules and interconvert into each other. The extent and magnitude of internal chain dynamics are limited only by molecular connectivity and excluded volume of the

chain¹⁰⁰. The most straightforward experimental way to distinguish between diffusion- and reaction-controlled processes is to examine the isothermal viscosity and temperature dependencies of the reaction.

Protein folding reactions represent complex processes since entropic and enthalpic terms contribute to the activation barrier¹⁰¹. The temperature dependence of protein folding reactions has been studied extensively starting with the experiments of Pohl¹⁰². Investigations on several small proteins^{103; 104; 105; 106} have shown that the influences of temperature on the rate-limiting steps in folding are complex, since large changes in the heat capacity contribute to the folding process. However little is known about the effect of temperature and solvent viscosity on the dynamics of the earliest steps in protein folding. The loop length dependence on intrachain contact formation in poly(glycine-serine) peptides revealed that the intrinsic chain dynamics are limited by different processes for motions over short and long distances³¹. To investigate both limits more precisely first we performed viscosity dependencies of intrachain contact formation in short and long poly(glycine-serine) peptides. The viscosity dependencies (Figure 14 A, B) revealed that in short poly(glycine-serine) peptides intrachain diffusion is not fully determined by solvent motions, exhibiting a fractional viscosity dependence described

with the empirical equation

$$k_c = k_c^0 \left(\frac{\eta}{\eta_0} \right)^\beta \quad (3.2)$$

where η_0 is the reference solvent viscosity in water and k_c^0 the rate constant of end-to-end contact formation at η_0 . The empirical fractional exponent β reflects the sensitivity of the reaction to solvent viscosity η . A β -value of -1 indicates that the observed rate constant is inversely proportional to solvent viscosity as found for diffusion-controlled reactions⁷⁸. A β -value of 0 indicates that the reaction rate constant is independent of solvent viscosity. Reactions, which are not influenced by solvent viscosity, are

determined fully by the activation barrier as found e.g. for the *cis-trans* peptidyl-prolyl bond isomerization (Table 4, p. 41).

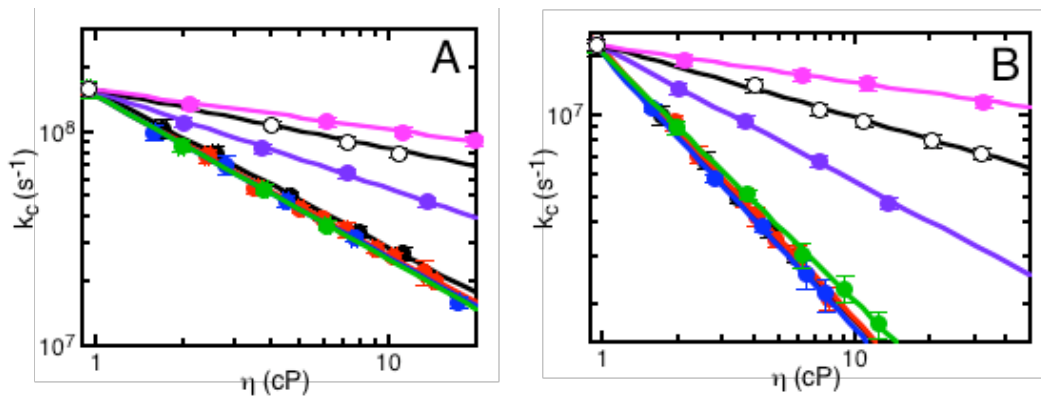


Figure 14: Viscosity dependencies of the observed end-to-end contact formation rate constants k_c in $(GS)_1$ -peptide (A) and $(GS)_{14}$ -peptide (B) in presence of viscous co-solutes with different molecular weight. As viscous co-solutes were chosen: ethylene glycol (●), glycerol (●), glucose (●), sucrose (●), polyethylene (PEG) 1500 (●), PEG 6000 (○) and PEG 20000 (●). The continuous lines show fits of the experimental data to equation 3.2 with the slopes β . All measurements were carried out at 22.5°C.

The solvent viscosity measured in a viscometer is defined as the macroscopic viscosity. However molecular interactions between reactant, solvent and co-solute are described by the microscopic viscosity¹⁴. Several investigations on diffusion processes showed that the macroscopic and microscopic viscosities are different if the viscous co-solute is larger than the reactant¹⁴. To test for effects due to changes in microscopic viscosity we used viscous co-solutes of different size to determine the effect of solvent viscosity on intrachain contact formation. For small viscous co-solutes β -values of -0.75 ± 0.03 were determined for end-to-end contact formation in the $(GS)_1$ -peptide, whereas the end-to-end intrachain dynamics in long flexible poly(glycine-serine) peptides scale inversely with solvent viscosity with β values of -1.00 ± 0.03 (Figure 14A, B). For larger co-solutes, such as poly(ethylene glycol) the β -values decrease significantly with increasing co-solute size. The results suggest that only small viscous co-solutes such as

ethylene glycol, glycerol, glucose and sucrose can be applied to determine the effect of solvent viscosity on the end-to-end contact formation rate constant in poly(glycine-serine) peptides.

To get more information on the fractional viscosity dependence we determined the length and temperature dependence on β (Figure 15). Indeed, with increasing chain length the fractional exponent β increases exponentially with chain length N and converges at the value -1 for polypeptide chains longer than 20 residues (Figure 15A). This result suggests that poly(glycine-serine) peptide chains with more than 20 peptide bonds behave like random coil chains and end-to-end contact formation becomes diffusion-controlled as suggested theoretically⁶¹. The effect of temperature on the fractional exponent β is shown in figure 15B. A slight increase of β with increasing temperature is observed. Over the chosen temperature range between 10 and 35°C the fractional exponent β increases from -0.74 ± 0.03 to -0.81 ± 0.04 . The effect of temperature on β is not large but significant. The results obtained from the viscosity dependence suggest that intrachain dynamics might be limited by additional barriers.

The effect of viscosity on intrachain diffusion in short and long polypeptides is independent of the size of the viscous co-solute, if the co-solute dimensions are significantly smaller than the dimension of the investigated peptides, like for ethylene glycol, glycerol, glucose or sucrose. In presence of small co-solutes identical β -values are determined. The use of larger viscous co-solutes such as polyethylene glycols leads to significant smaller β -values in short and long poly(glycine-serine) peptides (Figure 14 A, B). With increasing co-solute size the β -values decrease indicating that the macroscopic viscosity cannot describe accurately the effect of solvent viscosity on the reaction rate constant anymore.

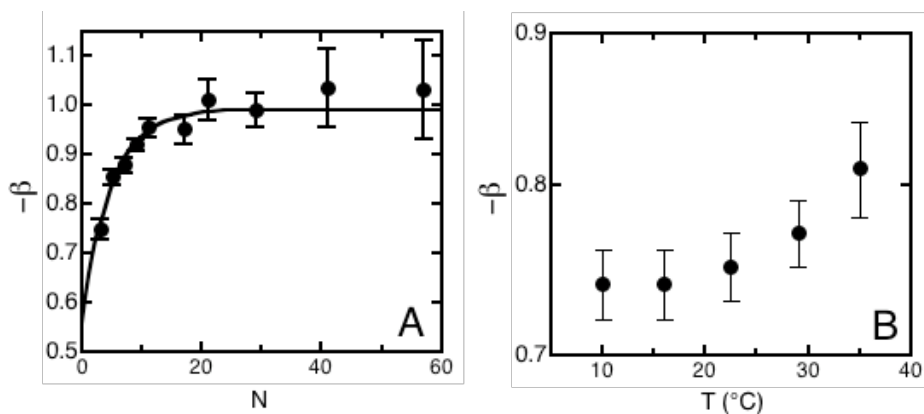


Figure 15: Effect of the empirical exponent β on chain length N of poly(glycine-serine) peptides (A) at 22.5°C and on temperature in the $(GS)_1$ -peptide (B). For all measurements glycerol was chosen as viscous co-solute. The solid line in (A) represents an empirical exponential fit of the equation:

$$-\beta(N) = (0.99 \pm 0.02) \cdot e^{-(0.20 \pm 0.04) \cdot N} - (0.43 \pm 0.06).$$

Since in long peptides the time constant of end-to-end contact formation is directly proportional to solvent viscosity it is suggested that the effect of the chosen small viscous co-solutes influence only the diffusive motion of the chain. Therefore a significant effect on the chain dimension of the poly(glycine-serine) peptides due to preferential hydration can be ruled out as found for proteins¹⁰⁷.

The origin of the fractional viscosity dependence for short $(GS)_1$ -peptides indicates the existence of enthalpic barriers, which limit intrachain diffusion. The determination of the effect of temperature on intrachain contact formation in short and long $(GS)_1$ -peptides should give insight about the barriers of the rate-limiting steps.

It is well known that the viscosity of water depends strongly on temperature⁸³. To determine effects of temperature on intrachain dynamics all observed rate constants for end-to-end contact formation were corrected against water viscosity changes due to temperature variation, using equation 3.2. The slight change of the fractional exponent β with temperature was neglected (Figure 15B). The temperature dependency on intrachain contact formation in $(GS)_1$ -peptides revealed that viscosity corrected rate

constants for end-to-end contact formation exhibit apparently Arrhenius activation behaviour ($\ln k_c^0 \propto \frac{E_a}{RT}$) in all investigated polypeptide chains (data not shown), suggesting that changes in heat capacity do not contribute to the activation barrier¹⁰¹.

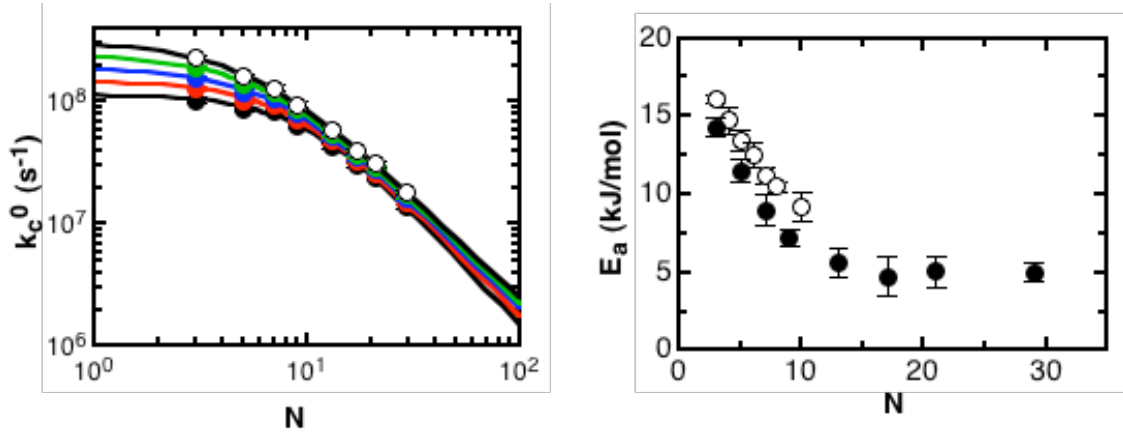


Figure 16: (A) Effect of temperature on viscosity-corrected rate constants k_c^0 for end-to-end contact formation in flexible poly(glycine-serine) peptides at 2.5°C (●), 12.5°C (●), 22.5°C (●), 32.5°C (●) and 42.5°C (○). The observed rate constants were corrected against changes of the water viscosity due to temperature changes using equation 3.2. (B) Effect of loop size on the apparent Arrhenius activation barrier E_a in (●) poly(glycine-serine) and (○) polyserine peptides.

Intrachain diffusion over short and long distances is influenced differently by temperature (Figure 16A). Intrachain dynamics over short distances are significantly stronger influenced than end-to-end contact formation over long distances. The effect of temperature on the loop size dependence of the rate constants suggests that poly(glycine-serine) peptide chains become more flexible with increasing temperature. This result is expected since the chain properties depend strongly on the nature of the molecular bond, which should be affected by temperature³².

For the short (GS)₁-peptide an Arrhenius activation barrier of 14.3±1.0 kJ/mol was determined, whereas intrachain motions over longer distances are only slightly affected by temperature with activation barriers of 5.0±1.0 kJ/mol (Figure 16B). The low activation barrier for intrachain contact formation in long poly(glycine-serine) peptides

indicates that the process is nearly diffusion-controlled. The less flexible polyserine chains show a slightly increased activation barrier compared to the more flexible poly(glycine-serine) peptide chains, which might originate from a decrease in conformational space due to the larger side-chain of the serine residue.

Comparison of the effect of temperature and viscosity on dynamics of other chemical reactions

The origin of the fractional viscosity dependence on loop closure dynamics might be explained due to the presence of apparent activation barriers. For several chemical dynamic processes like segmental polymer motions^{108; 109; 110}, protein fluctuations¹¹¹ and photoisomerization of small organic molecules^{112; 113} it is empirically observed that the relative solvent viscosity η decreases the reaction rate constants with a fractional power law as given in equation 3.2 (Table 4).

Table 4: Activation barriers and observed fractional exponents β of the viscosity dependence for different chemical reactions.

Investigated reaction	E_a (kJ/mol) ^a	$-\beta$	Technique	Ref
Conformational fluctuation in myoglobin due to binding of CO and O ₂	15-35	0.4-0.8	Laser flash photolysis	¹¹¹
Photochemical isomerization of cyanine dye (DODCl)	11.3-57.3	0.26-0.43	Flash photolysis	¹¹²
Polystyrene local chain dynamics	14±3	0.76±0.05	² H NMR relaxation study	¹⁰⁸
Segmental dynamics of polyisoprene	12±1	0.65±0.05	Fluorescence anisotropy measurements	¹⁰⁹
Polystyrene local chain dynamics	10.5±0.6	0.73±0.02	Electron spin resonance	¹¹⁰
Folding of GB1 ⁴¹⁻⁵⁶ hairpin	~ 0	0.92-0.95	T-jump, fluorescence	¹¹⁴
<i>Cis-trans</i> peptidyl-prolyl isomerization	~ 84±4 ^{95; 96}	0	Enzyme-coupled assay	this work
End-to-end contact formation in (GS) ₁ -peptide	12.5±1.0	0.75±0.03	TTET/ Laser flash photolysis	this work
End-to-end contact formation in (GS) ₁₄ -peptide	5.0±1.0	1.00±0.03	TTET/ Laser flash photolysis	this work
Bimolecular TTET form Xan to NAla	3.0±1.0	1.00±0.03	Laser flash photolysis	this work

a: E_a was calculated from viscosity corrected rate constants, accordingly equation 3.2.

Interestingly, there is apparently a correlation between the magnitude of the activation barrier and the fractional exponent β . For small activation barriers the β -value converges against -1 and for high activation barriers, as found for the *cis-trans* peptidyl-prolyl isomerization, the β -value is zero. Several factors are suggested to explain the experimentally observed fractional viscosity dependence of the reaction rate constant. As mentioned above macroscopic bulk diffusion is not an entirely adequate model to describe diffusion-controlled reactions and the Stokes-Einstein relation breaks down on molecular dimension¹¹⁵. One must expect that microscopic diffusion coefficients at molecular dimension will be different from the macroscopic coefficients of diffusion. Indeed, it is expected that only in the limit when the rotating group of the solute is bulky compared to solvent molecules the rotating group does feel the full effect of the macroscopic viscosity. However this argument can be ruled out for the interpretation of the observed β -value in viscosity dependence of end-to-end contact formation in short peptides, since the fractional exponent β is independent of co-solute size, if the dimension of the co-solute is significant smaller than the peptide (Figure 14A, B).

Another approach suggests that a barrier crossing process is not only determined by the friction coefficient but also by the frequency of the top of the transition state^{116; 117}. If an activated crossing process over a transition state becomes faster than solvent motion, not all solvent fluctuation modes will be effective in damping the barrier crossing¹¹⁸. In this regime solvent motions, which contribute to the macroscopic shear viscosity, will be too slow to have any influence on the barrier crossing. Indeed, a relation between activation barrier and fractional viscosity dependence is observed qualitatively for several different reactions (Table 4). In comparison to results obtained from other system the results suggest that end-to-end contact formation in short (glycine-serine) peptides is limited by an apparent activation barrier, leading to a partly diffusion-controlled reaction. For

longer poly(glycine-serine) chains the apparent activation barrier decreases and intrachain contact formation becomes nearly diffusion-controlled.

Comparison to activation parameter obtained from protein folding experiments

The effects of temperature on the barriers of the rate-limiting steps in protein folding reactions are still not well understood. Temperature dependencies of observed rate constants of protein folding reactions exhibit complex relations, suggesting that the changes in the heat capacity significantly influence the activation barriers^{104; 105; 106}. Determined activation parameters for refolding reactions vary over a large range between 30 and 85 kJ/mol (Table 5).

Table 5: Activation barriers observed in protein folding experiments compared to activation barriers found for end-to-end contact formation in short peptides.

Reaction	E_a (kJ/mol) ^a	E_a (kJ/mol) ^b	temperature (°C)	Reference
Refolding of horse cytochrome c	-	31±2 ^c	Temp.-independent	¹¹⁹
First transition in refolding of tendamistat (C45A/C73A)	51±1 ^e	68±1 ^d	25	¹⁰⁶
Local refolding of the GCN4leucine Zipper	-	86±4 ^d	Temp.-independent	¹²⁰
Refolding of cold shock protein CspB	-	32±2 ^d	25	¹⁰⁴
End-to-end contact formation in (Gly-Ser) ₁ peptides	14±1 ^c	28±1 ^c	Temp.-independent	This work
End-to-end contact formation in (Gly-Ser) ₁₄ peptides	5±1 ^c	21±1 ^c	Temp.-independent	This work

a: corrected against viscosity effects.

b: no correction against viscosity effects.

c: using Arrhenius equation¹²¹.

d: using Eyring equation¹²².

e: using Kramers equation¹¹⁶.

The comparison of the barriers observed in refolding experiments and the data obtained from intrachain dynamics is quite critical. The effect of solvent viscosity on the refolding rate constant is quite complex for most proteins and therefore the contributions arising from changes in solvent viscosity due to temperature changes are not included in the analysis of the temperature dependencies of protein folding reactions. Therefore

apparent activation barriers for intrachain contact formation without viscosity correction are compared to barriers determined in protein refolding. All measured activation barriers of refolding are higher than the activation barriers for intrachain diffusion. Especially, the barriers for intrachain dynamics over long distances are significantly lower than protein refolding barriers. Therefore it can be ruled out that intrachain contact formation over long distances is the rate-limiting step in protein refolding. Interestingly the barriers determined in refolding of horse cytochrome c and cold shock protein CspB are in similar magnitude as activation barriers observed for intrachain contact formation over short distances. However it is not clear, how intrachain dynamics over short distances contribute to the barriers of protein refolding, since these steps might be affected by several other chemical processes. The results indicate that intrachain motions over short distances are significantly limited by activation barriers of around 12-16 kJ/mol, whereas in long flexible poly(glycine-serine) peptide chains end-to-end contact formation is a virtually diffusional process, as expected from polymer theory^{61; 62}.

3.4.2 Kinetic isotope effects on the dynamics of loop formation in short peptides

Activation barriers of 12-16 kJ/mol significantly limit intrachain diffusion in short peptides. The magnitude of the measured Arrhenius barriers is comparable to rotational barriers determined in di- and tripeptides¹²³. The molecular origin of the barriers for end-to-end contact formation is not known. To obtain more information about the origin of these barriers we determined kinetic isotope effects (KIE) on local intrachain dynamics since the incorporation of deuterium on the C_α-position on the polypeptide backbone might influence chain dynamics. Deuterium labelling of the C_α-position next to a peptidyl-prolyl bond revealed a deuterium secondary kinetic isotope effect (²H-SKIE)

on the rate and equilibrium constant of *cis-trans* peptidyl-prolyl bond isomerization¹²⁴. The effect was explained as hyperconjugative in origin.

SKIE's are caused by isotopic substitutions which are neither made nor broken in the rate limiting step¹²⁵. The origin of ²H-SKIE is most likely hyperconjugative¹²⁶ and ²H-SKIE's are interpreted as arising mainly from rotational motions of isotopic substituents. Steric size differences due changes in bond length contribute less to SKIE. Generally, ²H-SKIE (k_H/k_D) are with values between 1.2 – 1.7 usually quite small. Differences in zero-point energy changes for H and D, the different populations of excited vibrational levels for H and D, and changes in the mass and moment of inertia might contribute to ²H-SKIE¹²⁶.

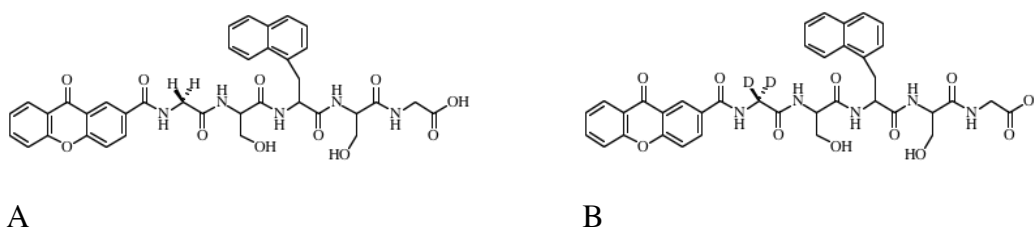


Figure 17: Comparison of the molecular structure of protonated (A) and deuterium labelled (B) (GlySer)₁-peptides.

To test if backbone modification due to deuterium labelling has an effect on the dynamics of end-to-end contact formation in short peptides we substituted the C_α-protons of the glycine residue by deuterium in the (GS)₁-peptide (Figure 17). The deuterium labelling at the C_α-position of a glycine residue might change the rotational barrier or influence the motion of neighbour atoms. For both peptides the end-to-end contact formation rate constants were determined at identical conditions. As shown in figure 18A end-to-end contact formation of protonated and deuterium labelled (GS)₁-peptides show identical intrachain dynamics at 279 K in water. Decreasing the temperature to 235 K does not result in a measurable kinetic isotope effect (Figure 18B).

Isotope labelling at the C_{α} -position of the glycine residue has no influence on end-to-end contact formation in short (GS)-peptides.

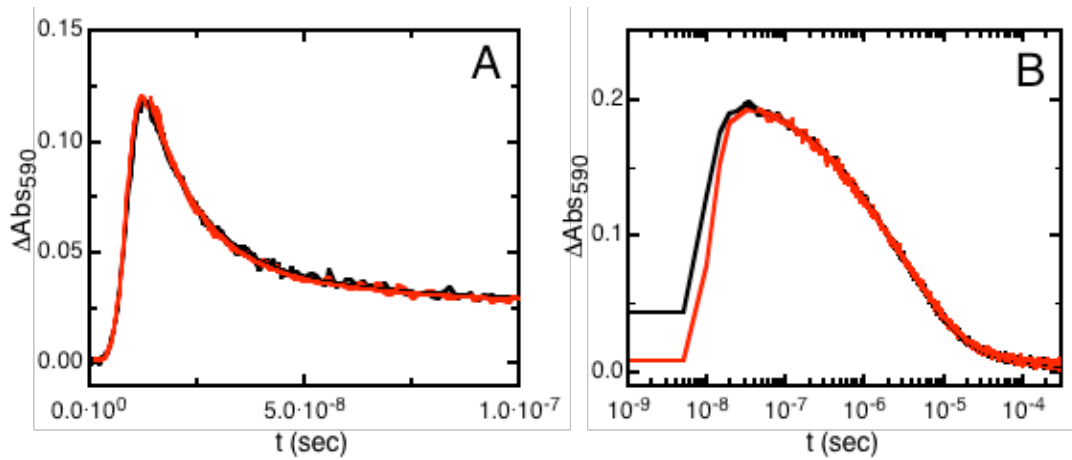


Figure 18: Comparison of the kinetics of end-to-end contact formation in $(GS)_1$ -peptides with protonated (—) and deuterium labelled (—) C_{α} -atom in the glycine residue. Transient TTET kinetics were measured in water, 279 K (A) and in 60 % glycerol, at 235 K (B).

There might be several reasons for this result. First, barriers of rotation around single bonds are only slightly influenced by isotopic deuterium substitutions. The barrier of internal rotation around a deuterium labelled methyl group CD_3 - is only slightly decreased compared to the protonated methyl group CH_3 - by 1 – 4 %¹²⁷.

Second, end-to-end contact formation in short peptides is mainly a diffusive process as shown by the viscosity dependencies (Figure 14A). A strong isotopic effect on local intrachain dynamics should be detected due to deuterium labelling if the diffusion coefficient of this reaction would be influenced. This would include that probably the averaged end-to-end distance of the short peptides is changed due to steric and hyperconjugative interactions¹²⁶.

As mentioned above single exponential kinetics for end-to-end contact formation in all short proline-free peptides are observed. This observation suggests that the interconversion of different conformations is fast compared to the loop closure

reaction^{61; 62} and argues against the notion that end-to-end contact formation in short peptides is influenced by single bond rotation.

3.5 Solvent effects on intrachain contact formation in unstructured peptides

End-to-end intrachain dynamics in short (GS)-peptides are 3-4 times faster in water compared to ethanol⁷³. Both solvents have similar viscosity at room temperature⁸³ suggesting that this effect is due to the nature of the solvent. Ethanol is a better solvent for polypeptide chains than water²³. The ethanol dependence on intrachain dynamics showed⁷³ that the observed rate constants k_{obs} decrease exponentially with co-solute concentration $[x]$ as

$$k_c = k_{H_2O} \cdot e^{-\frac{m_c [x]}{RT}} \quad (3.3)$$

where k_{H_2O} is rate constant for end-to-end diffusion at zero molar co-solute concentration and m_c is an empirical correlation coefficient, which describes changes in rate constants due to co-solute interactions. Such correlations between rate constant or linear free energy for unfolding (ΔG^0) and co-solute concentrations are usually found in protein folding experiments^{128; 129}. In analogy to protein folding investigations m_c -values for end-to-end contact formation can be determined. Co-solutes, such as guanidine hydrochloride, urea, dimethylformamide or trifluoroethanol significantly increase the solubility of peptides and proteins²³ and should affect intrachain contact formation in unstructured peptides.

3.5.1 Guanidine hydrochloride and urea

Guanidine hydrochloride (GdnHCl) and urea are frequently used to unfold proteins and characterize transition states in protein folding¹³⁰. It is assumed that the denaturants interact with the polypeptide backbone because of the similar structure like the peptide

unit in polypeptides (-CO-HN-C_α-) ^{131; 132} leading to changes in chain properties. For all investigated peptides an approximately linear relationship between the logarithm of the observed end-to-end contact formation rate constants $\log(k_{\text{obs}})$ and GdnHCl or urea concentrations is observed (Figure 19A). The m_c -values for GdnHCl are 1.4-1.5 times higher than for urea¹³³ and for both co-solutes m_c -values increase slightly with peptide length (Figure 19B). The decreased m_c -value for short chains might originate from increased chain stiffness in short (GS)-peptides, which leads to a decreased sensitivity to solvent viscosity.

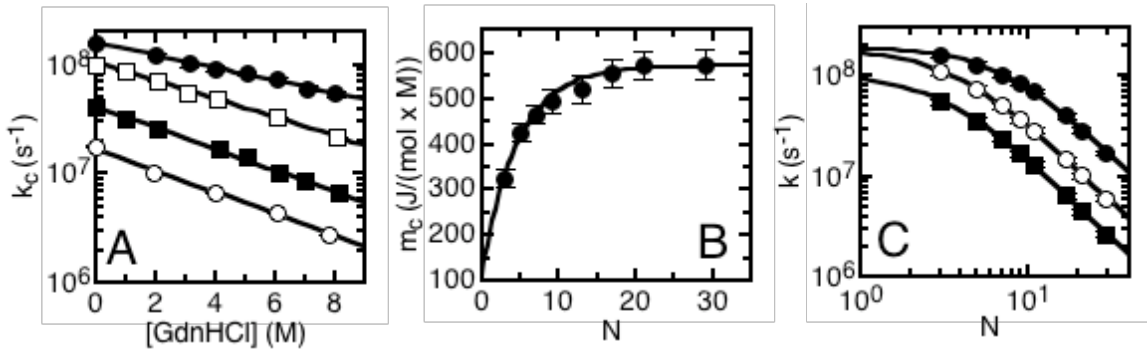


Figure 19: (A) GdnHCl dependence of end-to-end contact formation in various (GS)_n-peptide with $n = 1$ (●), $n = 3$ (□), $n = 8$ (■) and $n = 10$ (○). Solid lines represent fits of the data to equation 3.3. (B) Effect of the loop length on the m_c -value. The increase of the m_c -value with chain length for GdnHCl can be described empirically by an exponential equation:

$$m_c = (570 \pm 20) - (460 \pm 80) \cdot e^{-(0.21 \pm 0.05) \cdot N}$$
 (C) Loop length dependence of the rate constant of contact formation in water (●) and in 8 M GdnHCl (○). The black squares (■) represent rate constants, corrected for viscosity effects and are normalized to $\eta_0 = 0.94$ cP using equation 3.2. Solid lines are fits of the data to equation 3.1.

The distance dependence of contact formation in (GS)-peptides is slightly changed at 8 M GdnHCl compared to pure water (Figure 19C). In 8 M GdnHCl the switch from the distance independent dynamics over short chain segments to the distance-dependent regime for long peptides occurs at shorter distances suggesting that (GS)-peptides become more flexible in the presence of high GdnHCl concentrations. For long peptides

the effect of donor-acceptor distance on the rate constants of intrachain contact formation is similar in 8 M GdnHCl and water with $k_c \sim N^{-1.8 \pm 0.1}$.

Addition of 8M GdnHCl increases the solvent viscosity 2.3 times compared to water¹³⁴. The increase in solvent viscosity significantly affects end-to-end diffusion in (GS)-peptides as shown above. We used equation 3.2 to correct the observed rate constants against the contributions from increased solvent viscosity. The corrected rate constants in 8M GdnHCl still differ from the rate constants in pure water (Figure 19C). This effect is more pronounced for long chains, than for short chains. Interestingly, the limiting intrinsic chain dynamics for intrachain contact formation over short distances are not affected by the addition of high concentration of GdnHCl if the rate constants are corrected against viscosity. This indicates that the limiting intrinsic chain dynamics of end-to-end contact formation depend only on temperature and solvent viscosity. Furthermore the quality of the solvent does not influence the limiting intrinsic dynamics.

The decreased rate constants for intrachain diffusion in long polypeptides might be due to preferential interactions of GdnHCl with the polypeptide chain. Indeed, it could be shown that the effect of GdnHCl on the viscosity-corrected rate constants of end-to-end contact formation can be analyzed successfully by Schellman's weak binding model^{133; 135}. The better solvation of polypeptide chains in 8 M GdnHCl compared to water leads to an increase of the chain dimension due to preferential binding of GdnHCl to the backbone. Due to a swelling of the chain dimension the end-to-end distances increase leading to a decrease in contact formation rate constants (Möglich & Kiefhaber, manuscript in preparation).

Parts of these results are described in the following publications:

“Dynamics of Unfolded Polypeptide Chains as Model for the Earliest Steps in Protein Folding.” Florian Krieger, Beat Fierz, Oliver Bieri, Mario Drewello & Thomas Kiefhaber (2003) *J. Mol. Biol.* **332**, 265-274.

“Molecular basis of the effect of urea and guanidinium chloride on the dynamics of unfolded polypeptide chains.” Möglich A., Krieger F. and Kiefhaber T. (2005) *J. Mol. Biol.* **345**, 153-162.

3.5.2 Other co-solutes

It can be expected that most organic solvents that are miscible with water denature proteins, if the proteins remain soluble as the concentration of the denaturant is increased²³. However organic solvents are less applied to investigate folding mechanism due to limited spectroscopic properties, complex water-denaturant interactions or partial unfolding. In the following results obtained from the determination of end-to-end contact formation rate constants of a (GS)₄-peptide in aqueous 2,2,2-trifluoroethanol (TFE)- and N,N-dimethylformamide (DMF)-mixtures are discussed and compared to rate constants obtained from studies in ethanol⁷³. As shown above end-to-end contact formation in peptides depends strongly on solvent viscosity (Figure 14). Addition of the mentioned denaturants TFE, DMF and ethanol significantly changes solvent viscosity^{134; 136; 137}. To correct the contributions from increased solvent viscosity η apparent end-to-end contact formation rate constants k_c were analyzed using equation 3.2.

Denaturants such as aqueous TFE or neat DMF are good solvents for peptides¹³¹. There is rather an inverse tendency for ethanol-water mixtures, which lead to precipitation of proteins and peptides¹³¹. Indeed, it could be shown that addition of ethanol to aqueous protein solutions reduces the dielectric constant and therefore favours protein aggregation²³.

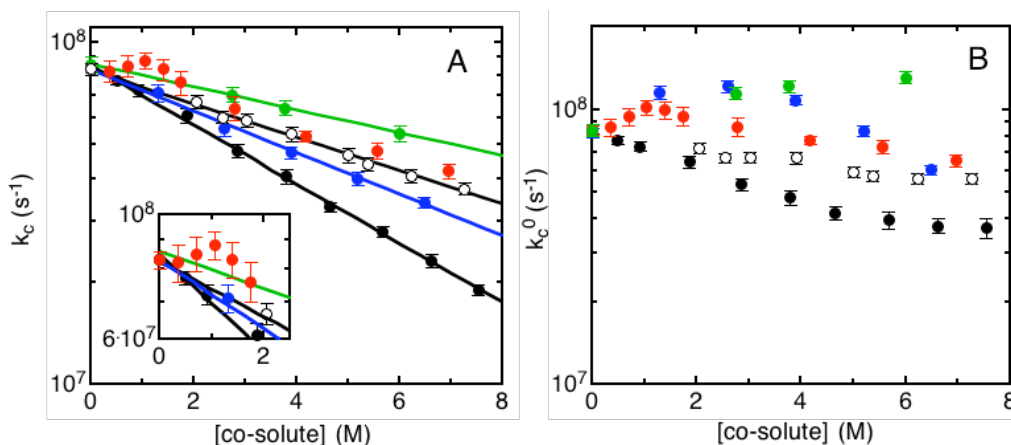


Figure 20: (A) Influence of various co-solutes (GdnHCl (●) urea (○), DMF (●), TFE (●), ethanol (●)) on end-to-end contact formation rate constants in a (GS)₄-peptide. Inset: Behaviour of end-to-end contact formation rate constants at low molarities of co-solutes. (B) Effects of various co-solutes on the viscosity-corrected rate constants of end-to-end contact formation k_c^0 . Rate constants were corrected against viscosity effects using equation 3.2.

As observed for ethanol⁷³ the logarithm of the rate constants of end-to-end contact formation in the (GS)₄-peptide scales linearly with increasing concentration of DMF (Figure 20A). The addition of DMF leads to a smaller change in the rate constants of intrachain contact formation compared to GdnHCl. Interestingly, addition of TFE leads to a completely different behaviour compared to GdnHCl, urea, DMF or ethanol. The logarithm of the observed rate constants does not decrease linearly anymore. Instead a complex relation between rate constant and TFE concentration is observed. With increasing TFE concentration the rate constants first decrease, then increase, reach a maximum around 1.2 M (10% (v/v)) and then decrease again (Figure 20A).

If the rate constants are corrected against solvent viscosity only the effect of urea is similar to that of GdnHCl (Figure 20B). For all other co-solutes the viscosity-corrected rate constants do not decrease monotone anymore, but reveal maxima in the rates of intrachain contact formation with increasing co-solute concentration (Figure 20B). The rate constants are normalized to viscosity suggesting that an additional solvation effect leads to the observed differences in rate constants. A possible origin for the differences

of intrachain dynamics in peptides due to co-solute effects might be changes in the chain dimension. The dimension of a flexible polymer chain depends strongly on the interaction with solvent molecules³². Good solvents increase the chain dimension, whereas poor solvents decrease the average end-to-end distance. The expansion factor α is defined as

$$\alpha = \frac{\sqrt{\langle r^2 \rangle}}{\sqrt{\langle r^2 \rangle_0}} \quad (3.5)$$

and relates the real dimension $\sqrt{\langle r^2 \rangle}$ of a polymer chain to the dimension $\sqrt{\langle r^2 \rangle_0}$ of an ideal polymer chain. The expansion factor α depends on chain length, temperature and solvent quality and adopts the value 1 for θ -conditions. At θ -conditions real polymer chains show apparently properties of ideal chains³². Changes in the average end-to-end distances lead to changes in the rate constant of end-to-end contact formation at “iso-viscous” conditions as shown theoretically⁶¹.

The obtained results on solvent effects suggest that the averaged end-to-end distance is significantly influenced by changes in solvent quality. Addition of low concentrations of ethanol and DMF decreases the solvent quality and might lead to a shrinking of the averaged end-to-end distance and to an increase in the rate constant of intrachain contact formation. With increasing DMF concentration intrachain dynamics are slowed down again, suggesting that now the peptide is better solved leading to an expansion of the chain dimension. The described results are in agreement with data obtained on the solubility of a short model peptide in aqueous ethanol and DMF solutions, which revealed that both co-solutes do not increase significantly the solubility of a model peptide at low co-solute molarities in water¹³¹.

A similar interpretation can be given for the behaviour of TFE. With increasing TFE concentration the viscosity-corrected rate constants for contact formation first increase,

reach a maximum around 1.2 M (10%, (v/v)) and then decrease again indicating that the average end-to-end distance of a peptide becomes smaller at around 10% of aqueous TFE (v/v). Interestingly, an increase of the refolding rate constants at around 10% of aqueous TFE (v/v) is observed in protein folding experiments^{138; 139}. It is suggested that the changes in refolding and unfolding rate constants observed at this TFE concentration are due to general global effects of aqueous TFE on the dimensions of proteins¹³⁹, contradicting that TFE stabilize only defined parts of proteins.

The three organic co-solutes ethanol, DMF and TFE show complex effects on the intrachain dynamics in a model peptide. The results suggest that the average end-to-end distances of the polypeptide chain are significantly changed. Increased rate constants of intrachain contact formation might be explained as a decrease in the average end-to-end distance. However to proof this relation, direct measurements of the average end-to-end distances at chosen conditions are necessary.

3.6 Applying triplet-triplet energy transfer to measure loop formation in natural protein sequences

Protein polypeptide sequences consist of twenty amino acids with different reactivities. To test whether the experimental results³¹ obtained on homo-polypeptides and host-guest studies can be generally applied to dynamics in natural loop sequences we determined the rate constants for end-to-end contact formation for sequences derived from two well characterized proteins, β -parvalbumin (PA) and protein G B1 domain (GB1) (Figure 21A-D).

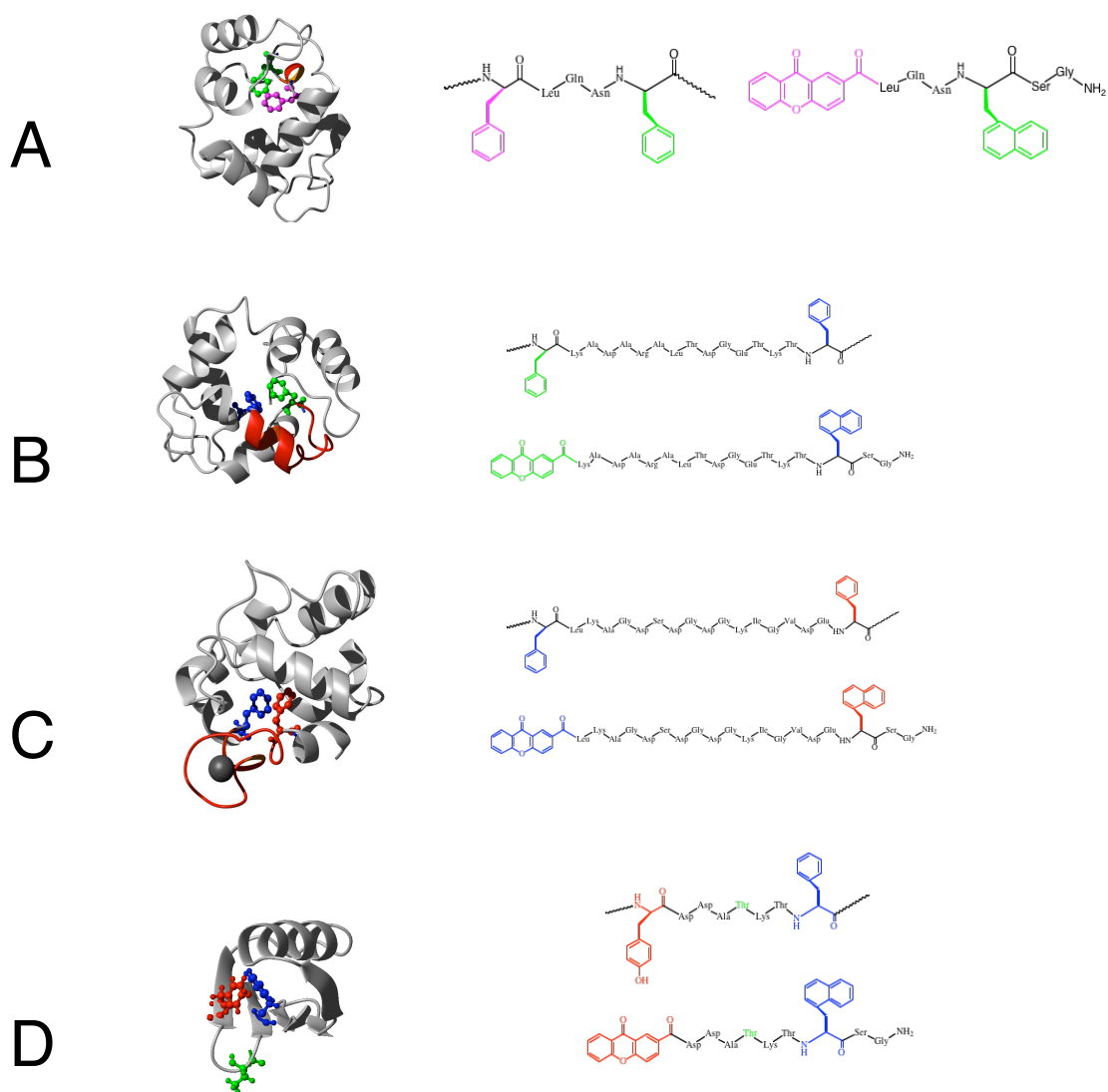


Figure 21: Molecular structure of peptide sequences, derived from carp muscle β -parvalbumin (PA) and protein G B1 domain (GB1). Coloured aromatic amino acids were replaced by triplet donor and acceptor, respectively. (A) PA⁶⁶⁻⁷⁰, (B) PA⁷⁰⁻⁸⁵, (C) PA⁸⁵⁻¹⁰², (D) GB1⁴⁵⁻⁵². Additionally, Thr49 (green) of GB1⁴⁵⁻⁵² fragment was replaced by a glycine and proline (D). The figures were prepared using the program MOLMOL¹⁴⁰ and the pdb files 4CPV¹⁴¹ for the PA sequences and 1GB1¹⁴² for the GB1 sequence.

The chosen sequences of lengths up to 18 amino acids are part of common local folds such as α -helices, β -hairpins or the helix-turn-helix structure. Since the chosen sequences are free of histidine, methionine, tryptophane and tyrosine it is possible to determine absolute rate constants for intrachain contact formation by TTET from Xan to NAla⁸⁵. Loop formation in natural unstructured protein sequences can be well described

by single exponential kinetics like in all other proline-free peptide sequences (Figure 22A). The rate constants of sequences, which contain mainly amino acids with small side-chains or which have a high glycine content like the PA⁸⁵⁻¹⁰² are in good agreement with rate constants obtained from polyserine peptides (Figure 22B).

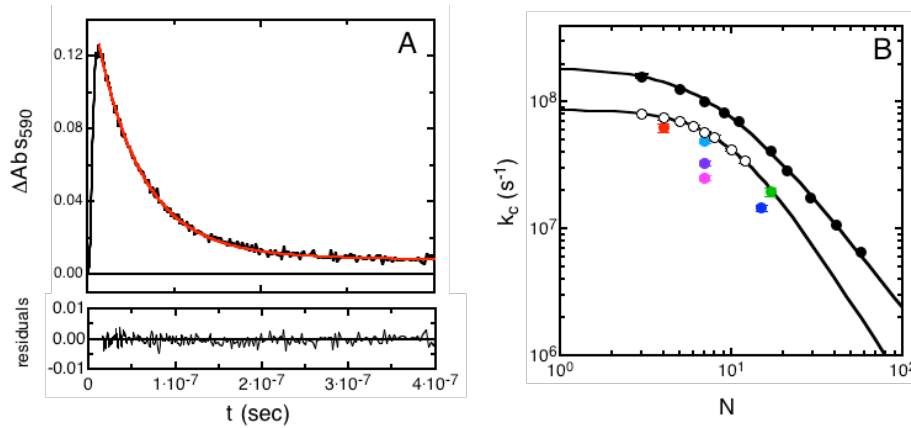


Figure 22: (A) Time-course of formation and decay of xanthone triplets in PA⁸⁵⁻¹⁰² fragment, measured in 10 mM sodium cacodylate, pH 7.0, 22.5 °C (—). End-to-end contact formation can be described by single exponential kinetics with a time constant of 54 ± 3 ns (—). (B) Comparison of the end-to-end contact formation rate constants at 22.5 °C in: poly(glycine-serine) (●), polyserine (○) and various protein fragments derived from β -parvalbumin (PA) and protein G B1 domain (GB1): (●) PA⁶⁶⁻⁷⁰; (●) PA⁷⁰⁻⁸⁵; (●) PA⁸⁵⁻¹⁰²; (●) GB1⁴⁵⁻⁵²; (●) GB1⁴⁵⁻⁵², T49G; (●) GB1⁴⁵⁻⁵², T49P, (solid lines, adapted from ref.³¹).

Intrachain contact formation rate constants of the PA⁷⁰⁻⁸⁵ and the GB1⁴⁵⁻⁵² sequences are significantly slower than rate constants obtained from polyserine peptides. This effect might be explained by the presence of various amino acids with large side-chains, which slow down intrachain dynamics as shown in host-guest studies^{31; 63}. Surprisingly, the PA⁶⁶⁻⁷⁰ sequence shows similar contact formation rate constants and a similar GdnHCl-dependence as the (Ser)₃-peptide within error, although the sequence contains amino acids with larger side-chains (Table 6).

To test whether local chain dynamics of natural loop sequences are influenced due to an incorporation of single glycine and proline residues we substituted Thr49 of the GB1

sequence (45-52) by these two amino acids (Table 6). As shown in figure 22B the incorporation of a glycine residue leads to an increase in the rate constants of end-to-end contact formation as also observed in polyserine peptides of similar length. End-to-end contact formation in the proline containing sequence is slowed down compared to the wild type sequence and can be described by single-exponential kinetics. As mentioned above a single proline induces complex kinetics for chain dynamics over short distances (Figure 12, p.32). Short peptides with a *cis* proline show faster end-to-end chain dynamics than ones with a *trans* proline. The single-exponential behaviour of end-to-end contact formation of the proline containing GB1 fragment suggests that conformations with a *cis* or *trans* proline show virtually identical dynamics as observed for proline containing model peptides with more than 5 peptide bonds (Figure 12).

Table 6: Rate constants and Arrhenius activation parameter of intrachain contact formation in natural unstructured peptides derived from carp muscle β -parvalbumin (PA) and protein G B1 domain (GB1).

Peptide sequence	N	k_c ($10^7 \cdot s^{-1}$) ^b	β^c	E_a (kJ/mol) ^d	A ($10^9 \cdot s^{-1}$) ^e
PA 66-70 ^a	4	6.3 ± 0.4	n.d.	n.d.-	n.d.
PA 70-85	15	1.5 ± 0.1	0.86 ± 0.03	14.3 ± 0.7	4.6 ± 1.9
PA 85-102	17	1.9 ± 0.1	0.85 ± 0.03	12.5 ± 0.7	2.7 ± 0.3
PA 85-102 (D/N,E/Q)	17	1.9 ± 0.1	0.84 ± 0.03	12.5 ± 0.7	2.7 ± 0.3
GB1 45-52 wt	7	3.2 ± 0.2	0.90 ± 0.05	15.7 ± 0.9	18 ± 7
GB1 45-52 T/P	7	2.4 ± 0.2	0.90 ± 0.05	15.7 ± 0.8	14 ± 4
GB1 45-52 T/G	7	5.0 ± 0.3	0.90 ± 0.04	15.5 ± 0.7	24 ± 7

a: is not soluble in water. The rate constant was determined from a GdnHCl dependence. The m_c -value is 380 ± 35 J/(mol·M) for GdnHCl.

b: determined in 10 mM potassium phosphate, pH 7.0, 22.5°C.

c: Fractional exponent obtained from viscosity dependencies (equation 3.2). Glycerol was used as viscous co-solute.

d: Arrhenius activation energy were determined from viscosity-corrected rate constants (equation 3.2).

e: Arrhenius pre-exponential factor, normalized to solvent viscosity of $\eta_0 = 1.0$ cP.

Viscosity and temperature dependencies revealed that intrachain diffusion in natural unstructured protein sequences is significantly stronger influenced by chain stiffness compared to flexible homo-polypeptides. For all investigated natural sequences apparent Arrhenius activation energies of around 12-15 kJ/mol were observed (Table 6). The

higher activation barrier might explain the smaller rate constants for intrachain dynamics in natural loop sequences compared to the data obtained from homopolypeptides. A possible explanation might be that the conformational space in natural peptide sequences is more restricted due to the larger amino acid side-chains. The results suggest that the earliest steps in protein folding are significantly limited by activation barriers.

All chosen protein fragments consist of charged amino acids except the PA⁶⁶⁻⁷⁰ sequence. Particularly the PA⁸⁵⁻¹⁰² sequence contains an alternating part of negative side-chains and uncharged residues between residue 89 and 96 (GDSDGDG) that represents the calcium-binding site of the EF-hand structure (Figure 21C). From polyelectrolytes, like poly(styrene sulfonate) or DNA at neutral conditions it is known that the backbone rigidity is supported by Coulomb repulsion of the side-chains since same charges distributed along a chain at high density repel each other to extend the otherwise flexible chain¹⁴³. To test if charges influence intrachain diffusion in natural peptide sequences we replaced the negatively charged amino acids Asp and Glu of the PA⁸⁵⁻¹⁰² sequence by the neutral amino acids Asn and Gln. Both sequences show the same end-to-end contact formation rate constants and Arrhenius activation parameter suggesting that charges in the chosen natural sequences do not influence chain dynamics (Table 6).

Parts of these results are described in detail in the following manuscripts:

“Intrachain diffusion in a protein loop fragment from carp parvalbumin.” Florian Krieger, Beat Fierz, Fabian Axthelm, Karin Joder, Dominique Meyer and Thomas Kiefhaber (2004) *Chem. Phys.* **307**, 209-215.

“Kinetics and Barriers of Loop Formation in Natural Protein Sequences.” Florian Krieger & Thomas Kiefhaber. To be submitted.

3.7 Intrachain diffusion of α -helices in helix-turn-helix motifs

In 1976 Karplus and Weaver suggested the diffusion-collision model²⁵ as a possible mechanism for protein folding²⁷. In this model the formed secondary structures have to interact by chain diffusion processes. It is assumed that the formation of α -helices or β -sheets on a polypeptide chain is rapid relative to the rate-limiting steps in protein folding. The diffusion-collision model reduces the dynamics of the folding process from a consideration of the properties of the random coil chain to that of the properties of secondarily structured parts of the protein and their interactions²⁷.

We wanted to test, how formation of secondary structures influences the rate constants for intrachain contact formation. We chose two different approaches to determine contact formation rate constants of putatively secondary structured parts in polypeptides by TTET: first, polypeptide sequences derived from the EF-hand of carp muscle β -parvalbumin and second, de-novo designed helix-turn-helix sequences.

3.7.1 Induction of secondary structure in the Ca^{2+} -binding EF-hand motif from carp muscle β -parvalbumin

Parvalbumin belongs to the family of EF-hand proteins and exhibits a very high affinity for Ca^{2+} and moderate affinity for Mg^{2+} , properties consistent with a Ca^{2+} -buffering role¹⁴⁴. The overall structure consists of six helices associated in three helix-turn-helix motifs, AB, CD and EF. The loops between the helices C and D and between the helices E and F bind one calcium ion each¹⁴¹. The parvalbumin helices are all amphiphilic, forming the core of the structure with their hydrophobic faces. 80 % of the hydrophobic side-chains buried in the protein core are located in helices.

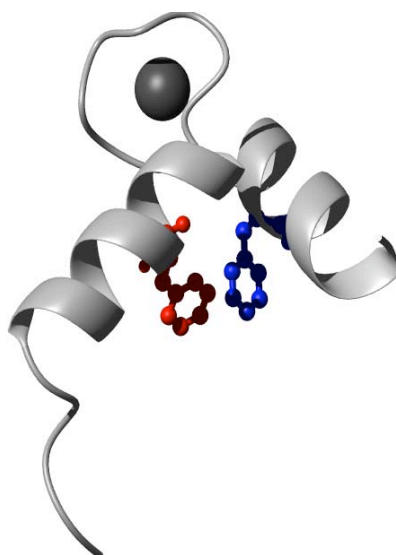


Figure 23: Ribbon diagram of the isolated EF-hand (PA⁷⁴⁻¹⁰⁸) of the carp muscle β -parvalbumin structure used in this study. The phenylalanines Phe85 (red) and Phe102 (blue) shown as ball-and-stick models have been replaced by the triplet donor and acceptor labels, xanthone acid and naphthylalanine, respectively. The bound calcium ion is plotted as a black ball. The figure was prepared using the program MOLMOL¹⁴⁰ and the pdb file 4CPV¹⁴¹.

First synthesis of a fragment of the EF-hand from carp muscle β -parvalbumin (PA⁸⁵⁻¹⁰²) labelled with triplet chromophores showed that the chosen peptide is unstructured, even in presence of high calcium concentrations. The short length of the chosen peptide fragment might be an origin for the missing structure. Therefore we synthesized the complete EF-hand with tail elongation on both sides of the triplet chromophores (Figure 23, Table 7). However, sequence elongation on both sides of the triplet chromophores does not cause any structure in presence of 100 mM CaCl₂ and also in absence of calcium ions, as determined by circular dichroism (Figure 24A). It is quite controversial whether the isolated EF-hand fragment of carp parvalbumin is structured and is able to specifically bind calcium ions. Earlier investigations showed that a 33-residue fragment (residues 76-108) of parvalbumin has a calcium binding constant of 300 (1/M)¹⁴⁵. However this binding constant is only seven times higher than the calcium binding constant of glutamic acid¹⁴⁶. Another work suggested that Arg-75 plays a critical role in

the structural organization and calcium binding activity of the molecule and that the isolated EF-hand fragment is unstructured and does not bind calcium¹⁴⁷. The structure and stability of carp muscle parvalbumin is relatively sensitive to deletion mutations. Deletion of several amino acids of the C-terminal part of the sequence (residue 103 - 108) has a dramatic effect on parvalbumin structure and stability¹⁴⁸.

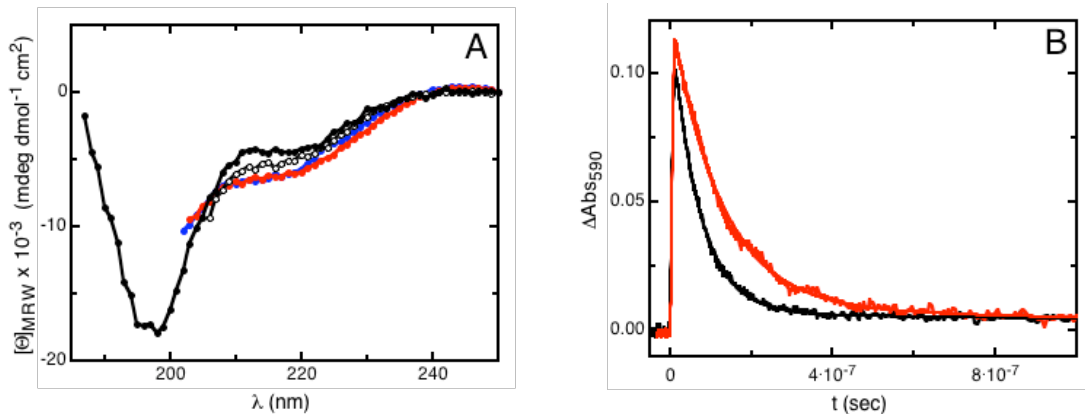


Figure 24: (A) CD spectra of PA⁸⁵⁻¹⁰⁸ (●), (●) and PA⁷⁴⁻¹⁰⁸ (○), (●) in presence of 100 mM CaCl₂ and without calcium, respectively. (B) Time course of formation and decay of xanthone triplets in PA⁸⁵⁻¹⁰⁸ (—) and PA⁷⁴⁻¹⁰⁸ (—). All data were recorded in 10 mM sodium cacodylate, pH 7.0 at 22.5°C, except the CD-spectra for PA⁷⁴⁻¹⁰⁸ (●) in absence of Ca²⁺, which was recorded in 10 mM potassium phosphate, pH 7.0.

The determination of intrachain dynamics showed that intrachain contact formation in all sequences derived from the parvalbumin EF-hand structure is well described by single exponential kinetics (Figure 24B). The observed rate constants decrease slightly with increasing tail length on both ends (Table 7) as observed for homo-polypeptide chains⁶³ (B.F. & T.K., unpublished data). Addition of calcium ions shows no effect on intrachain dynamics of the chosen sequences, suggesting that no structure is formed and no calcium ions are bound selectively. The synthesis of longer peptide fragments, like the PA sequence (residues 66-108) failed due to limitations in peptide synthesis.

Table 7: Sequences and intrachain diffusion rate constants of the investigated polypeptides from in carp muscle β -parvalbumin (PA). In the case of PA⁷⁴⁻¹⁰⁸ xanthone acid was introduced in the sequence via side-chain modification of the non-natural amino acid L- α,β -di-amino propionic acid (Dpr).

Peptide sequence	Amino acid sequence	Number of tail residues	k_c (s ⁻¹) ^a
PA ⁸⁵⁻¹⁰²	Xan-LKAGDSDGDGKIGVDE-NAla-SG-NH ₂	2	$1.8 \cdot 10^7$
PA ⁸⁵⁻¹⁰⁸	Xan-LKAGDSDGDGKIGVDE-NAla-TALVKA-COOH	6	$1.5 \cdot 10^7$
PA ⁷⁴⁻¹⁰⁸	Ac-ARALTDGETKT-Dpr(Xan)-LKAG-DSDGDGKIGVDE-NAla-TALVKA-COOH	11 + 6	$8.2 \cdot 10^6$

a: measured in 10 mM sodium cacodylate, pH 7.0, at 22.5°C.

3.7.2 Design of peptides with de novo helix-turn-helix motifs

The results obtained on the EF-hand fragment of parvalbumin showed that the incorporation of the triplet chromophores xanthone acid and naphthylalanine into structured natural sequences derived from protein structures failed, leading to structureless peptide fragments. To obtain peptides with significant secondary structure we decided to design a stable helix-turn-helix structure, based on polyalanine peptides. Marqusee and Baldwin described the first de novo-designed α -helical peptide system in water^{149; 150}. In this system the peptides mostly consist of alanine. Glutamic acid and lysine or arginine are inserted for solubility at the positions $i, i + 4$, which show stabilizing interactions (Table 8)⁴⁸.

We used such alanine-based peptides to introduce the triplet chromophores xanthone acid and naphthylalanine into helical peptide chains. Several sequences are given in table 8. Salt bridges were introduced in the sequences to stabilize the peptides. As shown in figure 25A the positions 2 and 6, 7 and 11, and 12 and 16 were chosen for stabilizing Glu-Lys salt bridges. The triplet labels were attached in the centre of the sequences at the position 8 or 9.

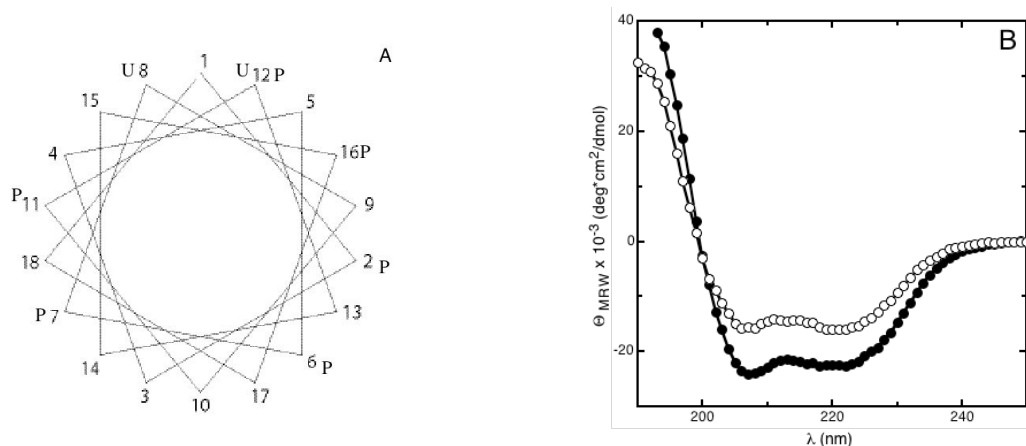


Figure 25: (A) Helical wheel to design an α -helix based on polyalanine peptide chains. P indicates preferential positions for charged amino acids. U indicates position, which lead to peptide aggregation. (B) Circular dichroic (CD) spectra of Ac-AE(A)₃RE-NAla-(A)₂RE(A)₃R-NH₂ (●) and Ac-AE(A)₃RE-A-Dpr(Xan)-ARE(A)₃R-NH₂ (○) measured in 5 mM KP_i, pH 7.0 ± 0.1 at 5 °C.

Table 8: Sequences and Θ_{MRW} at 222 nm of the investigated peptides as determined by CD. Xan: xanthone acid, NAla: naphthylalanine, Dpr: diamino-propionic acid. CD measurements were performed out in 5 mM KP_i, pH 7.0 ± 0.1 at 5 °C.

Sequence	$[\Theta]_{MRW,222nm}$ (deg·cm ² /dmol)	% helical content ^a	structure
Ac-AE(A) ₃ KEA-NAla-AKE(A) ₃ K-NH ₂	-13200	39.1	α -helix
Ac-AE(A) ₃ KE-NAla-(A) ₂ KE(A) ₃ K-NH ₂	-11500	34.1	α -helix
Ac-AE(A) ₃ RE-NAla-(A) ₂ RE(A) ₃ R-NH ₂	-20100	59.6	α -helix
Ac-AE(A) ₃ RE-A-Dpr(Xan)-ARE(A) ₃ R-NH ₂	-16100	47.7	α -helix
Ac-AE(A) ₃ KAE-NAla-(A) ₂ KAE(A) ₃ K-NH ₂	-7600	-	α -helix/ β -sheet ^b

a: Experimentat fractional content were determined from $[\Theta]_{MRW,222nm}$ values using - 40000·(1-2.5/n) and 0 deg·cm²/dmol as the values for 100 and 0% helix, respectively; n is the number of amino acid residues in the peptide^{151; 152}.

b: aggregation at high concentration and reduced solubility.

A first goal was to find conditions, at which the helical peptides are stable even at high peptide concentration. For several peptides with charges on position 2 and 6, 7 and 11, and 12 and 16 we could induce helical structure as indicated by circular dichroic measurements (Figure 25B). The helical content is concentration independent and

depends on the positions of the triplet chromophores and of the salt bridges. The helical content of the peptides is estimated to be between 34 and 60% at 5°C (Table 8).

As preliminary result the synthesis and a first characterization of a designed helix-turn-helix peptide of the sequence:

Ac-AE(A)₃RE-NAla-(A)₂RE(A)₃R-(GS)₅G-AE(A)₃RE-A-Dpr(Xan)ARE(A)₃R-NH₂ could be performed consisting of two segments with high helical propensity linked by a (GS)₅G-loop. The synthesis yield of this peptide was very low, however some amount of this peptide could be purified. Circular dichroic measurements revealed that this peptide has a helical content of 23% at 5°C (Figure 26A). The helical content is significantly lower than in the separated α -helices. This behaviour is expected since the unstructured part of the peptide contributes additionally to the CD signal. The temperature transition is reversible and at high temperature the peptide is fully unfolded (Figure 26A, B).

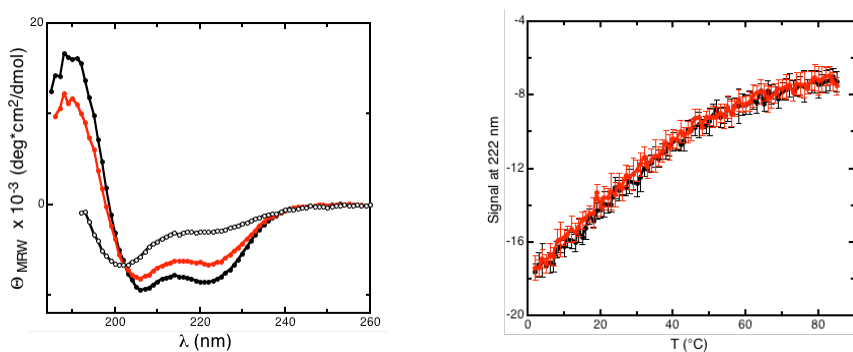


Figure 26: CD spectra (A) and temperature transition (B) of Ac-AE(A)₃RE-NAla-(A)₂RE(A)₃R-(GS)₅G-AE(A)₃RE-A-Dpr(Xan)-ARE(A)₃R-NH₂. CD spectra were recorded at 5°C (●), 22.5°C (●) and 85°C (○) in 5 mM KP_i, pH 7.0. The temperature transitions were recorded from 3 - 85°C (—) and from 85 - 3°C (---) in the same buffer in a cuvette with the layer thickness $d = 2\text{mm}$. The peptide concentration was 20.5 μM in all measurements.

However, it is not possible to estimate the helical content on both ends of the peptide. Furthermore the data give no insight if the two putative helices interact with each other. Both α -helices of the helix-turn-helix peptide contain several positive and negative

charges and therefore it cannot be ruled out that both helices interact due to electrostatic interactions. A decrease of the pH might give more insight in this problem since at acidic conditions both α -helices are only positively charged.

The determination of intrachain dynamics revealed that the intrachain contact formation kinetics of the helix-turn-helix peptide are more complex compared to the kinetics of intrachain contact formation in a long unstructured polypeptide (Figure 27A). For long unstructured polypeptides intrachain contact formation is well described by single-exponential kinetics (Figure 27A). For intrachain contact formation in the putative helix-turn-helix peptide a second kinetic phase with a time constant of $\tau_2 = 312$ ns is observed in addition to the main kinetic phase with a time constant of $\tau_3 = 115$ ns.

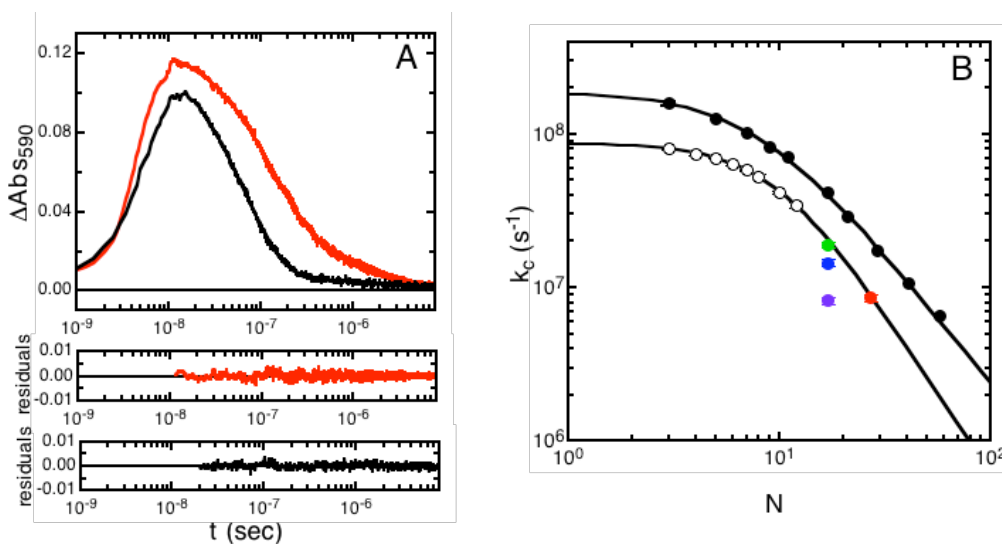


Figure 27: (A) Time course of formation and decay of xanthone triplets in the helix-turn-helix peptide (—) in 5 mM potassium phosphate, pH 7.0 at 22.5°C and in comparison with the unstructured PA^{85-108} -peptide (—). Intrachain contact formation of the PA^{85-108} -peptide is described by single-exponential kinetics with a rate constant of $1.5 \cdot 10^7 \text{ s}^{-1}$. Intrachain contact formation in the helix-turn-helix fragment is described by a double exponential fit with the following parameters: $A_1=0.026$, $k_1=3.2 \cdot 10^6 \text{ s}^{-1}$, $A_2=0.078$, $k_2=8.6 \cdot 10^6 \text{ s}^{-1}$. (B) Comparison of the end-to-end contact formation rate constants at 22.5°C in: poly(glycine-serine) (●), polyserine (○) and various protein fragments derived from β -parvalbumin PA^{85-102} (●), PA^{85-108} (●), PA^{74-108} (●) and the helix-turn-helix peptide HLH (●), solid lines, adapted from ref.³¹

Interestingly, the observed rate constant of the main phase agrees quite well with rate constants predicted for long polyserine peptides suggesting that the main phase describes intrachain contact formation of the unfolded polypeptide (Figure 27B). However, it is not clear, how to interpret the observed rate constants, since the complete kinetics of intrachain contact formation might be influenced due to a coupling of other conformational changes in the structured peptide. Due to the low synthesis yield of the helix-turn-helix peptide it was not possible to vary the experimental conditions in order to test the effect on the rate constants and amplitudes of the kinetic phases. For a better understanding of the mechanism of intrachain contact formation in the helix-turn-helix peptide an accurate kinetic investigation of intrachain diffusion is necessary.

3.7.3 Position effect of charges on the structure of alanine based peptides

During the design of helical peptides, we observed that the position of charges critically determines the structure of peptides. Interestingly, if charges are placed on the positions 2 and 6, 8 and 12, and 14 and 18 the peptide exhibits no helical but β -sheet structure formation and strong aggregation of the peptide is observed (Table 8). The incorporation of one alanine residue between two consecutive Glu-Lys salt bridges might change the topology of the peptide, leading to an alternation of charged residues. Every second and fourth amino acid of the sequence contains now a charge at neutral pH (underlined, ...KAEAAKAEAAK...). Similar results are also described for alanine-based peptides with alternating charges of the different moduli¹⁵³: (I) -A+A-A+A..., (II) -A-A+A+A..., (IV) -A-A-A-A... or +A+A+A+A... . It could be shown that the peptides have the ability to form β -sheets and α -helical structures, although the used amino acids have a high propensity to form α -helices¹⁵⁴. The authors concluded that β -sheet formation in the alanine based peptides is due to the architecture of the

peptide, in which the repetitive and charged residues form intermolecular ionic bonds on one side of the β -sheet, whereas the alanine residues on the opposite hydrophobic side form intermolecular van-der-Waals interactions. The results suggest that the alternating position of charges leads to the formation of β -sheets.

Table 9: Sequences of the investigated alanine based peptides.

Name	Sequence
KE	Ac-AEAAAKEAAAKEAAAKGY-NH ₂
KAE	Ac-AEAAK <u>A</u> EA <u>A</u> KA <u>A</u> EA <u>A</u> AKGY-NH ₂
KAAE	Ac-AEAAK <u>A</u> <u>A</u> EA <u>A</u> KA <u>A</u> <u>A</u> EA <u>A</u> AKGY-NH ₂

Underlined: alanine residues, which were placed between $i, i + 4$ side-chain interactions of the oppositely charged Glu and Lys. A tyrosine was attached at the C-terminus to determine the peptide concentration used in the experiment.

To investigate the β -sheet formation of alanine based peptides more accurately we synthesized peptides of the sequences, given in table 9. The chosen peptide sequences differ only in the configuration of the salt bridges. The peptide sequence KE is based on published work according to Marqusee and Baldwin^{149; 150} and forms α -helical structures at pH 2.5 and pH 7.0 (Table 10). The helical content at pH 2.5 is slightly decreased compared to pH 7.0 suggesting that at neutral pH the negatively charged glutamic acid interacts with positively charged lysine¹⁵⁰. Incorporation of an additional alanine between the $i, i + 4$ salt bridges E-K, as shown in table 9 (underlined), changes the properties of the peptide dramatically (Table 10). At pH 7 the peptide is virtually insoluble and strong aggregation is observed. Lowering the pH to 2.5 leads to formation of helical structure at low peptide concentration ($c < 100 \mu\text{M}$) and of β -sheet-structure at high peptide concentration ($c > 1 \text{ mM}$). Indeed, circular dichroic measurements reveal that the alanine peptide KAE exists in β -sheet-structure at pH 2.5 at high concentration, as found in earlier studies on similar peptides¹⁵³.

Incorporation of a second additional alanine residue between the salt bridges induces again α -helical structures at pH 7 and pH 2.5 in the KAAE-peptide. The solubility of

KAAE-peptide is elevated compared to the KAE-peptide. The results suggest that the position of the side-chains plays a critical role in determining helical content of alanine based peptides.

Table 10: Solubility and helix content of alanine based peptides used in this study.

Name	Solubility ^a	Structure	$\Theta_{\text{MRW},222}$ (deg•cm ² /dmol) ^b	Helical content (%) ^c	pH ^b
KE	++	α -helical	-24400	70.8	2.5 \pm 0.1
	++	α -helical	-26500	76.9	7.0 \pm 0.1
KAE	+/-	β -sheet	-19200 ^d	-	2.5 \pm 0.1
	-	aggregation	n.d.	-	7.0 \pm 0.1
KAAE	+	α -helical	-27000	76.1	2.5 \pm 0.1
	++	α -helical	-29000	81.8	7.0 \pm 0.1

a: as determined quantitatively in UV/VIS measurements. Solvation of the KAE-peptide at pH 7 induces strong aggregation.

b: Θ_{MRW} measured at 222 nm in 5 mM potassium phosphate at 5°C.

c: fractional content were determined experimentally from $[\Theta]_{\text{MRW},222\text{nm}}$ values using $-40000 \cdot (1-2.5/n)$ and 0 deg•cm²/dmol as the values for 100 and 0% helix, respectively; n is the number of amino acid residues in the peptide^{151; 152}.

d: an extinction coefficient of $\epsilon_{275}=1390 \text{ (M}\cdot\text{cm)}^{-1}$ was assumed¹⁵⁵.

The KAE peptide shows quite different spectroscopic properties at pH 2.5 compared to both other peptides, KE and KAAE. At high concentrations a red shifted band appears in the UV/VIS spectra (Figure 28A). It is suggested that this band might originate from interactions between tyrosines, such as charge transfer (CT) complex formation due to an intermolecular interaction between two or more peptide chains. Most absorption and emission processes involve just one molecule. In some cases two or more molecules may participate in cooperative absorption or emission. This means the change in absorption or emission can only be understood as arising from oligomeric structures⁷⁶. The formation of CT complexes results in the appearance of new electronic absorption bands, which are not observed in the separated compounds¹⁵⁵. Generally the new absorption band is broad and devoid of vibrational structure. This width is explained by the fact that the rather small binding energy of CT complexes allows many different

structural configurations of the complex to exist in equilibrium with one another⁷⁶. The absorption energy for each configuration will differ and cause a broadening of the band.

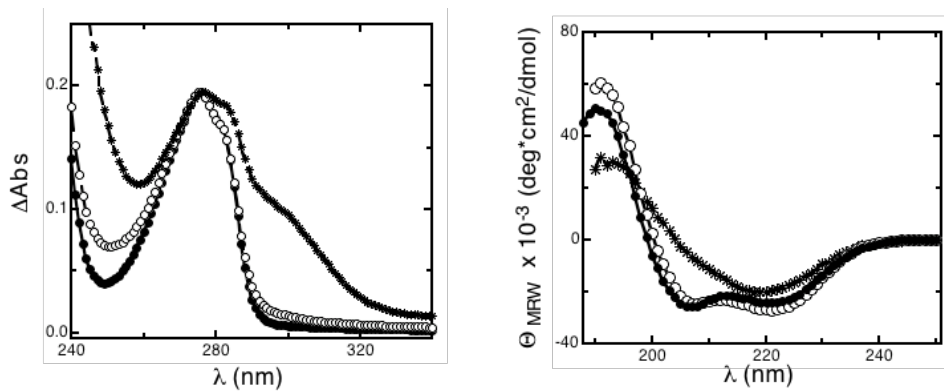


Figure 28: UV/VIS-spectra (A) and CD-spectra (B) of KE (●), KAE (*) and KAAE (○) peptides measured in 5 mM potassium phosphate, pH 2.5. The concentration of the stock solution of the peptides was more than 1mM. UV/VIS spectra were recorded after centrifugation of the peptide solution. UV.VIS spectra normalized at 275 nm.

The molecular origin of the different properties of the KAE-peptide compared to the other peptides is not known. The spectroscopic characterization of this peptide is quite complicated, since aggregation is observed at various conditions. The poor solubility of the peptide in water might originate from the configuration of the charges in the peptide. A main problem in studying aggregation phenomena of synthetic peptides might be the synthesis and purification of such peptides. Due to the poor solubility it cannot be ruled out that during the synthesis and purification aggregation already started, leading to irreversible denatured aggregates.

Materials and Methods

Xanthone acid synthesis

9-Oxoxanthen-2-carboxylic acid synthesis was performed as described previously³¹.

Synthesis of Fmoc-Gly-d₂

C_α- deuterium labelled glycine (Gly-d₂) was from Cambridge isotope Laboratories. The synthesis of Fmoc-Gly-d₂ was carried out as described¹⁵⁶ and purified on a C18-column by HPLC.

Peptide synthesis

All Dpr-free peptides (Dpr: diaminopropionic acid) were synthesized on acid labile Trityl-resins (Rapp-polymer) or Rink-amid resins (Novabiochem) and purified using standard Fmoc-chemistry¹⁵⁷. C-terminal blocked, Dpr containing peptides were synthesized on acid-stable Rink-amide resins (Novabiochem) with Dpr(Mtt) (Mtt: 4-Methyl-trityl, Novabiochem). The hyperacid sensitive Mtt protecting group was deprotected selectively in 1% TFA in dichloromethane (v/v). Xanthone acid was attached on the free amino group of the diaminopropionic acid with standard Fmoc-chemistry. Dpr containing peptides with a free C-terminus were synthesized on a acid-labile Trityl-resins with Dpr(ivDde) (ivDde: 4,4-dimethyl-2,6-dioxocyclohex-1-ylidene-3-methylbutyl, Novabiochem). The base sensitive ivDde protecting group was removed selectively in 2% hydrazine in dimethylformamide (v/v). Afterwards xanthone acid was attached as described above.

Laser flash experiments

Laser flash experiments were carried out as described previously³¹. The determination of end-to-end contact formation rate constants was performed at temperatures between 5 and 45°C. The temperature was directly measured in the cuvette.

Determination of solvent viscosity

Ethylene glycol, glycerol and glucose were from Aldrich, Sucrose and PEG 1500, 6000 and 20000 were from Merck. The viscosity was varied adding different polyolic co-solutes to the solutions. For all solutions, the viscosity was measured by using a falling ball viscometer (Haake) at the given temperature.

Measurements with DMF were carried out up to a DMF concentration of 6.5 M, which corresponds to 50 % aqueous DMF solution (v/v). Higher DMF concentrations do not allow measurements of intrachain dynamics with TTET, since xanthone triplets are quenched by DMF and triplet lifetimes of xanthone in pure DMF become smaller than 20 ns. Tabulated values for the viscosities of water⁸³, TFE¹³⁶, EtOH⁸³, DMF¹³⁷, urea¹³⁴ and GdnHCl¹³⁴ were used to correct observed end-to-end contact formation rate constants against solvent viscosity effects, accordingly equation 3.2.

Recording of the CD-spectra

After purification all peptides were dried by lyophilization. Afterwards the solid dried peptides were spectroscopically characterized. Aggregating peptides were solved in buffer solutions as described and centrifuged at 14000 rpm for 10 min or filtered before recording UV/VIS-spectra. CD-spectra were recorded as reported previously⁸⁵.

Determination of the rate constant of *cis-trans* peptidyl-prolyl bond isomerization

To determine the *cis-trans* peptidyl-prolyl bond isomerization rate constants of the model peptide Succinyl-Ala-Ala-Pro-Phe-pNA (pNA: para-nitroaniline) the chymotrypsin coupled assay was used, first described by Fischer¹⁵⁸. All kinetic measurements were performed on an Applied Photophysics SX.18MV instrument at $22.5 \pm 0.1^\circ\text{C}$. The absorption of the released p-NA was monitored at 385 nm. All measurements were carried out in 0.035 M 4-(2-hydroxyethyl)-piperazine-1-ethane sulfonic acid (HEPES) pH 8.5. Glycerol was added to increase the solvent viscosity. Typically, a stock solution of the peptide in dimethylsulfoxide (8mM) was prepared and diluted ten times in the reaction buffer. The diluted substrate and enzyme solutions (25mg/ml α -chymotrypsin) were mixed in a ratio of 1 + 10. Under all conditions the cleavage rate of the α -chymotrypsin catalyzed proteolysis was faster than the *cis-trans* peptidyl-prolyl bond isomerization.

4 Summary

This thesis deals with the dynamics of unfolded polypeptide chains as model for the earliest steps in protein folding. Starting from an ensemble of unfolded conformations a folding polypeptide chain has to form specific backbone and side-chain interactions to reach the native state. The rate at which two defined contacts are formed on a polypeptide chain is limited by intrachain diffusion. The characterization of rate constants of intrachain contact formation in polypeptides and their dependence on length, sequence and solvent effects give new insights for an understanding of the dynamics of the earliest steps in protein folding.

Until recently, little was known about absolute time scales of intramolecular contact formation in polypeptide chains. Direct measurements of fast intramolecular diffusion processes became possible with the development of fast diffusion-controlled electron transfer processes. In the presented work triplet-triplet energy transfer was used to characterize intrachain contact formation in homo-polypeptides and peptide fragments derived from natural protein sequences.

The transfer of triplet electrons between the triplet donor xanthone to the triplet acceptor naphthalene is diffusion-controlled as tested by measuring its temperature and viscosity dependencies. The results suggest that triplet-triplet energy transfer from xanthone to naphthalene provides the requirement to determine absolute intramolecular contact formation rate constants in polypeptide chains.

Intrachain contact formation in unstructured polypeptides is well described as a single exponential process. The loop-size dependence of the rate constants of intrachain contact formation revealed that intrachain motions over short and long distances are limited by different rate-limiting steps. In short peptide chains end-to-end contact formation is with a minimum time constant of 5-10 ns virtually independent of chain

length and limited by an activation barrier of 12-16 kJ/mol. In long flexible poly(glycine-serine) peptide chains with more than twenty peptide bonds N the rate constants decrease with $N^{-1.7 \pm 0.1}$ and end-to-end contact formation becomes nearly completely entropy-driven.

Glycine and proline residues significantly change local intrachain dynamics compared to all other amino acids. Glycine accelerates contact formation whereas short proline containing peptides reveal complex kinetics of contact formation. Local chain dynamics are accelerated by a *cis* and slowed down by a *trans* peptidyl-prolyl bond. The effects vanish in peptide chains if the sequence contains more than five amino acids on each side of a single glycine or a single proline residue.

The dynamics of loop formation are sensitive to the nature of the solvent. Good solvents, such as denaturants slow down intrachain dynamics compared to water. The effect of solvent composition on chain dynamics indicates that the chain properties of polypeptides strongly depend on the surrounding conditions.

Natural protein sequences are more complex than homo-polypeptide chains because they consist of 20 different amino acids. We determined the dynamics of loop formation in sequences derived from two proteins, carp muscle β -parvalbumin and protein G B1 domain. Compared to homo-polypeptides the intrachain dynamics in natural loop sequences are slowed down and higher activation barriers are determined. The results suggest that the dynamics of the earliest steps in protein folding are limited by significant activation barriers.

The results allow us to estimate an upper time scale for rates of contact formation in unstructured peptide chains. In glycine-rich sequences, which are often found in β -hairpins and turns a first contact over 3-4 peptide bonds will be formed within 10-15 ns. For glycine-free sequences local contact formation is slowed down to 15-50 ns

depending on the sequence. Due to the strong distance dependence of the rate constant of the end-to-end contact formation long-range interactions on an unfolded polypeptide chain over 50-60 peptide bonds will not be formed faster than in 500 ns.

5 Acknowledgements

This work was completed from January 2001 to November 2004 in the laboratory of Prof. Dr. Thomas Kiefhaber in the Department of Biophysical Chemistry at the Biozentrum of the University of Basel, Switzerland.

A **MERCI VIEL MOL** to all people, who contributed to this work!

First of all, I would like to thank Thomas Kiefhaber for guiding me through the exciting field of peptide dynamics and protein folding. I have learned a lot from his excellent scientific approach.

Many thanks to the former and current colleagues of the TK lab for ingenious discussions and the exceptionally nice atmosphere in and outside the lab, particularly thanks to Andy, Beat, Christophe and Stefan. Further I want to thank Gernot, Hans and Leo from the floor service group for the small and big facilities during my PhD thesis.

Many thanks to Joseph Wey for technical support in xanثone acid synthesis and thanks to Dr. Paul Jenö and Thierry Mini for mass spectrometry support.

I'm particularly grateful to Annett sharing with me the beauties of life.

Many thanks to the Swiss mountains showing me how barriers may look like.

...and special thanks to the NMR-spectrometer being outside of the BIOZENTRUM.

6 References

1. Anson, M. L. & Mirsky, A. E. (1934). The Equilibrium Between Active Native Trypsin and Inactive Denatured Trypsin. *J. Gen. Physiol.* 17, 393-408.
2. Neurath, H., Greenstein, J. P., Putnam, F. W. & Erickson, J. O. (1944). The chemistry of protein denaturation. *Chem. Rev.* 34, 157-265.
3. Mirsky, A. E. & Pauling, L. (1936). On the structure of native, denatured and coagulated proteins. *Proc. Natl. Acad. Sci. U S A* 22, 439-447.
4. Kendrew, J. C., Dickerson, R. E. & Strandberg, B. E. (1960). Structure of myoglobin: a three-dimensional Fourier synthesis at 2 Å resolution. *Nature* 185, 422-427.
5. Blake, C. C., Koenig, D. F., Mair, G. A., North, A. C., Phillips, D. C. & Sarma, V. R. (1965). Structure of hen egg-white Isozyme. A three-dimensional Fourier synthesis at 2 Å resolution. *Nature* 206, 757-763.
6. Anfinsen, C. B., Cooke, J. P. & Sela, M. (1962). Reversible reduction of disulfide bonds in polyalanyl ribonuclease. *J. Biol. Chem.* 237, 1825-1838.
7. Anfinsen, C. B. (1973). Principles that Govern the Folding of Protein Chains. *Science* 181, 223-230.
8. Levinthal, C. (1968). Are there pathways for protein folding? *J. Chim. Phys.* 65, 44.
9. Levinthal, C. (1969). *Mössbauer Spectroscopy in Biological Systems*, Allerton House, Monticello, Ill.
10. Jackson, S. E. (1998). How do small single-domain proteins fold? *Folding & Design* 3, R81-R91.
11. Zwanzig, R. (1997). Two-state models for protein folding. *Proc. Natl. Acad. Sci. USA* 94, 148-150.
12. Bieri, O., Wirz, J., Hellrung, B., Schutkowski, M., Drewello, M. & Kiefhaber, T. (1999). The speed limit for protein folding measured by triplet-triplet energy transfer. *Proc. Natl. Acad. Sci. USA* 96, 9597-9601.
13. Dunker, A. K., Lawson, J. D., Brown, C. J., Williams, R. M., Romero, P., Oh, J. S., Oldfield, C. J., Campen, A. M., Ratliff, C. M., Hipps, K. W., Ausio, J., Nissen, M. S., Reeves, R., Kang, C. H., Kissinger, C. R., Bailey, R. W., Griswold, M. D., Chiu, W., Garner, E. C. & Obradovic, Z. (2001). Intrinsically disordered protein. *J. Mol. Graphics Modell.* 19, 26-59.
14. Gutfreund, H. (1995). *Kinetics for the life sciences*, Cambridge University Press, Cambridge.
15. Ansari, A., Berendzen, J., Bowne, S. F., Frauenfelder, H., Iben, I. E. T., Sauke, T. B., Shyamsunder, E. & Young, R. D. (1985). Protein states and proteinquakes. *Proc. Natl. Acad. Sci. U S A* 82, 5000-5004.
16. Volkman, B. F., Lipson, D., Wemmer, D. E. & Kern, D. (2001). Two-state behavior in a single-domain signaling protein. *Science* 291, 2429-2433.

17. Neri, D., Billeter, M., Wider, G. & Wüthrich, K. (1992). NMR determination of residual structure in a urea-denatured protein, the 434-repressor. *Science* 257, 1559-1563.
18. Lapidus, L. J., Eaton, W. A. & Hofrichter, J. (2000). Measuring the rate of intramolecular contact formation in polypeptides. *Proc. Natl. Acad. Sci. USA* 97, 7220-7225.
19. Klein-Seetharaman, J., Oikawa, M., Grimshaw, S. B., Wirmer, J., Duchardt, E., Ueda, T., Imoto, T., Smith, L. J., Dobson, C. M. & Schwalbe, H. (2002). Long-range interactions within a nonnative protein. *Science* 295, 1719-1722.
20. Sánchez, I. E. & Kiefhaber, T. (2003). Hammond behavior versus ground state effects in protein folding: evidence for narrow free energy barriers and residual structure in unfolded states. *J. Mol. Biol.* 327, 867-884.
21. Dyson, H. J. & Wright, P. E. (2004). Unfolded proteins and protein folding studies by NMR. *Chem. Rev.* 104, 3607-3622.
22. Schwalbe, H., Fiebig, K. M., Buck, M., Jones, J. A., Grimshaw, S. B., Spencer, A., Glaser, S. J., Smith, L. J. & Dobson, C. M. (1997). Structural and Dynamical Properties of a Denatured Protein. Heteronuclear 3D NMR Experiments and Theoretical Simulations of Lysozyme in 8 M Urea. *Biochemistry* 36, 8977-8991.
23. Tanford, C. (1968). Protein denaturation. Part A. Characterization of the denatured state. *Adv. Prot. Chem.* 23, 121-217.
24. Haas, E., Katchalski-Katzir, E. & Steinberg, I. Z. (1978). Brownian motion at the ends of oligopeptid chains as estimated by energy transfer between chain ends. *Biopolymers* 17, 11-31.
25. Karplus, M. & Weaver, D. L. (1976). Protein-folding dynamics. *Nature* 260, 404-406.
26. Kim, P. S. & Baldwin, R. L. (1990). Intermediates in the folding reactions of small proteins. *Annu. Rev. Biochem.* 59, 631-660.
27. Karplus, M. & Weaver, D. L. (1994). Protein folding dynamics: The diffusion-collision model and experimental data. *Protein Sci.* 3, 650-668.
28. Ptitsyn, O. B. (1996). How molten is the molten globule? *Nat. Struct. Biol.* 3, 488-490.
29. Wetlaufer, D. B. (1973). Nucleation, rapid folding, and globular intrachain regions in proteins. *Proc. Natl. Acad. Sci. USA* 70, 697-701.
30. Abkevich, V. I., Gutin, A. M. & Shakhovich, E. I. (1994). Specific nucleus in the transition state for protein folding: evidence from the lattice model. *Biochemistry* 33, 10026-10036.
31. Krieger, F., Fierz, B., Bieri, O., Drewello, M. & Kiefhaber, T. (2003). Dynamics of unfolded polypeptide chains as model for the earliest steps in protein folding. *J. Mol. Biol.* 332, 265-274.
32. Flory, P. J. (1969). *Statistical Mechanics of Chain Molecules*, Hanser Publishers, Munich.
33. Brant, D. A. & Flory, P. J. (1965). The configuration of random polypeptides. I. Experimental results. *J. Am. Chem. Soc.* 87, 2788-2791.

34. Brant, D. A. & Flory, P. J. (1965). The configuration of random polypeptide chains. II. Theory. *J. Am. Chem. Soc.* 87, 2791-2800.
35. Miller, W. G., Brant, D. A. & Flory, P. J. (1967). Random coil configurations of polypeptide copolymers. *J. Mol. Biol.* 23, 67-80.
36. Schimmel, P. R. & Flory, P. J. (1968). Conformational energy and configurational statistics of copolypeptides containing L-proline. *J. Mol. Biol.* 34, 105-120.
37. Lapanje, S. & Tanford, C. (1967). Proteins as Random Coils. IV. Osmotic Pressures, Second Virial Coefficients, and Unperturbed dimensions in 6 M Guanidine Hydrochloride. *J. Am. Chem. Soc.* 89, 5030-50-33.
38. Damaschun, G., Damaschun, H., Gast, K., Misselwitz, R., Muller, J. J., Pfeil, W. & Zirwer, D. (1993). Cold Denaturation Induces Conformational Changes in Phosphoglycerate Kinase from Yeast. *Biochemistry* 32, 7739-7746.
39. Kohn, J. E., Millet, I. S., Jacob, J., Zagrovic, B., Dillon, T. M., Cingel, N., Dothager, R. S., Seifert, S., Thiyagarajan, P., Sosnick, T. R., Hasan, M. Z., Pande, V. S., Ruczinski, I., Doniach, S. & Plaxco, K. W. (2004). Random-coil behavior and the dimensions of chemically unfolded proteins. *Proc. Natl. Acad. Sci. U S A* 101, 12491-12496.
40. Pappu, R. V., Srinivasan, R. & Rose, G. D. (2000). The Flory isolated-pair hypothesis is not valid for polypeptide chains: Implications for protein folding. *Proc. Natl. Acad. Sci. USA* 97, 12565-12570.
41. Baldwin, R. L. (2002). A new perspective on unfolded proteins. *Adv. Prot. Chem* 62, 361-367.
42. Bieri, O. & Kiefhaber, T. (1999). Elementary steps in protein folding. *Biol. Chem.* 380, 923-929.
43. Schwarz, G. & Seelig, J. (1968). Kinetic properties and the electric field effect of the helix-coil transition of poly-(γ -benzyl-L-glutamate) determined from dielectric relaxation measurements. *Biopolymers* 6, 1263-1277.
44. Gruenewald, B., Nicola, C. U., Lustig, A. & Schwarz, G. (1979). Kinetics of the helix-coil transition of a polypeptide with non-ionic side chain groups, derived from ultrasonic relaxation measurements. *Biophys. Chem.* 9, 137-147.
45. Thompson, P., Eaton, W. & Hofrichter, J. (1997). Laser temperature jump study of the helix-coil kinetics of an alanine peptide interpreted with a "kinetic zipper" model. *Biochemistry* 36, 9200-9210.
46. Munoz, V., Thompson, P., Hofrichter, J. & Eaton, W. (1997). Folding dynamics and mechanism of β -hairpin formation. *Nature* 390, 196-199.
47. Pauling, L., Corey, R. B. & Branson, H. R. (1951). The structure of proteins- 2 Hydrogen-bonded helical configurations of the polypeptide chain. *Proc. Natl. Acad. Sci. U S A* 37, 205-211.
48. Scholtz, J. M. & Baldwin, R. L. (1992). The mechanism of alpha-helix formation by peptides. *Annu. Rev. Biophys. Biomol. Struct.* 21, 95-118.
49. Zimm, B. H. & Bragg, J. K. (1959). Theory of phase transition between helix and random coil in polypeptide chains. *J. Chem. Phys.* 31, 526-535.

50. Lifson, S. & Roig, A. (1961). On the theory of helix-coil transitions in polypeptides. *J. Chem. Phys.* 34, 1963-1974.
51. Schwarz, G. (1965). On the kinetics of the helix-coil transition of polypeptides in solution. *J. Mol. Biol.* 11, 64-77.
52. Blanco, F. J., Rivas, G. & Serrano, L. (1994). A short linear peptide that folds into a native stable β -hairpin in aqueous solution. *Nat. Struct. Biol.* 1, 584-590.
53. Blanco, F., Jimenez, M., Pineda, A., Rico, M. & Santoro, J. N., J. (1994). NMR Solution Structure of the Isolated N-Terminal Fragment of Protein-G B1 Domain. Evidence of Trifluorethanol Induced Native-like β -Hairpin Formation. *Biochemistry* 33, 6004-6014.
54. Munoz, V., Henry, E. R., Hofrichter, J. & Eaton, W. A. (1998). A statistical mechanical model for β -hairpin kinetics. *Proc. Natl. Acad. Sci. U S A* 95, 5872-5879.
55. Einstein, A. (1906). Zur Theorie der Brownschen Bewegung. *Ann. d. Phys.* 19, 371-381.
56. von Smoluchowski, M. (1906). Zur kinetischen Theorie der Brownschen Molekularbewegung und der Suspensionen. *Ann. d. Phys.* 21, 756-780.
57. von Smoluchowski, M. (1916). Drei Vorträge über Diffusion, Brownsche Molekularbewegung und Koagulation von Kolloidteilchen. *Phys. Z.* 17, 557-571.
58. Jacobsen, H. & Stockmayer, W. H. (1950). Intramolecular reaction in polycondensations. I. The theory of linear systems. *J. Chem. Phys.* 18, 1600-1606.
59. Flory, P. J., Suter, U. W. & Mutter, M. (1976). Macrocyclization equilibria. 1. Theory. *J. Am. Chem. Soc.* 98, 5733-5739.
60. Mutter, M. (1977). Macrocyclization equilibria of Polypeptides. *J. Am. Chem. Soc.* 99, 8307-8314.
61. Szabo, A., Schulten, K. & Schulten, Z. (1980). First passage time approach to diffusion controlled reactions. *J. Chem. Phys.* 72, 4350-4357.
62. Pastor, R. W., Zwanzig, R. & Szabo, A. (1996). Diffusion limited first contact of the ends of a polymer: comparison of theory with simulation. *J. Chem. Phys.* 105, 3878-3882.
63. Fierz, B. (2002). Diploma thesis, Basel.
64. Haas, E., Wilchek, M., Katchalski-Katzir, E. & Steinberg, I. Z. (1975). Distribution of end-to-end distances of oligopeptides in solution as estimated by energy transfer. *Proc. Natl. Acad. Sci. U S A* 72, 1807-1811.
65. Winnik, M. A. (1985). End-to-end cyclization of polymer chains. *Acc. Chem. Res.* 18, 73-79.
66. Cuniberti, C. & Perico, A. (1984). Intramolecular diffusion-controlled reactions and polymer dynamics. *Prog. Polym. Sci.* 10, 271-316.
67. Chang, I.-J., Lee, J. C., Winkler, J. R. & Gray, H. B. (2003). The protein-folding speed limit: Intrachain diffusion times set by electron-transfer rates in denatured

- RU(NH₃)₅(His-33)-Zn cytochrome c. *Proc. Natl. Acad. Sci. U S A* 100, 3838-3840.
68. Hudgins, R. R., Huang, F., Gramlich, G. & Nau, W. M. (2002). A fluorescence-based method for direct measurements of submicrosecond intramolecular contact formation in biopolymers: an exploratory study with polypeptides. *J. Am. Chem. Soc.* 124, 556-564.
 69. Neuweiler, H., Schulz, H., Bohmer, M., Enderlein, J. & Sauer, M. (2003). Measurements of Submicrosecond Intramolecular Contact Formation in Peptides at the Single-Molecule Level. *J. Am. Chem. Soc.* 125, 2269-2272.
 70. Hagen, S. J., Hofrichter, J., Szabo, A. & Eaton, W. A. (1996). Diffusion-limited contact formation in unfolded cytochrome c: Estimating the maximum rate of protein folding. *Proc. Natl. Acad. Sci. USA* 93, 11615-11617.
 71. Wagner, P. J. & Klán, P. (1999). Intramolecular triplet energy transfer in flexible molecules: electronic, dynamic, and structural aspects. *J. Am. Chem. Soc.* 121, 9626-9635.
 72. Camacho, C. J. & Thirumalai, D. (1995). Theoretical predictions of folding pathways by using the proximity rule, with applications to bovine pancreatic inhibitor. *Proc. Natl. Acad. Sci. USA* 92, 1277-1281.
 73. Bieri, O. (2000). PhD thesis, Basel.
 74. Bent, D. V. & Hayon, E. (1975). Excited state chemistry of aromatic amino acids and related peptides: III. Tryptophan. *J. Am. Chem. Soc.* 97, 2612-2619.
 75. Huang, F. & Nau, W. M. (2003). A conformational flexibility scale for amino acids in peptides. *Angew. Chem. Int. Ed. Engl.* 42, 2269-2272.
 76. Turro, N. J. (1991). *Modern Molecular Photochemistry*, University Science Books, Sausalito, CA.
 77. Dexter, D. L. (1953). A theory to sensitized luminescence in solids. *J. Chem. Phys.* 21, 836 - 850.
 78. Rice, S. A. (1985). *Diffusion-limited reactions*. Chemical Kinetics (Bamford, C. H., Tipper, C. F. H. & Compton, R. G., Eds.), 25, Elsevier, Amsterdam.
 79. Anderson Jr., R. W., Hochstrasser, R. M., Lutz, H. & Scott, G. W. (1974). Direct measurements of energy transfer between triplet states of molecules in liquids using picosecond pulses. *J. Chem. Phys.* 61, 2500-2506.
 80. Mohtat, N., Cozens, F., L. & Scaiano, J. C. (1998). Multistage exit of excited xanthone from micelles. *J. Phys. Chem. B* 102, 7557-7562.
 81. Scaiano, J. C. (1980). Solvent effects in the photochemistry of xanthone. *J. Am. Chem. Soc.* 102, 7747-7753.
 82. Grabner, G., Rechthaler, K., Mayer, B., Kohler, G. & Rotkiewicz, K. (2000). Solvent influence on the photolysis of naphthalene: fluorescence and triplet state properties in aqueous solutions and cyclodextrins. *J. Phys. Chem. A* 104, 1365-1376.
 83. (1972). *Handbook of Chemistry and Physics*. 53rd edit (Weast, R. C., Ed.), CRC, Cleveland, OH.

84. Satzger, H., Schmidt, B., Root, C., Zinth, W., Fierz, B., Krieger, F., Kiefhaber, T. & Gilch, P. (2004). Ultrafast quenching of the xanthone triplets by energy transfer: New insight into the intersystem crossing kinetics. *J. Phys. Chem. A*, in press.
85. Krieger, F., Fierz, B., Axthelm, F., Joder, K., Meyer, D. & Kiefhaber, T. (2004). Intrachain diffusion in a protein loop fragement from carp parvalbumin. *Chem. Phys.* 307, 209-215.
86. Bent, D. V. & Hayano, T. (1975). Excited state chemistry of aromatic amino acids and related peptides: I. Tyrosine. *J. Am. Chem. Soc.* 97, 2599-2606.
87. Ramseier, M., Senn, P. & Wirz, J. (2003). Photohydration of Benzophenone in Aqueous Acid. *J. Phys. Chem. A* 107, 3305-3315.
88. Bright, J. N., Woolf, T. B. & Hoh, J. H. (2001). Predicting properties of intrinsically unstructured proteins. *Prog. Biophys. Mol. Biol.* 76, 131-173.
89. de Gennes, P. G. (1979). *Scaling Concepts in Polymer Physics*, Cornell University Press, Ithaca, New York.
90. Wouterson, S., Mu, Y., Stock, G. & Hamm, P. (2001). Subpicosecond conformational dynamics of small peptides probed by two-dimensional vibrational spectroscopy. *Proc. Natl. Acad. Sci. U S A* 98, 11254-11258.
91. Sporlein, S., Carstens, H., Satzger, H., Renner, C., Behrendt, R., Moroder, L., Tavan, P., Zinth, W. & Wachtveitl, J. (2002). Ultrafast spectroscopy reveals subnanosecond peptide conformational dynamics and validates molecular dynamics simulation. *Proc. Natl. Acad. Sci. U S A* 99, 7998-8002.
92. Bredenbeck, J., Helbig, J., Sieg, A., Schrader, T., Zinth, W., Renner, C., Behrendt, R., Moroder, L., Wachtveitl, J. & Hamm, P. (2003). Picosecond conformational transition and equilibration of a cyclic peptide. *Proc. Natl. Acad. Sci. U S A* 100, 6452 - 6457.
93. Kubelka, J., Hofrichter, J. & Eaton, W. A. (2004). The protein folding 'speed limit'. *Curr. Opin. Struct. Biol.* 14, 76-88.
94. Ansari, A., Colleen, J. M., Eric, H. R., Hofrichter, J. & Eaton, W. A. (1992). The Role of Solvent Viscosity in the Dynamics of Protein Conformational Changes. *Science* 256, 1796 - 1798.
95. Reimer, U., Scherer, G., Drewello, M., Kruber, S., Schutkowski, M. & Fischer, G. (1998). Side-chain effects on peptidyl-prolyl cis/trans isomerization. *J. Mol. Biol.* 279, 449-460.
96. Brandts, J. F., Halvorson, H. R. & Brennan, M. (1975). Consideration of the possibility that the slow step in protein denaturation reactions is due to *cis-trans* isomerism of proline residues. *Biochemistry* 14, 4953-4963.
97. Wilmot, C. M. & Thornton, J. M. (1988). Analysis and prediction of the different types of β -turns in proteins. *J. Mol. Biol.* 203, 221-232.
98. Dumy, P., Keller, M., D.E., R., Rohwedder, B., Woehr, T. & Mutter, M. (1997). Pseudo-prolines as a molecular hinge: Reversible induction of cis amide bond into peptide backbones. *J. Am. Chem. Soc.* 119, 918-925.

99. Keller, M., Sager, C., Dumy, P., Schutkowski, M., Fischer, G. S. & Mutter, M. (1998). Enhancing the proline effect: Pseudo-prolines for tailoring cis/trans isomerization. *J. Am. Chem. Soc.* 120, 2714-2720.
100. Berg, O. G. & von Hippel, P. H. (1985). Diffusion-controlled Macromolecular Interactions. *Ann. Rev. Biophys. Biophys. Chem.* 14, 131 - 160.
101. Bieri, O. & Kiefhaber, T. (2000). Kinetic models in protein folding. In *Protein Folding: Frontiers in Molecular Biology* 2nd edit. (Pain, R., ed.), pp. 34-64. Oxford University Press, Oxford.
102. Pohl, F. M. (1976). Temperature-dependence of the kinetics of folding of chymotrypsinogen A. *FEBS Lett.* 65, 293-296.
103. Oliveberg, M., Tan, Y. J. & Fersht, A. R. (1995). Negative activation enthalpies in the kinetics of protein folding. *Proc. Natl. Acad. Sci. U S A* 92, 8926 - 8929.
104. Schindler, T. & Schmid, F. X. (1996). Thermodynamic characterization of an extremely rapid protein folding reaction. *Biochemistry* 51, 16833-16842.
105. Schönbrunner, N., Koller, K.-P. & Kiefhaber, T. (1997). Folding of the disulfide-bonded β -sheet protein tendamistat: Rapid two-state folding without hydrophobic collapse. *J. Mol. Biol.* 268, 526-538.
106. Bachmann, A. & Kiefhaber, T. (2001). Apparent two-state tendamistat folding is a sequential process along a defined route. *J. Mol. Biol.* 306, 375-386.
107. Timasheff, S. N. (1998). Control of protein stability and reactions by weakly interacting cosolvents: The simplicity of the complicated. *Adv. Prot. Chem* 51, 355-432.
108. Zhu, W. & Ediger, M. D. (1997). Viscosity dependence of polystyrene local dynamics in dilute solution. *Macromolecules* 30, 1205 - 1210.
109. Punched, B. J. & Adolf, D. B. (2002). Pressure and Temperature Dependence of the Dilute Solution Segmental Dynamics of Anthracene-Labeled Polyisoprene. *Macromolecules* 35, 3281-3287.
110. Pilar, J. & Labsky. (2003). Solvent Dependence of Polystyrene Local Segmental Dynamics in Dilute Solution by Spin-Label X Band ESR. *Macromolecules* 36, 913-920.
111. Beece, D., Eisenstein, L., Frauenfelder, H., Good, D., Marden, M. C., Reinisch, L., Reynolds, A. H., Sorensen, L. B. & Yue, K. T. (1980). Solvent viscosity and protein dynamics. *Biochemistry* 19, 5147 - 5157.
112. Velsko, P. S., Waldeck, D. H. & Fleming, G. R. (1983). Breakdown of Kramers theory description of photochemical isomerization and the possible involvement of frequency dependent friction. *J. Chem. Phys.* 78, 249-258.
113. Courtney, S. H. & Fleming, G. R. (1985). Photoisomerization of stilbene at low viscosity solvents: Comparison of isolated and solvent molecules. *J. Chem. Phys.* 83, 215-222.
114. Jas, G. S., Eaton, W. A. & Hofrichter, J. (2001). Effect of viscosity on the kinetics of α -helix and β -hairpin formation. *J. Phys. Chem. B* 105, 261 - 272.
115. Barshtein, G., Almagor, A., Yedgar, S. & Gavish, B. (1995). Inhomogeneity of viscous solutions. *Phys. Rev. E* 52, 555 - 557.

116. Kramers, H. A. (1940). Brownian motion in a field of force and the diffusion model of chemical reactions. *Physica* 4, 284-304.
117. Hanggi, P., Talkner, P. & Borkovec, M. (1990). Reaction-rate theory: fifty years after Kramers. *Rev. Mod. Phys.* 62, 251 - 341.
118. Grote, R. F. & Hynes, J. T. (1980). The stable states picture of chemical reaction. II. Rate constants for condensed and gas phase reaction models. *J. Chem. Phys.* 73, 2715 - 2732.
119. Shastry, W. C. R. & Roder, H. (1998). Evidence for barrier-limited protein folding kinetics on the microsecond time scale. *Nat Struct Biol* 5, 385-392.
120. Holtzer, M. E., Bretthorst, G. L., d'Avignon, D. A., Angeletti, R. H., Mints, L. & Holtzer, A. (2001). Temperature dependence of the folding and unfolding kinetics of the GCN4 leucine zipper via ¹³C(alpha)-NMR. *Biophys. J.* 80, 939-951.
121. Arrhenius, S. (1889). Über die Reaktionsgeschwindigkeit bei der Inversion von Rohrzucker durch Säuren. *Z. Phys. Chem.* 4, 226-248.
122. Eyring, H. (1935). The activated complex in chemical reactions. *J. Chem. Phys.* 3, 107-115.
123. Daragan, V. A. & Mayo, K. H. (1993). Tri- and Diglycine Backbone Rotational Dynamics Investigated by ¹³C NMR Multiplet Relaxation and Molecular Dynamics Simulations. *Biochemistry* 32, 11488-11499.
124. Fischer, G., Berger, E. & Bang, H. (1989). Kinetic β -deuterium isotope effects suggest a covalent mechanism for the protein folding enzyme peptidylprolyl cis/trans isomerase. *FEBS Lett.* 250, 267-270.
125. Jencks, W. P. (1969). *Catalysis in Chemistry and Enzymology*, McGraw-Hill Book Company, New York.
126. Hogg, J. L. (1978). Secondary Hydrogen Isotope Effects. In *Transition States of biochemical Processes*. (Gandour, R. D. & Schowen, R. L., eds.), pp. 201-224. Plenum Press, New York.
127. Hirota, E., Saito, S. & Endo, Y. (1979). Barrier to internal rotation in ethane from microwave spectrum of CH₃CHD₂. *J. Chem. Phys.* 71, 1183-1187.
128. Aune, K. C. & Tanford, C. (1969). Thermodynamics of the denaturation of lysozyme by guanidine hydrochloride. II Dependence on denaturant concentration at 25 degrees. *Biochemistry* 11, 4586-4590.
129. Santoro, M. M. & Bolen, D. W. (1988). Unfolding free energy changes determined by the linear extrapolation method. 1. Unfolding of phenylmethanesulfonyl alpha-chymotrypsin using different denaturants. *Biochemistry* 27, 8063-8068.
130. Pace, C. N. (1986). Determination and analysis of urea and guanidine hydrochloride denaturation curves. *Meth. Enzymol.* 131, 266-280.
131. Robinson, D. R. & Jencks, W. P. (1965). The effect of compounds of the urea-guanidinium class on the activity coefficient of acetyltetraglycine ethyl ester and related compounds. *J. Am. Chem. Soc.* 87, 2462-2470.

132. Nozaki, Y. & Tanford, C. (1963). Solubility of amino acids and related compounds in aqueous urea solutions. *J. Biol. Chem.* 238, 4074-4081.
133. Moglich, A., Krieger, F. & Kiefhaber, T. (2004). Molecular basis for the effect of urea and guanidinium chloride on the dynamics of unfolded polypeptide chains. *J. Mol. Biol.*, in press.
134. Perl, D., Jacob, M., Bánó, M., Stupák, M., Antalík, M. & Schmid, F. X. (2002). Thermodynamics of a diffusional protein folding reaction. *Biophys. Chem.* 2-3, 173-190.
135. Schellman, J. A. (2002). Fifty years of solvent denaturation. *Biophys. Chem.* 96, 91-101.
136. Kaiser, B., Laesecke, A. & Schmeck, M. (1989). Experimental study and correlation of the viscosity of 2,2,2-Trifluoroethanol (TFE)-water mixtures. *Int. J. Thermophys.* 10, 713-726.
137. Aminabhavi, T. M. & Gopalakrishna, B. (1996). Density, viscosity, refractive index, and speed of sound in aqueous mixtures of N,N-dimethylformamide, dimethyl sulfoxide, N,N-dimethylacetamide, acetonitrile, ethylene glycol, diethylene glycol, 1,4-dioxane, tetrahydrofuran, 2-methoxyethanol, and 2-Ethoxyethanol at 298.15 K. *J. Chem. Eng. Data* 40, 856-861.
138. Buck, M. (1998). Trifluoroethanol and colleagues: cosolvents come of age. recent studies with peptides and proteins. *Quart. Rev. Biophys.* 31, 297-335.
139. Main, E. R. G. & Jackson, S. E. (1999). Does trifluoroethanol affect folding pathways and can it be used as a probe of structure in transition states. *Nat. Struct. Biol.* 6, 831-835.
140. Koradi, R., Billeter, M. & Wuthrich, K. (1996). MOLMOL: a program for display and analysis of macromolecular structures. *J. Mol. Graphics* 14, 51-55.
141. Kumar, V. D., Lee, L. & Edwards, B. F. (1990). Refined crystal structure of calcium-liganded carp parvalbumin 4.25 at 1.5-Å resolutions. *Biochemistry* 29, 1404-1412.
142. Gronenborn, A. M., Filpula, D. R., Essig, N. Z., Achari, A., Whitlow, M., Wingfield, P. T. & Clore, G. M. (1991). A novel, highly stable fold of the immunoglobulin binding domain of streptococcal protein G. *Science* 253, 657-661.
143. Teraoka, I. (2001). *Polymer solutions: an introduction to physical properties.*, John Wiley & Sons, New York.
144. Kretsinger, R. H. (1980). Structure and evolution of calcium-modulated proteins. *CRC Crit. Rev. Biochem.* 8, 119-174.
145. Derancourt, J., Haitech, J. & Pechere, J.-F. (1978). Binding of calcium by parvalbumin fragments. *Biochim. Biophys. Acta* 532, 373-375.
146. Lumb, R. F. & Martell, A. E. (1953). Metal chelating tendencies of glutamic and aspartic acids. *J. Phys. Chem.* 57, 690-693.
147. Coffee, C. J. & Solano, C. (1976). Preparation and properties of carp muscle parvalbumin fragments A (residues 1-75) and B (residues 76-108). *Biochim. Biophys. Acta* 453, 67-80.

148. Corson, D. C., Williams, T. C., Kay, L. E. & Sykes, B. D. (1986). ¹H NMR Spectroscopic Studies of Calcium-Binding proteins. Stepwise Proteolysis of the C-Terminal α -Helix of a Helix-Loop-Helix Metal-Binding Domain. *Biochemistry* 25, 1817-1826.
149. Marqusee, S. & Baldwin, R. L. (1987). Helix stabilization by Glu...Lys+ salt bridges in short peptides of de novo design. *Proc. Natl. Acad. Sci. U S A* 84, 8898-8902.
150. Marqusee, S., Robbins, V. H. & Baldwin, R. L. (1989). Unusually stable helix formation in short alanine-based peptides. *Proc. Natl. Acad. Sci. U S A* 86, 5286-5290.
151. Chen, Y.-H., Yang, J. T. & Chau, K. H. (1974). Determination of the Helix and β Form of Proteins in Aqueous Solutions by Circular Dichroism. *Biochemistry* 13, 3350-3359.
152. Chakrabartty, A., Schellman, J. A. & Baldwin, R. L. (1991). Large differences in the helix propensities of alanine and glycine. *Nature* 351, 586-588.
153. Zhang, S. & Rich, A. (1997). Direct conversion of an oligopeptide from a β -sheet to an α -helix: A model for amyloid formation. *Proc. Natl. Acad. Sci. U S A* 94, 23-28.
154. Chakrabartty, A., Kortemme, T. & Baldwin, R. L. (1994). Helix propensities of the amino-acids measured in alanine-based peptides without helix-stabilizing side-chain interactions. *Protein Sci.* 3, 843-852.
155. Demchenko, A. P. (1986). *Ultraviolet Spectroscopy of Proteins*, Springer, Berlin.
156. Carpino, L. A. & Han, G. Y. (1972). The 9-Fluorenylmethoxycarbonyl Amino-Protecting Group. *J. Org. Chem.* 37, 3404-3409.
157. Chan, W. C. & White, P. D. (2000). *Fmoc Solid Phase Peptide Synthesis. A Practical Approach* (Hames, B. D., Ed.), Oxford University Press, Oxford.
158. Fischer, G., Bang, H. & Mech, C. (1984). Detection of enzyme catalysis for cis-trans-isomerization of peptide-bonds using proline containing peptides as substrates. *Biomed. Biochim. Acta* 43, 1101-1111.

Dynamics of Unfolded Polypeptide Chains as Model for the Earliest Steps in Protein Folding

Florian Krieger¹, Beat Fierz¹, Oliver Bieri¹, Mario Drewello² and Thomas Kiefhaber^{1*}

¹Department of Biophysical Chemistry, Biozentrum der Universität Basel
Klingelbergstr. 70, CH-4056 Basel, Switzerland

²Max-Planck Arbeitsgruppe "Enzymologie der Proteinfaltung", Weinbergweg 22a, D-06120 Halle, Germany

The rate of formation of intramolecular interactions in unfolded proteins determines how fast conformational space can be explored during folding. Characterization of the dynamics of unfolded proteins is therefore essential for the understanding of the earliest steps in protein folding. We used triplet–triplet energy transfer to measure formation of intrachain contacts in different unfolded polypeptide chains. The time constants ($1/k$) for contact formation over short distances are almost independent of chain length, with a maximum value of about 5 ns for flexible glycine-rich chains and of 12 ns for stiffer chains. The rates of contact formation over longer distances decrease with increasing chain length, indicating different rate-limiting steps for motions over short and long chain segments. The effect of the amino acid sequence on local chain dynamics was probed by using a series of host-guest peptides. Formation of local contacts is only sixfold slower around the stiffest amino acid (proline) compared to the most flexible amino acid (glycine). Good solvents for polypeptide chains like EtOH, GdmCl and urea were found to slow intrachain diffusion and to decrease chain stiffness. These data allow us to determine the time constants for formation of the earliest intrachain contacts during protein folding.

© 2003 Elsevier Ltd. All rights reserved.

Keywords: protein folding; chain dynamics; intrachain diffusion; triplet–triplet transfer; unfolded proteins

*Corresponding author

Introduction

Many biological polymers have to form specific intramolecular interactions to be able to fulfil their biological functions. The rate at which unstructured polymer chains can explore conformational space is limited by intrachain diffusion, i.e. by the rate at which two points on a polymer chain can make contact. Intrachain diffusion is particularly important during protein folding, which requires formation of a large number of specific long-range and short-range interactions. The knowledge of the rates of intrachain contact formation in polypeptide chains and their dependence on amino acid sequence and chain length is therefore essential for the understanding of the dynamics of the earliest steps in protein folding and for the charac-

terization of the free energy barriers for protein folding reactions. We have recently applied the method of triplet–triplet energy transfer (TTET) between thioxanthone and 1-naphthyl alanine (NAla) to measure intrachain contact formation in short unstructured polypeptide chains.¹ TTET is a diffusion-controlled process that allows direct measurements of the rates of intrachain diffusion. Flexible peptide chains with alternating glycine and serine residues showed exponential kinetics for end-to-end diffusion with time constants on the 15–20 ns time-scale for formation of an $i, i + 3$ contact.¹ Due to sensitivity of the triplet energy of the donor to solvent polarity, these experiments were carried out in ethanol. Later, similar experiments were reported for quenching of tryptophan triplet states by cysteine in water, which gave significantly slower rates for chain diffusion compared to TTET.² This reaction is, however, not diffusion-controlled and is accompanied by the formation of S• radicals, which does not allow direct model-free measurements of intrachain diffusion kinetics.^{2,3} Another experimental approach used

Abbreviations used: TTET, triplet–triplet energy transfer; NAla, 1-naphthyl alanine; DBO, 2,3 diazabicyclo[2.2.2]oct-2-ene.

E-mail address of the corresponding author: t.kiefhaber@unibas.ch

short poly(glycine-serine) peptides and measured intrachain diffusion by the quenching of 2,3 diaza-bicyclo[2.2.2]oct-2-ene (DBO) fluorescence by tryptophan in water.⁴ This system is close to the diffusion limit and gave similar contact rates in water as TTET from thioxanthone to naphthalene in ethanol.

All available data on polypeptide dynamics were obtained from short, glycine-rich polypeptide chains, which gave an estimate of the maximum rate of local contact formation in flexible chains. However, these experiments did not allow a reliable scaling of the contact rates to longer distances and they did not give information on the dynamics of stiffer chains. In addition, no direct measurements of intrachain diffusion have been reported for a completely diffusion-controlled system in water. Here, we present direct model-free TTET measurements of intrachain diffusion in longer poly(glycine-serine) chains with up to 56 amino acid residues between the points of contact formation. We further measured intrachain diffusion in stiffer glycine-free chains and tested the effect of various amino acids on local chain dynamics. Changing the triplet donor from thioxanthone to xanthone (9-oxoxanthen) enabled us to perform all measurements in water and to additionally test the influence of various solvents on chain dynamics.

Results

Intrachain triplet–triplet energy transfer (TTET)

TTET from a triplet donor to a triplet acceptor group involves transfer of two electrons (Dexter mechanism) and requires van der Waals contact between donor and acceptor.⁵ Efficient electron transfer occurs if the triplet donor has a high extinction coefficient, a high intersystem crossing efficiency and a higher triplet energy than the acceptor. In suitable donor/acceptor pairs, the transfer process is diffusion-controlled and can be followed easily by changes in the absorbance of the triplet states. This makes TTET a perfect tool to measure rates of contact formation between two reporter groups directly. When triplet donor and acceptor groups are attached to a polypeptide chain, the TTET kinetics are determined by the dynamics and by the conformational properties of the chain. In unfolded polypeptides, a large number of different conformations are present in equilibrium, which gives a wide distribution of donor–acceptor distances. Some polypeptide chains statistically have contact between donor and acceptor at the time of the laser flash. This will lead to instantaneous TTET in these molecules. The remaining molecules have to sample conformational space by intrachain diffusion, which will also lead to collisional TTET (Figure 1). The observed transfer kinetics in these molecules reflects the rate at which chain dynamics allows

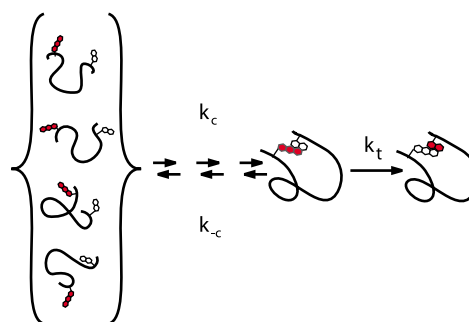
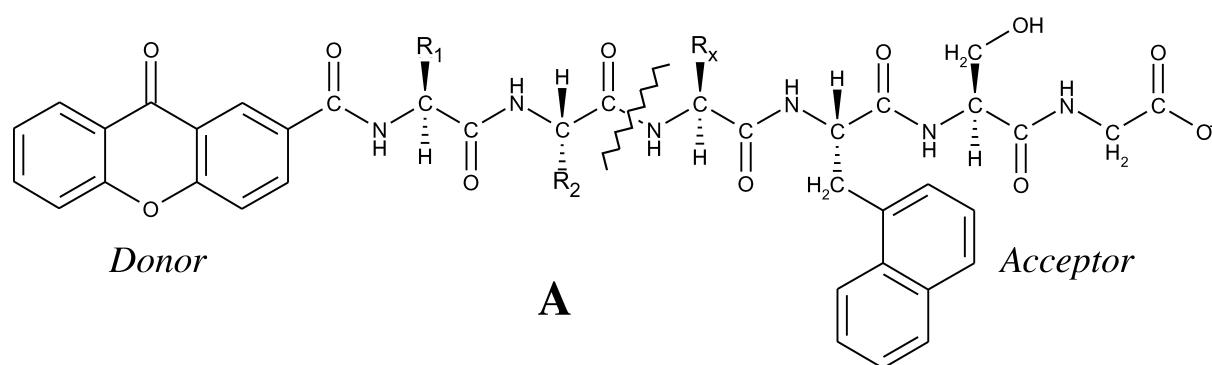


Figure 1. A representation of the triplet–triplet energy transfer (TTET) experiments. Triplet donor and acceptor groups are attached at specific positions on an unstructured polypeptide chain. Triplet states are produced in the donor group by a short laser flash and transferred to the acceptor upon encounter at van der Waals distance in a diffusion-controlled process. The experiments allow determination of the absolute rate constant for contact formation (k_c) in the ensemble of unfolded conformation, if the transfer process is much faster than chain dynamics ($k_t \gg k_c, k_{-c}$).

the two labeled points on the polypeptide chain to meet. This assumes that electron transfer is faster than chain diffusion ($k_t \gg k_c$ and k_{-c} ; Figure 1). We used naphthalene as a triplet acceptor and xanthone as a triplet donor. Xanthone has a significantly higher triplet energy than the previously used thioxanthone,⁶ which allows TTET to naphthalene in all solvents. The triplet states of xanthone are produced by a 4 ns laser flash at 355 nm. Formation of xanthone triplets from the excited singlet state is a rapid process with a time constant of about 8 ps and a high quantum yield for intersystem crossing ($\sim 99\%$).⁷ Xanthone triplets are long-lived ($t_{1/2} \approx 30 \mu\text{s}$ in water), which allows measurements of TTET processes from several nanoseconds (limited by the length of the laser pulse) up to about $10 \mu\text{s}$. TTET from xanthone to naphthylacetic acid has a bimolecular rate constant of $k_t = 4 \times 10^9 \text{ M}^{-1} \text{ s}^{-1}$ in water for donor and acceptor free in solution (data not shown), which is the value expected for a diffusion-controlled bimolecular reaction. This shows that electron transfer is significantly faster than formation of the encounter complex. In TTET measurements of xanthone dissolved in liquid 2-methyl naphthalene ($c = 7\text{M}$) transfer was complete within the dead-time of the 4 ns laser flash, which shows that transfer occurs in the subnanosecond time region ($k_t \gg 10^9 \text{ s}^{-1}$). This is consistent with data for similar TTET pairs, which showed time constants for transfer ($1/k_t$) in the 10–20 ps time region.⁸

We attached the xanthone group to the N terminus of various model peptides and used the non-natural amino acid NAla to introduce the naphthalene acceptor group near the C terminus. The length and the sequence of the polypeptide chain between donor and acceptor were varied, which gave peptides (A) of the general form:



where R_1 and R_2 represent any amino acid side-chain, and x indicates the number of amino acid residues between donor and acceptor. All peptides used in this study were unstructured, as judged by CD and NMR spectroscopy.

Intrachain contact formation in glycine-rich polypeptides in water

To compare the results obtained from the xanthone/naphthyl TTET pair in water with previous experimental results on short flexible peptides,^{1,4,9} we measured intrachain contact formation in poly(glycine-serine) peptides in water. In contrast to the earlier studies that were limited to chains of two to 15 amino acid residues, we were able to extend the distance between donor and acceptor to 57 amino acid residues. Figure 2 shows the kinetics of TTET from xanthone to naphthalene in a Xan-(Gly-Ser)₁₄-NAla-Gly-Ser peptide as a representative example. The formation of xanthone triplets is reflected by the laser-induced appearance of an absorbance band at 590 nm. Decay of the xanthone triplets is accompanied by the appearance of a new absorbance band at 420 nm, indicating formation of naphthalene triplets through intramolecular TTET. Both triplet states show a fine structure in their absorbance spectra in water, in accordance with previous reports.^{10,11} The absence of any absorbance bands between 450 nm and 500 nm indicates the absence of xanthone radicals throughout the experiment.¹² Monitoring the time-

dependent decrease in the xanthone triplet absorbance band at 590 nm (Figure 3) clearly shows that intrachain contact formation can be described by single-exponential kinetics, in agreement with predictions from theoretical studies.¹³ A fit of the data gives a time constant for contact formation ($\tau = 1/k_c$) of $57(\pm 3)$ ns for formation of an $i, i + 29$ contact (Figures 2 and 3). The observation of single-exponential kinetics indicates that: (i) the rates of interconversion between the different conformations in the ensemble of random coil states is fast and thus allows the chain to maintain the equilibrium distribution of the ensemble of conformations that has not made contact; and (ii) the number of equilibrium conformations that have donor and acceptor in van der Waals contact is small compared to the total number of states.¹³ Rapid interconversion between the different conformations in the ensemble of unfolded states is in agreement with sub-nanosecond relaxation kinetics in a cyclic azo-peptide after a photo-induced switch from the *cis* to the *trans* conformation.¹⁴ The TTET kinetics observed in our experiments cannot be due to intermolecular exchange processes. The peptide concentrations used in the experiments are typically in the 50–100 μM range, which should not allow bimolecular processes with half-times faster than 50 μs , even for the shortest peptides. This was confirmed in experiments using a mixture of peptides containing either xanthone or naphthalene both present at the same concentration as in the intramolecular transfer experiments. In these experiments, no TTET

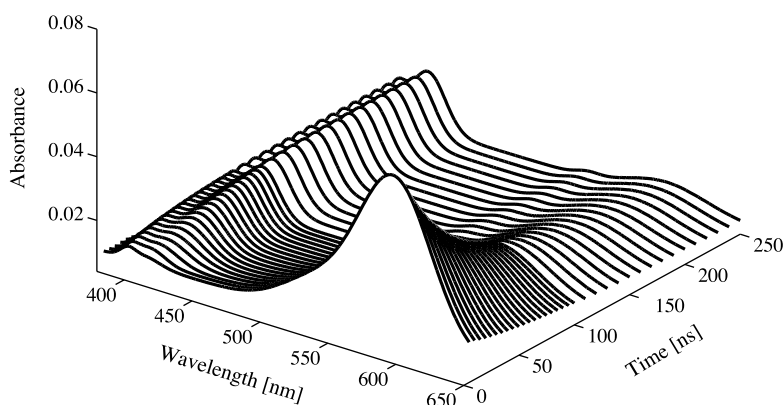


Figure 2. Time-dependent change in the absorbance spectrum of a Xan-(Gly-Ser)₁₄-NAla-Ser-Gly peptide after a 4 ns laser flash at 355 nm. The decay in the intensity of the xanthone triplet absorbance band around 590 nm is accompanied by a corresponding increase in the naphthalene triplet absorbance band around 420 nm.

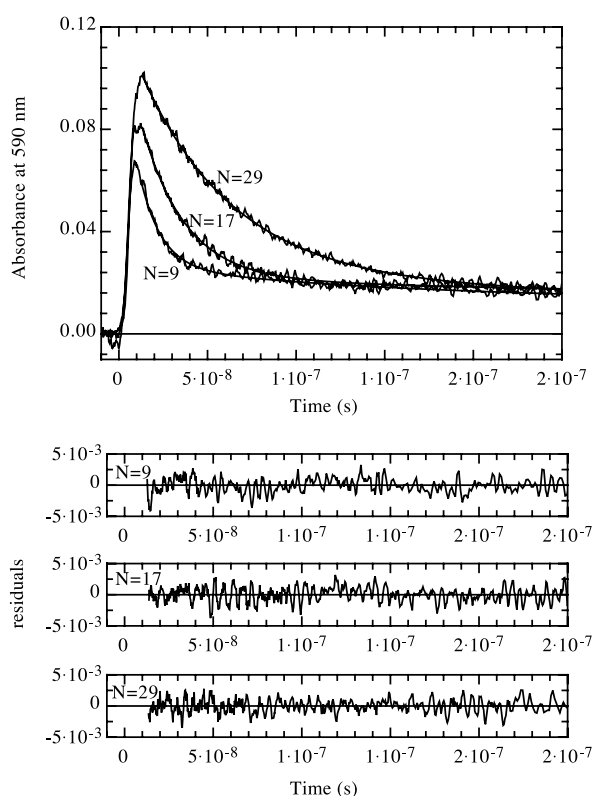


Figure 3. Time-course of formation and decay of xanthone triplets in peptides of the form Xan-(Gly-Ser)_n-NAla-Ser-Gly after a 4 ns laser flash at $t=0$ and measured by the change in absorbance of the xanthone triplets at 590 nm. Data for different numbers of peptide bonds (N) between donor and acceptor are displayed. Additionally, single-exponential fits of the data (continuous line) and the corresponding residuals are displayed. The fits gave time constants of $11.6(\pm 0.4)$ ns, $25.0(\pm 1.3)$ ns and $57.1(\pm 3.3)$ ns for $N = 9, 17$, and 29 , respectively. At the end of the exponential decay a small amount of xanthone groups remain in the triplet state. These triplets decay on a much slower time-scale with a rate corresponding to the intrinsic lifetime of xanthone triplets ($t_{1/2} \approx 30 \mu\text{s}$). For comparison, all curves were normalized to identical amplitudes from the fit of the fast decay, i.e. all curves meet at $t = 0$, corresponding to the start of the laserflash. Chain conformations that have donor and acceptor in close contact at the time-point of the laserflash should lead to fast TTET during the 4 ns laser flash. Comparison with the kinetics of triplet decay of free xanthone shows that this fraction is less than 2% for longer peptides. In short peptides this fraction is more difficult to determine due to an uncertainty in $t = 0$ compared to the fast time-scale of diffusional TTET in these peptides. We can, however, estimate that the fraction of rapid transfer does not exceed 10% even in the shortest peptides used in our studies.

from xanthone to naphthalene was observed (data not shown). We can further neglect contributions from through-bond transfer processes, since this cannot occur over distances beyond eight bonds.^{15,16} Even in our shortest peptides, donor and acceptor are separated by 11 bonds.

At the end of the fast exponential decay, a small

amount of xanthone remains in the triplet state (Figure 3). These triplets decay on a much slower time-scale with a rate corresponding to the intrinsic lifetime of xanthone triplets ($t_{1/2} \approx 30 \mu\text{s}$). In contrast to the fast triplet decay, which is limited by chain dynamics, the slow reaction is sensitive to oxygen concentration and is not accompanied by a corresponding increase in naphthalene triplets at 420 nm. This reaction might be due to a second triplet state of xanthone that cannot transfer its electrons to naphthalene. However, the slow reaction is more likely caused by a population of peptides that cannot form contact between donor and acceptor on the time-scale of our experiments, e.g. due to a small population of multimeric peptide associates. These triplets will decay with the intrinsic lifetime of the xanthone triplet state. This idea is supported by a decrease of the amplitude of the slow reaction observed in good solvents like urea and GdmCl solutions (data not shown).

To test the effect of donor-acceptor distance on the rates of intrachain diffusion, we measured TTET kinetics in peptides containing between one and 28 Gly-Ser units between donor and acceptor. This allows measurements of contact formation kinetics for distances ranging from $i, i+3$ to $i, i+57$, which covers the range of side-chain contacts in small proteins. Figure 3 displays three representative TTET kinetics for peptides of different

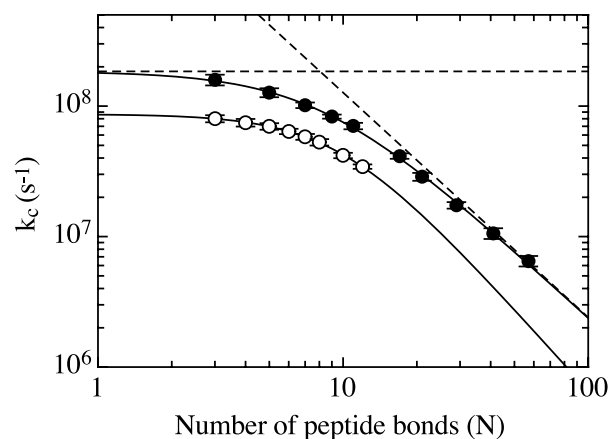


Figure 4. Effect of increasing chain length (N) on the rate constant of contact formation in a series of peptides with alternating glycine-serine (●) and with polyserine (○) between donor and acceptor. The continuous lines represent fits to the equation:

$$k_c = \frac{1}{1/k_0 + 1/(k_1 N^m)}$$

The fits give values of $k_0 = 1.8(\pm 0.2) \times 10^8 \text{ s}^{-1}$, $k_1 = 6.7(\pm 1.6) \times 10^9 \text{ s}^{-1}$ and $m = -1.72 \pm 0.08$ for the poly(glycine-serine) series, and values of $k_0 = 8.7(\pm 0.8) \times 10^7 \text{ s}^{-1}$, $k_1 = 1.0(\pm 0.8) \times 10^{10} \text{ s}^{-1}$ and $m = -2.1 \pm 0.3$ for polyserine. The broken lines represent the limiting regimes for contact formation in the poly(glycine-serine) peptides with a length-independent upper limit for contact formation of $k_0 = 1.8(\pm 0.2) \times 10^8 \text{ s}^{-1}$ and length-dependent parameters of $k_1 = 6.7 \times 10^9 \text{ s}^{-1}$ and $m = -1.72$.

length. For all polypeptide chains, single-exponential kinetics for contact formation are observed. Figure 4 shows that for long donor-acceptor distances ($x > 30$) the rate of contact formation decreases with $N^{-1.7 \pm 0.1}$ ($N = x + 1$ is the number of peptide bonds between donor and acceptor). This indicates a stronger distance-dependence than expected for purely entropy-controlled intrachain diffusion in ideal, freely-jointed Gaussian chains, which should scale with $k \sim N^{-1.5}$.^{13,17} However, Flory already pointed out that excluded volume effects should significantly influence the chain dimensions.¹⁸ Accounting for excluded volume effects in the end-to-end diffusion model presented by Szabo *et al.*¹³ leads to $k \sim N^{-1.8}$, which is nearly identical with the value found for the long poly(glycine-serine) chains. This indicates that the dimensions and the dynamics of unfolded polypeptide chains in water are influenced significantly by excluded volume effects, which is in agreement with recent results on conformational properties of polypeptides derived from simplified Ramachandran maps.¹⁹ Another source for the stronger distance dependence observed in our experiments might be an additional small enthalpic barrier, which was observed in the temperature dependence of intrachain diffusion (F.K. & T.K., unpublished results). For shorter chains ($i < 20$) the observed simple scaling law breaks down and contact formation becomes virtually independent of chain length, with a limiting value of $k_0 = 1.9 \times 10^8 \text{ s}^{-1}$. As pointed out above, the limiting rate for contact formation in short chains is not due to limits of triplet-triplet transfer (k_t in Figure 1), since this process occurs on the sub-nanosecond time-scale. Obviously, the intrinsic dynamics of polypeptide chains are limited by different processes for motions over short and over long segments. This is in agreement with polymer theory, which suggests that the properties of short chains are influenced strongly by chain stiffness, which leads to a breakdown of theoretically derived scaling laws for ideal chains.²⁰

To characterize the dimensions and the stiffness of polymer chains, Flory introduced the characteristic ratio (C_n):²⁰

$$C_n = \frac{\langle r^2 \rangle_0}{nl^2}$$

for a chain with n segments of length l . The characteristic ratio relates the increase in the mean square end-to-end distance $\langle r^2 \rangle_0$ of an unperturbed polymer chain, i.e. of a real chain without intramolecular interactions, to the change in $\langle r^2 \rangle_0$ of an ideal random walk or freely jointed chain, which grows with $\langle r^2 \rangle_0 = nl^2$. For short chains, C_n increases with chain length due to preferential chain propagation into one direction. For long chains, C_n reaches a constant limiting value (C_∞). In this limit, $\sqrt{\langle r^2 \rangle_0}$ grows proportional to $n^{1/2}$, like an ideal random-walk chain. However, in a real chain, $\sqrt{\langle r^2 \rangle_0}$ increases by a factor of C_∞ more

per segment compared to an ideal random-walk chain. The value of C_∞ gives the average number of consecutive chain segments that propagate in the same direction ("statistical segment"). Larger values of C_∞ therefore indicate stiffer chains. For polypeptide chains, polyglycine represents the most flexible chain with $C_\infty = 2$.^{21,22} The 1:1 mixture of glycine and serine used in our experiments is expected to have a C_∞ value of about 3 and C_∞ should be reached for $N > 10$.²³ Figure 4 shows, however, that the poly(glycine-serine) chains behave like random chains only over distances longer than 20–30 amino acid residues when the contact rates scale with $k_c \sim N^{-1.7}$. This indicates increased chain stiffness compared to the predicted value, which could be due to specific intrachain hydrogen bonds or van der Waals interactions and to excluded volume effects that restrict the number of chain conformations (see below).

Effect of amino acid sequence on intrachain diffusion

Both theoretical considerations^{21–23} and experimental results²⁴ show that polypeptide chains are especially flexible around glycyI residues. All amino acids except proline ($C_\infty > 100$) and glycine ($C_\infty = 2$) are predicted to have C_∞ values around 9, and C_∞ should be reached for intrachain distances longer than about 40–50 amino acid residues.^{21–23} This indicates increased chain stiffness and longer mean end-to-end distances ($\sqrt{\langle r^2 \rangle_0}$) compared to the poly(glycine-serine) chains. To test the effect of chain stiffness on peptide dynamics, we measured the rates of contact formation in polyserine peptides. Figure 4 shows the distance dependence of intrachain diffusion in a series of Xan-(Ser)_x-NAla-Ser-Gly peptides with $x = 2–11$ ($N = 3–12$). We observed single exponential kinetics for contact formation in all polyserine peptides. For short chains ($N < 5$), contact formation is virtually independent of chain length, with a limiting value of $k_0 = 8.7 \times 10^7 \text{ s}^{-1}$. This shows that the local dynamics in polyserine are about twofold slower than in the poly(glycine-serine) peptides, which seems to be small compared to the largely different properties expected for the stiffer polyserine chains. A possible origin of this rather small effect is that the decreased flexibility and the longer end-to-end distances in the polyserine chains are compensated by the decreased conformational space available for polyserine compared to poly(glycine-serine) peptides. For longer polyserine chains, contact formation slows with increasing chain length. The effect of increasing chain length on the rates of contact formation seems to be slightly larger in polyserine compared to the poly(glycine-serine) series ($k \sim N^{-2.1 \pm 0.2}$). However, due to limitations in peptide synthesis, we were not able to obtain longer peptides, which will be required to get a reliable scaling law for the polyserine peptides.

The kinetics of contact formation over short

Table 1. Effect of different amino acids on local peptide dynamics

Xaa	k_c (10^7 s $^{-1}$)
Gly	12 ± 1
Ala	8.0 ± 0.7
Ser	6.7 ± 0.7
Glu	5.4 ± 0.2
Arg	5.5 ± 0.7
His	4.9 ± 0.4
Ile	4.4 ± 0.3
<i>trans</i> Pro	2.0 ± 0.3
<i>cis</i> Pro	25 ± 5

The rate constants for end-to-end contact formation (k_c) were measured by TTET from xanthone (Xan) to 1-naphthyl alanine (NAla) in host guest peptides of the sequence Xan-Ser-Xaa-Ser-NAla-Ser-Gly with varying guest amino acid (Xaa). Conditions were 22.5 °C, pH 7 in water.

distances differ only by a factor of 2 for polyserine compared to poly(glycine-serine) (Figure 4). To investigate the effect of other amino acids on local chain dynamics, we measured TTET in short host-guest peptides of the canonical sequence Xan-Ser-Xaa-Ser-NAla-Ser-Gly using the guest amino acid residues Xaa = Gly, Ser, Ala, Ile, His, Glu, Arg and Pro. Table 1 shows that the amino acid side-chain indeed has only little effect on the rates of contact formation. All amino acids except proline and glycine show very similar dynamics. Interestingly, there is a small but significant difference in rate between short side-chains (Ala, Ser) and amino acids with longer side-chains (Ile, Glu, Arg, His). Obviously, chains that extend beyond the C $^{\beta}$ atom slightly decrease the rates of local chain dynamics, whereas charges do not influence the dynamics. Glycyl and prolyl residues show significantly different dynamics compared to all other amino acids, as expected from their largely different conformational properties.²² As seen before (Figure 4), glycine accelerates contact formation about twofold compared to serine. Proline shows slower and more complex kinetics of contact formation with two relaxation times of $1/k_1 = 2.5 \times 10^8$ s $^{-1}$ and $1/k_2 = 2 \times 10^7$ s $^{-1}$ and respective amplitudes of $A_1 = 80(\pm 5)\%$ and $A_2 = 20(\pm 5)\%$. This essentially reflects the *cis-trans* ratio at the Ser-Pro peptide bond in our host-guest peptide, which has a *cis* content of $16(\pm 2)\%$ as determined by 1D ^1H NMR spectroscopy using the method described by Reimer *et al.*²⁵ Since the rate of *cis-trans* isomerization is slow ($1/k \sim 20$ seconds at 22 °C) there is no equilibration between the two isomers on the time-scale of the TTET experiments. This allows the measurement of the dynamics of both the *trans* and the *cis* form. Our results show that the two isomers differ significantly in their dynamic properties of $i, i + 4$ contact formation.

Effect of solvent on chain dynamics

The dynamics of short poly(glycine-serine)

peptides measured with the xanthone–naphthalene TTET pair in water are about three times faster than the previously measured rates in the same peptides using the thioxanthone–naphthalene pair in EtOH,¹ although both systems were shown to be diffusion-controlled. To test whether this difference originates in the influence of the solvent on chain dynamics we measured the EtOH-dependence of intrachain diffusion using the xanthone–naphthalene pair. Figure 5(A) shows that the logarithm of k_c decreases linearly with increasing concentration of EtOH in a Xan-(Gly-Ser) $_4$ -NAla-Ser-Gly peptide. A possible origin of this effect is an influence of the solvent on the chain properties. EtOH is a better solvent for polypeptide chains than water and should thus lead to a more extended ensemble of unfolded states. To test whether good solvents generally slow intrachain diffusion, we measured the effect of GdmCl and urea on the contact rates in aqueous solutions. Both denaturants show similar effects on intrachain diffusion as EtOH. They lead to a linear decrease in $\ln k_c$ with increasing concentration of

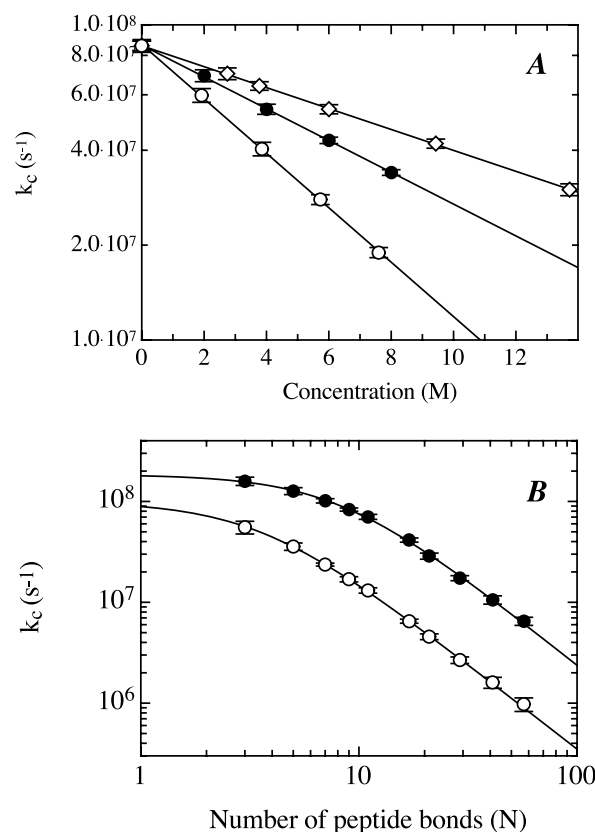


Figure 5. (A) Effect of various co-solvents on the dynamics of intrachain contact formation in a Xan-(Gly-Ser) $_4$ -NAla-Ser-Gly peptide. TTET was measured in aqueous solutions in the presence of ethanol (\diamond), urea (\bullet) and GdmCl (\circ). (B) Effect of donor–acceptor distance (N) on the rate constant of contact formation in a series of peptides with alternating glycine-serine between donor and acceptor. The rate constants for contact formation in water (\bullet) are compared to the values in 8 M GdmCl (\circ).

denaturant. Interestingly, the change in $\ln k_c$ with concentration of denaturant ($m_c = d \ln k_c / d[\text{Den}]$) is twofold higher for GdmCl compared to urea, which essentially corresponds to their relative strength in unfolding proteins.²⁶ Similar m_c -values for EtOH, urea and GdmCl are observed in a Xan-(Ser)₉-NAla-Ser-Gly peptide (data not shown). The observed effect is significantly stronger than expected from the increased solvent viscosity in concentrated GdmCl and urea solutions (F.K. & T.K., unpublished results).

Since the effect of GdmCl and urea on the chain dynamics suggests that these co-solvents change the chain properties significantly, we measured the effect of GdmCl on the distance-dependence of intrachain diffusion. Figure 5(B) shows that the distance-dependence of contact formation in the Xan-(Gly-Ser)_{*n*}-NAla-Ser-Gly series is changed at high concentrations of denaturant compared to water. In 8 M GdmCl, the switch from the distance-independent dynamics over short contact distances to the distance-dependent regime for longer peptides occurs at shorter distances. This indicates decreased chain stiffness in 8 M GdmCl compared to water, although the dynamics are slowed significantly in the presence of GdmCl both for short and for long donor-acceptor distances. For long peptides, the effect of donor-acceptor distance on the rates of contact formation is similar in 8 M GdmCl and water with $k_c \sim N^{-1.8 \pm 0.1}$.

Structure and dynamics of unfolded polypeptide chains

Our results show that both in the poly(glycine-serine) peptides and in polyserine, the rate-limiting steps for motions over short and over long distances are different. For long poly(glycine-serine) chains ($N > 30$) contact formation scales with $N^{-1.7 \pm 0.1}$ as expected for an ideal chain with contributions from excluded volume effects. For shorter chains the dynamics reach a limiting value due to limited chain stiffness. The less flexible polyserine chains show about twofold slower local dynamics. These results are in contrast to theoretical considerations based on the contributions of chain entropy to the process of loop formation, which predicted that intrachain contact formation should show a maximum rate for an $i, i+10$ interaction²⁷ and should decrease for longer and shorter chains. Our experiments rather show that the process of intrachain contacts continues to become faster for interactions shorter than $i, i+10$ in all peptides and asymptotically approaches a limiting value for short chains (Figure 4).

Denaturants like GdmCl and urea slow local chain dynamics but lead to more flexible chains that behave like ideal polymers already at shorter end-to-end distances (Figure 5(B)). These seemingly contradicting findings can be rationalized by considering the effect of denaturants on the conformational properties of polypeptide chains. Unlike water, solutions with high

concentrations of denaturants represent good solvents for polypeptide chains. This reduces the strength of intramolecular interactions like hydrogen bonds and van der Waals interactions relative to peptide-solvent interactions, which makes unstructured polypeptide chains more flexible and leads to a behavior expected for an unperturbed chain. In agreement with this interpretation, the distance-dependence of the rates of contact formation in 8 M GdmCl is close to the behavior of an unperturbed polypeptide chain predicted by Flory and co-workers.²³ The decreased rate of contact formation at high concentrations of denaturant can be explained, in part, by an increased solvent viscosity and by an increased end-to-end distance that is expected in good solvents compared to water. Additional effects like denaturant binding might also slow chain dynamics.

Another interesting feature of the dynamics of contact formation are single exponential kinetics seen in all peptides. This is in agreement with our earlier results from the thioxanthone-naphthalene pair in EtOH. Due to the additional barriers that limit Trp-triplet quenching by Cys and DBO quenching by Trp, no direct information on the dynamics of contact formation systems is available from these systems.²⁻⁴ For protein folding without significant energy barriers (“downhill protein folding”)²⁸ complex kinetics with “stretched exponential” behavior ($A = A_0 \exp - (kt)^\beta$) were predicted.²⁹ Our results from the kinetics of contact formation show that even the elementary steps of protein folding show simple exponential behavior. The dynamics of intrachain contact formation are kinetically similar to the proposed downhill scenario for protein folding: many states with similar free energy are in rapid equilibrium and many different routes to the formation of productive contacts exist. This suggests that protein folding in the absence of barriers should show exponential kinetics after a rapid equilibration process in the ensemble of unfolded conformations. Experimentally observed stretched exponential kinetics for protein folding on the milliseconds time-scale²⁹ might have a different origin. They could be caused by the presence of local energy minima and/or parallel folding pathways, e.g. due to prolyl and non-prolyl peptide isomerization in addition to the direct folding process. The observation of single-exponential folding kinetics in the absence of energy barriers is in agreement with theoretical considerations by Zwanzig.^{30,31}

The earliest steps in protein folding

Our results show that the amino acid sequence has only little effect on the local dynamics of polypeptide chains. All amino acids show very similar rates of end-to-end diffusion with time constants between 12 ns and 20 ns for the formation of $i, i+4$ contacts. Polypeptide chains are significantly more flexible around glycine ($\tau = 8$ ns) and stiffer around prolyl residues ($\tau = 50$ ns for the *trans*

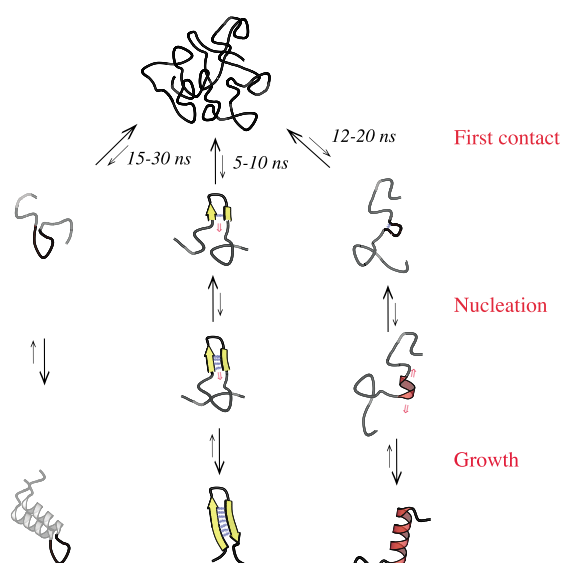


Figure 6. A representation of the time constants for the first steps in formation of loops, β -hairpins and α -helices during protein folding derived from the data measured by TTET in water.

isomer). Presumably due to the shorter distances resulting from a *cis* peptide bond the rates of $i, i + 4$ contact formation are fastest around the *cis* Ser-Pro bond ($\tau = 4$ ns). These results allow us to set an upper limit to the rates of formation of the first productive local contacts during protein folding (Figure 6). β -Hairpins, which are the most local structures in proteins, are often glycine and proline-rich.³² Gly-Ser is actually one of the most frequent sequences found in hairpin loops. Our results indicate that the time constants for the first steps in the formation of the tightest turns with $i, i + 3$ contacts are around 5 ns for Gly and *cis* Xaa-Pro (Table 1). In glycine and proline-free turns, these rates are slowed to about 10–20 ns, depending on the amino acid sequence. Formation of α -helices is most likely initiated by formation of a helical turn, which involves formation of an $i, i + 4$ interaction.^{33,34} Since helices are usually free of glycol and prolyl residues, the initiation cannot occur faster than about 12–20 ns (Figure 4 and Table 1). It should be noted that these rates do not represent the time constants for nucleation of helices and hairpins, which requires formation of more than one specific interaction and most likely encounters additional entropic and enthalpic barriers (Figure 6). Our results rather represent an upper limit for the kinetics of secondary structure formation and can thus be used to calculate the height of the free energy barriers once rates for helix and hairpin nucleation can be measured directly. The measured time constants for intrachain contact formation can, however, be used directly to predict the kinetics of loop formation, which are typically of the size of six to ten amino acid residues.³⁵ Contacts at the end of loops form

with time constants of about 15 ns for glycine-rich loops and 30–40 ns for stiffer loops.

These considerations show that proteins will not be able to fold faster than on the 10–20 ns time-scale due to the limitations set by chain dynamics. A value of 10^8 s⁻¹ therefore seems a reasonable pre-exponential factor (k_0) for determining the free energy barriers ($\Delta G^{0\ddagger}$) for folding reactions starting from the ensemble of unfolded states with:

$$k = k_0 e^{-\Delta G^{0\ddagger}/RT}$$

It is likely, however, that the pre-exponential factor changes along the reaction coordinate when intramolecular interactions are formed.

It may be argued that intrachain contacts in a polypeptide chain with a stronger bias towards folded structures can form faster than the observed dynamics in our unstructured model peptides. However, weak interactions like van der Waals contacts between side-chains and hydrogen bonds should dominate the early interactions during the folding process and our data will provide a good model for these earliest events in folding. A stronger energy bias towards the native state will mainly increase the strength or the number of these interactions but not their dynamics of formation.

The rates of intrachain diffusion measured by TTET in water are significantly faster than the values from the tryptophan–cysteine quenching pair² and for the DBO quenching by tryptophan.⁴ This is expected, since both systems are not completely diffusion-controlled. Especially, tryptophan triplet quenching by cysteine is limited by large additional barriers. It is further accompanied by radical formation and gives only minor absorbance changes, which introduces additional uncertainties in the data analysis.^{2,36} Thus, the rates for contact formation from this system had to be estimated using various assumptions. Even after correction, the rate constants obtained from these experiments are significantly smaller than the values we obtained using TTET from xanthone to naphthalene.⁹ These results show that it is essential to use diffusion-controlled systems in order to obtain accurate model-free rates on intrachain diffusion processes.

Materials and Methods

Peptide synthesis

All peptides were synthesized using standard fluoroenylmethoxycarbonyl (F-moc) chemistry on preloaded TentaGel S Trityl-resins (Rapp Polymer). 9-Oxoxanthene-2-carboxylic acid was synthesized as described³⁷ and attached to the N terminus of the peptides using F-moc chemistry. Peptides were purified by reversed-phase HPLC (RP-8-column). Purity of all peptides, which was checked by mass spectroscopy and HPLC, exceeded 98% for short peptides and 95% for long peptides. The mass was confirmed by electrospray ionization mass spectroscopy.

Laser flash experiments

Transient triplet absorption decay data were collected using a Laser Flash Photolysis Reaction Analyzer (LKS.60) from Applied Photophysics. Xanthone as triplet donor was excited selectively by using a Quantel Nd:YAG-Laser (354.6 nm, 4 ns pulse of 50 mJ). All solutions were degassed before the measurements to minimize triplet quenching by oxygen. Peptide concentrations were determined by UV absorbance at 343 nm using a molar absorption coefficient of $3900 \text{ M}^{-1} \text{ cm}^{-1}$ for xanthone in water and were typically in the 50–100 μM range. All measurements were performed at 22.5 °C and pH 7 in water. The data for contact formation for the Xan-Ser-Pro-Ser-NAla-Ser-Gly peptide (Table 1) were obtained by measuring kinetics at different concentrations of GdmCl between 0 M and 8 M, and extrapolating the data to zero denaturant. This improved the accuracy of the data for the faster kinetic phase corresponding to the dynamics of the *cis* Ser-Pro conformation. The concentrations of all guanidine hydrochloride and urea solutions were calculated from the refractive index.²⁶ The concentration of ethanol was adjusted by net weight. Transient triplet absorption of xanthone and of naphthalene was measured at 590 nm and 420 nm, respectively. Individual kinetics were typically measured five times, averaged and analyzed using ProFit (Quansoft, Zürich, Switzerland). All given rate constants were obtained by fitting the decay of xanthone triplets at 590 nm, which is more accurate than fitting the increase of naphthalene triplets at 420 nm due to small contributions of xanthone fluorescence in the early time-scales at this wavelength.

Acknowledgements

We thank Annett Bachmann and Andreas Möglich for discussion and comments on the manuscript. This work was supported by a grant from the Schweizerischen Nationalfonds (SNF).

References

- Bieri, O., Wirz, J., Hellrung, B., Schutkowski, M., Drewello, M. & Kiefhaber, T. (1999). The speed limit for protein folding measured by triplet-triplet energy transfer. *Proc. Natl Acad. Sci. USA*, **96**, 9597–9601.
- Lapidus, L. J., Eaton, W. A. & Hofrichter, J. (2000). Measuring the rate of intramolecular contact formation in polypeptides. *Proc. Natl Acad. Sci. USA*, **97**, 7220–7225.
- Yeh, I.-C. & Hummer, G. (2002). Peptide loop-closure kinetics from microsecond molecular dynamics simulations. *J. Am. Chem. Soc.* **124**, 6563–6568.
- Hudgins, R. R., Huang, F., Gramlich, G. & Nau, W. M. (2002). A fluorescence-based method for direct measurements of submicrosecond intramolecular contact formation in biopolymers: an exploratory study with polypeptides. *J. Am. Chem. Soc.* **124**, 556–564.
- Klessinger, M. & Michl, J. (1995). *Excited States and Photochemistry Of Organic Molecules*, VCH, Weinheim.
- Murov, S. L., Carmichael, I. & Hug, G. L. (1993). *Handbook of Photochemistry*, Marcel Dekker, New York.
- Damschen, D. E., Merritt, C. D., Perry, D. L., Scott, G. W. & Talley, L. D. (1978). Intersystem crossing kinetics of aromatic ketones in the condensed phase. *J. Phys. Chem.* **82**, 2268–2272.
- Anderson, R. W., Jr, Hochstrasser, R. M., Lutz, H. & Scott, G. W. (1974). Direct measurements of energy transfer between triplet states of molecules in liquids using picosecond pulses. *J. Chem. Phys.* **61**, 2500–2506.
- Lapidus, L. J., Steinbach, P. J., Eaton, W. A., Szabo, A. & Hofrichter, J. (2002). Effects of chain stiffness on the dynamics of loop formation in polypeptides. Appendix: testing a 1-dimensional diffusion model for peptide dynamics. *J. Phys. Chem. B*, **106**, 11628–11640.
- Ley, C., Morlet-Savary, F., Fouassier, J. P. & Jacques, P. (2000). The spectral shape dependence of xanthone triplet-triplet absorption on solvent polarity. *J. Photochem. Photobiol. A*, **137**, 87–92.
- Grabner, G., Rechthaler, K., Mayer, B., Köhler, G. & Rotkiewicz, K. (2000). Solvent influence on the photo-physics of naphthalene: fluorescence and triplet state properties in aqueous solutions and in cyclodextrin complexes. *J. Phys. Chem. A*, **104**, 1365–1376.
- Garner, A. & Wilkinson, F. (1976). Laser photolysis of the triplet state of xanthone and its ketyl radical in fluid solution. *J. Chem. Soc. Faraday Trans. 2* **72**, 1010–1020.
- Szabo, A., Schulten, K. & Schulten, Z. (1985). First passage time approach to diffusion controlled reactions. *J. Chem. Phys.* **72**, 4350–4357.
- Spörlein, S., Carsten, H., Satzger, H., Renner, C., Behrendt, R. & Moroder, L. (2002). Ultrafast spectroscopy reveals subnanosecond peptide conformational dynamics and validates molecular dynamics simulation. *Proc. Natl Acad. Sci. USA*, **99**, 7998–8002.
- Closs, G. L., Johnson, M. D., Miller, J. R. & Piotrowiak, P. (1989). A connection between intramolecular long-range electron, hole and triplet energy transfer. *J. Am. Chem. Soc.* **111**, 3751–3753.
- Wagner, P. J. & Klán, P. (1999). Intramolecular triplet energy transfer in flexible molecules: electronic, dynamic, and structural aspects. *J. Am. Chem. Soc.* **121**, 9626–9635.
- Jacobsen, H. & Stockmayer, W. H. (1950). Intramolecular reaction in polycondensations. I. The theory of linear systems. *J. Phys. Chem.* **18**, 1600–1606.
- Flory, P. J. (1953). *Principles of Polymer Chemistry*, Cornell University Press, Ithaca.
- Pappu, R. V., Srinivasan, R. & Rose, G. D. (2000). The Flory isolated-pair hypothesis is not valid for polypeptide chains: implications for protein folding. *Proc. Natl Acad. Sci. USA*, **97**, 12565–12570.
- Flory, P. J. (1969). *Statistical Mechanics of Chain Molecules*, Hanser Publishers, Munich.
- Brant, D. A. & Flory, P. J. (1965). The configuration of random polypeptide chains. II. Theory. *J. Am. Chem. Soc.* **87**, 2791–2800.
- Schimmel, P. R. & Flory, P. J. (1967). Conformational energy and configurational statistics of poly-L-proline. *Proc. Natl Acad. Sci. USA*, **58**, 52–59.
- Miller, W. G., Brant, D. A. & Flory, P. J. (1967). Random coil configurations of polypeptide chains. *J. Mol. Biol.* **23**, 67–80.
- Schwalbe, H., Fiebig, K. M., Buck, M., Jones, J. A., Grimshaw, S. B. & Spencer, A. (1997). Structural and

- dynamical properties of a denatured protein. Heteronuclear 3D NMR experiments and theoretical simulations of lysozyme in 8 M urea. *Biochemistry*, **36**, 8977–8991.
25. Reimer, U., Scherer, G., Drewello, M., Kruber, S., Schutkowski, M. & Fischer, G. (1998). Side-chain effects on peptidyl-prolyl *cis/trans* isomerization. *J. Mol. Biol.* **279**, 449–460.
 26. Pace, C. N. (1986). Determination and analysis of urea and guanidine hydrochloride denaturation curves. *Methods Enzymol.* **131**, 266–280.
 27. Camacho, C. J. & Thirumalai, D. (1995). Theoretical predictions of folding pathways by using the proximity rule, with applications to bovine pancreatic inhibitor. *Proc. Natl Acad. Sci. USA*, **92**, 1277–1281.
 28. Eaton, W. A., Thompson, P. A., Chan, C. K., Hagen, S. J. & Hofrichter, J. (1996). Fast events in protein folding. *Structure*, **4**, 1133–1139.
 29. Sabelko, J., Ervin, J. & Gruebele, M. (1999). Observation of strange kinetics in protein folding. *Proc. Natl Acad. Sci. USA*, **96**, 6031–6036.
 30. Zhou, H.-X. & Zwanzig, R. (1991). A rate process with an entropic barrier. *J. Chem. Phys.* **94**, 6147–6152.
 31. Zwanzig, R. (1997). Two-state models for protein folding. *Proc. Natl Acad. Sci. USA*, **94**, 148–150.
 32. Wilmot, C. M. & Thornton, J. M. (1988). Analysis and prediction of the different types of β -turns in proteins. *J. Mol. Biol.* **203**, 221–232.
 33. Zimm, B. H. & Bragg, J. K. (1959). Theory of phase transition between helix and random coil in polypeptide chains. *J. Chem. Phys.* **31**, 526–535.
 34. Lifson, S. & Roig, A. (1961). On the theory of helix-coil transitions in polypeptides. *J. Chem. Phys.* **34**, 1963–1974.
 35. Leszczynski, J. F. & Rose, G. D. (1986). Loops in globular proteins: a novel category of secondary structure. *Science*, **234**, 849–855.
 36. Lapidus, L. J., Eaton, W. A. & Hofrichter, J. (2001). Dynamics of intramolecular contact formation in polypeptides: distance dependence of quenching rates in a room-temperature glass. *Phys. Rev. Letters*, **258101-1–258101-4**.
 37. Graham, R. & Lewis, J. R. (1978). Synthesis of 9-Oxoxanthen-2-carboxylic acids. *J. Chem. Soc., Perkin Trans. 1*, 876–881.

Edited by R. Huber

(Received 30 April 2003; received in revised form 7 July 2003; accepted 7 July 2003)

Intrachain diffusion in a protein loop fragment from carp parvalbumin

Florian Krieger, Beat Fierz, Fabian Axthelm, Karin Joder,
Dominique Meyer, Thomas Kiefhaber *

Department of Biophysical Chemistry, Biozentrum der Universität Basel, Klingelbergstr. 70, CH-4056 Basel, Switzerland

Received 17 May 2004; accepted 27 May 2004

Available online 28 July 2004

Abstract

During protein folding a polypeptide chain has to form specific intrachain interactions starting from an ensemble of unfolded conformation. Thus, intrachain diffusion in unfolded polypeptide chains can be regarded as an elementary step in protein folding, which should determine the dynamics of the early stages in the folding process. We have previously applied exothermic triplet–triplet energy transfer from xanthone to naphthalene to determine rate constants for intrachain end-to-end contact formation in unstructured homo-polypeptide chains. Here we show that the method can be applied to determine absolute rate constants for intrachain diffusion in natural loop sequences, if they are free of methionine, tryptophan and tyrosine. We measured the rate of loop formation in an 18 amino acid polypeptide chain corresponding to a natural loop sequence from carp muscle β -parvalbumin (residues 85–102). Contact formation shows single exponential kinetics with a time constant ($\tau=1/k$) of 53 ± 3 ns at 22.5 °C in water. Comparison with the results on homo-polypeptide chains shows that this value agrees well with rates obtained earlier for a polyserine chain of the same length.

© 2004 Elsevier B.V. All rights reserved.

Keywords: Protein folding; Chain dynamics; Intrachain diffusion; Triplet–triplet energy transfer; Carp-parvalbumin; Calcium-binding

1. Introduction

During protein folding a polypeptide chain has to form specific intramolecular interactions to finally reach its biologically active native state. Starting from the ensemble of unfolded conformations a polypeptide chain has to sample a vast conformational space in search for energetically favourable conformations. To understand the dynamics of the folding process it is essential to characterize the structure, the dimensions and the dynamics of unfolded and partially folded states. A large

amount of structural information on unfolded proteins and folding intermediates has been obtained from nuclear magnetic resonance spectroscopy [1–3] and small-angle X-ray scattering [3–5]. Much less is known about the dynamics of conformational sampling in these states and about the rates of interconversion between different local minima on the free energy landscape. The rate at which an unstructured polypeptide chain can explore conformational space during the earliest steps in protein folding is limited by intrachain diffusion, i.e. by the maximum rate at which two specific points on the chain can make contact. We have recently applied the method of exothermic triplet–triplet energy transfer (TTET) from xanthone to naphthalene to directly measure end-to-end diffusion rates in unstructured polypeptide chains [6,7]. TTET is a two-electron transfer process (Dexter

* Corresponding author. Tel.: +41-61-267-2194; fax: +41-61-267-2189.

E-mail address: t.kiefhaber@unibas.ch (T. Kiefhaber).

mechanism) and requires van-der-Waals contact between donor and acceptor [8]. Formation of xanthone triplet states (k_T) and electron transfer from xanthone to naphthalene (k_{TTET}) both occur on the 1–5 ps time scale (H. Satzger, W. Zinth, B.F. and T.K., in preparation) and the bimolecular transfer process is diffusion-controlled ($k_T = 4 \times 10^9 \text{ M}^{-1} \text{ s}^{-1}$) [6,7]. Thus, this system allows us to determine absolute rate constants for intrachain diffusion processes slower than about 20 ps (Fig. 1). In earlier experiments the spectroscopic labels were placed at the ends of poly(glycine–serine) and polyserine chains [6,7]. The length-dependence of end-to-end diffusion in these homopolymers showed that diffusion over long and over short distances is limited by different processes (Fig. 2). For contact formation over short distances we observed that the dynamics of intrachain diffusion are virtually independent of chain length with limiting time constants ($1/k_0$) of 5 ns for poly(glycine–serine) chains and of 12 ns for polyserine [7]. Contact formation over longer distances decreases with increasing chain length and scales with $k \sim N^{-1.7}$ for poly(glycine–serine) and $k \sim N^{-2.1}$ for polyserine with N being the number of peptide bonds between donor and acceptor. The complete length-dependence of intrachain diffusion could thus be described by the following equation (Fig. 2)

$$k_c = \frac{1}{1/k_0 + 1/(k_i \cdot N^m)}. \quad (1)$$

This suggests that chain stiffness governs local chain dynamics whereas entropy-limited conformational search sets the limit for contact formation over longer distances. These results contradicted theoretical considerations, which postulated a maximum rate constant in the distance-dependence of end-to-end diffusion for formation of an $i, i+9$ contact and predicted slower contact rates for shorter distances [9]. The effect of the amino acid sequence on local chain dynamics was

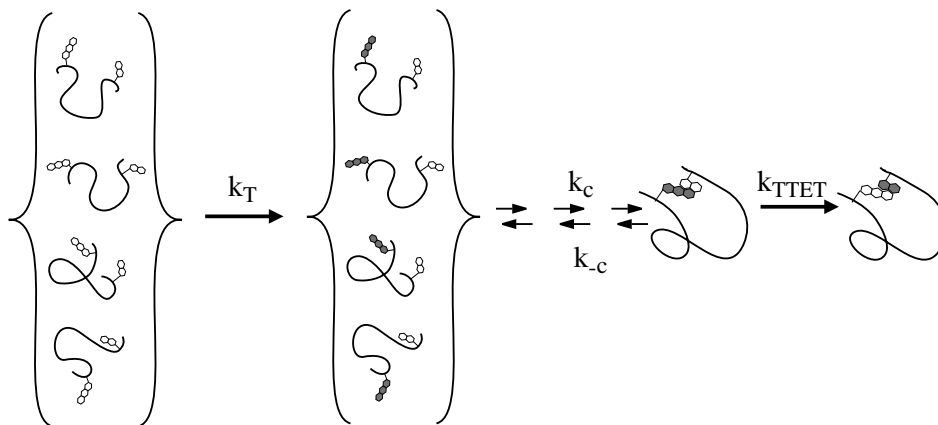


Fig. 1. Schematic of the triplet–triplet energy transfer (TTET) measurements applied to measure intrachain diffusion in polypeptide chains. The time constants for triplet formation ($1/k_T$) and electron transfer ($1/k_{TTET}$) are around 1–5 ps (H. Satzger, W. Zinth, B.F. and T.K., in preparation) and TTET is diffusion-controlled [7]. This allows measurements of absolute rate constants for intrachain diffusion (k_c) for processes slower than about 20 ps.

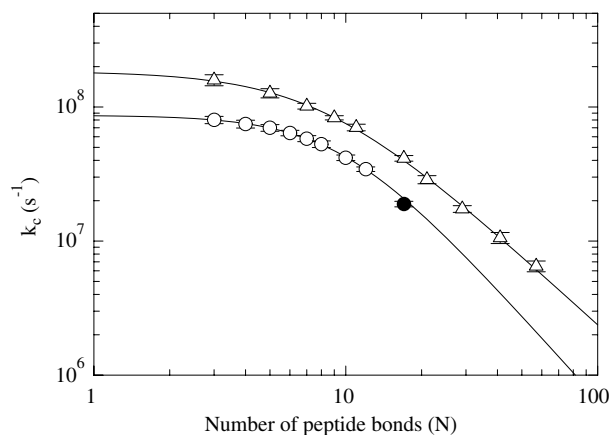


Fig. 2. Comparison of the end-to-end diffusion kinetics of the parvalbumin loop fragment 85–102 (●) with results on poly(glycine–serine) (Δ) and polyserine (○) homo-polypeptides. The experimentally measured rate constants for contact formation between xanthone and naphthalene (k_c) are plotted against the number of peptide bonds (N) between xanthone and naphthalene. The data were fitted to Eq. (1) and gave values of $k_0 = (1.8 \pm 0.2) \times 10^8 \text{ s}^{-1}$, $k_i = (6.7 \pm 1.6) \times 10^9 \text{ s}^{-1}$ and $m = -1.72 \pm 0.08$ for the poly(glycine–serine) series and of $k_0 = (8.7 \pm 0.8) \times 10^7 \text{ s}^{-1}$, $k_i = (1.0 \pm 0.8) \times 10^{10} \text{ s}^{-1}$ and $m = -2.1 \pm 0.3$ for polyserine. The data for the parvalbumin loop were not included in the fit. Data for poly(glycine–serine) and for polyserine as well as the fit of the data were taken from [7].

probed with a series of host–guest peptides and showed that formation of local contacts is only eightfold slower around the stiffest amino acid (proline) compared to the most flexible amino acid (glycine) [7]. Experiments carried out in different solvents further revealed that good solvents for polypeptide chains like ethanol and concentrated solutions of denaturants like urea and GdmCl significantly slow down chain dynamics compared to the dynamics in water [7]. This explained the faster dynamics observed in water [7] compared to our earlier studies, which were carried out in ethanol [6].

Our previous studies were mainly aimed at the scaling laws for intrachain diffusion in homo-polypeptide chains and at the effect of the amino sequence on local chain dynamics. However, polypeptide sequences from natural proteins are usually more complex and consist of different amino acids with a variety of side chains of different size and chemical properties. Here we test, whether the results from homo-polypeptides can be applied to dynamics of natural sequences. We measured intrachain contact formation between the ends of a loop region derived from a natural protein structure. Since triplet states may be quenched by a variety of chemical groups and triplet energy of xanthone may be transferred to aromatic amino acids we first determined the

effect of various amino acids on the xanthone triplet lifetime. The results show that methionine, tryptophan, tyrosine and deprotonated histidine efficiently interact with xanthone triplets. We then choose a loop sequence of β -carp parvalbumin (residues 85–102) as a model to study kinetics of intrachain diffusion. This loop sequence is perfectly suited to measure intrachain diffusion using the xanthone/naphthalene pair since none of its amino acids interferes with the TTET experiments and the loop brings together two phenylalanine residues in native carp parvalbumin, which have similar properties as our donor and acceptor group (Fig. 3). In addition, this loop binds Ca^{2+} in the native state, which allows us to test the influence of Ca^{2+} on the dynamics of the

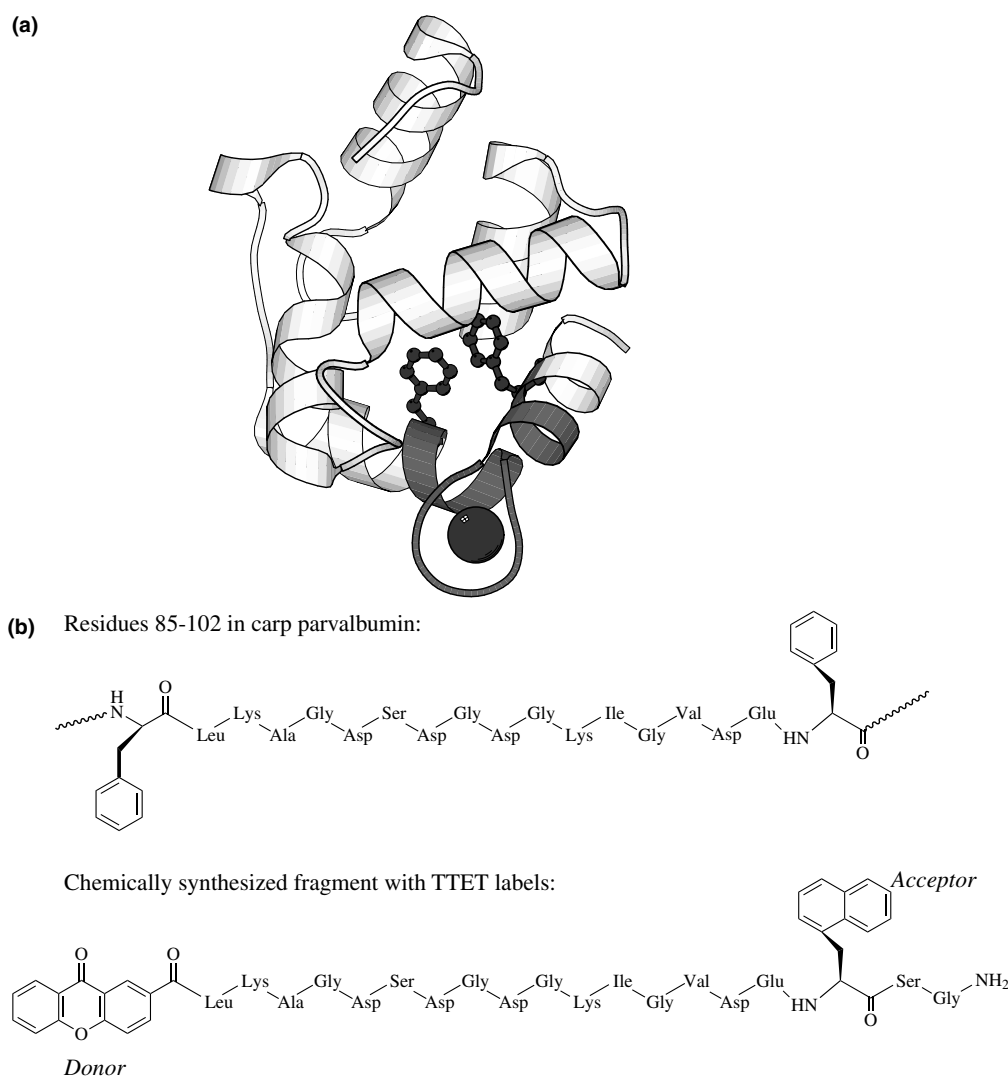


Fig. 3. (a) Ribbon diagram of the structure of carp muscle β -parvalbumin [13] with the loop region residues 85–102 shown in dark. Phe85 and Phe102 are shown as ball-and-stick models and the Ca^{2+} ion bound to the EF-loop is displayed as space-fill model. The phenylalanine residues have been replaced by the triplet donor and acceptor labels, xanthone acid and naphthylalanine, respectively. The figure was prepared using the program MolScript [20] and the pdb file 4CPV [13]. (b) Sequence of the β -carp muscle parvalbumin EF loop region (residues 85–102) and the synthesized fragment labeled with the two phenylalanine residues at positions 85 and 102 replaced by donor (xanthone) and acceptor (naphthylalanine) groups for triplet–triplet energy transfer.

isolated loop fragment. The results show that the rate constant of formation of this 18 amino acid loop corresponds well to the rate constant expected for a polyserine peptide of the same length. Ca^{2+} has no effect on the structure and chain dynamics, indicating that the binding constant for Ca^{2+} is significantly reduced compared to native parvalbumin.

2. Experimental

9-Oxoxanthene-2-carboxylic acid (xanthone acid) synthesis and synthesis of the sequence Xan–Leu–Lys–Ala–Gly–Asp–Ser–Asp–Gly–Asp–Gly–Lys–Ile–Gly–Val–Asp–Glu–NAla–Ser–Gly–NH₂ (NAla = 1-naphthylalanine) were performed as described previously [7]. Transient triplet absorption decay data were collected using a Laser Flash Photolysis Reaction Analyzer (LKS.60) from Applied Photophysics. Xanthone as triplet donor was excited selectively by using a Continuum Surelite Nd:YAG-Laser operating at 354.6 nm (~4 ns pulse width, ~100 mJ pulse energy). Peptide concentrations were determined by UV absorbance at 343 nm using a molar absorption coefficient of 3900 M⁻¹cm⁻¹ for xanthone in water and were typically in the 30–50 μM range. At these concentrations no intermolecular triplet–triplet energy transfer between two peptides occurs on the time scale of the experiments (5 ns to 50 μs) [6,7]. Transient triplet absorption of xanthone and of naphthalene was measured at 590 and 420 nm, respectively. All measurements were performed in 10 mM sodium cacodylate, pH 7.0 at 22.5 °C. 1 mM EDTA was added to remove traces of Ca^{2+} ions from the solutions, since this loop region binds Ca^{2+} with high affinity in native carp parvalbumin [10]. The effect of calcium on structure and dynamics of the loop was tested by adding various amounts of Ca^{2+} (1–100 mM) in the absence of EDTA. Sodium cacodylate, EDTA and Ca^{2+} have no influence on the xanthone triplet lifetime. Individual kinetics were measured five times, averaged and analyzed using ProFit software program. The kinetics were described by a double exponential fit. The faster reaction corresponds to TTET limited by intrachain diffusion and the slower reaction ($\tau = 1.2$ μs) represents spontaneous triplet decay of xanthone under the given conditions. The solutions were not degassed for the measurements described in this work, since the kinetics of contact formation occur on a much faster time scale than spontaneous triplet decay. Degassing the solution leads to an increase in the triplet lifetime of xanthone triplets from $\tau = 1.3$ to 10 μs, indicating that the triplet lifetime of xanthone in water is determined by oxygen quenching. The rate of intrachain diffusion is not affected by degassing the solutions [7]. The bimolecular rate constants for quenching and TTET for the different amino acids were determined by measuring xanthone

acid triplet lifetimes in the presence of the various amino acids under pseudo-first-order conditions with xanthone acid concentration was 35 μM and amino acid concentrations between 1 and 50 mM. For all amino acids which interacted with xanthone acid a linear dependence of the pseudo-first-order rate constant on the quencher concentration was observed. The slope of this dependence was used for calculating the bimolecular quenching constant (k_q). Circular dichroism (CD) measurements of the peptide fragment were performed at 22.5 °C using an Aviv 62DS spectropolarimeter equipped with a temperature-control unit.

3. Results and discussion

3.1. Effect of various amino acids on xanthone triplet lifetimes

To test how experimental results on intrachain contact formation obtained on poly(glycine–serine) and polyserine chains compare to the dynamics of natural sequences we measured the rates of intrachain contact formation between the ends of a natural loop sequence. We first tested for possible interference from amino acid side chains with the TTET reaction from xanthone to naphthalene by determining the effect of various amino acids on the xanthone triplet lifetimes. Table 1 shows bimolecular triplet quenching and TTET rate constants measured under pseudo-first-order conditions. The thioether group of methionine ($k_q = (2.0 \pm 0.1) \times 10^9$ M⁻¹s⁻¹) and the deprotonated imidazole ring of histidine ($k_q = (1.8 \pm 0.1) \times 10^9$ M⁻¹s⁻¹) quench xanthone triplets very efficiently with a rate constant close to the diffusion limit. The other amino acid side chains quench xanthone triplets either very inefficiently (Cys, His⁺, N-terminus)

Table 1
Interaction of different amino acids with the triplet state of xanthone^a

Amino acid	k_q (M ⁻¹ s ⁻¹) ^b	Conditions
Naphthyl acetic acid ^c	$(4.0 \pm 0.1) \times 10^9$	Water, pH 7
Trp ^c	$(3.0 \pm 0.1) \times 10^{9e}$	Water, pH 7
NAla ^c	$(2.8 \pm 0.1) \times 10^9$	Water, pH 7
Tyr ^c	$(2.5 \pm 0.1) \times 10^{9e}$	Water, pH 7
Met ^d	$(2.0 \pm 0.1) \times 10^9$	Water, pH 7
His ^d	$(1.8 \pm 0.1) \times 10^{9e}$	0.1 M KPP, pH 8
His ^{+d}	$(2.8 \pm 0.2) \times 10^7$	0.1 M NaOAc, pH 4
Cys ^d	$(5.1 \pm 0.2) \times 10^7$	Water, pH 7
N-terminal NH ^{3+d}	$(2.0 \pm 0.5) \times 10^6$	Water, pH 7

^a No effect on the xanthone triplet lifetime was observed for Ala, Arg⁺, Asn, Gly, Lys⁺, Phe, Ser, Asp, Asp⁻.

^b Bimolecular quenching constants were measured under pseudo-first-order conditions as described in Section 2.

^c Triplet–triplet energy transfer (TTET).

^d Triplet quenching.

^e Radical formation as side reaction.

or not at all (Ala, Arg, Asn, Asp, Gly, Lys, Ser, Phe). The aromatic amino acids tryptophan, tyrosine and phenylalanine are possible acceptors for xanthone triplets in TTET reactions. Table 1 shows that TTET between xanthone and tryptophan ($k_{\text{TTET}} = (3.0 \pm 0.1) \times 10^9 \text{ M}^{-1} \text{ s}^{-1}$) and tyrosine ($k_{\text{TTET}} = (2.5 \pm 0.1) \times 10^9 \text{ M}^{-1} \text{ s}^{-1}$) occur in diffusion-controlled reactions with virtually the same bimolecular rate constants as observed for TTET from xanthone to naphthylalanine ($k_{\text{TTET}} = (2.8 \pm 0.1) \times 10^9 \text{ M}^{-1} \text{ s}^{-1}$). However, TTET from xanthone to tryptophan and tyrosine are complex reactions with at least two observable rate constants. For both amino acids TTET is accompanied by radical formation [11,12], which explains the complex kinetics and makes them not suitable for the use as TTET acceptors in polypeptide chains. These results show that mainly methionine, tryptophan and tyrosine interfere with TTET from xanthone to naphthalene in intrachain diffusion experiments. Histidine containing sequences can be measured with this donor/acceptor pair if the pH of the solution is below 5.5.

3.2. Contact formation in an 18-residue loop fragment from β -parvalbumin

To measure intrachain contact formation in a naturally occurring protein loop we chemically synthesized a polypeptide fragment corresponding to residues 85–102 in β -carp parvalbumin (EF-loop; Fig. 3) as model system. None of the amino acids in this loop region influences the triplet states of xanthone (see Table 1). In the native state of parvalbumin the two phenylalanine residues at positions 85 and 102 are located near the C-terminus of the E-helix and near the N-terminus of the F-helix, respectively, and are in close contact with each other (Fig. 3(a)) [13]. We replaced Phe85 by xanthone and Phe102 by naphthylalanine (Fig. 3(b)) to measure intrachain diffusion between the ends of the EF-loop. The CD spectrum of the labelled peptide resembles a random coil spectrum (Fig. 4) and addition of 8 M urea has no influence on the CD signal, suggesting that the fragment is completely unfolded in water. This makes the parvalbumin EF-loop a suitable model to investigate the dynamics of intrachain contact formation in an unfolded polypeptide chain. In native carp parvalbumin the EF-loop binds calcium with a binding constant of $k_{\text{D}} = 10^{-8} \text{ M}^{-1}$ [10]. Addition of 1–100 mM CaCl_2 to the synthesized fragment did not induce structure formation (data not shown) indicating that the peptide fragment has a significantly reduced binding affinity for calcium compared to native parvalbumin.

Fig. 5 shows the time-course of TTET from xanthone to naphthalene in the synthetic parvalbumin EF-loop in the absence of Ca^{2+} . The kinetics of contact formation can be described by a single-exponential, as previously observed for all poly(glycine-serine) and polyserine pep-

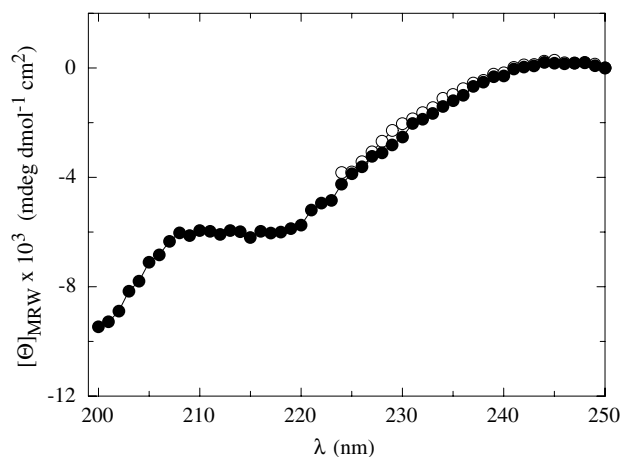


Fig. 4. Circular dichroism spectrum of chemically synthesized parvalbumin fragment 85–102 labeled with xanthone and naphthylalanine (see Fig. 3(b)). Spectra were recorded in 10 mM sodium cacodylate, 1 mM EDTA, pH 7.0 (●) and in 10 mM sodium cacodylate, 8 M urea, pH 7.0 (○).

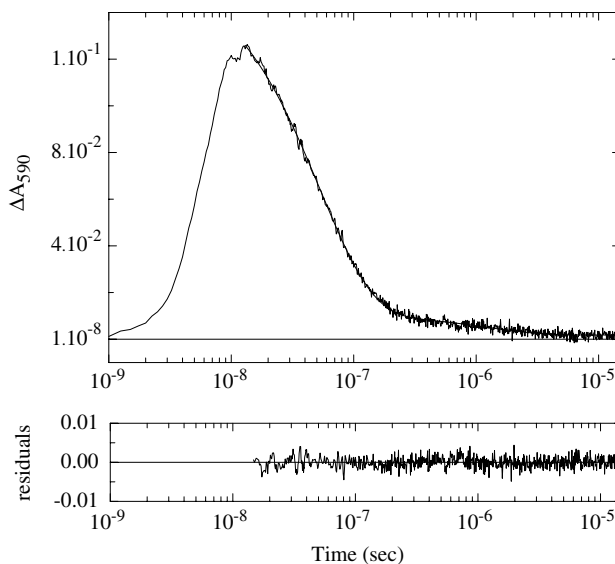


Fig. 5. Time-course of formation and decay of xanthone triplets in parvalbumin loop fragment 85–102 after a 4 ns laser flash at $t=0$. The change of xanthone triplet absorbance is measured at 590 nm. The kinetic data can be described by a double-exponential fit. The residuals of the double exponential fit are displayed. The fast main phase has a rate constant $\tau = 1/k$ of $54 \pm 3 \text{ ns}$ (95% amplitude) and corresponds to end-to-end contact formation. The second phase (5% amplitude) has the same rate constant as spontaneous triplet decay of xanthone in water ($\tau = 1.3 \pm 0.1 \mu\text{s}$). This reaction presumably corresponds to xanthone triplet decay in peptides, which are not able to form end-to-end contacts, e.g. due to intermolecular association reactions [7]. It is not observed in the presence of 8 M urea, which dissolves associates (data not shown). The experiments were carried out in 10 mM sodium cacodylate, 1 mM EDTA, pH 7.0 at 25 °C. EDTA was added to remove traces of Ca^{2+} .

tides. A fit of the data gives a time constant ($\tau = 1/k$) of $53 \pm 3 \text{ ns}$ for end-to-end contact formation. This value is in good agreement with the value expected from

TTET measurements on polyserine peptides (Fig. 2). For contact formation in a peptide with sixteen serine residues between xanthone and naphthalene the polyserine series predicts a time constant of 49 ± 4 ns for end-to-end contact formation [7]. Unfortunately, we cannot directly measure the rate to a Xan–(Ser)₁₆–NAla peptide due to limitations in peptide synthesis. The presence of four glycine residues in the EF-loop sequence should lead to faster chain dynamics in the parvalbumin loop compared to polyserine. However, residues with branched C_β-positions like valine and leucine show slightly slower local dynamics than alanine and serine [7,14], which might explain the good agreement with the polyserine data. In addition, electrostatic repulsion between the five negatively charged aspartate/glutamate residues might increase the chain dimensions and thus compensate for the effect of the flexible glycines. The presence of 10 mM CaCl₂ does not effect the kinetics of end-to-end diffusion in the parvalbumin loop ($\tau = 54 \pm 3$ ns; data not shown). This shows that the fragment is not able to bind calcium ions with high affinity in the absence of the native tertiary structure and confirms the results from the CD measurements. These results indicate that the specific conformation of this loop in the native state rather than merely its amino acid sequence is responsible for strong Ca²⁺-binding.

3.3. Comparison with dynamics of other natural and non-natural peptide sequences

Recently several other experimental systems have been used to measure the dynamics of intrachain diffusion in polypeptide chains. Quenching of tryptophan triplet states by cysteine in various peptide sequences gave significantly slower rates for chain diffusion compared to TTET from xanthone to naphthylalanine [15]. This reaction is, however, accompanied by the formation of S[•] radicals and it is far from the diffusion limit. Thus, it can not be used to measure absolute model-free rate constants for intrachain diffusion [15,16]. Significantly slower rate constants compared to a diffusion-controlled TTET are thus expected in this system. Another experimental approach used short poly(glycine–serine) peptides and measured intrachain diffusion by the quenching of 2,3 diazabicyclo[2.2.2]oct-2-ene (DBO) fluorescence by tryptophan in water [17]. This quenching reaction is close to the diffusion limit but due to the short fluorescence lifetime of DBO it is restricted to fast contact kinetics, i.e. to short fluorophor-quencher distances. The results from these experiments gave slightly slower contact rates compared to TTET from xanthone to naphthylalanine in the same peptides. Again, this is in agreement with the observation that DBO quenching by tryptophan is not completely diffusion-controlled. A third experimental approach investigated intrachain contact for-

mation in unfolded cytochrome *c* using electron transfer from a triplet excited Zn–porphyrine group to Ru bound to a specific histidine residue (His 33). This resulted in a loop size of 15 residues and yielded a time constant for electron transfer of 250 ns [18], which is significantly slower than our results on the 18 amino acid carp parvalbumin loop ($\tau = 54$ ns). However, the experiments on unfolded cytochrome *c* were performed in the presence of 5.4 M GdmCl, in order to completely unfold the protein. Our earlier experiments have shown that chemical denaturants like urea and GdmCl significantly decrease the rates of intrachain diffusion [7]. Both in poly(glycine–serine) and in polyserine peptides the presence of 5.4 M GdmCl slows down contact formation about threefold, which would result in a contact time constant of 80 ns in water for chain diffusion-limited electron transfer the 15 residues cytochrome *c* loop. The remaining small discrepancy between the dynamics in unfolded cytochrome *c* and the carp parvalbumin loop in water is probably due to the effect of end extensions on intrachain diffusion. We observed both in the poly(glycine–serine) and in the polyserine peptides that extensions of the polypeptide chains beyond the points of contact formation decreases the rates of intrachain diffusion (B.F. and T.K., manuscript in preparation). This effect will slow down contact formation in the 15 amino acid loop in unfolded cytochrome *c*, which comprises a total of 104 amino acids, compared to the synthesized parvalbumin loop, which just extends for just two additional amino acids beyond naphthylalanine (Fig. 3(b)). The rate constants for intrachain diffusion measured in unfolded cytochrome *c* using the Zn–porphyrine/Ru system are significantly faster than the dynamics estimated from intrachain diffusion in unfolded cytochrome *c* using methionine binding to heme as reporter reaction [19]. This can be explained by the reaction-controlled heme to methionine binding process, which does not allow the measurement of absolute rate constants for intrachain diffusion.

4. Conclusions

We showed that triplet–triplet energy transfer from xanthone to naphthylalanine can be used to measure absolute rate constants for end-to-end diffusion in natural loop sequences. This method will also be applicable to measure chain dynamics in unfolded and partially folded proteins provided the proteins are free of methionine, tryptophan and tyrosine. The results on the end-to-end contact formation in the EF-loop from carp parvalbumin revealed that our earlier measurements on polyserine peptides provide a good estimate for dynamics of natural protein loops. This is probably due to the compensating effect of flexible glycine residues, which leads

to faster dynamics and of C_β-branched side chains, which slow down chain dynamics.

References

- [1] D. Neri, M. Billeter, G. Wider, K. Wüthrich, *Science* 257 (1992) 1559.
- [2] J. Klein-Seetharaman, M. Oikawa, S.B. Grimshaw, J. Wirmer, E. Duchardt, T. Ueda, T. Imoto, L.J. Smith, C.M. Dobson, H. Schwalbe, *Science* 295 (2002) 1719.
- [3] W.Y. Choy, F.A. Mulder, K.A. Crowhurst, D.R. Muhandiram, I.S. Millet, S. Doniach, J.D. Forman-Kay, L.E. Kay, *J. Mol. Biol.* 316 (2002) 101.
- [4] G. Damaschun, H. Damaschun, K. Gast, D. Zirwer, *Biochemistry (Moscow)* 63 (1998) 259.
- [5] D. Segel, A. Bachmann, J. Hofrichter, K. Hodgson, S. Doniach, T. Kiefhaber, *J. Mol. Biol.* 288 (1999) 489.
- [6] O. Bieri, J. Wirz, B. Hellrung, M. Schutkowski, M. Drewello, T. Kiefhaber, *Proc. Natl. Acad. Sci. USA* 96 (1999) 9597.
- [7] F. Krieger, B. Fierz, O. Bieri, M. Drewello, T. Kiefhaber, *J. Mol. Biol.* 332 (2003) 265.
- [8] M. Klessinger, J. Michl, *Excited States and Photochemistry of Organic Molecules*, Weinheim/D, VCH, 1995.
- [9] C.J. Camacho, D. Thirumalai, *Proteins* 22 (1995) 27.
- [10] R.H. Kretsinger, *CRC Crit. Rev. Biochem.* 8 (1980) 119.
- [11] D.V. Bent, E. Hayon, *J. Am. Chem. Soc.* 97 (1975) 2612.
- [12] D.V. Bent, T. Hayano, *J. Am. Chem. Soc.* 97 (1975) 2599.
- [13] V.D. Kumar, L. Lee, B.F. Edwards, *Biochemistry* (1990) 1404.
- [14] F. Huang, W.M. Nau, *Angew. Chem. Int. Ed. Engl.* 42 (2003) 2269.
- [15] L.J. Lapidus, W.A. Eaton, J. Hofrichter, *Proc. Natl. Acad. Sci. USA* 97 (2000) 7220.
- [16] I.-C. Yeh, G. Hummer, *J. Am. Chem. Soc.* 124 (2002) 6563.
- [17] R.R. Hudgins, F. Huang, G. Gramlich, W.M. Nau, *J. Am. Chem. Soc.* 124 (2002) 556.
- [18] I.-J. Chang, J.C. Lee, J.R. Winkler, H.B. Gray, *Proc. Natl. Acad. Sci. USA* 100 (2003) 3838.
- [19] S.J. Hagen, J. Hofrichter, A. Szabo, W.A. Eaton, *Proc. Natl. Acad. Sci. USA* 93 (1996) 11615.
- [20] P. Kraulis, *J. Appl. Cryst.* 24 (1991) 946.

Ultrafast Quenching of the Xanthone Triplet by Energy Transfer: New Insight into the Intersystem Crossing Kinetics

H. Satzger,[†] B. Schmidt,[†] C. Root,[†] W. Zinth,[†] B. Fierz,[‡] F. Krieger,[‡] T. Kiefhaber,[‡] and P. Gilch^{*,†}

Department für Physik, Ludwig-Maximilians-Universität, Oettingenstrasse 67, D-80538 München, Germany, and Division of Biophysical Chemistry, Biozentrum, University of Basel, Klingelbergstrasse 50/70, CH-4056 Basel, Switzerland

Received: June 4, 2004; In Final Form: August 11, 2004

The formation and quenching of the triplet state of xanthone are studied by femtosecond techniques. As revealed by femtosecond fluorescence spectroscopy, the primarily excited $^1\pi\pi^*$ state decays within 1.5 ps. In a transient absorption experiment, this time constant is associated with a partial rise of a triplet signature. This rise has a second and slower component with a time constant of 12 ps. In the presence of high concentrations of the quencher 1-methylnaphthalene, the slow 12 ps rise component is absent. This finding gives strong evidence that the biphasic rise of the triplet absorption of xanthone is due to a sequential mechanism, namely, a $^1\pi\pi^* \rightarrow ^3n\pi^*$ with fast intersystem crossing followed by a $^3n\pi^* \rightarrow ^3\pi\pi^*$ internal conversion. Furthermore, an analysis of the concentration dependence of the quenching kinetics allows one to pin down the intrinsic transfer time of the triplet energy from xanthone to 1-methylnaphthalene to ~ 1 ps.

1. Introduction

Many biological polymers have to form specific intramolecular interactions to be able to fulfill their biological function. The rate at which polymers can explore conformational space is limited by intrachain diffusion, that is, by the rate at which two points on a polymer chain can make contact. Experimental methods are therefore needed which allow direct measurement of intrachain diffusion rates. Triplet–triplet energy transfer (TTET) between a triplet donor (D) and a triplet acceptor (A) has recently been applied to measure intrachain contact formation in polypeptide chains to gain information on the dynamics of the earliest steps in the protein folding process.^{1–3} In these experiments (see Figure 1), the two positions of interest on the polypeptide chain are labeled with a D/A pair. The donor is photoexcited, which marks the beginning of the detection period. When the donor meets the acceptor, an energy transfer process takes place transferring the triplet state to the acceptor. The transfer is detected by the concomitant absorbance changes. An ideal D/A system for such studies should have the following properties (see Figure 1): (i) The rate constant for formation of the excited donor state should be higher than the rate constant for intrachain diffusion ($k_{\text{ex}} \gg k_{\text{c}}$). (ii) The excited donor state should be very long-lived, so that the peptide dynamics can be traced throughout a large time window. (iii) The transfer rate constant between D and A should be higher than the rate constant for breaking of the contact ($k_{\text{TT}} \gg k_{\text{-c}}$), and (iv) the transfer rate should decay very rapidly with distance. A D/A system that matches these criteria is the couple xanthone (D) and naphthalene (A). Following photoexcitation, xanthone undergoes intersystem crossing (ISC) with a quantum yield close to 1⁴ preparing a long-lived triplet state ($\tau = 20 \mu\text{s}$ in water²). The energy of this state can be transferred to the naphthalene.

TTET occurs via a two-electron exchange process (Dexter mechanism); thus the transfer rate decays very rapidly with distance (and the transfer is only efficient for a close contact situation (van der Waals contact)).⁵ By incorporation of xanthone and naphthalene moieties in a peptide chain, intramolecular contact time constants on the nanosecond time scale could be determined.^{2,3} In these experiments, the donor triplet states were produced by a 4 ns laser flash and dynamics with time constants between 5 ns and several microseconds were observed, depending on the length and the sequence of the polypeptide chains.^{2,3} Tests for the presence of faster chain dynamics on the subnanosecond time scale require a characterization of the photodynamics of the xanthone/naphthalene TTET pair in detail. This yields information on the limitations of the method in the early time region set by triplet formation and TTET dynamics.

The first transient absorption experiments on xanthone with picosecond resolution^{6,7} yielded time constants for triplet formation in the 10 ps range. Later experiments with higher time resolution⁸ revealed a biphasic rise of the triplet absorption $A(t)$; that is, the rise could be described by

$$A(t) = A_1(1 - \exp(-t/\tau_1)) + A_2(1 - \exp(-t/\tau_2)) \quad (1)$$

with time constants $\tau_1 \approx 1$ ps and $\tau_2 \approx 10$ ps and amplitudes A_1 and A_2 . This biphasic behavior was attributed to a branched reaction scheme. In the proposed model (Figure 2a), the primarily excited $^1\pi\pi^*$ decays via two channels: internal conversion (IC) to the $^1n\pi^*$ state and ISC to the $^3n\pi^*$ state. The time constants of both processes were assumed to be similar. They should lead to the observed value of $\tau_1 \approx 1$ ps. Subsequently the $^1n\pi^*$ state is depleted by a second ISC process to the $^3\pi\pi^*$ state. This model obeys the El-Sayed rules for ISC processes:⁹ For an ISC process for which the type of the orbitals occupied is changed (e.g., $n\pi^* \rightarrow \pi\pi^*$) in the course of the spin flip, a high ISC rate is expected. By this simultaneous change of spin and orbital character of the electronic wave

* To whom correspondence should be addressed. Fax: +49-89/2180-9202. E-mail: Peter.Gilch@physik.uni-muenchen.de.

[†] Ludwig-Maximilians-Universität.

[‡] University of Basel.

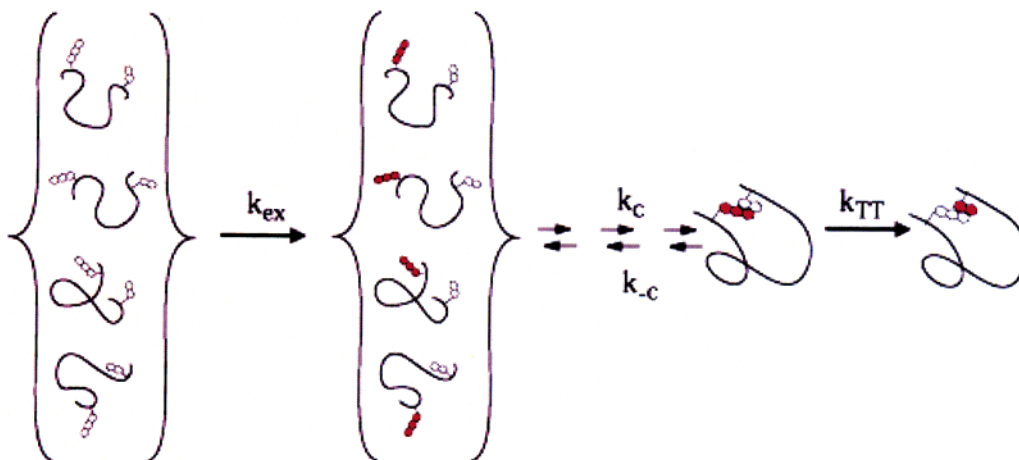


Figure 1. Principle of the experimental determination of contact rates in peptides. Two residues of a peptide chain are marked with an acceptor (A) and a donor (D), respectively. In a thermal ensemble of those marked peptides, the acceptor is photoexcited by a laser pulse. If during their diffusive motion the markers come into close contact, energy will be transferred indicating this contact.

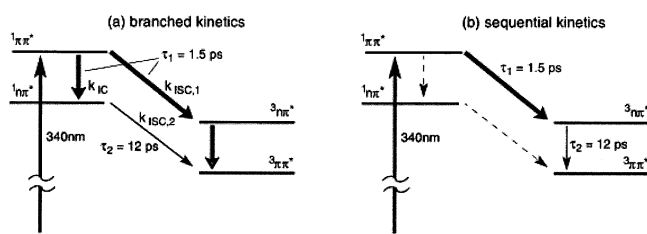


Figure 2. Jablonski diagram depicting the various excited states of xanthone and their potential decay patterns. (a) Branched kinetics. (b) Sequential kinetics. Rate constants for intersystem crossing processes are $k_{ISC,1,2}$, and those for internal conversion are $k_{IC,1,2}$. Only ISC processes allowed by the El-Sayed rules are considered. Thick arrows represent large rate constants.

function, the change of the spin angular momentum in the ISC can be compensated by a corresponding change of the orbital angular momentum. Obviously, the branched kinetic scheme proposed in ref 8 is in agreement with these rules. However, an alternative model with sequential kinetics (Figure 2b), that is, fast ISC from the $^1\pi\pi^*$ to the $^3n\pi^*$ state followed by slower internal conversion in the triplet manifold to the $^3\pi\pi^*$, cannot be ruled out. The authors of ref 8 assumed that only the final product state in the sequential process is spectroscopically visible at the single detection wavelength used in ref 8. Under this assumption, a sequential process would lead to amplitudes A_1 and A_2 with opposite sign. The observation of the same signs of A_1 and A_2 pointed to branched kinetics. However, if the intermediate state contributes to the observed time trace its extinction coefficient has to be taken into consideration and a decision between branched and sequential kinetics with the present information is not possible.

In this paper, we will readdress the ISC of xanthone and we will have a close look at the quenching of the xanthone triplet by naphthalene. We will deal with the problem by combining three experimental approaches. (i) The decay of the fluorescence of xanthone, that is, the depletion of its primarily excited state, is monitored using an ultrafast Kerr shutter. (ii) The formation of the triplet state of xanthone is followed by spectrally resolved transient absorption spectroscopy in a broad spectral range. (iii) Finally, the influence of a quencher, 1-methyl-naphthalene, on the absorption dynamics is investigated. These experiments not only aim at the dynamics of the quenching itself but will yield a deeper insight into the mechanism of the ISC of xanthone, a prerequisite for using this D/A pair as a sensor for the

investigation of peptide closure dynamics in the picosecond realm.

2. Experimental Section

The setup for femtosecond fluorescence experiments has been described before in detail in ref 10. Briefly, a Clark CPA 2001 laser/amplifier system (repetition rate of 1 kHz) pumped a two-stage noncollinear optical parametric amplifier (NOPA) tuned to 540 nm. The compressed output of the NOPA was frequency doubled in a 100 μm type I BBO crystal. The resulting 270 nm pulses with an energy of 0.2 μJ are used to excite the xanthone sample. The fluorescence light emitted by the sample is imaged onto a Kerr gate consisting of two wire grid polarizers and a fused silica plate. The Kerr gate is operated by femtosecond NIR (1100 nm) laser pulses obtained from a two-stage optical parametric amplifier (OPA) again pumped by the laser/amplifier system. The fluorescence light that passed the Kerr gate was dispersed by a spectrometer and detected by a liquid nitrogen cooled CCD camera. Parasitic stray light was suppressed by a UG 5 filter (to reject the NOPA fundamental at 540 nm) and KG 5 (to reject the ~ 1100 nm gate light). The time resolution of the setup in the near UV was ~ 200 fs (full width at half-maximum (fwhm)). The xanthone fluorescence emission lies in the blue wing of the detection window of the Kerr setup where its spectral sensitivity starts to drop drastically. This circumstance requires prolonged acquisition times which cause an instrumental artifact. The irradiation of the Kerr medium by the gate pulses induces a very weak background gradually rising with time. The rising background signal feigns a long-lived emission of xanthone. Experimental runs with the excitation light blocked, that is, in the absence of the xanthone emission, allow that contribution to be estimated and the intrinsic behavior of xanthone to be extracted.

Femtosecond absorption spectra were recorded with the setup detailed in ref 11. As in the fluorescence experiments, the sample was excited by the frequency doubled output of a NOPA pumped by a conventional 1 kHz Ti:Sa laser/amplifier system. Two wavelengths were used for the excitation process: 266 and 340 nm. The excitation energy was equal to ~ 400 nJ at the sample location for both wavelengths. Absorption changes induced by these pump pulses were probed by a white light continuum generated by focusing a portion of the amplifier fundamental into a CaF₂ plate. The white light was split into a probe and reference branch. Both probe and reference continua were spectrally dispersed and recorded in single shot fashion

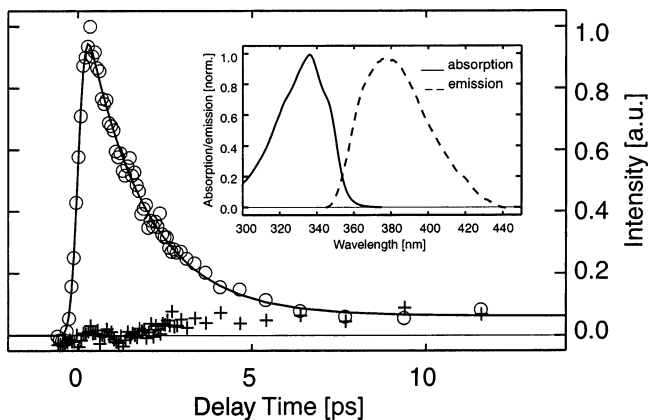


Figure 3. Decay of the fluorescence emission of xanthone at 414 nm (○) following 270 nm excitation recorded with a Kerr setup. The contribution of a slowly rising background emission is included (+). In this blank experiment, all parameters were identical to the experiment on the xanthone except for the blocked excitation light. The insert shows the absorption and fluorescence spectra of xanthone in ethanol.

by diode arrays.¹² The probe continuum overlaps in the sample with the pump light, and the reference continuum passes through a part of the sample not excited by the pump light. The pump–probe spectra were recorded with a magic angle setting.

For the femtosecond experiments, xanthone (Fluka) was dissolved in absolute ethanol (Merck, VLSI Selectipur). The concentrations of the sample solutions were in the range of 2–5 mM. The concentrations of 1-methylnaphthalene (Fluka) were adjusted by adding the respective volumes to the solutions. Xanthonic acid was synthesized and purified as described in ref 2. The concentrations of the acid in water and ethanol were about 0.5 mM. The solutions were exchanged between the laser shots by pumping them through fused silica flow cells (optical path length of 0.5 mm in the absorption experiment and 1 mm in the fluorescence experiment).

Transient absorption data in the nano- to microsecond realm were collected using a laser flash photolysis reaction analyzer (LKS.60) from Applied Photophysics. Xanthone was excited selectively by using a Quantel Nd:YAG laser (354.6 nm, 4 ns pulse of 50 mJ). For the pseudo-first-order quenching experiments, xanthone (Fluka, purum) was dissolved in ethanol (Merck, Uvasol) to a concentration of 0.05 mM measured by UV spectroscopy ($\epsilon_{335} = 6700 \text{ M}^{-1} \text{ cm}^{-1}$). 1-Methylnaphthalene (Fluka, purum; $\epsilon_{280} = 6000 \text{ M}^{-1} \text{ cm}^{-1}$) was added to concentrations ranging from 1 to 10 mM. The temperature was 295.6 K. The kinetics was typically measured five times, averaged, and analyzed using ProFit (Quansoft, Zürich, Switzerland). All rate constants were obtained by fitting the decay of the xanthone triplet at 610 nm.

Steady-state absorption and fluorescence spectra were recorded with a diode array spectrometer (Specord S100) and a Fluorolog 1680 0.22 m double spectrometer from Specs, respectively.

3. Results

3.1. Fluorescence Decay of Xanthone. When excited with UV light, xanthone emits a very weak fluorescence peaking around 380 nm (see the insert in Figure 3). A rough estimate for the lifetime of the excited $^1\pi\pi^*$ state may be obtained from the fluorescence quantum yield and the application of the Strickler–Berg relation¹³ yielding a lifetime of the $^1\pi\pi^*$ state of the order of $\tau_1 = 1$ ps. Direct experimental access to τ_1 is obtained by measuring the time dependence of the emission decay with the help of a Kerr shutter setup. A typical time trace

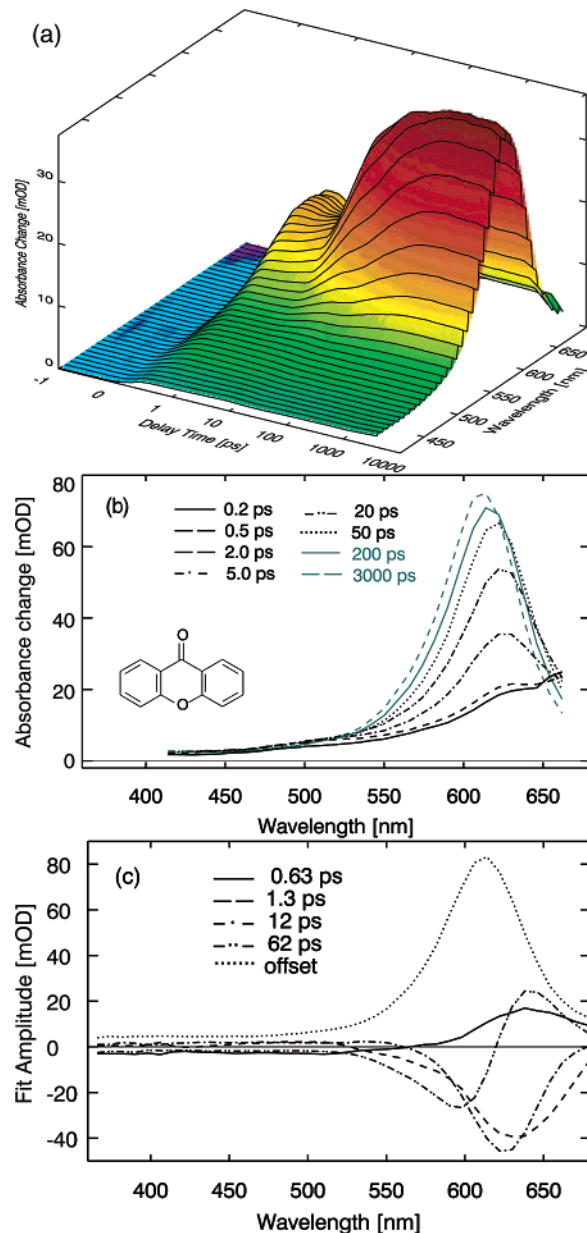


Figure 4. Transient absorption spectra of xanthone following 340 nm excitation. (a) Overview. (b) Transient spectra at distinct delay times. (c) Decay associated spectra.

is depicted in Figure 3. After an instrumentally limited rise, the fluorescence emission decays within ~ 1 ps. The weak background observed stems from emission of the Kerr medium induced by the gate pulses. This weak emission gradually rises with time and is responsible for the small offset (see the time trace recorded with the excitation light blocked in Figure 3). Thus, only the initial emission decay originates from the xanthone sample. An exponential fit considering the experimental response function yields a time constant of ~ 1.5 ps. Since the emission monitors the population of the excited $^1\pi\pi^*$ state, we can conclude that this state decays with $\tau_1 \sim 1.5$ ps. As this value is in line with the prediction based on fluorescence quantum yield, a contribution of a long-lived component, for which the Kerr setup is not very sensitive, can be ruled out.

3.2. Transient Absorption Spectroscopy. Absorption changes of xanthone with the femtosecond excitation tuned to the maximum of its $\pi\pi^*$ transition (340 nm) were recorded in a spectral region ranging from 370 to 700 nm (Figure 4a gives an overview of the data, and transient spectra at selected delay

times are depicted in Figure 4b). Immediately after photoexcitation, a weak and broad transient absorption peaking around 700 nm is formed (excited-state absorption from the $^1\pi\pi^*$ state). The decay of this signature within ~ 1 ps gives way to a stronger and sharper feature centered around 620 nm. The 620 nm band continues to rise during ~ 15 ps. On the time scale of our experiments (up to 3.5 ns), this feature remains constant in amplitude, though the center of the spectrum slightly shifts to the blue on the 50 ps time scale. To obtain quantitative information on the kinetics, the data were subject to a global fitting procedure. As a model function, a sum of exponentials (decaying absorption corresponds to positive amplitudes) convoluted with an instrumental response (Gaussian with a width of 200 fs fwhm) was used. A satisfactory agreement between data and fit was obtained using four exponentials and an offset to account for the triplet absorption persistent in the time range (up to ~ 3 ns) of the femtosecond experiment. The time constants derived are ~ 0.5 , 1.3, 12, and 62 ps, and the corresponding decay associated spectra are depicted in Figure 4c. The spectra associated with the time constants of 1.3 and 12 ps both have negative sign (indicative for the buildup of an absorption) throughout most of the spectral region covered and are similar though not identical in shape. The 62 ps spectrum is sigmoidal, representing a blue shift of the ~ 620 nm band with time. The spectrum of the offset peaks at 610 nm and has a width of 50 nm (fwhm). From comparison of this spectrum with the spectrum obtained by nanosecond spectroscopy,¹⁴ it can be safely assigned to the absorption of the relaxed triplet state of xanthone in ethanol. Transient absorption experiments have also been performed with the excitation tuned to 266 nm. The observed transient spectra and derived time constants are virtually identical to those in the experiments presented above (excitation wavelength of 340 nm). Thus, the different excitation wavelengths in the fluorescence and the absorption experiments do not seem to alter the photodynamics and the two may be directly compared. Very similar results were also obtained for xanthonic acid in ethanol (data not shown). Xanthonic acid in water shows similar kinetics in the time domain $\tau < 200$ ps. However, the blue shift of the triplet absorption band occurs here with a 300 ps time constant (data not shown).

3.3. Triplet Quenching by 1-Methyl-naphthalene (1-MN).

Naphthalene and derivatives are known to quench the triplet states of aromatic ketones such as benzophenone and xanthone at least partially via intermolecular energy transfer.^{15,16} The question of whether electron transfer can compete with energy transfer will be addressed below. In the following, the dependence of the photokinetics of xanthone on the concentration of the naphthalene derivative 1-MN is investigated. These experiments not only allow us to draw conclusions on the quenching dynamics but also help to elucidate the ISC mechanism of xanthone. Addition of 1-MN to a solution of xanthone reduces the lifetime of the triplet spectrum centered around 610 nm (see Figure 5). For low quencher concentrations of ≤ 0.01 M, the lifetime reduction can be traced by nanosecond laser flash photolysis. Here, the rate constant k_{TT} of the decay of the xanthone triplet absorption scales linearly with the quencher concentrations (Figure 7). Higher quencher concentrations require femtosecond techniques to time resolve the energy transfer dynamics. This is illustrated by a measurement performed on xanthone in the presence of 0.6 M (10 vol %) 1-MN (Figure 5). The 610 nm absorption with an “infinite” lifetime in pure ethanol now decays in ~ 100 ps. Synchronized with this decay, a spectrum centered around 420 nm rises. This spectrum can be safely assigned to the triplet state of 1-MN;^{17,18} that is,

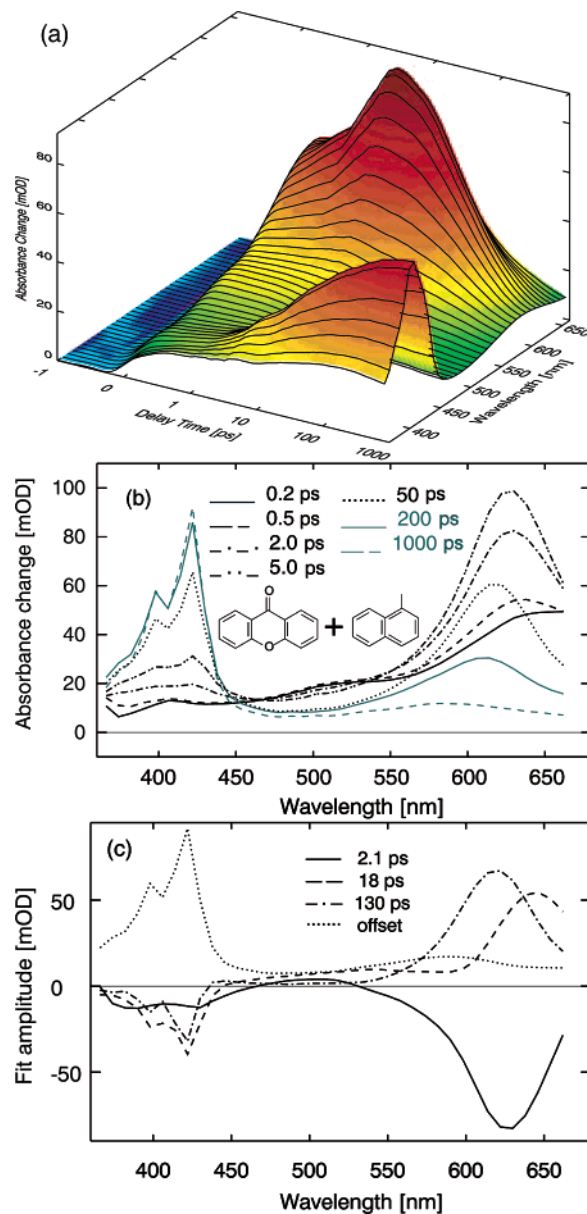


Figure 5. Transient absorption spectra of xanthone in the presence of 0.6 M 1-MN following 340 nm excitation. (a) Overview. (b) Transient spectra at distinct delay times. (c) Decay associated spectra.

as expected for an energy transfer process the decay of the xanthone triplet goes along with the population of the 1-MN triplet.

The temporal behavior of this transfer for various quencher concentrations can be nicely represented by two spectral integrals (Figure 6). The first integral (550–650 nm) covers the triplet absorption of xanthone (lower part of Figure 6), and the second integral (414–438 nm) that of the 1-MN. In the absence of a quencher, the biphasic rise of the xanthone triplet absorption is clearly represented by the “610 nm” integral. The “420 nm” integral is virtually “silent”. In the presence of 0.06 M 1-MN, the kinetics of the rise is not altered but now the 610 nm integral decays with a time constant of ~ 1.5 ns. In parallel, the 420 nm integral rises with the same time constant. Increasing the 1-MN concentration by a factor of 10 has the expected effect of accelerating the decay of the 610 nm integral and the rise of the 410 nm integral by roughly the same factor. A closer inspection of that rise shows that it is multiphasic and its description with exponentials requires at least three terms with

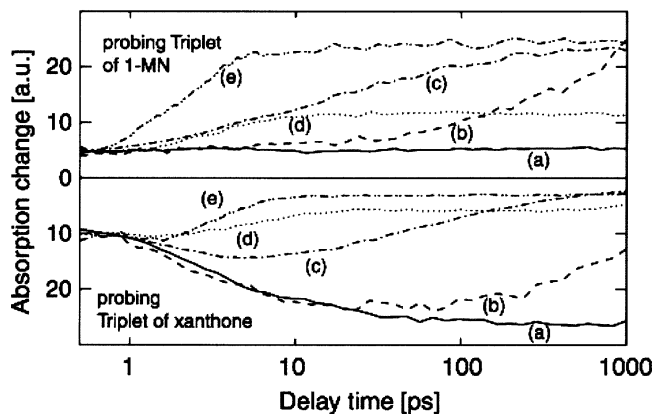


Figure 6. Time dependences of the spectral integrals covering the absorption of the 1-MN triplet (upper part) and that of the xanthone triplet absorption (lower part) for different 1-MN concentrations. For a better representation, the time traces in the lower part have been turned upside down. The 1-MN concentrations were (a) 0 M, (b) 0.06 M, (c) 0.6 M, (d) 3.0 M, and (e) neat 1-MN.

time constants of 2, 15, and 120 ps. The amplitudes of these terms are comparable. Note that the 2 ps time constant is shorter than the slow component observed for the rise of the xanthone triplet in the absence of a quencher. For the highest quencher concentration, that is, when performing the experiment in neat 1-MN, one obtains very fast and very simple quenching kinetics. The 420 nm integral rises in 2 ps, and the integral representing the xanthone triplet absorption decays in 2 ps to a small offset. Inspection of the spectrally resolved data (not shown) clearly demonstrates that this offset stems from a tail of the 1-MN triplet absorption and not from a residual xanthone triplet population.

Surprisingly, for a concentration of 3.0 M kinetics is found that deviates from the overall trend. In addition to the ultrafast kinetics which is comparable to the behavior in neat 1-MN, a longer lived component with a time constant of 1 ns appears that is absent even in the 0.6 M experiment. A possible origin for this anomaly will be addressed in the discussion.

4. Discussion

4.1. Intersystem Crossing in Xanthone. *Ultrafast Spectroscopy of Xanthone.* The time-resolved fluorescence data presented here allow us to exactly pin down the decay time of the primarily excited $^1\pi\pi^*$ state of xanthone. In previous time-resolved absorption experiments,⁸ this time could only be inferred indirectly from the rise of the triplet spectrum. Among the four excited states ($^1\pi\pi^*$, $^3\pi\pi^*$, $^1n\pi^*$, and $^3n\pi^*$) within energetic reach of the photoexcitation, only the $^1\pi\pi^*$ state shares an appreciable transition moment with the ground state. Transitions from the singlet ground state to the two triplet state are spin forbidden, and the transition to the $^1n\pi^*$ state is forbidden by symmetry. The extinction coefficient for that transition is $6 \text{ M}^{-1} \text{ cm}^{-1}$ as compared to $12\,000 \text{ M}^{-1} \text{ cm}^{-1}$ for the allowed transition to the $^1\pi\pi^*$ state.¹⁹ Thus, the only emissive state is the $^1\pi\pi^*$ state and the decay of the fluorescence with a time constant of $\tau \sim 1.5 \text{ ps}$ ³² can be safely assigned to the depletion of this state. If the 1.5 ps time constant characterizes the depletion of the $^1\pi\pi^*$ state, the additional 0.5 ps component present in the transient absorption experiment has to be associated with relaxation processes within the $^1\pi\pi^*$ state and not with its decay.

From the three states that might be populated in the course of the $^1\pi\pi^*$ decay, the $^3\pi\pi^*$ state can be excluded since such a transition would violate El-Sayed's rules. Thus, only the $^1\pi\pi^* \rightarrow ^1n\pi^*$ IC process (rate constant $k_{IC,1}$) and the $^1\pi\pi^* \rightarrow ^3n\pi^*$ ISC process (rate constant $k_{ISC,1}$) have to be taken into

consideration. The sum of the rate constants $k_{IC,1} + k_{ISC,1}$ is equal to the observed rate constant for the decay of the emissive state ($k_{IC,1} + k_{ISC,1} = 1/\tau_1 = (1.5 \text{ ps})^{-1}$). Partitioning that decay into the two channels requires additional information on the involved states. The global analysis of the transient absorption spectra associates the rise of a strong absorption band around 610 nm with the $\tau_1 = 1.5 \text{ ps}$ kinetic component. The shape and position of that spectrum are very close to those of the triplet spectrum of xanthone reported in the literature.⁴ Thus, it seems likely that the 1.5 ps process populates at least partially a triplet state. On the other hand, the singlet $^1n\pi^*$ state could have a very similar spectral signature to the $^3n\pi^*$ state and therefore a contribution or even a domination of the $^1\pi\pi^* \rightarrow ^1n\pi^*$ IC process cannot be excluded on the basis of these results.

The next process (time constant of 12 ps) in the biphasic rise of the triplet absorption is characterized by a spectrum which is slightly blue shifted and has a somewhat smaller width than the 1.5 ps spectrum. This spectrum represents the changes associated with the population of the finally reached triplet state. The sigmoidal feature with a characteristic time of 62 ps observed thereafter describes a blue shift of the triplet absorption. A similar effect has been found for the thio-analogue of xanthone (thioxanthone).²⁰ By examination of this shift for thioxanthone in several solvents, it was demonstrated that the characteristic times of this shift correlate with the dielectric responses of the solvents.²⁰ For ethanol, a value of 72 ps was derived which nicely matches the value reported here. Thus, the 62 ps time constant is not associated with a buildup of the triplet absorption and the 12 ps kinetic component thus describes the slowest part of this rise. Special care has to be taken for the assignment of this $\sim 12 \text{ ps}$ kinetic component. The assignment of the observed absorption changes depends on the reaction model. They may result either from an increase of the population in the triplet manifold, that is, from the $^1n\pi^* \rightarrow ^3\pi\pi^*$ transition (branched kinetics, Figure 2a), or from a $^3n\pi^* \rightarrow ^3\pi\pi^*$ transition (sequential kinetics, Figure 2b). In the latter case, the absorption spectra of the two triplet states have to be similar in shape but the extinction coefficient of the $^3\pi\pi^*$ state has to be approximately twice that of the $^3n\pi^*$ state. Since no independent access to these extinction coefficients is available, a final decision between branched and sequential kinetics on the spectroscopic results alone is still not possible. Here, the quenching experiments come into play.

Quenching Kinetics: Dynamic and Static Limit. For bimolecular quenching experiments, two limiting cases are known: the dynamic and static regimes. For low concentrations of the quencher (dynamic regime), the kinetics is pseudo-first-order (even the lowest quencher concentration applied in this paper is much higher than the xanthone concentration) with the rate constant k_{TT} given⁵ by

$$k_{TT} = k_q[1\text{-MN}] \quad (2)$$

The plot of the quenching rate constant versus the 1-MN concentration (Figure 7) confirms this simple law for concentrations lower than 0.06 M. Note that this plot includes results obtained by nanosecond and femtosecond experiments. The bimolecular rate constant k_q obtained by a linear fit of the data in Figure 7 is $7.2 \times 10^9 \text{ M}^{-1} \text{ s}^{-1}$ and is close to the value of $5.4 \times 10^9 \text{ M}^{-1} \text{ s}^{-1}$ predicted by a simple Smoluchowski–Stokes–Einstein treatment.²¹ This indicates that in this concentration regime the quenching is controlled by diffusion.

Simple single-exponential kinetics is also observed in neat 1-MN; here the formation of the 1-MN triplet can be described with one time constant of $\tau \sim 1.8 \text{ ps}$ (see the behavior of the

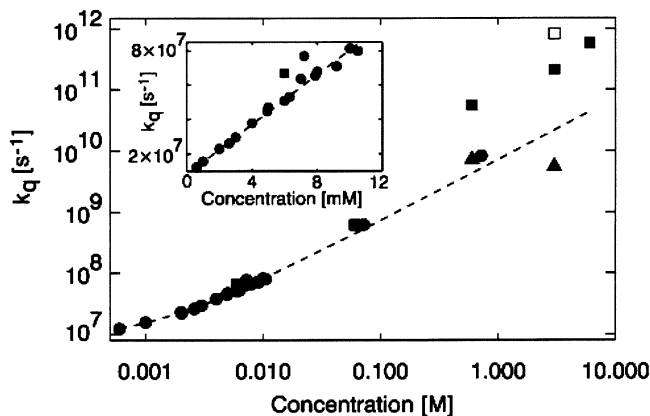


Figure 7. Dynamic quenching rate constant in dependence on the 1-MN concentration. The large (log–log) plot includes nanosecond (●) and femtosecond data (▲ for the slowest component, ■ and □ for the faster components). The insert focuses on the nanosecond experiments.

rise of the 1-MN triplet absorption in the upper part of Figure 6, curve e). In this static limit, the xanthone molecule is surrounded by quenchers at the instant of photoexcitation and no diffusional motion is required to induce the quenching process. The triplet quenching in this limit is faster than the slower component (12 ps) in the formation of the triplet absorption of xanthone in neat ethanol.

Quenching Kinetics: Intermediate Concentration Range and the 3.0 M Anomaly. Of course there is a smooth transition between the dynamic and the static regimes. A concentration of 0.6 M 1-MN nicely marks this interim domain where both regimes leave their kinetic imprint. The slow component of 120 ps obeys the prediction of eq 2, but faster components of 15 ps and even 2 ps for the rise of the 1-MN triplet absorption represent a static contribution to the quenching. Such complex kinetics is frequently observed for bimolecular photoprocesses provided that one of the reaction partners is present in high concentrations. Often these studies addressed the quenching of excited singlet states (for a review which includes theoretical description see ref 22), but an identical behavior should prevail in triplet quenching.

The 3.0 M experiment deviates from the smooth trend from a dynamic to a static behavior. On short time scales, the 2 ps component dominates the quenching kinetics. Thus, static quenching is the exclusive mechanism in that time domain. Quite surprisingly, in addition to this picosecond process the decay kinetics of the xanthone triplet contains a component with a time constant of ~ 3 ns. This time is longer than the 120 ps observed for the dynamic quenching in the 0.6 M experiment. At present, a final explanation of this anomaly cannot be given. It might have its origin in the mixing thermodynamics of 1-MN and ethanol in combination with preferential solvation. Aromatic hydrocarbons (like 1-MN) and alcohols are known to have positive mixing enthalpies ΔH_m ,²³ and this endothermicity usually peaks for a molar fraction of 0.5 of one of the constituents.²⁴ A concentration of 3.0 M 1-MN in ethanol corresponds to a molar fraction of 0.3. Typical peak values for ΔH_m are 1 kJ/mol,²⁴ which is close to the term of the entropy gain $T\Delta S_m$ which equals ~ 1.6 kJ/mol²⁵ for ideal solutions. Therefore, the free energy change of mixing ΔG_m will be close to zero and these solutions are at the edge of separation. It is to be expected that such a solution microscopically consists of clusters where 1-MN is enriched and others where ethanol is enriched. The xanthone molecules could then be partitioned between these clusters. The rather polar xanthone molecule might dominantly be solvated by ethanol. Examples for such a

preferential solvation are abundant in the literature (see e.g. ref 26 and references therein). With these findings, one could derive a model for the 3.0 M anomaly. A fraction of the xanthone molecules are situated in ethanol clusters. Their triplet states are quenched slowly as the 1-MN molecules have to diffuse into these clusters. The other fraction is surrounded by 1-MN molecules, and rapid static quenching is possible. This is of course only a hypothesis, and its validation requires further experiments. For the purpose of this paper, it can be set aside since on the time scale of interest here (picoseconds) the 3.0 M experiment is in line with those at other concentrations, particularly with that in neat 1-MN.

Quenching Mechanism: Energy Transfer versus Electron Transfer. In neat 1-MN, the formation of a 1-MN triplet with a time constant of 1.8 ps is faster than the slower component of the rise of the triplet absorption of xanthone in ethanol (12 ps). An interpretation of this surprising finding requires a discussion of the quenching mechanism. Two processes may contribute to the ultrafast quenching: triplet energy transfer and electron transfer generating a radical pair consisting of a 1-MN cation and a xanthone anion. Based on photocurrent measurements, Högemann and Vauthey excluded an electron-transfer quenching of triplet xanthone by 1-MN.¹⁶ For naphthalenes with a less positive oxidation potential, an electron transfer is found to occur. This behavior is explained by the energetics of the process. While the gain of free energy ΔG is essentially zero for the electron transfer between triplet xanthone and 1-MN, it is negative for naphthalenes with a less positive oxidation potential, and electron transfer is feasible. However, for an electron transfer starting from an excited singlet state of xanthone the free energy released during electron transfer is higher than for the triplet electron transfer. The energy of a radical pair is virtually independent of its spin state,²⁷ and the energy of the initial states is larger by the singlet triplet splitting of xanthone (~ 0.3 eV¹⁹). Therefore, starting from singlet excited xanthone electron transfer is exergonic by ~ 0.3 eV and fast electron transfer cannot be excluded. In the experiments presented in ref 16, due to the low (0.1 M) quencher concentrations such a singlet electron transfer could not occur since the intersystem crossing is finished before donor and acceptor meet. For the experiments in neat 1-MN presented here, one could imagine that electron transfer offers an additional decay channel for the excited singlet state.

Several observations allow quenching by electron transfer to be excluded. First, within the accuracy of transient absorption experiments the absorbance of triplet 1-MN after termination of the quenching is identical for different 1-MN concentrations. In other words, the yield of triplet 1-MN shows no dependence on the quencher concentration. Since electron transfer would constitute a competing quenching mechanism, one would expect a reduction of the triplet yield for high quencher concentrations. Second, in the experiments on xanthone in neat 1-MN no spectral signature of the xanthone anion nor of the 1-MN cation is detected. The respective signatures should be observed in a spectral range of 550–720 nm.¹⁶ After the ultrafast decay of the xanthone triplet spectrum within 1 ps, no longer lived signatures can be seen in this region. Thus the radical pair, if formed, has to be generated and has to recombine within less than 1 ps. Considering the highly different free energies which are released in both processes (-0.3 eV for the formation of the singlet radical pair and -3.18 eV¹⁶ for its recombination), simple calculations based on the Marcus theory exclude ultrafast kinetics for both processes. Therefore, for the following it seems justified to take only energy transfer mechanisms into account.

Quenching Mechanism: Branched versus Sequential Kinetics.

The absence of the 12 ps time constant in neat 1-MN can only be understood if all xanthone molecules undergo ISC with a time constant of 1.5 ps. Thus, the data are not compatible with a branched mechanism as described by Figure 2a. In that mechanism, the slow component would originate from the ISC between the $^1n\pi^*$ and the $^3\pi\pi^*$ state. A quenching of the $^1n\pi^*$ state by singlet–singlet energy transfer between excited xanthone and 1-MN is not feasible since the excited singlet state of 1-MN lies energetically well above that of xanthone.²¹ A singlet triplet energy transfer though energetically possible is spin-forbidden and will not occur between different molecules with such a short time constant. Therefore, the slow component of the triplet formation cannot be assigned to the depletion of the $^1n\pi^*$ and the branched mechanism must be excluded. In the sequential mechanism, the only deactivation channel of the primarily excited $^1\pi\pi^*$ state is the fast ISC process to the $^3n\pi^*$ state. That state then decays via slower (12 ps) IC to the $^3\pi\pi^*$ state. Already after the first transition the complete population resides in a state of triplet multiplicity. This state *can* be quenched via triplet–triplet energy transfer, and the disappearance of the longer lived component of the xanthone triplet rise in the quenching experiments is naturally explained.

What seems in conflict with this mechanism is the finding reported in ref 8 by Cavaleri et al. that a biphasic rise of the triplet absorption is found irrespective of the solvent polarity. With changing polarity, the energy gap between the $^3n\pi^*$ state and the $^3\pi\pi^*$ state is known to vary.¹⁹ Cavaleri et al. assume, in line with ref 19, that even an inversion occurs; that is, in polar solvents the $^3\pi\pi^*$ state is lowest in energy whereas it is the $^3n\pi^*$ state in nonpolar surroundings. Since El-Sayed's rules require that an ISC process starting in the $^1\pi\pi^*$ state will end up in a $^3n\pi^*$ state, in a nonpolar solvent the *lowest* triplet state would be directly populated. In our model which claims that the $^1\pi\pi^* \rightarrow ^1n\pi^*$ IC is slow, this would imply a single-exponential ISC behavior in nonpolar solvents, in contradiction to the findings of Cavaleri et al. However, the inversion of the triplet states has been questioned by Connors et al.^{28,29} They performed temperature-dependent phosphorescence experiments indicating that the energy gap between the triplet states is very small in nonpolar solvents but that the $^3\pi\pi^*$ state remains the state of lowest energy. The energy gap is smaller than kT at room temperature, and the $^3\pi\pi^*$ state can “behave” like a $^3n\pi^*$ state via thermal activation. Thus, the energetic ordering $^1\pi\pi^*$, $^3n\pi^*$, and $^3\pi\pi^*$ on which our sequential model depends seems to hold irrespective of the solvent polarity.

In the sequential mechanism proposed here, the ISC process from the $^1\pi\pi^*$ state to the $^3n\pi^*$ state is faster than the following IC process within the triplet manifold. This might be somewhat surprising. Yet, for El-Sayed allowed ISC processes the electronic couplings mediating the ISC transitions can be of the same order of magnitude as those of IC processes. For xanthone, this is supported by recent high-resolution spectra.³⁰ Still, one might argue that the IC process should be faster because of the energy gap law.⁵ The energy gap between the triplet states is smaller than the pertinent separation of the $^1\pi\pi^*$ state and the triplet states. As the energetic separations are rather small, a few 1000 cm^{-1} for the ISC and a few 100 cm^{-1} for the IC process, a breakdown of the energy gap law is at least not unlikely, since the energies released in the transitions approach the value of typical vibrational quanta. A thorough analysis would require a quantum chemical treatment not only of the Franck–Condon factors (replacing the energy gap term) but also of the respective coupling elements. Here, we just stress that

our experimental kinetic data point to an ISC process that is faster than IC.

4.2. Ultrafast Quenching Dynamics. The experiments on the triplet quenching at different 1-MN concentrations can also be used to obtain information on the quenching process itself. At low 1-MN concentrations, see Figures 5 and 6, exponential kinetics with a linear concentration dependence of the quenching rates is obtained. This is expected from the Smoluchowski–Stokes–Einstein treatment (see eq 2). At higher quencher concentrations, deviations from the single-exponential behavior result. At the highest concentrations, a single-exponential rise of the 1-MN triplet absorption is found with the time constant of the xanthone triplet formation (see above). This single exponentiality and the low transient concentration of the precursor state (see lower part Figure 6, curve e), the xanthone triplet, indicate that the transition rate $k^3_{X \rightarrow ^3(1-MN)} = 1/\tau^3_{X \rightarrow ^3(1-MN)}$ from the triplet of xanthone to 1-MN is considerably faster ($\tau < 1$ ps) than the formation time (1.5 ps) of the xanthone triplet itself. From the experiments with neat 1-MN as a solvent, we cannot decide whether the fast quenching rate is due to single, well-placed quenching molecules or due to the large number of 1-MN molecules in contact with the triplet donor.

A distinction between the two scenarios should be possible based on the quenching experiments at lower 1-MN concentrations. Lowering this concentration decreases the average number of acceptor molecules surrounding the xanthone donor. If the rate constant of the static quenching were dependent on that number, one would expect a decrease of the rate constant of the fastest component. Let us consider the experiments with a 1-MN concentration of 0.6 M. The rise of the 1-MN triplet absorption, which is directly related to the energy transfer, can be described by three time constants: 2 ps (relative contribution of 30%), 15 ps (40%), and 120 ps (30%). The process associated with the longest time constant obeys eq 2 and is therefore assigned to dynamic quenching. The 15 ps component is assigned to the transition regime between static and dynamic quenching. The 2 ps contribution finally is due to purely static quenching. This time constant is identical to the one observed for neat 1-MN, that is, with a quencher concentration 10 times higher. In other words, only the amplitude of the static quenching but not its rate constant decreases with concentration. Thus, the rate constant for the static quenching does not scale with the number of acceptors in close contact. It seems to be sufficient to have a single 1-MN in contact with triplet excited xanthone for energy transfer within ≤ 1 ps to occur. Such a behavior is not covered by the usual treatments of bimolecular quenching such as the Collins–Kimball approach³¹ and the more sophisticated models mentioned in ref 22.

5. Conclusion

The investigation of transient emission and absorption in a wide range of wavelengths on xanthone and xanthone/1-methylnaphthalene solutions yielded important new information on the ISC in xanthone and on triplet quenching. It was shown that triplet formation after optical excitation of xanthone is an ultrafast process. The reaction is sequential, proceeding from the $^1\pi\pi^*$ to the $^3n\pi^*$ state within 1.5 ps and followed by a slower (12 ps) internal conversion in the triplet manifold to the $^3\pi\pi^*$ state. 1-Methyl-naphthalene acts as an efficient and ultrafast quencher for the xanthone triplet state. Quenching occurs on the 1 ps time scale. The experiments show that the system xanthone/1-methyl-naphthalene can be used for the study of very fast structural dynamics of peptides even on the picosecond time scale.

Acknowledgment. Dedicated to Professor Ulrich E. Steiner on the occasion of his 60th birthday.

References and Notes

- (1) Bieri, O.; Wirz, J.; Hellrung, B.; Schutkowski, M.; Drewello M.; Kiefhaber, T. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 9597.
- (2) Krieger, F.; Fierz, B.; Bieri, O.; Drewello, M.; Kiefhaber, T. *J. Mol. Biol.* **2003**, *332*, 265.
- (3) Krieger, F.; Fierz, B.; Axthelm, F.; Joder, K.; Meyer, D.; Kiefhaber, T. *Chem. Phys.*, in press.
- (4) Scaiano, J. C. *J. Am. Chem. Soc.* **1980**, *102*, 7747.
- (5) *Modern Molecular Photochemistry*; Turro, N., Ed.; Benjamin/Cummings: Menlo Park, CA, 1978.
- (6) Damschen, D. E.; Merritt, C. D.; Perry, D. L.; Scott, G. W.; Talley, L. D. *J. Phys. Chem.* **1978**, *82*, 2268.
- (7) Greene, B. I.; Hochstrasser, R. M.; Weisman, R. B. *J. Chem. Phys.* **1979**, *70*, 1247.
- (8) Cavaleri, J.; Prater, K.; Bowman, R. *Chem. Phys. Lett.* **1996**, *259*, 495.
- (9) El-Sayed, M. *J. Chem. Phys.* **1963**, *38*, 2834.
- (10) Schmidt, B.; Laimgruber, S.; Zinth, W.; Gilch, P. *Appl. Phys. B* **2003**, *76*, 809.
- (11) Satzger, H.; Spörlein, S.; Root, C.; Wachtveitl, J.; Zinth, W.; Gilch, P. *Chem. Phys. Lett.* **2003**, *372*, 216.
- (12) Seel, M.; Wildermuth, E.; Zinth, W. *Meas. Sci. Technol.* **1997**, *8*, 449.
- (13) Strickler, S.; Berg, R. *J. Chem. Phys.* **1962**, *37*, 814.
- (14) Ley, C.; Morlet-Savary, F.; Fouassier, J. P.; Jacques, P. *J. Photochem. Photobiol., A* **2000**, *137*, 87.
- (15) Hochstrasser, R. A., Jr.; Lutz, H.; Scott, G. W. *J. Chem. Phys.* **1974**, *61*, 2500.
- (16) Högemann, C.; Vauthey, E. *J. Phys. Chem. A* **1998**, *102*, 10051.
- (17) Melhuish, W. *J. Chem. Phys.* **1969**, *50*, 2779.
- (18) Wang, X.; Kofron, W.; Kong, S.; Rajesh, C.; Modarelli, D.; Lim, E. *J. Phys. Chem. A* **2000**, *104*, 1461.
- (19) Pownall, H. J.; Huber, J. R. *J. Am. Chem. Soc.* **1971**, *93*, 6249.
- (20) Morlet-Savary, F.; Ley, C.; Jacques, P.; Wieder, F.; Fouassier, J. P. *J. Photochem. Photobiol., A* **1999**, *126*, 7.
- (21) *Handbook of Photochemistry*; Murov, S. L., Carmichael, I., Hug, G. L., Eds.; Marcel Dekker: New York, 1993.
- (22) Sikorski, M.; Krystkowiak, E.; Steer, R. *J. Photochem. Photobiol., A* **1998**, *117*, 1.
- (23) Ott, J.; Sipowska, J. *J. Chem. Eng. Data* **1996**, *41*, 987.
- (24) Lien, P.; Lin, H.; Lee, M. *Fluid Phase Equilib.* **2004**, *215*, 187.
- (25) Atkins, P.; de Paula, J. *Physical Chemistry*, 7th ed.; Oxford University Press: Oxford, 2002.
- (26) Jozefowicz, M.; Heldt, J. *Chem. Phys.* **2003**, *294*, 105.
- (27) Steiner, U. E.; Ulrich, T. *Chem. Rev.* **1989**, *89*, 51.
- (28) Connors, R. E.; Christian, W. R. *J. Phys. Chem.* **1982**, *86*, 1524.
- (29) Connors, R. E.; Sweeney, R. J.; Cerio, F. *J. Phys. Chem.* **1987**, *91*, 819.
- (30) Ohshima, Y.; Fujii, T.; Fujita, T.; Inaba, D.; Baba, M. *J. Phys. Chem. A* **2003**, *107*, 8851.
- (31) Collins, F. C.; Kimball, G. E. *J. Colloid. Sci.* **1949**, *4*, 425.
- (32) The time constant determined by transient absorption is 1.3 ps, and by fluorescence spectroscopy 1.5 ps. In the following we will refer to both as the 1.5 ps component.

Molecular Basis for the Effect of Urea and Guanidinium Chloride on the Dynamics of Unfolded Polypeptide Chains

Andreas Möglich, Florian Krieger and Thomas Kiefhaber*

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

Chemical denaturants are frequently used to unfold proteins and to characterize mechanisms and transition states of protein folding reactions. The molecular basis of the effect of urea and guanidinium chloride (GdmCl) on polypeptide chains is still not well understood. Models for denaturant–protein interaction include both direct binding and indirect changes in solvent properties. Here we report studies on the effect of urea and GdmCl on the rate constants (k_c) of end-to-end diffusion in unstructured poly(glycine-serine) chains of different length. Urea and GdmCl both lead to a linear decrease of $\ln k_c$ with denaturant concentration, as observed for the rate constants for protein folding. This suggests that the effect of denaturants on chain dynamics significantly contributes to the denaturant-dependence of folding rate constants for small proteins. We show that this linear dependency is the result of two additive non-linear effects, namely increased solvent viscosity and denaturant binding. The contribution from denaturant binding can be quantitatively described by Schellman's weak binding model with binding constants (K) of $0.62(\pm 0.01) \text{ M}^{-1}$ for GdmCl and $0.26(\pm 0.01) \text{ M}^{-1}$ for urea. In our model peptides the number of binding sites and the effect of a bound denaturant molecule on chain dynamics is identical for urea and GdmCl. The results further identify the polypeptide backbone as the major denaturant binding site and give an upper limit of a few nanoseconds for residence times of denaturant molecules on the polypeptide chain.

© 2004 Elsevier Ltd. All rights reserved.

Keywords: protein folding; polypeptide chain dynamics; solvent denaturation; denaturant binding; Schellman model

*Corresponding author

Introduction

Chemical denaturants like guanidinium chloride (GdmCl) and urea have long been known to unfold proteins^{1,2} by stabilizing the unfolded state compared to the native state.^{3–8} Unfolding by GdmCl and urea has since then become the standard method to determine protein stability and to investigate the kinetics and mechanisms of protein folding reactions. Empirically, a linear correlation between the free energy for unfolding (ΔG^0) and denaturant concentration ($[D]$) was found in the region of the unfolding transition.^{9,10} The

correlation coefficient of this linear free energy relationship is commonly termed m_{eq} -value ($m_{\text{eq}} = \partial \Delta G^0 / \partial [D]$). It reflects the denaturing strength of the denaturant. Linear relationships were also found between denaturant concentration and the logarithm of the rate constants for protein folding (k_f) and unfolding (k_u) indicating that the activation free energies for folding (ΔG_f^{\ddagger}) and unfolding (ΔG_u^{\ddagger}) are also linearly dependent on denaturant concentration.¹¹ Accordingly, kinetic m -values were defined as:

$$m_f = \frac{\partial \Delta G_f^{\ddagger}}{\partial [D]} \quad \text{and} \quad m_u = \frac{\partial \Delta G_u^{\ddagger}}{\partial [D]}$$

Despite the wealth of experimental data on the effect of chemical denaturants on protein folding and stability, the physical basis of denaturant action and the origin of the linear free energy relationships are still under discussion. Myers *et al.*¹² found a

Abbreviations used: GdmCl, guanidinium chloride; Xan, xanthone; NAla, naphthylalanine; TTET, triplet-triplet energy transfer.

E-mail address of the corresponding author: t.kiefhaber@unibas.ch

correlation between the denaturant m_{eq} -values and the change in solvent accessible surface area (ΔASA) upon unfolding. Mainly two models have been proposed to account for the destabilizing effect of denaturants. The transfer model correlates the effect of denaturants on protein stability to free energies of transfer of the polypeptide chain from water to denaturant solutions.^{5,13-17} These data showed that GdmCl and urea increase the solubility of most parts of a protein compared to water explaining the stabilization of the more solvent exposed unfolded state relative to the native state. An alternative model ascribes chemical denaturation to weak binding of denaturant molecules to the polypeptide chain^{8,9,18-23} and explains denaturation by the larger number of binding sites in the unfolded state compared to the native state. Both models are able to explain chemical denaturation curves of proteins in the region of the unfolding transition. Santoro & Bolen²⁴ compared the linear extrapolation method with the binding model and the transfer model by employing different denaturants and extrapolating the stabilities measured in the transition region to zero denaturant. Only the linear extrapolation method yielded identical ΔG^0 -values for all applied denaturants. This suggests the applicability of the empirical linear extrapolation method although to date no physical explanation has been given to explain the linearity over the complete range of denaturant concentrations.

The observed effects of denaturants on protein stability and folding are the result of differences between effects on the native state and on the unfolded state for m_{eq} -values and between the transition state and the respective ground states for m_{f} and m_{u} . This causes a major difficulty in the interpretation of the experimental data, since the effect of denaturants on a single state cannot be obtained directly. We recently reported that GdmCl and urea affect the rate constants for end-to-end diffusion (k_{c}) in unstructured polypeptide chains in the same way as they influence the microscopic rate constants for folding and unfolding, i.e. a linear correlation between $\ln k_{\text{c}}$ and $[D]$ was observed.²⁵ Here we report studies on the effect of urea and GdmCl on intra-chain diffusion in poly(glycine-serine) chains of different length, which serve as models for unstructured polypeptide chains. This allows us to specifically characterize the effect of denaturants on the ensemble of unfolded polypeptide chains and to assess the contributions of chain dynamics to m_{f} -values in protein folding.

Results

Effect of denaturants on end-to-end diffusion in unstructured polypeptide chains

We have previously used triplet-triplet energy transfer between xanthone (Xan) and naphthylalanine (NAla) to determine absolute rate constants for intra-chain diffusion (k_{c}) and the effect of

denaturants on chain dynamics in various unstructured polypeptide chains. To understand the molecular basis of denaturant action we investigated the effect of urea and GdmCl on unstructured polypeptide chains of varying length in more detail. Figure 1(A) shows representative kinetic traces for triplet-triplet energy transfer in a Xan-(Gly-Ser)₈-NAla-Ser-Arg-Gly peptide at various GdmCl concentrations. Single exponential kinetics for contact formation were observed in all measurements. The presence of GdmCl leads to a significant decrease in k_{c} from $4.0(\pm 0.1)\times 10^7 \text{ s}^{-1}$ in water to $1.7(\pm 0.1)\times 10^7 \text{ s}^{-1}$ in 8 M GdmCl. A similar effect was observed for urea (data not shown). A plot of $\ln k_{\text{c}}$

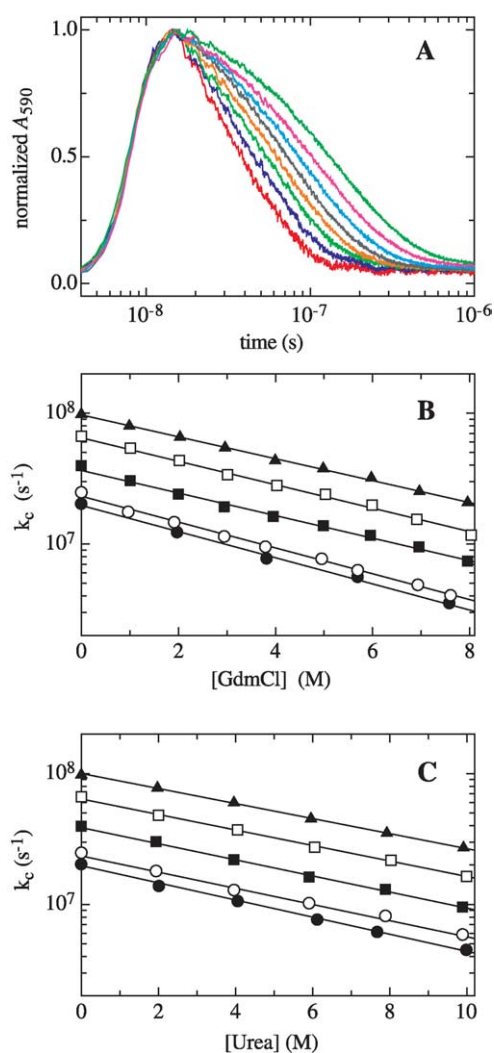


Figure 1. (A) Kinetics of end-to-end diffusion measured by triplet-triplet energy transfer in a Xan-(Gly-Ser)₈-NAla-Ser-Arg-Gly peptide as a function of GdmCl concentration (from left to right 1, 2, 3, 4, 5, 6, 7, and 8 M GdmCl). The reaction is followed by the decrease in xanthone triplet absorbance at 590 nm. (B) GdmCl dependence and (C) urea dependence of the rate constants for end-to-end contact formation (k_{c}) for various (Gly-Ser)_n peptides with $n=3$ (▲), 5 (□), 8 (■), 12 (○) and 14 (●) measured by triplet-triplet energy transfer as shown in (A). Solid lines represent fits of the data to equation (13).

versus $[D]$ reveals approximately linear relationships both for GdmCl and urea (Figure 1(B) and (C)). In analogy to protein folding, we determined m_c -values ($m_c = -RT\partial \ln k_c/\partial [D]$) for the effect of urea and GdmCl on chain dynamics (Figure 1(B) and (C)). For the (Gly-Ser)₈ peptide a fit to equation (13) gave m_c values of 0.49 (kJ/mol)/M for GdmCl and 0.35 (kJ/mol)/M for urea (Table 1). To investigate the effect of chain length on the m_c -values we measured the urea and GdmCl dependence of end-to-end diffusion in various Xan-(Gly-Ser)_{*n*}-NAla-Ser-Arg-Gly peptides with $n = 3, 5, 8, 12, 14$. For all peptides an approximately linear relationship between $\ln k_c$ and $[D]$ is observed (Figure 1). Both for urea and for GdmCl the m_c -values increase slightly with peptide length (Table 1).

Viscosity-corrected chain dynamics

Addition of denaturants significantly increases solvent viscosity,²⁶ which is known to affect end-to-end diffusion in unstructured peptides.²⁷ To determine the contributions from increased solvent viscosity to the m_c -values we measured chain dynamics in various glycerol/water mixtures (Figure 2). All peptides show a linear correlation between $\ln k_c$ and $\ln \eta$. The effect of η on k_c was analyzed using:

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

where η_0 is the reference solvent viscosity at zero denaturant concentration, k_c^0 is the rate constant at end-to-end diffusion at η_0 , k_c is the rate constant at a given viscosity η and the exponent β reflects the sensitivity of the reaction to solvent viscosity. A β -value of -1 indicates that k_c is inversely proportional to solvent viscosity and a β -value of 0 indicates that k_c is independent of solvent viscosity. Fitting the data to equation (1) gives β -values ranging from -0.90 for the shortest peptide to -0.99 for the longest peptide (Table 1). To test for contributions from specific interactions of the co-solutes with the peptides we also used ethylene glycol, glucose and sucrose as co-solutes to increase solvent viscosity. All co-solutes gave the same β -values within error, indicating that the

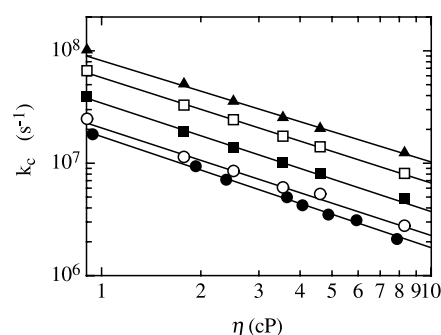


Figure 2. Effect of solvent viscosity (η) on k_c for the series of (Gly-Ser)_{*n*} peptides. Peptides and symbols are the same as shown in Figure 1(B) and (C). Measurements were performed in glycerol/water mixtures. Continuous lines show fits of the experimental data to equation (2) with the slopes (β) given in Table 1.

observed effect is due to an increase in solvent viscosity and not caused by specific co-solute-peptide interactions (data not shown and F.K. & T.K., unpublished results). To correct for the contributions of changes in η we used the β -values in combination with the viscosities of the respective denaturant solutions to calculate viscosity-corrected rate constants for end-to-end diffusion (k'_c) according to equation (2):

$$k'_c = k_c \left(\frac{\eta}{\eta_0} \right)^{-\beta} \quad (2)$$

This correction reduces the denaturant dependence of chain dynamics. However, the corrected k_c -values still depend on denaturant concentration (Figure 3(A) and (B)) but asymptotically approach a limiting value at high denaturant concentrations. This indicates that the effect of denaturants on end-to-end diffusion is partly due to a non-specific increase in solvent viscosity and partly due to a denaturant-specific effect, which becomes saturated at high denaturant concentrations. Correction for viscosity effects completely eliminates the denaturant-dependence of bimolecular triplet-triplet energy transfer (TTET) from xanthone to naphthylalanine,

Table 1. Denaturant dependence of end-to-end diffusion in Xan-(Gly-Ser)_{*n*}-Ser-Arg-Gly peptides

Peptide	β^a	Urea			GdmCl		
		m_c^b ((J/mol)/M)	K_{ex}^c	γ^d	m_c ((J/mol)/M)	K_{ex}	γ
(Gly-Ser) ₃	-0.90 ± 0.01	325.2 ± 2.5	13.1 ± 0.8	0.70 ± 0.01	474.1 ± 1.3	34.9 ± 1.1	0.65 ± 0.01
(Gly-Ser) ₅	-0.94 ± 0.01	335.3 ± 1.6	13.6 ± 0.6	0.68 ± 0.01	507.2 ± 1.9	23.8 ± 0.6	0.72 ± 0.01
(Gly-Ser) ₈	-0.97 ± 0.01	348.3 ± 2.1	13.3 ± 0.6	0.69 ± 0.01	488.1 ± 1.8	33.8 ± 0.9	0.66 ± 0.01
(Gly-Ser) ₁₂	-0.96 ± 0.01	349.8 ± 3.3	15.6 ± 1.0	0.68 ± 0.01	559.0 ± 3.0	26.0 ± 0.9	0.77 ± 0.01
(Gly-Ser) ₁₄	-0.99 ± 0.04	370.2 ± 3.4	14.5 ± 1.2	0.71 ± 0.02	565.7 ± 4.9	27.4 ± 1.5	0.73 ± 0.01
Global	–	–	13.8 ± 0.3	0.69 ± 0.01	–	32.6 ± 0.4	0.68 ± 0.01

^a Viscosity dependence of k_c ; see equation (3).

^b $m_c = RT \partial \ln k_c / \partial [D]$; see equation (13).

^c Equilibrium constant for exchange of a water molecule with a denaturant molecule; see equation (3).

^d Effect of a bound denaturant molecule on chain dynamics; see equation (4).

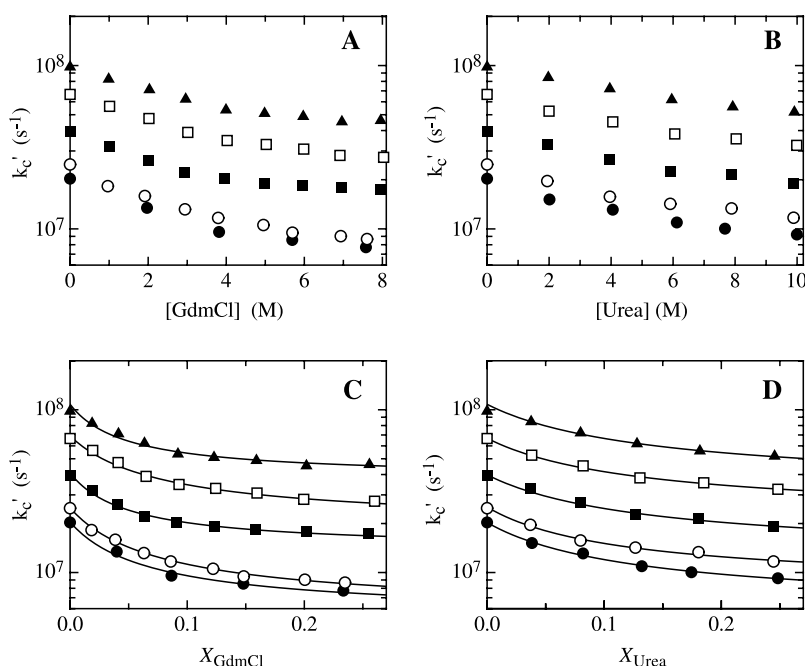


Figure 3. Effect of GdmCl (A) and urea (B) on the viscosity-corrected rate constants of end-to-end contact formation (k'_c) displayed on a molar scale. (C) and (D) The same data, respectively, on a mole fraction scale. In addition, the results of global fits of the data to the Schellman "site exchange" model are shown in (C) and (D). The results of the global fits and of all individual fits are given in Table 1. The (Gly-Ser)_n peptides and symbols are the same as for Figure 1.

indicating that denaturants do not influence the electron transfer reaction.

Schellman's weak binding model versus transfer model

We analyzed the denaturant-specific effect of urea and GdmCl on chain dynamics with the "site exchange" formalism, which was developed by Schellman to describe weak interactions of denaturants and other osmolytes with proteins.^{8,22} Assuming independent binding sites on a polypeptide chain, the average fractional binding site occupancy, \bar{v} , can be expressed as a function of the mole fraction of denaturant, X_D :

$$\bar{v} = \frac{(K_{\text{ex}} - 1)X_D}{(K_{\text{ex}} - 1)X_D + 1} \quad (3)$$

K_{ex} is the equilibrium constant for the exchange of a water molecule with a denaturant molecule at a single binding site. Further details and the derivation of equation (3) are discussed by Schellman.^{8,22} To analyze the effect of denaturants on k'_c in terms of the weak binding model we converted the denaturant concentrations to the mole fraction scale (Figure 3(C) and (D)). We made the assumption that the effect of denaturant binding on chain dynamics is proportional to the fractional binding site occupancy (\bar{v}), which means that all denaturant binding sites on the polypeptide chains are identical and independent of each other. These assumptions lead to equation (4):

$$k'_c = k_c^0 - \gamma k_c^0 \bar{v} \quad (4)$$

where k_c^0 is the rate constant for intra-chain diffusion in the absence of denaturant. The proportionality constant, γ , reflects the effect exerted by bound denaturant molecules on chain dynamics.

Fitting the data for urea and GdmCl to equations (3) and (4) shows that the denaturant dependence of k'_c is well-described by Schellman's weak binding model for all peptides (Figure 3(C) and (D); Table 1). Interestingly, both K_{ex} and γ do not vary systematically with peptide length (Table 1), which prompted us to fit the data for all peptides globally for each denaturant. This gave K_{ex} -values of 32.6 ± 0.4 and 13.8 ± 0.3 for binding of GdmCl and urea, respectively (Table 1), which corresponds to respective binding constants (K) of $0.62(\pm 0.01) \text{ M}^{-1}$ and $0.26(\pm 0.01) \text{ M}^{-1}$ on the molar scale. Remarkably, binding of GdmCl and urea slows down contact formation by the same amount as indicated by the virtually identical γ -values of 0.68 ± 0.01 for GdmCl and 0.69 ± 0.01 for urea (Table 1).

To also test whether the transfer model is able to describe the effect of denaturants on k'_c we compared the data to predictions from ΔG_{tr}^0 -values for different amino acids reported by Nozaki & Tanford.^{13,16,17} The correlation between $\ln k'_c$ and ΔG_{tr}^0 shows significant deviations from linearity especially for GdmCl (Figure 4). In addition, the slope of the correlation changes significantly with peptide length, both for urea and for GdmCl. This is not compatible with a general effect of transfer free energies on the chain properties. Similar results were obtained when only the ΔG_{tr}^0 -values for the polypeptide backbone were considered.

Discussion

Direct and indirect effects of denaturants on polypeptide dynamics

Our results show that the effect of denaturants on intra-chain diffusion in unstructured polypeptides

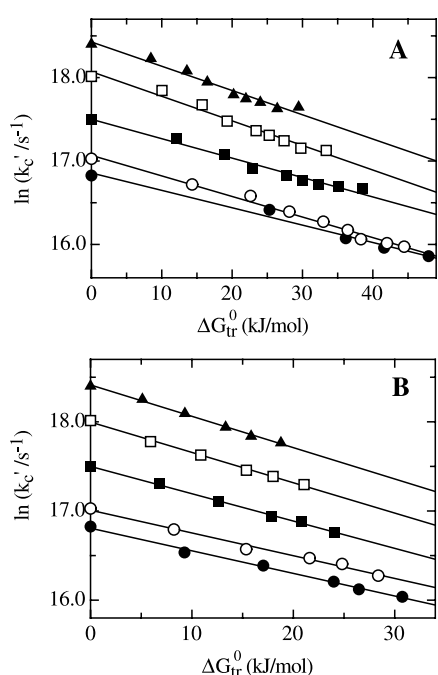


Figure 4. Correlation between k'_c and ΔG_{tr}^0 for GdmCl (A) and urea (B) calculated from the ΔG_{tr}^0 -values given by Nozaki & Tanford (see Materials and Methods). The (Gly-Ser) $_n$ peptides and symbols are the same as for Figure 1. The linear fits of the data give slopes ($-d \ln k'_c / d \ln K_{tr} = d \ln k'_c / d \ln \Delta G_{tr}^0 / RT$) of $-7.2(\pm 0.5) \times 10^{-2}$, $-7.2(\pm 0.4) \times 10^{-2}$, $-5.7(\pm 0.3) \times 10^{-2}$, $-6.1(\pm 0.2) \times 10^{-2}$ and $-5.1(\pm 0.4) \times 10^{-2}$ for $n=3,5,8,12$ and 14, respectively, for GdmCl and of $-8.6(\pm 0.3) \times 10^{-2}$, $-8.3(\pm 0.2) \times 10^{-2}$, $-7.5(\pm 0.2) \times 10^{-2}$, $-6.3(\pm 0.3) \times 10^{-2}$, $-6.2(\pm 0.2) \times 10^{-2}$, respectively, for urea.

can be dissected into two components. Denaturants indirectly slow down chain dynamics by increasing solvent viscosity,²⁷ and denaturants directly interact with polypeptide chains. The individual contributions from both components have non-linear effects on the denaturant dependence of $\ln k_c$. k_c is proportional to η^β (see equation (1)) with β -values between -0.90 and -0.99 for the different peptides (Figure 2). Solvent viscosity increases approximately hyperbolically with denaturant concentration,²⁶ which leads to a corresponding decrease in $\ln k_c$ with denaturant concentration. The effect of denaturant binding, in contrast, strongly changes $\ln k_c$ at low denaturant concentrations and approaches a limiting value at high concentrations (Figure 3). The sum of both effects results in an approximately linear denaturant dependence of $\ln k_c$, as seen in Figure 1. This model predicts that the linearity will break down at very high denaturant concentrations, which are, however, experimentally not accessible. A closer inspection of the additive effects of viscosity and binding reveals small deviations from linearity in the experimentally accessible concentration range. These small deviations are indeed seen both in the curve

calculated from the fitted parameters and in the experimental data (Figure 5) supporting the model that solvent viscosity and weak denaturant binding are the molecular origin of the observed effect of denaturants on end-to-end diffusion.

The effect of the direct denaturant-polypeptide interaction on chain dynamics can be quantitatively described by Schellman's "site exchange" model (Figure 3). The site exchange model gives K_{ex} -values of 32.6 ± 0.4 and 13.8 ± 0.3 for binding of GdmCl and urea, respectively (Table 1), which corresponds to respective binding constants (K) of $0.62(\pm 0.01) \text{ M}^{-1}$ and $0.26(\pm 0.01) \text{ M}^{-1}$ on the molar scale. Urea and GdmCl have virtually identical γ -values (0.69 ± 0.01 versus 0.68 ± 0.01 ; Table 1), which reflect the effect of bound denaturant molecules on chain dynamics (see equation (4)). This indicates that the polypeptide chains have the same number of binding sites for urea and GdmCl and that bound urea and GdmCl molecules decrease the rate constant for intra-chain contact formation by the same amount. The different efficiency of GdmCl and urea in slowing down chain dynamics is thus solely based on their different binding affinities to polypeptide chains. Full binding site occupancy decreases the rate of contact formation by about 3.3-fold for both denaturants. The binding constants and γ -values for urea and GdmCl are independent of the chain length, indicating that the slightly higher m_c -values for longer chains are due to their higher sensitivity to changes in solvent viscosity (Figure 2 and Table 1). The relative hydrophobicity should be significantly higher for the shorter peptides compared to the longer peptides, due to the contributions of the aromatic TTET labels. The insensitivity of the binding constant for chain length indicates only little contribution from side-chains and argues for the polypeptide backbone as the major denaturant binding site. This is consistent with the observation that the m_{eq} -values for protein unfolding are

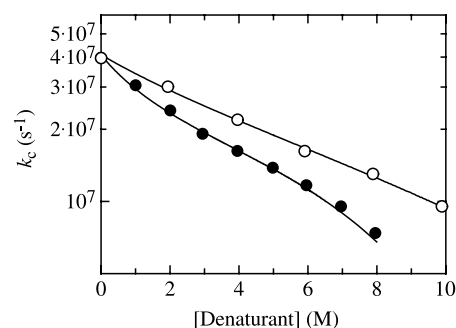


Figure 5. Comparison of the experimental data for the GdmCl (●) and urea (○) dependence of $\ln k_c$ in the (Gly-Ser) $_8$ peptide with the simulated curve assuming additive effects of solvent viscosity and denaturant binding on chain dynamics. The curve was simulated using equations (3) and (4) with the experimentally determined values for the viscosity dependence ($\beta = -0.97$) and denaturant-binding ($K_{ex}(\text{urea}) = 13.3$; $K_{ex}(\text{GdmCl}) = 33.8$).

approximately independent of the amino acid sequence¹² and with recent results from molecular dynamics simulations, which suggested the polypeptide backbone as the major site for interaction of denaturants with unfolded proteins.^{28,29}

Comparison to binding constants measured in other systems

The binding constants of $0.26(\pm 0.01) \text{ M}^{-1}$ for urea and $0.62(\pm 0.01) \text{ M}^{-1}$ for GdmCl measured for their effect on end-to-end diffusion represent specific binding constants of the denaturants to the unfolded polypeptide chain. These values are in good agreement with values determined for denaturant binding to model compounds. Robinson & Jencks¹⁵ reported solubility measurements and transfer studies of acetyl tetraglycyl ester from water to urea and GdmCl solutions and obtained binding constants of 0.29 M^{-1} for urea and 0.90 M^{-1} for GdmCl at 25°C by using a classical binding model. Tanford & Nozaki determined transfer free energies of various amino acids and peptide bond mimics from water to solutions of various denaturant concentrations.^{13,16,17} From these data Pace⁶ calculated binding constants in the range of 0.08 M^{-1} to 0.29 M^{-1} for urea and 0.51 M^{-1} to 1.03 M^{-1} for GdmCl. The values found for binding to peptide groups are 0.09 M^{-1} for urea and 0.60 M^{-1} for GdmCl. The binding constants determined from the effect of GdmCl on chain dynamics also agree well with calorimetrically determined values for GdmCl binding to unfolded proteins, which were 0.60 M^{-1} at 25°C for RNase A and lysozyme³⁰ and 1.16 M^{-1} for apocytochrome *c*.³¹ However, calorimetrically determined binding constants for interactions of urea with unfolded proteins are significantly lower ($K=0.06 \text{ M}^{-1}$ at 25°C) than those derived from the effect of urea on chain dynamics.³⁰

Comparison of our data to binding constants derived from unfolding transitions in proteins is not straightforward. Analysis of unfolding transitions usually assumes identical binding constants for denaturants on the unfolded and native state but different number of denaturant binding sites as the origin of the transitions. Still, the binding constants determined for the effect of urea and GdmCl on the helix-coil transition in short model peptides agree well with the values derived from chain dynamics. Applying the weak binding model to analyze the effect of urea on the helix-coil transition in short peptides gave a binding constant of 0.14 M^{-1} .³² Using the weak binding model to analyze the data reported by Scholtz and co-workers for the effect of GdmCl in the same peptides³³ we obtained values of 0.67 M^{-1} and 0.58 M^{-1} in the presence 10 mM and 100 mM NaCl, respectively.

A major question in the analysis of the effect of denaturants on proteins is whether there is direct evidence for binding or whether high denaturant concentrations non-specifically lead to the presence of denaturants near the polypeptide chain. Recent

magnetization transfer NMR studies on native and unfolded proteins have shown direct evidence for specific protein-denaturant interactions. For urea-binding to native BPTI a binding constant in the range of 0.1 M^{-1} to 0.36 M^{-1} was determined at 36°C .³⁴ Experiments on urea binding to unfolded intestinal fatty acid-binding protein revealed both strong and weak binding sites.³⁵ The experimental data for the weak binding sites are in agreement with binding constants between 0.05 M^{-1} and 0.2 M^{-1} , which are similar to the value of 0.26 M^{-1} found in our experiments. However, in the analysis of the magnetization transfer data the binding constant was fixed to the indicated values and not included in the data fitting. For GdmCl binding to unfolded and reduced α -lactalbumin an average binding constant of $0.16(\pm 0.07) \text{ M}^{-1}$ was reported,³⁶ which is significantly lower than the value of 0.62 M^{-1} found in our studies.

Residence time of denaturant molecules on the polypeptide chain

The observation of single exponential kinetics for contact formation in all peptides and at all denaturant concentrations suggests residence times of a few nanoseconds or less for denaturants bound to the unfolded chain. Exchange of denaturant molecules slower than 5 ns would lead to heterogeneous populations of unfolded molecules and result in complex kinetics. Fast exchange of bound denaturants is in agreement with results from NMR studies on native and unfolded proteins. Binding of urea to native BPTI and Pec-60 showed average residence times of a few nanoseconds at 4°C .³⁴ For urea and GdmCl-binding to weak binding sites in unfolded fatty acid binding protein average residence times of around 1 ns or shorter were reported.³⁵ Short residence times of urea on unfolded CI2 were also observed in molecular dynamics simulations.²⁹

Implications for protein folding *m*-values

Our results demonstrate that the rate constants for intra-chain diffusion in unfolded polypeptide chains significantly depend on denaturant concentration and exhibit the same linear relationship between $\ln k_c$ and $[D]$ as rate constants for protein folding reactions. This will contribute to the denaturant dependence of the rate constants for protein folding (m_f -values). In general, the rate constant (k) of a reaction contains contributions from the prefactor (k_0) and the activation free energy ($\Delta G^{0\dagger}$) according to:

$$k = k_0 e^{-\Delta G^{0\dagger}/RT} \quad (5)$$

It is very difficult to estimate absolute values for pre-exponential factors for the folding k_0^f and unfolding reaction k_0^u . They are most likely different for each protein, since the pre-exponential factors depend on the shape of the potential and on the

dynamics in the individual kinetic states (unfolded protein, native protein and folding intermediates). According to Kramers theory for barrier crossing reactions in the high friction limit, which should correspond to most reactions in solution and also to protein folding reactions, the rate of a reaction depends on the friction (γ) encountered by motions on a free energy landscape, by the shape of the barriers and by the frequency of motion (ω) on the free energy landscape according to equation (6):^{37,38}

$$k = \frac{\omega_0 \omega_B}{2\pi\gamma} e^{(-E_b/kT)} \quad (6)$$

ω_0 and ω_B denote the frequencies of motion in the starting well and on top of the barrier, respectively. This shows that the pre-exponential factor for protein folding depends on the shape of the potential and on the dynamics in the individual wells (states). ω_0 will be different for the refolding reaction starting from unfolded protein (ω_0^U) and for the unfolding reaction starting from native protein (ω_0^N) and hence the pre-factors will also be different for the forward and backward reaction (Figure 6). Eaton and co-workers measured an effective viscosity of 4.1 cP for internal motions in native myoglobin, in agreement with increased friction and/or decreased frequency of chain motions in native proteins compared to unfolded proteins.³⁹

In the light of these considerations our results suggest that the denaturant dependence of chain dynamics has major effects on the pre-factors for protein folding reactions. For a two-state folding reaction:



equation (8) is valid under all conditions:⁴⁰

$$K_{eq} = \frac{k_f}{k_u} \quad (8)$$

Thus, the equilibrium constant for folding (K_{eq})

contains contributions from the pre-exponential factors for folding (k_0^f) and unfolding (k_0^u) and from the free energies of activation, $\Delta G_f^{0\dagger}$ and $\Delta G_u^{0\dagger}$, according to:⁴¹

$$K_{eq} = \frac{k_0^f e^{-\Delta G_f^{0\dagger}/RT}}{k_0^u e^{-\Delta G_u^{0\dagger}/RT}} \quad (9)$$

The contribution of differences in the pre-factors for folding and unfolding to ΔG^0 ($= -RT \ln K_{eq}$) are consequently:

$$\begin{aligned} \Delta G^0 &= \Delta G_f^{0\dagger} - \Delta G_u^{0\dagger} + \Delta G_{pref}^0 \\ \Delta G_{pref}^0 &= RT \ln \frac{k_0^u}{k_0^f} \end{aligned} \quad (10)$$

Differences in k_0 for folding and unfolding can thus significantly contribute to apparent protein stabilities. A fivefold difference in k_0 between unfolded and native state, e.g. caused by differences in ω_0 (see Figure 6) would contribute 4 kJ/mol to the experimentally determined ΔG^0 -values at 25 °C.⁴¹ Consequently, changes in k_0^f caused by solvent-induced changes in the dynamics of the unfolded state, will influence both k_f and K_{eq} , even if the height of the barrier is not changed. Experimentally it can thus not be distinguished, whether a solvent-induced change in a rate constant arises from changes in the free energy barriers or from changes in the pre-factors, since equation (8) always holds. The effect of denaturants on the pre-factor for folding (m_f^0) will contribute to the observed kinetic and equilibrium m -values, $m_f(\text{app})$ and $m_{eq}(\text{app})$, according to equation (11):

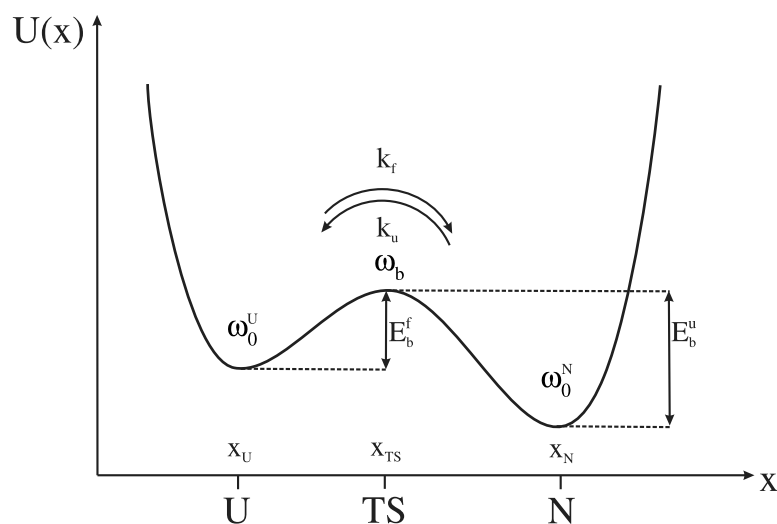


Figure 6. Schematic representation of a barrier crossing reaction on a double-well potential surface with two states N and U separated by an energy barrier. Escape from state U to N occurs via the rates k_f and k_u . E_b^f and E_b^u are the activation energies for the forward and back reaction. ω_0^U , ω_0^N and ω_b denote the frequencies of motion in the unfolded state, the native state and at the top of the barrier, respectively. The drawing was adopted from Hänggi *et al.*³⁸ and adapted for a schematic protein folding reaction.

$$\begin{aligned}\frac{\partial \Delta G_f^{0\dagger}(\text{app})}{\partial [D]} &= \frac{\partial \Delta G_f^{0\dagger}}{\partial [D]} - RT \frac{\partial \ln(k_f^0)}{\partial [D]} = m_f - m_f^0 = m_f(\text{app}) \\ \frac{\partial \Delta G_u^{0\dagger}(\text{app})}{\partial [D]} &= \frac{\partial \Delta G_u^{0\dagger}}{\partial [D]} - RT \frac{\partial \ln(k_u^0)}{\partial [D]} = m_u - m_u^0 = m_u(\text{app}) \\ \frac{\partial \Delta G^0(\text{app})}{\partial [D]} &= \frac{\partial \Delta G_u^{0\dagger}(\text{app})}{\partial [D]} - \frac{\partial \Delta G_f^{0\dagger}(\text{app})}{\partial [D]} = m_u - m_f + (m_f^0 - m_u^0) = m_{\text{eq}}(\text{app})\end{aligned}\quad (11)$$

This allows to estimate the contributions of the denaturant dependence of intra-chain diffusion (m_c) to experimentally determined m -values for small proteins. The m_c -values for peptide dynamics vary only little with chain length and show values around 0.35 (kJ/mol)/M for urea and 0.50 (kJ/mol)/M for GdmCl. Typical m_f -values for folding of the smallest fast folding proteins with chain length between 40 and 50 amino acid residues are around 1.0 (kJ/mol)/M for urea indicating that up to 30% of the measured m_f -values may arise from contributions of chain dynamics. Larger two-state folders consisting of 70–100 amino acid residues typically show m_f -values around 3 (kJ/mol)/M for urea and 5 (kJ/mol)/M for GdmCl,^{42,43} indicating that the denaturant dependence of chain dynamics constitutes up to 10% of the experimental m_f -values. It is difficult to judge the effect of denaturants on the dynamics of the native state, which will contribute to the experimental m_u -values. However, we have observed that the effect of denaturants on chain dynamics is largely based on increased end-to-end distances (A.M. & T.K., unpublished results). This would argue for only little effects of denaturant binding on the internal dynamics of the native state and the transition state structures, which were shown to be structurally well-defined and robust against changes in denaturant concentration.^{43–46} In addition, native proteins have less denaturant binding sites and the dynamics of native proteins, which contribute to k_0 for the unfolding reaction (see equation (6) and Figure 6), are strongly influenced by intra-molecular interactions. This reduces the frequency of chain motions as observed for myoglobin.³⁹ These motions should thus be less sensitive to the presence of denaturants. Consequently, the m_u -values are likely to have only little contributions from the effects of denaturants on chain dynamics ($m_u^0 \approx 0$) and m_{eq} -values should contain similar contributions from m_f^0 as the m_f -values. (see equation (11)).

Comparison of kinetic and equilibrium m -values is frequently used to characterize protein folding transition states according to the rate-equilibrium free energy relationship:^{11,44,45}

$$\alpha_D = \frac{m_f}{m_{\text{eq}}}\quad (12)$$

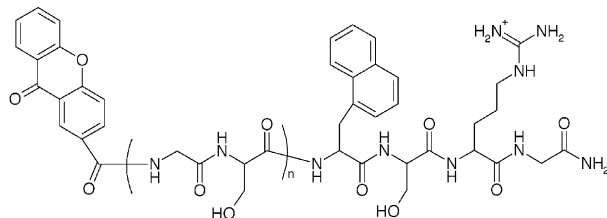
α_D is interpreted as a measure for the solvent accessibility of the transition state. According to our results, the contributions from chain dynamics to the m_f -values of small proteins lead to apparently

higher α_D -values and thus to apparently less solvent exposed transition states. This might explain the commonly observed higher α_D -values compared to α_C -values ($\alpha_C = \Delta C_p^{0\dagger}/\Delta C_p^0$), which are also believed to monitor the solvent accessibility of transition states.^{44,45} Temperature dependence of intra-chain diffusion in unstructured peptides showed that this process is not associated with a measurable change in heat capacity (F.K. & T.K., unpublished results). This suggests that α_C -values give a more reliable picture of the solvent accessibility of the transition state than α_D -values.

The ratio of binding constants determined for the binding of GdmCl and urea to unfolded chains is 2.4 (Table 1), which is virtually identical to the average ratio of 2.3 found for the m_{eq} -values of the two denaturants,¹² whereas the ratio of m_c -values for GdmCl and urea is significantly lower (1.5). This indicates that denaturant binding is the major source for the effect of denaturants on protein stability. We cannot compare these values to m_f -values for urea and GdmCl in protein folding reactions due to the small number of proteins that were characterized both with urea and GdmCl. However, the contribution of chain dynamics on m_f -values should depend on the location of the transition state along the reaction coordinate so a larger scattering in the ratio of urea and GdmCl m_f -value is expected.

Materials and Methods

Xanthonic acid and peptides were synthesized and purified as described.²⁵ The peptides are based on repetitive units of glycine-serine and have the canonical sequence Xan-(Gly-Ser)_n-NAla-Ser-Arg-Gly:



A peptide with n repeats of (Gly-Ser) between the two chromophores xanthone (Xan) and naphthylalanine (NAla) is referred to as (Gly-Ser)_n peptide. All measurements were performed in degassed solutions containing 10 mM potassium phosphate (pH 7) at 22.5 °C. Peptide concentrations were between 30 μM and 100 μM. Denaturant concentrations were determined by measuring the refractive index according to Pace.⁶

Triplet-triplet energy transfer (TTET) measurements

were performed as described by Krieger *et al.*²⁵ Xanthone triplet states were generated by a 4 ns laser pulse (Nd:YAG, 355 nm) and transient absorbance of the xanthone triplet state was monitored at 590 nm. The decay of the xanthone triplet absorbance is accompanied by an increase in the naphthyl triplet absorbance band at 420 nm indicating exchange of two electrons between xanthone and naphthalene upon contact formation between the chromophores.²⁵ Generally, four to eight kinetic traces at 590 nm were averaged and analyzed by non-linear least-squares fitting of the experimental data to exponential functions. Single exponential kinetics were observed for contact formation in all peptides. Errors were less than 5%. Data analysis was carried out using the computer programs ProFit (Quantum Soft, Zürich, Switzerland) and Matlab (The MathWorks, Natick, MA, USA).

The m_c -values for end-to-end diffusion were determined by fitting the data to equation (13):

$$\ln k_c = \ln k_c^0 - \frac{m_c[D]}{RT} \quad (13)$$

where k_c^0 is the rate constant for end-to-end diffusion at zero denaturant concentration and k_c is the respective value at the denaturant concentration $[D]$. Viscosities of denaturant solutions were calculated according to Perl *et al.*²⁶ These values agree well with viscosities measured in a falling ball viscosimeter (Haake, Germany). Viscosities of water/glycerol solutions were determined with a falling ball viscometer. Transfer free energies (ΔG_{tr}^0) from water to solutions containing various concentrations of urea or GdmCl were calculated from the values given by Nozaki & Tanford.^{13,16,17} The values for Trp were used to calculate ΔG_{tr}^0 for naphthylalanine and 1.5 times the value for Trp was used to calculate the ΔG_{tr}^0 value for xanthone. The values for Ser were approximated by the values of Thr minus the difference in ΔG_{tr}^0 between Leu and Val.

Acknowledgements

We thank Annett Bachmann, Beat Fierz and Manuela Schätzle for discussion and comments on the manuscript and Marty Scholtz for sending us data from Smith & Scholtz.³³ Mass spectrometry was conducted by Thierry Mini and Paul Jenö. Xanthonic acid was synthesized by Joseph Wey.

References

1. Spiro, K. (1900). Über die Beeinflussung der Eiweiss-coagulation durch stickstoffhaltige Substanzen. *Z. Physiol. Chem.* **30**, 182–199.
2. Ramsden, W. (1902). Some new properties of urea. *J. Physiol.* **28**, 23–27.
3. Simpson, R. B. & Kauzman, W. (1953). The kinetics of protein denaturation. I. The behavior of the optical rotation of ovalbumin in urea solutions. *J. Am. Chem. Soc.* **75**, 5139–5152.
4. Frensdorff, K. H., Watson, M. T. & Kauzman, W. (1953). The kinetics of protein denaturation. V. The viscosity of urea solutions of serum albumin. *J. Am. Chem. Soc.* **75**, 5167–5172.
5. Tanford, C. (1964). Isothermal unfolding of globular proteins in aqueous urea solutions. *J. Am. Chem. Soc.* **86**, 2050–2059.
6. Pace, C. N. (1986). Determination and analysis of urea and guanidine hydrochloride denaturation curves. *Methods Enzymol.* **131**, 266–280.
7. Makhatadze, G. I. (1999). Thermodynamics of protein interactions with urea and guanidinium hydrochloride. *J. Phys. Chem.* **103**, 4781–4785.
8. Schellman, J. A. (2002). Fifty years of solvent denaturation. *Biophys. Chem.* **96**, 91–101.
9. Aune, K. C. & Tanford, C. (1969). Thermodynamics of the denaturation of lysozyme by guanidine hydrochloride. II Dependence on denaturant concentration at 25 degrees. *Biochemistry*, **11**, 4586–4590.
10. Greene, R. F. J. & Pace, C. N. (1974). Urea and guanidine-hydrochloride denaturation of ribonuclease, lysozyme, alpha-chymotrypsin and beta-lactoglobulin. *J. Biol. Chem.* **249**, 5388–5393.
11. Tanford, C. (1970). Protein denaturation. Part C. Theoretical models for the mechanism of denaturation. *Adv. Protein Chem.* **24**, 1–95.
12. Myers, J. K., Pace, C. N. & Scholtz, J. M. (1995). Denaturant m values and heat capacity changes: relation to changes in accessible surface areas of protein unfolding. *Protein Sci.* **4**, 2138–2148.
13. Nozaki, Y. & Tanford, C. (1963). The solubility of amino acids and related compounds in aqueous urea solutions. *J. Biol. Chem.* **238**, 4074–4081.
14. Wetlaufer, D. B., Malik, S. K., Stoller, L. & Coffin, R. L. (1964). Non-polar group participation in the denaturation of proteins by urea and guanidinium salts. Model compound studies. *J. Am. Chem. Soc.* **86**, 508–514.
15. Robinson, D. R. & Jencks, W. P. (1965). The effect of compounds of the urea-guanidinium class on the activity coefficient of acetyltetraglycine ethyl ester and related compounds. *J. Am. Chem. Soc.* **87**, 2462–2470.
16. Nozaki, Y. & Tanford, C. (1970). The solubility of amino acids, diglycine, and triglycine in aqueous guanidine hydrochloride solutions. *J. Biol. Chem.* **245**, 1648–1652.
17. Nozaki, Y. & Tanford, C. (1971). The solubility of amino acids and two glycine peptides in aqueous ethanol and dioxane solutions. Establishment of a hydrophobicity scale. *J. Biol. Chem.* **246**, 2211–2217.
18. Schellman, J. A. (1955). The stability of hydrogen-bonded peptide structures in aqueous solutions. *Compt. rend. Carlsberg Lab. Ser. Chim.* **29**, 230–259.
19. Schellman, J. A. (1958). The factors affecting the stability of hydrogen-bonded polypeptide structures in solution. *J. Phys. Chem.* **62**, 1485–1494.
20. Schellman, J. A. (1975). Macromolecular binding. *Biopolymers*, **14**, 999–1018.
21. Schellman, J. A. (1978). Solvent denaturation. *Biopolymers*, **17**, 1305–1322.
22. Schellman, J. A. (1987). Selective binding and solvent denaturation. *Biopolymers*, **26**, 549–559.
23. Timasheff, S. N. (2002). Thermodynamic binding and site occupancy in the light of the Schellman exchange concept. *Biophys. Chem.* **101–102**, 99–111.
24. Santoro, M. M. & Bolen, D. W. (1988). Unfolding free energy changes determined by the linear extrapolation method. 1. Unfolding of phenylmethanesulfonyl alpha-chymotrypsin using different denaturants. *Biochemistry*, **27**, 8063–8068.
25. Krieger, F., Fierz, B., Bieri, O., Drewello, M. &

- Kiefhaber, T. (2003). Dynamics of unfolded polypeptide chains as model for the earliest steps in protein folding. *J. Mol. Biol.* **332**, 265–274.
26. Perl, D., Jacob, M., Bánó, M., Stupák, M., Antalík, M. & Schmid, F. X. (2002). Thermodynamics of a diffusional protein folding reaction. *Biophys. Chem.* **2–3**, 173–190.
27. Bieri, O., Wirz, J., Hellrung, B., Schutkowski, M., Drewello, M. & Kiefhaber, T. (1999). The speed limit for protein folding measured by triplet-triplet energy transfer. *Proc. Natl Acad. Sci. USA*, **96**, 9597–9601.
28. Tobi, D., Elber, R. & Thirumalai, D. (2003). The dominant interaction between peptide and urea is electrostatic in nature: a molecular dynamics simulation study. *Biopolymers*, **68**, 359–369.
29. Bennion, B. J. & Daggett, V. (2003). The molecular basis for the chemical denaturation of proteins by urea. *Proc. Natl Acad. Sci. USA*, **100**, 5142–5147.
30. Makhatadze, G. I. & Privalov, P. L. (1992). Protein interactions with urea and guanidinium chloride. A calorimetric study. *J. Mol. Biol.* **226**, 491–505.
31. Pfeil, W., Welfle, K. & Bychkova, V. E. (1991). Guanidine hydrochloride titration of the unfolded apocytocrome c studied by calorimetry. *Studia Biophysica*, **140**, 5–12.
32. Scholtz, J. M., Barrick, D., York, E. J., Stewart, J. M. & Baldwin, R. L. (1995). Urea unfolding of peptide helices as a model for interpreting protein unfolding. *Proc. Natl Acad. Sci. USA*, **92**, 185–189.
33. Smith, J. S. & Scholtz, J. M. (1996). Guanidine hydrochloride unfolding of peptide helices: separation of denaturant and salt effects. *Biochemistry*, **35**, 7292–7297.
34. Liepinsh, E. & Otting, G. (1994). Specificity of urea binding to proteins. *J. Am. Chem. Soc.* **116**, 9670–9674.
35. Modig, K., Kurian, E., Prendergast, F. G. & Halle, B. (2003). Water and urea interactions with the native and unfolded forms of a beta-barrel protein. *Protein Sci.* **12**, 2768–2781.
36. Denisov, V. P., Jonsson, B. H. & Halle, B. (1999). Hydration of denatured and molten globule proteins. *Nature Struct. Biol.* **6**, 253–260.
37. Kramers, H. A. (1940). Brownian motion in a field of force and the diffusion model of chemical reactions. *Physica*, **4**, 284–304.
38. Hänggi, P., Talkner, P. & Borkovec, M. (1990). Reaction-rate theory: fifty years after Kramers. *Rev. Mod. Phys.* **62**, 251–341.
39. Ansari, A., Jones, C. M., Henry, E., Hofrichter, J. & Eaton, W. A. (1992). The role of solvent viscosity in the dynamics of protein conformational changes. *Science*, **256**, 1796–1798.
40. van't Hoff, J. H. (1884). *Etudes de dynamique*. Muller, Amsterdam.
41. Bieri, O. & Kiefhaber, T. (2000). Kinetic models in protein folding. In *Protein Folding: Frontiers in Molecular Biology* (Pain, R., ed.) 2nd edit., pp. 34–64, Oxford University Press, Oxford.
42. Jackson, S. E. (1998). How do small single-domain proteins fold? *Fold. Des.* **3**, R81–R91.
43. Sánchez, I. E. & Kiefhaber, T. (2003). Evidence for sequential barriers and obligatory intermediates in apparent two-state protein folding. *J. Mol. Biol.* **325**, 367–376.
44. Sánchez, I. E. & Kiefhaber, T. (2003). Non-linear rate-equilibrium free energy relationships and Hammond behavior in protein folding. *Biophys. Chem.* **100**, 397–407.
45. Sánchez, I. E. & Kiefhaber, T. (2003). Hammond behavior versus ground state effects in protein folding: evidence for narrow free energy barriers and residual structure in unfolded states. *J. Mol. Biol.* **327**, 867–884.
46. Sánchez, I. E. & Kiefhaber, T. (2003). Origin of unusual phi-values in protein folding: evidence against specific nucleation sites. *J. Mol. Biol.* **334**, 1077–1085.

Edited by F. Schmid

(Received 7 July 2004; received in revised form 7 October 2004; accepted 13 October 2004)

7.5 Effect of Proline and Glycine Residue on the Dynamics and Barriers of Loop Formation in Polypeptide Chains (Krieger et al., submitted)

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

Florian Krieger, Andreas Möglich and Thomas Kiefhaber*

Biozentrum der Universität Basel, Department of Biophysical Chemistry,
Klingelbergstr. 70 CH-4056 Basel, Switzerland

Classification: Biophysics, Chemistry

*corresponding author. Phone: ++41-61-267 2194; fax: ++41-61-267 2189; e-mail:

t.kiefhaber@unibas.ch

ABSTRACT

Glycine and proline residues are frequently found in turn and loop sequences of proteins and are believed to play an important role during chain compaction early in protein folding. We investigated the effect of proline and glycine on the dynamics of loop formation in various unstructured polypeptide chains. Our results show that both amino acids significantly influence the kinetics of formation of short loops consisting of 3 to 10 residues, which is the typical size of turns and loops in native proteins. Formation of longer loops is not affected by the presence of glycine and proline in the central loop region. Formation of short loops is significantly slower around *trans* prolyl peptide bonds and faster around glycine residues compared to any other amino acid. However, formation of short loops occurs fastest around *cis* prolyl peptide bonds with a time constant of 6 ns for formation of a 5 residue loop. This effect can be ascribed to a higher Arrhenius pre-exponential factor, which leads to fast kinetics despite a higher Arrhenius activation energy for loop formation around *cis* isomers compared to *trans* isomers. Monte Carlo simulation of energetically favorable conformations in 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 in *cis* proline isomers and to a smaller loss in conformational entropy upon loop formation, which explains the unusually high pre-exponential factor.

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 non-covalent 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 are the properties of the ensemble of unfolded conformations as the starting point of the reaction. Several experimental techniques like nuclear magnetic resonance spectroscopy (1-4) and small-angle X-ray scattering techniques (5-7) have been applied to analyze structural properties of unfolded polypeptide chains. Less is known about the dynamics in unfolded polypeptide chain, which determine the formation of interactions during the early stages of folding. Especially the formation of loop and turn structures are important processes, since they allow the polypeptide chain to explore energetically favorable interactions and lead to compaction of the chain through formation of secondary structures like β -hairpins, turns and loops. Here we investigate the effect of glycine and proline residues on the loop closure dynamics in unfolded polypeptide chains. These residues frequently occur in turn and loop sequences (8, 9) and thus may play a major role early in folding during compaction of the polypeptide chain.

In recent years several experimental systems have been applied to obtain information on the time scale of intrachain contact formation in polypeptides (10-16). We have used triplet-triplet energy transfer (TTET) from xanthone (Xan) to naphthylalanine (NAIa) 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 two groups (18). Formation of the xanthone triplet state and electron transfer between xanthone and naphthalene occur on the low picosecond time scale (19) and the transfer process is diffusion-controlled (15, 17). TTET thus allows measurements of absolute rate constants for diffusional processes slower than the photochemical processes, i.e. slower than about 10 ps for the xanthone/naphthalene system (Figure 1A) (19). In earlier experiments the triplet donor and acceptor groups were introduced near the ends of various repetitive sequences or of fragments of natural proteins (12, 15, 17). This enabled us to determine of 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 acids 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 proline and glycine residues on the kinetics of loop formation. We show that the heterogeneity in the proline containing peptides can be assigned to dynamics of *cis* and *trans* Xaa-pro isomers. Loops of varying length are used to assess the range 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 simulations using statistical enumeration of conformational space.

MATERIALS AND METHODS

9-Oxoxanthen-2-carboxylic acid (xanthone acid) synthesis and peptide synthesis were carried out as described previously (15). The synthesis of pseudo-proline (ps-Pro)-containing peptides were performed using a Fmoc-Val-[$\Psi^{\text{Me,Me}}\text{Ser}$] building block (Novabiochem) on a acid sensitive preloaded trityl resin (TentaGel S, Rapp Polymer) and trityl protecting groups. Cleavage conditions were 1 % trifluoroacetic acid (Fluka) in dichloromethane (Fluka). All peptides were purified by preparative HPLC. Purity and mass of the peptides were checked by analytical HPLC and mass spectrometry, respectively. Laser flash 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 with less 10% of the total amplitude 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 ($\tau > 20 \mu\text{s}$) and may be due to oligomeric structures which do not allow TTET (15). NMR spectra were recorded on a Bruker ARX600 spectrometer with a proton resonance frequency at 600 MHz. The *cis* content of the proline containing peptides was determined according to the method described by Reimer et al. (20).

Viscosity dependence. The viscosity was varied by adding 0–70 % (v/v) glycerol (Aldrich) to the solutions. For all viscous solutions, the solvent viscosity η was measured by using a falling sphere viscometer (Haake) at 22.5°C. The viscosity dependencies of the end-to-end contact formation rate constants were analyzed using the empirical equation

$$k_c = k_c^0 \cdot \left(\frac{\eta_0}{\eta} \right)^\beta \quad [1]$$

where k_c is the observed contact formation rate constant, η_0 is the reference viscosity of 1.0 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. The same values as measured for glycerol were obtained for glucose, sucrose and ethylene glycol, indicating a genuine effect of solvent viscosity on the dynamics.

Temperature dependence. In all laser flash experiments the temperature was measured directly in the cuvette. For the determination of activation energies the observed rate constants were corrected to account for the effect of temperature on water viscosity. For the corrections the β -values determined in the viscosity-dependence (eq. [1]) were used and the data analyzed according to the Arrhenius equation:

$$k_c^0(T) = k_c(T) \cdot \left(\frac{\eta(T)}{\eta_0} \right)^\beta = A \cdot e^{-\frac{E_a}{RT}} \quad [2]$$

where A is Arrhenius pre-exponential factor at $\eta_0=1.0$ cP and E_a the Arrhenius activation energy. Tabulated values were used to determine the solvent viscosity of water at given temperatures $\eta(T)$ (21).

Simulations of chain conformations. All-atom simulations of oligopeptides were carried out using a hard-sphere model. Hard-sphere contact radii, bond lengths, and bond angles corresponded to the values recently used by Rose and co-workers (22).

Atom contact radii were taken from Hopfinger (23) and softened by multiplying with a factor of 0.95. Bond lengths and angles were obtained from Engh and Huber (24). 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° (25). 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 molecules were checked for steric overlaps using the above hard-sphere contact radii. Following Pappu et al. (22) 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 any 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. The Ramachandran maps (Fig. 5A) for the effect of *cis* and *trans* prolyl isomers on conformational energy were calculated for the central serine residue in Ser-Ser-Pro peptides (26) 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.

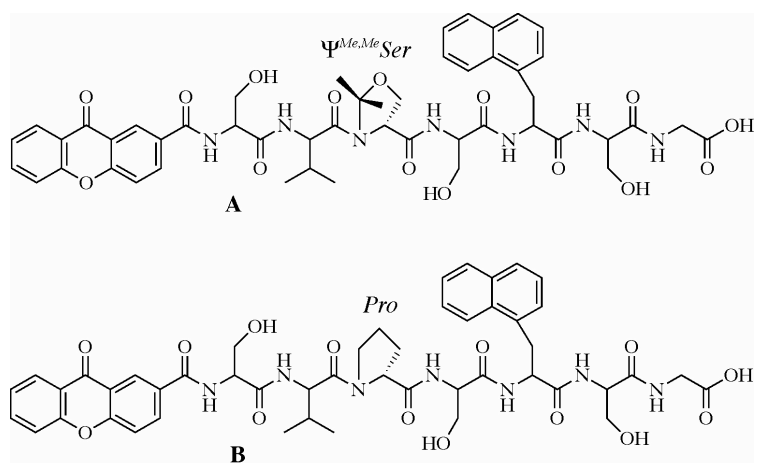
The simulation program was written in ANSI C and compiled using the Bloodshed Dev-C++ development environment (<http://www.bloodshed.net>). Compiled programs were executed on IBM compatible computers running Windows.

RESULTS AND DISCUSSION

Kinetics of loop formation in short proline and glycine containing peptides. To investigate the effect of proline and glycine residues on chain dynamics in more detail we performed measurements of the kinetics of loop formation in host-guest peptides of the structure $Xan-(Ser)_x-Xaa-(Ser)_y-NAla-Ser-Gly$ with $x,y = 1-7$ and $Xaa = Gly, Ser$ or Pro . Figure 1B compares TTET kinetics in the $Xan-Ser-Ser-Ser-NAla-Ser-Gly$ peptide (SSS peptide) 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 as they frequently occur at the end of β -hairpins. Intrachain loop formation in the SSS and SGS peptides can be described by single exponential kinetics with rate constants of $(1.6\pm 0.1)\cdot 10^7 s^{-1}$ and $(2.5\pm 0.1)\cdot 10^7 s^{-1}$, respectively, in the presence of 52% glycerol. Under then same conditions, double exponential kinetics are observed for loop formation in the SPS peptide with rate constants of $(2.7\pm 0.3)\cdot 10^7 s^{-1}$ and $(4.7\pm 0.2)\cdot 10^7 s^{-1}$ and respective amplitudes of $20 \pm 5\%$ and $80 \pm 5\%$. The faster kinetic phase in the SPS peptides is slightly faster than the dynamics 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 containing peptide 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%. Glycerol slows down chain dynamics due to increased solvent viscosity and thus allows a more reliable determination of the rate constant and amplitude of the fastest reaction in water (Figure 2A). The rate constants for loop formation in the SPS peptide in water are $(2.0\pm 0.1)\cdot 10^7 s^{-1}$ and $(1.2\pm 0.2)\cdot 10^8 s^{-1}$ (see Figure 2A and table 1). For the SSS and SGS peptides rate constants in water of $(6.7\pm 0.7)\cdot 10^7 s^{-1}$ and $(1.2\pm 0.1)\cdot 10^8 s^{-1}$, respectively were determined (table 1). Glycerol has no effect on the amplitudes of the two kinetic phases (Figure 2B). The average

amplitudes for the faster and slower process measured at the various glycerol concentrations are $17.0\pm 2.2\%$ and $83.0\pm 2.2\%$, respectively.

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 proline moiety. Studies on model peptides showed that the *cis* content depends on the preceding amino acid and varies between 7 and 38% (20) in contrast to about 0.15-0.5 % *cis* isomer at non-prolyl peptide bonds (27). The observed $17\pm 2\%$ 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 $16\pm 2\%$ as determined by 1D ^1H -NMR spectroscopy (data not shown). However, also rotation around the ψ -angle of the amino acid following the proline residue was shown to be slow with a rotation barrier around 50 kJ/mol. This might also give rise to the observed heterogeneity in the kinetics of loop formation in proline containing peptides. To discriminate between *cis-trans* isomers of Xaa-Pro bonds and slow rotation around ψ -angle as molecular origin for the complex dynamics we investigated the effect of the *cis* content on the kinetics of loop formation. Pseudo-proline ($[\Psi^{\text{Me,Me}}\text{Ser}]$) is known to increase the *cis* content of a Val-Pro peptide bond to 80–90% due to steric conflicts induced by the methyl groups in the *trans* conformation (28, 29). We introduced pseudo-proline (psPro) into two different peptides of the sequence Xan-Val- $[\Psi^{\text{Me,Me}}\text{Ser}]$ -Ser-NAla-Ser-Gly (V(psP)S-peptide) and Xan-Ser-Val- $[\Psi^{\text{Me,Me}}\text{Ser}]$ -Ser-NAla-Ser-Gly (SV(psP)S peptide, peptide **A**) and compared the resulting kinetics of loop formation to the kinetics of the corresponding proline containing peptides (peptide **B**).



Both pseudo-proline containing peptides show double exponential kinetics of contact formation with significantly increased amplitude (70-80%) of the faster reaction compared to the corresponding proline containing peptides (Figure 3A and Table 1). This shows that the faster reaction corresponds to the dynamics of peptides with a *cis* isomer. Since *cis-trans* isomerization is slow ($1/k \sim 20$ s at 22 °C) (20, 29, 30) there is no equilibration between the two isomers on the time scale of the TTET experiments and the dynamics of both isomers can be resolved.

Temperature-dependence of loop formation. Information on the origin of the faster dynamics in chains with *cis* proline 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 (eq. [2]) yielded significantly different apparent activation energies of 47.1 kJ/mol for the *cis* isomer and 38.6 kJ/mol for the *trans* isomer. Since the dynamics were shown to be sensitive to solvent viscosity (Fig. 2) the rate constants measured at the different temperatures had to be corrected for contributions from the effect of temperature on water viscosity to obtain the actual activation energies (see Materials and Methods). This resulted in viscosity corrected

activation energies (E_a) of 29.8 and 20.0 kJ/mol for loop formation in peptides with a *cis* and a *trans* Val-Pro peptide bond, respectively (Fig. 3B). The fit further revealed largely different pre-exponential factors (A) of $1.6 \cdot 10^{13} \text{ s}^{-1}$ and $2.9 \cdot 10^{10} \text{ s}^{-1}$ for loop formation in the *cis* and *trans* isomer, respectively. These results show that loop formation in the *cis* isomer encounters higher barriers compared to the *trans* isomer but is faster at room temperature due to a significantly higher pre-exponential factor. This indicates favorable entropic contributions to rate constant for loop closure in the *cis* isomer. A significantly higher activation energy of $42.9 \pm 3.4 \text{ kJ/mol}$ and a higher pre-exponential factor of $2.7 \pm 4 \cdot 10^{15} \text{ s}^{-1}$ is 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 peptides the pre-exponential factors for the loop closure dynamics in the *trans* 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* proline isomers mainly influence the height of the energy barrier for loop formation but not the entropic contributions. Linear Arrhenius plots were observed for all peptides, which show that loop formation is not associated with significant changes in heat capacity ($\Delta C_p \approx 0$). The results on the dynamics in the different proline containing peptides further reveal that the amino acid preceding the proline residue significantly influences the dynamics of both the *cis* and the *trans* isomer. In all short peptides Val-Pro bonds lead to about 2 fold slower loop formation kinetics compared to Ser-Pro bonds. This argues for steric effect of large side chains on the intrachain dynamics.

Effect of loop-size. Our 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 proline or glycine residue at a central position and varied the loop size. Figure 4 compares the kinetics of loop formation in glycine and proline containing peptides with the results from polyserine and poly(glycine-serine) loops. The effect of a proline and glycine residue decreases with increasing distance from the contact labels. A glycine residue has no effect on loop formation if the sites of contact formation are further than 5 amino acids from the glycine on both sides. 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 (Fig. 4). A more complex behavior is observed for proline residues. Only for $x, y \leq 2$ the dynamics of peptides with *cis* and *trans* isomers can be resolved, suggesting that the dynamics of the *cis* and *trans* isomers are similar in longer peptides (Fig. 4). Interestingly, in short peptides the *cis* isomer always shows faster loop formation than the *trans* isomer, irrespective of whether donor and acceptor are separated by an even or an uneven number of peptide bonds (Fig. 4 and Table 1). Formation of short loops becomes faster with increasing loop size in chains with *cis* isomers and the rate constant reaches a maximum of $k=1.5 \cdot 10^8 \text{ s}^{-1}$ for $N=5$. This is the fastest rate constant measured for loop formation in any sequence of this length up to date. The rate constant for loop formation in peptides with *trans* proline isomers 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 (Fig. 4). These results show that single glycine and proline residues affect chain dynamics only very locally for loops

shorter than 11 residues. However, this should have pronounced effect of the formation of loops and turns in proteins, which frequently contain proline or glycine residues. The average loop size in proteins is 6-10 residues (31) and β -turns typically have $i,i+3$ or $i,i+4$ interaction (9). Our results show that glycine and proline residues significantly alter the dynamics for formation of these loops.

Effect of *cis* and *trans* prolyl peptide bonds on chain properties. The temperature-dependence of the loop closing reaction showed that the faster dynamics in *cis* proline isomers are due to a significantly increased pre-exponential factor compared to *trans* isomers arguing for entropically favored loop formation in *cis* isomers. A likely origin of this effect are differences in the chain dimensions and/or conformations between *cis* and *trans* isomers, which prompted us to evaluate sterically allowed conformations for our model polypeptide chains. Using Monte Carlo simulations of conformational space introduced by Rose and co-workers (22, 32) 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 (Fig. 5). In these calculations all bond lengths, bond angles and peptide torsion angles were held at the values given by Pappu et al. (22). 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 favourable conformations for each peptide (22). The results show that the energetically favorable conformational space of a *cis* proline isomer in a Ser-Pro bond is reduced compared to the *trans* isomer (Fig. 6A). An extended analysis of non-prolyl *cis* peptide bonds in protein structures also revealed that the geometrical parameters of *cis* peptide bonds in native proteins are significantly different from those of *trans* peptide bonds and that *cis* peptide bonds have drastically restricted conformational space (33).

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 (Fig. 5B, C). 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 (Fig. 5B). These results show that a single *cis* Xaa-Pro peptide bond significantly influences the available conformational space of a polypeptide chain and the chain dimensions. This is compatible with earlier calculations on the conformations of poly(L-proline), which also showed that the presence of a small fraction of *cis* proline isomers drastically reduces the chain dimensions (34). A similar behavior was observed for stereoirregular vinyl polymers (35) and in stereoirregular 1,4-polybutadiene and 1,4-polyisoprene (36), 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 upon loop formation. 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 due to steric clashes on a restricted conformational space (Fig. 5B).

ACKNOWLEDGEMENTS

We thank Gerd Scherer for recording NMR spectra, Thierry Mini and Paul Jenö for mass spectroscopy, Beat Fierz for discussion and Andreas Reiner for comments on the manuscript.

Table 1: Rate constants and activation parameter for intrachain contact formation in various peptides

Peptide	k_1 (10^6 s^{-1}) ^a	A_1 (s^{-1}) ^b	$E_a(1)$ (kJ/mol) ^c	β_1 ^d	k_2 (10^6 s^{-1}) ^a	A_2 (s^{-1}) ^b	$E_a(2)$ (kJ/mol) ^c	β_2 ^d	% cis ^e
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	n.d. ^f	n.d.	-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	n.d.	n.d.	0.83±0.03	-	-	-	-	-
VPS	10±1	(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(psP)S	9.7±0.4	n.d.	n.d.	0.80±0.05	98±5	n.d.	n.d.	-0.78±0.06	72±5
SSPS	28±2	n.d.	n.d.	0.83±0.04	150±20	n.d.	n.d.	-0.87±0.06	18±4
SVPS	16±1	n.d.	n.d.	0.86±0.02	71±5	n.d.	n.d.	-0.96±0.04	18±2
SV(psP)S	15±1	n.d.	n.d.	0.83±0.06	75±5	n.d.	n.d.	-0.98±0.03	78±6
SSSSPSSSS	37±2	(2.4±0.7)·10 ⁹	10.3±0.7	0.93±0.03	-	-	-	-	-
SSSSSSSSS	40±2	(1.7±0.8)·10 ⁹	9.2±1.0	0.92±0.03	-	-	-	-	-
SSSSGSSSS	48±3	(1.6±0.8)·10 ⁹	8.7±1.2	0.93±0.02	-	-	-	-	-

^a k_1 and k_2 at $T = 22.5^\circ\text{C}$ in, in water ($\eta = 0.94 \text{ cP}$).

^bArrhenius pre-exponential factor at $\eta=1 \text{ cP}$ (water at 20°C) according to equation [2].

^cArrhenius activation energy according to equation [2].

^d 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 equation [1]. The origin of the fractional β -values will be discussed elsewhere (F.K. & T.K., in preparation).

^eThe *cis* content reflects the relative amplitude of the fast kinetic phases. The values obtained from measurements at different viscosities were averaged.

^fLoop formation in the *cis* isomer of the SPS peptide was too fast to reliably determine activation energies.

Figures

Figure 1: (A): Schematic of triplet-triplet energy transfer measurements in unfolded polypeptide chains. The time constant for formation of xanthyone triplets (k_T) is about 2 ps and the time constant for triplet-triplet energy transfer is about 1 ps. This allows direct measurement of time constants for chain diffusion processes (k_c) slower than about 10 ps.(19, 37). (B): Time course of formation and decay of xanthyone triplets in peptides of the form Xan-(Ser)₃-NAla-Ser-Gly (—), Xan-Ser-Gly-Ser-NAla-Ser-Gly (—) and Xan-Ser-Pro-Ser-NAla-Ser-Gly (—), measured in 52% glycerol, 10 mM potassium phosphate, pH 7.0 ($\eta=5.90$ cP, $T=22.5^\circ\text{C}$). Intrachain contact formation in Xan-(Ser)₃-NAla-Ser-Gly and Xan-Ser-Gly-Ser-NAla-Ser-Gly can be described by single exponential kinetics with a time constant of $\tau(=1/k)=64\pm 3$ ns and 40 ± 2 ns, respectively. Xan-Ser-Pro-Ser-NAla-Ser-Gly (—) shows more complex kinetics for intrachain contact formation with two relaxation times of $\tau_1=220\pm 15$ ns and $\tau_2=37\pm 3$ ns for contact formation and respective amplitudes of $A_1=20\pm 5\%$ and $A_2=80\pm 5\%$.

Figure 2: Viscosity dependence of the rate constants (A) and amplitudes (B) of the kinetics of loop formation in the *cis* (○) and *trans* (●) prolyl isomers of Xan-Ser-Pro-Ser-Nal-Ser-Gly with glycerol as cosolvent. The solid lines represent fits to equation [2] with values of $k_1=1.9\pm 0.1\cdot 10^7$ s⁻¹, $\beta_1 = 0.75\pm 0.02$, $k_2=1.1\pm 0.2\cdot 10^8$ s⁻¹, $\beta_2 = 0.77 \pm 0.04$. The *cis*-content determined from the averaged amplitudes of the kinetics at all viscosities is 17 ± 2 %.

Figure 3: (A) Kinetics of intrachain contact formation in Xan-Val-**Xaa**-Ser-NAla-Ser-Gly peptides with Xaa= [$\Psi^{\text{Me,Me}}\text{Ser}$] (peptide **A**; —) or Pro (peptide **B**, —) in 10 mM

potassium phosphate, pH 7. In addition the residuals for double exponential fits are shown. The results are given in table 1. (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 activation energies and preexponential factors (A) indicated in table 1. The rate constants were corrected for effects of temperature on solvent viscosity (see eq. [2], Materials and Methods).

Figure 4: Effect of increasing chain length (N) on the rate constant of contact formation in various peptides. The effect of loop-size on the rate constants for loop formation in polyserine (○) and poly(glycine-serine) peptides (●), is compared to the rate constants in $(\text{Ser})_x\text{-Gly-(Ser)}_y$ (●), *cis* (○) and *trans* (●) isomers of $(\text{Ser})_x\text{-Pro-(Ser)}_y$ with x,y varying from 0 to 7. Data for polyserine and poly(Gly-Ser) were taken from ref. (15). The lines represent the loop-size dependence of intrachain loop formation formation in the homopolymers as described in ref (15).

Figure 5: (A) Ramachandran maps for the serine residues preceding either the *cis* or the *trans* prolyl isomer in Ser-Ser-Pro peptides. The contour levels denote free energies of 0.5, 1, 2 and 4 $k_B T$. The crosses indicate the conformation of lowest free energy. (B) Normalized end-to-end distance distribution in *cis* (—) and *trans* (—) prolyl isomers of Ser-Pro-Ser and Ser₄-Pro-Ser₄ peptides compared to Ser₃ and Ser₉ peptides, respectively (—). (C) 10 Representative structures of *cis* and *trans* prolyl isomers of a Ser₃-Pro-Ser₃ peptide. The structures are overlaid with in respect of identical orientation of the central prolyl residue. Structures of the peptides were simulated and end-to-end distances calculated as described in Materials and Methods.

REFERENCES

1. Neri, D., Billeter, M., Wider, G. & Wüthrich, K. (1992) *Science* **257**, 1559-1563.
2. Logan, T. M., Theriault, Y. & Fesik, S. W. (1994) *J. Mol. Biol.* **236**, 637-648.
3. Kortemme, T., Kelly, M. J., Kay, L. E., Forman-Kay, J. D. & Serrano, L. (2000) *J. Mol. Biol.* **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. (2002) *Science* **295**, 1719-1722.
5. Damaschun, G., Damaschun, H., Gast, K. & Zirwer, D. (1998) *Biochemistry (Moscow)* **63**, 259-275.
6. Bachmann, A. & Kiefhaber, T. (2002) *Biophys. Chem.* **96**, 141-151.
7. Millett, I. S., Doniach, S. & Plaxco, K. W. (2002) *Adv. Protein Chem.* **62**, 241-262.
8. Richardson, J. S. (1981) *Advan Prot Chem* **34**, 167-339.
9. Wilmot, C. M. & Thornton, J. M. (1988) *J. Mol. Biol.* **203**, 221-232.
10. Haas, E., Katchalski-Katzir, E. & Steinberg, I. Z. (1978) *Biopolymers* **17**, 11-31.
11. Hagen, S. J., Hofrichter, J., Szabo, A. & Eaton, W. A. (1996) *Proc. Natl. Acad. Sci. USA* **93**, 11615-11617.
12. Bieri, O., Wirz, J., Hellrung, B., Schutkowski, M., Drewello, M. & Kiefhaber, T. (1999) *Proc. Natl. Acad. Sci. USA* **96**, 9597-9601.
13. Lapidus, L. J., Eaton, W. A. & Hofrichter, J. (2000) *Proc. Natl. Acad. Sci. USA* **97**, 7220-7225.
14. Hudgins, R. R., Huang, F., Gramlich, G. & Nau, W. M. (2002) *J. Am. Chem. Soc.* **124**, 556-564.

15. Krieger, F., Fierz, B., Bieri, O., Drewello, M. & Kiefhaber, T. (2003) *J. Mol. Biol.* **332**, 265-274.
16. Chang, I.-J., Lee, J. C., Winkler, J. R. & Gray, H. B. (2003) *Proc. Natl. Acad. Sci. USA* **100**, 3838-3840.
17. Krieger, F., Fierz, B., Axthelm, F., Joder, K., Meyer, D. & Kiefhaber, T. (2004) *Chemical Physics* **307**, 209-215.
18. Klessinger, M. & Michl, J. (1995) *Excited states and photochemistry of organic molecules* (VCH, Weinheim/D).
19. Satzger, H., Schmidt, B., Root, C., Zinth, W., Fierz, B., Krieger, F., Kiefhaber, T. & Gilch, P. (2004) *J. Phys. Chem.* **in press**.
20. Reimer, U., Scherer, G., Drewello, M., Kruber, S., Schutkowski, M. & Fischer, G. (1998) *J. Mol. Biol.* **279**, 449-460.
21. (1972) *Handbook of Chemistry and Physics* (CRC, Cleveland, OH).
22. Pappu, R. V., Srinivasan, R. & Rose, G. D. (2000) *Proc. Natl. Acad. Sci. USA* **97**, 12565-12570.
23. Hopfinger, A. J. (1973) *Conformational Properties of Macromolecules* (Academic Press, New York).
24. Engh, R. A. & Huber, R. (1991) *Acta Cryst.* **47**, 392-400.
25. Schimmel, P. R. & Flory, P. J. (1967) *Proc. Natl. Acad. Sci. USA* **58**, 52-59.
26. Schimmel, P. R. & Flory, P. J. (1968) *J. Mol. Biol.* **34**, 105-120.
27. Scherer, G., Kramer, M. L., Schutkowski, M., Reimer, U. & Fischer, G. (1998) *J. Am. Chem. Soc.* **120**, 5568-5574.
28. Dumy, P., Keller, M., D.E., R., Rohwedder, B., Woehr, T. & Mutter, M. (1997) *J. Am. Chem. Soc.* **119**, 918-925.

29. Keller, M., Sager, C., Dumy, P., Schutkowski, M., Fischer, G. S. & Mutter, M. (1998) *J. Am. Chem. Soc.* **120**, 2714-2720.
30. Brandts, J. F., Halvorson, H. R. & Brennan, M. (1975) *Biochemistry* **14**, 4953-4963.
31. Leszczynski, J. F. & Rose, G. D. (1986) *Science* **234**, 849-855.
32. Fitzkee, N. C. & Rose, G. D. (2004) *Protein Sci.* **13**, 633-639.
33. Jabs, A., Weiss, M. S. & Hilgenfeld, R. (1999) *J. Mol. Biol.* **286**, 291-304.
34. Tanaka, S. & Scheraga, H. A. (1975) *Macromolecules* **8**, 623-631.
35. Flory, P. J. (1969) *Statistical Mechanics of Chain Molecules* (Hanser Publishers, Munich).
36. Tanaka, S. & Nakajima, A. (1972) *Polym. J.* **3**, 600-506.
37. Fierz, B. & Kiefhaber, T. (2004) in *Protein Folding Handbook*, eds. Buchner, J. & Kiefhaber, T. (WILEY-VCH Verlag GmbH & Co KGaA, Weinheim), pp. 845-851.

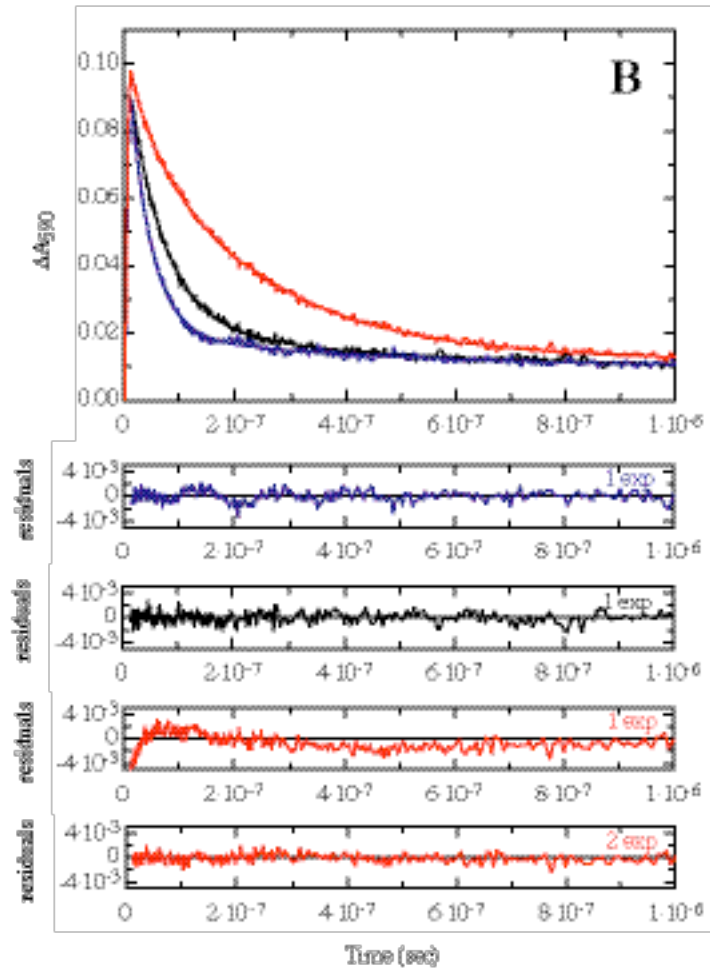
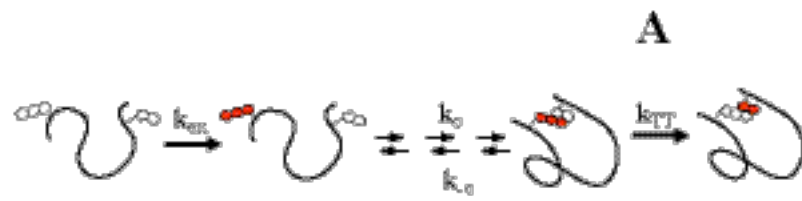


Figure 1

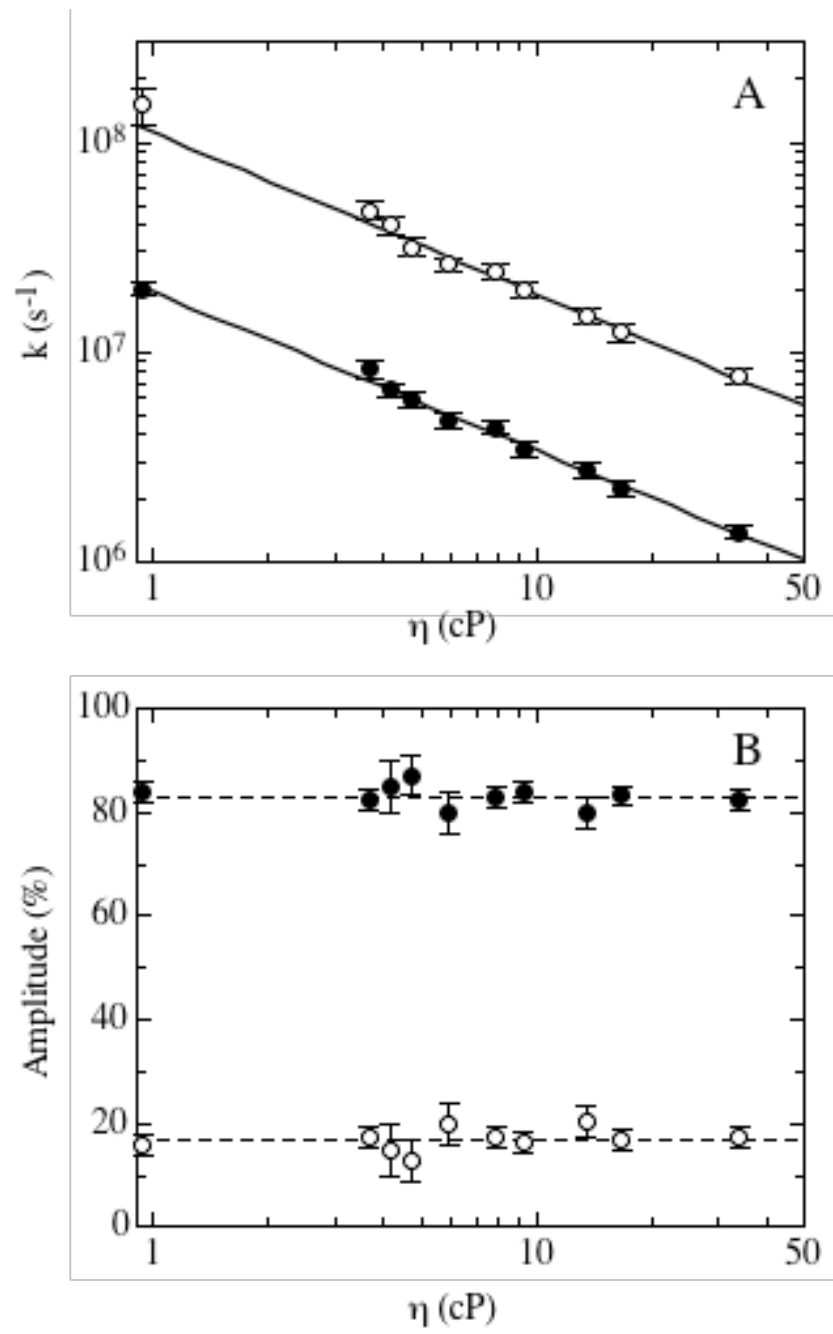


Figure 2

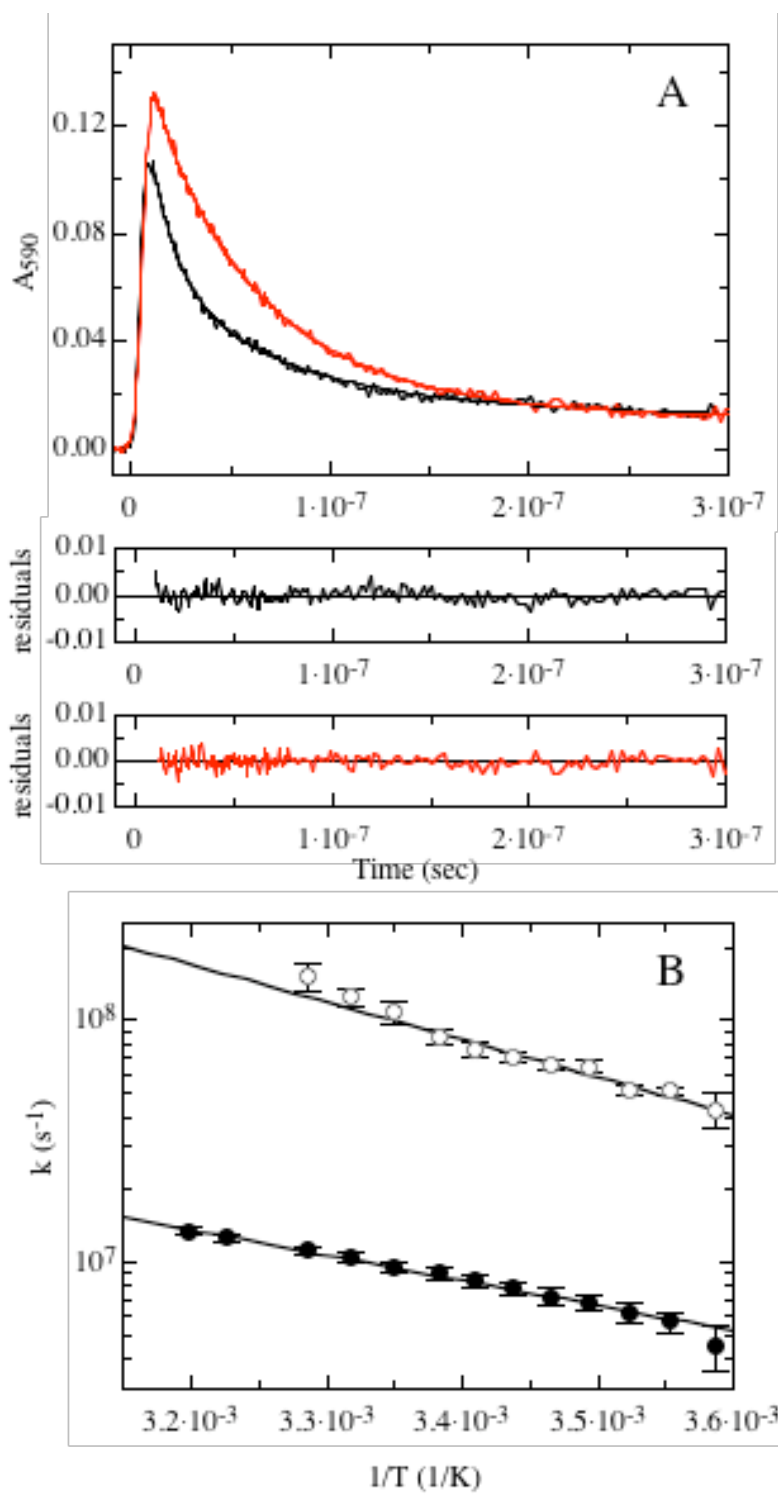


Figure 3

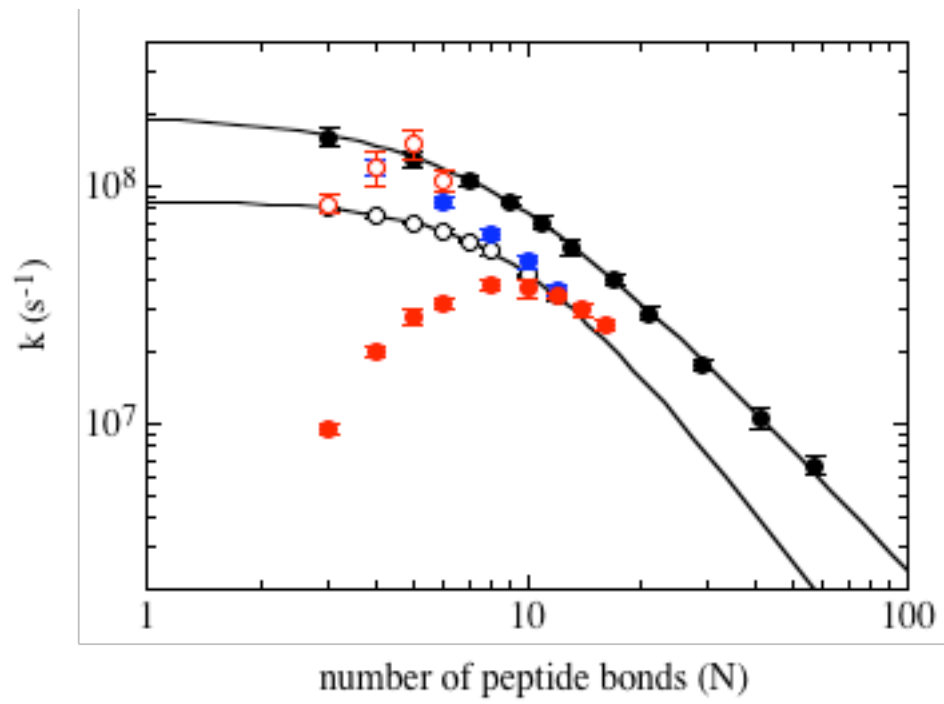


Figure 4

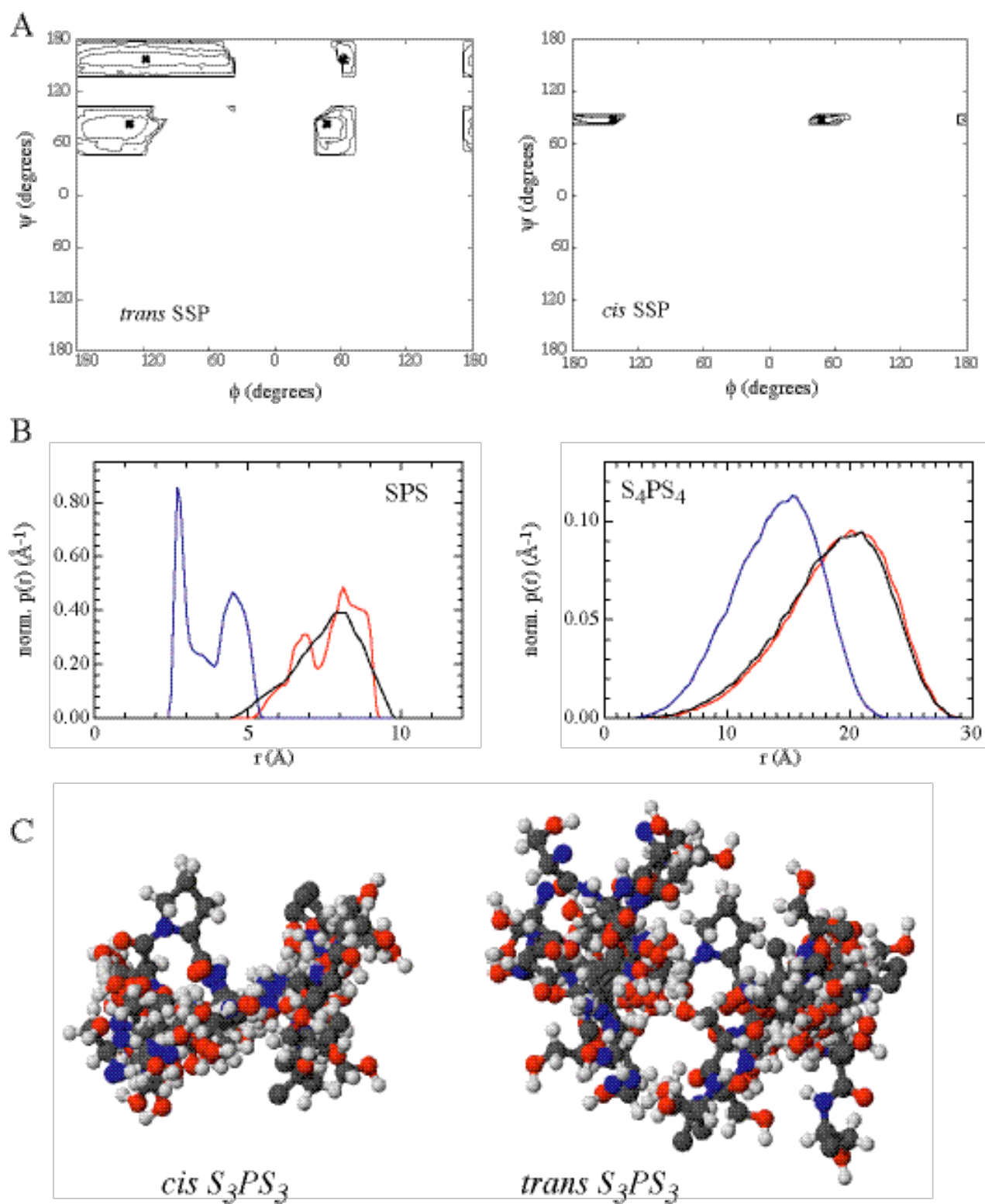


Figure 5

7.6 Kinetics and Barriers of Loop Formation in Natural Protein Sequences

(Krieger and Kiefhaber, to be submitted)

Kinetics and Barriers of Loop Formation in Natural Loop Sequences

Florian Krieger & Thomas Kiefhaber*

Biozentrum der Universität Basel, Department of Biophysical Chemistry,
Klingelbergstr. 70 CH-4056 Basel, Switzerland

*Correspondence should be addressed to T.K.: phone: ++41-61-267 2194; fax: ++41-61-267 2189; e-mail: t.kiefhaber@unibas.ch

Abstract

The rate of intramolecular contact formation during protein folding is limited by intrachain diffusion. Characterization of intrachain dynamics of unstructured polypeptides is particularly important to understand the earliest events in protein folding. We used triplet-triplet energy transfer to determine rate constants for loop formation in various natural protein sequences from carp muscle β -parvalbumin and protein G B1 domain with. The chosen sequences of various lengths up to 18 amino acids are part of different folding motifs such as α -helices, β -hairpins, loops or helix-turn-helix structures in the native state of the proteins. Rate constants for intrachain contact formation, their viscosity dependence and activation parameters obtained from the temperature dependence are compared to data from homo-polypeptide chains. Determined rate constants agree well with data on homo-polypeptides and host-guest studies. The results indicate that intrachain diffusion in natural sequences depend significantly on the amino acid composition. Sequences with large amino acid side-chains decrease intrachain diffusion over short and long distances and increase the activation barriers for end-to-end contact formation. The activation barriers for intrachain contact formation are with 12 – 16 kJ/mol similar to activation barriers determined for fast protein folding processes. The results indicate that intrachain diffusion significantly contribute to the earliest steps in protein folding,

Keywords: Protein folding; polypeptide dynamics; triplet-triplet energy transfer; natural loop sequence; parvalbumin; protein G B1 domain; temperature dependence

Introduction

A protein folding reaction is a complex intramolecular chemical process, which can include reactions ranging from nanosecond time scale to minutes¹. Starting from the ensemble of unfolded conformations a folding polypeptide chain samples the vast conformational space in search for the energetically preferred native state. The folding process ends at the well-defined, three-dimensional, native structure with the complete set of short- and long-range interactions. For a complete characterization of a protein folding reaction it is essential to have knowledge regarding the structure, the dimensions and the dynamics of all states along the reaction pathway. The characterization of the unfolded state as starting point of the refolding reaction is particularly important. A large amount of structural information of the denatured state has been obtained by nuclear magnetic resonance spectroscopy^{2; 3; 4; 5; 6} and small-angle X-ray scattering techniques^{7; 8}. Less is known about dynamics of the unfolded states of a protein folding reaction, particularly the influences of sequence and solvent conditions on dynamics of the polypeptide chain. The formation of a first intramolecular contact between two separated residues is limited by intrachain diffusion. Dynamics of the earliest conformational events related to protein folding such as loop formation, intramolecular side-chain and backbone interactions on a polypeptide chain^{9; 10; 11; 12} or the formation of secondary structure elements^{13; 14} set an upper speed limit for the first steps in protein folding¹⁵.

Recently, several experimental system have we been applied to measure intrachain dynamics in unstructured polypeptide chains^{9; 10; 11; 12}. We applied the method of triplet-triplet energy transfer (TTET) to directly measure intrachain dynamics in unstructured polypeptides^{9; 12; 16}. TTET from xanthone to naphthalene is a diffusion-controlled reaction and provides the requirements to determine absolute end-to-end contact

formation rate constants for a *van-der-Waals*-interaction on polypeptide chains¹⁷. In earlier studies the spectroscopic labels xanthone acid (Xan) and naphthylalanine (NAla) were linked at the ends of flexible peptide chains with alternating glycine and serine residues and at the ends of polyserine peptide chains¹². Except for short proline containing peptides end-to-end contact formation is well-described by single exponential kinetics in all investigated peptides^{9; 12}. The loop-size dependence of the rate constant of intrachain contact formation revealed that intrachain motions over short and long distances are limited by different rate-limiting steps¹². The maximum rate constant of intrachain contact formation in short flexible glycine-serine peptides reaches a limiting value of $1.8 \cdot 10^8 \text{s}^{-1}$ ¹². With increasing chain length end-to-end contact formation is slowed down and for long poly(glycine-serine) chains the rate constants scale with the power law $k \sim N^{-1.7 \pm 0.1}$, with N as the number of peptide bonds between donor and acceptor. The effect of the amino acid sequence on local chain dynamics was probed by using a series of host-guest peptides¹². All amino acids except glycine and proline influence intrachain dynamics slightly. There is a general tendency that peptide chains with large amino acid side-chains like Glu, Arg and Ile show slower intrachain rate constants than peptides with Ala and Ser. Formation of local contacts has shown to be strongly influenced by proline and glycine residues. Glycine increases the dynamics of end-to-end contact formation. Introducing a proline residue into the sequence leads to more complex kinetics. A proline with *trans* peptidyl-prolyl bond slows down local chain dynamics whereas a *cis* proline isomer increases local intrachain dynamics (Krieger et al., submitted). It could be shown that the stereoirregularity of the *cis* peptidyl-prolyl bond induce a chain reversal leading to more compact conformations of the chain. This effect vanishes in peptide chains if the sequence contains more than five amino acids sidewise a single glycine or a single proline respectively. Temperature and

viscosity dependencies of end-to-end contact formation of homo-polypeptides revealed that motions over short distances are limited by enthalpic barriers of 12 – 16 kJ/mol, whereas for long flexible poly(glycine-serine) peptides end-to-end contact formation is nearly completely entropy driven (F. K. & T. K., in preparation).

In the studies described above we used homo-polypeptides and host-guest peptides to determine systematically the length¹², sequence¹² and solvent¹⁸ dependencies of the end-to-end contact formation rate constants in unstructured polypeptide chains. However protein polypeptide chains are more complex than simple homo-polypeptide chains because they consist of 20 different amino acids with a variety of side-chains of quite different chemical reactivities. It could be shown that TTET from xanthone to naphthalene can be applied to natural protein sequences, if they are free of tryptophane, tyrosine, methionine and histidine, since these amino acids interact with xanthone triplets¹⁶. In a preliminary study we could successfully apply intramolecular TTET to determine rate constants of intrachain contact formation in a natural loop sequence of carp muscle parvalbumin. To test whether the experimental results obtained from the studies on model homo-polypeptides can be generally applied to dynamics of protein polypeptide chains we characterized the dynamics and barriers for contact formation between the ends of various natural sequences from two well characterized proteins, carp muscle β -parvalbumin and protein G B1 domain, under various conditions.

Results

Fragments from Parvalbumin and protein G B1 domain as model systems to study intrachain dynamics with triplet-triplet energy transfer

Carp muscle β -parvalbumin belongs to the family of EF-hand proteins and exhibits a very high affinity for Ca^{2+} and moderate affinity for Mg^{2+} , properties consistent with a

Ca²⁺-buffering role¹⁹. The overall structure consists of six helices associate in three helix-loop-helix motifs, AB, CD and EF. The loops between the helices C and D and between the helices E and F bind one calcium ion. The parvalbumin helices are all amphiphilic, with their hydrophobic faces forming the core of the structure²⁰. 80 % of the hydrophobic side-chains buried in the protein core are located in α -helices. Carp muscle β -parvalbumin (PA) does not contain any tyrosine, tryptophane or methionine residues²¹ and therefore represents a perfect system to study intrachain diffusion by TTET from xanthone to naphthalene¹⁶. We have synthesized three fragments from different structures of parvalbumin, PA⁶⁶⁻⁷⁰, PA⁷⁰⁻⁸⁵ and PA⁸⁵⁻¹⁰² (Figure 1A-C, Table 1). PA⁶⁶⁻⁷⁰ is part of the D-helix, PA⁷⁰⁻⁸⁵ is part of the loop connecting the D- and E-helices and PA⁸⁵⁻¹⁰² is part of the calcium binding side of the EF-hand. All loops bring phenylalanine residues at the N- and C-terminus in close contact in the native structure. The flanking phenylalanines 66, 70, 85 and 102 are part of the hydrophobic core and were replaced by the triplet donor xanthone acid and acceptor naphthylalanine, respectively.

Residue 41-56 of the protein G B1 domain (GB1) form a stable β -hairpin in solution and became a reference system to study β -hairpin folding and stability. It has been the subject of various studies like temperature-jump experiments^{14; 22}, simulations^{23;24} and modeling²⁵ of folding pathways. The stability of the β -hairpin is attributed to hydrophobic cluster formation between the valine side-chain (residue 43) and the rings of three aromatic residues (Tyr45, Phe52, Trp54). We synthesized a peptide fragment derived from the protein G B1 domain residues 45-52 (Figure2, Table 1). Tyrosine 45 and phenylalanine 52 are part of the hydrophobic cluster of the isolated stable β -hairpin and were replaced by the triplet donor and acceptor, respectively.

Circular dichroic measurements revealed that the chosen protein fragments are unstructured in water (data not shown). TTET measurements revealed that end-to-end contact formation in all protein fragments can be described as a single-exponential process as previously observed for all investigations on proline-free peptide chains^{9;12}. Figure 3A shows the time-course of TTET from xanthone to naphthylalanine in the PA⁷⁰⁻⁸⁵ fragment. End-to-end contact formation of PA⁷⁰⁻⁸⁵ is significantly slower than expected from results obtained on polyserine peptides and also from the previously described data on PA⁸⁵⁻¹⁰² (Figure 4). The PA⁷⁰⁻⁸⁵ sequence contains a number of amino acids with large side-chains, such as Lys, Arg, Leu and Thr and only one glycine residue. Amino acids with β -branched side-chains were shown to slow down intrachain dynamics compared to amino acids with small side-chains¹². These results suggest that amino acids with large side-chains decrease the flexibility of the peptide chains leading to significantly slower contact formation rate constants than observed for polyserine peptides. The rate constants for intrachain contact formation in the short helical fragment PA⁶⁶⁻⁷⁰ agrees well with results obtained from polyserine peptides, although the sequence consist of amino acids with larger side-chains than the (Ser)₃-peptide (Figure 4, Table2). For sequences consisting of amino acids with small side-chains or with high glycine content the intrachain rate constants of polyserine predict quite well intrachain contact formation rate constants of natural loop sequences.

The GB1 fragment GB1⁴⁵⁻⁵² consists of several amino acids with large side-chains. As shown in figure 4 end-to-end chain dynamics in the GB1⁴⁵⁻⁵²-fragment are significantly slowed down compared to polyserine peptide chains of the same length (Figure 4, Table 2). The 2-fold decrease in the rate constant of contact formation is compatible with the presence of large amino acid side-chains. Glycine and proline residues are frequently found in hairpin loops²⁶. We observed in homo-polypeptide chains that the local chain

dynamics are strongly influenced by glycine or proline residues (F. Krieger et al., submitted). To test if natural protein sequences show similar behavior we replaced Thr49 in the GB1⁴⁵⁻⁵² sequence against a glycine and proline (Table 1). The substitution of a single glycine residue against a threonine residue into the GB1⁴⁵⁻⁵² sequence accelerates intrachain contact formation, whereas a proline residue slows down intrachain dynamics (Figure 3B and Figure 4B). As observed for host-guest peptides short proline containing peptides exhibit complex kinetics on local intrachain dynamics, suggesting that peptides with a *cis* proline isomer show faster end-to-end contact formation kinetics than peptides with a *trans* proline isomer (F. Krieger et al., submitted). However end-to-end contact formation of the proline containing GB1⁴⁵⁻⁵² fragment is described by single-exponential kinetics, suggesting that peptides with *cis* and *trans* proline isomers have similar intrachain dynamics.

Except the PA⁶⁶⁻⁷⁰ fragment all chosen protein polypeptide sequences consist of many charged amino acids, which might influence intrachain dynamics (Table 1). The PA⁸⁵⁻¹⁰² sequence contains a stretch of alternating negatively charged side-chains and uncharged residues between residues 89 and 96 (GDSDGDG), which represents the calcium-binding motif of the EF-hand structure (Figure 1C). From polyelectrolytes like poly(styrene-sulfonate) or DNA at neutral pH it is known that the backbone rigidity is supported by Coulomb repulsion of the side chains, since same charges distributed along a chain at high density repel each other to extend the otherwise flexible chain²⁷. To investigate whether charged side-chains influence intrachain dynamics in the highly charged PA⁸⁵⁻¹⁰² loop sequences we replaced the negatively charged side chains amino acids aspartic acid (D) and glutamic acid (E) by the uncharged amino acid side chains asparagines (N) and glutamine (Q) (Table 1). The variant PA⁸⁵⁻¹⁰² fragment shows virtually the same dynamic properties as the wild type fragment, indicating that charged

amino acid side-chain have no or only a little effect on the intrachain dynamics of natural protein sequences.

Effect of viscosity and temperature on the intrachain dynamics of natural protein sequences

The characterization of end-to-end contact formation in poly(glycine-serine) peptide chains revealed that intrachain dynamics over short distances is not fully governed by solvent motion and limited by activation barriers of 12-14 kJ/mol. In contrast, formation of loops with more than 20 amino acids is a diffusion-controlled process and nearly completely entropy driven. Generally, the effect of increasing solvent viscosity on the kinetics of intrachain loop formation can be described by following relation:

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

where η_0 is the reference solvent viscosity at 20°C and k_c^0 the rate constant of end-to-end contact formation at η_0 . β -values of -0.75 were determined for the short (GS)₁-peptide and of -1.0 for long peptides (F.K & T. K to be submitted). The fractional exponent β reflects the sensitivity of the reaction to solvent viscosity. For a fully diffusion-controlled reaction the exponent adopts the value -1. A value of 0 indicates that the reaction rate constant is independent of solvent viscosity and becomes fully reaction-controlled. Values between 0 and 1 can be interpreted as a coupling between solvent motions and the crossing over an activation barrier, leading to a partly diffusion-controlled mechanism^{28; 29}. For all peptides derived from natural loop structures intrachain contact formation is partly diffusion-controlled with β -values between 0.84 and 0.90 (Figure5A). Similar results could be obtained for short glycine-serine peptides and polyserine peptides of similar length (Table2) (F.K & T. K to be submitted).

To verify the partly diffusion-controlled mechanism of intrachain dynamics in the chosen protein fragments we performed the temperature dependency on the end-to-end contact formation rate constants. Increasing the temperature leads to a decrease in the viscosity of water³⁰. To correct for the contributions of changes in solvent viscosity due to temperature changes we used β -values in combination with the viscosities of the corresponding temperature to calculate the viscosity-corrected rate constants for end-to-end contact formation k_c^0 according to equation (2):

$$k_c^0(T) = k_c(T) \cdot \left(\frac{\eta_0}{\eta(T)} \right)^\beta \quad (2).$$

This correction reduces the apparent activation barrier by 17.7 kJ/mol, since the viscosity of water changes due to temperature changes³⁰. The viscosity-corrected rate constants for loop formation k_c^0 still dependent significantly on temperature as shown in figure 5B. All temperature dependencies of the investigated peptides can be described by an Arrhenius type equation (3) in the denoted temperature range:

$$k_c^0(T) = A \cdot e^{-\frac{E_a}{RT}} \quad (3).$$

A is the Arrhenius pre-exponential factor, independent of solvent viscosity and E_a represents the Arrhenius activation energy. The fits yielded apparent Arrhenius activation energies for the GB1⁴⁵⁻⁵² fragments E_a of 15.7 ± 0.9 kJ/mol and a pre-exponential factors A of $1.4 - 2.8 \times 10^{10}$ for a reference viscosity η_0 of 1.0 cP (Table2). The activation barriers and pre-exponential factors of end-to-end contact formation in GB1⁴⁵⁻⁵² fragments are significantly higher compared to results obtained from the polyserine peptide of the same length ((Ser)₆, Table 2). Interestingly the replacement of the Thr49 residue by a glycine or proline residue does not change the Arrhenius activation energy significantly, but the pre-exponential factor is strongly influenced (Table 2).

With increasing chain length the apparent activation barriers for end-to-end contact formation decrease as observed in poly(glycine-serine) and polyserine peptides, suggesting that intrachain motions over long distances are stronger influenced by the solvent motion and becomes more entropy driven. In both the PA⁷⁰⁻⁸⁵ and the PA⁸⁵⁻¹⁰² fragments intrachain loop formation is limited by activation barriers of 14.3 ± 0.7 and 12.5 ± 0.7 kJ/mol, respectively (Table 2). Intrachain dynamics in the long protein fragments PA⁷⁰⁻⁸⁵ and PA⁸⁵⁻¹⁰² are significantly stronger influenced by temperature than intrachain dynamics in polyserine peptides like (Ser)₉. Unfortunately we cannot directly measure end-to-end contact formation of a Xan-(Ser)₁₆-Nal peptide for direct comparison due to limitation in peptide synthesis.

Discussion

Intrachain contact formation in unstructured protein polypeptide sequences

Natural protein sequences are more complex than simple homo-polypeptides, since they consist of twenty different amino acids with a variety of chemical reactivities. Therefore a natural unstructured protein sequence cannot be regarded as a regular homogeneous sequence and the application of simple scaling laws derived from polymer theoretical models breaks down. The investigated natural peptide sequences contain between four and 18 amino acids. The results are comparable with predictions obtained from data on homo-polypeptides and host-guest peptides, except the data of the PA⁶⁶⁻⁷⁰ fragment. Generally, it is observed that amino acid sequences with small side-chains show faster contact formation rates than sequences with β -branched side-chains¹². The heterogeneous nature of protein polypeptide chain leads to local changes of the persistence length³¹, which is a quantity for the stiffness of a chain. Persistence length of 5 – 10 amino acids were experimentally determined depending on the protein and

conditions. Thus about 5-10 amino acids can be regarded as the bead size of an unstructured protein chain, however with quite different properties. Indeed our results show, that local chain dynamics in unstructured protein polypeptide sequences are quite different. Depending on the sequence the time constant of the formation of an $i, i + 6$ interaction will be estimated between 10 ns for the (GS)₃-peptide and 40 ns for the GB1⁴⁵⁻⁵² T49P. It is expected that with increasing chain length unstructured protein polypeptide chains adopt more similar averaged properties. End-to-end contact formation over 16 and 18 residues in protein sequences from parvalbumin is significantly slowed down compared to short sequences. Time constants of 66 and 53 ns are determined for end-to-end contact formation in the PA⁷⁰⁻⁸⁵- and PA⁸⁵⁻¹⁰²-fragments, respectively. Both fragments are parts of loops and helices and have a quite similar length.

Intrachain contact formation is particularly important over short distances, since the formation of secondary structures is mostly due to interactions over few residues. The strong length dependence of intrachain contact formation over long distances decreases the probability to form a first elementary contact. An intramolecular contact over more than 25 amino acid residues will not be formed faster than in 100 ns and over 60 residues will not be formed faster than in 500 ns (Figure 4)¹². In this time range parts of the folding polypeptide chain might already adopt secondary structure.

Origin of decreased contact formation rates

Our results show that the intrinsic dynamic properties of protein polypeptide chains behave different as expected from simple scaling laws, originally developed from polymer theory^{32; 33}. Polymer theories are based on models with significantly less detail and high degree of simplifications. The heterogeneity of protein sequences has a strong

influence on the conformational properties of the chain, leading to more complex dynamical and structural behavior. The temperature and viscosity dependencies showed that intrachain contact formation in protein sequences over both short and longer distances ($N \leq 18$) is not fully diffusion-controlled and significantly influenced by activation barriers of 12-16 kJ/mol. To form an end-to-end contact the moving ends of a polypeptide chain have to sample conformations, which might be energetically unfavorable, leading to a limited diffusion process as observed for GB1⁴⁵⁻⁵²- and PA-fragments. Interestingly the substitutions of a threonine by a glycine or proline residue in the β -hairpin sequences 45-52 of GB1 leads to apparent changes in contact formation rate constants and the pre-exponential factor, but not in the activation barrier (Table 2). Glycine and proline significantly change the average end-to-end distance of short unstructured peptide chains³³. It is therefore expected that the decrease in the average end-to-end distance leads to an increase in end-to-end contact formation rate constants³⁴, as observed in variants of the GB1⁴⁵⁻⁵² fragment. However the replacement of the Thr49 by a glycine or proline do not affect the barrier of intrachain contact formation. The pre-exponential factor of all three GB1-fragments is increased compared to the (Ser)₆-peptide. This increase might indicate that intrachain motion in the GB1-fragments are less entropic-driven compared to (Ser)₆ due to a stronger restriction of conformational space.

End-to-end contact formation in natural protein sequences is mainly influenced by the property of the backbone. The incorporation of an additional glycine or proline significantly changes local chain dynamics. The incorporation of larger side-chains has a cumulative effect, suggesting that this effect is significant, if several side-chains are changed. Interestingly the deletion of charges in repetitive glycine-aspartic acid units

does not affect intrachain dynamics, suggesting that charges have only a slight influence on the dynamics of natural polypeptide chains.

Intrachain diffusion in GB1⁴⁵⁻⁵²-fragment and implication for the first steps in β -hairpin formation.

The presented results reveal that intrachain contact formation in GB1⁴⁵⁻⁵² is strongly influenced by the properties of the peptide chain. The chosen sequence is part of a stable β -hairpin sequence of the protein G B1 domain. Using temperature jump techniques on β -hairpin folding, Munoz et al. determined rate constants for hairpin unfolding. Assuming a two-state behavior between folded and unfolded conformations a folding time constant of 6 μ s at 297 K was estimated¹⁴. To explain the apparent two-state behavior a statistical mechanical zipper model was developed, suggesting that the most probable first step in β -hairpin formation begins with the formation of the β -turn (residues 46-51)²⁵. However, it is controversial to assume a two-state model for such small peptides and up to date no direct determinations of rate constants for β -hairpin folding are available to prove these results. The temperature jump experiments revealed rate constants for the overall folding of the hairpin. Our results suggest that the time scale of the first step in formation of the GB1⁴¹⁻⁵⁶ β -hairpin, namely the formation of an $i, i + 4$ interaction in the β -turn, can be estimated to 25-30 ns. This time range is significantly faster than the estimated rate constant for β -hairpin folding. Therefore the loop of the GB1 hairpin (residues 41-56) cannot be formed faster than in 30 ns.

Interestingly the presented data contradict previous results, which suggest that folding of the β -hairpin GB1⁴¹⁻⁵⁶ is nearly diffusion-controlled as shown with the dependence of viscosity on the folding rate constants²². Using a simple statistical mechanical model, the intrinsic activation energy for the elementary step of rotating a pair of backbone

angles was estimated to be zero and the folding is fully governed by solvent motion. The discussed data were obtained from temperature jump experiments assuming a two-state model for folding and unfolding of the β -hairpin GB1⁴¹⁻⁵⁶ and same activation barriers for folding and unfolding.

With the presented results we show that at least the formation of a first intramolecular contact is limited by an apparent activation barrier of 16 kJ/mol. Although the investigated GB1⁴⁵⁻⁵² sequences is shorter than the GB1⁴¹⁻⁵⁶ sequences it is expected that the formation of a first intramolecular contact over seven peptide bonds in the complete GB1⁴¹⁻⁵⁶ sequence is also strongly limited by activation barriers, which might originate from the intrinsic stiffness of the chain. However, at the moment it is not clear, how this barrier influence the overall folding process of small β -hairpin peptides, since no folding data are available, which describe accurately the effect of temperature on folding kinetics.

Furthermore the results suggest that the rate constant for hairpin formation should significantly be affected by insertion of proline or glycine residue.

Materials and Methods

9-Oxoxanthen-2-carboxylic acid synthesis was performed as reported previously¹². All peptides were synthesized on a Rink-amide resin using standard Fmoc-chemistry. The formation and decay of xanthon triplets were measured at 590 nm using transient absorption technique as described previously¹². All measurements were performed at pH 7.0 in 10 mM sodium cacodylate for parvalbumin sequences and in 10 mM potassium phosphate for GB1 sequences. Viscosity dependencies were measured at $22.5 \pm 0.1^\circ\text{C}$. Solvent viscosity was varied adding different amount of glycerol to solutions. For all glycerol solutions, the solvent viscosity was measured by using a falling ball viscometer (Haake) at $22.5 \pm 0.1^\circ\text{C}$. The temperature dependencies were carried out in range from 1 – 50°C . Individual kinetics were measured five times, averaged and analyzed using ProFit software program. The kinetic rate constants for the temperature dependencies were corrected against viscosity using tabular values for the viscosity of water³⁰. Measurements of the circular dichroism (CD) of the peptide fragments were performed at 22.5°C using an Aviv 62DS spectropolarimeter equipped with a temperature-control unit.

Acknowledgements

We thank Beat Fierz and Andi Möglich for discussions and Thierry Mini and Paul Jenö for peptide mass analysis. Xanthonic acid was synthesized by Joseph Wey.

Tables

Table 1: Sequences and structural context of the investigated sequences in native proteins. PA: carp muscle β -parvalbumin, GB1: protein G B1 domain, HLH: Helix-loop-helix structure.

Peptide sequence	Amino acid sequence	Native structure of
PA 66-70	Xan-LQN-NAlaSG-NH ₂	α -Helix
PA 70-85	Xan-KADARALTDGETKT-NAlaSG-NH ₂	HLH
PA 85-102	Xan-LKAGDSDGDGKIGVDE-NAlaSG-NH ₂	HLH
PA85-102 (D/N,E/Q)	Xan- LKAGNSNGNGKIGVNQ-NAlaSG-NH ₂	HLH
GB1 45-52 wt	Xan-DDATKT-NAlaSG-NH ₂	β -Hairpin
GB1 45-52 T/P	Xan-DDAPKT-NAlaSG-NH ₂	β -Hairpin
GB1 45-52 T/G	Xan-DDAGKT-NAlaSG-NH ₂	β -Hairpin

Table 2: Observed rate constants and activation parameters for loop formation in natural unstructured peptides, derived from protein G B1 domain (GB1) and β -parvalbumin (PA) and direct comparison to results obtained from polyserine peptides.

Peptide sequence	N	k_c (s ⁻¹) ^b	β	E_a (kJ/mol) ^c	A (10 ⁹ ·s ⁻¹) ^c
PA66-70 ^a	4	$6.3 \pm 0.4 \times 10^7$	n.d.	n.d.	n.d.
S ₃	4	$7.4 \pm 0.5 \times 10^7$	0.83 ± 0.03	14.7 ± 1.0	28 ± 9
GB1 45-52 wt	7	$3.2 \pm 0.2 \times 10^7$	0.90 ± 0.05	15.7 ± 0.9	18 ± 7
S ₆	7	$5.8 \pm 0.4 \times 10^7$	0.89 ± 0.05	11.2 ± 0.6	5.2 ± 1.1
GB1 45-52 T/P	7	$2.4 \pm 0.2 \times 10^7$	0.90 ± 0.05	15.7 ± 0.8	14 ± 4
GB1 45-52 T/G	7	$5.0 \pm 0.3 \times 10^7$	0.90 ± 0.04	15.5 ± 0.7	24 ± 7
PA70-85	15	$1.5 \pm 0.1 \times 10^7$	0.86 ± 0.03	14.3 ± 0.7	4.6 ± 1.9
PA85-102	17	$1.9 \pm 0.1 \times 10^7$	0.85 ± 0.03	12.5 ± 0.7	2.7 ± 0.3
PA85-102 (D/N,E/Q)	17	$1.9 \pm 0.1 \times 10^7$	0.84 ± 0.03	12.5 ± 0.7	2.7 ± 0.3

a: not soluble in water, The rate constant was obtained from the guanidine hydrochloride dependence and extrapolation to zero molar denaturant concentration as reported^{12; 18}.

b: measured at 22.5°C.

c: according to equation 3 and normalized to a viscosity η_0 of 1.0 cP,

Figures

Figure 1: Ribbon diagrams of carp muscle β -parvalbumin²⁰ (PA) and the molecular structures of the fragments used in this study. The figures were prepared using the program MOLMOL³⁵ and the pdb file 4CPV²⁰. (A) Ribbon diagram of PA with the helical region 66-70 (red) and sequences of the synthesized fragments with the two flanking Phe66 (pink) and Phe70 (green). (B) Ribbon diagram of PA and sequences of the synthesized fragments with the loop region 70-85 (red) and the two flanking Phe70 (green) and Phe85 (blue). (C) Ribbon diagram of PA and sequences of the synthesized fragments with the loop region 85-102 (red) and the two flanking Phe85 (blue) and Phe102 (red). The flanking phenylalanines Phe66 (pink), Phe70 (green), Phe85 (blue) and Phe102 (red) have been replaced by the triplet donor and acceptor labels, xanthone acid and naphthylalanine, respectively.

Figure 2: (A) Solution structure of the protein G B1 domain (residue 1-56)³⁶ and (B) molecular structure of the synthesized fragment GB1⁴⁵⁻⁵². The flanking Tyr45 (red) and Phe(52) (blue) have been replaced by the triplet donor and acceptor labels, xanthone acid and naphthylalanine, respectively. In a comparative study Thr49 (green) were additionally replaced by a glycine and proline. The figure was prepared using the program MOLMOL³⁵ and the pdb file 1GB1³⁶.

Figure 3: (A) Time-course of formation and decay of xanthone triplets in PA⁷⁰⁻⁸⁵-fragment and (B) GB1⁴⁵⁻⁵² (—), T49G, (—) GB1⁴⁵⁻⁵² and T49P (—) after a 4 ns laser flash at t = 0. The change of xanthone triplet absorbance is measured at 590 nm. End-to-end contact formation is described as single exponential process with time constants of (A) $\tau = 1/k = 66 \pm 3$ ns for PA⁷⁰⁻⁸⁵ and (B) of $\tau = 30 \pm 2$ ns (GB⁴⁵⁻⁵²), 40 ± 3 ns (GB⁴⁵⁻⁵² T49P) 20 ± 1 ns and (GB⁴⁵⁻⁵² T49G).

Figure 4: Comparison of the end-to-end diffusion rate constants of various protein sequences derived from β -parvalbumin (PA) and protein G B1 domain (GB1) used in this study (Fig.1, 2, Tab.1) with rate constants obtained from poly(glycine-serine) peptides (●) and polyserine peptides (○): (●) PA⁶⁶⁻⁷⁰; (●) PA⁷⁰⁻⁸⁵; (●) PA⁸⁵⁻¹⁰²; (●) GB1⁴⁵⁻⁵²; (●) GB1⁴⁵⁻⁵², T49G; (●) GB1⁴⁵⁻⁵², T49P. The experimentally observed rate constants for contact formation are plotted against the number of peptide bonds (N) between xanthone and naphthalene. The solid lines represent fits according ref.¹².

Figure 5: Viscosity dependencies (A) and temperature dependencies (B) of end-to-end contact formation in sequences derived from protein G B1 domain (GB1): (●) GB1⁴⁵⁻⁵², (●) GB1⁴⁵⁻⁵²T49G and (●) GB1⁴⁵⁻⁵²T49P. For the viscosity dependence (A) observed rate constant were analyzed according to equation 1. The slopes $-\beta$ of the fit of the viscosity dependencies are given in table 2. For the temperature dependencies (B) observed rate constants were corrected against viscosity according to equation (2) and analyzed using an Arrhenius equation. The fit yielded activation barriers of 15.7 ± 0.9 kJ/mol and pre-exponential factors k_0 of $1.8 \pm 0.7 \times 10^{10} \text{ s}^{-1}$ for GB1⁴⁵⁻⁵², $2.4 \pm 0.7 \times 10^{10} \text{ s}^{-1}$ for GB1⁴⁵⁻⁵²T49G and $1.4 \pm 0.4 \times 10^{10} \text{ s}^{-1}$ for GB1⁴⁵⁻⁵²T49P.

References

1. Bieri, O. & Kiefhaber, T. (1999). Elementary steps in protein folding. *Biol. Chem.* 380, 923-929.
2. Shortle, D. (2002). The expanded denatured state: an ensemble of conformations trapped in a locally encoded topology space. *Adv. Prot. Chem* 62, 1-23.
3. Neri, D., Billeter, M., Wider, G. & Wüthrich, K. (1992). NMR determination of residual structure in a urea-denatured protein, the 434-repressor. *Science* 257, 1559-1563.
4. Dyson, H. J. & Wright, P. E. (2004). Unfolded proteins and protein folding studies by NMR. *Chem. Rev.* 104, 3607-3622.
5. Wilkins, D. K., Grimshaw, S. B., Receveur, V., Dobson, C. M., Jones, J. A. & Smith, L. J. (1999). Hydrodynamic radii of native and denatured proteins measured by pulse field gradient NMR techniques. *Biochemistry* 38, 16424-16431.
6. Schwalbe, H., Fiebig, K. M., Buck, M., Jones, J. A., Grimshaw, S. B., Spencer, A., Glaser, S. J., Smith, L. J. & Dobson, C. M. (1997). Structural and Dynamical Properties of a Denatured Protein. Heteronuclear 3D NMR Experiments and Theoretical Simulations of Lysozyme in 8 M Urea. *Biochemistry* 36, 8977-8991.
7. Millet, I. S., Doniach, S. & W., P. K. (2002). Towards a taxonomy of the denatured state: Small angle scattering studies of unfolded proteins. *Adv. Prot. Chem* 62.
8. Damaschun, G., Damaschun, H., Gast, K., Misselwitz, R., Muller, J. J., Pfeil, W. & Zirwer, D. (1993). Cold Denaturation Induces Conformational Changes in Phosphoglycerate Kinase from Yeast. *Biochemistry* 32, 7739-7746.

9. Bieri, O., Wirz, J., Hellrung, B., Schutkowski, M., Drewello, M. & Kiefhaber, T. (1999). The speed limit for protein folding measured by triplet-triplet energy transfer. *Proc. Natl. Acad. Sci. USA* 96, 9597-9601.
10. Lapidus, L. J., Eaton, W. A. & Hofrichter, J. (2000). Measuring the rate of intramolecular contact formation in polypeptides. *Proc. Natl. Acad. Sci. USA* 97, 7220-7225.
11. Hudgins, R. R., Huang, F., Gramlich, G. & Nau, W. M. (2002). A fluorescence-based method for direct measurements of submicrosecond intramolecular contact formation in biopolymers: an exploratory study with polypeptides. *J. Am. Chem. Soc.* 124, 556-564.
12. Krieger, F., Fierz, B., Bieri, O., Drewello, M. & Kiefhaber, T. (2003). Dynamics of unfolded polypeptide chains as model for the earliest steps in protein folding. *J. Mol. Biol.* 332, 265-274.
13. Thompson, P., Eaton, W. & Hofrichter, J. (1997). Laser temperature jump study of the helix-coil kinetics of an alanine peptide interpreted with a "kinetic zipper" model. *Biochemistry* 36, 9200-9210.
14. Munoz, V., Thompson, P., Hofrichter, J. & Eaton, W. (1997). Folding dynamics and mechanism of β -hairpin formation. *Nature* 390, 196-199.
15. Kubelka, J., Hofrichter, J. & Eaton, W. A. (2004). The protein folding 'speed limit'. *Curr. Opin. Struct. Biol.* 14, 76-88.
16. Krieger, F., Fierz, B., Axthelm, F., Joder, K., Meyer, D. & Kiefhaber, T. (2004). Intrachain diffusion in a protein loop fragment from carp parvalbumin. *Chem. Phys.* 307, 209-215.
17. Satzger, H., Schmidt, B., Root, C., Zinth, W., Fierz, B., Krieger, F., Kiefhaber, T. & Gilch, P. (2004). Ultrafast quenching of the xanthone triplets by energy

- transfer: New insight into the intersystem crossing kinetics. *J. Phys. Chem. A*, in press.
18. Moglich, A., Krieger, F. & Kiefhaber, T. (2004). Molecular basis for the effect of urea and guanidinium chloride on the dynamics of unfolded polypeptide chains. *J. Mol. Biol.*, 345, 153-162.
 19. Kretsinger, R. H. (1980). Structure and evolution of calcium-modulated proteins. *CRC Crit. Rev. Biochem.* 8, 119-174.
 20. Kumar, V. D., Lee, L. & Edwards, B. F. (1990). Refined crystal structure of calcium-liganded carp parvalbumin 4.25 at 1.5-Å resolutions. *Biochemistry* 29, 1404-1412.
 21. Coffee, C. J. & Bradshaw, R. A. (1973). Carp muscle calcium-binding protein. I. Characterization of the tryptic peptides and the complete amino acid sequence. *J. Biol. Chem.* 248, 3305-3312.
 22. Jas, G. S., Eaton, W. A. & Hofrichter, J. (2001). Effect of viscosity on the kinetics of α -helix and β -hairpin formation. *J. Phys. Chem. B* 105, 261 - 272.
 23. Dinner, A. R., Lazaridis, T. & Karplus, M. (1999). Understanding β -hairpin formation. *Proc Natl Acad Sci U S A* 96, 9068-9073.
 24. Pande, V. S. & Rokhsar, D. S. (1999). Molecular dynamics simulations of unfolding and refolding of a β -hairpin fragment of protein G. *Proc Natl Acad Sci U S A* 99, 9062-9067.
 25. Munoz, V., Henry, E. R., Hofrichter, J. & Eaton, W. A. (1998). A statistical mechanical model for β -hairpin kinetics. *Proc Natl Acad Sci U S A* 95, 5872-5879.
 26. Wilmot, C. M. & Thornton, J. M. (1988). Analysis and prediction of the different types of β -turns in proteins. *J. Mol. Biol.* 203, 221-232.

27. Teraoka, I. (2001). *Polymer solutions: an introduction to physical properties.*, John Wiley & Sons, New York.
28. Grote, R. F. & Hynes, J. T. (1980). The stable states picture of chemical reaction. II. Rate constants for condensed and gas phase reaction models. *J. Chem. Phys.* 73, 2715 - 2732.
29. Hanggi, P., Talkner, P. & Borkovec, M. (1990). Reaction-rate theory: fifty years after Kramers. *Rev. Mod. Phys.* 62, 251 - 341.
30. (1972). *Handbook of Chemistry and Physics*. 53rd edit (Weast, R. C., Ed.), CRC, Cleveland, OH.
31. Bright, J. N., Woolf, T. B. & Hoh, J. H. (2001). Predicting properties of intrinsically unstructured proteins. *Prog. Biophys. Mol. Biol.* 76, 131-173.
32. de Gennes, P. G. (1979). *Scaling Concepts in Polymer Physics*, Cornell University Press, Ithaca, New York.
33. Flory, P. J. (1969). *Statistical Mechanics of Chain Molecules*, Hanser Publishers, Munich.
34. Szabo, A., Schulten, K. & Schulten, Z. (1980). First passage time approach to diffusion controlled reactions. *J. Chem. Phys.* 72, 4350-4357.
35. Koradi, R., Billeter, M. & Wuthrich, K. (1996). MOLMOL: a program for display and analysis of macromolecular structures. *J. Mol. Graphics* 14, 51-55.
36. Gronenborn, A. M., Filpula, D. R., Essig, N. Z., Achari, A., Whitlow, M., Wingfield, P. T. & Clore, G. M. (1991). A novel, highly stable fold of the immunoglobulin binding domain of streptococcal protein G. *Science* 253, 657-661.

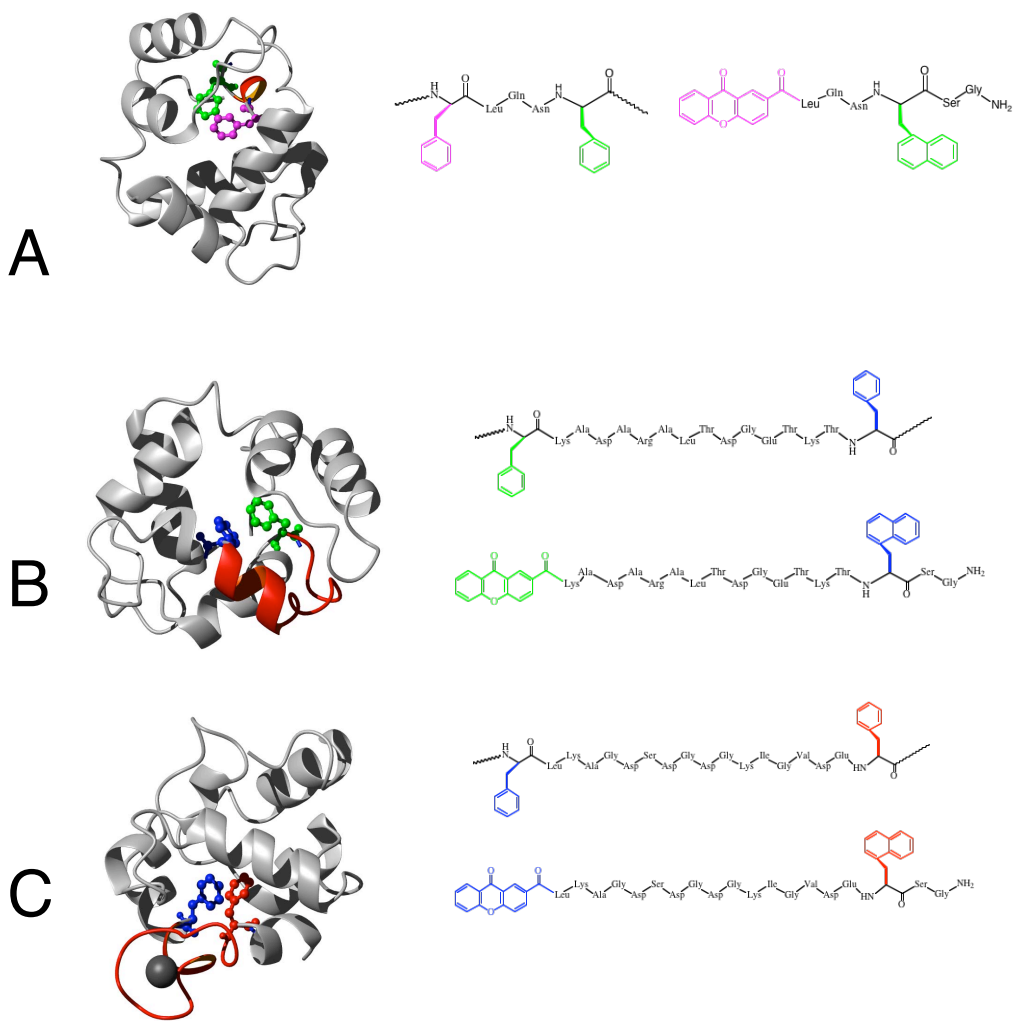


Figure 1

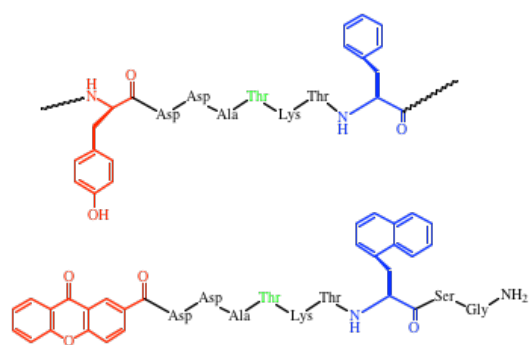
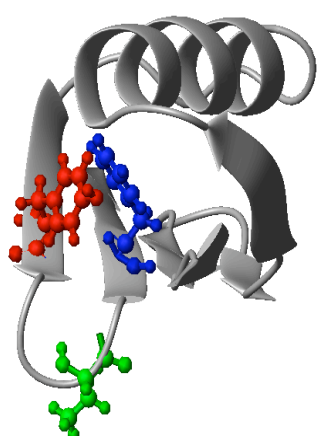


Figure 2

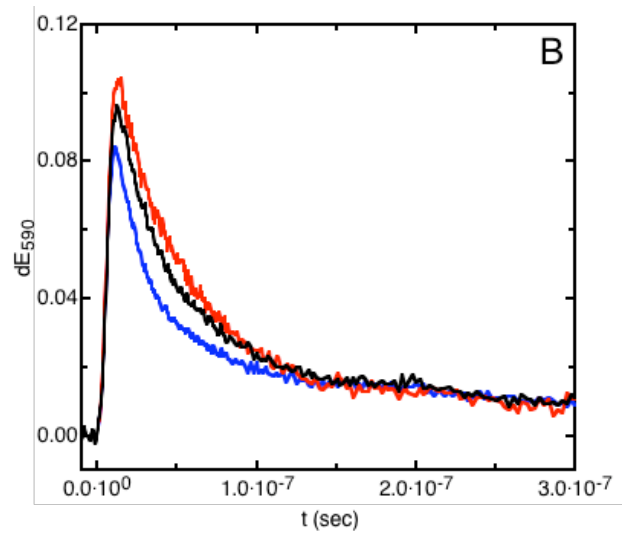
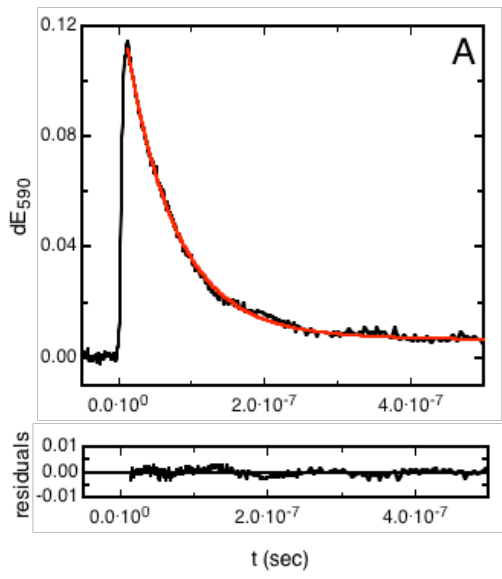


Figure 3

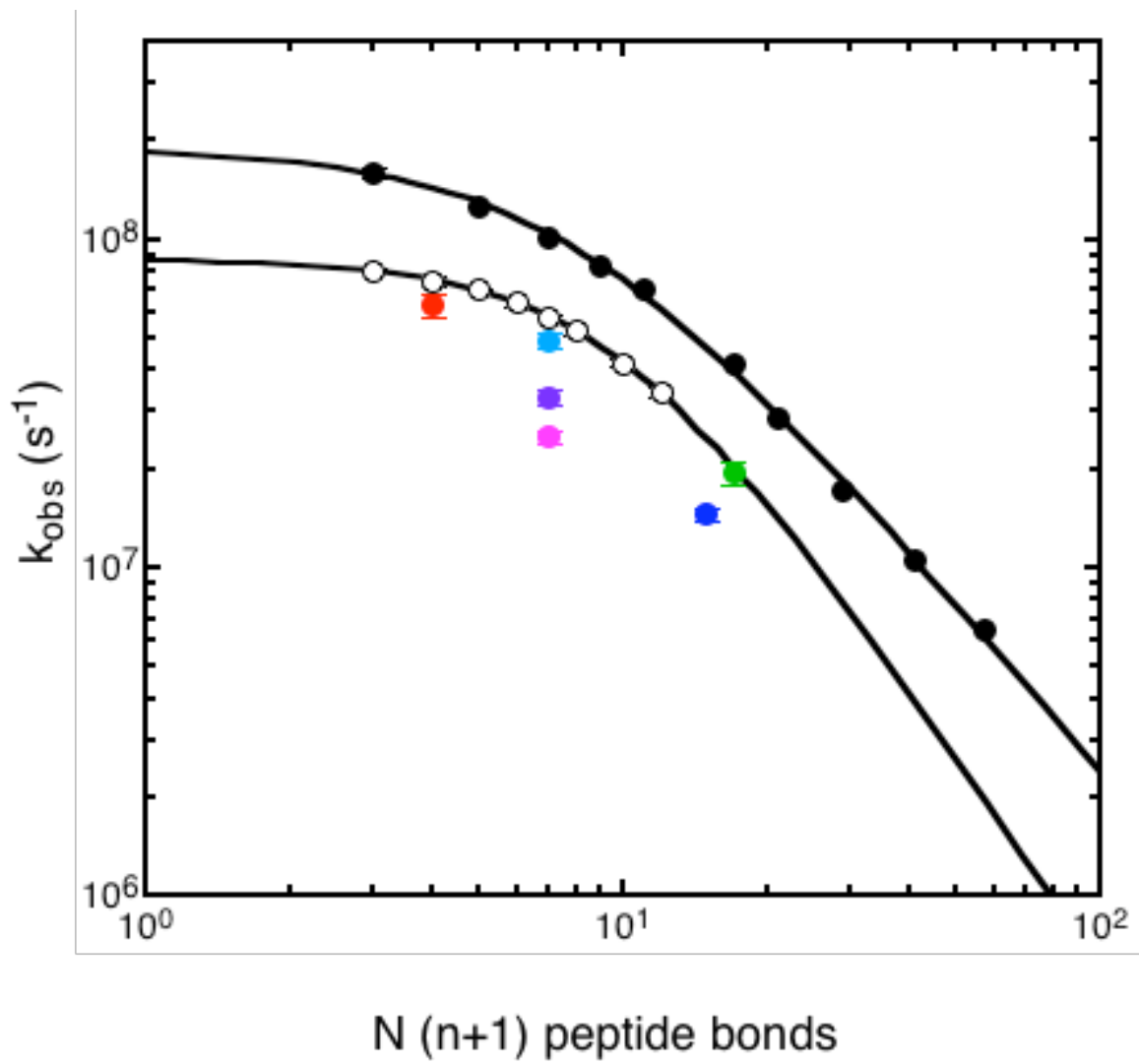
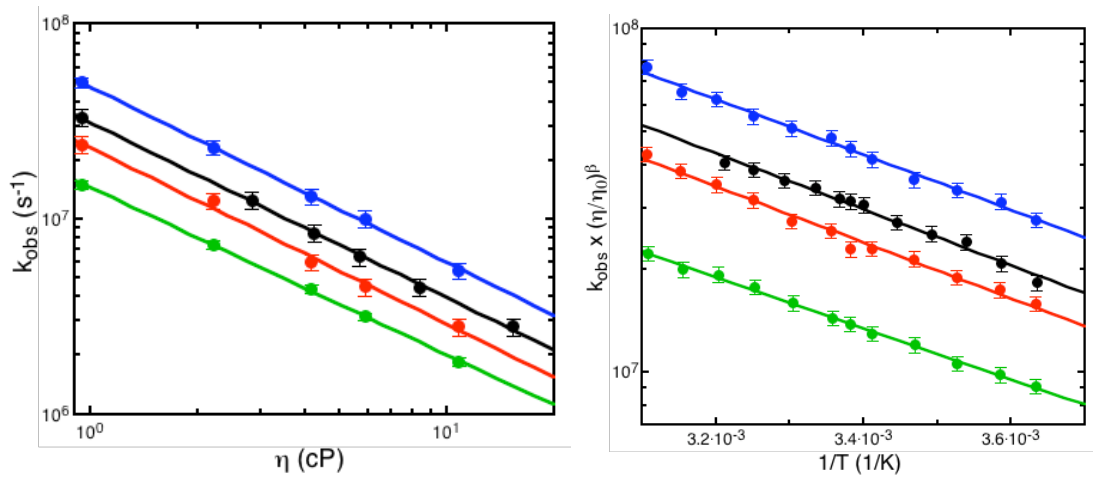


Figure 4



A

B

Figure 5

8. Curriculum vitae

Florian Stefan Willibald Krieger

Biozentrum der Universität Basel

Abteilung Biophysikalische Chemie phone: ++41-(0)61-26722007

Klingelbergstrasse 70 Fax: ++41-61-267 2189

CH-4056 Basel, Switzerland e-mail: florian.krieger@unibas.ch

Personal Information

Marital status: single Date of birth: 13.11.1974

Nationality: german Place of birth: Freiburg/Brsg

Education

January 2001 – December 2004

Biozentrum, University of Basel, Switzerland

Ph. D. in biophysics in December 2004

Ph. D. thesis supervisor: Prof. Dr. Thomas. Kiefhaber

Title: “Characterization of Intrachain Diffusion of Unfolded Polypeptide Chains as an Elementary Step in Protein Folding”

September 2000 – December 2000

Martin-Luther University of Halle-Wittenberg, Germany

Scientific assistant at the Institute of Biochemistry, Department of Enzymology,

December 1999 – September 2000

Martin-Luther University of Halle-Wittenberg, Germany

Diploma thesis in September 2000

Diploma supervisor: Dr. Stephan. König

Title: Isolation und Characterization of Pyruvate Decarboxylase from *Kluyveromyces lactis*,

October 1995 – November 1999

Martin-Luther University of Halle-Wittenberg, Germany

Studies in biochemistry

September 1985 – June 1994

Helmholtz-Gymnasium Karlsruhe, Germany

Abitur, in June 1994

September 1981– July 1985

Weinbrenner-Grundschule in Karlsruhe, Germany